Journal of Natural Sciences and Mathematics
Qassim University, Vol. 10, No. 2, pp 137-152 (July 2017/Shawwal 1438H.)

A study of some new inequalities for Galerkin method for model diffusion problem and finite element basis

Ali Allahem

Department of Mathematics, College of Sciences,

Qassim University, Saudi Arabia.

E-mail: a.allahem@qu.edu.sa

Abstract. Our work is concerned to Galerkin projection method for numerical solution of partial differential equations (PDEs). In particular those which include polynomial basis function and finite element basis. Polynomial basis function should be avoided and the reason behind that is presented. Also, we will see that Galerkin method plays an important role on integrals of functions that can easy be evaluated on the domain. In addition, Galerkin method presents high-order approximation. Smart problems are presented for clarification and signification on their properties, and proof of some tricky inequalities.

Keywords: boundary value problem, Diffusion problem, Galerkin method, Polynomial basis, Finite element basis.

1. Introduction

We introduce the Galerkin method through the classic model diffusion problem in D space dimensions,

$$\frac{-d^2u}{dx^2} = f(x), \quad 0 < x < 1, \quad u(0) = u(1) = 0. \tag{1}$$

We start the interest for purposes of introduction when the dimension is one. We denote u to the exact solution of (1) and $\tilde{u}$ to its numerical solution. This is known as a two points boundary value problem (BVP)[1,2,3]. For example, this is a model of heat diffusion in a rod of length 1, where the end and fixed are at temperature 0, and we heat the rod with a driving term $f(x)$.

A domain is a bounded open set, which identifies the physical setting of the problem. In this case

$$I = (0,1) = \{x: 0 < x < 1\}.$$

In this work, we are interested to the one dimensional problem, hence our domain will be an open interval I =(a,b) $\subset$ $\mathbb{R}$. Also, $\bar{I}$ is closure of I. So if I = (a,b), then the closure $\bar{I}$= [a,b] and boundary ∂I = {a,b}.

Consider $L^2(I)$ as the space of measurable functions on open interval I, with inner product $\langle f, g\rangle = \int_0^1 f(x)g(x)dx$. Hence, $\|f\| = (\int_0^1 f(x)^2dx)^{1/2} < \infty$. We work in the space $L^2(I)$ as a related space rather than set of continuous functions.

Consider the well-known space $L^2(I = (0,1))$ of square integrable functions on I, it is precisely defined and based on Lebesgue integration theory. Also, $L^2(I)$ is a Hilbert space with norm $\|f\| = \sqrt{\langle f, f\rangle}$, $\forall\, f \in L^2(I)$ and inner product $\langle f, g\rangle = \int_0^1 f(x)g(x)dx$. It is remarkable that $\|f\| < \infty$, $\forall\, f$, that comes directly from the definition of the space, where $\int_0^1 f(x)^2dx < \infty$.

We need to assign a meaning to the derivative If of an $f \in L^2(D)$ even though f which may not have a derivative in the traditional sense. Consider $\varphi \in C_c^\infty(I)$ is set of infinitely differentiable functions in I with support of φ and denoted by $supp(\varphi) \subset I$ where

$$supp(\varphi) = \overline{\{x \in I: \varphi(x) \neq 0\}}.$$

We say v is the weak derivative of $u \in L^2(I)$ if

$$\int_D v(x)\varphi(x)dx = -\int_D u(x)\acute{\varphi}(x)dx \qquad \forall\varphi \in C_c^\infty(I) \tag{2}$$

We write $v = \acute{D}u$ "generalized derivative of u". Similarly, we can define $D^{\alpha}u$ as a function v that obeys

$$\int_D v(x)\varphi(x)dx = (-1)^{\alpha}\int_D u(x)\varphi^{(\alpha)}(x)dx \qquad \forall\varphi \in C_c^{\infty}(I)$$

Note that, if $u(x)$ is differentiable almost everywhere (roughly, except at a finite number of points), it is not true that u has a generalized derivative as function $f: Df \to \mathbb{R}$.

Model diffusion problem

We look at ways to reformulating I. There are two important methods.

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

$$\int_0^1 DuD\varphi dx = \int_0^1 f(x)\varphi(x)dx \qquad \forall\varphi \in H_0^1(I)$$

V stands for variation.

(M) Let $F: H_0^1(I) \to \mathbb{R}$ be

$$F(u) = \frac{1}{2}\int_0^1 (Du)^{\frac{1}{2}}dx - \int_0^1 f(x)u(x)dx\,,$$

then find minimum for F over $u \in H_0^1(I)$ [12,11].

We will now state three important theorems that study in the reformulations of the model diffusion problem.

Theorem 1. *Consider $u \in C^2(0,1)$ such that*

$$-u_{xx} = f(x), \quad 0 < x < 1, \quad u(0) = u(1) = 0.$$

Then, u solves (V). Find $u \in H_0^1(I)$ with I =(0,1) such that

$$\int_0^1 u_x\varphi_x dx = \int_0^1 f(x)\varphi(x)dx$$

for all test functions $\varphi \in H_0^1(I)$[4].

Theorem 2. *A solution u of (V) is equivalent to a solution u of (M) with I =(0,1) [4].*

In order to see the existence and uniqueness of solution (V), we develop the Riesz representation. Firstly, we define a linear function on a Hilbert space H such that $l: H \to \mathbb{R}$ where

$$l(\alpha x + \beta y) = \gamma l(x) + \beta l(y),$$

for $x, y \in H$ and $\alpha, \beta \in \mathbb{R}$. Secondly, a linear function l on H is bounded if, for some $k > 0, l(x) \leq k\|x\|$ for any $x \in H$. Equivalently, $l(x) \leq k$ for any $x \in H$ with $\|x\| = 1$.

Theorem 3. *(Riesz Representation) Let H be a Hilbert space. Every bounded linear function l on H can be written as*

$$l(x) = \langle x, y \rangle,$$

for some $y \in H$. Furthermore, y is unique [6,10].

2. Theory

We drive a large class of numerical method for

$$-u_{xx} = f(x), \quad u(0) = u(1) = 0, \tag{3}$$

called the Galerkin method. Recall the weak form and consider $f \in L^2(0,1)$ and $V = H_0^1(0,1)$ find $u \in V$ such that

$$a(u, \varphi) = l(\varphi),$$

where $a(u, v) = \int_0^1 u_x v_x dx$ and $l(\varphi) = \int_0^1 f\ \varphi\ dx$ for all $\varphi \in V$.

The idea is to introduce a finite dimensional subspace V_k of the infinite dimensional space V. Let $\tilde{u}$ be the solution of (3), then find $\tilde{u} \in V_k$such that

$$a(\tilde{u}\ , \varphi) = l(\varphi), \tag{4}$$

for all $\varphi \in V_k$. The $\tilde{u}$ defined the Galerkin approximation [9].

Let V_k is equal to span$\{\psi_1, \psi_2, \dots, \psi_k\}$ where ψ_k are under independent basis function in V. For example, let $\psi_1(x) = x^j(1 - x)$ for $j = 1, 2, \dots, k$. Then $\psi_j(x)$ are smooth and have boundary conditions, and hence $u \in V_k$can be written as

$$u = \sum_{j=1}^{k} \alpha_j\, \psi_j, \tag{5}$$

for some $\alpha \in \mathbb{R}$. We seek a linear system that defines α_j. Thus, we substitute (5) into (4) and we get

$$a(\sum_{j=1}^{k} \alpha_j\, \psi_j, \varphi) = l(\varphi), \quad \forall \psi \in V_k.$$

For the reason a is bilinear, the form becomes

$$\sum_{j=1}^{k} \alpha_j \, a(\psi_j, \varphi) = l(\varphi)$$

Let $\psi_i = \varphi$, then

$$\sum_{j=1}^{k} \alpha_j \, a(\psi_j, \psi_i) = l(\psi_i), \quad i = 1, 2, \dots, k \qquad (6)$$

Define a $k \times k$ matrix A with entries $a_{ij} = a(\psi_j, \psi_i)$ and a vector $b \in \mathbb{R}^k$ with entries $b_i = l(\psi_i)$ and unknown $\underline{x} \in \mathbb{R}^k$ as

$$\underline{x} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_k \end{bmatrix}$$

Then the equation (6) can be written as $A\underline{x} = \underline{b}$. In finite element, A is the stiffness matrix and $\underline{b}$ is the load vector. Let us look at properties of A [13].

Theorem 4. *The stiffness matrix A is symmetric and positive definite.*

Proof. a is symmetric

$$a(u, v) = \int_0^1 u_x v_x dx = \int_0^1 v_x u_x dx = a(v, u)$$

and hence $a_{ij} = a_{ji}$. Also, A is positive definite if

$$\underline{x}^T A \underline{x} > 0, \qquad for \ \underline{x} \neq 0.$$

Thus,

$$\begin{aligned}
\underline{x}^T A \underline{x} &= \sum_{j=1}^{k} \sum_{i=1}^{k} x_j a_{ij} x_i \\
&= \sum_{j=1}^{k} \sum_{i=1}^{k} x_j a(\psi_j, \psi_i) x_i \\
&= \sum_{j=1}^{k} x_j a(\sum_{i=1}^{k} x_i \psi_j, \psi_i) \\
&= a(\sum_{i=1}^{k} x_i \, \psi_i \, , \sum_{j=1}^{k} x_j \psi_j) \\
&= a(\tilde{u}, \tilde{u})
\end{aligned}$$

as $\tilde{u} = \sum_{j=1}^{k} x_j \psi_j$. Recall, $\|u\|_E = a(u, u)^{1/2}$ that is so-called " energy norm" and a norm on V. In particular,

$$\|u\|_E = 0 \iff u = 0,$$

Hence,

$$\begin{aligned}\underline{x}^T A\underline{x} &\Leftrightarrow a(\tilde{u},\tilde{u}) = 0,\\ &\Leftrightarrow \tilde{u} = 0,\\ &\Leftrightarrow x = 0,\end{aligned}$$

as ψ_j are linearly independent. ■

Thus, two consequences,

• A is non-singular and hence Galerkin approximation $\tilde{u}$ is well defined.

• Special algorithm (such Cholesky and Conjugate Gradient) are available to solve.

Consequently, $A\underline{x} = \underline{b}$ is efficiently. Note that, the minimisation from the diffusion problem (M) also leads to numerical method. Now, we look at key approximate property of Galerkin method.

Theorem 5. *(Best Approximation) If u is solution $a(u,\psi) = l(\psi),\quad \forall\psi \in V$ and $\tilde{u}$ solves*

$$a(\tilde{u},\psi_k) = l(\psi_k),\quad \forall\psi_k \in V_k. \tag{7}$$

Then

$$\|u-\tilde{u}\|_E \le \|u-\psi_k\|_E,\ \ \forall\psi_k \in V_k \tag{8}$$

In other words, $\tilde{u}$ is the best way approximating u in V_k (has least error in energy norm) [5].

Proof. As we know that $V_k \subset V$ So, the equation (8) implies

$$a(u,\psi_k) = l(\psi_k),$$

for any $\psi_k \in V_k$. Take difference of (7) and (9)

$$a(u-\tilde{u},\psi_k) = 0.$$

Hence, $u-\tilde{u}$ is orthogonal to the space V_k, with respect to $a(.,.)$.

Now,

$$\begin{aligned}\|u-\tilde{u}\|_E^2 &= a(u-\tilde{u},u-\tilde{u}),\\ &= a(u-\tilde{u},u) - a(u-\tilde{u},\tilde{u})\end{aligned}$$

Note that by orthogonal property $a(u-\tilde{u},\tilde{u}) = 0$ as $\tilde{u} \in V_k$. Further, for any $\psi \in V$, $a(u-\tilde{u},\psi_k) = 0$. Thus, we can replace last term

$$\begin{aligned}\|u-\tilde{u}\|_E^2 &= a(u-\tilde{u}) - a(u-\psi_k),\\ &= a(u-\tilde{u},u-\psi_k)\end{aligned}$$

Now Cauchy-Schwartz inequality says

$$a(u,v) = \int_0^1 u_x v_x dx \le \|u_x\|.\|v_x\| = \|u\|_E.\|v\|_E.$$

We have

$$\|u-\tilde{u}\|_E^2 \le \|u-\tilde{u}\|_E . \|u-\psi_k\|_E.$$

Divide through to get

$$\|u-\tilde{u}\|_E \le \|u-\psi_k\|_E,$$

for all $\psi_k \in V_k$. ■

Let us look at best choice of V_k.

Polynomial Basis

Assume that $V_k = span\{x^j(x-1): j = 1,2,\dots,k\}$ and note that if $u \in V_k$, then $u(0) = u(1) = 0$, and $u \in C^1(0,1)$. Hence,

$$u \in H_0^1(0,1) = V = \{u \in L^2(0,1): \frac{du}{dx} \in L^2(0,1), u(0) = u(1) = 0\}$$

Let us run through an example for clarification. Consider $f(x) = 1$ and $k = 2$. Then $V_2 = \{x(x-1), x^2(x-1)\}$. Seek $\tilde{u} = \alpha_1 x(x-1) + \alpha_2 x^2(x-1)$ such that $a(\tilde{u},\psi) = l(\psi_2)$ for any $\psi_2 \in V_2$. Let us call $\varphi_1 = x(x-1)$ and $\varphi_2 = x^2(x-1)$, and hence

$$\tilde{u} = \alpha_1\varphi_1 + \alpha_2\varphi_2$$

Substituting

$$a(\alpha_1\psi_1 + \alpha_2\psi_{2,}\varphi_2) = l(\varphi_2),$$
$$\alpha_1 a(\psi_1,\varphi_2) + \alpha_2 a(\psi_{2,}\varphi_2) = l(\varphi_2),$$

Since $\psi_1 = \varphi_1$ and $\psi_2 = \varphi_2$, then

$$\alpha_1 a(\psi_1,\psi_1) + \alpha_2 a(\psi_1\psi_1) = l(\psi_1),$$
$$\alpha_1 a(\psi_1,\psi_2) + \alpha_2 a(\psi_2\psi_2) = l(\psi_2).$$

This is the linear system $A\underline{x} = \underline{b}$ where the stiffness matrix A has entries

$$a_{11} = \int_0^1 (\frac{d\psi_1}{dx})^2 dx = \int_0^1 (2x-1)^2 dx = 1/3$$
$$a_{12} = \int_0^1 (\frac{d\psi_1}{dx}\frac{d\psi_2}{dx}) dx = \int_0^1 (2x-1)(3x^2-2x) dx = 1/6$$
$$a_{21} = a_{12}$$
$$a_{22} = \int_0^1 (\frac{d\psi_2}{dx})^2 dx = \int_0^1 (3x^2-2x)^2 dx = 2/15,$$

The load vector is

$$b_1 = \int_0^1 f(x)\psi_1 dx = \int_0^1 (x^2 - x)dx = -1/6,$$
$$b_2 = \int_0^1 f(x)\psi_2 dx = \int_0^1 (x^3 - x^2)dx = -1/12.$$

Hence, the linear system is

$$\begin{bmatrix} 1/3 & 1/6 \\ 1/6 & 2/15 \end{bmatrix} \cdot \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} = \begin{bmatrix} -1/6 \\ -1/12 \end{bmatrix}$$

This gives $\alpha_1 = -1/2$ and $\alpha_2 = 0$. Thus, $\tilde{u} = -1/2x(x-1)$.

In fact, $\tilde{u} = u$ is the true solution. This can be predicted from Best Approximation Theorem because $u \in V_2$ in the Galerkin approximation picks the best approximation to $u \in V_2$, which is seriously u ($\|u - u\|_E = 0$). It is clear that $u \in V_2$ as $-u_{xx} = 1$, and hence u quadratic.

In general, Polynomial basis functions are to be avoided for two reasons:

- The stiffness matrix is dense ($a_{ij} \neq 0$ for most i, j).

- The condition number of matrix A grows rapidly as the dimension k is increased. It behaves as Hilbert matrix [8].

Finite element Basis

In the model problem (3), we define D =(0,1) into a grid $0 < x_0 < x_1 < \cdots < x_{k+1} = 1$. Let us assume grid is sparse uniformly $x_{j+1} - x_j = h$ for $j = 0,1,\dots,k$. Define basis function $N_j(x)$ which is piecewise linear on grid (on each (x_{j+1}, x_j) to is linear) such that

$$N_j(x) = \begin{cases} 1 & if \quad i = j, \\ 0 & if \quad i \neq j. \end{cases}$$

Note that $N_j(x) \in H_0^1$ for $j = 0,1,\dots,k$ because $N_j(0) = N_j(1) = 0$ and

$$\frac{d}{dx} N_j(x_i) = \begin{cases} 1/h & x_{j-1} < x < x_j \\ -1/h & x_j < x < x_{j+1} \\ 0 & other \end{cases} \tag{9}$$

Thus, $\frac{d}{dx} N_j(x_i) \in L^2(0,1)$ and $N_j(x)$ are linearly independent.

Let $V_k = span\,\{N_1, N_2, \dots, N_k\}$, then V_k has dimension k and is a subspace of $V = H_0^1(0,1)$. We often write $V_h = V_k$ to emphasis grid space h. Note that N_0, N_{k+1} are not included in V_h. Let us construct the stiffness matrix $A = a_{ij}$ where $a_{ij} = a(N_i, N_j)$.

- Case $i = j$:

$a_{ij} = a(N_i, N_j) = \int_0^1 (\frac{dN_i}{dx})^2 dx$ for $i = 0,1,\dots,k$. Thus, by (9) we get

$$a_{ij} = \int_{x_{i-1}}^{x_{i+1}} (\frac{1}{h})^2 dx = 2/h$$

- Case $i = j - 1$:

$$a_{ij} = \int_{x_{i-1}}^{x_{i+1}} (\frac{dN_i}{dx})(\frac{dN_j}{dx}) dx = \int_{x_{i-1}}^{x_{i+1}} (\frac{-1}{h})(\frac{1}{h}) dx = \frac{-1}{h}$$

- Case $i = j + 1$:

$$a_{ij} = a_{ji} = \frac{-1}{h}$$

- Case $|i - j| > 1$:

Note that $N_j(x) = N_j(x) = 0$ for all x. Hence, $a_{ij} = 0$.

Finally, we have

$$A = \frac{-1}{h}\begin{bmatrix} 2 & -1 & & & & \\ -1 & 2 & -1 & & & \\ & -1 & 2 & -1 & & \\ & & -1 & 2 & -1 & \\ & & & -1 & 2 & -1 \\ & & & & -1 & 2 \end{bmatrix}$$

As usual, the matrix is symmetric and positive definite. For the finite element basis, the stiffness matrix is sparse [1]. Next, it is important to show that the finite element solution converges to the weak solution in the limit $h \to 0$.

Convergence

We look at the effect of increasing number k of finite element basis function on the approximation error.

Theorem 6. *Let u be true solution and ũ be Galerkin approximation in* V_h *, then*

$$||u - \tilde{u}||_E \le h||f||$$

Proof. Let u^* denotes to the piecewise linear interpellant of u at grid points $x_0, x_1, \dots, x_{k+1}$. This means u^* is linear on (x_i, x_{i+1}) for $i = 0,1, \dots, k$ and $u^*(x_i) = y(x_i)$ for $i = 0,1, \dots, k$. Note that $u^* \in V_h$, but u^* may not be equal to Galerkin finite element solution $\tilde{u}$. By best approximation

$$||u - \tilde{u}||_E \le ||u - u^*||_E$$

as $u^* \in V_h$. Next step, we estimate $||u - \tilde{u}||_E$ in terms of h. Let $e = u - u^*$, then $e(x_i) = 0$ for $i = 0,1, \dots, k + 1$. By Poincare Inequality,

$$\int_0^1 \varphi^2 dx \le \int_0^1 \varphi_x^2 dx \quad if \quad \varphi(0) = \varphi(1) = 0.$$

Change variable and let $y = x_i + xh$ and hence $dy = hdx$. Thus, when $x = 0$, then $y = x_i$ and when $x = 1$, then $y = x_{i+1}$. So,

$$\int_{x_{i-1}}^{x_{i+1}} \varphi^2 \frac{dy}{h} \leq \int_{x_{i-1}}^{x_{i+1}} h^2 \varphi_y^2 \frac{dy}{h} \leq h \int_{x_{i-1}}^{x_{i+1}} \varphi_y^2 \frac{dy}{h}$$

Consequently,

$$\int_{x_{i-1}}^{x_{i+1}} \varphi^2 \frac{dy}{h} \leq h^2 \int_{x_{i-1}}^{x_{i+1}} \varphi_y^2 dy$$

We can substitute $\varphi = \frac{de}{dx}$. This produces

$$\int_{x_{i-1}}^{x_{i+1}} (\frac{de}{dx})^2 \frac{dx}{h} \leq h^2 \int_{x_{i-1}}^{x_{i+1}} (\frac{d^2e}{dx^2})^2 dx$$

By integrate, $i = 0$ up to $i = k$, we get

$$\int_0^1 (\frac{de}{dx})^2 \frac{dx}{h} \leq h^2 \int_0^1 (\frac{d^2e}{dx^2})^2 dx$$

Notice, L.H.S is $\|e\|_E^2$. Now, $e = u - u^*$ and hence $\frac{d^2e}{dx^2} = \frac{d^2u}{dx^2}$ as u^* piecewise linear. Thus,

$$\|e\|_E^2 \leq h^2 \int_0^1 \left(\frac{d^2u}{dx^2}\right)^2 dx = h^2 \int_0^1 f^2 dx$$

as $-u_{xx} = f$ and finally $\|e\|_E^2 = \|u - \tilde{u}\|_E \leq h\|f\|$. ■

As result, we have linear convergence in $\|.\|_E$ and quadratic convergence in $\|.\|(L^2(0,1)$ norm) [7].

Theorem 7. *[14] Let u be true solution and $\tilde{u}$ be Galerkin approximation in V_h , then*

$$\|u - \tilde{u}\|_E \leq h^2\|f\|$$

Proof. Let ω solve the dual problem

$$\begin{cases} -\dfrac{d^2\omega}{dx^2} = u - \tilde{u} \quad on\ (0,1) \\ \quad \omega(0) = \omega(1) = 0. \end{cases} \tag{10}$$

Then by (10),

$$\|u - \tilde{u}\|^2 = (u - \tilde{u}, u - \tilde{u}) = \left(u - \tilde{u}, -\frac{d^2\omega}{dx^2}\right).$$

Integrate by parts,

$$\int_0^1 (u-\tilde{u})(-\frac{d^2\omega}{dx^2})dx = [u-\tilde{u}\left(-\frac{d\omega}{dx}\right)]_0^1 + \int_0^1 \frac{d}{dx}(u-\tilde{u})\frac{d}{dx}(\omega)dx$$

As $(u-\tilde{u})(x) = 0$ at $x = 0$,

$$\left(u-\tilde{u})(-\frac{d^2\omega}{dx^2}\right) = a(u-\tilde{u}, N).$$

Hence,

$$\|u-\tilde{u}\|^2 = a(u-\tilde{u}, \omega).$$

Introduce $\omega^* \in V_h$ defined to the piecewise linear interpolant of ω at $x_0, x_1, \dots, x_{k+1}$. Recall $a(u-\tilde{u}, \varphi_k)$ for any $\varphi_k \in V_k = V_h$ where $\tilde{u}$ is Galerkin approximation to u (see Best approximation proof). In particular,

$$a(u-\tilde{u}, \omega^*) = 0.$$

with (10) this implies

$$\|u-\tilde{u}\|^2 = a(u-\tilde{u}, \omega-\omega^*) = \int_0^1 \frac{d}{dx}(u-\tilde{u})\frac{d}{dx}(\omega-\omega^*)dx.$$

Apply, Cauchy-Schwarz inequality,

$$\|u-\tilde{u}\|^2 \le a(u-\tilde{u}, \omega-\omega^*) = \int_0^1 \frac{d}{dx}(u-\tilde{u})\frac{d}{dx}(\omega-\omega^*)dx\,. \qquad (11)$$
$$= \|u-\tilde{u}\|^2 . \|\omega-\omega^*\|^2$$

Now,

$$\|\omega-\omega^*\|^2 \le \left\|\frac{d^2\omega}{dx^2}\right\| \le h\|u-\tilde{u}\| \;\; as \;\; \frac{d^2\omega}{dx^2} = u-\tilde{u}, \qquad (12)$$

and hence, $\|u-\tilde{u}\| \le \|u-\tilde{u}\|_E$ by Poincare inequality as $u-\tilde{u} \in H_0^1(0,1)$. Also, we have $\|u-\tilde{u}\| \le h\|f\|$ by previous theorem. Putting all together form (12), we get

$$\|\omega-\omega^*\|^2 \le h\|u-\tilde{u}\| \le h.h\|f\|.$$

By the equation of (11), we get

$$\|u-\tilde{u}\|^2 \le \|u-\tilde{u}\|_E . \|\omega-\omega^*\|^2$$

Thus,

$$\|u-\tilde{u}\|_E \le h^2\|f\|$$

■

Problem

Consider the Boundary Value Problem (BVP):

$$-(p(x)\acute{u}(x))\acute{} + r(x)u(x) = f(x), \qquad (13)$$

on $[a, b]$ with boundary conditions

$$-p(a)\acute{u}(a) + \alpha u(a) = A, \qquad p(b)\acute{u}(b) + \beta u(b) = B,$$

for $\alpha, \beta \geq 0$ and $A, B \in \mathbb{R}$. Assume that

$$p \in C^1(0,1), \;\; r \in C(0,1) \;\; and \;\; f \in L^2(0,1),$$

for some $c_0 > 0$

$$p(x) \geq c_0, \;\; r(x) \geq 0, \qquad \forall x \in [0,1].$$

To show that the weak formulation of the BVP, let $V = H^1(0,1)$. Also, $u \in V$ such that $a(u, \phi) = l(\phi)$ for all test function $\phi \in V$, where

$$a(u, \phi) = \int_a^b \left[-(p(x)\acute{u}(x))\acute{}\acute{\phi}(x) + r(x)u(x)\acute{\phi}(x)\right]dx + \alpha u(a)\phi(a) + \beta u(b)\phi(b),$$

and

$$l(v) = \int_a^b f(x)\phi(x)dx + A\phi(a) + B\phi(b),$$

Multiply by a test function $\phi \in V$ and integrate by parts:

$$\int_a^b \phi(x)\left[-(p(x)\acute{u}(x))\acute{} + r(x)u(x)\right]dx = \int_a^b f(x)\phi(x)dx\,,$$

$$[-(p(x)\acute{u}(x))\phi(x)]_a^b + \int_a^b \left[p(x)\acute{u}(x)\acute{\phi}(x) + r(x)u(x)\phi(x)\right]dx = \int_a^b f(x)\phi(x)dx\,,$$

Apply the boundary conditions to simplify the first term:

$$\begin{aligned}[-(p(x)\acute{u}(x))\phi(x)]_a^b &= p(a)\acute{u}(a)\phi(a) - p(b)\acute{u}(b)\phi(b) \\ &= \alpha u(a)\phi(a) - A\phi(a) - \beta u(b) - B\phi(b) + \beta u(b)\phi(b)\end{aligned}$$

We conclude that the weak form is as follows: find $u \in V$ such that $a(u, \phi) = l(\phi)$ for all $\phi \in V$.

Finite element approximation of the BVP based on piecewise elements on the subdivision $a < x_0 < x_1 < \cdots < x_n = b$. We may assume that the elements are uniform with spacing h. The method gives rise to a set of $n + 1$ equations in $n + 1$ unknowns. To formulate a finite element approximation on the subdivision $a < x_0 < x_1 < \cdots < x_n = b$, we introduce the basis functions $N_0(x), N_1(x), \dots, N_n(x)$, which are the piecewise linear functions on the subdivision with $N_j(x_k)$ for all $k \neq j$ and $N_j(x_j) = 1$. Let $V = span\{N_j(x): j = \{0, \dots, n\}$. The FEM method is to find $u_n \in V_n$. such that

$$a(u_n, \phi) = l(\phi),$$

for all $\phi \in V_h$. Write $u_n(x) = \alpha_0 N_0(x) + \alpha_1 N_1(x) + \cdots + \alpha_n N_n(x)$, for coefficient $\alpha_0, \alpha_1, \dots, \alpha_n$ to be determined. We find the linear system satisfied by α_j: Substitute $\phi(x) = N_j(x)$, then

$$a(u_n, N_n) = \sum_{j=1}^{n} \alpha_j \, a\big(N_j, N_k\big) = l(f, N_k).$$

Let $a_{jk} = a\big(N_j, N_k\big)$ and $b_k = l(f, N_k)$, then we solve

$$A \begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_n \end{bmatrix},$$

where A is $(n+1) \times (n+1)$ matrix with entries a_{ij}. Moreover, the linear system has a unique solution, we show the matrix A is positive definite. To see that assume

$$\boldsymbol{\alpha^T A \alpha} = u_n \, \boldsymbol{A} \, u_n = a(u_n, u_n), \quad \boldsymbol{\alpha} = (\alpha_0, \alpha_1, \dots, \alpha_n)$$

We show $a(u_n, u_n)^{1/2}$defines a norm on V_h. The key is that for u non-zero

$$a(u, u) = \int_0^1 p(x)\acute{u}(x)^2 + r(x)u(x)^2 dx + \alpha u(a)^2 + \beta u(b)^2 > 0$$

as $\alpha, \beta \geq 0$ and $p(x) \geq c_0 \geq 0$ and $r(x) \geq 0$. Hence, we have the condition that $a(u_n, u_n) > 0$ when u_n is non-trivial. Then if $\boldsymbol{\alpha} \neq \boldsymbol{0}$, $\boldsymbol{\alpha^T A \alpha} > \boldsymbol{0}$ and the matrix A is positive definite and hence non-singular. Consequently, the Galerkin system has a unique solution. Furthermore, The matrix is symmetric as $a(u, v) = a(v, u)$ and tridiagonal because the support of $N_i(x)$ is $[x_{i-1}, x_{i+1}] \cap [a, b]$. Hence, the product $N_i(x)N_j(x)$is zero if $|i - j| > 1$.

If we consider the Boundary Value Problem (13) on [0,1] with boundary conditions

$$u(0) = u(1) = 0$$

for $\alpha, \beta \geq 0$ and $A, B \in \mathbb{R}$. Assume that

$$p \in C^1(0,1), \;\; r \in C(0,1) \;\; and \;\; f \in L^2(0,1),$$

and for some $c_0 > 0$

$$p(x) \geq c_0, \;\; r(x) \geq 0, \qquad \forall x \in [0,1]$$

Given that u^h denotes the piecewise linear finite element approximation to u on uniform elements of width h, it is easy to show that

$$\|u - u^h\|_{H_0^1(0,1)} \leq C_1 h \|u''\|,$$

with $C_1^2 = \frac{1}{c_0}[\|p\|_\infty + h^2 \|r\|_\infty]$. The best approximation property asserts that

$$a(u-u^h,u-u^h) \leq a(u-\phi,u-\phi),$$

for any $\phi \in V^h$. This is a generalisation of the property derived in class for the diffusion problem. Let v^h denote the piecewise linear interpolant to u at grid points x_j . With $e = u - v^h$, we have

$$\|e'\| \leq h\|e_{xx}\|, \qquad \|e\| \leq h^2\|e_{xx}\|.$$

We use $\|u\|_\infty$ to denote $sup_{0\leq x\leq 1}|u(x)|$. Hence,

$$\begin{aligned} a(e,e) &= \int_0^1 p(x)\acute{u}(x)^2 + r(x)u(x)^2 dx, \\ &\leq \|p\|_\infty\|e'\|^2 + h^2\|r\|_\infty\|e\|^2, \\ &\leq \|p\|_\infty h^2\|e''\|^2 + h^4\|r\|_\infty\|e''\|^2, \\ &= h^2[\|p\|_\infty + h^2\|r\|_\infty]\|e''\|^2. \end{aligned}$$

Now, $e'' = u''$ as $e = u - u^h$ and u^h is piecewise linear. Thus

$$a(e,e) = h^2[\|p\|_\infty + h^2\|r\|_\infty]\|u''\|^2$$

Also, note

$$\begin{aligned} a(u-u^h,u-u^h) &= \int_0^1 p(x)(u'(x)-u^{h'}(x))^2 + r(x)(u(x)u^h(x))^2 dx \\ &\geq c_0\|u-u^h\|^2{}_{H_0^1(0,1)} \end{aligned}$$

Hence,

$$\|u-u^h\|^2{}_{H_0^1(0,1)} \leq \frac{1}{c_0}[\|p\|_\infty + h^2\|r\|_\infty]h^2\|u''\|^2$$

Thus, we have proved the result with $C_1^2 = \frac{1}{c_0}[\|p\|_\infty + h^2\|r\|_\infty]$ (assuming $h \geq 1$). Consequently, the following inequality

$$c_0\|u'\|^2 \leq \|f\|\|u\|$$

and

$$\|u''\| \leq \left\|\frac{p'}{p}\right\|_\infty \|f\| + \left\|\frac{r}{p}\right\|_\infty \|f\| + \left\|\frac{1}{p}\right\| \|f\|$$

hold. To see that, from $-(pu')' + ru = f$, we get $-pu'' - p'u' + ru = f$ and

$$\begin{aligned} \|u''\| &= \left\|\frac{p'}{p}u' - \frac{r}{p}u + \frac{1}{p}f\right\|, \\ &\leq \left\|\frac{p'}{p}u'\right\| - \left\|\frac{r}{p}u\right\| + \left\|\frac{1}{p}f\right\|, \\ &\leq \left\|\frac{p'}{p}\right\|_\infty \|u'\| - \left\|\frac{r}{p}\right\|_\infty \|u\| + \left\|\frac{1}{p}\right\|_\infty \|f\|. \end{aligned}$$

From the weak form,

$$\int_0^1 p(x)\acute{u}(x)^2 + r(x)u(x)^2 dx = \int_0^1 f(x)u(x)dx$$

Using Poincare's inequality (as $u \in H_0^1(0,1)$) and Cauchy-Schwarz and $p(x) \geq c_0$, we get $c_0\|u'\|^2 \leq \|f\|\|u\| \leq \|f\|\|u'\|$ and hence

$$\|u\| \leq \|u'\| \leq \frac{1}{c_0}\|f\|$$

Further, we can show that

$$\|u - u^h\|_{H_0^1(0,1)} \leq C_2 h\|f\|,$$

where C_2 should be specified. Now,

$$\|u''\| \leq \left\|\frac{p'}{p}\right\|_\infty \frac{1}{c_0}\|f\| - \left\|\frac{r}{p}\right\|_\infty \|f\| + \left\|\frac{1}{p}\right\|_\infty \|f\|$$

Hence,

$$C_2 = C_1\left(\frac{1}{c_0}\left\|\frac{p'}{p}\right\|_\infty + \frac{1}{c_0}\left\|\frac{r}{p}\right\|_\infty + \left\|\frac{1}{p}\right\|_\infty\right)$$

Now, we provide an example to calculate the right hand side in these inequality such that $(x) = 1, r(x) = 0, f(x) = 1$ and $h = 10^{-3}$. So, we get $c_0 = 1, \|r\|_\infty = 0$ and $\|p\|_\infty = 0$. Hence, $C_1 = 1$ and $C_1 = C_2 = 1$. Hence

$$\|u\|_{H_0^1(0,1)} \leq h = 10^{-3}.$$

Conclusion

We have seen that Galerkin method plays an important role on integrals of functions that can easy be evaluated on the domain. We showed that Galerkin method provides high-order approximation. Tricky problem presented with proof of some smarted inequalities for clarification and signification on it properties.

Acknowledgement

The author is supported by Qassim University and the Ministry of Education of Saudi Arabia.

References

[1] Philippe G Ciarlet. The finite element method for elliptic problems. SIAM, 2002.

[2] Thomas JR Hughes. Multiscale phenomena: Green's functions, the Dirichlet-to-Neumann formulation, subgrid scale models, bubbles and the origins of

stabilized methods. Computer methods in applied mechanics and engineering, 127(1-4):387-401, 1995.

[3] Claes Johnson. Numerical solution of partial differential equations by the finite element method. Courier Corporation, 2012. [4] Beatrice Riviere. Discontinuous Galerkin methods for solving elliptic and parabolic equations: theory and implementation, SIAM, 2008.

[5] Augusto Cesar Galeao and Eduardo Gomes Dutra Do Carmo. A consistent approximate upwind Petrov-Galerkin method for convection- dominated problems. Computer Methods in Applied Mechanics and Engineering, 68(1):83-95, 1988.

[6] Ivar Stakgold. Boundary Value Problems of Mathematical Physics: Volume 1, SIAM, 2000.

[7] Raytcho Lazarov, Sergey Repin, and Satyendra K Tomar. Functional a posteriori error estimates for discontinuous Galerkin approximations of elliptic problems. Numerical Methods for Partial Differential Equations, 25(4):952-971, 2009. 16

[8] Ted Belytschko, Yun Yun Lu, and Lei Gu. Element-free Galerkin methods. International journal for numerical methods in engineering, 37(2):229-256, 1994.

[9] Erik Burman and Peter Hansbo. Edge stabilization for Galerkin approximations of convection-diffusion-reaction problems, Computer Methods in Applied Mechanics and Engineering, 193(15):1437-1453, 2004.

[10] Cem Celik and Melda Duman. Crank-Nicolson method for the fractional diffusion equation with the Riesz fractional derivative, Journal of Computational Physics, 231(4):1743-1750, 2012.

[11] Konstantin Lipnikov, Mikhail Shashkov, Daniil Svyatskiy, and Yu Vassilevski. Monotone finite volume schemes for diffusion equations on unstructured triangular and shape-regular polygonal meshes. Journal of Computational Physics, 227(1):492-512, 2007.

[12] J Tinsley Oden, Ivo Babuska, and Carlos Erik Baumann. A discontinuous hpfinite element method for diffusion problems. Journal of computational physics, 146(2):491-519, 1998.

[13] AJ Wathen. Realistic eigenvalue bounds for the Galerkin mass matrix. IMA Journal of Numerical Analysis, 7(4):449-457, 1987.

[14] Elman, Howard C., David J. Silvester, and Andrew J. Wathen. Finite elements and fast iterative solvers: with applications in incompressible fluid dynamics, Numerical Mathematics & Scientific Computation, 2014.