ACCURACY ISSUES IN SOLVING SADDLE-POINT PROBLEMS

Mario Arioli

m.arioli@rl.ac.uk

STFC-Rutherford Appleton Laboratory

Introduction

- Why saddle-point problems?

Introduction

- Why saddle-point problems?
  - Saddle-point problems are the framework of several applications (G. Strang Siam review 1988)

Introduction

- Why saddle-point problems?
  - Saddle-point problems are the framework of several applications (G. Strang Siam review 1988)
  - Lagrange systems, optimization, mixed-hybrid finite-element for elasticity, fluid-dynamic, etc...

Introduction

- Why saddle-point problems?
  - Saddle-point problems are the framework of several applications (G. Strang Siam review 1988)
  - Lagrange systems, optimization, mixed-hybrid finite-element for elasticity, fluid-dynamic, etc...
- Why accuracy?

Introduction

- Why saddle-point problems?
  - Saddle-point problems are the framework of several applications (G. Strang Siam review 1988)
  - Lagrange systems, optimization, mixed-hybrid finite-element for elasticity, fluid-dynamic, etc...
- Why accuracy?
  - An approach based only on roundoff error analysis of the linear algebra is not sufficient for complex problems with highly discontinuous coefficients

Introduction

- Why saddle-point problems?
  - Saddle-point problems are the framework of several applications (G. Strang Siam review 1988)
  - Lagrange systems, optimization, mixed-hybrid finite-element for elasticity, fluid-dynamic, etc...
- Why accuracy?
  - An approach based only on roundoff error analysis of the linear algebra is not sufficient for complex problems with highly discontinuous coefficients
  - A GLOBAL approach taking into account the accuracy of the model, the accuracy of the approximation, and the accuracy of the numerical solver is required

Introduction

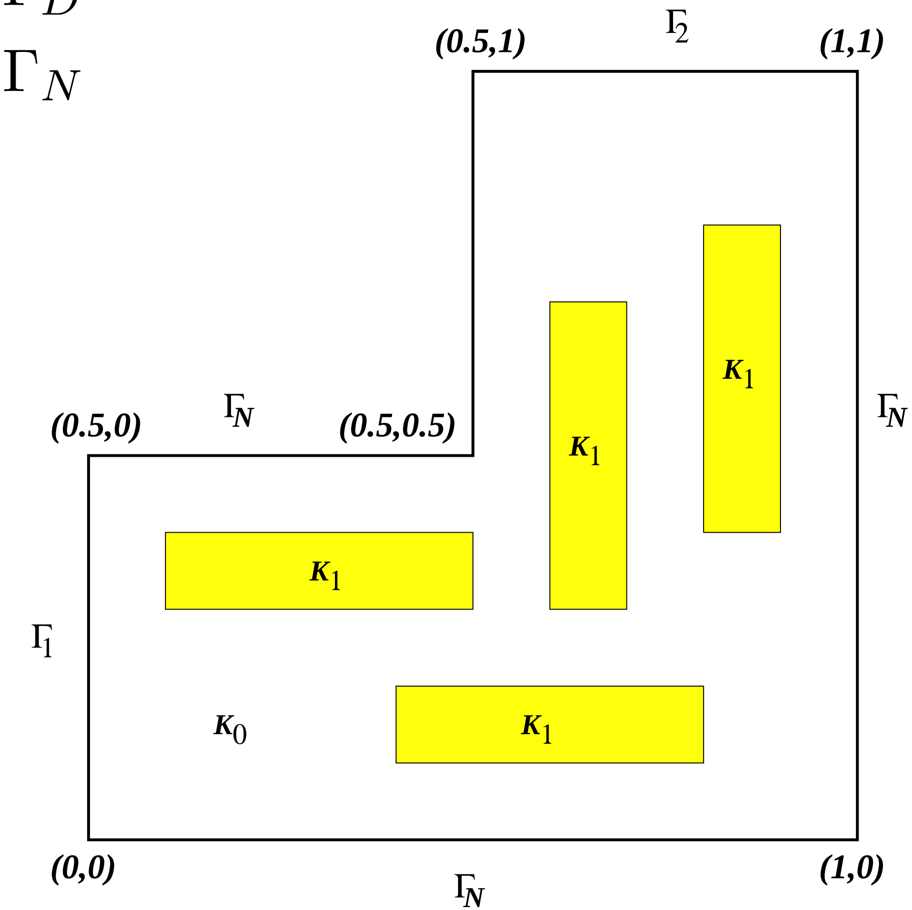
- Why saddle-point problems?
  - Saddle-point problems are the framework of several applications (G. Strang Siam review 1988)
  - Lagrange systems, optimization, mixed-hybrid finite-element for elasticity, fluid-dynamic, etc...
- Why accuracy?
  - An approach based only on roundoff error analysis of the linear algebra is not sufficient for complex problems with highly discontinuous coefficients
  - A GLOBAL approach taking into account the accuracy of the model, the accuracy of the approximation, and the accuracy of the numerical solver is required
  - We need to measure the distance between the exact solution of the model and the “numbers” we have in the computer

Darcy's equations: formulation

$$
\begin{aligned}
\mathbf{u}(\mathbf{x}) &= -\mathfrak{K}(\mathbf{x})\,\nabla\, p(\mathbf{x}), & \mathbf{x} \in \Omega \\
\operatorname{div}\, \mathbf{u}(\mathbf{x}) &= f(\mathbf{x}), & \mathbf{x} \in \Omega \\
p(\mathbf{x}) &= g_D(\mathbf{x}), & \mathbf{x} \in \Gamma_D \\
\mathbf{u} \cdot \mathbf{n} &= g_N(\mathbf{x}), & \mathbf{x} \in \Gamma_N
\end{aligned}
$$

Darcy's equations: formulation

$$
\begin{aligned}
\mathbf{u}(\mathbf{x}) &= -\mathfrak{K}(\mathbf{x})\,\nabla\, p(\mathbf{x}), && \mathbf{x} \in \Omega \\
\operatorname{div}\,\mathbf{u}(\mathbf{x}) &= f(\mathbf{x}), && \mathbf{x} \in \Omega \\
p(\mathbf{x}) &= g_D(\mathbf{x}), && \mathbf{x} \in \Gamma_D \\
\mathbf{u}\cdot\mathbf{n} &= g_N(\mathbf{x}), && \mathbf{x} \in \Gamma_N
\end{aligned}
$$



Darcy's equations: weak form

$$\mathcal{V} = \left\{ \mathbf{q} \,|\, \mathbf{q} \in \left(L^2(\Omega)\right)^2, \mathrm{div}\,\mathbf{q} \in L^2(\Omega), \mathbf{q}\cdot\underline{n}|_{\Gamma_N} = 0 \right\},$$

$L^2(\Omega)$

find $\mathbf{v} \in \mathcal{V}$ *and* $p \in L^2(\Omega)$ *such that:*

$$\begin{cases} \displaystyle\int_\Omega \mathfrak{K}^{-1}\mathbf{v}\cdot\mathbf{w}\,d\mathbf{x} - \int_\Omega p\,\mathrm{div}\,\mathbf{w}\,d\mathbf{x} &= \displaystyle -\int_{\Gamma_D} g_D \mathbf{w}\cdot\underline{n}\,ds \quad \forall \mathbf{w} \in \mathcal{V}, \\ \displaystyle\int_\Omega (\mathrm{div}\,\mathbf{v})\phi\,d\mathbf{x} &= \displaystyle\int_\Omega f\phi\,d\mathbf{x} \quad \forall \phi \in L^2(\Omega). \end{cases}$$

Darcy's equations: mixed finite element

$\mathfrak{T}_{\mathfrak{h}} = \{T\}$ a family of conforming triangulations covering the computational domain Ω with $\mathfrak{h} = \max_{T \in \mathfrak{T}_{\mathfrak{h}}} \mathsf{diam}(T)$

Darcy's equations: mixed finite element

$\mathfrak{T}_{\mathfrak{h}} = \{T\}$ a family of conforming triangulations covering the computational domain Ω with $\mathfrak{h} = \max_{T \in \mathfrak{T}_{\mathfrak{h}}} \mathsf{diam}(T)$

$$V_{\mathfrak{h}} = \left\{ \mathbf{w}(\mathbf{x}) : \Omega \to \mathbf{R}^2,\, \mathbf{w}(\mathbf{x})|_T = \alpha \mathbf{x} + \mathbf{z},\, \alpha \in \mathbf{R}\, \forall T \in \mathfrak{T}_{\mathfrak{h}},\, \mathbf{w} \cdot \underline{n}|_{\Gamma_N} = 0 \right\},$$

$$Q_{\mathfrak{h}} = \{ \phi(\mathbf{x}) : \Omega \to \mathbf{R},\, \phi(\mathbf{x})|_T = \mathsf{const},\, \forall T \in \mathfrak{T}_{\mathfrak{h}} \},$$

find $\mathbf{v}_{\mathfrak{h}} \in V_{\mathfrak{h}}$ *and* $p_{\mathfrak{h}} \in Q_{\mathfrak{h}}$ *such that:*

$$\begin{cases} \int_\Omega \mathfrak{K}^{-1} \mathbf{v}_{\mathfrak{h}} \cdot \mathbf{w}\, d\mathbf{x} - \int_\Omega p_{\mathfrak{h}} \operatorname{div} \mathbf{w}\, d\mathbf{x} = -\int_{\Gamma_D} g_D \mathbf{w} \cdot \underline{n}\, ds & \forall \mathbf{w} \in V_{\mathfrak{h}}, \\ \int_\Omega (\operatorname{div} \mathbf{v}_{\mathfrak{h}}) \phi\, d\mathbf{x} = \int_\Omega f \phi\, d\mathbf{x} & \forall \phi \in Q_{\mathfrak{h}}. \end{cases}$$

Darcy's equations: mixed finite element

- Every family of triangulations is *regular* in the sense of Ciarlet, no triangle in $\mathfrak{T}_{\mathfrak{h}}$ can have more than one edge on the boundary Γ nor that a triangle can exist with a vertex on Γ_D and any other vertex on Γ_N.

Darcy's equations: mixed finite element

- Every family of triangulations is *regular* in the sense of Ciarlet, no triangle in $\mathfrak{T}_{\mathfrak{h}}$ can have more than one edge on the boundary Γ nor that a triangle can exist with a vertex on Γ_D and any other vertex on Γ_N.
- $V_{\mathfrak{h}}$ is the space of the lowest-order Raviart-Thomas vector fields defined on Ω by using $\mathfrak{T}_{\mathfrak{h}}$
  $Q_{\mathfrak{h}}$ is the space of the piecewise constant functions defined on Ω by using $\mathfrak{T}_{\mathfrak{h}}$.

Darcy's equations: mixed finite element

- Every family of triangulations is *regular* in the sense of Ciarlet, no triangle in $\mathfrak{T}_{\mathfrak{h}}$ can have more than one edge on the boundary Γ nor that a triangle can exist with a vertex on Γ_D and any other vertex on Γ_N.
- $V_{\mathfrak{h}}$ is the space of the lowest-order Raviart-Thomas vector fields defined on Ω by using $\mathfrak{T}_{\mathfrak{h}}$
  $Q_{\mathfrak{h}}$ is the space of the piecewise constant functions defined on Ω by using $\mathfrak{T}_{\mathfrak{h}}$.
- Easy to couple with conservation laws approximated by FINITE VOLUME methods

Darcy's equations: mixed finite element

- Every family of triangulations is *regular* in the sense of Ciarlet, no triangle in $\mathfrak{T}_{\mathfrak{h}}$ can have more than one edge on the boundary Γ nor that a triangle can exist with a vertex on Γ_D and any other vertex on Γ_N.
- $V_{\mathfrak{h}}$ is the space of the lowest-order Raviart-Thomas vector fields defined on Ω by using $\mathfrak{T}_{\mathfrak{h}}$
  $Q_{\mathfrak{h}}$ is the space of the piecewise constant functions defined on Ω by using $\mathfrak{T}_{\mathfrak{h}}$.
- Easy to couple with conservation laws approximated by FINITE VOLUME methods
- Sufficiently accurate for the problem ($\mathfrak{K}$ can be only given as random field in black-oil problems)

Darcy's equations: saddle-point system

$$\begin{bmatrix} M & A \\ A^T & 0 \end{bmatrix} \begin{bmatrix} u \\ p \end{bmatrix} = \begin{bmatrix} q \\ b \end{bmatrix},$$

$$(M)_{e_i e_k} = \int_\Omega \mathfrak{K}^{-1} \mathbf{w}_{e_i} \cdot \mathbf{w}_{e_k} \, d\mathbf{x} \,, \quad (A)_{e_i j} = -\int_\Omega \operatorname{div} \mathbf{w}_{e_i} \phi_j \, d\mathbf{x} \,,$$
$$(q)_{e_i} = \int_\Gamma g_D \mathbf{w}_{e_i} \cdot \underline{n} \, ds \,, \qquad (b)_j = -\int_\Omega f \phi_j \, d\mathbf{x} \,.$$

Darcy's equations: saddle-point system

$$\begin{bmatrix} M & A \\ A^T & 0 \end{bmatrix} \begin{bmatrix} u \\ p \end{bmatrix} = \begin{bmatrix} q \\ b \end{bmatrix},$$

$$(M)_{e_i e_k} = \int_\Omega \mathfrak{K}^{-1} \mathbf{w}_{e_i} \cdot \mathbf{w}_{e_k} \, d\mathbf{x}\,, \quad (A)_{e_i j} = -\int_\Omega \operatorname{div} \mathbf{w}_{e_i} \phi_j \, d\mathbf{x}\,,$$
$$(q)_{e_i} = \int_\Gamma g_D \mathbf{w}_{e_i} \cdot \underline{n} \, ds\,, \qquad (b)_j = -\int_\Omega f \phi_j \, d\mathbf{x}\,.$$

$$\boldsymbol{\mathcal{A}} = \begin{bmatrix} M & A \\ A^T & 0 \end{bmatrix}$$

From Brezzi-Fortin (Springer, 1991), we have

- $\boldsymbol{\mathcal{A}}$ invertible $\forall \mathfrak{h}$
- We can choose norms s.t. $||u|| + ||p|| \leq C(||q|| + ||b||)$ with C independent of $\mathfrak{h}$ i.e. $\boldsymbol{\mathcal{A}}$ uniformly stable

Darcy's equations: saddle-point system

$$\begin{bmatrix} M & A \\ A^T & 0 \end{bmatrix} \begin{bmatrix} u \\ p \end{bmatrix} = \begin{bmatrix} q \\ b \end{bmatrix},$$

$$(M)_{e_i e_k} = \int_\Omega \mathfrak{K}^{-1} \mathbf{w}_{e_i} \cdot \mathbf{w}_{e_k} \, d\mathbf{x} \,, \quad (A)_{e_i j} = -\int_\Omega \operatorname{div} \mathbf{w}_{e_i} \phi_j \, d\mathbf{x} \,,$$
$$(q)_{e_i} = \int_\Gamma g_D \mathbf{w}_{e_i} \cdot \underline{n} \, ds \,, \qquad (b)_j = -\int_\Omega f \phi_j \, d\mathbf{x} \,.$$

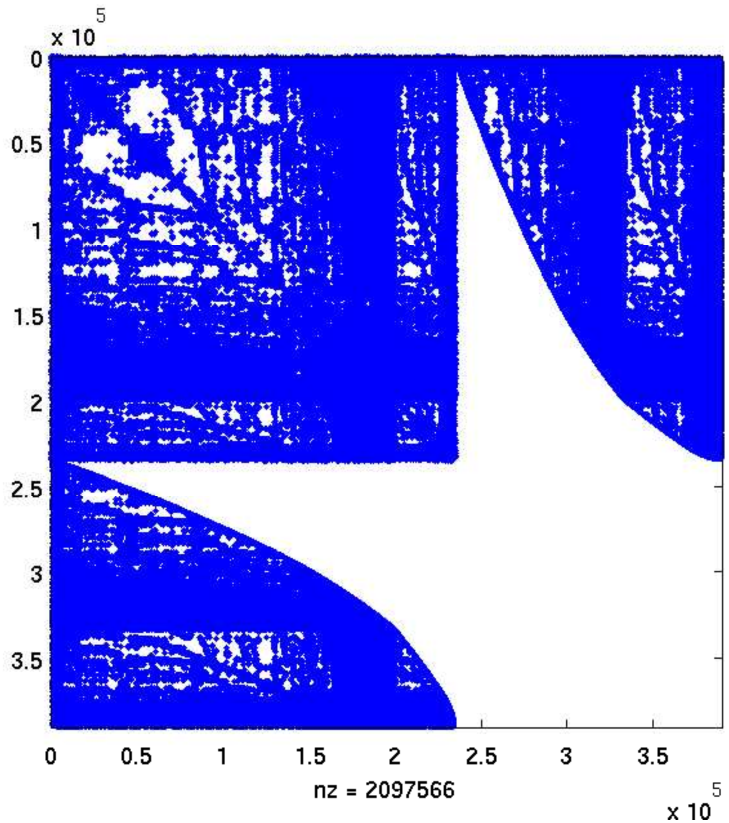
$$\boldsymbol{\mathcal{A}} = \begin{bmatrix} M & A \\ A^T & 0 \end{bmatrix}$$

From Brezzi-Fortin (Springer, 1991), we have

- $\boldsymbol{\mathcal{A}}$ invertible $\forall \mathfrak{h}$
- We can choose norms s.t. $||u|| + ||p|| \leq C(||q|| + ||b||)$ with C independent of $\mathfrak{h}$ i.e. $\boldsymbol{\mathcal{A}}$ uniformly stable

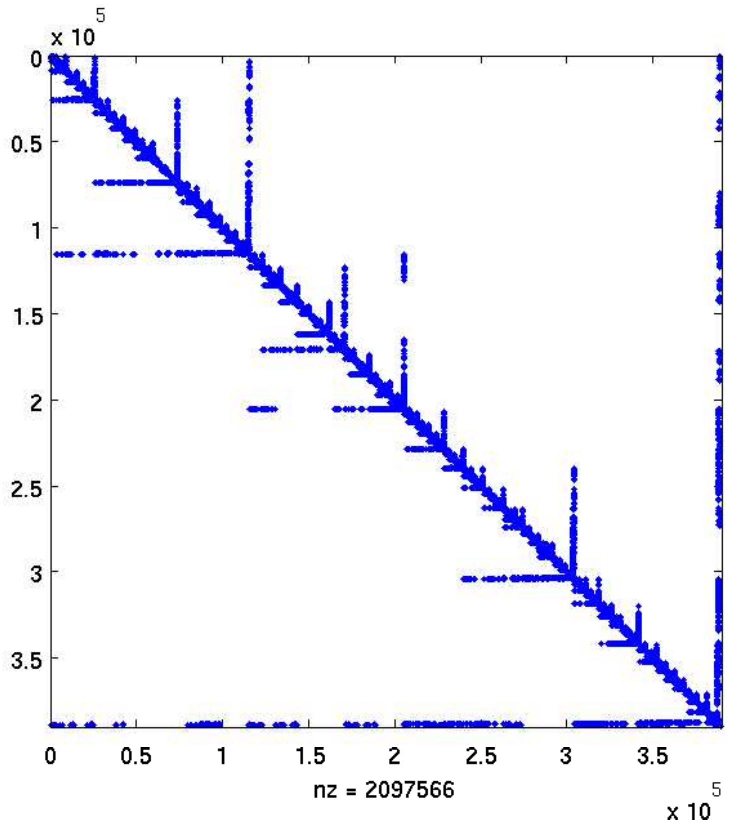
Solution methods

- Direct solvers
  $P^T \boldsymbol{\mathcal{A}} P = L^T D L$ (MA57)



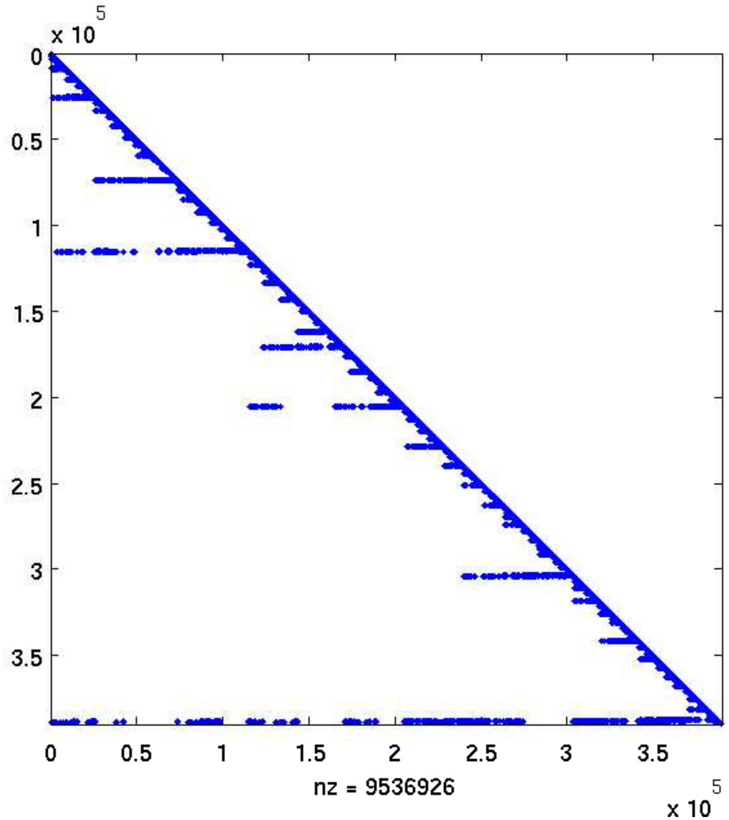
Solution methods

- Direct solvers
  $P^T \boldsymbol{\mathcal{A}} P = L^T D L$ (MA57)



Solution methods

- Direct solvers
  $P^T \boldsymbol{\mathcal{A}} P = L^T D L$ (MA57)



Solution methods

- Direct solvers
  $P^T \mathcal{A} P = L^T D L$ (MA57)
- Preconditioned Iterative solvers (Krylov methods)

Krylov methods for $Hx = g$

Find $\tilde{x} \in \mathcal{K}$ such that $g - H\tilde{x} \perp \mathcal{L}$

Solution methods

- Direct solvers
  $P^T \boldsymbol{\mathcal{A}} P = L^T D L$ (MA57)
- Preconditioned Iterative solvers (Krylov methods)

Krylov methods for $Hx = g$

Find $\tilde{x} \in \mathcal{K}$ such that $g - H\tilde{x} \perp \mathcal{L}$

If H SPD

$\mathcal{L} = \mathcal{K} \rightarrow \min_{z \in \mathcal{K}} ||g - Hz||_{H^{-1}}$

$(= \min_{z \in \mathcal{K}} ||x - z||_H)$ (CG)

Solution methods

- Direct solvers
  $P^T \boldsymbol{\mathcal{A}} P = L^T D L$ (MA57)
- Preconditioned Iterative solvers (Krylov methods)

Krylov methods for $Hx = g$

Find $\tilde{x} \in \mathcal{K}$ such that $g - H\tilde{x} \perp \mathcal{L}$

If H SPD

$\mathcal{L} = \mathcal{K} \rightarrow \min_{z \in \mathcal{K}} ||g - Hz||_{H^{-1}}$

$(= \min_{z \in \mathcal{K}} ||x - z||_H)$ (CG)

If $\mathcal{L} = H\,\mathcal{K} \rightarrow \min_{z \in \mathcal{K}} ||g - Hz||_2$ (GMRES)

Solution methods

- Direct solvers
  $P^T \boldsymbol{\mathcal{A}} P = L^T D L$ (MA57)
- Preconditioned Iterative solvers (Krylov methods)

Krylov methods for $Hx = g$

Find $\tilde{x} \in \mathcal{K}$ such that $g - H\tilde{x} \perp \mathcal{L}$

If H SPD

$\mathcal{L} = \mathcal{K} \rightarrow \min_{z \in \mathcal{K}} ||g - Hz||_{H^{-1}}$

$(= \min_{z \in \mathcal{K}} ||x - z||_H)$ (CG)

If $\mathcal{L} = H\,\mathcal{K} \rightarrow \min_{z \in \mathcal{K}} ||g - Hz||_2$ (GMRES)

Standard choice $\mathcal{K} = span\{g, Hg, \cdots, H^i g\}$

Solution methods

- Direct solvers
  $P^T \boldsymbol{\mathcal{A}} P = L^T D L$ (MA57)
- Preconditioned Iterative solvers (Krylov methods)
- Hybrid methods

$$\begin{bmatrix} M & A \\ A^T & 0 \end{bmatrix} \begin{bmatrix} u \\ p \end{bmatrix} = \begin{bmatrix} q \\ b \end{bmatrix},$$

Solution methods

- Direct solvers
  $P^T \boldsymbol{\mathcal{A}} P = L^T D L$ (MA57)
- Preconditioned Iterative solvers (Krylov methods)
- Hybrid methods
  - Schur complement (or Uzawa's method)

$$\begin{bmatrix} M & A \\ A^T & 0 \end{bmatrix} \begin{bmatrix} u \\ p \end{bmatrix} = \begin{bmatrix} q \\ b \end{bmatrix},$$

$$\Rightarrow \; -A^T M^{-1} A P = b - A^T M^{-1} q$$

Solution methods

- Direct solvers
  $P^T \boldsymbol{\mathcal{A}} P = L^T D L$ (MA57)
- Preconditioned Iterative solvers (Krylov methods)
- Hybrid methods
  - Schur complement (or Uzawa's method)
  - Null Space method

$$\begin{bmatrix} M & A \\ A^T & 0 \end{bmatrix} \begin{bmatrix} u \\ p \end{bmatrix} = \begin{bmatrix} q \\ b \end{bmatrix},$$

$\Rightarrow \; -A^T M^{-1} A P = b - A^T M^{-1} q$

$\Rightarrow \; Z^T M Z w = s$ with $Z^T A = 0$

Null Space algorithm

$Y^T A = I_m$ and $Z^T A = 0_{n-m,m}$.

Null Space Algorithm:

1. $u_0 = Yb$,
2. $Z^T M Z\, w = Z^T q - Z^T M u_0 = s$,
3. $u = u_0 + Zw$,
4. $p = Y^T q - Y^T M u$.

Null Space algorithm

$$Y^T A = I_m \qquad \text{and} \qquad Z^T A = 0_{n-m,m}.$$

Null Space Algorithm:

1. $u_0 = Yb,$
2. $Z^T M Z\, w = Z^T q - Z^T M u_0 = s,$
3. $u = u_0 + Zw,$
4. $p = Y^T q - Y^T M u.$

- $\mathcal{M} = Z^T M Z$ symmetric positive definite

Null Space algorithm

$$Y^T A = I_m \qquad \text{and} \qquad Z^T A = 0_{n-m,m}.$$

Null Space Algorithm:

1. $u_0 = Yb,$
2. $Z^T M Z\, w = Z^T q - Z^T M u_0 = s,$
3. $u = u_0 + Zw,$
4. $p = Y^T q - Y^T M u.$

- $\boldsymbol{\mathcal{M}} = Z^T M Z$ symmetric positive definite
- Y and Z can be computed either by a sparse version of the QR decomposition (`MA49`) where Z is stored as product of sparse matrices or by an implicit LU decomposition (A. SIMAX 2000 and SIMAX 2001, A. & Manzini BIT 2003 and ETNA 2006). We do not need to compute explicitly $\boldsymbol{\mathcal{M}}$ and the product $\boldsymbol{\mathcal{M}}w$ can be computed by $Z^T(M(Zw))$

The conjugate gradient method

$\boldsymbol{\mathcal{M}} = Z^T M Z$ symmetric positive definite

$\mathcal{H} = (\mathbf{R}^N, ||\cdot||_{\boldsymbol{\mathcal{M}}})$ and $\mathcal{H}' = (\mathbf{R}^N, ||\cdot||_{\boldsymbol{\mathcal{M}}^{-1}})$

At each step k the conjugate gradient method minimizes the ENERGY NORM of the error $\delta w^{(k)} = w - w^{(k)}$ on a Krylov space $w^{(0)} + \mathcal{K}_k$:

$$\min_{w^{(k)} \in \, w^{(0)} + \mathcal{K}_k} \|\delta w^{(k)}\|^2_{\boldsymbol{\mathcal{M}}}$$

$$\|\delta w^{(k)}\|_{\boldsymbol{\mathcal{M}}} = \|\rho_{w^{(k)}}\|_{\mathcal{H}'} = \|r^{(k)}\|_{\boldsymbol{\mathcal{M}}^{-1}}$$

$$r^{(k)} = s - \boldsymbol{\mathcal{M}} w^{(k)}$$

The stopping criteria

- Classic Criterion:

$$\text{IF } \|\boldsymbol{\mathcal{M}} w^{(k)} - s\|_2 \leq \sqrt{\varepsilon}\|s\|_2 \text{ THEN STOP },$$

$\varepsilon = 10^{-16}$

The stopping criteria

- Classic Criterion:

$$\text{IF } \|\boldsymbol{\mathcal{M}} w^{(k)} - s\|_2 \leq \sqrt{\varepsilon}\|s\|_2 \text{ THEN STOP },$$

$\varepsilon = 10^{-16}$

- New Criterion:

$$\text{IF } \|\boldsymbol{\mathcal{M}} w^{(k)} - s\|_{\boldsymbol{\mathcal{M}}^{-1}} \leq \eta\|s\|_{\boldsymbol{\mathcal{M}}^{-1}} \text{ THEN STOP },$$

with $\eta < 1$ an a-priori threshold fixed by the user.

The stopping criteria

- Classic Criterion:

$$\text{IF } \|\boldsymbol{\mathcal{M}} w^{(k)} - s\|_2 \leq \sqrt{\varepsilon}\|s\|_2 \text{ THEN STOP },$$

$\varepsilon = 10^{-16}$

- New Criterion:

$$\text{IF } \|\boldsymbol{\mathcal{M}} w^{(k)} - s\|_{\boldsymbol{\mathcal{M}}^{-1}} \leq \eta\|s\|_{\boldsymbol{\mathcal{M}}^{-1}} \text{ THEN STOP },$$

with $\eta < 1$ an a-priori threshold fixed by the user.

The choice of η will depend on the properties of the problem that we want to solve, and, in the practical cases, $\eta = \mathcal{O}(\mathfrak{h})$.

The stopping criteria cont.

$$\|\mathcal{M}w^{(k)} - s\|_{\mathcal{M}^{-1}} = \|w^{(k)} - w\|_{\mathcal{M}}$$

The stopping criteria cont.

$$\|\mathcal{M}w^{(k)} - s\|_{\mathcal{M}^{-1}} = \|w^{(k)} - w\|_{\mathcal{M}}$$

- At step k of CG we can compute an accurate lower bound ξ_k of the error at step $k-d$ using the numerical values already computed by CG. The choice of a value for d depends on preconditioner and ill-conditioning (Hestenes-Stiefel rule (1952), Strakoš and Tichý, ETNA(2002), BIT(2006), A. Numer. Math. 2004).
  numerically stable in finite-precision arithmetic

The stopping criteria cont.

$$\|\mathcal{M}w^{(k)} - s\|_{\mathcal{M}^{-1}} = \|w^{(k)} - w\|_{\mathcal{M}}$$

- At step k of CG we can compute an accurate lower bound ξ_k of the error at step $k-d$ using the numerical values already computed by CG. The choice of a value for d depends on preconditioner and ill-conditioning (Hestenes-Stiefel rule (1952), Strakoš and Tichý, ETNA(2002), BIT(2006), A. Numer. Math. 2004). numerically stable in finite-precision arithmetic
- If $||w - w^{(0)}||^2_{\mathcal{M}} \leq ||w||^2_{\mathcal{M}}$ we can compute a stable lower bound ρ_k for

$$s^T\mathcal{M}^{-1}s = w^T\mathcal{M}u \geq \rho_k.$$

The stopping criteria cont.

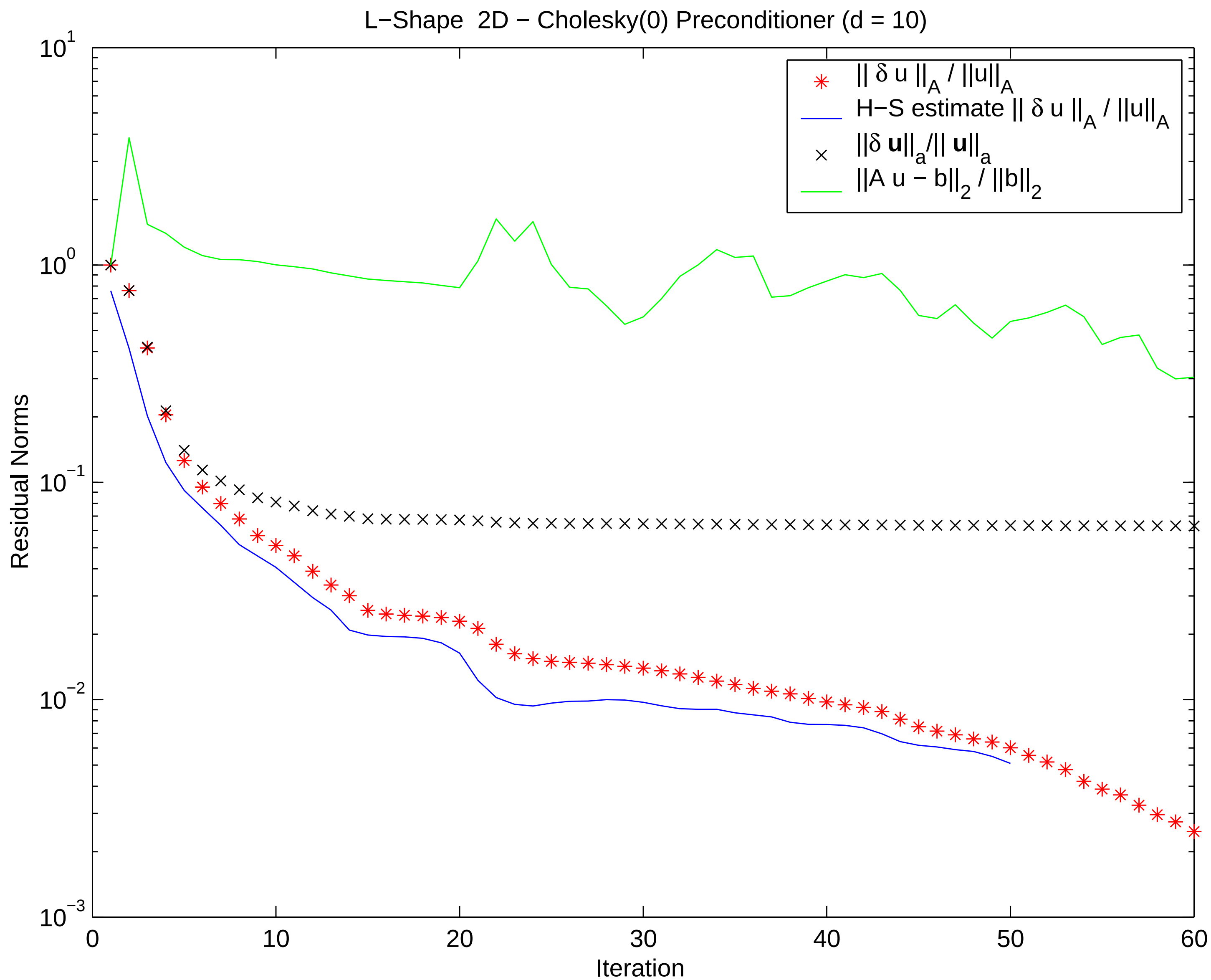
$$\|\mathcal{M}w^{(k)} - s\|_{\mathcal{M}^{-1}} = \|w^{(k)} - w\|_{\mathcal{M}}$$

- At step k of CG we can compute an accurate lower bound ξ_k of the error at step $k-d$ using the numerical values already computed by CG. The choice of a value for d depends on preconditioner and ill-conditioning (Hestenes-Stiefel rule (1952), Strakoš and Tichý, ETNA(2002), BIT(2006), A. Numer. Math. 2004).
  numerically stable in finite-precision arithmetic
- If $||w - w^{(0)}||^2_{\mathcal{M}} \leq ||w||^2_{\mathcal{M}}$ we can compute a stable lower bound ρ_k for

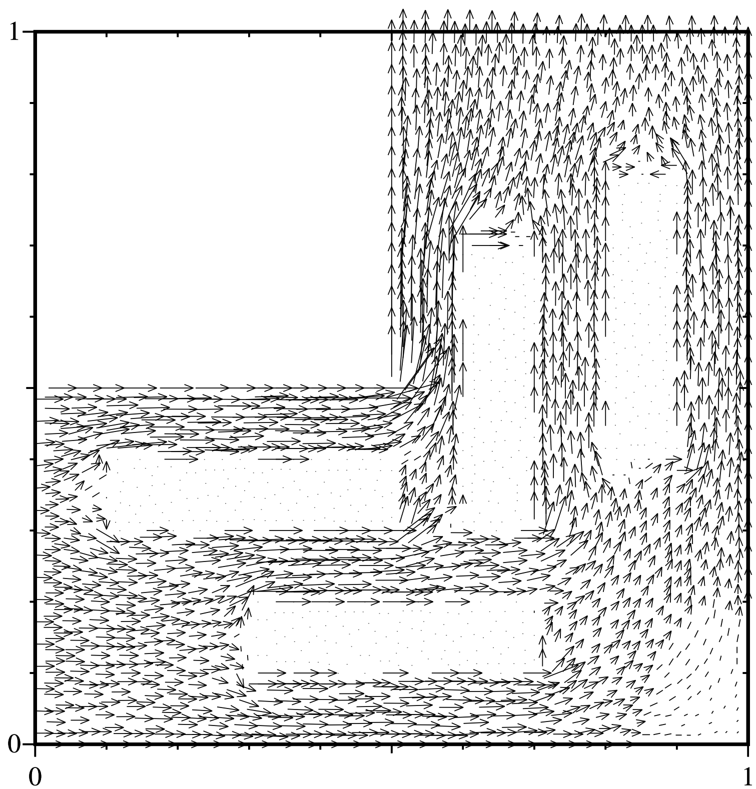
$$s^T \mathcal{M}^{-1} s = w^T \mathcal{M} u \geq \rho_k.$$

Therefore, we use the following stopping criterion (HSL_MI31)

IF $\xi_k \leq \eta^2 \rho_k$ THEN STOP .

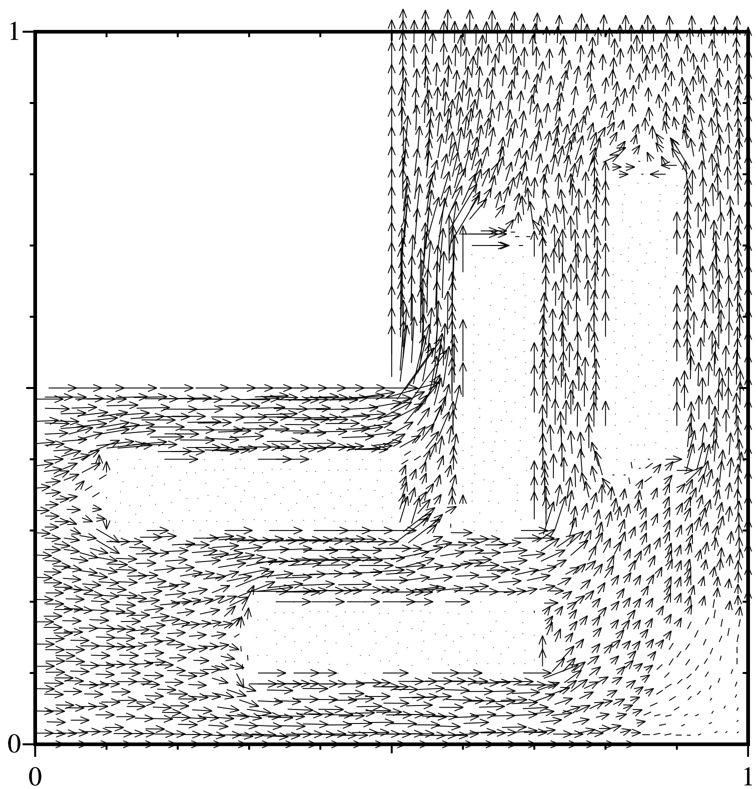

L−Shape 2D − Cholesky(0) Preconditioner (d = 10)
|| δ u ||A / ||u||A
H−S estimate || δ u ||A / ||u||A
||δ u||a/|| u||a
||A u − b||2 / ||b||2
Residual Norms
Iteration
10^1
10^0
10^−1
10^−2
10^−3
0
10
20
30
40
50
60

Numerical results ($K_1 = 10^{-8}$)

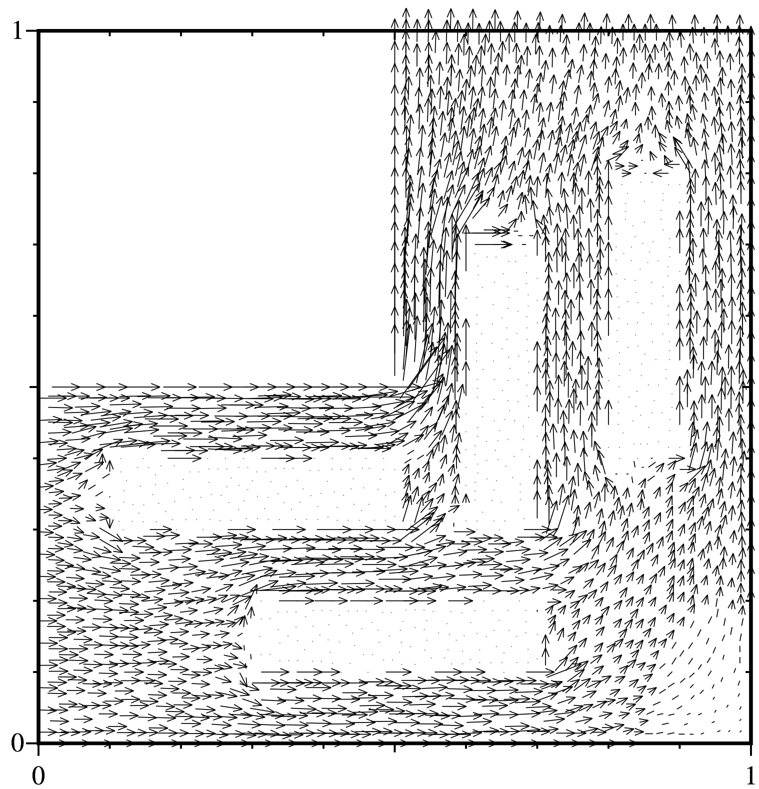


Direct solver solution

Numerical results ($K_1 = 10^{-8}$)



Direct solver solution



Null Space solution ($\eta = \mathcal{O}(\mathfrak{h})$)

Comments

- Roundoff?

Comments

- Roundoff?

  We can incorporate the roundoff error analysis in the bounds (A. SIMAX 2000, A. and Baldini SIMAX 2003)

Comments

- Roundoff?
  - We can incorporate the roundoff error analysis in the bounds (A. SIMAX 2000, A. and Baldini SIMAX 2003)
  - We can separate the two components of the error: roundoff and approximation. Usually, the latest is the dominant one

Comments

- Roundoff?
  - We can incorporate the roundoff error analysis in the bounds (A. SIMAX 2000, A. and Baldini SIMAX 2003)
  - We can separate the two components of the error: roundoff and approximation. Usually, the latest is the dominant one
- The theorems can be extended to other linear and non-linear cases (A., Loghin, Wathen Numer.Math 2005, A. and Loghin ETNA)

Comments

- Roundoff?
  - We can incorporate the roundoff error analysis in the bounds (A. SIMAX 2000, A. and Baldini SIMAX 2003)
  - We can separate the two components of the error: roundoff and approximation. Usually, the latest is the dominant one
- The theorems can be extended to other linear and non-linear cases (A., Loghin, Wathen Numer.Math 2005, A. and Loghin ETNA)
  - Advection-diffusion

Comments

- Roundoff?
  - We can incorporate the roundoff error analysis in the bounds (A. SIMAX 2000, A. and Baldini SIMAX 2003)
  - We can separate the two components of the error: roundoff and approximation. Usually, the latest is the dominant one
- The theorems can be extended to other linear and non-linear cases (A., Loghin, Wathen Numer.Math 2005, A. and Loghin ETNA)
  - Advection-diffusion
  - Navier-Stokes

Backward error

Theorem 0.0 *Let* u *be the solution of the weak formulation and let* $\mathbf{u}, u_h = \Pi_h \mathbf{u}$ *satisfy*

$$\boldsymbol{\mathcal{A}}\mathbf{u} = \mathbf{f}; \quad \frac{\|u - u_h\|_{\mathcal{H}}}{\|u_h\|_{\mathcal{H}}} \leq C(h).$$

Then, under B-B conditions, $\tilde{u}_h = \Pi_h \tilde{\mathbf{u}}$ *satisfies*

$$\frac{\|u - \tilde{u}_h\|_{\mathcal{H}}}{\|\tilde{u}_h\|_{\mathcal{H}}} \leq \tilde{C}(h) = O(C(h))$$

if

$$\frac{\|\mathbf{f} - \boldsymbol{\mathcal{A}}\tilde{\mathbf{u}}\|_{H^{-1}}}{\|\tilde{\mathbf{u}}\|_{H}} \leq \eta C(h) C_2,$$

for some $\eta \in (0,1)$.

$$\left(\mathcal{H} = \mathcal{V} \times L^2\right)$$

A., Loghin, and Wathen Numer. Math. 2005, and A. and Loghin ETNA 2006

Conclusions

A GLOBAL ACCURACY APPROACH INVOLVES A CORRECT MODELLING OF THE PHYSICAL PROBLEM, AND ADEQUATE APPROXIMATION METHOD AND NUMERICAL SOLVER. THIS MAKES THE DIFFERENCE BETWEEN CRUNCHING NUMBERS AND RELIABLE SIMULATION