1D Model Problem: Galerkin and Ritz Methods:

So far, we have discussed various forms of our given one-dimensional model problem. Today, we discuss how to arrive at approximations of our model problem by discretizing the variational and minimization forms.

To begin, recall the variational form of our model problem:

$$(V) \begin{cases} \text{Find } u \in \mathcal{A} \text{ such that:} \\ \quad \int_0^L K u_{,x} w_{,x} \, dx = \int_0^L f w \, dx \\ \text{for all } w \in \mathcal{V}. \end{cases}$$

Above:

$$\mathcal{V} := \{ v \in H^1((0,L)) : v(0) = v(L) = 0 \}$$

is the space of test functions and:

$$\mathcal{A} := \{ v \in H^1((0,L)) : v(0) = g_0, v(L) = g_L \}$$

is the set of trial solutions. We refer to $\mathcal{V}$ as a space as it is a vector space as linear combinations of members of $\mathcal{V}$ also belong to $\mathcal{V}$, that is:

$$\sum_{i=1}^{m} c_i v_i \in \mathcal{V}$$

for any $\{v_i\}_{i=1}^m \subset \mathcal{V}$ and $\{c_i\}_{i=1}^m \subset \mathbb{R}$. Likewise, $\mathcal{U}$ is referred to as a set as linear combinations of members of $\mathcal{U}$ are generally not members of $\mathcal{U}$. To see this, note for two functions v_1, v_2 of $\mathcal{U}$:

$$v_1(0) + v_2(0) = 2g_0$$
$$v_1(L) + v_2(L) = 2g_L$$

Thus, unless $g_0 = g_L = 0$, then $v_1 + v_2 \notin \mathcal{U}$. However, while $\mathcal{U}$ is not technically not a vector space, it still has vector space-like structure. In particular, given $g \in \mathcal{U}$, we can express any $u \in \mathcal{U}$ as:

$$u = v + g$$

where $v = u - g \in \mathcal{V}$. We write:

$$\mathcal{U} = \mathcal{V} + g$$

to denote the fact that $\mathcal{U}$ is composed of all sums of the form $v+g$ where $v \in \mathcal{V}$. This is referred to as set addition.

Now, note that both $\mathcal{U}$ and $\mathcal{V}$ are infinite-dimensional. To arrive at an approximation of our model problem using a Galerkin method, we approximate $\mathcal{U}$ and $\mathcal{V}$ using finite-dimensional sets $\mathcal{U}^h \subset \mathcal{U}$ and $\mathcal{V}^h \subset \mathcal{V}$. In particular, we choose $\mathcal{V}^h$ to be a finite-dimensional subspace of $\mathcal{V}$, we choose $\mathcal{U}^h$ to be $\mathcal{V}^h + g^h$

where $\hat{\mathcal{V}}^h$ is a finite-dimensional subspace of $\mathcal{V}$ and g^h is a member of $\mathcal{G}$, and we require that the dimensions of $\mathcal{V}^h$ and $\hat{\mathcal{V}}^h$ be equal. We denote the dimension of $\mathcal{V}^h$ and $\hat{\mathcal{V}}^h$ as n:

$$n := \dim(\mathcal{V}^h) = \dim(\hat{\mathcal{V}}^h)$$

The Galerkin approximation of our model problem associated with the finite-dimensional set of trial solutions $\mathcal{G}^h$ and finite-dimensional space of test functions $\mathcal{V}^h$ then reads:

$$(G) \begin{cases} \text{Find } u^h \in \mathcal{G}^h \text{ such that:} \\ \displaystyle\int_0^L K u^h_{,x} w^h_{,x} \, dx = \int_0^L f w^h \, dx \\ \text{for all } w^h \in \mathcal{V}^h. \end{cases}$$

A solution to Problem (G) is referred to as a Galerkin solution. It should be emphasized that while we presented above the construction of a Galerkin method for the variational form of our model problem, the same procedure can be employed to arrive at a Galerkin method for the strong or weak weak forms of our model problem. In fact, we can construct Galerkin methods for any weak form of a given partial differential equation. The only required ingredient is the selection of a suitable set of trial solutions and a suitable space of weighting functions.

There are two general classes of Galerkin methods:

Bubnov-Galerkin Methods: Bubnov-Galerkin methods correspond to the selection $\mathcal{V}^h = \hat{\mathcal{V}}^h$.

Petrov-Galerkin Methods: Petrov-Galerkin methods correspond to the selection $\mathcal{V}^h \neq \hat{\mathcal{V}}^h$.

Bubnov-Galerkin methods are more common than Petrov-Galerkin methods and are the focus of this class. Petrov-Galerkin methods are commonly employed in the modeling of transport phenomena, where Bubnov-Galerkin methods are known to be unstable. It should be noted that for our given model problem, Bubnov-Galerkin methods are always well-posed while Petrov-Galerkin methods are subject to a so-called inf-sup condition. We will later prove the well-posedness of Bubnov-Galerkin methods for our model problem, and we will later discuss inf-sup conditions in the context of mixed finite element methods for plane strain linear elastostatics.

Note the key ingredient in any Galerkin method is the selection of suitable test and trial functions. Indeed, there are many possible choices. For instance, a natural choice for the finite-dimensional space of test functions is the space of polynomials of degree k

satisfying the prescribed boundary conditions:

$$\mathcal{V}^h := \left\{ v^h = \sum_{A=0}^{k} c_A x^k \;:\; \{c_A\}_{A=1}^{k} \subset \mathbb{R} \text{ and } v^h(0) = v^h(L) = 0 \right\}$$

Another choice is the space of Fourier series satisfying the prescribed boundary conditions truncated after k terms:

$$\mathcal{V}^h := \left\{ v^h = \sum_{A=1}^{k} c_A \sin\left(\frac{A\pi x}{L}\right) : \{c_A\}_{A=1}^{k} \subset \mathbb{R} \right\}$$

Both selections above yield spectrally accurate approximations of smooth weak solutions - their error in standard norms decay to zero spectrally fast as the degree k or number of terms k increases. As such, Galerkin methods employing polynomial or Fourier test and trial functions are commonly referred to as spectral methods. However, spectral methods suffer from a number of pronounced disadvantages. First of all, they produce spurious overshoots and undershoots when applied to problems exhibiting layer-like phenomena. Second of all, they are quite inaccurate when applied to rough solutions. Finally, and perhaps most importantly, they are difficult to extend to domains with complex geometry, especially in the three-dimensional setting. For these reasons, we will turn to finite element test and trial functions in this class.

While it may not be apparent yet, Problem (G) is simply a linear system of n equations for n unknowns. To see this, first recall that each finite-dimensional vector space admits

a finite basis whose size is equal to the space's dimension. Therefore, the space $\mathcal{V}^h$ admits a basis $\{\psi_A\}_{A=1}^n$ and the space $\hat{\mathcal{V}}^h$ admits a basis $\{\phi_B\}_{B=1}^n$. Moreover, each weighting function $w^h \in \mathcal{V}^h$ may be written as:

$$w^h = \sum_{A=1}^{n} c_A \psi_A$$

where the coefficients $\{c_A\}_{A=1}^n$ are arbitrary real numbers, and the solution $u^h \in \mathcal{S}^h$ to Problem (G) may be written as:

$$u^h = \sum_{B=1}^{n} d_B \phi_B + g^h$$

where the coefficients $\{d_B\}_{B=1}^n$ are unknown degrees of freedom to be determined. If we plug the above expressions for w^h and u^h into Problem (G), we obtain the equation:

$$\int_0^L K\left(\sum_{B=1}^{n} d_B \phi_B + g^h\right)_{,x} \left(\sum_{A=1}^{n} c_A \psi_A\right)_{,x} dx = \int_0^L f\left(\sum_{A=1}^{n} c_A \psi_A\right) dx$$

Since differentiation is a linear operator, we have:

$$\int_0^L K\left(\sum_{B=1}^{n} d_B \phi_{B,x} + g^h_{,x}\right)\left(\sum_{A=1}^{n} c_A \psi_{A,x}\right) dx = \int_0^L f\left(\sum_{A=1}^{n} c_A \psi_A\right) dx$$

By linearity, we also have:

$$\sum_{A=1}^{n} c_A \int_0^L K\left(\sum_{B=1}^{n} d_B \phi_{B,x} + g^h_{,x}\right) \psi_{A,x}\, dx = \sum_{A=1}^{n} c_A \int_0^L f\, \psi_A\, dx$$

Re-arranging terms yields:

$$\sum_{A=1}^{n} c_A \left(\int_0^L K \left(\sum_{B=1}^{n} d_B \phi_{B,x} + g^h_{,x} \right) \psi_{A,x} \, dx - \int_0^L f \psi_A \, dx \right) = 0$$

Since the coefficients $\{c_A\}_{A=1}^{n}$ are arbitrary, it must hold that:

$$\int_0^L K \left(\sum_{B=1}^{n} d_B \phi_{B,x} + g^h_{,x} \right) \psi_{A,x} \, dx - \int_0^L f \psi_A \, dx = 0$$

for $A = 1, \ldots, n$. Finally, exploit linearity once more to write:

$$\sum_{B=1}^{n} \left(\int_0^L K \psi_{A,x} \phi_{B,x} \, dx \right) d_B = \int_0^L f \psi_A \, dx - \int_0^L K g^h_{,x} \psi_{A,x} \, dx$$

for $A = 1, \ldots, n$. If we define $\underline{\underline{K}}$ to be the $n \times n$ matrix with entries:

$$K_{AB} = \int_0^L K \psi_{A,x} \phi_{B,x} \, dx$$

$\underline{F}$ to be the $n \times 1$ vector with entries:

$$F_A = \int_0^L f \psi_A \, dx - \int_0^L K g^h_{,x} \psi_{A,x} \, dx$$

and $\underline{d}$ to be the $n \times 1$ vector with entries d_B, then we see have arrived at a linear system of n equations for n unknowns as promised. In particular, we have arrived at the problem:

$$(L) \begin{cases} \text{Find } \underline{d} \in \mathbb{R}^n \text{ such that:} \\ \underline{\underline{K}}\,\underline{d} = \underline{F} \end{cases}$$

Problem (L) is equivalent to Problem (G). Indeed, if we know the solution $\underline{d} \in \mathbb{R}^n$ to Problem (L), the solution $u^h \in \mathcal{A}^h$ to Problem (G) is simply:

$$u^h = \sum_{B=1}^{n} d_B \phi_B + g^h$$

This is the power of Galerkin methods - once an algebraic system of equations is solved, one attains the Galerkin solution. Of course, this statement hides a lot of details. In particular:

(i) One first has to choose $\mathcal{V}^h$, $\tilde{\mathcal{V}}^h$, and g^h.

(ii) One next has to choose basis functions $\{\psi_A\}_{A=1}^n$ and $\{\phi_B\}_{B=1}^n$.

(iii) One then has to evaluate the integral expressions defining $\underline{\underline{K}}$ and $\underline{F}$.

(iv) One finally has to solve the linear system $\underline{\underline{K}}\,\underline{d} = \underline{F}$ for $\underline{d}$.

We will discuss all of the above steps in quite some detail for the particular case of a Bubnov-Galerkin finite element method. We will repeat this discussion for a series of increasingly complex

applications in the remainder of this class.

EXAMPLE

To illustrate the above steps for an example problem, consider the simple case when:

$$K = 1, \quad f = 1, \quad L = 1, \quad g_0 = 0, \quad g_L = 0$$

The variational form of the problem at hand is then:

$$(V) \begin{cases} \text{Find } u \in \mathcal{V} \text{ such that:} \\ \int_0^1 u_{,x} w_{,x} \, dx = \int_0^1 w \, dx \\ \text{for all } v \in \mathcal{V} \text{ where:} \\ \mathcal{V} := \{ v \in H^1((0,1)) : v(0) = v(1) = 0 \} \end{cases}$$

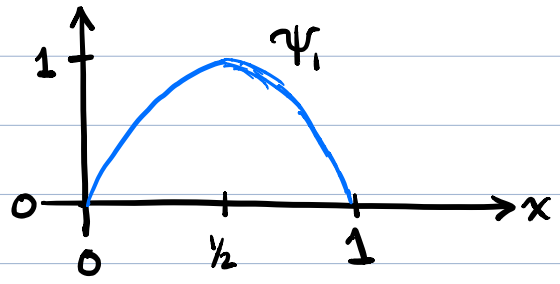
Suppose a Bubnov-Galerkin method is used to approximate the above problem using quadratic test and trial functions. Then the finite-dimensional space of test functions is one-dimensional and equal to:

$$\mathcal{V}^h := \{ v^h = c x (1-x) : c \in \mathbb{R} \}$$

and a corresponding basis for $\mathcal{V}^h$ consists of just one function. We choose for this function:

$$\psi_1 = 4x(1-x)$$

Visually:



Since we have a Bubnov-Galerkin method, $\phi_1 = \psi_1$, and since $g_a = g_b = 0$, the solution to Problem (G) is simply equal to:

$$u^h(x) = d_1 \psi_1(x)$$

where d_1 is the solution of Problem (L) which reduces to a single linear equation in the current context:

$$K_{11} d_1 = F_1$$

where:

$$K_{11} = \int_0^1 \psi_{1,x} \psi_{1,x} \, dx$$

$$F_1 = \int_0^1 \psi_1 \, dx$$

A quick calculation yields:

$$K_{11} = 16/3, \quad F_1 = 2/3$$

So:

$$d_1 = F_1 / K_{11} = 1/8$$

It follows then that:

$$u^h(x) = d_1 \psi_1(x) = \tfrac{1}{2} x(1-x)$$

Coincidentally, the above is not only the Galerkin solution for this case but also the exact solution to Problem (V)!

The matrix $\underline{\underline{K}}$, the vector $\underline{d}$, and the vector $\underline{F}$ appearing in Problem (L) are commonly referred to as the stiffness matrix, displacement vector, and forcing vector, respectively. This is due to the origins of finite element analysis in the structural engineering community.

Note that Problem (L) is well-posed if and only if $\underline{\underline{K}}$ is invertible. This in turn depends on the selections of $\mathcal{V}^h$ and $\hat{\mathcal{V}}^h$. When a Bubnov-Galerkin method is employed, then $\mathcal{V}^h = \hat{\mathcal{V}}^h$, and one can choose $\phi_A = \psi_A$ for $A = 1, \dots, n$. In this case, the stiffness matrix $\underline{\underline{K}}$ satisfies two key properties:

Property 1: $\underline{\underline{K}}$ is symmetric:

$$\underline{\underline{K}} = \underline{\underline{K}}^T$$

Property 2: $\underline{\underline{K}}$ is positive-definite:

$$\underline{x}^T \underline{\underline{K}} \underline{x} > 0 \quad \text{for all } \underline{x} \in \mathbb{R}^n$$

As $\underline{\underline{K}}$ is symmetric and positive-definite, it is invertible. Thus, Problem (L) is well-posed, and as a consequence, so is Problem (G), as previously claimed. Since the test and trial basis functions are the same for a Bubnov-Galerkin method, we will use the same notation for these for the remainder of the class, namely $\{N_A\}_{A=1}^{n}$.

So far, we have discussed discretization of the variational form of our model problem. Now, we discuss discretization of the minimization form of our model problem:

$$(M) \begin{cases} \text{Find:} \\ \quad u := \underset{v \in \mathcal{A}}{\operatorname{argmin}} \; E(v) \end{cases}$$

Above:

$$E(v) := \frac{1}{2} \int_0^L \kappa (v,_x)^2 \, dx - \int_0^L f v \, dx$$

To arrive at an approximation of the minimization form using a Ritz method, we again approximate $\mathcal{A}$ using a finite-dimensional set $\mathcal{A}^h \subset \mathcal{A}$ of the form $\mathcal{V}^h + g^h$ where $\mathcal{V}^h$ is a finite-dimensional subspace of $\mathcal{V}$ and g^h is a member of $\mathcal{A}$. The Ritz approximation of our model problem using the given finite-dimensional set of trial solutions $\mathcal{A}^h$ then takes the form:

$$(R) \quad \begin{cases} \text{Find:} \\ \qquad u^h := \underset{v^h \in \mathcal{A}^h}{\operatorname{argmin}} \; E(v^h) \end{cases}$$

The solution of Problem (R) is called a Ritz solution. The minimizer of Problem (R), just like the minimizer of Problem (M), may be found by taking the first variation of $E(u^h)$ with respect to a direction $w^h \in \mathring{\mathcal{V}}^h$ and setting it to zero:

$$\delta_{w^h} E(u^h) = 0 \qquad \text{for all } w^h \in \mathring{\mathcal{V}}^h$$

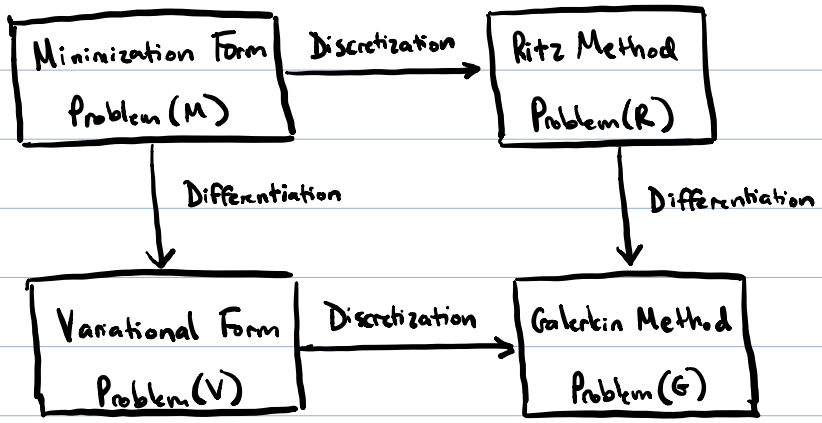
A series of calculations shows that:

$$\delta_{w^h} E(u^h) = \int_0^L K u^h_{,x} w^h_{,x} \, dx - \int_0^L f w^h \, dx$$

Thus a solution $u^h \in \mathcal{A}^h$ of Problem (R) satisfies:

$$\int_0^L K u^h_{,x} w^h_{,x} \, dx = \int_0^L f w^h \, dx$$

for all $w^h \in \mathcal{V}^h$. Therefore, the solution $u^h \in \mathcal{A}^h$ of Problem (R) is also a solution to Problem (G) for the specific choice of $\mathcal{V}^h = \mathring{\mathcal{V}}^h$. That is, a Ritz method is simply a Bubnov - Galerkin method in disguise! The above result is a consequence of the fact that discretization (i.e., approximating the space of test functions) and differentiation (i.e., taking a variation) commute in this context:



Since Ritz methods are simply Bubnov-Galerkin methods in disguise, the latter are often referred to as Ritz-Galerkin methods. In fact, Boris Galerkin himself referred to Bubnov-Galerkin methods as Ritz methods. However, I prefer the terminology Bubnov-Galerkin as Bubnov-Galerkin methods can be constructed for any partial differential equation of interest while Ritz methods can only be constructed for partial differential equations associated with a total potential energy minimization principle.

As Ritz methods are simply Bubnov Galerkin methods in disguise, they also give rise to linear systems of the form:

$$\underline{\underline{K}}\,\underline{d} = \underline{F}$$

To see this, first note for a Ritz method:

$$\delta_{N_A} E(u^h) = 0$$

for all $A = 1, \ldots, N$. Since E is a quadratic functional, we can exactly express the above first variation with a two-term Taylor series expansion about g^h:

$$\delta_{N_A} E(u^h) = \delta_{N_A} E(g^h) + \delta_{N_B}\left(\delta_{N_A} E(g^h)\right) d_B$$

Thus defining:

$$K_{AB} = \delta_{N_B}\left(\delta_{N_A} E(g^h)\right) \quad \text{and} \quad F_A = \delta_{N_A} E(g^h)$$

we have arrived at the stated linear system. In fact, it is easily shown this linear system is the same as that stated in Problem (L) for a Bubnov-Galerkin method. Finally note:

$$\delta_{N_B}\left(\delta_{N_A} F(g^h)\right) = \delta_A\left(\delta_B F(g^h)\right)$$

for any functional $F: \mathcal{A} \to \mathbb{R}$. Thus the stiffness matrix associated with any Ritz method is symmetric, not just the Ritz method presented here for our 1D model problem.

In this class, we will nearly exclusively concern ourselves with Bubnov-Galerkin methods. As such, unless otherwise specified, when I say Galerkin method, I mean Bubnov-Galerkin method. As a final note, observe that a Galerkin solution is only an approximation of the true solution of the problem at hand. To arrive at an accurate approximation, one must use a sufficiently rich set of trial solutions.