1 Week 01: 20 January 2009

These notes will cover the numerical solution of boundary value problems (BVPs) for elliptic partial differential equations (PDEs) by finite element methods. We'll review the basic problems of the subject and the minimum existence theory necessary to make numerical methods worthwhile. Then we'll treat the construction, analysis and selection of numerical methods. We will also learn to use and build software packages which implement effective methods. The recommended text [EG04] for the course is rather dense, and should be read carefully. Notes will be available from the URL `http://www.math.berkeley.edu/~strain`.

2 A one-dimensional example

Let u be a solution of the second-order linear variable-coefficient symmetric two-point boundary value problem

$$Lu := -(pu')' + qu = f \qquad a < x < b$$

with homogeneous mixed boundary conditions

$$u(a) = u'(b) = 0.$$

Problems with nonhomogeneous boundary conditions $u(a) = \alpha$ and $u'(b) = \beta$ can be reduced to this form by subtracting any function satisfying the boundary conditions from u, and modifying the right-hand side f approriately. Higher-dimensional problems often employ similar extension techniques to homogenize boundary conditions. Suppose for simplicity that the leading coefficient p is a positive C^1 function with $p(x) \geq \sigma > 0$ on $[a, b]$, while $q \geq 0$ and f are C^0. (A function $f : [a, b] \to R$ is continuous if $\lim_{x \to x_0} f(x) = f(x_0)$ for all $x \in [a, b]$, and the vector space of all such functions is denoted C^0. A function is in the space C^1 if f and its derivative f' are both in C^0. The spaces C^2, C^3 and so forth are defined similarly.)

We begin by converting the equation to a "dual" form. The idea of duality is very general, and means different things in different contexts. For example, in Euclidean n-space R^n, the primal viewpoint says for example that a vector x is the zero vector 0 iff all its components x_j are 0. The dual viewpoint observes that $x_j = e_j^T x$ where e_j is the jth standard basis vector, and thus that a vector x is 0 iff all other vectors v have $v^T x = 0$. In the context of differential equations, duality regards the solution as a vector in an infinite-dimensional space and "test functions" as the dual vectors which reveal propoerties of the solution. Point values are the components of functions, and the inner product involves integration rather than summation.

Thus multiply the differential equation by any C^1 function v with $v(a) = 0$, integrate by parts and use the boundary conditions to get the variational form

$$\int_a^b pu'v' + quv =: b(u,v) = \int_a^b fv =: (f,v).$$

Here (f,v) denotes the $L^2 = H^0$ inner product on the space

$$L^2 = \{f : [a,b] \to R | \|f\|_0^2 := \int_a^b |f|^2 < \infty\}$$

and b is a bilinear form on the space

$$H_0^1 = \{u \in L^2 | u' \in L^2, u(a) = 0\}.$$

It turns out that b is actually an equivalent inner product on H_0^1, because the associated norm $\sqrt{b(u,u)}$ satisfies an equivalence

$$\epsilon \left((b-a)^2\|u'\|^2 + \|u\|^2\right) \leq b(u,u) \leq C \left((b-a)^2\|u'\|^2 + \|u\|^2\right)$$

with the usual (carefully dimensionalized) H^1 norm

$$\|u\|_1^2 = \left((b-a)^2\|u'\|^2 + \|u\|^2\right).$$

This variety of inner product structures yields many useful features connected with orthogonality, and makes the spaces H^0, H^1 and so forth much more useful than the simpler classical spaces C^0, C^1 and so forth.

Indeed, the Cauchy-Schwarz inequality $(f,v) \leq \|f\|\|v\|$ implies

$$|u(x)|^2 = |u(x) - u(a)|^2 = |\int_a^x u'(x)dx|^2 \leq \int_a^x 1^2 dx \int_a^x |u'(x)|^2 dx \leq (b-a)\|u'\|^2.$$

Integrating over $[a,b]$ and taking square roots gives the simplest "Sobolev inequality"

$$\|u\| \leq (b-a)\|u'\|$$

and a lower bound or "coercivity inequality" guaranteeing that b generates an equivalent norm on H_0^1:

$$b(u,u) \geq \sigma\|u'\|^2 \geq \frac{\sigma}{2(b-a)^2}\left((b-a)^2\|u'\|^2 + \|u\|^2\right) =: \epsilon\|u\|_1^2$$

for all $u \in H_0^1$. Note that without the boundary condition $u(a) = 0$ or some similar way of tying u down, no such inequality is possible.

Next we derive an a priori estimate

$$\|u\|_2^2 := (b-a)^4\|u''\|^2 + (b-a)^2\|u'\|^2 + \|u\|^2 \leq C\|f\|^2.$$

Such estimates characterize elliptic problems, and are essential to the success of classical finite element methods. Indeed, $b(u,u) = (f,u)$, so

$$\epsilon\|u\|_1^2 \leq (f,u) \leq \|f\|\|u\| \leq \|f\|\|u\|_1,$$

so dividing by $\|u\|_1$ gives

$$\|u\|_1 le \frac{1}{\epsilon}\|f\|.$$

On the other hand, the differential equation expresses u'' in terms of lower-order derivatives and the data, so

$$\|u''\|^{\leq}\frac{1}{\sigma}\left(\max|p'|\|u'\| + \max|q|\|u\| + \|f\|\right).$$

Combining these two inequalities bounds the solution in the H^2 norm, in terms of the data and coefficient bounds. This a priori estimate is helpful in guaranteeing the convergence of finite element methods.

Returning to the variational characterization of the solution, we observe that the quadratic form

$$a(v) = b(v,v) - 2(f,v)$$

is minimized over H_0^1 at the solution u. Indeed,

$$\begin{aligned} a(u+w) &= b(u,u) + 2b(u,w) + b(w,w) - 2(f,u) - 2(f,w) \\ &= a(u) + 2(b(u,w) - (f,w)) + b(w,w) \\ &\geq a(u) \end{aligned}$$

since $b(u,w) - (f,w) = 0$ for all $w \in H_0^1$.

The simplest finite element method for this problem minimizes a over a subspace S^h consisting of piecewise linear functions

$$v^h(x) = \sum_{k=1}^{N} q_k \phi_k(x)$$

where x_j are grid points on $a = x_0 < x_2 < \cdots < x_N = b$, and $\phi_k(x)$ are hat functions, piecewise linear functions with $\phi_k(x_j) = \delta_{kj}$ and minimal support $[\max(a, x_{k-1}), \min(b, x_{k+1})]$. The sum starts at $j = 1$ because the functions in the subspace all vanish at $x_0 = a$. An element of the subspace is a continuous piecewise-linear function with values q_k at grid point x_k. Note that this approach corresponds to choosing a globally continuous function and making it satisfy something like the differential equation: the opposite choice is also possible, where we piece together local solutions of the differential equation and make them globally continuous.

This approach dominates the minimum over H_0^1 since S^h is a subspace, and converges to the minimum (if we're lucky) as $h \to 0$. Since S^h is finite-dimensional, it is now a discrete problem: minimize the quadratic form

$$a(v^h) = \sum_{jk} q_j q_k b(\phi_j, \phi_k) - 2 \sum_k q_k (f, \phi_k) = q^T B q - 2 q^T F.$$

Since b is symmetric, so is B, and $b(v^h, v^h) = q^T B q \geq 0$. Since $b(u, u) \geq \epsilon \|u\|_1^2$, the matrix B is positive definite. Thus there is a unique solution to the problem of minimizing a over S^h, given by $Bq = F$ where $u^h = \sum_k q_k \phi_k$. Now we characterize the solution, by the indirect method of the calculus of variations (find conditions that the solution satisfies and solve them instead): the following conditions are equivalent.

1. $a(u^h) = \min a(v^h)$ over $v^h \in S^h$.
2. $b(u^h, v^h) = (f, v^h)$ for all $v^h \in S^h$.
3. $b(u - u^h, v^h) = 0$ for all $v^h \in S^h$.
4. $b(u - u^h, u - u^h) \leq b(u - v^h, u - v^h)$ for all $v^h \in S^h$.

Indeed, for any real number δ

$$a(u^h + \delta v^h) = a(u^h) + 2\delta(b(u^h, v^h) - (f, v^h)) + \delta^2 b(u^h, v^h) \geq a(u^h)$$

for all v^h iff (2) holds. Assuming (2), part (3) is obvious by subtraction, and assuming (3), we have

$$b(u - v^h, u - v^h) = b(u - u^h, u - u^h) + 2b(u - u^h, w^h) + b(w^h, w^h)$$

where $w^h = u^h - v^h$ so by (3), part (4) holds. Thus u^h is the perpendicular projection on S^h of u in the inner product b. The fourth characterization gives us an idea for a convergence proof. Choose v^h close to the exact solution and use it to bound the error.

$$\epsilon \|u - u^h\|_1^2 \leq b(u - u^h, u - u^h) \leq b(u - v^h, u - v^h) \leq C\|u - v^h\|_1^2.$$

Take v^h to be the piecewise linear function which interpolates the exact solution u at the grid points x_j. Let $e = u - v^h$ be the interpolation error. Then

$$\|e\|_1^2 = \sum_{j=1}^{N} (b-a)^2 \int_{x_{j-1}}^{x_j} |e'|^2 + \int_{x_{j-1}}^{x_j} |e|^2.$$

Note that $e(x_j) = 0$ and $e'' = u''$ in between grid points. Let z_j be a point between x_{j-1} and x_j where $e'(z_j) = 0$. Then

$$|e'(x)| = |\int_{z_j}^{x} e''| = |\int_{z_j}^{x} u''| \leq \int_{x_{j-1}}^{x_j} |u''|.$$

Summing up,

$$\int_a^b |e'|^2 \leq h^2 \int_a^b |u''|^2.$$

Since $e(a) = 0$, we also know that

$$\int_a^b |e|^2 \leq (b-a)^2 \int_a^b |e'|^2,$$

so

$$\sqrt{\epsilon} C \|u - u_h\|_1^2 \leq \|e\|_1^2 \leq 2C(b-a)^2 h^2 \|f\|^2$$

where we have used $\|u\|_2 \leq K\|f\|$.

Unfortunately this shows only first-order convergence. To get second-order convergence for the L^2 norm we use Nitsche's trick. Define $e^h = u - u^h \in C^0$ and let c be the solution of $Lc = e^h$. Let c^h be the finite element approximation to c. Then

$$b(c, v) = (e^h, v)$$

for all $v \in H_0^1$, so in particular

$$b(c, e^h) = (e^h, e^h)$$

since $e^h \in H_0^1$. Also we have

$$b(c^h, v^h) = (e^h, v^h)$$

for all $v^h \in S^h$. But by (3) we have $b(c^h, u - u^h) = 0$ so altogether,

$$\begin{aligned}
(e^h, e^h) &= b(c, e^h) \\
&= b(c - c^h, e^h) \\
&\leq \sqrt{b(c - c^h, c - c^h)}\sqrt{b(e^h, e^h)} \\
&\leq \sqrt{Ch^2\|e^h\|^2}\sqrt{Ch^2\|f\|^2} \\
&\leq Ch^2\|e^h\|\|f\|
\end{aligned}$$

giving second-order convergence.

References

[EG04] A. Ern and J.-L. Guermond. *Theory and Practice of Finite Elements.* Springer-Verlag, New York, 2004.