Stopping criteria in the context of finite-element discretization

Mario Arioli

m.arioli@rl.ac.uk

CCLRC-Rutherford Appleton Laboratory

Outline

- Backward error and norms
- Krylov methods and stopping criteria
- Elliptic problems
- Summary and open problems
- Conclusions

Linear systems

$$\mathbf{A}\mathbf{u} = \mathbf{b}$$

$$\mathbf{A} \in \mathbf{R}^{\mathrm{N}\times\mathrm{N}}$$

Linear systems: variational framework

- Find $\mathbf{u} \in \mathcal{H}$ such that for all $\mathbf{v} \in \mathcal{H}$

$$a(\mathbf{u}, \mathbf{v}) = L(\mathbf{v}) \qquad (L(\cdot) \in \mathcal{H}' \text{ dual space of } \mathcal{H})$$

Linear systems: variational framework

- Find $\mathbf{u} \in \mathcal{H}$ such that for all $\mathbf{v} \in \mathcal{H}$

$$a(\mathbf{u}, \mathbf{v}) = L(\mathbf{v}) \qquad (L(\cdot) \in \mathcal{H}' \text{ dual space of } \mathcal{H})$$

- Existence and uniqueness: $\forall \mathbf{v}, \mathbf{w} \in \mathcal{H}$

$$\begin{aligned} a(\mathbf{w}, \mathbf{v}) &\leq C_1 \|\mathbf{w}\|_{\mathcal{H}} \|\mathbf{v}\|_{\mathcal{H}} \\ \sup_{\mathbf{w} \in \mathcal{H} \setminus \{0\}} \frac{a(\mathbf{w}, \mathbf{v})}{\|\mathbf{w}\|_{\mathcal{H}}} &\geq C_2 \|\mathbf{v}\|_{\mathcal{H}} \end{aligned}$$

Linear systems: variational framework

- Find $\mathbf{u} \in \mathcal{H}$ such that for all $\mathbf{v} \in \mathcal{H}$

$$a(\mathbf{u}, \mathbf{v}) = L(\mathbf{v}) \qquad (L(\cdot) \in \mathcal{H}' \text{ dual space of } \mathcal{H})$$

- Existence and uniqueness: $\forall \mathbf{v}, \mathbf{w} \in \mathcal{H}$

$$\begin{aligned} a(\mathbf{w}, \mathbf{v}) &\leq C_1 \|\mathbf{w}\|_{\mathcal{H}} \|\mathbf{v}\|_{\mathcal{H}} \\ \sup_{\mathbf{w} \in \mathcal{H} \setminus \{0\}} \frac{a(\mathbf{w}, \mathbf{v})}{\|\mathbf{w}\|_{\mathcal{H}}} &\geq C_2 \|\mathbf{v}\|_{\mathcal{H}} \end{aligned}$$

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

CCLRC Rutherford Appleton Laboratory

Linear systems: variational framework

- Find $\mathbf{u} \in \mathcal{H}$ such that for all $\mathbf{v} \in \mathcal{H}$

$$a(\mathbf{u}, \mathbf{v}) = L(\mathbf{v}) \qquad (L(\cdot) \in \mathcal{H}' \text{ dual space of } \mathcal{H})$$

- Existence and uniqueness: $\forall \mathbf{v}, \mathbf{w} \in \mathcal{H}$

$$\begin{aligned} a(\mathbf{w}, \mathbf{v}) &\leq C_1 \|\mathbf{w}\|_{\mathcal{H}} \|\mathbf{v}\|_{\mathcal{H}} \\ \sup_{\mathbf{w} \in \mathcal{H} \setminus \{0\}} \frac{a(\mathbf{w}, \mathbf{v})}{\|\mathbf{w}\|_{\mathcal{H}}} &\geq C_2 \|\mathbf{v}\|_{\mathcal{H}} \end{aligned}$$

- $\mathcal{H} = (\mathbf{R}^N, || \cdot ||_{\mathbf{H}})$ and $\mathcal{H}' = (\mathbf{R}^N, || \cdot ||_{\mathbf{H}^{-1}})$ $\mathbf{H}$ SPD
- $\mathbf{A}\mathbf{w} = \sigma \mathbf{H} \mathbf{v}$, $\mathbf{A}^T \mathbf{v} = \sigma \mathbf{H} \mathbf{w}$ and $\sigma \in (C_2, C_1)$
- C_2 and C_1 independent of N.

Backward error

We have the following equivalence:

$$\left.\begin{array}{l}\exists b \in \mathcal{BL}(\mathcal{H}), \exists \delta L \in \mathcal{H}' \text{ such that:}\\ a(\tilde{\mathbf{u}},\mathbf{v}) + b(\tilde{\mathbf{u}},\mathbf{v}) = (L+\delta L)(\mathbf{v}),\\ \forall \mathbf{v} \in \mathcal{H}, \text{ and}\\ \|b(\cdot,\cdot)\|_{\mathcal{BL}(\mathcal{H})} \le \alpha, \|\delta L\|_{\mathcal{H}'} \le \beta\end{array}\right\} \Leftrightarrow \left\{\begin{array}{l}\|\rho_{\tilde{\mathbf{u}}}\|_{\mathcal{H}'} \le \alpha \|\tilde{\mathbf{u}}\|_{\mathcal{H}} + \beta\\ \text{where } \rho_{\tilde{\mathbf{u}}} \in \mathcal{H}' \text{ is defined by}\\ \langle \rho_{\tilde{\mathbf{u}}}, \mathbf{v}\rangle_{\mathcal{H}',\mathcal{H}} = a(\tilde{\mathbf{u}},\mathbf{v}) - L(\mathbf{v}),\\ \forall \mathbf{v} \in \mathcal{H}\end{array}\right.$$

Rigal and Gaches (1967), A., Noulard, and Russo (2001)

The symmetric case: conjugate gradient method

$\mathbf{A}$ symmetric positive definite

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

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

$$\min_{\mathbf{u}^{(k)} \in\, \mathbf{u}^{(0)} + \mathcal{K}_k} \|\delta\mathbf{u}^{(k)}\|^2_{\mathbf{A}}$$

$$\|\delta\mathbf{u}^{(k)}\|_{\mathbf{A}} = \|\rho_{\mathbf{u}^{(k)}}\|_{\mathcal{H}} = \|r^{(k)}\|_{\mathbf{A}^{-1}}$$

$$r^{(k)} = \mathbf{b} - \mathbf{A}\mathbf{u}^{(k)}$$

The symmetric case: stopping criteria

- Classic Criterion:

$$\text{IF } \ \|\mathbf{A}\mathbf{u}^{(k)} - \mathbf{b}\|_2 \le \sqrt{\varepsilon}\|\mathbf{b}\|_2 \ \text{ THEN STOP },$$

The symmetric case: stopping criteria

- Classic Criterion:

$$\text{IF } \|\mathbf{Au}^{(k)} - \mathbf{b}\|_2 \leq \sqrt{\varepsilon}\|\mathbf{b}\|_2 \text{ THEN STOP },$$

- New Criterion:

$$\text{IF } \|\mathbf{Au}^{(k)} - \mathbf{b}\|_{\mathbf{A}^{-1}} \leq \eta\|\mathbf{b}\|_{A^{-1}} \text{ THEN STOP },$$

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

The symmetric case: stopping criteria

- Classic Criterion:

$$\text{IF } \|\mathbf{A}\mathbf{u}^{(k)} - \mathbf{b}\|_2 \leq \sqrt{\varepsilon}\|\mathbf{b}\|_2 \text{ THEN STOP },$$

- New Criterion:

$$\text{IF } \|\mathbf{A}\mathbf{u}^{(k)} - \mathbf{b}\|_{\mathbf{A}^{-1}} \leq \eta\|\mathbf{b}\|_{A^{-1}} \text{ THEN STOP },$$

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, η can be frequently much larger than ε , the roundoff unit of the computer finite precision arithmetic.

The symmetric case: stopping criteria cont.

$$\|\mathbf{A}\mathbf{u}^{(k)} - \mathbf{b}\|_{\mathbf{A}^{-1}}$$

The symmetric case: stopping criteria cont.

$$\|\mathbf{A}\mathbf{u}^{(k)} - \mathbf{b}\|_{\mathbf{A}^{-1}}$$

- Hestenes-Stiefel rule (1952) (see Strakoš and Tichý, ETNA(2002), BIT(2006)) numerically stable in finite-precision arithmetic

The symmetric case: stopping criteria cont.

$$\|\mathbf{Au}^{(k)} - \mathbf{b}\|_{\mathbf{A}^{-1}}$$

- Hestenes-Stiefel rule (1952) (see Strakoš and Tichý, ETNA(2002), BIT(2006)) numerically stable in finite-precision arithmetic
- Gauss quadrature rules (Golub and Meurant, 1997)

The symmetric case: stopping criteria cont.

$$\|\mathbf{Au}^{(k)} - \mathbf{b}\|_{\mathbf{A}^{-1}}$$

- Hestenes-Stiefel rule (1952) (see Strakoš and Tichý, ETNA(2002), BIT(2006)) numerically stable in finite-precision arithmetic
- Gauss quadrature rules (Golub and Meurant, 1997)
  - Gauss equivalent to Hestenes-Stiefel rule (Strakoš and Tichý). The Gauss quadrature does not require any a-priori knowledge of the smallest and the biggest eigenvalues and computes a lower bound of $\|\mathbf{Au}^{(k)} - \mathbf{b}\|_{\mathbf{A}^{-1}}$.

The symmetric case: stopping criteria cont.

$$\|\mathbf{A}\mathbf{u}^{(k)} - \mathbf{b}\|_{\mathbf{A}^{-1}}$$

- Hestenes-Stiefel rule (1952) (see Strakoš and Tichý, ETNA(2002), BIT(2006)) numerically stable in finite-precision arithmetic
- Gauss quadrature rules (Golub and Meurant, 1997)
  - Gauss equivalent to Hestenes-Stiefel rule (Strakoš and Tichý). The Gauss quadrature does not require any a-priori knowledge of the smallest and the biggest eigenvalues and computes a lower bound of $\|\mathbf{A}\mathbf{u}^{(k)} - \mathbf{b}\|_{\mathbf{A}^{-1}}$.
  - Gauss-Lobatto and Gauss-Radau. They compute lower and upper bounds using the extremes eigenvalues of $\mathbf{A}$.

The symmetric case: Hestenes-Stiefel rule

During the conjugate gradient iterates, we compute the scalar α_k and the conjugate vectors $\mathbf{p}^{(k)}$ ($\mathbf{p}^{(j)T}\mathbf{A}\mathbf{p}^{(i)} = 0,\ j \neq i$) and the residuals $r^{(k)}$. Thus,

$$\mathbf{u} = \sum_{j=1}^{\mathbf{N}} \alpha_j \mathbf{p}^{(j)}$$

and

$$\|\delta \mathbf{u}^{(k)}\|_{\mathbf{A}}^2 = \|\mathbf{A}\mathbf{u}^{(k)} - \mathbf{b}\|_{\mathbf{A}^{-1}}^2 = e_{\mathbf{A}}^2 = \sum_{j=k+1}^{\mathbf{N}} \alpha_j r^{(j)T} r^{(j)}$$

The symmetric case: Hestenes-Stiefel rule

Under the assumption that $e_{\mathbf{A}}^{(k+d)} << e_{\mathbf{A}}^{(k)}$, where the integer d denotes a suitable delay, the Hestenes and Stiefel estimate ξ_k will be

$$\xi_k = \sum_{j=k+1}^{k+d} \alpha_j r^{(j)T} r^{(j)}.$$

The symmetric case: Hestenes-Stiefel rule

Under the assumption that $e_{\mathbf{A}}^{(k+d)} << e_{\mathbf{A}}^{(k)}$, where the integer d denotes a suitable delay, the Hestenes and Stiefel estimate ξ_k will be

$$\xi_k = \sum_{j=k+1}^{k+d} \alpha_j r^{(j)T} r^{(j)}.$$

The choice of a value for d depends on preconditioner and ill-conditioning.

$\mathbf{b}^T\mathbf{A}^{-1}\mathbf{b}$

If $||\mathbf{u}-\mathbf{u}^{(0)}||^2_{\mathbf{A}} \leq ||\mathbf{u}||^2_{\mathbf{A}}$

$$\mathbf{b}^T\mathbf{A}^{-1}\mathbf{b} = \mathbf{u}^T\mathbf{A}u \geq \sum_{j=0}^{k} \alpha_j r^{(j)T} r^{(j)} + \mathbf{b}^T\mathbf{u}^{(0)} + \mathbf{r}^{(0)T}\mathbf{u}^{(0)} = \rho_k.$$

is a stable lower bound (see Strakoš and Tichý BIT(2006))
Therefore, we use the following stopping criterion

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

Preconditioning

The dual norm of the preconditioned residual is equal to the dual norm of the original residual.

A. Numer. Math. 2004, Meurant Numer Alg. 2004

Continuous problem

$$a(u,v) = \int_\Omega \mathfrak{K}(x)\nabla u \cdot \nabla v d\mathbf{x}, \quad \forall u,v \in H_0^1(\Omega)$$

$\forall u, v \in H_0^1(\Omega)$, $\exists \gamma \in \mathbf{R}_+$ and $\exists M \in \mathbf{R}_+$ such that

$$\begin{aligned} \gamma ||u||_{1,\Omega}^2 \leq \ & a(u,u) \\ & a(u,v) \ \leq M||u||_{1,\Omega}||v||_{1,\Omega} \ , \end{aligned}$$

$L(v) = \int_\Omega f v d\mathbf{x}$, $L(v) \in H^{-1}(\Omega)$.

$$(P) \begin{cases} \text{Find } u \in H_0^1(\Omega) \text{ such that} \\ a(u,v) = L(v), \ \ \forall v \in H_0^1(\Omega), \end{cases}$$

has a unique solution.

Finite-element approximation

- Weak formulation

$$\begin{cases} \text{Find } u_h \in \mathcal{H}_h \text{ such that} \\ a_h(u_h, v_h) = L_h(v_h), \quad \forall v_h \in \mathcal{H}_h, \end{cases}$$

Finite element methods choose $\mathcal{H}_h$ to be a space of functions v_h defined on a subdivision Ω_h of Ω into simplices T of diameter h_T ; h denotes a piecewise constant function defined on Ω_h via $h|_T = h_T$.

Finite-element approximation

- Weak formulation

$$\begin{cases} \text{Find } u_h \in \mathcal{H}_h \text{ such that} \\ a_h(u_h, v_h) = L_h(v_h), \quad \forall v_h \in \mathcal{H}_h, \end{cases}$$

Finite element methods choose $\mathcal{H}_h$ to be a space of functions v_h defined on a subdivision Ω_h of Ω into simplices T of diameter h_T ; h denotes a piecewise constant function defined on Ω_h via $h|_T = h_T$.

- Existence and uniqueness: $\mathcal{H}_h \subset \mathcal{H} = H_0^1(\Omega)$.
- Error Estimate: $\|u - u_h\|_{\mathcal{H}} \leq C(h)$

Finite-element approximation

- Weak formulation

$$\begin{cases} \text{Find } u_h \in \mathcal{H}_h \text{ such that} \\ a_h(u_h, v_h) = L_h(v_h), \quad \forall v_h \in \mathcal{H}_h, \end{cases}$$

  Finite element methods choose $\mathcal{H}_h$ to be a space of functions v_h defined on a subdivision Ω_h of Ω into simplices T of diameter h_T ; h denotes a piecewise constant function defined on Ω_h via $h|_T = h_T$.

- Existence and uniqueness: $\mathcal{H}_h \subset \mathcal{H} = H_0^1(\Omega)$.
- Error Estimate: $\|u - u_h\|_{\mathcal{H}} \leq C(h)$
- Solve $\mathbf{A}\mathbf{u}_h = \mathbf{b}$

Finite-element approximation

- Weak formulation

$$\begin{cases} \text{Find } u_h \in \mathcal{H}_h \text{ such that} \\ a_h(u_h, v_h) = L_h(v_h), \quad \forall v_h \in \mathcal{H}_h, \end{cases}$$

Finite element methods choose $\mathcal{H}_h$ to be a space of functions v_h defined on a subdivision Ω_h of Ω into simplices T of diameter h_T ; h denotes a piecewise constant function defined on Ω_h via $h|_T = h_T$.

- Existence and uniqueness: $\mathcal{H}_h \subset \mathcal{H} = H_0^1(\Omega)$.
- Error Estimate: $\|u - u_h\|_{\mathcal{H}} \leq C(h)$
- Solve $\mathbf{A}\mathbf{u}_h = \mathbf{b}$

$$\text{Note: } \|v_h\|_{\mathcal{H}_h} = \|\mathbf{v}\|_{\mathbf{H}}.$$

Finite-element framework

Finally, assuming $h < 1$ and $t > 0$, and choosing $\eta = \mathcal{O}(h)$, we have

$$\|u - u_h^{(k)}\|_{\mathcal{H}} \leq C^*(h^t)\|u\|_{\mathcal{H}} + 2\|u - u_h\|_{\mathcal{H}} \leq C(h).$$

where

- $u(\mathbf{x})$ is the exact solution of the variational problem,
- $u_h(\mathbf{x})$ is the exact solution of the approximate problem,
- $u_h^{(k)}(\mathbf{x}) = \sum_{i=1}^{N} \mathbf{u}_h^{(k)} \phi_i(\mathbf{x})$ is the approximate solution at step k. ($\phi_i(\mathbf{x})$ are the basis functions)

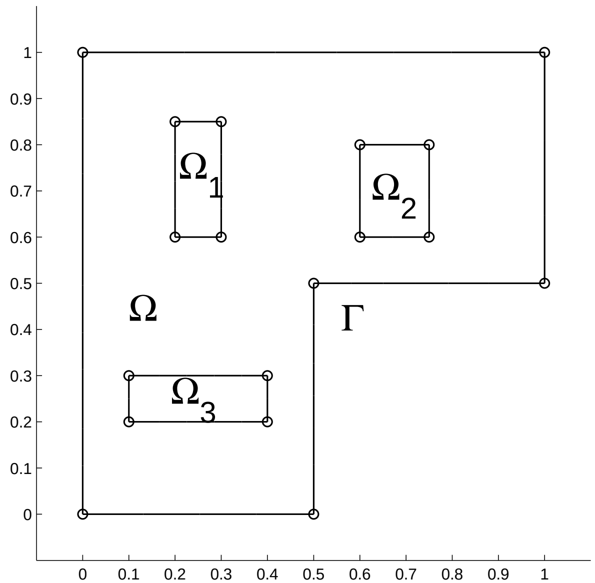
A. Numer. Math. 2004

Test problem

Problem

$$\mathfrak{K}(\mathbf{x}) = \begin{cases} 1 & \mathbf{x} \in \Omega \setminus \{\Omega_1 \cup \Omega_2 \cup \Omega_3\}, \\ 10^6 & \mathbf{x} \in \Omega_1, \\ 10^4 & \mathbf{x} \in \Omega_2, \\ 10^2 & \mathbf{x} \in \Omega_3. \end{cases}$$

$$L(v) = \int_\Omega 10 v d\mathbf{x}, \quad \forall v \in H_0^1(\Omega)$$



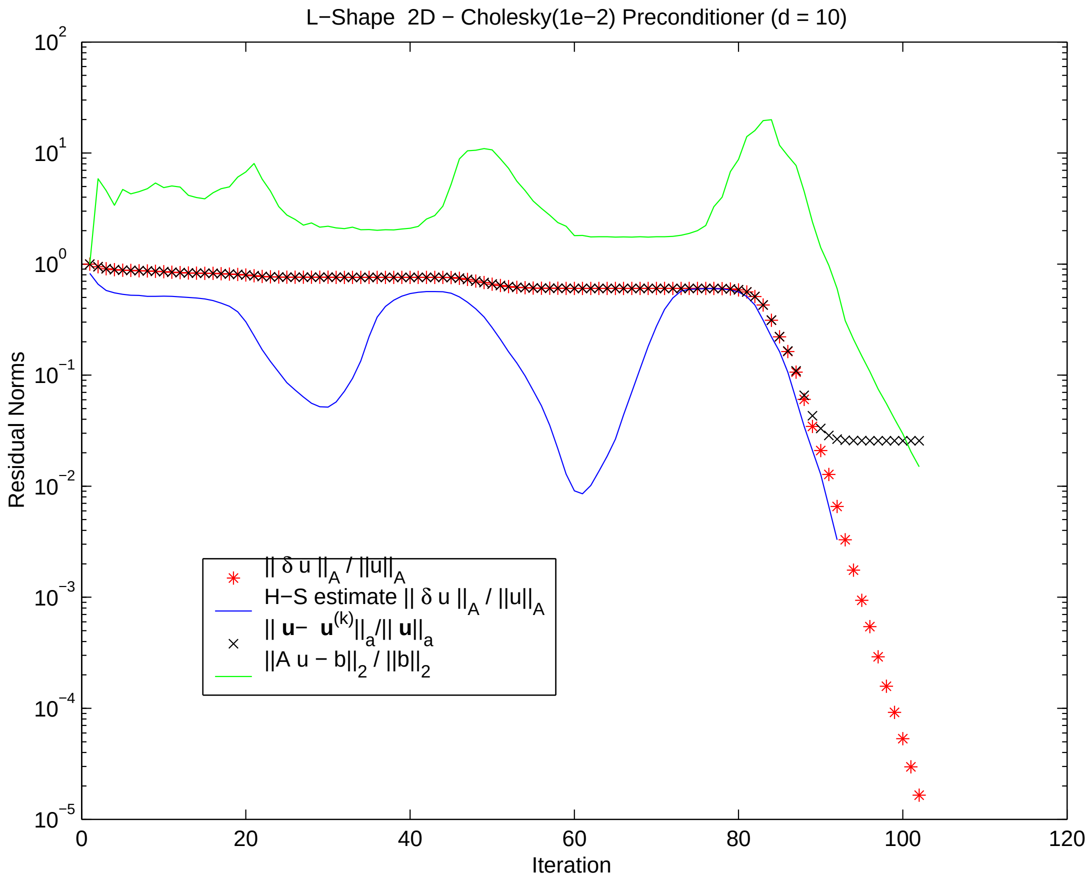
Preconditioners: estimates for $\kappa(\mathbf{M}^{-1}\mathbf{A})$

M	$\kappa(\mathbf{M}^{-1}\mathbf{A})$
I	$1.8\ 10^{10}$
Jacobi	$1.5\ 10^{9}$
Inc. Cholesky(0)	$4.3\ 10^{8}$

$\eta^2 \approx 10^{-5}$ and $\mathbf{N} = 29619.$

The condition numbers of the preconditioned matrices $\mathbf{M}^{-1}\mathbf{A}$ for the second problem are are still very high, and only the incomplete Cholesky preconditioner with drop tolerance 10^{-2} is an effective choice.

Example



Behaviour of the norms of the residual for the incomplete Cholesky preconditioner with drop tolerance 10^{-2} and $d = 10$.

The positive definite problem

- $a(u,v) \neq a(v,u)$
- $\mathbf{A}$ asymmetric but positive definite
- $\mathbf{H} = \frac{1}{2}(\mathbf{A}^T + \mathbf{A})$ SPD

See A., Login, and Wathen Numer. Math. 2005

Backward error

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

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

Then $\tilde{u}_h = \Pi_h \tilde{\mathbf{u}}$ *satisfies*

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

if

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

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

How to calculate $\|\mathbf{r}^k\|_{\mathbf{H}^{-1}}$?

- Solve preconditioned system

$$\mathbf{H}^{-1/2}\mathbf{A}\mathbf{H}^{-1/2}\hat{\mathbf{u}} = \mathbf{H}^{-1/2}\mathbf{f}$$

  - ► $\|\hat{\mathbf{r}}^k\|_{l_2} = \|\mathbf{r}^k\|_{H^{-1}}$
  - ► 3-term recurrence.
- Approximate it from Krylov subspace information.
- Concus & Golub, Widlund: 3-term recurrences for nonsymmetric problems
  - ► work in $\mathbf{H}$-inner product
  - ► do not minimize the residual norm.

Numerical experiments

- Discretization:
  linear elements on uniform & adaptive meshes. Error estimate:

$$|u - u_h|_1 \le C h^{s-1} \|u\|_s, \qquad 1 \le s \le 2.$$

Numerical experiments

- Discretization:
  linear elements on uniform & adaptive meshes. Error estimate:

$$|u - u_h|_1 \leq C h^{s-1} \|u\|_s, \qquad 1 \leq s \leq 2.$$

- Estimation of parameters

$$h \sim \frac{\|\mathbf{u}^k\|_{\mathbf{M}}}{\|\mathbf{u}^k\|_{l_2}}, \quad \|u\|_s \sim \|\mathbf{A}\mathbf{u}^k\|_{l_2}$$

Numerical experiments

Stopping criteria and estimates

- Residual dual norm: $\|\mathbf{r}^k\|_{\mathbf{H}^{-1}}$

Numerical experiments

Stopping criteria and estimates

- Residual dual norm: $\|\mathbf{r}^k\|_{\mathbf{H}^{-1}}$
- Energy estimate $\|\mathbf{u}^k - \mathbf{u}^{k-1}\|_H \leq C_2 h^2 \|\mathbf{A}\mathbf{u}^k\|_{l_2}$

Advection-diffusion problem

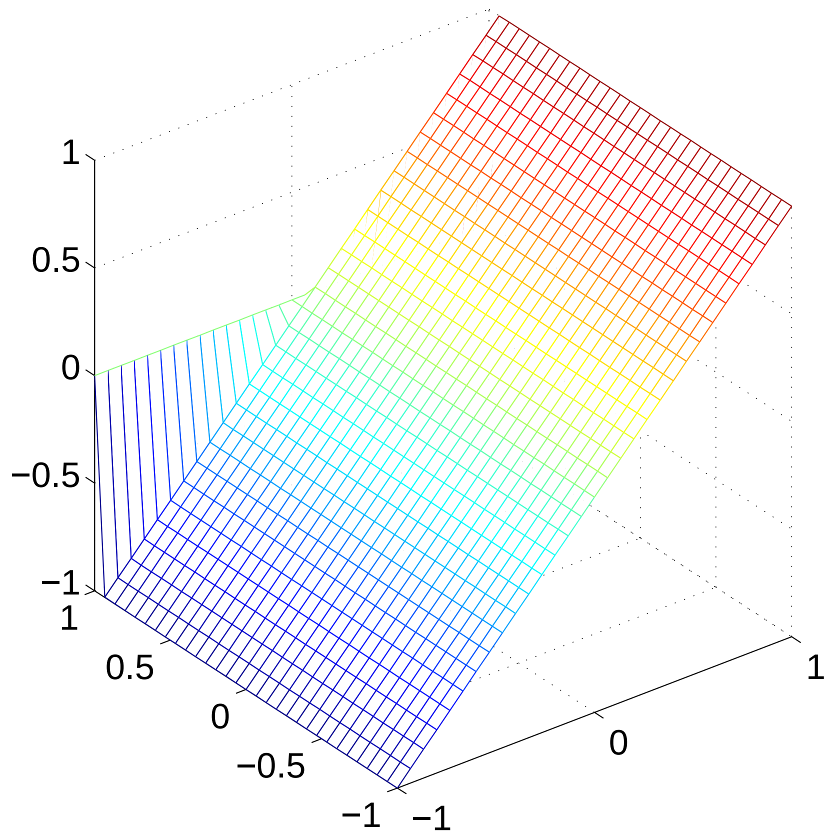
$$
\begin{aligned}
-\varepsilon \Delta u + \mathbf{b} \cdot \nabla u &= f \qquad \text{in } \Omega \\
u &= g \qquad \text{on } \Gamma.
\end{aligned}
$$

$$\mathbf{b} = (2y(1-x^2), -2x(1-y^2)),$$

$$u(x,y) = x \left(\frac{1 - e^{\frac{y-1}{\varepsilon}}}{1 - e^{-\frac{2}{\varepsilon}}} \right),$$

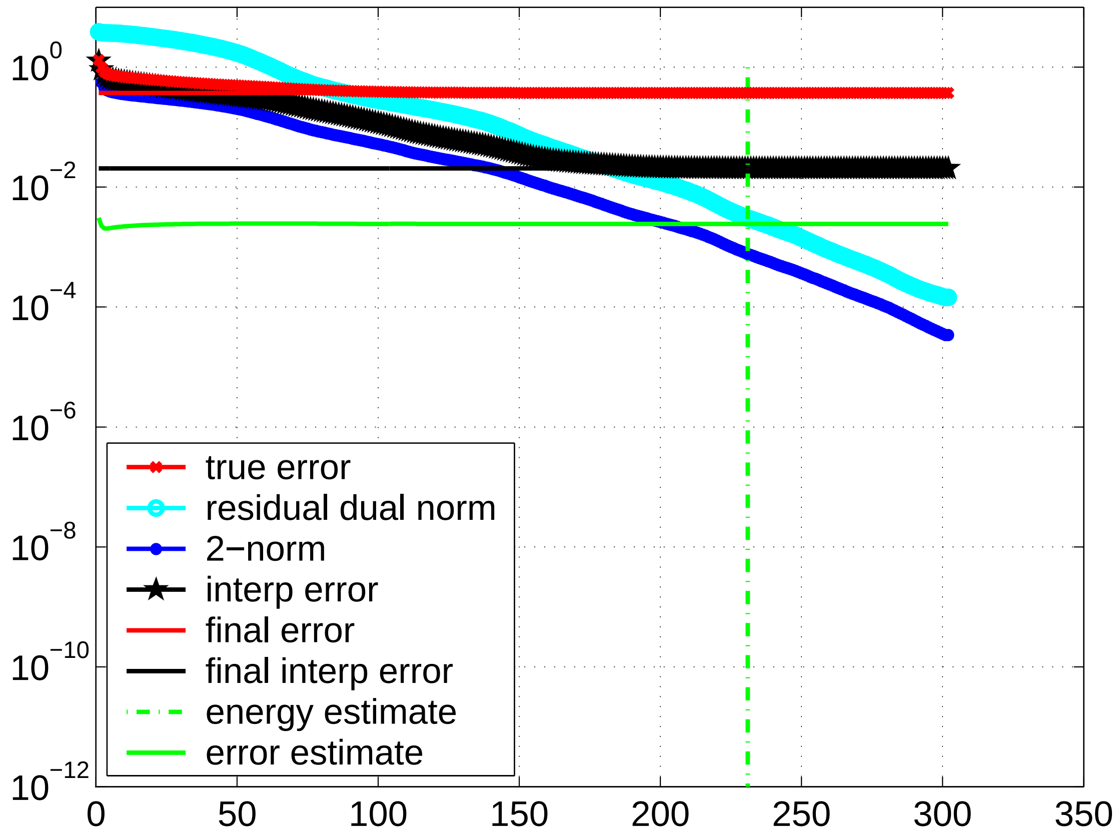
$$\|v_h\|^2_{\mathcal{H}_h} = \varepsilon |v_h|_1^2 + \sum_{T \in \mathcal{T}^h} \delta_T \|\mathbf{b} \cdot \nabla v_h\|^2_{0,T}$$

Advection-diffusion problem



$$\varepsilon = 10^{-2}$$

Advection-diffusion problem



Uniform mesh; $\varepsilon = 10^{-2}$

Summary, open problems, and other applications

- The natural norm for FEM problems is the residual dual norm;
- Estimation is possible;
- PCG + HS is stable;
- For non-symmetric problems, preconditioning with the norm is useful;
- Reliable estimates for finite element error are required;
- Application to indefinite/saddle-point problems. Go to the talk of Daniel LOGHIN!!!

Summary, open problems, and other applications

- The natural norm for FEM problems is the residual dual norm;
- Estimation is possible;
- PCG + HS is stable;
- For non-symmetric problems, preconditioning with the norm is useful;
- Reliable estimates for finite element error are required;
- Application to indefinite/saddle-point problems. Go to the talk of Daniel LOGHIN!!!
- Block-CG ?
- Least-squares ?

CCLRC
Rutherford Appleton Laboratory

Conclusions

FINAL MESSAGE: DO NOT ACCURATELY COMPUTE THE SOLUTION OF AN INACCURATE PROBLEM

Linear regression and Least-squares

$$\mathbf{A}\mathbf{u} = \mathbf{b} + \mathbf{e}, \qquad \mathbf{e} \in \mathcal{N}(0, \sigma\mathbf{I})$$

$\mathbf{A} \in \mathbf{R}^{M \times N}$, $M \geq N$, and $\mathbf{A}$ full rank.

Least-squares

$$\min_{\mathbf{u}} ||\mathbf{A}\mathbf{u} - \mathbf{b}||$$

Normal equations

$$\mathbf{A}^T\mathbf{A}\mathbf{u} = \mathbf{A}^T\mathbf{b}$$

Linear regression and Least-squares

Backward error

Let $\tilde{\mathbf{u}}$ be an approximation of $\mathbf{u}$.
What is the minimum $\delta\mathbf{b}$ such that

$$\mathbf{A}^T\mathbf{A}\tilde{\mathbf{u}} = \mathbf{A}^T(\mathbf{b} + \delta\mathbf{b})?$$

Linear regression and Least-squares

Backward error

Let $\tilde{\mathbf{u}}$ be an approximation of $\mathbf{u}$.
What is the minimum $\delta\mathbf{b}$ such that

$$\mathbf{A}^T\mathbf{A}\tilde{\mathbf{u}} = \mathbf{A}^T(\mathbf{b} + \delta\mathbf{b})?$$

$$\delta\mathbf{b} = \mathbf{A}(\mathbf{A}^T\mathbf{A})^{-1}R \quad ||\delta\mathbf{b}||_2^2 = ||R||^2_{(\mathbf{A}^T\mathbf{A})^{-1}} \quad R = \mathbf{A}^T\mathbf{A}\tilde{\mathbf{u}} - \mathbf{A}^T\mathbf{b}$$

Linear regression and Least-squares

Backward error

Let $\tilde{\mathbf{u}}$ be an approximation of $\mathbf{u}$.
What is the minimum $\delta\mathbf{b}$ such that

$$\mathbf{A}^T\mathbf{A}\tilde{\mathbf{u}} = \mathbf{A}^T(\mathbf{b} + \delta\mathbf{b})?$$

$$\delta\mathbf{b} = \mathbf{A}(\mathbf{A}^T\mathbf{A})^{-1}R \quad ||\delta\mathbf{b}||_2^2 = ||R||^2_{(\mathbf{A}^T\mathbf{A})^{-1}} \quad R = \mathbf{A}^T\mathbf{A}\tilde{\mathbf{u}} - \mathbf{A}^T\mathbf{b}$$

PCG applied to the normal equations can be stopped with Hestenes-Stiefel rule

$$\text{IF } \|\delta\mathbf{b}\|_2^2 \leq \eta\|\mathbf{A}\mathbf{u}^{(k)} - \mathbf{b}\|_2^2 \text{ THEN STOP },$$

with η s.t.

$$Prob(\|\delta\mathbf{b}\|_2^2 > \|\mathbf{e}\|_2^2) = \eta$$

using the χ^2 distribution.

A. and Gratton work in progress