THE GALERKIN METHOD AND ITS VARIANTS

ADDIS ABABA UNIVERSTIY
COLLEGE OF NATURAL SCIENCE
DEPARTEMENT OF MATHEMATICS

By:
GLORIA DANIEL KENYI

June,2016
Addis Ababa, Ethiopia

THE GALERKIN METHOD AND ITS VARIANTS

ADDIS ABABA UNIVERSTIY
COLLEGE OF NATURAL SCIENCE
DEPARTEMENT OF MATHEMATICS

By:
GLORIA DANIEL KENYI
ADVISOR: TSEGAYE GEDIF(Ph.D)

This project submitted to the department of Mathematics.
Presented in partial fulfilment for the requirements
of master degree of Science in(Differential equations)

June,2016
Addis Ababa, Ethiopia

Addis Ababa University
Department of Mathematics

The undersigned hereby certify that they have read and recommend to the school of graduate studies for acceptance of a project entitled **The Galerkin Method and its variants** by **Gloria Daniel Kenyi** in partial fulfillment of the requirements for the master degree in Differential equations.
Signed by the Examining Committee:

Dr: Tsegaye Gedif	______________	______________
Advisor	signature	Date
Dr: Yibeltal Yitayew	______________	______________
Examiner	signature	Date
Dr: Mengistu Goa	______________	______________
Examiner	signature	Date

____

Chair of Department or Graduate program coordinator

Table of contents

Dedication

I dedicated this project to my Late father Daniel Kenyi and uncle Morris Gore, for their efforts to raise and educate me. I also dedicated this project to my Mother and my husband for all their encouragement and successive help.

____
signature of Author

Acknowledgements

First and foremost I would like to give my thanks to the God who gave me the chance to live and have all such experiences in learning up to this level, which I didn't think of it. Also I would like to thank the Ministry of Higher Education,Sciences and Technology of South Sudan for the opportunity, In this respect. I would like to thank Addis Ababa university, especially mathematics department for their assistance, and appreciation to my advisor Dr: Tsegaye G. For his unreserved help during my project report writing. I am also grateful to my family and especially my husband for all their encouragement and their successive help morally and financially in my study at Addis Ababa University. Finally, I would like to thanks all my friends who comment and help me in all activities.

Abstract

This project is about the Galerkin method and its variants.Its variation are the Petrov-Galerkin method,the generalized Galerkin method and Conjugate gradient method. The Galerkin method provides numerical method to solving elliptic boundary value problem and its application in error estimate of Galerkin solution of linear elliptic boundary value problem. This project also including the numerical solution of a one dimensional poisson with boundary condition Idea of Galerkin discretization Replace V in a(u,v)=$\ell(v) \quad \forall u, v \in V$ with a finite dimensional subspace V_N, is also viewed and some examples are used to illustrates some numerical methods for solving boundary value problems for the Partial Differential Equations(PDEs)to obtained the result.
Key works: Dirichlet Boundary condition - Galerkin method, Poisson equation, Numerical solution using Matlab.

Basic Nations

$\mathbb{R}$	The set of all real numbers.
$\Re^2$	The euclidean space of n- dimensional for $n > 1$
$\mathbb{N}$	The set of all natural numbers.
BVP	Boundary Value Problem
BC	Boundary Condition
FEM	Finite Element Method
PDE	Partial Differential Equation
Ω	open set in R^2
$\partial\Omega$	the boundary of Ω
$\bar{\Omega}$	Closure of Ω
Γ	$=\partial\Omega$
Γ_D	part of the boundary on with Dirichlet condition
Γ_N	part of the boundary on with Neumann condition
$\triangle$	Laplace operator
∇	Gradient operator
$\frac{\partial}{\partial x_i}$	partial derivative with respect to x_i
$\dim(A)$	dimension of A
$\det(A)$	determinant of square matrix A
V_N	finite dimensional subspace
N	mesh or grid size
I	interval in R
I_i	subintervals
x^T	Transpose of the vector $x \in R^n$, $d \in N$
$K(x,t)$	Kernel
K	Kernel
H	Hilbert space
$\langle .,. \rangle$	inner product
$\parallel . \parallel$	norm
$\parallel u \parallel$	$=(u,u)^{\frac{1}{2}}$
$\tilde{u}(x)$	approximation solution

Introduction

Galerkin methods are a class of methods for converting a continuous operator problem (such as a differential equation) to a discrete problem. In principle, it is the equivalent of applying the method of variation, by converting the equation to a weak formulation. Typically one then applies some constraints on the function space to characterize the space with a finite set of basis functions. The approach is usually credited to the Russian mathematician Galerkin but the method was discovered by the Swiss mathematician Walther Ritz [1] to whom Galerkin refers. Often when referring to a Galerkin method, one also gives the name along with typical approximation methods used, such as Galerkin method, Petrov-Galerkin method or Ritz-Galerkin method.A problem in weak formulation Let us introduce Galerkin's method with an abstract problem posed as a weak formulation on a Hilbert space V, namely, find $u \in V$ such that for all $v \in V$, a(u,v) = f(v). Here a(.,.) is a bilinear form (the exact requirements on a(.,.) will be specified later) and f is a bounded linear functional on V. Galerkin method and it is application to computing approximate solution of integral equation the emphasis is on Galerkin method with an orthogonal basis. Introduce the Galerkin method in the framework of Hibert space. Given a computation example that illustrates the importance of choosing the right basis for approximating finite dimensional subspace. Then consider the solution of an integral equation whose exact solution. In chapter two discussed the variational formulations of boundary value problems of partial differential equations in Hilbert spaces. The existence and uniqueness of an appropriately defined weak solution will be discussed. The approximation of this solution with the help of finite-dimensional spaces is called Ritz method or Galerkin method. Some basic properties of this method will be proved. In this chapter, a Hilbert space V will be considered with inner product a(.,.) $:V * V \leftrightarrow R$ and norm $\| v \|_V = a(v,v)^{\frac{1}{2}}$. The finite element method (FEM)is numerical method or technique for solving problems. Which are described by partial differential equation or can be formulation or formulated as functional minimization. A domain of interest is represented as an assembly of finite elements. Approximating functions in finite elements or determined. Finite element method is called also Galerkin method can be applied to many engineering problems that are governed by a differential equation. Need systematic approaches to generate finite element equations include weighted residual method and energy method. Ordinary differential equation, partial differential equation (second-order or fourth-order) can be solve using the weighted residual method in particular using Galerkin method.

Chapter 1

preliminaries

1.1 Definitions

Definition 1.1.1. ***Hilbert space:***_A complete inner product space V is called Hibert space. In other words a Hilbert space is inner product space in which a cauchy sequence is a ways convergent, we will usually denote Hilbert space by H._
Hilbert space is complete Linear space with a scalar product.

Definition 1.1.2. _Let_ V _be a normed space. A sequence_ $\{v_n\} \subset V$ _is called a Cauchy sequence if for any_ $\epsilon > 0$ _there exists a number_ $N(\epsilon)$ _such that_

$$\|v_m - v_n\| < \epsilon \quad \forall\ m,n\ >\ N(\epsilon)$$

Definition 1.1.3. _A normed space is said to be complete if every Cauchy sequence from the space converges to an element in the space. A complete normed space is called a Banach space._

Definition 1.1.4. ***Bounded bilinear form, coercive bilinear form, V-elliptic:***
Bilinear form :
Let a(.,.):V $\times$ _V_ $\to$ _R be a bilinear form on the Banach space V. Then it is bounded if_

$$| b(u,v) | \leq M \parallel u \parallel_V \parallel v \parallel_V \quad \forall u,v \in V \quad M > o$$

$$(Continuity) \tag{1.1}$$

where the constant M is independent of u and v.
The bilinear form is coercive or V-elliptic if

$$a(u,u) \geq m \parallel u \parallel_V^2 \quad \forall u \in V \quad m > o$$

$$(Coercivity) \tag{1.2}$$

where the constant m is independent of u.

Definition 1.1.5. *We say that the bilinear form a satisfies the* $\inf - \sup$*conditions on V if there exists*$\alpha > 0$ *such that*

$$\sup_{v \in V} \frac{a(u,v)}{\parallel v \parallel_V} \geq \alpha \parallel u \parallel_V \quad \forall u \in V and \tag{1.3}$$

$$\sup_{u \in V} \frac{a(u,v)}{\parallel u \parallel_V} \geq \alpha \parallel v \parallel_V \quad \forall v \in V \tag{1.4}$$

***Remark**:*

- *Clearly, if a is symmetric both conditions are the same.*
- *Note that condition (1.3) (and analogously (1.4)) can be written as*

$$\inf_{u \in V} \sup_{v \in V} \frac{a(u,v)}{\parallel u \parallel_V} > 0$$

which justifies the usual terminology.

- *If a is coercive it satisfies the* $\inf - \sup$ *conditions. In fact,*

$$\sup_{u \in V} \frac{a(u,v)}{\parallel v \parallel_V} \geq \frac{a(u,v)}{\parallel u \parallel_V} \geq \alpha \parallel u \parallel_V$$

$$a(u,v) = f(v) \qquad \forall v \in V \tag{1.5}$$

⋆ *set* $V = H_0^1(\Omega)$ *and define*

- *a bilinear form a(.,.) :* $V * V \to R$ *such that*

$$a(u,v) = \int_\Omega \nabla u . \nabla v dx \tag{1.6}$$

- *a Linear functional l:*$V \to R$ *such that*

$$L(v) = \int_\Omega f v dx \tag{1.7}$$

Definition 1.1.6. ***Inner product,(Norm):*** *A vector space V is called an inner product space for every pair of element u∈ V,v∈ V we can define a complex number, denoted by (u,v) and called inner product of u and v, with the following properties.*

- *Linearity:*

$$(\lambda u + \mu v, w) = \lambda(u,w) + \mu(v,w) \qquad for u, w \in V \tag{1.8}$$

And λ *complex. The vector u,v are called orthogonal if (u,v)=0. for every* $u \in V$ *we define the nonnegative number*
$\parallel u \parallel$, *by* $\parallel u \parallel = (u,u)^{1/2}$
is called norm of u.

- *symmetric proptery:*

$$\langle u, v \rangle = \langle v, u \rangle \tag{1.9}$$

and

$$(u,v) = (\overline{v,u}) \qquad \forall u, v \in V \tag{1.10}$$

where the bar denote complex conjugate

- *positive Definite property:for any* $u \in V$

$$\langle u, u \rangle \geq 0, \langle u, u \rangle = 0 \tag{1.11}$$

if and only if u=0

Definition 1.1.7. ***Banach space*** *:*
Is a vector space with a norm such that every cauchy sequence converges to a Limit in the space. We also say the space is complete

Theorem 1.1.1. *(Cauchy-Schwarz Inequality)*
Let V be an inner product space. Then

$$|(u, v)| \leq \|u\|\|v\| \qquad \forall u, v \in V$$

and the equality holds if and only if u and v are linearly dependent.

Theorem 1.1.2. *(Riesz representation theorem)*
Let V be a real or complex Hilbert space, $\ell \in V'$. Then there is a unique $u \in V$ for which

$$\ell(v) = (v, u) \qquad \forall v \in V \tag{1.12}$$

Inaddition,

$$\|\ell\| = \|u\| \tag{1.13}$$

Theorem 1.1.3. *(Lax-Milgram Lemma). Let a(.,.) :V×V → R be a bounded and coercive bilinear form on the Hilbert space V. Then for each bounded linear functional $f \in V'$ there is exactly one $u \in V$ with.*

$$a(u, v) = \ell(v)\forall v \in V \tag{1.14}$$

Chapter 2

Weak Formulation Of Elliptic Boundary Value Problem

The weak formulation are an important tool for the analysis of mathematical equation that permit the transfer of concepts of Linear algebra to solve problem in other field such as partial differential equations. In a weak formulation, an equation is on longer required to hold absolutely(and this is not even well defined) and has instead weak solution only with respect to certain 'test vector ' or ' test function' this is equivalent to formulating the problem to require a solution in the sense of a distribution.

2.1 Review on Partial differential equation

Definition 2.1.1. *The second order partial differential equation is called linear if it is of the form*

$$\sum_{i,j=1}^{n} a_{ij}(x)D_{ij}u + \sum_{i=1}^{n} b_i(x)D_i u + c(x)u + d(x) = 0$$

A partial differential equation and assume that F is differentiable in the η. We extend F to the whole space of $n \times n$ by say $F = (x, z, \xi, \eta) = F(x, z, \xi, \frac{1}{2}(A + A^T))$, where $(x, z, \xi, A) \in \Omega \times \mathbb{R} \times \mathbb{R}^n \times s^{n\times n} \to \mathbb{R}$ the $n \times n$ matrix $[F_{ij}(\varpi)]_{n\times n}$ is symmetric where $F_{ij} := \frac{\partial F}{\partial \eta_{ij}}$

Definition 2.1.2. *The equation is said to be elliptic at a point* $\varpi = (x, z, \xi, \eta) \in \Gamma = \Omega \times \mathbb{R} \times \mathbb{R}^n \times s^{n\times n}$ *if and only if the matrix* $[F_{ij}(\varpi)]_{n\times n}$ *is positive definite, that is* $\sum_{i,j=1}^{n} \frac{\partial F}{\partial \eta_{ij}} \zeta_i \zeta_j > 0$ *for all* $\zeta \in \mathbb{R}^n \backslash \{0\}$.

Definition 2.1.3. *(well-posedness) A problem is said to be well-posed If*

1. *it has a unique solution,*
2. *the solution depends continuously on the given data.*

otherwise the problem is ill-posed.

2.2 A model boundary value problem

Assume that Ω is an open boundary set in R^d and its boundary $\Gamma = \partial\Omega$ is Lipschitz continuous.
We will use $\partial u/\partial n$ to denote the normal derivative of u on Γ. We use the following model boundary value problem as an illustrative example:

$$\begin{cases} -\Delta u = f, & in\ \Omega \\ u = 0, & on\ \Gamma. \end{cases} \tag{2.1}$$

Here Δ denotes the Laplacian operator.
The differential equation in (2.1) is called the Poisson equation. The Poisson equation can be used to describe many physical processes, e.g., steady state heat conduction, electrostatics, deformation of a thin elastic membrane. A classical solution of the problem (2.1) is a smooth function $u \in C^2(\Omega) \cap C(\overline{\Omega})$ which satisfies the differential equation $(2.1)_1$ and the boundary condition $(2.1)_2$ pointwise. Necessarily we have to assume $f \in C(\Omega)$, but this condition or even the stronger condition $f \in C(\overline{\Omega})$ does not guarantee the existence of a classical solution of the problem. One purpose of the introduction of the weak formulation is to remove the high smoothness requirement on the solution and as a result it is easier to have the existence of a (weak) solution.

To derive the weak formulation corresponding to(2.1), we temporarily assume it has a classical solution $u \in C^2(\Omega) \cap C(\overline{\Omega})$. We multiply the differential equation $(2.1)_1$ by an arbitrary function $v \in C_0^\infty(\Omega)$ (so-called smooth test functions), and integrate the relation on Ω,

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

. An integration by parts for the integral on the left side yields (recall that $v = 0$ *on* Γ

$$\int_\Omega \nabla u \cdot \nabla v\, dx = \int_\Omega f v\, dx \tag{2.2}$$

This relation was proved under the assumptions $u \in C^2(\Omega) \cap C(\overline{\Omega})$ and $v \in C_0^\infty(\Omega)$. However, for all the terms in the relation (2.2) to make sense, we only need to require the following regularities of u and v that is $u, v \in H^1(\Omega)$, assuming $f \in L_2(\Omega)$. Recalling the homogeneous Dirichlet boundary condition $(2.1)_2$, we thus seek a solution
$u \in H_0^1(\Omega)$ satisfying the relation (2.2) for any $v \in C_0^\infty(\Omega)$. Since $C_0^\infty(\Omega)$ is dense in $H_0^1(\Omega)$, the relation (2.2) is then valid for any $v \in H_0^1(\Omega)$. Therefore, the weak formulation of the boundary value problem (2.1) is

$$u \in H_0^1(\Omega), \quad \int_\Omega \nabla u \cdot \nabla v\, dx = \int_\Omega f v\, dx \quad \forall v \in H_0^1(\Omega) \tag{2.3}$$

Actually, we can even weaken the assumption, $f \in L_2(\Omega)$. It is enough for us to assume $f \in H^{-1}(\Omega) = (H_0^1(\Omega))'$, as long as we interpret the integral $\int_\Omega f v\, dx$ as the duality pairing $\langle f, v\rangle$ between $H^{-1}(\Omega)$ and $H_0^1(\Omega)$. We adopt the convention of using $\int_\Omega f v\, dx$ for $\langle f, v\rangle$ when $f \in H^{-1}(\Omega)$ and $v \in H_0^1(\Omega)$.
We have shown that if u is a classical solution of (2.1), then it is also a solution of the weak formulation (2.3). Conversely, suppose u is a weak solution with the additional

regularity $u \in C^2(\Omega) \cap C(\overline{\Omega})$ and $f \in C(\Omega)$. Then for any $v \in C_0^\infty(\Omega) \subset H_0^1(\Omega)$ from (2.3) we obtain

$$\int_\Omega (-\Delta u - f) v \, dx = 0$$

Then we must have $-\Delta u = f$ in Ω, i.e., the differential equation $(2.1)_1$ is satisfied. Also u satisfies the homogeneous Dirichlet boundary condition pointwisely.

Thus we have shown that the boundary value problem (2.1) and the variational problem (2.2) are formally equivalent. In case the weak solution u does not have the regularity $u \in C^2(\Omega) \cap C(\overline{\Omega})$. We will say u formally solves the boundary value problem (2.1).

We let $V = H_0^1(\Omega)$, $a(\cdot,\cdot) : V \times V \to \mathbb{R}$ the bilinear form defined by

$$a(u,v) = \int_\Omega \nabla u \cdot \nabla v \, dx \quad for \quad u, v \in V,$$

and $\ell : V \to \mathbb{R}$ the linear functional defined by

$$\ell(v) = \int_\Omega f v \, dx \quad for \quad v \in V.$$

Then the weak formulation of the problem (2.1) is to find $u \in V$ such that

$$a(u,v) = \ell(v) \qquad \forall v \in V \tag{2.4}$$

We define a differential operator A associated with the boundary value problem (2.1) by

$$A : H_0^1(\Omega) \to H^{-1}(\Omega), \quad \langle u, v \rangle = a(u,v) \quad \forall u, v \in H_0^1(\Omega).$$

Here, $\langle \cdot, \cdot \rangle$ denotes the duality pairing between $H_0^1(\Omega)$ and $H^{-1}(\Omega)$. Then the problem (2.4) can be viewed as a linear operator equation

$$Au = \ell \quad in \quad H^{-1}(\Omega)$$

A formulation of the type (2.1) in the form of a partial differential equation and a set of boundary conditions is referred to as a classical formulation of a boundary value problem, whereas a formulation of the type(2.4) is known as a weak formulation. One advantage of weak formulations over classical formulations is that questions related to existence and uniqueness of solutions can be answered more satisfactorily.

2.3 Some general results on existence and uniqueness

We first present some general ideas and results on existence and uniqueness for a linear operator equation of the form

$$u \in V, \quad Lu = f, \tag{2.5}$$

where $L : \mathcal{D}(L) \subset V \to W$, V and W are Hilbert spaces, and $f \in W$. Notice that the solvability of the equation is equivalent to the condition $\mathcal{R}(L) = W$, whereas the uniqueness of a solution is equivalent to the condition $\mathcal{N}(L) = \{0\}$.

A very basic existence result is the following theorem.

Theorem 2.3.1. *Let V and W be Hilbert spaces, $L : \mathcal{D}(L) \subset V \to W$ a linear operator. Then $\mathcal{R}(L) = W$ if and only if $\mathcal{R}(L)$ is closed and $\mathcal{R}(L)^{\perp} = \{0\}$.*

Proof. If $\mathcal{R}(L) = W$, then obviously $\mathcal{R}(L)$ is closed and $\mathcal{R}(L)^{\perp} = \{0\}$. Now assume $\mathcal{R}(L)$ is closed and $\mathcal{R}(L)^{\perp} = \{0\}$, but $\mathcal{R}(L) \neq W$. Then $\mathcal{R}(L)$ is a closed subspace of W. Let $w \in W \backslash \mathcal{R}(L)$. By Theorem(**??**), the compact set $\{w\}$ and the closed convex set $\mathcal{R}(L)$ can be strictly separated by a closed hyperplane, i.e., there exists a $w^* \in W'$ such that $\langle w^*, w \rangle > 0$ and $\langle w^*, Lv \rangle \leq 0$ for all $v \in D(L)$. Since L is a linear operator, $\mathcal{D}(L)$ is a subspace of V. Hence, $\langle w^*, Lv \rangle = 0$ for all $v \in \mathcal{D}(L)$. Therefore, $0 \neq w^* \in \mathcal{R}(L)^{\perp}$. This is a contradiction. □

Example 2.3.1. *As a concrete example, we consider the weak formulation of the model elliptic boundary value problem (2.5). Here, $\Omega \subset \mathbb{R}^d$ is an open bounded set with a Lipschitz boundary $\partial\Omega$, $V = H_0^1(\Omega)$ with the norm $\|v\|_V = |v|_{H^1(\Omega)}$, and $V' = H^{-1}(\Omega)$. Given $f \in H^{-1}(\Omega)$, consider the problem*

$$\begin{cases} -\Delta u = f & \text{in } \Omega \\ u = 0 & \text{on } \partial\Omega \end{cases} \tag{2.6}$$

We define the operator $L : V \to V'$ by

$$\langle Lu, v \rangle = \int_\Omega \nabla u \nabla v \, dx \qquad u, v \in V.$$

Then L is linear, continuous, and strongly monotone; indeed, we have

$$\langle Lv, v \rangle = \|v\|_V^2 \qquad \forall v \in V.$$

Proof.

$$\begin{aligned} \langle L(\alpha u + \beta w), v \rangle &= \int_\Omega \nabla(\alpha u + \beta w) \cdot \nabla v \, dx \\ &= \int_\Omega (\alpha \nabla u + \beta \nabla w) \cdot \nabla v \, dx \\ &= \int_\Omega (\alpha \nabla u \nabla v + \beta \nabla w \nabla v) \, dx \\ &= \alpha \int_\Omega \nabla u \nabla v \, dx + \beta \int_\Omega \nabla w \nabla v \, dx \\ &= \alpha \langle Lu, v \rangle + \beta \langle Lw, v \rangle \\ &= \langle \alpha Lu, v \rangle + \langle \beta Lw, v \rangle \\ &= \langle \alpha Lu + \beta Lw, v \rangle \end{aligned}$$

Thus, $L(\alpha u + \beta w) = \alpha Lu + \beta Lw \quad \forall u,\ v \text{ and } w \in V$ and $\forall \alpha, \beta \in \mathbb{R}$. Hence L is linear. Next, let us show continuity

$$\begin{aligned} \langle Lu, v \rangle &= \int_\Omega \nabla u \cdot \nabla v \, dx \\ &\leq \int_\Omega |\nabla u \cdot \nabla v| \, dx \\ &\leq \int_\Omega \|\nabla u\| \|\nabla v\| \, dx \quad (\textit{by Chauchy Schwarz inequality for } \mathbb{R}^2) \\ &\leq \|\nabla u\|_{L_2(\Omega)} \|\nabla v\|_{L_2(\Omega)} \quad (\textit{by Chauchy Schwarz inequality for } L_2(\Omega)) \end{aligned}$$

Since,

$$\begin{aligned}\|\nabla u\|^2_{L_2(\Omega)} &= \int_\Omega \nabla u \cdot \nabla v \\ &\leq \int_\Omega (\nabla u \cdot \nabla u + u^2) \\ &= \|u\|^2_V\end{aligned}$$

and similarily for v, it follows that

$$\langle Lu, v\rangle \leq \|u\|_V \|v\|_V \qquad for\ all\ u, v \in V$$

implies,

$$\sup_{v\neq 0} \frac{\langle Lu, v\rangle}{\|v\|_V} \leq \|u\|_V.$$

Thus, $\|Lu\| \leq \|u\|_V$. Hence L is bounded. Therefore L is continuous.
Now let us show strong monotonocity of L. From Poincaré's inequality we get

$$\langle Lu, u\rangle \geq c^2 \|u\|^2_{H^1(\Omega)} \quad for\ \ all\ u \in H^1_0(\Omega)$$

Therefore, L is strongly monotone. Thus from Example (**??**), for any $f \in H^{-1}(\Omega)$, there is a unique $u \in H^1_0(\Omega)$ such that

$$\int_\Omega \nabla u \cdot \nabla v\, dx = \langle f, v\rangle \quad \forall v \in V.$$

i.e., the boundary value problem (2.6) has a unique weak solution. □

2.4 The Lax-Milgram Lemma

The Lax-Milgram Lemma is employed frequently in the study of linear elliptic boundary value problems. For a real Banach space V, Let us first explore the relation between a linear operator A:$V \to V'$ and a bilinear form a:$V \times V \to R$ related by

$$\langle Au, v\rangle = a(u, v) \quad \forall u, v \in V \tag{2.7}$$

The bilinear form a(.,.) is continuous if and only if there exists $M > 0$ such that

$$\mid a(u, v) \mid \leq M \parallel u \parallel \parallel v \parallel \quad \forall u, v \in V \tag{2.8}$$

Theorem 2.4.1. *There exists a one-to-one correspondence between linear continuous operators $A : V \to V'$ and continuous bilinear forms $a : V \times V \to R$, given by the formula(2.5)*

Theorem 2.4.2. *Assume K is a nonempty, closed,convex subset of the Hilbert space V, a(.,.) : $V \times V \to R$ bilinear, symmetric, bounded and V -elliptic, $L \in V'$. Let*

$$E(v) = \frac{1}{2}a(u,v) - L(v), v \in V \tag{2.9}$$

Then there exists a unique u ∈K such that

$$E(u) = \inf_{v \in K} E(v) \tag{2.10}$$

which is also the unique solution of the variational inequality

$$a(u, v-u) \geq L(v-u), \quad \forall u, v \in K, \tag{2.11}$$

Or

$$a(u,v) = L(v), \quad \forall u, v \in K$$

in the special case where K is a subspace

Theorem 2.4.3. *(The Lax-Milgram Lemma) Assume V is a Hilbert space, a(.,.) is a bounded, Velliptic bilinear form on V and $L \in V'$. Then there is a unique solution of the problem*

$$a(u,v) = L(v) \quad \forall v, u \in V \tag{2.12}$$

Before proving ,consider the case V= R and a simple linear equation with the corresponding weak formulation
$x \in R$,ax=L
$x \in R$,axy=Ly $\quad \forall y \in R$
To ensure existence of solutions we need:
$o < a < \infty$.i.e, the bilinear form a(x,y)= axy must be continuous and R-elliptic,
$\mid L \mid < \infty$ i.e, the linear functional l(y) = ly must be bounded

Proof. (Existence)
Let $A : V \to V'$ be the linear operator associated with the bilinear form $a(\cdot, \cdot)$; see $\langle Av, v \rangle$=a(u,v). Then A is bounded and strongly positive: $\forall v \in V$,

$$\|Av\| \leq M\|v\|,$$

$$\langle Av, v \rangle \geq \alpha \|v\|^2.$$

Denote $\mathcal{J} : V' \to V$ the isometric dual mapping from the Riesz representation theorem. Then

$$a(u,v) = \langle Au, v \rangle = (\mathcal{J}Au, v) \quad \forall u, v \in V,$$

and

$$\|\mathcal{J}Au\| = \|Au\| \quad \forall u \in V.$$

And, Let $L = \mathcal{J}A : V \to V$. We recall that $\mathcal{R}(L) = V$ if and only if $\mathcal{R}(L)$ is closed and $\mathcal{R}(L)^{\perp} = \{0\}$.
To show $\mathcal{R}(L)$ is closed, we let $\{u_n\} \subset \mathcal{R}(L)$ be a sequence converging to u. Then $u_n = \mathcal{J}Aw_n$ for some $w_n \in V$. We have

$$\|u_n - u_m\| = \|\mathcal{J}A(w_n - w_m)\| = \|A(w_n - w_m)\| \geq \alpha \|w_n - w_m\|.$$

Hence $\{w_n\}$ is a Cauchy sequence and so has a limit $w \in V$. Then

$$\|u_n - \mathcal{J}Aw\| = \|\mathcal{J}A(w_n - w)\| = \|A(w_n - w)\| \leq M\|w_n - w\| \to 0.$$

Hence, $u = \mathcal{J}Aw \in \mathcal{R}(L)$ and $\mathcal{R}(L)$ is closed.
Now suppose $u \in \mathcal{R}(L)^{\perp}$. Then for any $v \in V$,

$$0 = (\mathcal{J}Av, u) = a(v, u).$$

Taking $v = u$ above, we have $a(u, u) = 0$. By the V-ellipticity of $a(\cdot, \cdot)$, we conclude $u = 0$. So we can invoke Theorem (2.3.1) to obtain the existence of a solution.
(Uniqueness)

Suppose u_1, u_2 are two solutions of , i.e;

$$a(u_1, v) = \ell(v),$$

$$a(u_2, v) = \ell(v) \quad \forall v \in V.$$

Subtraction and linearity give

$$a(u_2, v) - a(u_1, v) = a(u_2 - u_1, v) = 0 \quad \forall v \in V.$$

In particular, choose $v = u_2 - u_1$, then we see

$$a(u_2 - u_1, u_2 - u_1) = 0.$$

By the V-ellipticity of $a(\cdot, \cdot)$,

$$\begin{aligned} & 0 \geq \alpha\|u_2 - u_1\|_V^2, \\ \Rightarrow \quad & 0 \geq \|u_2 - u_1\|_V, \\ \Rightarrow \quad & u_2 = u_1 \end{aligned}$$

□

2.5 Weak formulations of linear elliptic boundary value problems

In this section, we formulate and analyze weak formulations of some linear elliptic boundary value problems. To present the ideas clearly, we will frequently use boundary value problems associated with the Poisson equation

$$-\Delta u = f$$

and the Helmholtz equation

$$-\Delta u + u = f$$

as examples.

2.5.1 Problems with homogeneous Dirichlet boundary conditions

So far, we have studied the model elliptic boundary value problem corresponding to the Poisson equation with the homogeneous Dirichlet boundary condition

$$-\Delta u = f \quad in\ \Omega, \tag{2.13}$$

$$u = 0 \quad in\ \Gamma, \tag{2.14}$$

where $f \in L_2(\Omega)$. The weak formulation of the problem is

$$u \in V, \qquad a(u,v) = \ell(v) \qquad \forall v \in V. \tag{2.15}$$

Here

$$\begin{aligned} V &= H_0^1(\Omega), \\ a(u,v) &= \int_\Omega \nabla u \cdot \nabla v \, dx \quad for\ u, v \in V \\ \ell(v) &= \int_\Omega f\, v \, dx \quad for\ v \in V \end{aligned}$$

The problem (2.15) has a unique solution $u \in V$ by the Lax-Milgram Lemma. Dirichlet boundary conditions are also called essential boundary conditions since they are explicitly required by the weak formulation.

2.5.2 Problems with non-homogeneous Dirichlet boundary conditions

Suppose that instead of (2.14) the boundary condition is

$$u = g \quad on\ \Gamma. \tag{2.16}$$

To derive a weak formulation, we first assume the boundary value problem (2.13)-(2.16) has a classical solution $u \in C^2(\Omega) \cap C(\overline{\Omega})$. Multiplying the equation (2.13) by a test function v with certain smoothness which validates the following calculations, and integrating over Ω, we have

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

Integrate by parts,

$$-\int_\Gamma \frac{\partial u}{\partial \nu} v \, ds + \int_\Omega \nabla u \cdot \nabla v \, dx = \int_\Omega f\, v \, dx$$

We now assume $v = 0$ on Γ so that the boundary integral term vanishes; the boundary integral term would otherwise be difficult to deal with under the expected regularity condition $u \in H^1(\Omega)$ on the weak solution. Thus we arrive at the relation

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

if v is smooth and $v = 0$ on Γ. For each term in the above relation to make sense, we assume $f \in L_2(\Omega)$, and let $u \in H^1(\Omega)$ and $v \in H_0^1(\Omega)$. Recall that the solution u should satisfy the boundary condition $u = g$ on Γ. We observe that it is necessary to assume $g \in H^{1/2}(\Omega)$. Finally, we obtain the weak formulation for the boundary value problem (2.13)-(2.16):

$$u \in H^1(\Omega),\ u = g \quad on\ \Gamma, \quad \int_\Omega \nabla u \cdot \nabla v \, dx = \int_\Omega f\, v \, dx \quad \forall v \in H_0^1(\Omega). \tag{2.17}$$

For the weak formulation (2.17), though, we cannot apply Lax-Milgram Lemma directly since the trial function u and the test function v do not lie in the same space. There is a standard way to get rid of this problem. Since $g \in H^{1/2}(\Omega)$ and $\gamma(H^1(\Omega)) = H^{1/2}(\Omega)$ (recall that γ is the trace operator), we have the existence of a function $G \in H^1(\Omega)$ such that $\gamma G = g$. We remark that finding the function G in practice may be nontrivial. Thus, setting

$$u = w + G,$$

the problem may be transformed into one of seeking w such that

$$w \in H_0^1(\Omega), \quad \int_\Omega \nabla w \cdot \nabla v \, dx = \int_\Omega (f\, v - \nabla G \cdot \nabla v) \, dx \quad \forall v \in H_0^1(\Omega) \tag{2.18}$$

The classical form of the boundary value problem for w is

$$\begin{aligned} -\Delta w =& f + \Delta G \quad in\ \Omega, \\ w =& 0 \quad on\ \Gamma. \end{aligned}$$

Applying the Lax-Milgram Lemma, we have a unique solution $w \in H_0^1(\Omega)$ of the problem (2.18). Then we set $u = w + G$ to get a solution u of the problem (2.17). Notice that the choice of the function G is not unique, so the uniqueness of the solution u of the problem (2.17) does not follow from the above argument. Nevertheless we can show the uniqueness of u by a standard approach. Assume both u_1 and u_2 are solution of the problem (2.17). Then the difference $u_1 - u_2$ satisfies

$$u_1 - u_2 \in H_0^1(\Omega), \quad \int_\Omega \nabla(u_1 - u_2) \cdot \nabla v \, dx = 0 \quad \forall v \in H_0^1(\Omega)$$

Taking $v = u_1 - u_2$, we obtain

$$\int_\Omega |\nabla(u_1 - u_2)|^2 \, dx = 0$$

Thus, $\nabla(u_1 - u_2) = 0$ a.e. in Ω, and hence $u_1 - u_2 = c$ a.e. in Ω. Using the boundary condition $u_1 - u_2 = 0$ a.e. Γ, we see that $u_1 = u_2$ a.e. in Ω.

2.5.3 Problems with Neumann boundary conditions

Consider next the Neumann problem of determining u which satisfies

$$\begin{cases} -\Delta u + u = f \quad in\ \Omega, \\ \partial u / \partial v = g \quad on\ \Gamma. \end{cases} \tag{2.19}$$

Here f and g are given functions in Ω and on Γ, respectively, and $\partial/\partial\nu$ denotes the normal derivative on Γ. Again we first derive a weak formulation.
Assume $u \in C^2(\Omega) \cap C^1(\overline{\Omega})$ is a classical solution of the problem (2.19). Multiplying (2.19)1 by an arbitrary test function v with certain smoothness for the following calculations to make sense, integrating over Ω and performing an integration by parts, we obtain

$$\int_\Omega (\nabla u \cdot \nabla v + u\,v)\,dx = \int_\Omega f\,v\,dx + \int_\Gamma \frac{\partial u}{\partial \nu} v\,ds$$

Then, substitution of the Neumann boundary condition (2.19)2 in the boundary term leads to the relation

$$\int_\Omega (\nabla u \cdot \nabla v + u\,v)\,dx = \int_\Omega f\,v\,dx + \int_\Gamma g\,v\,ds$$

Assume $f \in L_2(\Omega)$, $g \in L_2(\Gamma)$. For each term in the above relation to make sense, it is natural to choose the space $H^1(\Omega)$ for both the trial function u and the test function v. Thus, the weak formulation of the boundary value problem (2.19) is

$$u \in H^1(\Omega) \quad \int_\Omega (\nabla u \cdot \nabla v + u\,v)\,dx = \int_\Omega fv\,dx + \int_\Gamma g\,v\,ds \quad \forall v \in H^1(\Omega) \tag{2.20}$$

This problem has the form (2.15), where $V = H^1(\Omega)$, $a(\cdot,\cdot)$ and $\ell(\cdot)$ are defined by

$$a(u,v) = \int_\Omega (\nabla u \cdot \nabla v + u\,v)\,dx,$$
$$\ell(v) = \int_\Omega f\,v\,dx + \int_\Gamma g\,v\,ds,$$

respectively. Applying the Lax-Milgram Lemma, we can show that the weak formulation (2.20) has a unique solution $u \in H^1(\Omega)$.
It is more delicate to study the pure Neumann problem for the Poisson equation

$$\begin{cases} -\Delta u = f & in\ \Omega, \\ \partial u/\partial v = g & on\ \Gamma. \end{cases} \tag{2.21}$$

where $f \in L_2(\Omega)$ and $g \in L_2(\Gamma)$ are given. In general, the problem (2.21) does not have a solution, and when the problem has a solution u, any function of the form $u + c, c \in \mathbb{R}$, is a solution. This suggests that in formulating the weak version of this problem we should restrict ourselves to the subspace

$$V = \{v \in H^1(\Omega) \mid \int_\Omega v\,dx = 0\}$$

Formally, the corresponding weak formulation is

$$u \in H^1(\Omega) \quad \int_\Omega \nabla u \cdot \nabla v\,dx = \int_\Omega f\,v\,dx + \int_\Gamma g\,v\,ds \quad \forall v \in H^1(\Omega) \tag{2.22}$$

An application of Equivalent Norm Theorem shows that over the space V, $|\cdot|_1$ is a norm equivalent to the norm $\|\cdot\|_1$. The bilinear form $a(u,v) = \int_\Omega \nabla u \cdot \nabla v\,dx$ is both continuous and V-elliptic. So there is a unique solution to the problem

$$u \in V \quad \int_\Omega \nabla u \cdot \nabla v\,dx = \int_\Omega f\,v\,dx + \int_\Gamma g\,v\,ds \quad \forall v \in V.$$

2.5.4 Problems with non-homogeneous Dirichlet and Neumann boundary conditions

$$\begin{cases} -\Delta u(x) = f(x) \quad in\ \Omega \ \ , -\Delta u(x) - f(x) = 0 \\ \\ u = g \quad on\ \Gamma_D \\ \frac{\partial u}{\partial n} = h \quad on\ \Gamma_N \end{cases} \tag{2.23}$$

convert to homogeneous problem
u=ω+û
where
$\omega \equiv$ known function w=g on Γ_D,
û$\equiv$ new function that we look for

$$V = \{v \in H^1(\Omega) \ \ and \ \ v = 0 \ \ on \ \ \Gamma_D\}$$

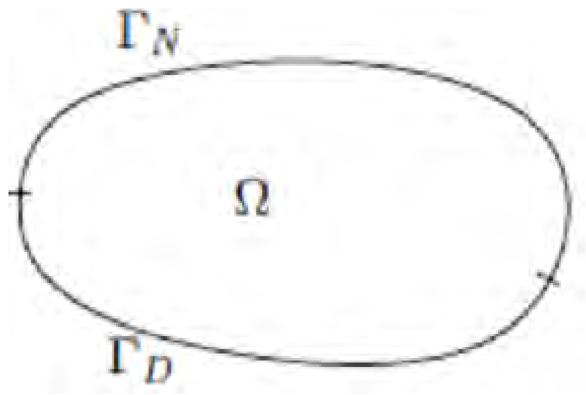
- Multiply by test function v and integrate.

$$\int -\Delta u(x)v(x)d\Omega - \int f(x)v(x)d\Omega = 0 \tag{2.24}$$

- Reduce order for (Lu,v)

$$\int -\Delta u(x)v(x)d\Omega = \int_\Omega \nabla u(x).\nabla v(x)d\Omega - \int_{\partial\Omega} \frac{\partial u}{\partial n} v(x)d\Gamma \tag{2.25}$$

- Apply boundary conditions



$$\int_{\partial\Omega} \frac{\partial u}{\partial n} v(x)d\Gamma = \int_{\Gamma_N} \frac{\partial u}{\partial n} v(x)d\Gamma + \int_{\Gamma_D} \frac{\partial u}{\partial n} v(x)d\Gamma = \int_{\Gamma_N} hv(x)d\Gamma \tag{2.26}$$

Therefore:

$$\int_\Omega \nabla u(x).\nabla v(x)d\Omega = \int f(x)v(x)d\Omega + \int_{\Gamma_N} hv(x)d\Gamma \tag{2.27}$$

But
u=ω+û

$$\int_\Omega \nabla(\omega(x) + \hat{u}(x)).\nabla v(x)d\Omega = \int f(x)v(x)d\Omega + \int_{\Gamma_N} hv(x)d\Gamma \tag{2.28}$$

$$\int_{\Omega} \nabla\omega(x).\nabla v(x)d\Omega + \nabla\hat{u}(x).\nabla v(x)d\Omega = \int f(x)v(x)d\Omega + \int_{\Gamma_N} hv(x)d\Gamma \tag{2.29}$$

$$\int_{\Omega} \nabla\hat{u}(x).\nabla v(x)d\Omega = \int f(x)v(x)d\Omega + \int_{\Gamma_N} hv(x)d\Gamma - \int_{\Omega} \nabla\omega(x).\nabla v(x)d \tag{2.30}$$

$$\int_{\Gamma_N} hv(x)d\Gamma \quad from \quad natural(NeumannBC)$$

$$\int_{\Omega} \nabla\omega(x).\nabla v(x)d\Gamma \quad from \quad essential(DirichletBC)$$

Under suitable assumptions, $f \in L^2(\Omega)$and $g \in L^2(_N)$, we can again apply the Lax-Milgram Lemma to conclude that the weak problem has a unique solution.Dirichlet and Neumann boundary conditions are called Rabin or Mixed boundary conditions.

Chapter 3

The Galerkin method and its variants

3.1 Introduction

In this chapter, we discuss some numerical methods for solving boundary value problems. These are the Galerkin method and its variants: the Petrov-Galerkin method and the generalized Galerkin method.In Section (3.5),we rephrase the conjugate gradient method ,discussed in for solving variational equations. The Galerkin method provides a general framework for approximation of operator equation which includes the finite element method as spacial case. partial differential equation of applied mathematics ,heat equation, laplace equation, wave equation, Poisson equation it also includes numerical method that are employed to solve the above mentioned partial differential equations. The most important discretization schemes are introduced e.g finite difference method and finite element method. New discussion the Galerkin method for a linear operator equation in a form directly applicable to the study of the finite element method the Galerkin method provides a natural framework for finding finite -dimensional approximation of weak solutions of elliptic boundary value problems.

3.2 The Galerkin method

The basic problem of interest is following: given a function $f \in v$ (the dual space v) to find $u \in V$ for which.

$$a(u, v) = \ell(v) \quad \forall v \in V \tag{3.1}$$

The Lax- milgram Lemma. the existence of a unique solution u to the problem.
In case of partial differential equation or integral equation, the space V is infinite dimensional, hence it is impossible to find an exact solution to problem a(u,v) = $\ell(v)$.
In order to construct an approximate solution, it is natural to consider a finite dimensional approximation to a(u,v)=$\ell(v)$ for instance, we consider the finite dimensional subspace $V_N \subset V$, an N-dimensional subspace of the space V, and the problem a(u,v)=$\ell(v)$ ontoV_N that is

$$a(u_N, v) = \ell(v) \quad \forall u_N \in V_N, v \in V_N \tag{3.2}$$

The resulting sequence of approximate solution converges in some sense to the solution of the original problem. Since a(.,.) is bounded and V-elliptic as before. Lax- Milgram grants the existence of a unique solution $u_N \in V_N$ for $L \in V'$
we rewrite problem

$$a(u_N, v) = \ell(v) \quad \forall u_N \in V_N, v \in V_N$$

Discrete (approximated) problem

If the problem is well-posed one can try to find an approximated solution u_N by solving the so-called discrete problem which is an approximation of the corresponding variational problem.
Find $u_N \in U_N$ so that

$$a_N(u_N, v_N) = \ell_N(v_N) \quad \forall v_N \in V_N \tag{3.3}$$

where :
a is an approximation of bilinear form A
u_N is approximation solution
v is discrete test function form the discrete test space
ℓ is an approximation of Linear form L
V_N is finite dimensional space of functions, called approximation space
As a Linear system supposing ${\Phi_i}_{i=1}^N$ be a basis of the finite dimensional subspace V_N, we write

$$u_N = \sum_{j=1}^{n} \xi_j \Phi_j$$

where$\Phi_j(x)$ is trial function
take $v \in V_N$ to be the basis functions Φ_i. Accordingly, we get the Linear system.

$$A\xi = b \tag{3.4}$$

where:
$\xi \in R^N$, ξ =$[\xi_1, \xi_2, ..., \xi_n]^T$is the unknown vector.
$(A)_{i,j}$ = $[a(\Phi_i, \Phi_j)]$=$\int \nabla\Phi_i.\nabla\Phi_j \quad \in R^{N}*^{N}$ is the stiffness matrix or Galerkin matrix .
b = $(L(\Phi_i))$=$\int \Phi_j f$ =$\langle\Phi_i, f\rangle \in R^N$ is the Load vector or Right hand side vector.
The approximate solution u_N differs from the exact solution u, to increase the accuracy of the solution, we look for solution i Larger finite dimensional subspace. Accordingly, corresponding to a sequence of subspace $V_1 \subset V_2 \subset V_3 ... \subset V$,$V_j < \infty$, j=1,2,3... we compute the approximate solution $u_i \in V_i$, i= 1,2,3,... this solution procedure generates the so-called the Galerkin method. Now consider the special case when a(u,v) is symmetric,

$$a(u, v) = a(v, u) \quad \forall u \in V$$

problem a(u,v) $= \ell(v)$ is equivalent to the minimization programme

$$E(u) = \inf_{v \in V} E(v), \quad u \in V \tag{3.5}$$

where the energy functional

$$E(v) = \frac{1}{2} a(v, v) - f(v) \tag{3.6}$$

Now with a finite dimensional subspace $V_N \subset V$ chosen, it is equally natural to develop a numerical method by minimizing the energy functional over the finite dimensional subspace V_N of V,

$$E(u_N) = \inf_{v \in V_N} E(v) \quad u_N \in V_N \tag{3.7}$$

it is easy to verify that the two approximate problems $a(u_N, v) = \ell(v)$ and $E(u_N) = \inf_{v \in V_N} E(v)$ are equivalent.
The method based on minimizing the energy functional over finite dimensional subspaces is called the Ritz method. We see that the Galerkin method is more general than the Ritz method, while when both methods are applicable, they are equivalent. Usually, the Galerkin method is also called the Ritz-Galerkin method. Or for many problems, the Ritz and Galerkin methods are theoretically equivalent.
• The Ritz method is based on the minimization form, and optimization techniques can be used to solve the problem.
• The Galerkin method usually has weaker requirements than the Ritz method. Not every problem has a minimization form, whereas almost all problems have some kind of weak form. How to choose suitable weak form and the convergence of different methods are all important issues for finite element methods.

The Ritz method

Although not every problem has a minimization form, the Ritz method was one of the earliest and has proven to be one of the most successful. For the model problem (3.8), the minimization form is

$$\min_{v(x) \in H_0^1(0,1)} F(v) : F(v) = \frac{1}{2} \int_0^1 (v')^2 dx - \int_0^1 f v dx \tag{3.8}$$

As before, we look for an approximate solution of the form $u_N(x) = \sum_{j=1}^{N} \alpha_j \Phi_j$
Substituting this into the functional form gives

$$F(u_N) = \frac{1}{2} \int_0^1 (\sum_{j=1}^{N} \alpha_j \Phi_j')^2 dx - \int_0^1 f \sum_{j=1}^{N} \alpha_j \Phi_j dx \tag{3.9}$$

which is a multivariate function of $\alpha_1, \alpha_1, , \alpha_{n-1}$ and can be written as
The necessary condition for a global minimum (also a local minimum) is

$$\frac{\partial F}{\partial \alpha_1} = 0, \frac{\partial F}{\partial \alpha_2} = 0, .., \frac{\partial F}{\partial \alpha_i} = 0, .., \frac{\partial F}{\partial \alpha_{n-1}} = 0$$

Thus taking the partial derivatives with respect to $_j$ we have

$$\frac{\partial F}{\partial \alpha_1} = \int_0^1 (\sum_{j=1}^N \alpha_j \Phi'_j)\Phi'_1 dx - \int_0^1 f\Phi_1 dx \tag{3.10}$$

$$\vdots \qquad \vdots \qquad \vdots$$

$$\frac{\partial F}{\partial \alpha_i} = \int_0^1 (\sum_{j=1}^N \alpha_j \Phi'_j)\Phi'_i dx - \int_0^1 f\Phi_i dx, i = 1, 2, ..n-1 \tag{3.11}$$

and on exchanging the order of integration and summation

$$\int_0^1 (\sum_{j=1}^N \Phi_j \Phi'_j)\Phi'_i dx - \int_0^1 f\Phi_i dx, i = 1, 2, ..n-1 \tag{3.12}$$

This is the same system of equations that follow from the Galerkin method with the weak form, i.e.,

$$\int_0^1 u'v'dx = \int_0^1 fvdx \tag{3.13}$$

immediately gives

$$\int_0^1 (\sum_{j=1}^N \alpha_j \Phi'_j)\Phi'_i dx = \int_0^1 f\Phi_i dx, i = 1, 2, ..n-1 \tag{3.14}$$

Different formulations for the 1D model

Let us consider the model.

$$-u''(x) = f(x), u(0) = 0, u(1) = 0, \tag{3.15}$$

There are at least three different formulations to consider for this problem:
• the (D)-form, the original differential equation
• the (V)-form, the variational form or weak form

$$\int_0^1 u'v'dx = \int_0^1 fvdx \tag{3.16}$$

for any test function $v \in H_0^1(0,1)$, the Sobolev space for functions in integral forms like the C^1 space for functions, and as indicated above, the corresponding FE method is often called the Galerkin method and
• the (M)-form, the minimization form

$$\min_{v(x) \in H_0^1(0,1)} \{\int_0^1 (\frac{1}{2}(v')^2 - fv)dx\} \tag{3.17}$$

Under certain assumptions these three different formulations are equivalent.

Theorem 3.2.1. *Representation theorem of Riesz*
Let $f \in V'$ be a continuous and linear functional, then there is a uniquely determined $u \in V$ with . a(u,v)=f(v)
In addition u is the unique solution of the variational problem.
F(v)=$\frac{1}{2}a(v,v) - f(v)$

Proof. First the existence of a solution u of the variational problem will be proved. Since f is continuous, it holds

$$| f(v) | \leq c \parallel v \parallel_V \quad \forall v \in V$$
$$F(v) \geq \frac{1}{2} \parallel v \parallel \parallel_V \geq -\frac{1}{2}c^2,$$

where in the last estimate the necessary criterion for a local minimum of the expression of the First estimate is used. Hence,the function F(.) is bounded from below and

$$d = \inf_{v \in V} F(v)$$

exists
Let $\{v_k\}_{k \in N}$ be a sequence with $F(v_k) \to d \quad$ for $k \to \infty$

$$\parallel v_k - v_L \parallel_V^2 + \parallel v_k + v_L \parallel_V^2 = 2 \parallel v_k \parallel_V^2 + 2 \parallel v_L \parallel_V^2$$

using the Linearity of f(.) and $d \leq F(v)$ For all $v \in V$ one obtains.

$$\parallel v_k - v_L \parallel_V^2 = 2 \parallel v_k \parallel_V^2 + 2 \parallel v_L \parallel_V^2 - 4 \parallel \frac{v_k + v_L}{2} \parallel^2 - 4f(v_k) - 4f(v_L) + 8f(\frac{v_k + v_L}{2})$$
$$= 4F(v_k + 4F(v_L) - 8F(\frac{v_k + v_L}{2})$$
$$\leq 4F(v_k) + 4F(v_L) - 8d \to 0$$

For k,L $\to \infty$, Hence $\{v_k\}_{k \in N}$ is a cauchy sequence.
because V is complete space, there exists a limit u of this sequence with $u \in V$. F(.) is continuous, it is f(u) =d and u is a solution of the variational problem. In the next step, it will be shown that each solution of the variational problem F(v)= $\frac{1}{2}$ a(v,v) - f(v) is also a solution of a(u,v) = f(v) it is

$$\Phi(\xi) F(u + \xi v) = \frac{1}{2} a(u + \xi v, u + \xi v) - f(u + \xi v)$$
$$= \frac{1}{2} a(u, u) + \xi a(u, v) + \frac{\xi^2}{2} a(v, v) - f(u) - \xi f(v)$$

if u is a minimum of the variational problem . then the function $\Phi(\xi)$ has a local minimum at ξ= 0 then necessary condition for a local minimum leads to
$\Phi(0)$ = a(u,v)- f(v) for all $v \in$ V
Finally, the uniqueness of the solution of the equation a(u,v)=f(v) if the solution of this equation is unique, Then the existence of two solution of the variational problem

$$F(v) = \frac{1}{2} a(v, v) - f(v)$$

Let u_1, u_1 be two solution of this equation a(u,v) = f(v).
$a(u_1 - u_1, v) = 0$ for all $v \in V$
this equation hold ,
In particular $u_1 - u_1$=v, Hence $\| u_1 - u_1 \|_V = 0$ such that $u_1 = u_1$

□

Proposition 3.2.1. *Cea's inequality, error estimate Let b:V $\times$ V $\to$ R be a bounded and coercive bilinear form on the Hibert space V and Let $f \in V'$ be a bounded linear functional. Let u be the solution of a(u,v)=f(v) and u_N be the solution of $a(u_N, v)$= f(v), then the following error estimate holds* $\| u - u_N \|_V \leq = \frac{M}{m} \inf_{v \in V_N} \| u - v \|_V$

Proof. where the constant $M > 0$, $m > 0$ can be chosen to be M=1 and m= 1 or $\frac{M}{m}$=c subtracting

$$a(u_N, v) = f(v) \quad from \quad a(u, v) = f(v)$$

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

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

$$a(u - u_N, v) = 0 \quad v \in V_N$$

$$a(v, v) = 0 \quad \Leftrightarrow v = 0$$

And we can defined energy norm
$\| v \|_V = \sqrt{a(v, v)}$

$$\| u - u_N \|_V^2 = a(u - u_N, u - u_N) \quad (using the Linearity)$$

$$\| u - u_N \|_V^2 = a(u - u_N, u - v) + a(u - u_N, v - u_N)$$

$$\| u - u_N \|_V^2 = a(u - u_N, u - v)$$

Now use inequality

$$\| u - u_N \|_V^2 \leq \| u - u_N \|_V \| u - v \|_V \quad for \quad any \quad a \quad bitrary \quad v \in V_N$$

Now there are two cases,
• case
u = u_N , the error =0
• case
the error $\neq 0$ we defined by

$$\| u - u_N \|_V^2 \leq \| u - u_N \|_V \| u - v \|_V \tag{3.18}$$

$$\| u - u_N \|_V \leq \| u - v \|_V \tag{3.19}$$

$$\| u - u_N \|_V \leq \inf_{v \in V_N} \| u - v \|_V \tag{3.20}$$

□

3.3 The Galerkin Method for the one dimensional examples

Example 1: Consider the boundary value problem

$$\begin{cases} -u'' = f, & in \quad (0,1) \\ u = (0) = u(1) = 0, \end{cases} \tag{3.21}$$

$$(-u'')vdx = fvdx$$

$$\int_0^1 (-u'')vdx = \int_0^1 fvdx$$

Left hand side and integration by parts [$\int udv$=uv- $\int vdu$]

$$\int_0^1 u''vdx = \int_0^1 u'v'dx \tag{3.22}$$

is the weak formulation
where V $=H_0^1$
Applying the Lax- Milgram Lemma the weak problem has a unique solution. To develop a Galerkin method we need to choose a finite-dimensional subspace of basis function That satisfies the boundary condition at x=0, x=1 . Consider one whose basis vector satisfies the boundary condition of the boundary value problem.
$V_N = span\{x^i(1-x)|i = 1, 2, 3....N\}$
approximating

$$u_N = \sum_{j=1}^{n} \xi_j x^i(1-x) \tag{3.23}$$

where
the coefficients ξ_i are determined by Galerkin equation or Linear system
take v to be the basis functions $x^i(1-x)$ we obtain the following Linear system of equation, is

$$A\xi = b \tag{3.24}$$

N	Condition number(A)
3	891.6637
4	2.4233e+ 04
5	6.5617e+ 05
6	1.7919e+ 07
7	4.9532e+ 08
8	1.3867e+ 10
9	3.9288e+ 11
10	1.1282e+ 13

where:
ξ=$(\xi_1, \xi_2, \xi_3, ..., \xi_N)^\top$ is vector of unknowns
A=is stiffness(coefficient) matrix has(i, j)
b=is vector whose $i-th$ component is

$$b = \int_0^1 fvdx = \int_0^1 f(x)x^i(1-x)dx \tag{3.25}$$

$$\int_0^1 u''vdx = \int_0^1 u'v'dx \tag{3.26}$$

$$\int_0^1 [x^i(1-x)]'[x^j(1-x)]'dx = \frac{(i+1)(j+1)}{i+j+1} + \frac{(i+2)(j+2)}{i+j+3} - \frac{(i+1)(j+2)+(i+2)(j+1)}{i+j+2} \tag{3.27}$$

The coefficient matrix is fairly ill-conditioned, and as a matter of fact, using the cond(X, p) command of Matlab, it is fairly easy to calculate the condition numbers of the matrices A, for instance in 2- norm,

Remarks:

• The weak formulation is also called variational formulation.

• As usual in mathematics, 'weak' means that something holds for all appropriately chosen test functions.

• The above examples clearly illustrate that it is very important to choose appropriate basis function or (polynomial basis function) to the
Finite dimensional subspaces or trigonometric polynomial basis function.

• The Galerkin method is not just a numerical scheme for approximating solution to a differential or integral equations. By possing to the limit, we can even prove some existence results.

Example 1 the Galerkin method for the 1D two-point BVP

$$-u''(x) = f(x), 0 < x < 1, u(0) = 0, u(1) = 0, \tag{3.28}$$

• Construct a variational or weak formulation, by multiplying both sides of the differential equation by a test function v(x) satisfying the boundary conditions (BC)
v(0)= 0,v(1) = 0 to get

$$u''v = fv$$

and then integrating from 0 to 1 (using integration by parts): thus

$$\int_0^1 (-u''v)dx = -u'v \mid_0^1 + \int_0^1 u'v'dx \tag{3.29}$$

$$= \int_0^1 u'v'dx \tag{3.30}$$

$$\int_0^1 u'v'dx = \int_0^1 fvdx \tag{3.31}$$

The weak form

• Generate mesh, e.g., a uniform Cartesian mesh x_i= i h, i = 0, 1, ... , n, where h = 1/n, defining the intervals $[x_{i-1}, x_i]$, i = 1, 2, ... , n.

• Construct a set of basis functions based on the mesh, such as the piecewise linear functions (i = 1, 2, ... , n - 1)

$$\Phi_i(x) = \begin{cases} \frac{x-x_i-1}{h} & if x_{i-1} \leq x \leq x_i, \\ \frac{x_{i+1}-x}{h} & if x_i \leq x \leq x_{i+1}, \\ 0 \quad otherwise \end{cases}$$

Or

using the Galerkin method and using the piecewise linear functions (figure 3.1) as basis:$\Phi_j(x) =$

$$\begin{cases} \frac{1}{\triangle}(x - x_{j-1}) & for x_{j-1} \leq x \leq x_j \\ \frac{1}{\triangle}(x_{j+1} - x) & for x_j \leq x \leq x_{j+1} \\ o \quad otherwise \end{cases}$$

$$with \triangle = x_{j+1} - x_j, \forall j, x_0 < x_1 < x_2 < .... < x_N < x_{N+1} = 1.$$

often called the hat functions.

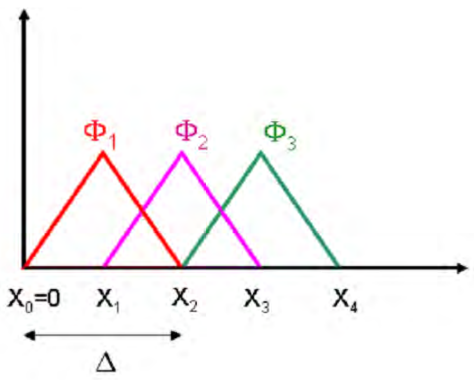


Figure 3.1: Linear basis functions

• Represent the approximate (Galerkin) solution by the linear combination of basis functions

$$u_N(x) = \sum_{j=1}^{N} c_j \Phi_j(x) \tag{3.32}$$

where the coefficients c_j are the unknowns to be determined.On assuming the hat basis functions,obviously $u_h(x)$ is also a piecewise linear function, although this is not usually the case for the true solution u(x). Other basis functions are considered later. We then derive a linear system of equations for the coefficients by substituting the approximate solution $u_N(x)$ for the exact solution u(x) in the weak form $\int_0^1 u'v'dx$=$\int_0^1 fvdx$

$$\int_0^1 u'_N v' dx = \int_0^1 fvdx \tag{3.33}$$

$$\int_0^1 u'_N(x)v'dx = \sum_{j=1}^{n-1} c_j \int_0^1 \Phi_j(x)v'dx = \int_0^1 fvdx \tag{3.34}$$

$$\sum_{j=1}^{n-1} c_j \int_0^1 \Phi_i'(x)\Phi_j'(x)dx = \int_0^1 f\Phi_j'(x)dx \tag{3.35}$$

Now the test function v(x) is chosen to be Φ_1,Φ_2, Φ_3... , Φ_{n-1} successively,to get the system of linear equations get the system of linear equations (noting that further errors are introduced):

$$\begin{gathered}
(\int_0^1 \Phi_1'\Phi_1' dx)c_1 + ... + (\int_0^1 \Phi_1'\Phi_{n-1}' dx)c_{n-1} = (\int_0^1 f\Phi_1' dx) \\
(\int_0^1 \Phi_2'\Phi_1' dx)c_1 + ... + (\int_0^1 \Phi_2'\Phi_{n-1}' dx)c_{n-1} = (\int_0^1 f\Phi_2' dx) \\
.... \qquad .... \qquad .... \qquad .... \\
(\int_0^1 \Phi_{n-1}'\Phi_1' dx)c_1 + ... + (\int_0^1 \Phi_{n-1}'\Phi_{n-1}' dx)c_{n-1} = (\int_0^1 f\Phi_{n-1}' dx)
\end{gathered} \tag{3.36}$$

or in matrix-vector form:

$$\begin{bmatrix}
a(\Phi_1,\Phi_1) & a(\Phi_1,\Phi_2) & ..... & a(\Phi_1,\Phi_{n-1}) \\
a(\Phi_2,\Phi_1) & a(\Phi_2,\Phi_2) & ..... & a(\Phi_2,\Phi_{n-1}) \\
\vdots & \vdots & ...... & \vdots \\
a(\Phi_{n-1},\Phi_1) & a(\Phi_{n-1},\Phi_2) & ..... & a(\Phi_{n-1},\Phi_{n-1})
\end{bmatrix} \tag{3.37}$$

where

$$a(\Phi_i,\Phi_j) = \int_0^1 \Phi_i'\Phi_j' dx, (f,\Phi_i) = \int_0^1 f\Phi_i dx \tag{3.38}$$

$$\begin{bmatrix}
\frac{2}{h} & -\frac{1}{2} & . & . & . & . & . & . \\
-\frac{1}{h} & \frac{2}{h} & -\frac{1}{h} & . & . & . & . & . \\
. & -\frac{1}{h} & \frac{2}{h} & -\frac{1}{h} & . & . & . & . \\
. & . & . & . & . & . & . & . \\
. & . & . & . & . & . & . & . \\
. & . & . & . & . & -\frac{1}{h} & \frac{2}{h} & -\frac{1}{h} \\
. & . & . & . & . & . & -\frac{1}{h} & \frac{2}{h}
\end{bmatrix}
\begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ . \\ . \\ . \\ c_{n-1} \end{bmatrix}
=
\begin{bmatrix} (f,\Phi_1) \\ (f,\Phi_2) \\ (f,\Phi_3) \\ . \\ . \\ . \\ (f,\Phi_{n-1}) \end{bmatrix} \tag{3.39}$$

• Solve the linear system($A\xi = b$) of equations for the coefficients and hence obtain the approximate solution $u_N(x) = \sum_{i=1}^{n} c_i\Phi_i(x)$

Carry out the error analysis (sometimes, a posteriori error analysis).

Approximate solution

• Approximate solution is a linear combination of trial functions.

• It satisfies the essential BC(Dirichlet boundary conditions), but not natural BC(Neumann boundary conditions)

• The approximate solution may not satisfy the differential equation(DE) exactly.

Exact solution

• Boundary value problem: differential equation + boundary conditions.
• Displacements in a uniaxial bar subject to a distributed.
In general, it is difficult to find the exact solution when the domain and/or boundary conditions are complicated
• Sometimes the solution may not exists even if the problem is well defined.

$$u_N = \sum_{i=1}^{n} c_i \Phi_i(x)$$

where : $\Phi_i(x)$ is trial function

$$\frac{\partial^2 u}{\partial x^2} + \rho(x) = f(x) \tag{3.40}$$

$$u(0) = o, \frac{\partial u}{\partial x}(1) = 1 \tag{3.41}$$

$$\int_0^1 f(x)\Phi_i(x)dx = 0, i = 1, 2, ...N$$

$$\int_0^1 (\frac{\partial^2 u'}{\partial x^2} + \rho(x))\Phi_i(x)dx = 0, i = 1, 2, ...N \tag{3.42}$$

$$\int_0^1 (\frac{\partial^2 u'}{\partial x^2} dx = -\int_0^1 \rho(x))\Phi_i(x)dx, i = 1, 2, ...N \tag{3.43}$$

using integration by parts

$$\frac{\partial u}{\partial x}\Phi_i \mid_o^1 - \int_0^1 (\frac{\partial u}{\partial x}\frac{\partial \Phi_i}{\partial x} dx = -\int_0^1 f(x))\Phi_i(x)dx \tag{3.44}$$

Apply natural boundary condition.

$$\int_0^1 \frac{\partial \Phi_i}{\partial x}\frac{\partial u'}{\partial x} dx = \int_0^1 f(x))\Phi_i(x)dx + \frac{\partial u}{\partial x}(1)\Phi_i(1) - \frac{\partial u}{\partial x}(0)\Phi_i(0) \tag{3.45}$$

some order of differentiation for both trial function and approximate solution.
substitute the approximate solution.

$$\int_0^1 \frac{\partial \Phi_i}{\partial x} \sum_{i=1}^{n} c_i \Phi_i(x) \frac{\partial \Phi_j}{\partial x} dx = \int_0^1 f(x))\Phi_i(x)dx + \frac{\partial u}{\partial x}(1)\Phi_i(1) - \frac{\partial u}{\partial x}(0)\Phi_i(0) \tag{3.46}$$

write in matrix form.

$$\sum_{j=1}^{n} k_{ij}c_j = F_i, i = 1, 2, ..., N \tag{3.47}$$

$$k_{ij} = A_{ij} = \int_0^1 \frac{\partial \Phi_i}{\partial x}\frac{\partial \Phi_j}{\partial x} dx \tag{3.48}$$

$$F_i = b_i = \int_0^1 f(x))\Phi_i(x)dx + \frac{\partial u}{\partial x}(1)\Phi_i(1) - \frac{\partial u}{\partial x}(0)\Phi_i(0) \tag{3.49}$$

coefficient matrix is symmetric $k_{ij} = k_{ji}$.
N equation with unkown coefficients.
Eaxmple 2
Consider the differential equation.

$$\frac{\partial^2 u}{\partial x^2} + 1 = 0, o \leq x \leq 1 \tag{3.50}$$

$$u(0) = o, \frac{\partial u}{\partial x}(1) = 1$$

solution:
trial function
$\Phi_1(x)$=x
$\Phi_2(x) = x^2$
$\Phi_1'(x)$=1
$\Phi_2'(x)$=2x
Approximate solution (satisfies the essential boundary condition)

$$u_N = \sum_{i=1}^{n} c_i\Phi_i(x) = c_1x + c_2x^2$$

coefficient matrix and R H S vector.

$$k_{11} = \int_0^1 (\Phi_1')^2 dx = 1$$

$$k_{12} = k_{21} = \int_0^1 (\Phi_1'\Phi_2')dx = 1$$

$$k_{22} = \int_0^1 (\Phi_2')^2 dx = \frac{4}{3}$$

$$F_1 = \int_0^1 \Phi_1(x)dx + \frac{\partial u}{\partial x}(1)\Phi_1(1) - \frac{\partial u}{\partial x}(0)\Phi_1(0) = \frac{3}{2}$$

$$F_2 = \int_0^1 \Phi_2(x)dx + \frac{\partial u}{\partial x}(1)\Phi_2(1) - \frac{\partial u}{\partial x}(0)\Phi_2(0) = \frac{4}{3}$$

matrix equation.

$$[k] = \begin{bmatrix} 3 & 3 \\ 3 & 4 \end{bmatrix}$$

$$F = \frac{1}{6}\begin{bmatrix} 9 \\ 8 \end{bmatrix}$$

$$c = [k]^{-1}F = \begin{bmatrix} 2 \\ -\frac{1}{2} \end{bmatrix}$$

Approximate solution

$$\tilde{u}(x) = 2x - \frac{x^2}{2} \tag{3.51}$$

Approximate solution is also the exact solution because the linear combination of the trial functions can represent the exact solution
Example 3
Solve the differential equation.

$$D(u(x)) = u''(x) + u(x) + 2x(1-x) = 0 u(0) = 0, u(1) = 0 \tag{3.52}$$

with the boundary condition:
solution:
Choose trial function $\tilde{u}(x)$=$\Phi_0(x)$+$\sum\limits_{i=1}^{n} c_i\Phi_i(x)$
we make n=3 and Φ_0=0
Φ_1 = x(x-1)
$\Phi_2 = x^2(x-1)^2$
$\Phi_3 = x^3(x-1)^3$
The weight functions are the same as the basis functions $\Phi_i(x)$ Substitute the trial function u(x) into

$$\int_0^1 \Phi_j(Du)dx = \int_0^1 \Phi_j(x)D[\Phi_0(x) + \sum_{i=1}^{n} c_i\Phi_i(x)]dx = 0 \tag{3.53}$$

i=1,2,3; we have three equations with three unknown coefficients
$-\frac{1}{15} - \frac{3c_1}{10} + \frac{5c_2}{84} - \frac{4c_3}{315}$=0

$\frac{1}{70}+\frac{5c_1}{84}-\frac{11c_2}{630}+\frac{61c_3}{13860}$=0

$-\frac{1}{315}-\frac{4c_1}{315}+\frac{61c_2}{1360}-\frac{73c_3}{60060}$=0

Solve this linear equation set, get

$c_1=\frac{1370}{7397}\simeq 0.18521$

$c_2=\frac{50688}{273689}\simeq 0.185203$

$c_3=-\frac{132}{21053}\simeq -0.00626989$

Obtain the approximation solution

u(x)=$\sum_{i=1}^{3} c_i\Phi_i(x)$

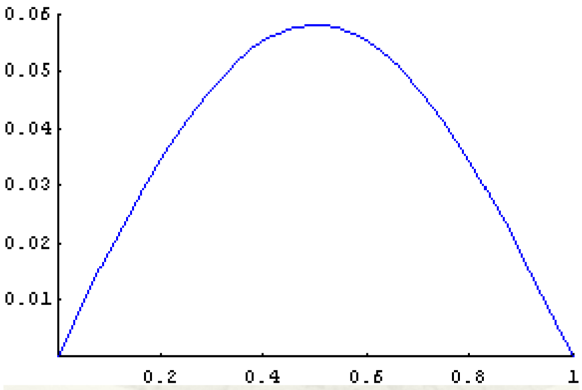


Figure 3.2: Galerkin solution

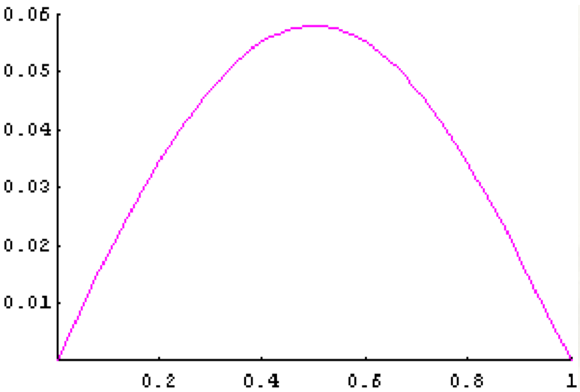


Figure 3.3: Analytic solution

Example 4
consider the boundary value problem

$$\begin{cases} u''(x) + u(x) = f \\ u(0) = 0, u(1) = 0 \end{cases} \tag{3.54}$$

where f$\in L^2(0,1)$and V=$H_0^1(\Omega)$=$\{v \in H_0^1(0,1)|v(0) = 0\}$
The weak form of the problem is

$$-u''v + uv = fv \tag{3.55}$$

$$\int_0^1 u''v + \int_0^1 uv = \int_0^1 fv \tag{3.56}$$

using integration by part

$$\int_0^1 u''vdx = uv|_0^1 - \int_0^1 u'v'dx \tag{3.57}$$

$$\int_0^1 u''vdx = u(1)v(1) - u(0)v(0) - \int_0^1 u'v'dx \tag{3.58}$$

Since v(0)=0,v(1)=0

$$\int_0^1 u'v'dx + \int_0^1 uvdx = \int_0^1 fvdx \quad \forall u, v \in V \tag{3.59}$$

where v=$H_0^1(0,1)$. Applying the Lax-Milgram Lemma, the weak problem has a unique solution.
To develop a Galerkin method, we need to choose a finite-dimensional subspace of V
updating Equation(3.72), we will obtain the weak form of the problem, find $u \in H_0^1(\Omega)$ using Galerkin method **Remark:**u_N can also be written as

$$u_N = \sum_{i=1}^{n} \xi_i \Phi_i(x) = c_1\Phi_1 + c_2\Phi_2 + ...c_n\Phi_n$$

the coefficient $\{\xi_i\}_{i=1}^N$ are determined by the Galerkin method

$$\int_0^1 \sum_{i=1}^{n} \xi_i \Phi_i'(x)v'dx + \int_0^1 \sum_{i=1}^{n} \xi_i \Phi_i(x)vdx = \int_0^1 fvdx \tag{3.60}$$

$$\sum_{j=1}^{n} \xi_i \int_0^1 \Phi_j'(x)\Phi_i'dx + \sum_{j=1}^{n} \xi_i \int_0^1 \Phi_i(x)\Phi_j dx = \int_0^1 f\Phi_j dx \tag{3.61}$$

Write Equation (3.74) as a system of linear equations,

$$\sum_{i=1}^{n} \xi_i \int_0^1 \Phi_i'(x)\Phi_1' dx + \sum_{i=1}^{n} \xi_i \int_0^1 \Phi_i(x)\Phi_1 dx = \int_0^1 f\Phi_1 dx \tag{3.62}$$

$$\vdots \qquad \vdots \qquad \vdots$$

$$\sum_{i=1}^{n} \xi_i \int_0^1 \Phi_i'(x)\Phi_n' dx + \sum_{i=1}^{n} \xi_i \int_0^1 \Phi_i(x)\Phi_n dx = \int_0^1 f\Phi_n dx$$

which

$$\Phi_i(x) = \begin{cases} \frac{x-x_i-1}{x_i-x_{i-1}} & x \in [x_{i-1}, x_i] \\ \frac{x_{i+1}-x}{x_{i+1}-x_i} & x \in [x_i, x_{i+1}] \\ 0 \quad otherwise \end{cases} \tag{3.63}$$

We can also write system of equations (3.75) in matrix form

$$A = \frac{1}{h} \begin{bmatrix} 2 & -1 & 0 & 0 & . & . & . & 0 \\ -1 & 2 & -1 & 0 & . & . & . & 0 \\ 0 & -1 & 2 & -1 & . & . & . & 0 \\ . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . \\ 0 & 0 & . & . & . & . & -1 & 2 \end{bmatrix}$$

3.4 The Petrov-Galerkin method

The Petrov-Galerkin method for a linear boundary value problem can be developed based on the framework of the generalized Lax-Milgram Lemma Let U and V be two real Hilbert spaces, a:$V \times V \to R$ a bilinear form, and $v \in V$.The problem to be solved is

$$a(u, v) = L(v) \quad \forall u \in U, v \in V \tag{3.64}$$

From the generalized Lax-Milgram Lemma, the problem (3.7) has a unique solution $u \in U$, if the following conditions are satisfied:(1.1),(1.3) where $U_N \subset U$, $V_N \subset V$ be finite dimensional subspace of U and V, and $dim(U_N)$= $dim(V_N)$ = N.

Such problem(3.7) can be solved approximately using the Petrov-Galerkin method given by

$$a(u_N, v_N) = L(v_N) \quad \forall u_N \in U_N, v_N \in V_N \tag{3.65}$$

Theorem 3.4.1. *Let assume on the spaces U, V , U_N and V_N, and the forms a(.,.) and L(.). Assume further that there exists a constant $_N > 0$, such that*

$$\sup_{0 \neq v_N \in V_N} a(\frac{(u_N, v_N)}{\| v_N \|_V} \geq \alpha \| u_N \|_U \quad \forall u_N \in U_N \tag{3.66}$$

Then the discrete problem (3.44) has a unique solution u_N, and we have the error estimate

$$\| u - u_N \|_U \leq (1 + \frac{m}{\alpha_N}) \inf_{w_N \in U_N} \| u - w_N \|_U \tag{3.67}$$

Proof. From the generalized Lax-Milgram Lemma, we conclude immediately that under the stated assumptions, the problem (3.78) has a unique solution u_N. Subtracting (3.78) from (3.77) with v = $v_N \in V_N$, we obtain the error relation

$$a(uu_N, v_N) = 0 \quad \forall v_N \in V_N. \tag{3.68}$$

Now for any $w_N \in V_N$, we write

$$\| u - u_N \|_U \leq \| u - w_N \|_U + \| u_N - w_N \| U. \tag{3.69}$$

Using the condition (3.79), we have

$$\alpha_N \| u_N - w_N \|_U \leq \sup_{0 \neq v_N \in V_N} a \frac{(u_N, v_N)}{\| v_N - w_N, v_N \|_V} \tag{3.70}$$

Using the error relation (3.81), we then obtain □

$$\alpha_N \| u_N - w_N \|_U \leq \sup_{0 \neq v_N \in V_N} a \frac{(u_N, v_N)}{\| v_N - w_N, v_N \|_V} \tag{3.71}$$

The right hand side can be bounded by M $\| uw_N \| U$. Therefore

$$\| u_N - w_N \|_U \frac{M}{\alpha_N} \| u - w_N \|_U \tag{3.72}$$

This inequality and (3.82) imply the estimate (3.80).

3.5 Generalized Galerkin method

In the Galerkin method discussed in chapter 3, the finite dimensional space V_N is assumed to be a subspace of V. The resulting numerical method is called an internal approximation method. For certain problems, we will need to relax this assumption and to allow the variational $V_N \subset V$. This, for instance, is the case for non conforming method. There are situations where considerations of other variational. Two such situations are when a general curved domain is approximated by a polygonal domain and when numerical quadratures are used to compute the integrals defining the bilinear form and the linear form. These considerations lead to the following framework of a generalized Galerkin method for the approximate solution of the problem (1.3)

$$a(u_N, v_N) = L(v_N) \quad \forall u_N \in U_N, v_N \in V_N \tag{3.73}$$

where,V_N is a finite dimensional space, but it is no longer assumed to be a subspace of V, the bilinear form $a_N(.,.)$ and the linear form $L_N(.)$ are suitable approximations of a(., .) and L(.). We have the following result related to the approximation method (3.9).

Theorem 3.5.1. *Assume a discretization dependent norm* $\| \, . \, \|_N$, *the approximation bilinear form* $a_N(.,.)$ *and the linear form* $L_N(.)$ *are defined on the space*
$V + V_N = \{w | w = v + v_N, v \in V, v_N \in V_N\}$.
Then the problem (3.9) has a unique solution $u_N \in V_N$, *and we have the error estimate*

$$\| u - u_N \|_N \leq (1 + \frac{M}{\alpha_0}) \inf_{w_N \in V_N} \| u - w_N \|_N + \frac{1}{\alpha_0} \sup_{v_N \in V_N} \frac{| a_N(u, v_N) - L_N(v_N)}{\| v_N \|_N} \tag{3.74}$$

Proof. : the problem (3.9) are unique solvability , Let us derive the error estimate (3.10). For any $w_N \in V_N$, we write of the approximate solution u_N, we have

$$\| u_N - u_N \|_N \leq \| u - w_N \|_N + \| w_N - u_N \|_N \tag{3.75}$$

Using the assumptions on the approximate bilinear form and the definition of the approximate solution u_N , we have

$$\alpha_0 \parallel w_N - u_N \parallel_N^2 \leq a_N(w_N - u_N, w_N - u_N) \tag{3.76}$$

$$= a_N(w_N - u, w_N - u_N) + a_N(u, w_N - u_N) - L_N(w_N - u_N) \tag{3.77}$$

$$\leq M \parallel w_N - u \parallel_N \parallel w_N - u_N \parallel_N + \mid a_N(u, w_N - u_N) - L_N(w_N - u_N) \mid \tag{3.78}$$

Thus

$$\frac{\alpha_0 \parallel w_N - u_N \parallel_N^2}{\parallel w_N - u_N \parallel_N} \leq \frac{M \parallel w_N - u \parallel_N \parallel w_N - u_N \parallel_N}{\parallel w_N - u_N \parallel_N} + \frac{|a_N(u, w_N - u_N) - L_N(w_N - u_N)|}{\parallel w_N - u_N \parallel_N} \tag{3.79}$$

$$\alpha_0 \parallel w_N - u_N \parallel_N \leq M \parallel w_N - u \parallel_N + \frac{\mid a_N(u, w_N - u_N) - L_N(w_N - u_N) \mid}{\parallel w_N - u_N \parallel_N} \tag{3.80}$$

We replace $w_N - u_N$ by v_N and take the supremum of the second term of the right hand side with respect to $v_N \in V_N$ to obtain (3.10).
The estimate (3.10) is a Strang type estimate for the effect of the variational $'crimes'$ on the numerical solution. We notice that in the bound of the estimate (3.10), the first term is on the approximation property of the solution u by functions from the finite dimensional space V_N , while the second term describes the extent to which the exact solution u satisfies the approximate problem.

$$\parallel u - u_N \parallel_N \leq (1 + \frac{M}{\alpha_0}) \inf_{w_N \in V_N} \parallel u - w_N \parallel_N + \frac{1}{\alpha_0} \sup_{v_N \in V_N} \frac{\mid a_N(u, v_N) - L_N(v_N)}{\parallel v_N \parallel_N}$$

□

3.6 Conjugate gradient method: variational formulation

The conjugate gradient method is an iteration method which was originally devised for solving finite linear systems which were symmetric and positive definite. The method has since been generalized in a number of directions, and in this section we consider its generalization to operator equations.

$$Au = f \tag{3.81}$$

In Section 3.5 , we present the variational formulation of the conjugate gradient method assume A is a bounded, positive definite, self adjoint linear operator on a Hilbert space V.
For this purpose, let V be a real Hilbert space with inner product (.,.) and norm $\parallel . \parallel$, and let a(.,.) : $V \times V \to R$ be a continuous, symmetric, V -elliptic bilinear form. Recall that the continuity assumption implies the existence of a constant $M > 0$ such that By definition [1.1,1.2] , Given $L \in v'$ the variational problem is

$$a(u, v) = L(v) \quad \forall v \in V \tag{3.82}$$

Under the stated assumptions, by the Lax-Milgram lemma the problem (3.18) has a unique solution $u \in V$. The conjugate gradient method for solving the variational problem (3.18). For this, we need to write it as an operator equation of the form Au = f. Introduce an operator A : $V \to V$ and an element $f \in V$ defined by

$$(Au, v) = a(u, v) \quad \forall u, v \in V \tag{3.83}$$

$$(f, v) = L(v) \quad \forall v \in V \tag{3.84}$$

The existence of A and f is ensured by (the Riesz representation theorem). From the assumptions on the bilinear form a(.,.) A is bounded, self-adjoint, and positive definite:

$$\parallel A \parallel \leq M$$

$$(Au, v) = (u, Av) \quad \forall u, v \in V,$$

$$(Av, v) \geq \alpha \parallel v \parallel \quad \forall v \in V.$$

Also, $\parallel f \parallel = \parallel L \parallel$ between the V -norm of f and $V' - norm$ of L. With A and f given by (3.19) and (3.20), the problem (3.18) is equivalent to the operator equation

$$Au = f$$

Now the conjugate gradient method for solving Au = f in V is defined as follows. Let u_0 be an initial guess for the solution u^*. Define r_0 = f - Au_0 and $s_0 = r_0$. For k $\geq$ 0, define

$$\begin{aligned} u_{k+1} &= u_k + \alpha_k s_k, \quad \alpha_k = \frac{\parallel r_k \parallel^2}{(As_k s_k)} \\ r_{k+1} &= f - Au_{k+1} \\ s_{k+1} &= r_{k+1} + \beta_k s_k, \quad \beta_k = \frac{\parallel r_{k+1} \parallel^2}{\parallel r_k \parallel^2}. \end{aligned} \tag{3.85}$$

Chapter 4

Test Example

4.1 Solve Example by using Galerkin method 1D problem

$$-u''(x) = f(x), 0 < x < 1, u(0) = 0, u(1) = 0,$$

4.1.1 Matlab code for Galerkin method

```
% Solution using the Galerkin methods of the boundary
% value problem:
%   - u'' = f    in (0,1)
%       u(0) = u(1) = 0

function galerkin_01


% Viewing the basis functions
Nx = 100;
dx = 1.0 / (Nx - 1);
X = 0.0 : dx : 1.0;

figure(1);
plot( X, phi( 1, X ), X, phi( 2, X ), X, phi( 3, X ), X,
phi( 10, X ) );
legend( 'order 1', 'order 2', 'order 3', 'order 10' );



% Solving the system;
```

```
N1 = 50;
dx1 = 1.0 / N1;
X1 = 0.0 : dx1 : 1.0;

A = Assemble( N1 );
rhs = RHS( N1, 10 * N1 );

y = A \ rhs';

% plotting the solution
figure(2);
plot(X, f( X ), '-o', X, sol(X, y, N1), '-x',
X, sol0( X ), '-');
legend('RHS', 'Approx. Sol', 'Exact Sol');

% visualizing the system matrix;
figure(3); surf( log(abs(A)) );

% Analyzing the condition number;
Nmin = 1e1;
Nmax = 1e3;

n = Nmin;
i = 1;
while ( n < Nmax )
  C(i,1) = n;
  C(i,2) = cond( Assemble( n ) );
  n = n * 2;
  i = i + 1;
end;

figure(4); loglog(C(:,1), C(:,2),'-x' );
legend('Galerkin method',);

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
```

```
% The RHS function
function [fun] = f( x )

  fun = - x;

% The RHS vector in the basis chosen;
function [G] = RHS( N, Ni )

dxi = 1.0 / Ni;
Xi = 0.0 : dxi : 1.0;

  for k=1:N
    s = 0.0;
    for m=2:Ni;
     s = s + f( Xi(m) ) * phi( k, Xi(m) );
    end;
    G(k) = ( s + 0.5 * (f( Xi(1) ) * phi( k, Xi(1) ) +
    f( Xi(Ni+1) ) * phi( k, Xi(Ni+1) ) ) ) * dxi;

  end;



% The basis function
function [fun] = phi( k, x )

  fun = x.^k .* ( 1.0 - x );


% Assembling the matrix;
function [A] = Assemble( N )

  for k=1:N
    for m=1:N
% $$$     A(j,k) = (j+1)*(k+1)/(j+k+1) - ((j+1)*(k+2)+(j+2)
% $$$*(k+1))/(j+k+2)
% $$$            ... + (j+2)*(k+2)/(j+k+3);
```

```
    A(k,m) = 2*k*m / (k^3+3*k^2*m+3*k*m^2-k-m+m^3);
  end;
end;


% Calculating the solution based on the expansion coeffients
%in the
% given basis
function [y] = sol( X, G, N )

  y = 0.0;
  for k=1:N
    y = y + G(k) * phi( k, X );
  end;


% Analytical solution for the problem f=-x;
function [y] = sol0( X )

  y = 1.0 / 6.0 * X .* ( X.^2 - 1.0 );
```

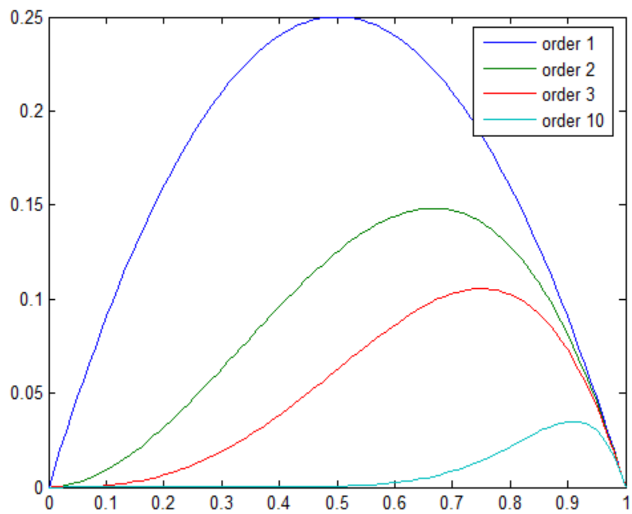


Figure 4.1: basis function

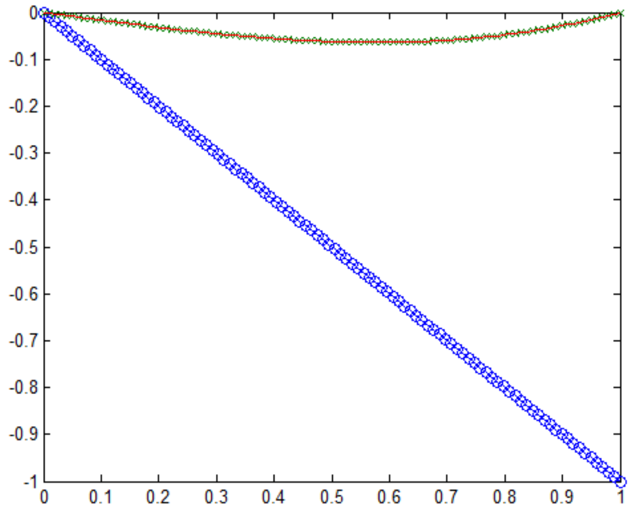


Figure 4.2:

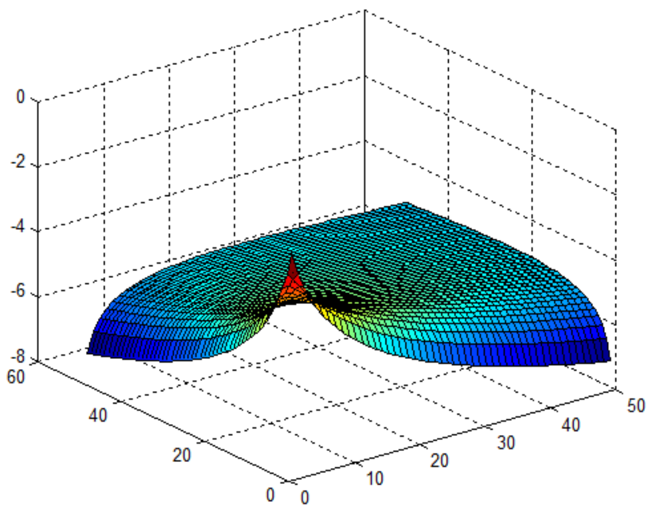


Figure 4.3:

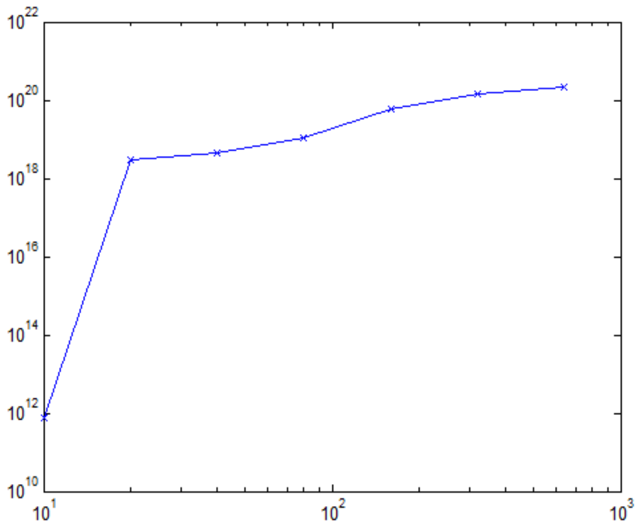


Figure 4.4:

4.1.2 matlab code for Galerkin method

```
% Solution using the Galerkin methods of the boundary
% value problem:
%   - u'' + u = f    in (0,1)
%      u(0) = 0; u'(1) = b;

function fem_01

% Size of the problem;
N1 = 10;

% Viewing the basis functions
Nx = 100;
dx = 1.0 / (Nx - 1);
X = 0.0 : dx : 1.0;

figure(1);
plot( X, phi( N1, 1, X ), X, phi( N1, 2, X ), X,
phi( N1, 3, X ), ...
    X, phi( N1, 10, X ) );
legend( 'order 1', 'order 2', 'order 3', 'order 10' );
```

```
% Solving the system;
dx1 = 1.0 / N1;
X1 = 0.0 : dx1 : 1.0;

A = Assemble( N1 );
rhs = RHS( N1, 10 * N1 );

y = A \ rhs';

% plotting the solution
figure(2);
plot(X, f( X ), '-o', X, sol(X, y, N1), '-x',
X, sol0( X ), '-');
legend('RHS', 'Approx. Sol', 'Exact Sol');

% visualizing the system matrix;
%figure(3); surf( log(abs(A)) );
figure(3); surf( (abs(A)) );

% Analyzing the condition number;
Nmin = 1e1;
Nmax = 1e3;

n = Nmin;
i = 1;
while ( n < Nmax )
  C(i,1) = n;
  C(i,2) = cond( Assemble( n ) );
  C(i,3) = cond( Assemble0( n ) );
  n = n * 2;
  i = i + 1;
end;

figure(4); loglog(C(:,1), C(:,2),'-x', C(:,1), C(:,3),'-o' );
legend('Galerkin method','x^k*(1-x)');


%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
```

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% The RHS function
function [fun] = f( x )

  fun = x;

% The RHS vector in the basis chosen;
function [G] = RHS( N, Ni )

dxi = 1.0 / Ni;
Xi = 0.0 : dxi : 1.0;

  for k=1:N
    s = 0.0;
    for m=2:Ni;
     s = s + f( Xi(m) ) * phi( N, k, Xi(m) );
    end;
    G(k) = ( s + 0.5 * (f( Xi(1) ) * phi( N, k, Xi(1) )
    + f( Xi(Ni+1) ...
                                    ) * phi( N, k, Xi(Ni+1) ) ) )
  end;

% Adding the boundary term;
  G(N) = G(N) + 1.0;



% The basis functions
function [fun] = phi( N, k, x )

  Nx = length( x );

  h = 1.0 / N;
  for i=1:Nx
    if ( (x(i) > (k-1) * h) & (x(i) < k * h) )
      fun(i) = ( x(i) - (k-1) * h ) / h;
```

```
    elseif ( (x(i) > k * h) & (x(i) < (k+1) * h) )
      fun(i) = ( (k+1) * h - x(i) ) / h;
    else
      fun(i) = 0.0;
    end;
  end;


% Assembling the matrix;
function [A] = Assemble( N )

  A(1:N,1:N) = 0.0;
  h = 1.0 / N;
  for k=1:N-1
% The main diagonal
    A(k,k)   = 2.0 * h / 3.0 + 2.0 / h;
% The upper diagonal
    A(k,k+1) = h / 6.0 - 1.0 / h;
% The lower diagonal
    A(k+1,k) = h / 6.0 - 1.0 / h;
  end;
% THe last element;
  A(N,N) = h / 3.0 + 1.0 / h;



% Calculating the solution based on the expansion coeffient
%in the
% given basis
function [y] = sol( X, G, N )

  y = 0.0;
  for k=1:N
    y = y + G(k) * phi( N, k, X );
  end;


% Analytical solution for the problem f=-x;
```

```
function [y] = sol0( X )

  y = X;


% Assembling the matrix (using the  basis functions);
function [A] = Assemble0( N )

  for k=1:N
    for m=1:N
% $$$         A(j,k) = (j+1)*(k+1)/(j+k+1) - ((j+1)*(k+2)+(j+2)
% $$$*(k+1))/(j+k+2) ...
% $$$            + (j+2)*(k+2)/(j+k+3);

%       A(k,m) = 2*k*m / (k^3+3*k^2*m+3*k*m^2-k-m+m^3);
    A(k,m) = 2*(k*m^3+m^2+k^3*m-k+8*k*m+5*k*m^2-m+k^2+
    2*k^2*m^2+5*k^2*m)/(-6*k-6*m+5*k^3+5*m^3+m^5-10*k
    *m+k^5-5*k^2-5*m^2+20*k^3*m+30*k^2*m^2+15*k^2*m+
    15*k*m^2+20*k*m^3+10*k^3*m^2+5*k^4*m+10*k^2*m^3+
    5*k*m^4+5*k^4+5*m^4);
    end;
  end;
```

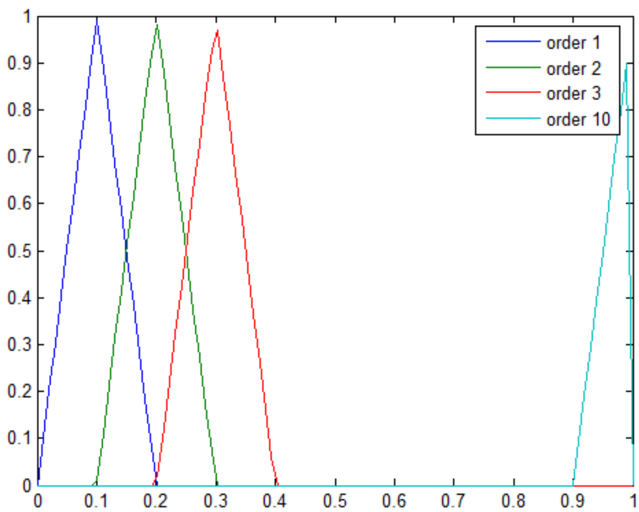


Figure 4.5: basis function

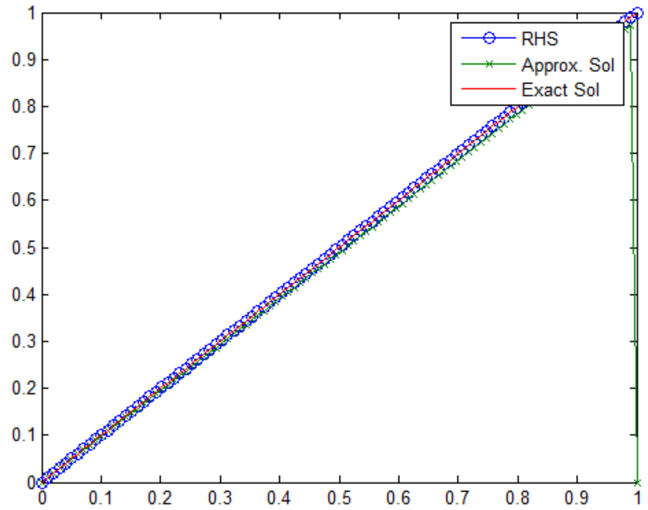


Figure 4.6:

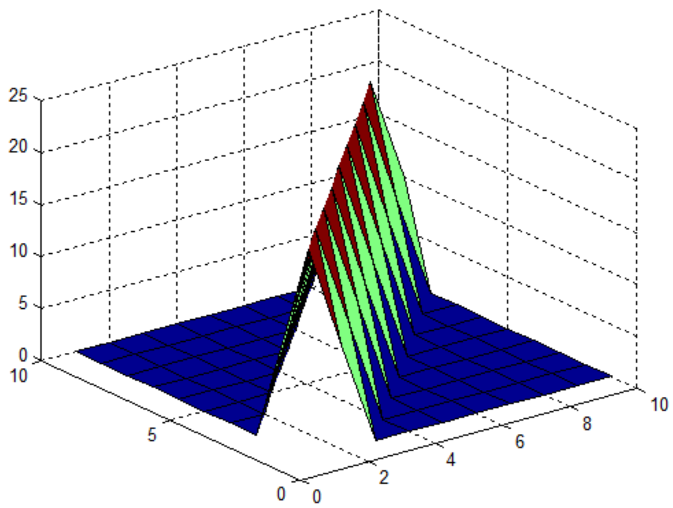


Figure 4.7:

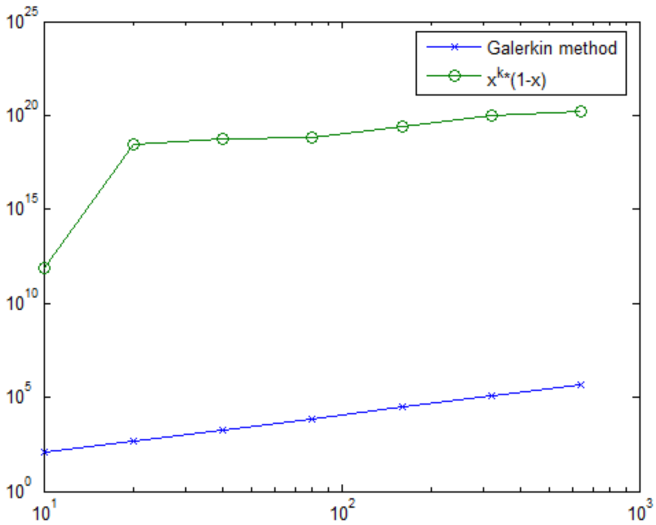


Figure 4.8:

Chapter 5

Conclusion

In this project, we have discussed the ideas about Galerkin method and its variants. Tt has a theoretical foundation and estimate the error in the approximate solution. The generalized formulation, its dependence on the boundary condition, treatment of Dirichlet and Neumann condition, the approximate solution and hence the error of approximation is determined by Galerkin method. This paper presents the basis understanding of Galerkin method and the methodology to solve any problem of differential equation. The main idea of the Galerkin method second order it to choose the basis function and then expressing the unknown as combination of the basis function. Finally a Galerkin matrix is generated and solution is obtained, calculations become difficult and hence the computation time also increase, Galerkin method can be executed computationally on Matlab.

References

[1] Atkinson/Han: Theoretical Numerical Analysis: A Functional Analysis Framework, 2nd ed.

[2] S. G. Mikhlin, ”Variational methods in Mathematical Physics”, Pergamon Press, 1964

[3] Adams, R. A., 1975: Sobolev spaces. Academic Press [A subsidiary of Harcourt Brace Jovanovich, Publishers], New York-London, xviii+268 pp., pure and Applied Mathematics, Vol. 65.

[4] P. G. Ciarlet. The finite element method for elliptic problems. North Holland, 1978, and SIAM Classic in Applied Mathematics 40, 2002.

[5] S.C. Brenner and L.R. Scott. The Mathematical Theory of Finite Element Methods. Springer New York, 2002.

[6] J. Adams, P. Swarztrauber, and R. Sweet. Fishpack: Efficient Fortran subprograms for the solution of separable elliptic partial differential equations. http://www.netlib.org/fishpack/.

[7] A. Ern and J.-L. Guermond, Theory and Practice of Finite Elements, Springer (AMS 159), (2003).