Finite Element Method For Numerically Solving PDE's

STEPHEN ROBERTS
DEPARTMENT OF MATHEMATICS
SCHOOL OF MATHEMATICAL SCIENCES, A.N.U.

Contents

1 Overview of the Finite Element Method

1.1 Introduction

In this section we will describe the Finite Element Method (FEM), a numerical method which provides an efficient and mathematically satisfying method of approximating the solution of elliptic partial differential equations.[1]

We will only give a brief overview of the method in the context of a standard model problem, namely the Poisson equation for 2 dimensional domain Ω.

Hence we will restrict our attention to the model problem:

$$\begin{aligned} -\Delta u &= f \quad \text{in } \Omega \subset \mathbb{R}^2 \\ u &= 0 \quad \text{on } \partial\Omega. \end{aligned} \tag{1}$$

Here to simplify the presentation we will assume that the boundary $\partial\Omega$ is piece-wise linear and convex (see Figure 1).

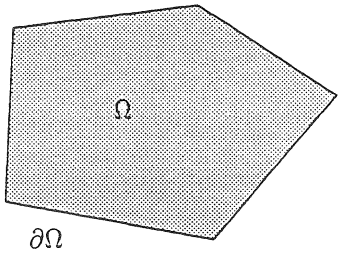

Figure 1: Typical domain Ω for model problem.

The choice of type of domain (convex with Lipschitz boundary) allows us to assume the standard regularity properties[2] for our model equation, i.e.

$$\|u\|_{H^2(\Omega)} \leq C\|f\|_{L^2(\Omega)}.$$

In addition, the choice of piece-wise linear boundary removes a difficulty with approximating the boundary when we "triangulate" our domain to derive our FEM.

Our plan is to give an overview of the techniques used in developing and analyzing a FEM for the approximate solution of our model problem. We break down our discussion into the following topics:

[1] In fact FEM's have been developed to solve parabolic and hyperbolic equations. The methods described here are most easily generalized to the parabolic case. In the hyperbolic case new ideas are necessary to deal with the possibility of discontinuous solutions.

[2] We use the notation $H^m(\Omega)$ to denote the Sobolov space $W^{m,2}(\Omega)$ (see John Urbas' PDE notes).

1. An "equivalent" variational formulation for our model problem is derived (§1.2).

2. A discretization is obtained by restricting the solution and variations to a finite dimensional subspace (§1.3).

3. A bound on the error of the discrete problem is obtained in terms of the approximation properties of the finite dimensional subspace (§1.4).

4. The finite element spaces consisting of piece-wise polynomial functions are introduced (§1.5).

5. The approximation properties of these finite element spaces are investigated (§1.6).

6. The convergence of our finite element method is proved (§1.7).

7. Finally we show how to form a matrix equation which is equivalent to the discrete problem and comment on some of the numerical methods used to solve such matrix problems (§1.8).

Comments on Further Reading

There are many good references on FEM's, but the following give a good range from the Mathematically demanding to the relatively practical. The text by Ciarlet [Cia78] (recently updated and republished as a part of [CL91]) is the classical reference for the mathematical theory of FEM's.

The recent book by Brenner and Scott [BS94] gives a very good mathematical introduction to most aspects of the mathematical theory of FEM's. The results are not as general as those found in Ciarlet's work, though they cover a number of more recent areas (such as Multigrid analysis).

Johnson has produced a very nice text [Joh90], which covers the theoretical and practical issues of FEM's. Johnson not only describes the basic ideas of the mathematical analysis of FEM's for elliptic equations, but also for parabolic and hyperbolic equations. The practical problems of implementing and solving the discrete problems are discussed at length.

1.2 Variational Formulation

A variational formulation of equation (1) can be obtained using the standard method "multiple by a test function and integrate by parts".

Consider $u \in H^2(\Omega) \cap H_0^1(\Omega)$ and $v \in H_0^1(\Omega)$. If u satisfies equation (1), then

$$\int_\Omega -\Delta u \, v \, dx = \int_\Omega f v \, dx.$$

Integration by parts, and the assumption that v vanishes on the boundary gives

$$\int_\Omega \nabla u \cdot \nabla v \, dx = \int_\Omega f v \, dx. \tag{2}$$

Note that u satisfies the "essential" boundary condition $u|_{\partial\Omega} = 0$ since $u \in H_0^1(\Omega)$.

Model Variational Problem

The boundary condition $u|_{\partial\Omega} = 0$ and equation (2) is well defined if $u \in H_0^1(\Omega)$. Hence we are led to the Variational Problem:

Find $u \in H_0^1(\Omega)$ such that

$$\int_\Omega \nabla u \cdot \nabla v \, dx = \int_\Omega f v \, dx$$

for all $v \in H_0^1(\Omega)$.

Abstract Variational Problem:

In fact we can consider more general variational problems.

Consider a Hilbert space V with a symmetric, continuous,[3] coercive,[4] bilinear form $a(\cdot,\cdot)$ on $V \times V$ and a continuous linear functional $F(\cdot)$ on V. The Abstract Variational Problem can be written in the form:

Find $u \in V$ such that

$$a(u,v) = F(v) \quad \text{for all } v \in V.$$

In our model case, the variational space V is the Hilbert space $H_0^1(\Omega)$, the bilinear form $a(\cdot,\cdot)$ is given by

$$a(u,v) = \int_\Omega \nabla u \cdot \nabla v \, dx$$

[3]Continuity of the bilinear form implies there exists a constant C such that $|a(u,v)| \leq C\|u\|_V\|v\|_V$ for all $u, v \in V$.

[4]Coercivity implies the existence of a constant α such that $a(u,u) \geq \alpha\|u\|_V^2$ for all $u \in V$.

and the linear form $F(\cdot)$ by[5]

$$F(v) = \int_\Omega fv \, dx.$$

The conditions on $a(\cdot,\cdot)$ imply that it is an inner product on V. Hence the Riesz Representation Theorem implies that there exists a unique $u \in V$ such that

$$a(u,v) = F(v) \quad \text{for all } v \in V.$$

1.3 Discretization

The discretization of the problem is easily obtained by restricting the solution and variations to a finite dimensional[6] subspace $V^h \subseteq V$.

Discrete Variational Problem:

For a finite dimensional subspace $V^h \subseteq V$ the discrete variational problem is:

Find $u^h \in V^h$ such that

$$a(u^h,v) = F(v) \quad \text{for all } v \in V^h.$$

This variational problem is equivalent to a finite system of linear equations and so is amenable to numerical solution. Using the Riesz Representation Theorem it follows that this finite dimensional problem has a unique solution in V^h.

1.4 Abstract Error Bound

Even without a specific choice of finite dimensional space V^h it is possible to obtain a bound on the error of the approximation. The approximate solution u^h is the projection of the exact solution u onto the subspace V^h, with respect to the inner product induced by the bilinear form $a(\cdot,\cdot)$, since

$$a(u-u^h,v) = 0 \quad \text{for all } v \in V^h.$$

By Pythagoras' theorem

$$a(u-v,u-v) = a(u-u^h,u-u^h) + a(u^h-v,u^h-v)$$

[5]It is left to the reader to verify the properties of $a(\cdot,\cdot)$ and $F(\cdot)$. Hint: Coercivity of $a(\cdot,\cdot)$ follows from the Poincaré inequality.

[6]There are many possible choices of finite dimensional spaces. For instance we could choose polynomial spaces (Chebyshev Polynomials) or trigonometric polynomial spaces. In the case of Finite Element methods, it is common to choose piece-wise polynomial spaces (see §1.5).

for all $v \in V^h$.

Hence we have

$$\|u - u^h\|_E = \inf_{v \in V^h} \|u - v\|_E$$

where $\|v\|_E = a(v,v)^{\frac{1}{2}}$ and is often known as the Energy norm.

In our particular case, the $\|\cdot\|_E$ and $\|\cdot\|_{H^1(\Omega)}$ norms are equivalent, and so we conclude that

$$\|u - u^h\|_{H^1(\Omega)} \leq C \inf_{v \in V^h} \|u - v\|_{H^1(\Omega)}. \tag{3}$$

1.5 Finite Element Spaces

The problem now is to choose finite dimensional spaces in which:

1. The quantity $\inf_{v \in V^h} \|u - v\|_a$ is small or controllable.
2. The space V^h is easy to work with.
3. The numerical problem of solving the finite dimensional linear equation is amenable.

We will consider the case in which the finite dimensional subspace consists of continuous piece-wise polynomial functions[7] with respect to a triangulation $\mathcal{T}$ of the domain Ω (see Figure 2).[8]

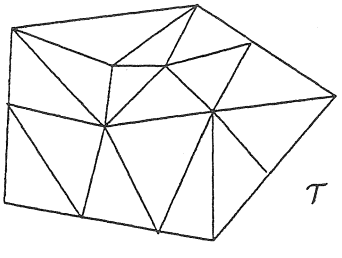

Figure 2: A Typical triangulation of the domain Ω.

A triangulation $\mathcal{T}$ consists of a finite collection of closed triangles ($\subseteq \Omega$), such that:

[7] Alternative finite dimensional spaces can be constructed from piecewise rational functions.

[8] For problems involving higher degree derivatives it is useful to ensure that the piece-wise polynomial spaces have more smoothness. For instance, in the case of the biharmonic equation the space $H^2(\Omega)$ is used as the variational space and $C^1(\Omega)$ piece-wise polynomial spaces are often used as the finite dimensional subspaces.

1. If $K, L \in \mathcal{T}$ then $K \cap L$ is either empty, a common vertex of K and L, a common edge of K and L, or $K = L$.

2. $\bigcup_{K \in \mathcal{T}} K = \bar{\Omega}$.

We define the discretization size h for the triangulation $\mathcal{T}$ by

$$h = \frac{\max\{\text{diam}(K) : K \in \mathcal{T}\}}{\text{diam}(\Omega)}.$$

We will often work with a collection of triangulations $\{\mathcal{T}^h : 0 < h \leq 1\}$ where the notation $\mathcal{T}^h$ is used to explicitly refer to the discretization size of a particular triangulation.

It is necessary to enforce conditions on the triangles of a triangulation to ensure good approximation behaviour of the finite dimensional spaces. So we will suppose that our triangulations $\{\mathcal{T}^h\}$ are non-degenerate in the sense that there exists a $\rho > 0$ such that

$$\text{diam}(B_K) \geq \rho \text{diam}(K)$$

for all triangles K in all triangulations $\{\mathcal{T}^h : 0 < h \leq 1\}$. Here B_K is the largest ball contained in a triangle K. In this situation no triangle in any of the triangulations can become too "long and thin" (see Figure 3).

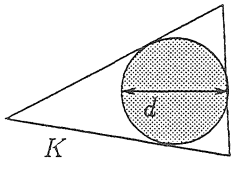

Figure 3: A non-degenerate triangle K with $d \geq \rho \text{diam}(K)$.

For each particular triangulation $\mathcal{T}^h$ we can consider the associated space of continuous piece-wise polynomials of degree $\leq q$. That is we can consider the spaces V^h of continuous functions v defined on Ω such that $v|_K \in P_q$ for all triangles $K \in \mathcal{T}^h$, where P_q denotes the polynomials of at most degree $\leq q$ (see Figure 4).

From now on, whenever we refer to a piece-wise function, we will implicitly mean piece-wise with respect to the associated triangulation.

The choice of sub-regions (triangles in our case) and corresponding finite dimensional function space (piece-wise polynomial functions) constitutes a particular choice of "element" and hence defines the particular finite element space which is being used.

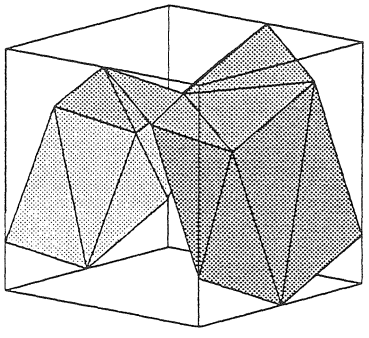

Figure 4: A typical piece-wise affine function.

1.6 Approximation Errors

To obtain error estimates and a convergence theorem for our method, it is important to understand the approximation behaviour of piece-wise polynomial spaces.

We will restrict our discussion to the case of the piece-wise affine space with respect to the triangulation $\mathcal{T}^h$. Let us denote the set of vertices of this triangulation by $\{x_i^h : i = 1, ..., n\}$. We can define an interpolation operator $I^h : H^2(\Omega) \to V^h$ by defining $I^h u$ to be that piece-wise affine function such that $I^h u(x_i) = u(x_i^h)$ for all the vertices x_i^h of our specific triangulation $\mathcal{T}^h$.

It can be shown for $m = 0, 1, 2$ that

$$|u - I^h u|_{H^m(\Omega)} \leq C h^{2-m} |u|_{H^2(\Omega)}.$$

Results of this form can be obtained using weighted Taylor's series (see [BS94, chapt 4]).

We will describe an alternative argument provided by Ciarlet ([CL91, p. 126], [Hac92, p. 185]). We need to derive estimates on individual triangles. So let's be specific and work on the triangle $\hat{K}$ with vertices at $(0,0)$, $(1,0)$ and $(0,1)$. In this case it can be shown that

$$\|v\|_{H^2(\hat{K})} \leq C \left(|v(0,0)| + |v(1,0)| + |v(0,1)| + |v|_{H^2(\hat{K})} \right) \tag{4}$$

for all $v \in H^2(\hat{K})$.

Leaving aside some details, if such a C didn't exist, there would exist a $v \in H^2(\hat{K})$ such that $\|v\|_{H^2(\hat{K})} = 1$, and

$$|v(0,0)| + |v(1,0)| + |v(0,1)| + |v|_{H^2(\hat{K})} = 0.$$

But $|v|_{H^2(\hat{K})} = 0$ implies that $v \in P_1(\hat{K})$. Now $v \in P_1(\hat{K})$ and

$$|v(0,0)| + |v(1,0)| + |v(0,1)| = 0$$

implies $v = 0$, in contradiction to $\|v\|_{H^2(\hat{K})} = 1$.

Choose a $q \in P_1(\hat{K})$ such that $v + q$ is zero at each of the vertices of the triangle. Then by equation (4)

$$\|v + q\|_{H^2(\hat{K})} \leq C|v|_{H^2(\hat{K})}.$$

Consequently

$$\inf_{p \in P_1(\hat{K})} \|v + p\|_{H^2(\hat{K})} \leq C|v|_{H^2(\hat{K})}.$$

Now our interpolation operator satisfies $I^h p = p$ for all $p \in P_1(\hat{K})$. Also for $m = 0, 1, 2$, the operator $v \mapsto v - I^h v$ from $H^2(\hat{K})$ to $H^m(\hat{K})$ is bounded. So for $m = 0, 1, 2$,

$$\begin{aligned} |u - I^h u|_{H^m(\hat{K})} &= \inf_{p \in P_1(\hat{K})} |u + p - I^h(u + p)|_{H^m(\hat{K})} \\ &\leq C \inf_{p \in P_1(\hat{K})} \|u + p\|_{H^2(\hat{K})} \\ &\leq C|u|_{H^2(\hat{K})}. \end{aligned}$$

We can affinely transform this canonical triangle $\hat{K}$ into any triangle K in our triangulation. Provided the triangulation is non-degenerate we can use a scaling argument to show that

$$|u - I^h u|_{H^m(K)} \leq Ch^{2-m}|u|_{H^2(K)}$$

for any $K \in \mathcal{T}^h$. The result for the full domain follows by summing the results for each of the individual triangles.[9]

1.7 Convergence of Method

We can use the abstract error bound (equation (3)) together with the interpolation estimate to conclude the convergence of our FEM.

In particular,

$$\begin{aligned} \|u - u^h\|_{H^1(\Omega)} &\leq C \inf_{v \in V^h} \|u - v\|_{H^1(\Omega)} \\ &\leq C\|u - I^h u\|_{H^1(\Omega)} \\ &\leq Ch\|u\|_{H^2(\Omega)} \\ &\leq Ch\|f\|_{L^2(\Omega)}. \end{aligned}$$

[9] Similar results of the form $|u - I^h u|_{H^m(\Omega)} \leq Ch^{q-m}|u|_{H^q(\Omega)}$ hold for piece-wise polynomial interpolation of degree $q - 1$.

So our method is first order accurate in the $H^1(\Omega)$ norm.

It is natural to also consider convergence in the $L^2(\Omega)$ norm.[10] From the approximation results we would expect to obtain second order accuracy. This turns out to be true in our case.

To prove convergence in the $L^2(\Omega)$ norm we need to use a more complicated argument known as the "Aubin-Nitsche trick" involving a duality argument. We consider the "dual problem" involving the error term as the right hand side of the equation,

$$\begin{aligned} -\Delta w &= u - u^h \quad \text{in } \Omega \subset \mathbb{R}^2 \\ w &= 0 \quad \text{on } \partial\Omega. \end{aligned}$$

In variational form we have

$$a(v, w) = \int_\Omega (u - u^h) v \, dx \quad \text{for all } v \in H_0^1(\Omega). \tag{5}$$

where as usual

$$a(u, v) = \int_\Omega \nabla u \cdot \nabla v \, dx.$$

Note that we have written the variational equation as a dual problem (position of w and v reversed in the equation). Of course this is unnecessary in this symmetric case, but the same argument follows in the un-symmetric case provided the dual problem satisfies the regularity result

$$\|w\|_{H^2(\Omega)} \leq C \|u - u^h\|_{L^2(\Omega)}.$$

Our assumptions on the domain Ω guarantees this regularity result.

Our previous H^1 convergence result implies that

$$\|w - w^h\|_{H^1(\Omega)} \leq Ch \|w\|_{H^2(\Omega)}.$$

Now with the choice $v = u - u^h$ in equation (5) we have

$$\begin{aligned} \|u - u^h\|^2_{L^2(\Omega)} &= a(u - u^h, w) \\ &= a(u - u^h, w - w^h) \\ &\leq C \|u - u^h\|_{H^1(\Omega)} \|w - w^h\|_{H^1(\Omega)} \\ &\leq Ch \|u\|_{H^2(\Omega)} Ch \|w\|_{H^2(\Omega)} \\ &\leq Ch^2 \|u\|_{H^2(\Omega)} \|u - u^h\|_{L^2(\Omega)}. \end{aligned}$$

[10] In many situations (for instance in the analysis of non-linear problems) it is also useful to measure the error in the $L^\infty(\Omega)$ norm. For piece-wise affine spaces in two dimensions it can be shown that $\|u - u^h\|_{L^\infty(\Omega)} \leq Ch^2 |\ln h| \|u\|_{W^{2,\infty}(\Omega)}$. It turns out that the error analysis in this case is more difficult. It is necessary to use so called weighted norms to estimate expression involving "discrete" Green's functions (see [BS94, chapt 7]).

This leads to our result

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

So provided the exact solutions of our model problem satisfy the standard regularity result, our piece-wise affine FEM will provide a method which is first order accurate when measured in the energy norm and second order accurate when measured in the $L^2(\Omega)$ norm.

1.8 Computational Issues

To actually implement a FEM we need to explicitly form some equations. For this we need to find a basis for our finite dimensional spaces.

Let us denote the set of vertices of the triangulation $\mathcal{T}^h$ by $\{x_i^h : i = 1, ..., n\}$ numbered in such a way that $x_i^h \in \partial\Omega$ if $i > N$. That is, there are N interior vertices and $n - N$ boundary vertices.

We can easily identify a basis for V^h when the space consists of continuous piece-wise affine functions which are zero on the boundary. Let ϕ_i^h be the unique continuous piece-wise affine function which satisfies $\phi_i^h(x_j^h) = \delta_{ij}$ (see Figure 5). The set of functions $\{\phi_i^h : i = 1, ..., N\}$ forms a basis for V^h.

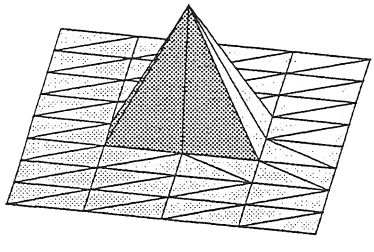

Figure 5: A typical piece-wise affine basis function.

The discrete equation can now be written explicitly as:

Find $u^h = \sum_{j=1}^N u_j^h \phi_j^h$ such that

$$a(u^h, \phi_i^h) = F(\phi_i^h) \quad \text{for all } i = 1, ..., N.$$

This leads to the following system of equations for the coefficients u_j^h.

$$\sum_{j=1}^{N} a(\phi_j^h, \phi_i^h) u_j^h = F(\phi_i^h) \quad \text{for all } i = 1, ..., N.$$

If we define an $N \times N$ matrix A^h via

$$A_{ij}^h = a(\phi_j^h, \phi_i^h)$$

and vectors $U^h = (u_1^h, ..., u_N^h)'$ and $F^h = (F(\phi_1^h), ..., F(\phi_N^h))'$, then the discrete equation is equivalent to the matrix equation

$$A^h U^h = F^h.$$

Of course for our model problem the components of the matrix A^h and the vector F^h are obtained via simple integrations.

So our problem has been reduced to an explicit matrix problem. There are many efficient methods for solving such equations. First it should be observed that the matrix is sparse in the sense that most of the matrix entries are zero. Typically we need to deal with matrices on the order of $10,000 \times 10,000$, but with only 100,000 non-zero entries.

It is possible to use sparse direct methods in this case (Gaussian Elimination in which sparsity is taken into account). Typically it will take $\mathcal{O}(N^2)$ operations to solve the matrix equation using (banded) direct methods (as compared to $\mathcal{O}(N^3)$ operations for a full matrix problem).

Iterative methods can also be employed. An important measure of the difficulty of applying iterative methods to a matrix A is the condition number $\kappa(A)$ of the matrix, define by $\kappa(A) = |\lambda_{max}(A)|/|\lambda_{min}(A)|$, the ratio of the largest to the smallest eigenvalue of the matrix A. Typically we have $\kappa(A^h) = \mathcal{O}(N)$.

The Conjugate Gradient (CG) method is a competitive method for positive definite matrices of this size (A is positive definite). It takes $\mathcal{O}(N\kappa(A)^{\frac{1}{2}})$ operations to obtain a solution of a specified accuracy when using the CG method. Hence the standard CG method will take $\mathcal{O}(N^{\frac{3}{2}})$ operations to obtain a required accuracy.

Multigrid methods provide very efficient methods for solving matrix problems associated with elliptic problems. Multigrid methods work with a heirarchy of triangulations and build up a solution on the finest triangulation by recursively solving problems on progressively coaser triangulations and then smoothly transferring the coarse triangulation solutions back to the finer triangulations. In this way both local and global information about the solution can be obtained in one iterative sweep of the multigrid method. These methods are "optimal" in the

sense that they require only $\mathcal{O}(N)$ operations to obtain a solution to a required accuracy. The tutorial book by Briggs [Bri87] provides a simple introduction to the multigrid method. An introduction to the mathematical analysis of the method can be found in [BS94, chapt. 8].

References

[Bri87] W. L. Briggs. *A Multigrid Tutorial.* SIAM, 1987.

[BS94] S. C. Brenner and L. R. Scott. *The Mathematical Theory of Finite Element Methods*, volume 15 of *Texts in Applied Mathematics.* Springer-Verlag, 1994.

[Cia78] P. G. Ciarlet. *The Finite Element Method for Elliptic Problems.* North-Holland, 1978. (paperback edition 1980).

[CL91] P. G. Ciarlet and J. L. Lions. *Handbook of Numerical Analysis*, volume II. North-Holland, 1991.

[Hac92] W. Hackbusch. *Elliptic Differential Equations, Theory and Numerical Treatment*, volume 18 of *Springer Series in Computational Mathematics.* Springer-Verlag, 1992.

[Joh90] C. Johnson. *Numerical Solution of Partial Differential Equations by the Finite Element Method.* Cambridge University Press, Cambridge, 1990.