Weighted Residual Methods

A. Salih
Department of Aerospace Engineering
Indian Institute of Space Science and Technology, Thiruvananthapuram
– December 2016 –

1 Introduction

Weighted residual method is a generic class of method developed to obtain approximate solution to the differential equations of the form

$$\mathscr{L}(\phi) + f = 0 \quad \text{in} \quad D \tag{1}$$

where $\phi(\mathbf{x})$ is the dependent variable and is unknown and $f(\mathbf{x})$ is a known function. $\mathscr{L}$ denotes the differential operator involving spatial derivative of ϕ, which specifies the actual form of the differential equation.

Weighted residual method involves two major steps. In the first step, an approximate solution based on the general behavior of the dependent variable is assumed. The assumed solution is often selected so as to satisfy the boundary conditions for ϕ. This assumed solution is then substituted in the differential equation. Since the assumed solution is only approximate, it does not in general satisfy the differential equation and hence results in an error or what we call a *residual*. The residual is then made to vanish in some average sense over the *entire* solution domain to produce a system of algebraic equations. The second step is to solve the system of equations resulting from the first step subject to the prescribed boundary condition to yield the approximate solution sought.

Let $\psi(\mathbf{x}) \approx \phi(\mathbf{x})$, is an approximate solution to the differential equation (15). When $\psi(\mathbf{x})$ is substituted in the differential equation (15), it is unlikely that the equation is satisfied. That is, we have

$$\mathscr{L}(\psi) + f \neq 0.$$

Or we may write

$$\mathscr{L}(\psi) + f = R \tag{2}$$

where $R(\mathbf{x})$ is a measure of error commonly referred to as the residual.

Multiply equation (15) by an arbitrary *weight function* $w(\mathbf{x})$ and integrating over the domain D to obtain

$$\int_D w[\mathscr{L}(\phi) + f]\, dD = 0. \tag{3}$$

Equations (15) and (3) equivalent. Replacing ϕ by ψ in equation (3) results in

$$\int_D w(\mathbf{x})\, [\mathscr{L}(\psi) + f]\; dD = \int_D w(\mathbf{x})\, R(\mathbf{x})\, dD \neq 0. \tag{4}$$

The integral in (4) gives the weighted average of the residual over the solution domain. In weighted residual method we force this integral to vanish over the solution domain. That is,

$$\int_D w(x)\, R(x)\, dD = 0. \tag{5}$$

We now seek the approximate solution in the form a generalized Fourier series, say

$$\psi(\mathbf{x}) = \sum_{i=1}^{n} c_i N_i(\mathbf{x}) = c_1 N_1(\mathbf{x}) + c_2 N_2(\mathbf{x}) + \cdots\cdots + c_n N_n(\mathbf{x}). \tag{6a}$$

In vector form

$$\psi(\mathbf{x}) = \mathbf{C}^T \mathbf{N}^T = (\mathbf{N}\mathbf{C})^T = \mathbf{N}\mathbf{C} \tag{6b}$$

where $\mathbf{N}$ is the row vector

$$\mathbf{N} = \begin{bmatrix} N_1 & N_2 & \cdots\cdots & N_n \end{bmatrix}$$

and $\mathbf{C}$ is the column vector

$$\mathbf{C} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}.$$

Here c_i's are unknown coefficients called *fitting coefficients* and n is the number of fitting coefficients. $N_i(\mathbf{x})$'s are assumed to be linearly independent functions of $\mathbf{x}$ and are called *trial functions*. The trial functions can be polynomials, trigonometric functions etc. The trial functions are usually chosen in such a way that the assumed function $\psi(\mathbf{x})$ satisfies the global boundary conditions for $\phi(\mathbf{x})$, although this not strictly necessary and certainly not always possible.

Polynomial Approximation. One of the simplest choices for a trial function is a polynomial, for a one-dimensional problem which can be obtained by taking $N_i(x) = x^i$. The result is

$$\psi(x) = \sum_{i=0}^{n} c_i x^i = c_0 + c_1 x + \cdots + c_n x^n.$$

This produces a smooth solution, but it suffers the same limitations as Lagrange interpolation. A particularly significant flaw is that this choice need not converge to $\phi(x)$ as n increases.

Trigonometric Approximation. Another often used set of trial function is trigonometric approximation based on Fourier series. An example is a Fourier sine series obtained by taking $N_k(x) = \sin\frac{k\pi x}{L}$. For a one-dimensional problem,

$$\psi(x) = \sum_{k=1}^{n} c_k \sin\frac{k\pi x}{L}.$$

Because $\sin(k\pi x/L)$ at $x = 0$ and $\sin(k\pi x/L)$ at $x = L$, this expansion requires the boundary conditions $y(0) = y(L) = 0$. This is not much of a restriction, because one can always make the change of variables so that the boundary conditions become homogeneous.

With the selection of $\psi(x)$ as the series expansion (6), it is evident that the residual R depends on the unknown parameters c_i's in the expansion:

$$R = R(\mathbf{x}; \mathbf{C}).$$

If the number of trial functions n is sufficiently large, then in principle, the unknown parameters c_i's can be chosen so that the residual R is small over the domain.

Weight functions. In general the weight function $w(\mathbf{x})$ may be written as

$$w(\mathbf{x}) = \sum_{i=1}^{n} a_i w_i = a_1 w_1 + a_2 w_2 + \cdots\cdots + a_n w_n = \mathbf{a}\mathbf{w} \tag{7}$$

where $\mathbf{a}$ and $\mathbf{w}$ are row and column vector given respectively by

$$\mathbf{a} = \begin{bmatrix} a_1 & a_2 & \cdots\cdots & a_n \end{bmatrix}, \qquad \mathbf{w} = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix}.$$

Here w_i's are known functions of $\mathbf{x}$ and a_i's are constant parameters. Substituting $w(x) = \mathbf{a}\mathbf{w}$ in the weighted residual equation (5) to yield

$$\mathbf{a} \int_D \mathbf{w} R \, dD = 0.$$

Since $\mathbf{a}$ is a constant vector, we have

$$\int_D \mathbf{w} R \, dD = 0 \tag{8a}$$

or

$$\begin{aligned} \int_D w_1 R \, dD &= 0 \\ \vdots \qquad & \quad \vdots \\ \int_D w_n R \, dD &= 0 \end{aligned} \tag{8b}$$

Now we have n equations to determine unknown coefficients c_i's. Finally, inserting $\psi = \mathbf{N}\mathbf{C}$ in equation (2) yields

$$R = \mathscr{L}(\mathbf{N}\mathbf{C}) + f = \mathscr{L}(\mathbf{N})\,\mathbf{C} + f \tag{9}$$

and hence the condition (8a) becomes

$$\left[\int_D \mathbf{w} \mathscr{L}(\mathbf{N}) \, dD \right] \mathbf{C} = -\int_D \mathbf{w} f \, dD. \tag{10a}$$

Introducing matrix $\mathbf{K}$ and $\mathbf{f}$ as

$$\mathbf{K} = \int_D \mathbf{w} \mathscr{L}(\mathbf{N}) \, dD, \qquad \mathbf{f} = -\int_D \mathbf{w} f \, dD$$

allows us to write equation (8) in compact form as

$$\mathbf{K}\mathbf{C} = \mathbf{f} \tag{10b}$$

which may be expanded as

$$\begin{bmatrix} \int_D w_1 \mathscr{L}(N_1)\, dD & \int_D w_1 \mathscr{L}(N_2)\, dD & \cdots & \int_D w_1 \mathscr{L}(N_n)\, dD \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ \int_D w_n \mathscr{L}(N_1)\, dD & \int_D w_n \mathscr{L}(N_2)\, dD & \cdots & \int_D w_n \mathscr{L}(N_n)\, dD \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} = - \begin{bmatrix} \int_D w_1 f \, dD \\ \int_D w_2 f \, dD \\ \vdots \\ \int_D w_n f \, dD \end{bmatrix}. \tag{10c}$$

The system of equation given by (10) can be solved for n unknown coefficients c_i's provided that a suitable weight function w is selected.

With regards to the selection of weight function, we have several choices. Hence, depending upon nature of weight function, we have different types of weighted residual methods. Some of the standard methods are:

1. Point Collocation Method
2. Subdomain Collocation Method
3. Least Square Method
4. Galerkin Method

2 Point Collocation Method

In point collocation method, the weight function is selected in such a way that the residual can be set equal to zero at n distinct points in the domain. This can be achieved by choosing weight function as the displaced Dirac delta function. So, for one-dimensional case,

$$w_i = \delta(x - x_i) = \begin{cases} \infty, & \text{if } x = x_i \\ 0, & \text{else} \end{cases} \tag{11}$$

where the fixed points $x_i \in [a,b]$, $(i = 1,2,\cdots,n)$ are called collocation points. The number of collocation points selected must be equal to the number of unknown coefficients c_i's in the definition of approximating function, $\psi(x)$.

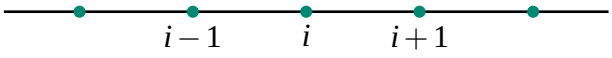

Figure 1: Collocation points in a one-dimensional domain.

The displaced Dirac delta function has the property that

$$(w_i, R) = \int_a^b \delta(x - x_i)\, R\, dx = R(x_i).$$

Thus, from equation (8) we have

$$R(x_i) = 0, \quad i = 1,2,\cdots,n \tag{12}$$

i.e., the residual $R(x)$ is forced to be zero at n collocation points. For the point collocation method the linear system of equation (10) takes the form

$$\begin{bmatrix} \mathscr{L}(N_1(x_1)) & \mathscr{L}(N_2(x_1)) & \cdots & \mathscr{L}(N_n(x_1)) \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ \mathscr{L}(N_1(x_n)) & \mathscr{L}(N_2(x_n)) & \cdots & \mathscr{L}(N_n(x_n)) \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} = - \begin{bmatrix} f(x_1) \\ f(x_2) \\ \vdots \\ f(x_n) \end{bmatrix}. \tag{13}$$

Note: It can be shown that the point collocation method is equivalent to the classical finite difference method.

Example 1

Let us illustrate the application of point collocation method using a simple physical problem. We consider a simply supported beam subjected to concentrated moments at both ends. The problem is governed by the following differential equation

$$EI\frac{d^2y}{dx^2} - M_0 = 0, \qquad x \in [0,L] \tag{14}$$

with boundary conditions (support condition in this case)

$$y(0) = 0 \qquad \& \qquad y(L) = 0.$$

Here, the coefficient EI represents the resistance of the beam to deflection, M_0 is the applied moment, and L is the length of the beam.

$$y(x) = \frac{M_0}{2EI}x(x-L) \tag{15}$$

The analytical solution of the problem in the interval $[0,L]$ is

Figure 2: Simply supported beam subject to bending moments

$$y(x) = -\frac{M_0}{2EI}x(L-x). \tag{16}$$

The negative sign in the expression shows that the displacement is negative for positive values of bending moment, M_0.

Trigonometric approximation to deflection curve. Let us pretend that we do not know the solution and select the approximating function $u(x)$ as a sinusoidal function of the form

$$u(x) = A\sin Bx$$

where A and B are constants. The function which satisfies the prescribed boundary conditions can be obtained by the application of boundary conditions to the chosen approximating function. Thus, we have

$$u(x) = A\sin\frac{\pi x}{L} = c_1N_1 \tag{17}$$

where $c_1 = A$ and $N_1 = \sin\frac{\pi x}{L}$. The second derivative of the assumed function,

$$\frac{d^2u}{dx^2} = c_1\frac{d^2N_1}{dx^2} = -\frac{A\pi^2}{L^2}\sin\frac{\pi x}{L}$$

Substitution the above expression for the second derivative into the (14) gives the residual R. That is,

$$R(x;A) = -EI\frac{A\pi^2}{L^2}\sin\frac{\pi x}{L} - M_0.$$

Since the approximating function contains just one fitting coefficient, we need to select only one collocation point in the domain $[0,L]$ and force residual to zero there. We do not know which point will be the best choice, so we arbitrarily select collocation point at $x = L/2$. By equation (12), we have

$$R(L/2) = -EI\frac{A\pi^2}{L^2}\sin\frac{\pi}{2} - M_0 = 0.$$

Solving for the unknown coefficient A, we obtain

$$A = -\frac{M_0 L^2}{EI\pi^2}.$$

Thus, the approximate solution in the interval $[0,L]$ is

$$u(x) = -\frac{M_0 L^2}{EI\pi^2}\sin\frac{\pi x}{L}. \tag{18}$$

Figure 3 shows that the approximate solution $u(x)$ agrees well with the exact solution $y(x)$ over the interval $[0,L]$. Note that if we had selected the collocation point other than at $x = L/2$, a different approximate solution would have been obtained.

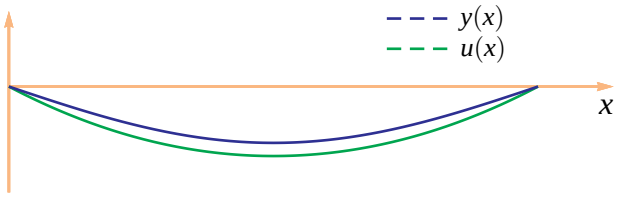

Figure 3: Beam deflection problem – result of point collocation method.

Polynomial approximation to deflection curve. Here we select a second degree polynomial of the form

$$u(x) = a + bx + cx^2.$$

The function which satisfies the prescribed boundary conditions can be obtained by the application of boundary conditions to the chosen approximating function. Thus, we have

$$u(x) = cx(x-L) = c_1 N_1 \tag{19}$$

where $c_1 = c$ and $N_1 = x(x-L)$. The second derivative,

$$\frac{d^2u}{dx^2} = 2c$$

The residual R is then given by

$$R(x;c) = EI \times 2c - M_0.$$

Here $R(x)$ is independent of x, so that the residual can be set to zero at every point in the interval automatically. Therefore,

$$EI \times 2c - M_0 = 0.$$

Solving for the unknown coefficient c, we get

$$c = \frac{M_0}{2EI}.$$

Thus, the approximate solution is

$$u(x) = -\frac{M_0}{2EI}x(L-x). \tag{20}$$

It may be noted that selection of a second degree polynomial yields exact solution since the selected polynomial represents the exact behaviour of the deflection curve.

Example 2

It is interesting to note that one-dimensional steady state heat conduction problem with uniform heat generation is similar to the beam deflection problem discussed above. The governing differential equation for the heat conduction problem is given by

$$k\frac{d^2T}{dx^2} + S = 0, \qquad x \in [0, L] \tag{21}$$

with the boundary conditions

$$T(0) = T(L) = 0$$

where S is the uniform the rate of heat generation per unit volume of the material with thermal constant conductivity k. The exact solution of the problem is

$$T(x) = \frac{S}{2k}x(L-x) \tag{22}$$

Exact solution will be obtained if second degree polynomial is selected as the trial function.

Example 3

We will now take a fluid mechanics problem which is governed by a second-order linear ordinary differential equation similar to that of beam deflection problem and steady state heat conduction problem discussed earlier. Consider the fully developed flow between infinite parallel plates. The plates are separated by a distance h, as shown in figure. The length of the plates in z-direction is assumed to be very large compared to h, with no variation of any fluid property in this direction. With this assumption, we have $\partial/\partial z = 0$. The flow is assumed to be steady, incompressible, and unidirectional with velocity components $v = w = 0$. Since the flow under consideration is unidirectional it satisfies the condition for parallel flows. The continuity and x-momentum equation are given by

$$\frac{\partial u}{\partial x} = 0$$

$$\rho\frac{\partial u}{\partial t} = \rho g_x - \frac{\partial p}{\partial x} + \mu\left(\frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2}\right)$$

$$0 = \rho g_y - \frac{\partial p}{\partial y}$$

$$0 = \rho g_z - \frac{\partial p}{\partial z}$$

For steady flow in the absence of gravitational force, the system reduces to

$$0 = -\frac{\partial p}{\partial x} + \mu \frac{\partial^2 u}{\partial y^2}$$
$$0 = \frac{\partial p}{\partial y}$$
$$0 = \frac{\partial p}{\partial z}$$

From the continuity equation we can infer that the velocity u is not a function of stream-wise direction, x. In other words, the flow is same in any x-location. The phrase *fully developed flow* is often used to describe this situation. Thus, in the fully developed flow, u is function of only y; i.e., $u = u(y)$.

The y and z-momentum equations show that the pressure is independent of y and z coordinates. Thus, pressure could be a function of x alone, i.e.,

$$p = p(x)$$

The x-momentum equation can be written as

$$\frac{d^2u}{dy^2} - \frac{1}{\mu}\frac{dp}{dx} = 0 \qquad x \in [0,h] \tag{23}$$

Since the left-hand side varies only with y and the right-hand side varies only with x, it follows that both sides must be equal to the same constant. Hence, the pressure gradient dp/dx is a constant. This equation can be integrated twice and no-slip boundary conditions can then be applied to obtain the analytical solution

$$u(y) = -\frac{1}{2\mu}\frac{dp}{dx} y(h-y) \tag{24}$$

Figure shows the parabolic velocity profile. Exact solution will be obtained if second degree polynomial is selected as the trial function.

Example 4

Solve the differential equation

$$\frac{d^2y}{dx^2} + y = x, \qquad x \in [0,2]$$

with the boundary conditions

$$y(0) = 0, \qquad y(2) = 5$$

using point collocation method. The exact solution of the problem is

$$y(x) = \frac{3}{\sin 2}\sin x + x$$

over the interval $[0,2]$.

To solve the problem using point collocation method, we use a polynomial trial function $u(x)$ of degree 3 in the form

$$u(x) = 2.5x + c_2 x(x-2) + c_3 x^2(x-2) = 2.5N_1 + c_2 N_2 + c_3 N_3.$$

Here we have three linearly independent trial functions $N_1 = x$, $N_2 = x(x-2)$, and $N_3 = x^2(x-2)$. The boundary conditions are met by the first term, and other terms are so selected that they are equal to zero at the boundaries so that $u(x)$ also meets the boundary conditions.[1]

The residual is obtained after substituting $u(x)$ for $y(x)$ in the differential equation,

$$R(x) = \frac{d^2u}{dx^2} + u - x.$$

From the $u(x)$ defined, we have

$$\frac{d^2u}{dx^2} = 2c_2 + c_3(6x-4).$$

Therefore, the residual becomes

$$R(x) = 2c_2 + c_3(6x-4) + 2.5x + c_2x(x-2) + c_3x^2(x-2) - x.$$

Since the trial function contains two unknown fitting coefficients, we can force the residual to be zero at two distinct points in $[0,2]$. We do not know which two points will be the best choices, so we arbitrarily select collocation points at $x = 0.7$ and $x = 1.3$. (Note that these points are more or less equally spaced in the interval). Setting the residual zero at these points gives a pair of equation for the constants c_2 and c_3:

$$1090c_2 - 437c_3 + 1050 = 0,$$
$$1090c_2 + 2617c_3 + 1950 = 0.$$

or in matrix form

$$\begin{pmatrix} 1.09 & -0.437 \\ 1.09 & 2.617 \end{pmatrix} \begin{bmatrix} c_2 \\ c_3 \end{bmatrix} = - \begin{bmatrix} 1.05 \\ 1.95 \end{bmatrix}.$$

Solving the above set of equations for c_2 and c_3 and substitute in the assumed trial function to obtain

$$\begin{aligned} u(x) &= \left(\frac{5}{2}\right)x - \left(\frac{60000}{55481}\right)x(x-2) - \left(\frac{900}{3054}\right)x^2(x-2) \\ &= -\left(\frac{900}{3054}\right)x^3 - \left(\frac{13895700}{28239829}\right)x^2 + \left(\frac{517405}{110962}\right)x. \end{aligned}$$

Figure 4 shows that the approximate solution $u(x)$ agrees well with the exact solution $y(x)$ over the interval $[0,2]$.

It is instructive to note the following points about point collocation method:

- Point collocation method does not automatically produce symmetric coefficient matrix which is a desirable property when the solution of the equation is sought. Also, symmetry has nothing to do with the type of approximate solution ϕ selected.
- Setting the residual to zero at discrete points does not mean that the errors in those points are actually zero.
- Computational effort required in the point collocation method is minimal.

[1] It is customary to match the boundary conditions with the initial term(s) of $u(x)$ and then make the succeeding terms equal to zero at the boundaries.

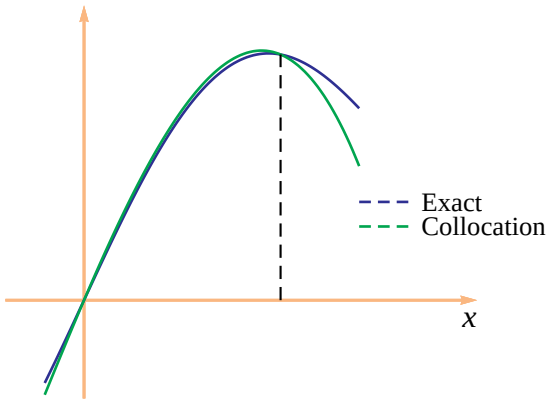

Figure 4: Comparison of point collocation and exact solutions of problem #4.

3 Subdomain Collocation Method

In the subdomain collocation method, we divide the physical domain into a number of non-overlapping subdomains. Number of subdomain n is taken as equal to the number of unknown coefficients in the approximating function. Now, each weight function is selected as unity over a specific subdomain and set equal to zero over other the other parts. That is, for one-dimensional problems,

$$w_i = \begin{cases} 1, & \text{if } x_i \le x \le x_{i+1} \\ 0, & \text{else} \end{cases} \qquad (i = 1, 2, \cdots, n) \tag{25}$$

Thus, equation (10) may be written as

$$\int_a^b w_i R(x)\, dx = \int_{x_i}^{x_{i+1}} R(x)\, dx = 0, \quad (i = 1, 2, \cdots, n). \tag{26}$$

This means that the average of the residual over each of n subdomains is forced to be zero. Or, in other words, differential equation is satisfied on the average in each of the n subdomains. For the subdomain collocation method the linear system of equation (8) takes the form

$$\begin{bmatrix} \int_{x_1}^{x_2} \mathscr{L}(N_1)dx & \int_{x_1}^{x_2} \mathscr{L}(N_2)dx & \cdots & \int_{x_1}^{x_2} \mathscr{L}(N_n)dx \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ \int_{x_n}^{x_{n+1}} \mathscr{L}(N_1)dx & \int_{x_n}^{x_{n+1}} \mathscr{L}(N_2)dx & \cdots & \int_{x_n}^{x_{n+1}} \mathscr{L}(N_n)dx \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} = - \begin{bmatrix} \int_{x_1}^{x_2} f\, dx \\ \int_{x_2}^{x_3} f\, dx \\ \vdots \\ \int_{x_n}^{x_{n+1}} f\, dx \end{bmatrix}. \tag{27}$$

Note: It can be shown that the subdomain collocation method is equivalent to the widely used *finite volume method* in computational fluid dynamics.

Example 5

Now, let us illustrate the application of subdomain collocation method using the beam deflection problem considered earlier. The governing differential equation is given by

$$EI\frac{d^2y}{dx^2} - M_0 = 0$$

with boundary conditions

$$y(0) = 0 \qquad \& \qquad y(L) = 0.$$

Trigonometric approximation to deflection curve. The sinusoidal trial function that satisfies the specified boundary conditions is given by

$$u(x) = A\sin\frac{\pi x}{L} = c_1 N_1$$

where $N_1 = \sin\frac{\pi x}{L}$ and the residual

$$R(x;A) = -EI\frac{A\pi^2}{L^2}\sin\frac{\pi x}{L} - M_0.$$

Since there is just one unknown coefficient in the approximating function, we have only one subdomain which is the domain itself. Thus, equation (26) becomes

$$\int_0^L R(x)dx = \int_0^L \left(-EI\frac{A\pi^2}{L^2}\sin\frac{\pi x}{L} - M_0\right)dx = 0.$$

The integration yields the following equation

$$-\left(\frac{2EI\pi}{L}\right)A - M_0 L = 0$$

which can be solved for A to obtain

$$A = -\frac{M_0 L^2}{2\pi EI}.$$

Thus, the approximate solution is

$$u(x) = -\frac{M_0 L^2}{2\pi EI}\sin\frac{\pi x}{L}.$$

This approximate solution is also found to be in close agreement with the exact solution. However, a comparison of the above results with that of point collocation method shows that the approximate solutions are different.

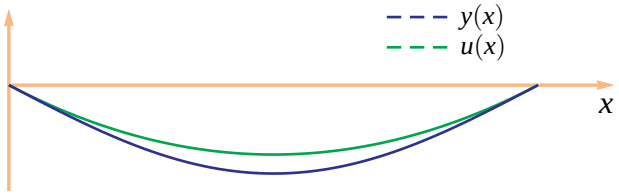

Figure 5: Beam deflection problem – result of subdomain collocation method.

Polynomial approximation to deflection curve. The second degree polynomial trial function that satisfies the specified boundary conditions is given by

$$u(x) = cx(x-L) = c_1 N_1$$

where $N_1 = x(x-L)$ and the residual

$$R(x;c) = EI \times 2c - M_0.$$

Integrating the residual over $[0,L]$

$$\int_0^L R(x)dx = \int_0^L (2EIc - M_0)\,dx = 0$$

which yields

$$(2EIc - M_0)\,L = 0.$$

Solving for c, we have

$$c = \frac{M_0}{2EI}.$$

and thus, the approximate solution is

$$u(x) = -\frac{M_0}{2EI}x(L-x).$$

As in the case of point collocation method, selection of a second degree polynomial as approximating function results in exact solution.

4 Least Square Method

In the least square weighted residual method, the weight functions are chosen to be the derivatives of residual with respect to unknown fitting coefficients c_i's of the approximate solution. So, we set

$$w_i = \frac{\partial R}{\partial c_i}, \qquad (i = 1,2,\cdots,n). \tag{28}$$

Thus, for a one-dimensional problem in the interval $[a,b]$, the weighted residual integral given by equation (8) becomes

$$\int_a^b w_i R(x)\,dx = \int_a^b \frac{\partial R}{\partial c_i}R(x)\,dx = 0, \quad (i = 1,2,\cdots,n). \tag{29}$$

The motivation for this choice of weight function is that we have the following equation

$$\frac{\partial}{\partial c_i}\int_a^b R^2(x)\,dx = 0$$

which implies that the 'average squared residual' in the interval $[a,b]$ is to be minimized with respect to fitting coefficients c_i. Driving the average squared residual to zero will drive the residual R to zero. Since, we have from equation (9), $\partial R/\partial c_i = \mathscr{L}(N_i)$,

$$\frac{\partial R}{\partial c_i} = \mathscr{L}(N_i)$$

for the least square method the linear system of equation (10) takes the form

$$\begin{bmatrix} \int_a^b \mathscr{L}(N_1)\mathscr{L}(N_1)dx & \int_a^b \mathscr{L}(N_1)\mathscr{L}(N_2)dx & \cdots & \int_a^b \mathscr{L}(N_1)\mathscr{L}(N_n)dx \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ \int_a^b \mathscr{L}(N_n)\mathscr{L}(N_1)dx & \int_a^b \mathscr{L}(N_n)\mathscr{L}(N_2)dx & \cdots & \int_a^b \mathscr{L}(N_n)\mathscr{L}(N_n)dx \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} = -\begin{bmatrix} \int_a^b \mathscr{L}(N_1)f\,dx \\ \int_a^b \mathscr{L}(N_2)f\,dx \\ \vdots \\ \int_a^b \mathscr{L}(N_n)f\,dx \end{bmatrix}. \tag{30}$$

Example 6

Trigonometric approximation to deflection curve. We again consider the beam deflection problem. The trigonometric trial function is given by

$$u(x) = A\sin\frac{\pi x}{L} = c_1 N_1$$

where $N_1 = \sin\frac{\pi x}{L}$ and the residual

$$R(x;A) = -EI\frac{A\pi^2}{L^2}\sin\frac{\pi x}{L} - M_0$$

and its derivative,

$$\frac{\partial R}{\partial A} = -EI\frac{\pi^2}{L^2}\sin\frac{\pi x}{L}.$$

The weighted residual equation (25) can now be written as

$$\int_0^L \frac{\partial R}{\partial A}R(x)dx = \int_0^L -EI\frac{\pi^2}{L^2}\sin\frac{\pi x}{L}\left(-EI\frac{A\pi^2}{L^2}\sin\frac{\pi x}{L} - M_0\right)dx = 0.$$

The integration yields the following equation

$$\left(\frac{EI\pi^2}{2L}\right)A + \frac{2M_0L}{\pi} = 0.$$

Solving for A, we have

$$A = -\frac{4M_0L^2}{\pi^3 EI}$$

and thus, the approximate solution is

$$u(x) = -\frac{4M_0L^2}{\pi^3 EI}\sin\frac{\pi x}{L}.$$

Figure 6 shows that the approximate solution $u(x)$ agrees well with the exact solution $y(x)$ over the interval $[0,L]$ and is found to be slightly more accurate than the solution using point collocation method.

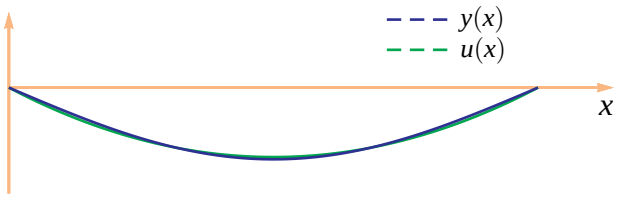

Figure 6: Beam deflection problem – result of least square method.

Polynomial approximation to deflection curve. We use the second degree polynomial trial function

$$u(x) = cx(x-L) = c_1 N_1$$

where $N_1 = x(x-L)$ and the residual

$$R(x;c) = EI \times 2c - M_0$$

and its derivative,

$$\frac{\partial R}{\partial c} = 2EI.$$

The weighted residual equation (25) can now be written as

$$\int_0^L \frac{\partial R}{\partial c} R(x) dx = \int_0^L 2EI\,(2EIc - M_0)\,dx = 0.$$

The integration yields the following equation

$$2EI\,(2EIc - M_0)\,L = 0.$$

Solving for c, we have

$$c = \frac{M_0}{2EI}$$

and thus, the approximate solution is

$$u(x) = -\frac{M_0}{2EI} x(L-x).$$

As in the case of other two methods, selection of a second-order polynomial as approximating function results in exact solution.

Following points about least square method may be noted:

- Least square method always produces symmetric coefficient matrix regardless of the differential operator $\mathscr{L}$ and approximate solution ϕ. Further, this method also produces positive definite matrix since diagonal entries are always positive.
- Least square method is often computationally expensive.

5 Galerkin Method

In Galerkin version of weighted residual method, the weight functions are chosen to be the trial functions themselves. This is the method we usually used for developing finite element equations for field problems. So, in Galerkin method we set

$$w_i = N_i, \qquad (i = 1, 2, \cdots, n). \tag{31}$$

The unknown coefficients in the approximate solution are determined by setting the integral over D of the weighted residual to zero. For one-dimensional problem in the interval $[a,b]$, this procedure will results

$$\int_a^b w_i R(x)\,dx = \int_a^b N_i R(x)\,dx = 0, \quad (i = 1, 2, \cdots, n). \tag{32}$$

For the Galerkin method the linear system of equation (8) takes the form

$$\begin{bmatrix} \int_a^b N_1 \mathscr{L}(N_1) dx & \int_a^b N_1 \mathscr{L}(N_2) dx & \cdots & \int_a^b N_1 \mathscr{L}(N_n) dx \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ \int_a^b N_n \mathscr{L}(N_1) dx & \int_a^b N_n \mathscr{L}(N_2) dx & \cdots & \int_a^b N_n \mathscr{L}(N_n) dx \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} = - \begin{bmatrix} \int_a^b N_1 f\,dx \\ \int_a^b N_2 f\,dx \\ \vdots \\ \int_a^b N_n f\,dx \end{bmatrix}. \tag{33}$$

Following points about Galerkin method may be noted:

- Galerkin method produces symmetric positive definite coefficient matrix if the differential operator is self-adjoint.
- Galerkin method requires less computational effort compared to the least square method.

Example 7

Trigonometric approximation to deflection curve. Yet again we consider the now familiar beam deflection problem. The trigonometric trial function is given by

$$u(x) = A\sin\frac{\pi x}{L} = c_1 N_1$$

where $N_1 = \sin\frac{\pi x}{L}$ and the residual

$$R(x;A) = -EI\frac{A\pi^2}{L^2}\sin\frac{\pi x}{L} - M_0.$$

The unknown coefficients in the approximate solution are determined by setting the integral over $[0,L]$ of the weighted residual to zero. The weighted residual equation give by (32) can now be written as

$$\int_0^L N_1 R(x)dx = \int_0^L \sin\frac{\pi x}{L}\left(-EI\frac{A\pi^2}{L^2}\sin\frac{\pi x}{L} - M_0\right)dx = 0.$$

The integration yields the following equation

$$\left(\frac{EI\pi^2}{2L}\right)A + \frac{2M_0L}{\pi} = 0.$$

Solving for A, we have

$$A = -\frac{4M_0L^2}{\pi^3 EI}$$

and thus, the approximate solution is

$$u(x) = -\frac{4M_0L^2}{\pi^3 EI}\sin\frac{\pi x}{L}.$$

Figure 7 shows that the approximate solution $u(x)$ agrees well with the exact solution $y(x)$ over the interval $[0,L]$ and is found to be slightly more accurate than the solution using point collocation method.

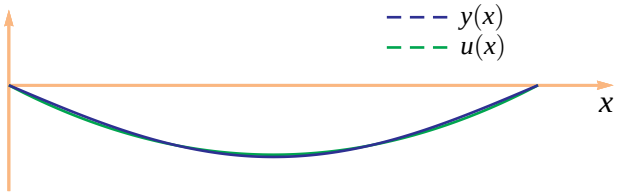

Figure 7: Beam deflection problem – result of Galerkin method.

Polynomial approximation to deflection curve. We use the second degree polynomial trial function

$$u(x) = cx(x-L) = c_1 N_1$$

where $N_1 = x(x-L)$ and the residual

$$R(x;c) = EI \times 2c - M_0.$$

Integrating the weighted residual over $[0,L]$

$$\int_0^L N_1 R(x)dx = \int_0^L x(x-L)\,(2EIc2c - M_0)\,dx = 0$$

to obtain the following equation

$$(2EIc - M_0)\left(\frac{L^3}{3} - \frac{L^3}{2}\right) = 0.$$

Solving for c, we have

$$c = \frac{M_0}{2EI}$$

and thus, the approximate solution is

$$u(x) = -\frac{M_0}{2EI}x(L-x).$$

As in the case of other methods, selection of a second-order polynomial as approximating function results in exact solution.

Example 8

Solve the differential equation

$$\frac{d^2y}{dx^2} + y = x, \qquad x \in [0,2]$$

with the boundary conditions

$$y(0) = 0, \qquad y(2) = 5$$

using Galerkin method.

We use the same trial function $u(x)$ as with the point collocation method:

$$u(x) = 2.5x + c_2x(x-2) + c_3x^2(x-2) = 2.5N_1 + c_2N_2 + c_3N_3$$

so that $N_2 = x(x-2)$ and $N_3 = x^2(x-2)$. The residual of the differential equation is given by

$$R(x) = \frac{d^2u}{dx^2} + u - x.$$

After duly substituting u and u'' in the above residual equation, we get

$$R(x) = 2c_2 + c_3(6x-4) + 2.5x + c_2x(x-2) + c_3x^2(x-2) - x.$$

The unknown coefficients in the approximate solution are determined by using equation (32):

$$\int_0^2 x(x-2)\,R(x)\,dx = 0$$
$$\int_0^2 x^2(x-2)\,R(x)\,dx = 0$$

which gives the two algebraic equations for c_2 and c_3:

$$4c_2 + 4c_3 = -5$$
$$2c_2 + 4c_3 = -3$$

Solving the above set of equations for c_2 and c_3 and substitute in the assumed trial function to obtain

$$\begin{aligned} u(x) &= \left(\frac{5}{2}\right)x - x(x-2) - \left(\frac{1}{4}\right)x^2(x-2) \\ &= -\left(\frac{1}{4}\right)x^3 - \left(\frac{1}{2}\right)x^2 + \left(\frac{9}{2}\right)x. \end{aligned}$$

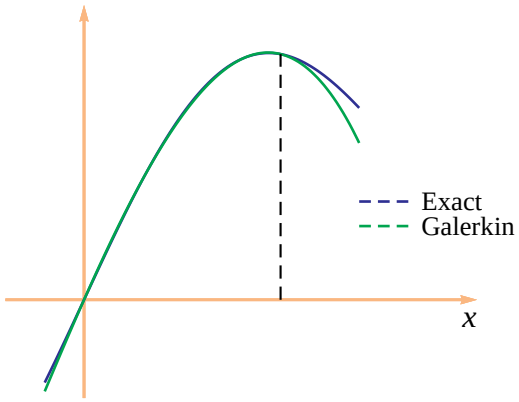

Figure 8: Comparison of Galerkin and exact solutions of problem #8.

Figure 8 shows that the approximate solution $u(x)$ agrees very well with the exact solution $y(x)$ over the interval $[0,2]$.

So we have used several types of weighted residual method for solving boundary value problems. It can be seen that, for the beam deflection problem, the application of all the method yield the exact solution, if second or higher degree polynomial is selected as the approximating function. This is because, the actual behavior of the deflection curve is parabolic, i.e., a second degree polynomial. However, the selection of a sinusoidal function as approximating function yield different solutions for different method except for leat square and Galerkin methods. Now, the question naturally arises is which method gives the most accurate results. Unfortunately, there is no conclusive answer for this. The error depend on the approximating function and the differential equation to be solved. However, for most problems, the Galerkin method gives the best results.

Before we close this discussion, we will develop the Galerkin formulation for the boundary-value problem governed by the generic second-order linear ordinary differential equation. Those differential equations which we have already considered are all could be viewed as special cases of this generic equation.

Example 9

Consider the following linear boundary value problem governed by the following generic second-order linear ordinary differential equation:

$$\frac{d^2y}{dx^2} + Q(x)y = F(x), \qquad x \in [0,1] \tag{34}$$

with the Dirichlet boundary conditions

$$y(0) = 0, \qquad y(1) = Y$$

To use Galerkin method to solve the above boundary value problem, we use a polynomial trial function $u(x)$ of degree 3 in the form

$$u(x) = c_1 N_1(x) + c_2 N_2(x) + c_3 N_3(x) = c_1 x + c_2 x(x-1) + c_3 x^2(x-1)$$

The trial functions N_1, N_2, and N_3 are linearly independent. Applying the boundary conditions yields $c_1 = Y$. Thus, the approximate solution is given by

$$u(x) = Yx + c_2 x(x-1) + c_3 x^2(x-1) = u(x; c_2, c_3) \tag{35}$$

The residual is obtained after substituting $u(x)$ for $y(x)$ in the differential equation (20)

$$R(x) = \frac{d^2u}{dx^2} + Q(x)u - F(x) \tag{36}$$

The second derivative u'' is obtained from equation (21):

$$\frac{d^2u}{dx^2} = 2c_2 + c_3(6x-2)$$

Therefore, the residual becomes

$$R(x) = 2c_2 + c_3(6x-2) + Q\left[Yx + c_2 x(x-1) + c_3 x^2(x-1)\right] - F \tag{37}$$

In Galerkin method, we choose the weighting function as the trial functions, thus:

$$w_2 = N_2 = x(x-1) \qquad \text{and} \qquad w_3 = N_3 = x^2(x-1)$$

The unknown coefficients in the approximate solution are determined by setting the integral of the weighted residual to zero.

$$\int_0^1 x(x-1)\left\{2c_2 + c_3(6x-2) + Q\left[Yx + c_2 x(x-1) + c_3 x^2(x-1)\right] - F\right\} dx = 0 \tag{38a}$$

$$\int_0^1 x^2(x-1)\left\{2c_2 + c_3(6x-2) + Q\left[Yx + c_2 x(x-1) + c_3 x^2(x-1)\right] - F\right\} dx = 0 \tag{38b}$$

Integration can be performed after substituting the functions $Q(x)$ and $F(x)$ to obtain the algebraic equations for unknowns c_2 and c_2. If Q and F are constants, it is easy to carry out the integration. The result is:

$$c_2\left(\frac{1}{3} - \frac{Q}{30}\right) + c_3\left(\frac{1}{6} - \frac{Q}{60}\right) = -\frac{QY}{12} + \frac{F}{6} \tag{39a}$$

$$c_2\left(\frac{1}{6} - \frac{Q}{60}\right) + c_3\left(\frac{2}{15} - \frac{Q}{105}\right) = -\frac{QY}{20} + \frac{F}{12} \tag{39b}$$

Note: It must be emphasized that the Galerkin method is not FEM. In fact, Galerkin method was available much before the concept of FEM is introduced. The essential difference between the Galerkin method and FEM is that unlike in Galerkin method, the approximating function in FEM is not defined over the whole physical domain; it is only defined over the individual elements which constitutes the physical domain. In standard FEM the Galerkin method is often used to derive the element equations.

.