FEM and Sparse Linear System Solving

Lecture 4, October 13, 2017: Finite element errors

http://people.inf.ethz.ch/arbenz/FEM16

Peter Arbenz
Computer Science Department, ETH Zürich
E-mail: arbenz@inf.ethz.ch

- The finite element method
  - Introduction, model problems.
  - 1D problems. Piecewise polynomials in 1D.
  - 2D problems. Triangulations. Piecewise polynomials in 2D.
  - Variational formulations. Galerkin finite element method.
  - Theory of errors/error estimation.
  - Adaptive mesh refinement
  - Discontinuous Galerkin finite element method.
  - Some problems from fluid mechanics.
- Direct solvers for sparse systems
- Iterative solvers for sparse systems

The Poisson problem

As a *model problem*, let us consider the problem

$$\begin{aligned} -\Delta u(\boldsymbol{x}) &= f(\boldsymbol{x}), \quad \boldsymbol{x} \in \Omega, \\ u(\boldsymbol{x}) &= g_D(\boldsymbol{x}), \quad \boldsymbol{x} \in \Gamma_D = \partial\Omega_D, \\ \frac{\partial u(\boldsymbol{x})}{\partial n} &= g_N(\boldsymbol{x}), \quad \boldsymbol{x} \in \Gamma_N = \partial\Omega_N. \end{aligned} \tag{*}$$

Here, $\partial\Omega = \Gamma_D \cup \Gamma_N$. On Γ_D *Dirichlet* boundary conditions are imposed. On Γ_N *Neumann* boundary conditions are imposed.

Weak form of the model problem

$$\underbrace{\int_\Omega \textbf{grad}\, u \cdot \textbf{grad}\, v \; d\boldsymbol{x}}_{a(u,v)} = \underbrace{\int_\Omega f\, v\; d\boldsymbol{x} + \int_{\Gamma_N} g_N\, v\; d\mathbf{s}}_{\ell(v) := (f,v) + (g_n, v)_{\partial\Omega_N}}$$

Weak forms

With these notations the weak forms of the Poisson problem is

$$\boxed{\begin{array}{l}\text{Find } u \in \mathcal{H}^1_E(\Omega) \text{ such that} \\ \qquad\qquad a(u,v) = \ell(v) \qquad \text{for all } v \in \mathcal{H}^1_{E_0}(\Omega)\end{array}} \tag{1}$$

The discrete problem is now

$$\boxed{\begin{array}{l}\text{Find } u_h \in S^h_E \text{ such that} \\ \qquad\qquad a(u_h,v_h) = \ell(v_h) \qquad \text{for all } v_h \in S^h_0\end{array}} \tag{2}$$

Error

We assume that $S_E^h \subset \mathcal{H}_E^1(\Omega)$. Then $e := u - u_h \in \mathcal{H}_{E_0}^1(\Omega)$ is called the error. From (1) we get

$$a(e, v) = a(u, v) - a(u_h, v) = \ell(v) - a(u_h, v).$$

If we plug in for v an arbitrary $v_h \in S_0^h \subset \mathcal{H}_{E_0}^1(\Omega)$ then we obtain the *Galerkin orthogonality property*

$$a(e, v_h) = a(u - u_h, v_h) = 0 \qquad \text{for all } v_h \in S_0^h \tag{3}$$

That is, the error $e \in \mathcal{H}_{E_0}^1(\Omega)$ is orthogonal to S_0^h with respect to the energy inner product $a(u, v)$.

Notice, that $a(u, u)^{1/2} = \|\mathbf{grad}\, u\|$ is a norm if the Dirichlet boundary is nonempty.

Best approximation property

Theorem:

$$\|\mathbf{grad}\, u - \mathbf{grad}\, u_h\| = \min\{\|\mathbf{grad}\, u - \mathbf{grad}\, v_h\| : v_h \in S_E^h\} \quad (4)$$

Proof: Identical to the 1D case. See Larson–Bengzon.

We will later use this theorem to estimate the error by a clever choice of v_h.

Some inequalities

Definition: $\mathcal{H}^1(\Omega)$-norm.

$$\|u\|_{1,\Omega} := \left(\|u\|^2 + \|\mathbf{grad}\ u\|^2\right)^{1/2} = \int_\Omega \left(u^2 + \left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2\right)^{1/2}$$

Definition: $\mathcal{H}^1_0(\Omega) = \{v \in \mathcal{H}^1(\Omega) \ : \ v|_{\partial\Omega} = 0\}$.

Theorem: (Poincaré inequality)
Let $\Omega \subset \mathbb{R}^2$ be a bounded domain. Then, there is constant $C = C(\Omega)$, such that for any $v \in \mathcal{H}^1_0$,

$$\|v\| \leq C\ \|\mathbf{grad}\ v\| \qquad \text{for all } v \in \mathcal{H}^1_0(\Omega).$$

Proof: Larson–Bengzon, p. 75.

Some inequalities (cont.)

The restriction $v|_{\partial\Omega}$ of v to the boundary $\partial\Omega$ of Ω is called the trace of v.

Theorem: (Trace inequality)
Let $\Omega \subset \mathbb{R}^2$ be a bounded domain with smooth or convex polygonal boundary $\partial\Omega$. Then, there is constant $C = C(\Omega)$, such that for any $v \in \mathcal{H}^1(\Omega)$,

$$\|v\|_{L^2(\partial\Omega)} \leq C \left(\|v\|^2_{L^2(\Omega)} + \|\mathbf{grad}\ v\|^2_{L^2(\Omega)} \right)^{1/2} = C\, \|v\|_{1,\Omega}$$

Proof: Larson–Bengzon, p. 75.

Some inequalities (cont.)

Definition: $\mathcal{H}^2(\Omega)$-norm: $\|u\|_{2,\Omega} := \left(\|u\|_{1,\Omega}^2 + \|D^2u\|^2\right)^{1/2}$
where

$$D^2u = \left(\frac{\partial^2 u}{\partial x^2}, \frac{\partial^2 u}{\partial x \partial y}, \frac{\partial^2 u}{\partial y \partial x}, \frac{\partial^2 u}{\partial y^2}\right)$$

Clearly, $\mathcal{H}^2(\Omega) \subset \mathcal{H}^1(\Omega)$.

Theorem: (Elliptic regularity) Let $\Omega \subset \mathbb{R}^2$ be a bounded convex domain with polygonal boundary or a general domain with smooth boundary. Then, there is a constant $C = C(\Omega)$, such that for any sufficiently smooth function v with $v = 0$ or $\frac{\partial v}{\partial \boldsymbol{n}} = 0$ on $\partial\Omega$,

$$\|D^2 v\|_{L^2(\Omega)} \leq C\|\Delta v\|_{L^2(\Omega)}$$

Proof: K. Eriksson *et al.*: Computational differential equations.

Some inequalities (cont.)

Definition: A sequence of triangular/ tetrahedral grids $\{\mathcal{T}_h\}$ is said to be *quasi-uniform* if there exists a constant $\rho > 0$ such that $\min_{\triangle_k} h_k = \underline{h} \geq \rho h = \rho \max_{\triangle_k} h_k$ for every grid in the sequence.

Theorem: (Inverse estimate) On a quasi-uniform mesh any $v_h \in S_E^h$ satisfies the inverse estimate

$$\|\mathbf{grad}\, v_h\|_{L^2(\Omega)} \leq Ch^{-1}\|v_h\|_{L^2(\Omega)}$$

Proof: Larson–Bengzon, p. 77.

Main convergence theorem

Theorem: If the variational problem (1) is solved using a mesh of linear triangular elements and if a minimal angle condition is satisfied (see below) then there is a constant C_1 such that

$$\|\mathbf{grad}\,(u - u_h)\| \leq C_1 \; h \; \|D^2 u\| \tag{5}$$

where $\|D^2 u\|$ measures the elliptic regularity of the solution, and h is the longest triangle edge in the mesh.

Idea of proof:

1. Replace u_h by the interpolant $\pi_h u$ of u in S_E^h and use best approximation property.
2. Do estimates elementwise using the transformations to the reference triangle.

Proof

First we note that $u \in \mathcal{H}^2(\Omega)$ implies $u \in C^0(\Omega)$, i.e., it makes sense to talk of function values $u(\boldsymbol{x})$.

(In contrast, $v \in \mathcal{H}^1(\Omega)$ may not be continuous. There are notions of boundary traces, e.g. $\|u\|_{\partial\Omega} \leq C\|u\|_{1,\Omega}$.)

Let $\pi_h u \in S_E^h$ that *interpolates* the exact solution at all nodes of the triangulation $\mathcal{T}_h$. Then,

$$\|\mathbf{grad}\,(u - u_h)\|^2 \leq \|\mathbf{grad}\,(u - \pi_h u)\|^2 = \sum_{\triangle_k \in \mathcal{T}_h} \|\mathbf{grad}\,(u - \pi_h u)\|^2_{\triangle_k}. \tag{6}$$

We now use the local-global mapping from the reference triangle $\triangle_*$ to the actual triangle $\triangle_k$.

Lemma 11: $\|\mathbf{grad}\,(u - \pi_h u)\|^2_{\triangle_k} \leq 2\frac{h_k^2}{|\triangle_k|}\|\mathbf{grad}\,(\bar{u} - \pi_h \bar{u})\|^2_{\triangle_*}$
Here $\bar{u}$ is the image of u on $\triangle_*$.

Lemma 12: (without proof)
$\|\mathbf{grad}\,(\bar{u} - \pi_h \bar{u})\|^2_{\triangle_*} \leq C\|D^2(\bar{u} - \pi_h \bar{u})\|^2_{\triangle_*} \equiv C\|D^2\bar{u}\|^2_{\triangle_*}$

Lemma 13: $\|D^2\bar{u}\|^2_{\triangle_*} \leq 8h_k^2\frac{h_k^2}{|\triangle_k|}\|D^2 u\|^2_{\triangle_k}$

Altogether, we have

$$\|\mathbf{grad}\,(u - \pi_h u)\|^2_{\triangle_k} \leq 16Ch_k^2\frac{h_k^4}{|\triangle_k|^2}\|D^2 u\|^2_{\triangle_k} \tag{7}$$

Summing over all elements and maximizing ($h = \max h_k$) almost gives the result; but we still have to discuss what happens with $\frac{h_k^2}{|\triangle_k|}$ which is the quotient of two areas.

Proof of Lemma 11

Let $e_k = u - \pi_h u$ and $\bar{e}_k$ be the mapped function on $\triangle_*$. Then,

$$\begin{aligned}
\|\mathbf{grad}\, e\|^2_{\triangle_k} &= \int_{\triangle_k} \left(\left(\frac{\partial e}{\partial x}\right)^2 + \left(\frac{\partial e}{\partial y}\right)^2 \right) dx\, dy \\
&= \int_{\triangle_*} \left(\left(\frac{\partial \bar{e}}{\partial x}\right)^2 + \left(\frac{\partial \bar{e}}{\partial y}\right)^2 \right) 2|\triangle_k|\, d\xi\, d\eta \\
&= \frac{1}{2|\triangle_k|} \int_{\triangle_*} \left(\left(b_2 \frac{\partial \bar{e}}{\partial \xi} + b_3 \frac{\partial \bar{e}}{\partial \eta}\right)^2 + \left(c_2 \frac{\partial \bar{e}}{\partial \xi} + c_3 \frac{\partial \bar{e}}{\partial \eta}\right)^2 \right) d\xi\, d\eta \\
&\leq \frac{4h_k^2}{2|\triangle_k|} \int_{\triangle_*} \left(\left(\frac{\partial \bar{e}}{\partial \xi}\right)^2 + \left(\frac{\partial \bar{e}}{\partial \eta}\right)^2 \right) d\xi\, d\eta
\end{aligned}$$

Here, we used the Schwarz inequality and $b_i, c_i \leq h_k$. $\square$

Proof of Lemma 13

$$
\begin{aligned}
\|D^2\bar{u}\|^2_{\triangle_*} &= \int_{\triangle_*} \left(\left(\frac{\partial^2 \bar{u}}{\partial \xi^2} \right)^2 + 2 \left(\frac{\partial^2 \bar{u}}{\partial \xi \partial \eta} \right)^2 + \left(\frac{\partial^2 \bar{u}}{\partial \eta^2} \right)^2 \right) \, d\xi \, d\eta \\
&= \int_{\triangle_k} \left(\left(\frac{\partial^2 u}{\partial \xi^2} \right)^2 + 2 \left(\frac{\partial^2 u}{\partial \xi \partial \eta} \right)^2 + \left(\frac{\partial^2 u}{\partial \eta^2} \right)^2 \right) \frac{1}{2|\triangle_k|} \, dx \, dy
\end{aligned}
$$

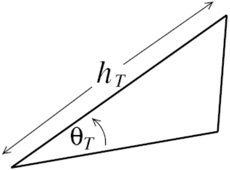
We deal with each of the three (four) terms individually. As an example the first one is treated on the following page.

Proof of Lemma 13 (cont.)

$$\begin{aligned}
\left(\frac{\partial}{\partial \xi}\left(\frac{\partial u}{\partial \xi}\right)\right)^2 &= \left(c_3 \frac{\partial}{\partial x}\left(\frac{\partial u}{\partial \xi}\right) - b_3 \frac{\partial}{\partial y}\left(\frac{\partial u}{\partial \xi}\right)\right)^2 \\
&= \left(c_3^2 \frac{\partial^2 u}{\partial x^2} - c_3 b_3 \frac{\partial^2 u}{\partial x \partial y} - b_3 c_3 \frac{\partial^2 u}{\partial y \partial x} + b_3^2 \frac{\partial^2 u}{\partial y^2}\right)^2 \\
&\leq \left(c_3^4 + 2 c_3^2 b_3^2 + b_3^4\right) \cdot \\
&\quad \cdot \left(\left(\frac{\partial^2 u}{\partial x^2}\right)^2 + \left(\frac{\partial^2 u}{\partial x \partial y}\right)^2 + \left(\frac{\partial^2 u}{\partial y \partial x}\right)^2 + \left(\frac{\partial^2 u}{\partial y^2}\right)^2\right) \\
&= \underbrace{\left(c_3^2 + b_3^2\right)^2}_{\leq 4 h_k^4} \underbrace{\left(\left(\frac{\partial^2 u}{\partial x^2}\right)^2 + 2\left(\frac{\partial^2 u}{\partial x \partial y}\right)^2 + \left(\frac{\partial^2 u}{\partial y^2}\right)^2\right)}_{|D^2 u|^2}. \qquad \square
\end{aligned}$$

Minimum angle condition I

Bounds in Lemmata 11 and 13 involve the *triangle aspect ratio* $\frac{h_k^2}{|\triangle_k|}$. Keeping the aspect ratio small is equivalent to a minimum angle condition



Proposition: For any triangle we have

$$\frac{h_T^2}{4}\sin\theta_T \le |\triangle_T| \le \frac{h_T^2}{2}\sin\theta_T$$

where $0 < \theta_T \le \pi/3$ is the smallest of the interior angles.
Proof. By elementary trigonometrics.

Remark: Similarly, for rectangles there is an aspect ratio condition.

Minimum angle condition II

Combining equations (6) and (7) we get

$$\|\mathbf{grad}\,(u-u_h)\|^2 \leq C \sum_{\triangle_k \in \mathcal{T}_h} \frac{1}{\sin^2\theta_k} h_k^2\, \|D^2 u\|^2_{\triangle_k} \qquad (8)$$

Definition: A sequence of triangular grids $\{\mathcal{T}_h\}$ is said to be *shape regular* if there exists a minimum angle $\theta_* \neq 0$ such that every element in $\mathcal{T}_h$ satisfies $\theta_T \geq \theta_*$.

Then (8) becomes (with $h_k \leq h$)

$$\|\mathbf{grad}\,(u-u_h)\|^2 \leq C(\theta_*)\, h^2 \sum_{\triangle_k \in \mathcal{T}_h} \|D^2 u\|^2_{\triangle_k} = C\, h^2 \|D^2 u\|^2 \qquad (9)$$

L_2 error

In a similar way we can prove that

$$\|u - \pi_h u\|^2 \leq C \sum_{\triangle_k \in \mathcal{T}_h} h_k^4 \|D^2 u\|_{\triangle_k}^2 \qquad (10)$$

Notice, that we do not bound $\|u - u_h\|$.

Remark: In a similar way error estimates can be deduced for the bilinear square (rectangular) element, or, in 3D, for the linear tetrahedron and the trilinear cube.

Errors for higher order elements

Using a higher-order finite element approximation space $\mathbb{P}_m$ or $\mathbb{Q}_m$ with $m \geq 2$ leads to the higher order convergence bound

$$\|\mathbf{grad}\,(u - u_h)\| \leq C_m h^m \|D^{m+1}u\|. \tag{11}$$

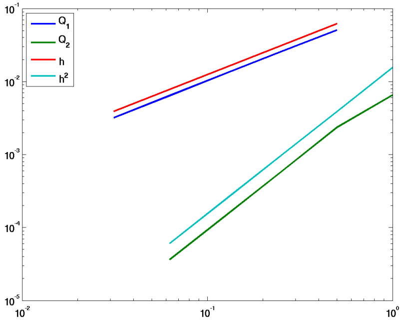
So, we get m-th order convergence as long as the regularity of the solution is good enough. Note that $\|D^{m+1}u\| < \infty$ if and only if all the $m+1$-st derivatives of u are in $L_2(\Omega)$.

The error in the L_2 norm is one order higher, in general.

Numerical tests

Table 1.1 *Energy error E_h for Example 1.1.3: ℓ is the grid refinement level; h is $2^{1-\ell}$ for $\boldsymbol{Q}_1$ approximation, and $2^{2-\ell}$ for $\boldsymbol{Q}_2$ approximation.*

ℓ	$\boldsymbol{Q}_1$	$\boldsymbol{Q}_2$	n
2	5.102×10^{-2}	6.537×10^{-3}	9
3	2.569×10^{-2}	2.368×10^{-3}	49
4	1.287×10^{-2}	5.859×10^{-4}	225
5	6.437×10^{-3}	1.460×10^{-4}	961
6	3.219×10^{-3}	3.646×10^{-5}	3969



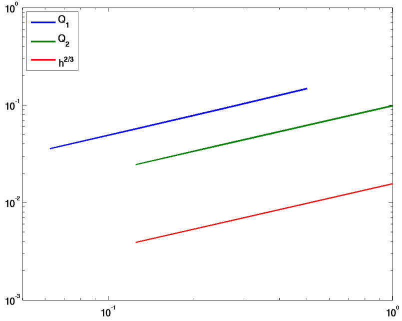
Data from the book by Elman, Silvester and Wathen.
Q_1, Q_2 denote piecewise bilinear/biquadratic square elements.

Numerical tests (cont.)

If the weak solution is not smooth, then the superiority of the $\mathbf{Q}_2$ method over the simpler $\mathbf{Q}_1$ method is not so clear.

Table 1.2 *Energy error E_h for Example 1.1.4.*

ℓ	Q_1	Q_2	n
2	1.478×10^{-1}	9.860×10^{-2}	33
3	9.162×10^{-2}	6.207×10^{-2}	161
4	5.714×10^{-2}	3.909×10^{-2}	705
5	3.577×10^{-2}	2.462×10^{-2}	2945



Numerical tests (cont.)

An appropriately stretched grid of $\mathbf{Q}_2$ elements with 705 degrees of freedom gives better accuracy than that obtained using a uniform grid with 2945 dofs.

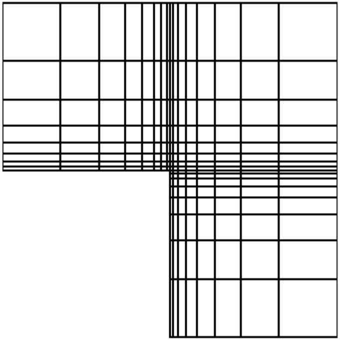


Table 1.3 *Energy error E_h for stretched grid solutions of Example* 1.1.4*:* $\ell = 4$.

α	$\boldsymbol{Q}_1$	$\boldsymbol{Q}_2$
1	9.162×10^{-2}	6.207×10^{-2}
5/4	6.614×10^{-2}	3.723×10^{-2}
3/2	7.046×10^{-2}	2.460×10^{-2}
2	1.032×10^{-1}	2.819×10^{-2}

Properties of the Stiffness Matrix

Theorem: The stiffness matrix A is SPD. □

Theorem: The condition number of the stiffness matrix A satisfies

$$\kappa(A) = \|A\|\|A^{-1}\| = \frac{\lambda_{\max}(A)}{\lambda_{\min}(A)} \leq Ch^{-2}.$$

Proof: Poincaré inequality $\|s\|^2_{L^2(\Omega)} \leq C_1\|\mathbf{grad}\, s\|^2_{L^2(\Omega)}$ in matrix notation:

$$\boldsymbol{x}^T M \boldsymbol{x} \leq C_1\, \boldsymbol{x}^T A \boldsymbol{x}$$

where M is mass matrix and $\boldsymbol{x}$ vector of nodal values of s.
Inverse estimate $\|\mathbf{grad}\, s\|^2_{L^2(\Omega)} \leq C_2 h^{-2}\|s\|^2_{L^2(\Omega)}$ in matrix notation

$$\boldsymbol{x}^T A \boldsymbol{x} \leq C_2 h^{-2} \boldsymbol{x}^T M \boldsymbol{x}.$$

Properties of the Stiffness Matrix (cont.)

Thus,

$$C_1^{-1}\frac{\mathbf{x}^T M\mathbf{x}}{\mathbf{x}^T\mathbf{x}} \le \frac{\mathbf{x}^T A\mathbf{x}}{\mathbf{x}^T\mathbf{x}} \le C_2 h^{-2}\frac{\mathbf{x}^T M\mathbf{x}}{\mathbf{x}^T\mathbf{x}}$$

Larson–Bengzon, Thm. 3.5: there are constants C_3 and C_4 s.t.

$$C_3 h^2 \le \frac{\mathbf{x}^T M\mathbf{x}}{\mathbf{x}^T\mathbf{x}} \le C_4 h^2.$$

Thus, the extremal eigenvalues of A are bounded by

$$(C_3/C_1)h^2 \le \lambda_{\min}(A) \quad \text{and} \quad \lambda_{\max}(A) \le C_2 C_4.$$

Combining this gives

$$\kappa(A) = \frac{C_1 C_2 C_4}{C_3 h^2} \equiv C\, h^{-2}. \quad \square$$

A priori error estimation [redux]

Definition: A sequence of triangular grids $\{\mathcal{T}_h\}$ is said to be *shape regular* if there exists a minimum angle $\theta_* \neq 0$ such that every element in $\mathcal{T}_h$ satisfies $\theta_T \geq \theta_*$.

Proposition: For piecewise linear finite elements, the convergence rate of the gradient of the error is linear

$$\|\mathbf{grad}\,(u - u_h)\|^2 \leq C(\theta_*) \sum_{T \in \mathcal{T}_h} h_T^2 \|D^2 u\|_T^2 = C\, h^2 \|D^2 u\|^2 \qquad (9)$$

where $h = \max h_T$.

Observation: The error can vary much in elements because of the term $\|D^2 u\|_T$.

A posteriori error estimation

- Many interesting problems have singularities or at least behave very differently in various parts of the domain. This motivates the concept of *a posteriori* error estimation.
- Given FE subdivision $\mathcal{T}_h$ and solution u_h: we want to compute a local (element-wise) *error estimator* η_T that approximates the local error in the energy norm $\|\mathbf{grad}\,(u-u_h)\|_T$. Ultimate goal: *modify mesh such that the (estimated) error is distributed equally among the elements*.
- Want to have garanteed accuracy, i.e.,

$$\|\mathbf{grad}\,(u-u_h)\|^2 \equiv \sum_{T\in\mathcal{T}_h} \|\mathbf{grad}\,(u-u_h)\|_T^2 \leq C(\theta_*) \sum_{T\in\mathcal{T}_h} \eta_T^2. \tag{12}$$

- η_T should be cheap to compute.
- η_T should be 'close' to the real error.

Localization of the error

The FE error $e = u - u_h$ is characterized by (Galerkin orthogonality)

$$a(e, v) = \ell(v) - a(u_h, v) \qquad \text{for all } v \in H^1_{E_0}. \tag{3}$$

Now, we assume (for simplicity) that the Neumann boundary conditions are homogeneous (if there are any), i.e., $g_N = 0$. Then we rewrite (3) as

$$\sum_{T \in \mathcal{T}_h} a(e, v)_T = \sum_{T \in \mathcal{T}_h} (f, v)_T - \sum_{T \in \mathcal{T}_h} a(u_h, v)_T \tag{13}$$

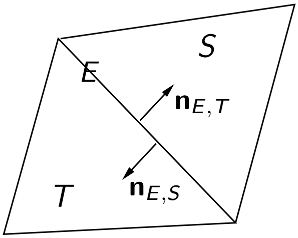
where subscripts T denote element-wise quantities.

Localization of the error (cont.)

We integrate the last term in (13) by parts,

$$a(u_h, v)_T = (\Delta u_h, v)_T - \sum_{E \in \mathcal{E}(T)} (\mathbf{grad}\, u_h \cdot \boldsymbol{n}_{E,T}, v)_E$$

Here, $\mathcal{E}(T)$ denotes the set of edges of element T and $\mathbf{grad}\, u_h \cdot \boldsymbol{n}_{E,T}$ is the outward-pointing normal flux.



Localization of the error (cont.)

Let us define the *flux jump*

$$\left[\frac{\partial v}{\partial n}\right] := (\mathbf{grad}\ v|_T - \mathbf{grad}\ v|_S)\cdot \boldsymbol{n}_{E,T} = (\mathbf{grad}\ v|_S - \mathbf{grad}\ v|_T)\cdot \boldsymbol{n}_{E,S}$$

Then, (13) becomes

$$\sum_{T\in\mathcal{T}_h} a(e,v)_T = \sum_{T\in\mathcal{T}_h}\left[(f+\Delta u_h, v)_T - \frac{1}{2}\sum_{E\in\mathcal{E}(T)}\left(\left[\frac{\partial u_h}{\partial n}\right], v\right)_E\right] \tag{14}$$

e has two distinct components

1. an *element interior residual* $R_T := \{f+\Delta u_h\}|_T$
2. an *inter-element flux jump* $R_E := [\partial u_h/\partial n]$

Estimate of the first term in (14)

We set $v = e - \pi e$. Then, using Cauchy-Schwarz inequality and standard interpolation error estimate:

$$\begin{aligned}(f + \Delta u_h, e - \pi e)_T &\le \|f + \Delta u_h\|_T \|e - \pi e\|_T \\ &\le \|f + \Delta u_h\|_T C h_T \|\mathbf{grad}\,(e - \pi e)\|_T\end{aligned}$$

Note: For piecewise linear triangle elements we have

$$R_T := \{f + \Delta u_h\}|_T = \{f\}_T$$

R_T can be approximated by a constant R_T^0 by projecting f on the piecewise constant functions.

Estimate of the second term in (14)

Here, we use the *scaled* trace inequality

$$\|v\|_{L^2(\partial T)} \le C \left(h_T^{-1} \|v\|^2_{L^2(T)} + h_T \|\mathbf{grad}\, v\|^2_{L^2(T)} \right)^{1/2}.$$

with $v = e - \pi e$:

$$\|e - \pi e\|_{L^2(\partial T)} \le C \left(h_T^{-1} \|e - \pi e\|^2_{L^2(T)} + h_T \|\mathbf{grad}\,(e - \pi e)\|^2_{L^2(T)} \right)^{1/2}$$
$$\overset{\text{Poincaré}}{\le} C \left(h_T \|\mathbf{grad}\,(e - \pi e)\|^2_{L^2(T)} \right)^{1/2} \le C h_T^{1/2} \|\mathbf{grad}\, e\|_{L^2(T)}.$$

$$\sum_{E \in \mathcal{E}(T)} \left(\left[\frac{\partial u_h}{\partial n} \right], e - \pi e \right)_E = \int_{\partial T} \left[\frac{\partial u_h}{\partial n} \right] (e - \pi e)\, ds$$
$$\le \left\| \left[\frac{\partial u_h}{\partial n} \right] \right\|_{\partial T} \|e - \pi e\|_{\partial T} \le C \left\| \left[\frac{\partial u_h}{\partial n} \right] \right\|_{\partial T} h_T^{1/2} \|\mathbf{grad}\, e\|_{L^2(T)}.$$

A posteriori error estimation

Therorem The finite element solution u_h of (1), satisfies the estimate

$$a(u - u_h, u - u_h) = \|\mathbf{grad}\,(u - u_h)\|^2 \leq C \sum_{T \in \mathcal{T}_h} \eta_T^2$$

where the element residual $\eta_T(u_h)$ is defined by

$$\eta_T(u_h) = h_T \|f + \Delta u_h\|_{L^2(T)} + \frac{1}{2} h_T^{1/2} \left\| \left[\frac{\partial u_h}{\partial n} \right] \right\|_{\partial T}$$

Here, $\left[\frac{\partial u_h}{\partial n}\right]$ denotes the jump in the normal derivative of u_h on the (interior) edges of element T. Also, since u_h is linear on T, $\Delta u_h = 0$.

Flux jump operator

We have to distinguish between edges in the interior of Ω, (element transitions), on the Dirichlet boundary, and on the Neumann boundary.
We set

$$R_E^* = \begin{cases} \frac{1}{2}\left[\partial u_h/\partial n\right], & E \text{ in interior of } \Omega, \\ -\mathbf{grad}\, u_h \cdot \boldsymbol{n}_{E,T}, & E \subset \text{on Neumann boundary}, \\ 0, & E \subset \text{Dirichlet boundary}. \end{cases}$$

Remark: In case of nonhomogeneous Neumann boundary conditions, the second term must be modified.

Adaptive mesh refinement

- Mesh refinement in two and three dimensions is much more complicated than in one dimension.
- Important issues to consider:
  1. invalid elements (i.e., with hanging nodes) are not allowed,
  2. refine as few elements as possible,
  3. minimal angle as large as possible.
- For triangle and tetrahedral meshes the most popular algorithms used for mesh refinement are:
  1. Rivara refinement, or, longest edge bisection.
  2. Regular refinement, or, red-green refinement.

  Both approaches available in MATLAB pdetool box.

For more info see Larson & Bengzon.

Adaptive finite elements using MATLAB

Here is a cook book for the procedure in MATLAB.

- Create a (coarse) initial mesh

```
g = 'cardg';     % predefined geometry of a cardioid
[p,e,t] = initmesh(g,'hmax',0.25);
```

- Compute the finite element solution u_h

```
[A,~,b] = assema(p,t,...);
...              %  Introduce the boundary conditions
x = A\b;
```

- Compute element residuals η_T for each triangle.

```
eta = pdejmps(p,t,...);
```

Adaptive finite elements using MATLAB (cont.)

- Select the elements to refine

```
elements = find(eta > tol)';
```

- Do the actual refinement, e.g., using regular refinement

```
[p,e,t] = refinemesh(g,p,e,t,elements,'regular');
```

On the next slide there is code to solve

$$-\Delta u = 1 \text{ in } \Omega, \qquad u = 0 \text{ on } \partial\Omega.$$

MATLAB routines are designed to solve

$$-\mathrm{div}\,(c\,\mathbf{grad}\,u(\boldsymbol{x})) + a\,u(\boldsymbol{x}) = f(\boldsymbol{x})$$

e.g., `[A,M,b] = assema(p,t,c,a,f);`.

The complete MATLAB procedure

```
function AdaptivePoissonSolver2D()
g = 'cardg';                             % set up geometry
[p,e,t] = initmesh(g,'hmax',0.25); % create initial mesh
while size(t,2) < 10000  % while not too many elemts, do
  [A,~,b] = assema(p,t,1,0,1);  % assemble st. matrix A,
                                %       and load vector b
  np = size(p,2);               %     get number of nodes
  fixed = unique([e(1,:) e(2,:)]);  % enforce bndry cond
  free = setdiff([1:np],fixed);
  A = A(free,free);   b = b(free);
  xi = zeros(np,1);
  xi(free) = A\b;                  % solve for FE solution
  eta = pdejmps(p,t,1,0,1,xi,1,1,1); % element residuals
  tol = 0.8*max(eta);             % set selection criteria
  elements = find(eta > tol)';        % elements to refine
              % refine elements using regular refinement
  [p,e,t] = refinemesh(g,p,e,t,elements,'regular');
end
```

Discontinous Finite Element Spaces

Let $\mathcal{K} = \{K\}$ be a mesh of Ω. Define the space of *discontinuous* piecewise linear functions

$$V_h = \{v : v|_K \in \mathbb{P}_1(K),\ \forall K \in \mathcal{K}\} \ \not\subset C^0(\Omega).$$

Let $\mathcal{E}_I$ be set of interior edges.

With each interior edge E we associate a fixed unit normal $\boldsymbol{n}$. We denote by K^+/K^- the element for which $\boldsymbol{n}/-\boldsymbol{n}$ is the exterior normal. (On $\partial\Omega$, $\boldsymbol{n}$ is the exterior unit normal to Ω.)

We define the *jump* and the *average* of a function $v \in V_h$ at the edge E by

$$[v] = v^+ - v^-, \qquad \langle v \rangle = \frac{v^+ + v^-}{2}.$$

Symmetric Interior Penalty Method

Model problem

$$
\begin{aligned}
-\Delta u(\boldsymbol{x}) &= f(\boldsymbol{x}), \qquad && \text{in } \Omega, \\
u(\boldsymbol{x}) &= 0, && \text{on } \partial\Omega.
\end{aligned}
$$

To derive a discontinuous Galerkin (DG) method we multiply the equation with a test function $v \in V_h$. Integration by parts on each element gives

$$
\begin{aligned}
(f, v) &= \sum_{K \in \mathcal{K}} (\mathbf{grad}\, u, \mathbf{grad}\, v)_K - (\boldsymbol{n} \cdot \mathbf{grad}\, u, v)_{\partial K} \\
&= \sum_{K \in \mathcal{K}} (\mathbf{grad}\, u, \mathbf{grad}\, v)_K - \sum_{E \in \mathcal{E}_I} (\boldsymbol{n} \cdot \mathbf{grad}\, u, [v])_E - (\boldsymbol{n} \cdot \mathbf{grad}\, u, v)_{\partial\Omega}
\end{aligned}
$$

Here, we used the continuity of $\boldsymbol{n} \cdot \mathbf{grad}\, u$, i.e. $[\boldsymbol{n} \cdot \mathbf{grad}\, u] = 0$.

Symmetric Interior Penalty Method (cont.)

To make sense of this expression also for $u \in V_h$ we replace the normal flux $\boldsymbol{n} \cdot \mathbf{grad}\, u$ by a discrete counterpart

$$\langle \boldsymbol{n} \cdot \mathbf{grad}\, u \rangle - \beta h^{-1}[u],$$

where β is a positive parameter.

Inserting the discrete flux we arrive at

$$\begin{aligned}(f, v) = \sum_{K \in \mathcal{K}} (\mathbf{grad}\, u, \mathbf{grad}\, v)_K - \sum_{E \in \mathcal{E}_I} (\langle \boldsymbol{n} \cdot \mathbf{grad}\, u \rangle, [v])_E - (\boldsymbol{n} \cdot \mathbf{grad}\, u, v)_{\partial\Omega} \\ + \sum_{E \in \mathcal{E}_I} (\beta h^{-1}[u], [v])_E + (\beta h^{-1} u, v)_{\partial\Omega}\end{aligned}$$

Symmetric Interior Penalty Method (cont.)

The following term is zero when u is the exact solution

$$\sum_{E\in\mathcal{E}_I}([u],\langle \boldsymbol{n}\cdot\mathbf{grad}\ v\rangle)_E + (u,\boldsymbol{n}\cdot\mathbf{grad}\ v)_{\partial\Omega}.$$

Therefore, we can subtract it from the right side to obtain a symmetric form.

$$\begin{aligned}
a_h(u,v) = &\sum_{K\in\mathcal{K}}(\mathbf{grad}\ u,\mathbf{grad}\ v)_K - \sum_{E\in\mathcal{E}_I}(\langle \boldsymbol{n}\cdot\mathbf{grad}\ u\rangle,[v])_E \\
&- (\boldsymbol{n}\cdot\mathbf{grad}\ u,v)_{\partial\Omega} - \sum_{E\in\mathcal{E}_I}([u],\langle \boldsymbol{n}\cdot\mathbf{grad}\ v\rangle)_E - (u,\boldsymbol{n}\cdot\mathbf{grad}\ v)_{\partial\Omega} \\
&+ \sum_{E\in\mathcal{E}_I}(\beta h^{-1}[u],[v])_E + (\beta h^{-1}[u],[v])_{\partial\Omega} \\
l_h(v) = &(f,v),
\end{aligned}$$

Symmetric Interior Penalty Method (cont.)

With the above definitions of $a_h(u, v)$ and $l_h(v)$, the finite element method reads:

$$\boxed{\begin{aligned}&\text{Find } u_h \in V_h \text{ such that}\\ &a_h(u_h, v) = l_h(v), \qquad \text{for all } v \in V_h.\end{aligned}} \tag{15}$$

This dG method is called the Nitsches method or the Symmetric Interior Penalty Galerkin method (SIPG).

This method is consistent, and satisfies the Galerkin orthogonality condition

$$a_h(u - u_h, v) = 0, \qquad \text{for all } v \in V_h. \tag{16}$$

A Priori Error Estimates

For the analysis of the method we define the following energy type norm

$$|||v|||^2 \equiv \sum_{K\in\mathcal{K}} \|\mathbf{grad}\ v\|_K^2 + \sum_{E\in\mathcal{E}_I} h\|\langle \boldsymbol{n}\cdot\mathbf{grad}\ v\rangle\|_E^2 + h\|\boldsymbol{n}\cdot\mathbf{grad}\ v\|_{\partial\Omega}^2$$
$$+ \sum_{E\in\mathcal{E}_I} h^{-1}\|[v]\|_E^2 + h^{-1}\|v\|_{\partial\Omega}^2$$

Then, the error of (15) satisfies the estimate

$$|||u-u_k|||^2 \leq Ch|u|^2_{\mathcal{H}^2(\Omega)}.$$

Implementation issues

In the DG methods the number of degrees of freedom are the number of elements times the number of element degree (assuming these are the same for each element).

It is easy to form the element integrals. They become blocks on the diagonal of A.

The edge flux matrix $S^E = (\langle \boldsymbol{n} \cdot \mathbf{grad}\, \varphi_i \rangle, [\varphi_j])_E$ and the edge penalty matrix $P^E = (\beta h^{-1}[\varphi_i], [\varphi_j])_E$ are more difficult to compute and assemble.

To this end, we traverse all elements and find for each edge the neighboring element. (We have to be careful not to treat an edge twice.)

Implementation issues (cont.)

Let $E = \overline{K^+} \cap \overline{K^-}$ with element basis functions $\varphi_i^{\pm}$, $i = 1, 2, 3$. Then,

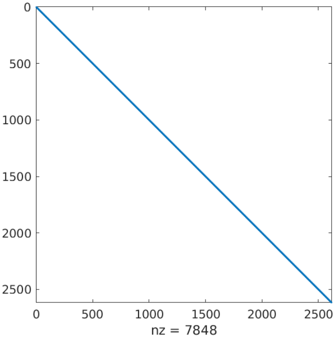
$$\varphi^E = \begin{bmatrix} \varphi_1^+ \\ \varphi_2^+ \\ \varphi_3^+ \\ \varphi_1^- \\ \varphi_2^- \\ \varphi_3^- \end{bmatrix}, \qquad [\varphi^E] = \begin{bmatrix} \varphi_1^+ \\ \varphi_2^+ \\ \varphi_3^+ \\ -\varphi_1^- \\ -\varphi_2^- \\ -\varphi_3^- \end{bmatrix},$$

since all 'hats' $\varphi_i^{\pm}$ are zero outside their associated element $K^{\pm}$. Furthermore,

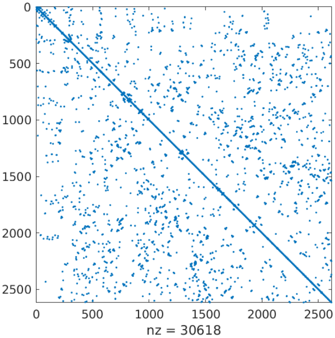
$$\langle \boldsymbol{n} \cdot \mathbf{grad}\, \varphi^E \rangle = \frac{1}{2} \boldsymbol{n}^+ \cdot \mathbf{grad}\, \varphi^E.$$

Line integrals can be computed by the Simpson rule.

Experiments

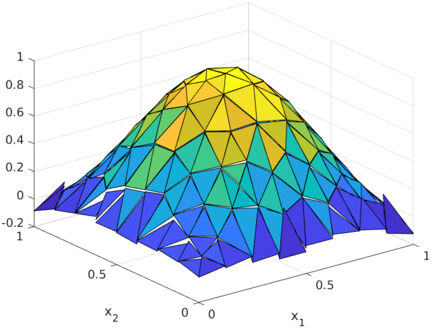


Spy of A

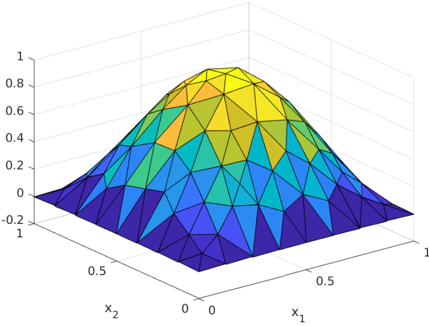


$A - S^E - (S^E)^T + (\beta/h)P^E$

Experiments (cont.)

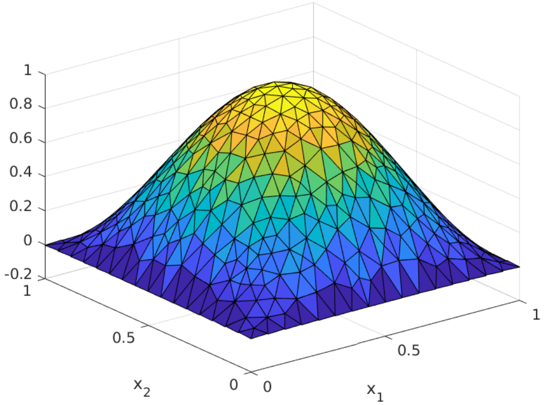


u_h with $\beta = 3$ and $h = 0.125$.



u_h with $\beta = 36$ and $h = 0.125$.

Experiments (cont.)



u_h with $\beta = 9$ and $h = 0.0625$.

Exercise 4:

http://people.inf.ethz.ch/arbenz/FEM17/pdfs/ex4.pdf