The Weak Form

In section 5.4 we used the so-called *strong form* of a BVP

$$-\frac{d}{dx}\left(k(x)\frac{du}{dx}\right) = f(x)$$

$$u(0) = u(l) = 0$$

to produce a *weak form* of the same problem.

$$\int_0^l k(x)\frac{du}{dx}\frac{dv}{dx}dx - \int_0^l f(x)\,v\,dx = 0$$

This integral equation must hold true for all functions v in $C_D^2[0,l]$. For convenience below, we introduce a *bilinear form*

$$a(u,v) = \int_0^l k(x)\frac{du}{dx}\frac{dv}{dx}dx$$

and express this problem in terms of that bilinear form and the usual inner product on $C_D^2[0,l]$,

$$(f,v) = \int_0^l f(x)\,v\,dx$$

In this new notation, the problem reduces to the problem of finding the u in $C_D^2[0,l]$ that satisfies

$$a(u,v) = (f,v)$$

for all v in $C_D^2[0,l]$.

The Galerkin Method

The Galerkin method is a method that seeks to construct approximate solutions for the weak form of the BVP. We make an approximation by selecting a subspace V_N of our original vector space $V = C_D^2[0,l]$ and seeking to find the function u_N in V_N that satisfies

$$a(u_N,v) = (f,v)$$

for all functions v in V_N.

Since

$$a(u-u_N,v) = a(u,v) - a(u_N,v) = (f,v) - (f,v) = 0$$

for all v in V_N, we see from an application of the projection theorem that u_N is the function in V_N that is closest to our target function u, at least with respect to the norm

$$\|y\|_a = \sqrt{a(y,y)}$$

Here now is an outline of the full method.

1. Find a basis $\{\varphi_1, \varphi_2, \ldots, \varphi_N\}$ for V_N.

2. Express u_N as a combination of those basis elements.

$$u_N = c_1\,\varphi_1 + c_2\,\varphi_2 + \cdots + c_N\,\varphi_N$$

3. Let v be any function in V_N. We can also write that as a combination of basis elements.

$$v = d_1\,\varphi_1 + d_2\,\varphi_2 + \cdots + d_N\,\varphi_N$$

4. By the linearity of both $a(u_N, v)$ and (f, v) we have that

$$a(u_N, d_1\,\varphi_1 + d_2\,\varphi_2 + \cdots + d_N\,\varphi_N) = (f, d_1\,\varphi_1 + d_2\,\varphi_2 + \cdots + d_N\,\varphi_N)$$

$$d_1\,a(u_N,\varphi_1) + d_2\,a(u_N,\varphi_2) + \cdots + d_N\,a(u_N,\varphi_N) = d_1\,(f,\varphi_1) + d_2\,(f,\varphi_2) + \cdots + d_N\,(f,\varphi_N)$$

5. If we can force $a(u_N,\varphi_j) = (f,\varphi_j)$ for all j we will be done.

6. The requirement in (5) coupled with (2) gives us a list of N problems to solve:

$$a(u_N,\varphi_j) = \sum_{i=1}^{N} c_i\,a(\varphi_i,\varphi_j) = (f,\varphi_j)$$

7. By introducing $K_{i,j} = a(\varphi_i,\varphi_j)$ and $f_j = (f,\varphi_j)$ these N problems can be rewritten as a single matrix equation.

$$K\,\mathbf{c} = \mathbf{f}$$

8. Solving this equation for the vector of coefficients $\mathbf{c}$ allows us to solve for u_N.

Observations about the Galerkin Method

1. Generally speaking, the approach is usually to pick the basis $\{\varphi_1, \varphi_2, \ldots, \varphi_N\}$ first and let the basis determine the subspace V_N.

2. The best basis vectors to use are vectors which are orthogonal with respect to the inner product $a(\ ,\ )$. An orthogonal basis leads directly to a K matrix that is diagonal, which in turn makes the matrix equation $K\,\mathbf{c} = \mathbf{f}$ easiest to solve.

3. If we can't get the basis vectors to be orthogonal, we can at least try to pick vectors that are "sort of" orthogonal in the sense that most of the off-diagonal entries in K are 0. This also leads to matrix equations $K\,\mathbf{c} = \mathbf{f}$ that are somewhat easier to solve.

An Example

Consider the BVP

$$-\frac{d^2}{dx^2} u(x) = f(x)$$

$$u(0) = u(l) = 0$$

We have already tackled this problem via the method of Fourier series. In that method we tried to express the function $u(x)$ as a combination of eigenvectors

$$\varphi_j(x) = \sin\left(\frac{n \pi}{l} x\right)$$

We will use those same vectors as the basis for our space V_N in the Galerkin method. Since $k(x) = 1$, we have that

$$a(u, v) = \int_0^l \frac{d u(x)}{dx} \frac{d v(x)}{dx} \, dx$$

Fortunately, it turns out that the vectors in question are orthogonal with respect to this inner product. This ends up making in the K matrix diagonal. Indeed, the solution we come to

$$u_N(x) = \sum_{n=1}^{N} c_n \varphi_n(x)$$

where

$$c_n = \frac{(f, \varphi_n)}{(\varphi_n, \varphi_n)}$$

is exactly the solution that the Fourier series method would produce. From this we see that the Galerkin method is actually a generalization of the Fourier series method.

Further Examples

More extensive examples are going to require more extensive calculation, so I will now switch over to Mathematica to show further examples of this method.