NUCLEAR ENGINEERING, UC BERKELEY

ME 280A Finite Element Analysis HOMEWORK 3: Potential and efficient solution techniques

Xin Wang

October 21, 2014

1 INTRODUCTION

The objective of this project is to solve a 1-D differential equation using linear elements.

$$\begin{aligned} \frac{d}{dx}(A_1(x)\frac{du}{dx}) &= f(x) \\ f(x) &= 256sin(\frac{3\pi kx}{4})cos(16\pi x) \\ A_1(x) &= piecewise \\ L &= 1 \\ u(0) &= 0 \\ A1(L)\frac{du}{dx}(L) &= 1 \end{aligned} \tag{1}$$

A_1 is a piece wise function, defined as bellow:

$$\begin{aligned} For \quad 0.0 \le x < 0.1: \quad A_1 = 2.0 \\ For \quad 0.1 \le x < 0.2: \quad A_1 = 2.5 \\ For \quad 0.2 \le x < 0.3: \quad A_1 = 1.25 \\ For \quad 0.3 \le x < 0.4: \quad A_1 = 0.25 \\ For \quad 0.4 \le x < 0.5: \quad A_1 = 4.0 \\ For \quad 0.5 \le x < 0.6: \quad A_1 = 1.75 \\ For \quad 0.6 \le x < 0.7: \quad A_1 = 0.5 \\ For \quad 0.7 \le x < 0.8: \quad A_1 = 0.75 \\ For \quad 0.8 \le x < 0.9: \quad A_1 = 3.25 \\ For \quad 0.9 \le x \le 1.0: \quad A_1 = 1.0 \end{aligned} \tag{2}$$

2 ANALYTICAL SOLUTION

The general solution for this equation on every segment can be found after two integrations:

$$\begin{aligned} A_1(x)\frac{du}{dx} &= \int 256sin(\frac{3\pi kx}{4})cos(16\pi x)dx \\ &= \int 128[sin(\frac{67\pi x}{4}) - sin(\frac{61\pi x}{4})] \\ &= \frac{512(67cos(\frac{61\pi x}{4}) - 61cos(\frac{67\pi x}{4}))}{4087\pi} + C1 \end{aligned} \tag{3}$$

$$\frac{du}{dx} = \frac{1}{A_1(x)}\left(\frac{512(67cos(\frac{61\pi x}{4}) - 61cos(\frac{67\pi x}{4}))}{4087\pi} + C1\right) \tag{4}$$

Integrating $\frac{du}{dx}$ and applying the boundary condition at x=0, we obtain the analytical solution for the differencial equation:

$$u(x) = \frac{\frac{137216sin(\frac{61\pi x}{4})}{61\pi} - \frac{124928sin(\frac{67\pi x}{4}))}{67\pi}}{4087\pi} + \frac{C1x}{A_1(x)} + C2 \tag{5}$$

At the interface of each segmentsx_N(N=1,...,10), the function u and its derivation is continuous, i.e.,

$$\begin{aligned} u(x_N^+) &= u(x_N^-) \\ A1(x_N^+)\frac{du}{dx}(x_N^+) &= A1(x_N^1)\frac{du}{dx}(x_N^-) \end{aligned} \tag{6}$$

In particular, at the left endpoint, $u(0) = u_0 = 0$ and at the right end point, $A1(L)\frac{du}{dx}(L) = A1(L)(\frac{du}{dx})_N = 1$.

Applying these boundary conditions and continuation conditions at the interfaces, we can solve the equation system and obtain the analytical solution:

$$\begin{aligned} For \quad & 0.0 \le x < 0.1: \quad C_1 = 1.1692 \\ For \quad & 0.1 \le x < 0.2: \quad C_1 = 1.1692 \\ For \quad & 0.2 \le x < 0.3: \quad C_1 = 1.1692 \\ For \quad & 0.3 \le x < 0.4: \quad C_1 = 1.1692 \\ For \quad & 0.4 \le x < 0.5: \quad C_1 = 1.1692 \\ For \quad & 0.5 \le x < 0.6: \quad C_1 = 1.1692 \\ For \quad & 0.6 \le x < 0.7: \quad C_1 = 1.1692 \\ For \quad & 0.7 \le x < 0.8: \quad C_1 = 1.1692 \\ For \quad & 0.8 \le x < 0.9: \quad C_1 = 1.1692 \\ For \quad & 0.9 \le x \le 1.0: \quad C_1 = 1.1692 \end{aligned} \tag{7}$$

$$
\begin{aligned}
For \quad & 0.0 \le x < 0.1: \quad C_2 = 0 \\
For \quad & 0.1 \le x < 0.2: \quad C_2 = 0.0101 \\
For \quad & 0.2 \le x < 0.3: \quad C_2 = -0.0964 \\
For \quad & 0.3 \le x < 0.4: \quad C_2 == 1.4040 \\
For \quad & 0.4 \le x < 0.5: \quad C_2 = 0.2742 \\
For \quad & 0.5 \le x < 0.6: \quad C_2 = 0.1166 \\
For \quad & 0.6 \le x < 0.7: \quad C_2 = -0.8391 \\
For \quad & 0.7 \le x < 0.8: \quad C_2 = -0.2384 \\
For \quad & 0.8 \le x < 0.9: \quad C_2 = 0.7997 \\
For \quad & 0.9 \le x \le 1.0: \quad C_2 = 0.0931
\end{aligned} \tag{8}
$$

3 FINITE ELEMENT METHOD

The first step of FEM is to derive the weak form of the differential equation:

Find u, u|Γu =d, such that $\forall v$, v|Γu =0

$$
\int_\Omega \frac{dv}{dx} A_1(x) \frac{du}{dx} dx = \int_\Omega f v dx + t v \mid_{\Gamma t} \tag{9}
$$

We approximate the real solution u by

$$
u(x) = \sum_{j=1}^{N} a_j \phi_j(x) \tag{10}
$$

and we choose the test function v with the same approximation functions

$$
v(x) = \sum_{i=1}^{N} b_i \phi_i(x) \tag{11}
$$

where N is the number of degree of freedom(number of nodes).

Then the equation becomes:

$$
\int_\Omega \frac{d}{dx} (\sum_{j=1}^{N+1} a_j \phi_j(x)) A_1(x) \frac{d}{dx} (\sum_{i=1}^{N+1} b_i \phi_i(x)) dx = \int_\Omega f (\sum_{i=1}^{N+1} b_i \phi_i(x)) dx + (\sum_{i=1}^{N+1} b_i \phi_i(x) t) \mid_{\Gamma t}, \forall b_i \tag{12}
$$

We can regroup the terms into:

$$\sum_{i=1}^{N+1} b_i \int_\Omega (\sum_{j=1}^{N+1} a_j \frac{d}{dx}\phi_j(x) A_1(x) \frac{d}{dx}\phi_i(x)) dx = \sum_{i=1}^{N+1} b_i \int_\Omega f\phi_i(x) dx + \sum_{i=1}^{N+1} b_i (\phi_i(x)\, t) \mid_{\Gamma t} \quad (13)$$

As the equation should be valid for any b_i, we obtain the matrix system to solve:

$$\begin{aligned} K_{ij} &= \int_\Omega \frac{d}{dx}\phi_j(x) A_1(x) \frac{d}{dx}\phi_i(x) dx \\ R_i &= \int_\Omega f\phi_i(x) dx + \phi_i(x)\, t \mid_{\Gamma t} \\ Ka &= R \end{aligned} \quad (14)$$

In this homework, piece-wise linear, quadratic or cubic basis functions are used. The numerical computation is carried over the corresponding master elements and mapped to the global elements.

Shape functions are defined for elements with different polynomial orders(figure **??**).

Linear shape functions are:

$$\begin{aligned} \hat{\phi}_1 &= \frac{1-\xi}{2} \\ \hat{\phi}_2 &= \frac{1+\xi}{2} \end{aligned} \quad (15)$$

and

$$\begin{aligned} \frac{d\hat{\phi}_1}{d\xi} &= -\frac{1}{2} \\ \frac{d\hat{\phi}_2}{d\xi} &= \frac{1}{2} \end{aligned} \quad (16)$$

Corresponding global coordinate x can be calculate x from ξ:

$$x = \sum \chi_i \hat{\phi}_i \quad (17)$$

where χ_i are the coordinates of the nodes in the global element.

The method to compute the matrices K and R element by element and assembling procedures are the same as in homework 1. The difference for higher order elements is the size of Ke and Re.

Matrix K is assembled from symmetric 2x2 element matrices. The economical data structure used to store only the three elements in Ke.

$$KE_{table} = \begin{matrix} K_{11}^1 & K_{12}^1 & K_{22}^1 \\ K_{11}^2 & K_{12}^2 & K_{22}^2 \\ & \dots & \\ K_{11}^{Ne} & K_{12}^{Ne} & K_{22}^{Ne} \end{matrix} \tag{18}$$

The multiplication operation K * T is ...

3.1 Applying boundary condition to KE table and R

As discussed in the previous homeworks, for Dirichlet BC on u(0):

$$\begin{aligned} K(1,1) = & \quad 1, K(1,2) = 0 \\ R(1) = & \quad u(0) \\ \Longleftrightarrow & \\ R(1) = & \quad u(0) \\ R(2) = & \quad R(2) - KE_{table}(1,2) * u(0) \\ KE_{table}(1,1) = & \quad 1, KE_{table}(1,2) = 0 \end{aligned}$$

For Dirichlet BC on u(L):

$$\begin{aligned} K(N,N) = & \quad 1, K(N,1:N-1) = 0 \\ R(N) = & \quad u(L) \\ \Longleftrightarrow & \\ R(N) = & \quad u(L) \\ R(N-1) = & \quad R(N-1) - KE_{table}(N-1,2) * u(L) \\ KE_{table}(N-1,3) = & \quad 1, KE_{table}(N-1,2) = 0 \end{aligned}$$

After the modification, a = K/R results in for the line corresponds to Dirichlet boundary condition:

a(1) = R(1)/K[1,1] = R(1)

or

a(N) = R(N)/K[N,N] = R(N)

3.1.1 Neumann boundary condition

Neumann boundary condition gives the value of traction $A1(x)\frac{du}{dx}$ on an end point.

For Neumann BC on x=0:

$R(1) = R(1) + A_1 \frac{du}{dx}[x = 0]$

For Neumann BC on x=L:

$R(N) = R(N) + A_1 \frac{du}{dx}[x = L]$

4 PRECONDITIONED CONJUGATE-GRADIENT SOLVER

There are two approaches to solve the equation for [K]a = R, direct and iterative. The number of operation in direct solvers are $O(N^3)$. They are employed when the number of unknowns are not very large, and there are multiple load vectors. Iterative solver are prefered when the systems are very large. Conjugate gradient type iterative techniques need $O(N^2)$ operations.

We define a potential(eq20) whose derivative is the residual of the equation to solve. So the minimizer of the potential Π is the solution to the discrete system.

$$\bigtriangledown \Pi = [K]\{a\} - \{R\} \tag{19}$$

$$\Pi = \frac{1}{2}\{a\}^T[K]\{a\} - \{a\}^T\{R\} \tag{20}$$

$$\begin{aligned} \{r\}^i = -\bigtriangledown \Pi = \{R\} - [K]\{a\}^i \\ \{a\}^{i+1} = \{a\}^i + \lambda^i \{r\}^i \end{aligned} \tag{21}$$

The idea of steepest descent is to find this minimizer of potential by going to the direction of the gradiant of the potential, i.e. $\{a\}^{i+1} = \{a\}^i + \lambda^i \{z\}^i$, where $\{z\}^i = \{r\}^i + \theta^i \{z\}^{i-1}$.

The coefficient θ is chosen so that z is [K] conjugate to $\{z\}^{i-1}$, i.e.,

$$\begin{aligned} \{z\}^{T,i}[K]\{z\}^{i-1} = 0 \\ \theta^i = -\frac{\{r\}^{T,i}[K]\{z\}^{i-1}}{\{z\}^{T,i-1}[K]\{z\}^{i-1}} \end{aligned} \tag{22}$$

Looking for λ to minimize the potential, we force $\frac{\partial \Pi}{\partial \lambda^i} = 0$ and solve for λ:

$$\lambda^i = \frac{\{z\}^{T,i}\{r\}^i}{\{z\}^{T,i}[K]\{z\}^i} \tag{23}$$

4.1 PRECONDITIONING OF [K]

The rate of convergence of the CG method is related to the condition number

$$\|\{a\}-\{a\}^i\|_K \leq (\frac{\sqrt{C([K])-1}}{\sqrt{C([K])+1}})^i \|\{a\}-\{a\}^1\|_K \tag{24}$$

To reduce the condition number, preconditioning of [K] is done by forming the transformation of variables:

$$\begin{aligned} \{a\} &= [T]\{A\} \\ [\bar{K}] &= [T]^T[K][T] \end{aligned} \tag{25}$$

Diagonal preconditioning is an inexpensive widely used, where [T] is defined from the diaganol of [K] as:

$$T_{ij} = \frac{1}{\sqrt{K_{ij}}}, i, j = 1, ...ndof \tag{26}$$

4.2 ITERATION

Starting from an initial guess for $\{a\}^0$, we can calculate the residual defined as in equation 27.

$$\begin{aligned} \{r\}^0 &= -\bigtriangledown \Pi = \{R\} - [K]\{a\}^0 \\ \{z\}^0 &= \{r\}^0 \\ \lambda^0 &= \frac{\{z\}^{T,0}\{r\}^0}{\{z\}^{T,0}[K]\{z\}^0} \\ \{a\}^1 &= \{a\}^0 + \lambda^0\{z\}^0 \end{aligned} \tag{27}$$

For iteration i >0,

$$\begin{aligned} \{r\}^i &= \{R\} - [K]\{a\}^i \\ \theta^i &= -\frac{\{r\}^{T,i}[K]\{z\}^{i-1}}{\{z\}^{T,i-1}[K]\{z\}^{i-1}} \\ \{z\} &= \{r\}^i + \theta^i\{z\}^{i-1} \\ \lambda^i &= \frac{\{z\}^{T,i}\{r\}^i}{\{z\}^{T,i}[K]\{z\}^i} \\ \{a\}^{i+1} &= \{a\}^i + \lambda^i\{z\}^i \end{aligned} \tag{28}$$

and compute the error for each iteration as:

$$e^i = \frac{\|\{a\}^i - \{a\}^i\|_K}{\|\{a\}^i\|_k} = \frac{|\lambda^i| \|\{z\}\|_K}{\|\{a\}^i_K\|} \tag{29}$$

If $e^i < \tau$ stop, otherwise reiterate.

5 Error Calculations

The error is defined as

$$\begin{aligned} e^N &= \frac{\|u-u^N\|_{A_1(\Omega)}}{\|u\|_{A_1(\Omega)}} \\ \|u\|_{A_1(\Omega)} &= \sqrt{\int_\Omega \frac{du}{dx} A_1 \frac{du}{dx} dx} = \sqrt{\int_\Omega A_1 (\frac{du}{dx})^2 dx} \\ \|u-u^N\|_{A_1(\Omega)} &= \sqrt{\int_\Omega A_1 (\frac{d(u-u^N)}{dx})^2 dx} \end{aligned} \tag{30}$$

To compute the error numerically, we calculate the two following quantities element by element, and then assemble them to obtain the overall error:

$$\begin{aligned} \|u\|^2_{A_1(\Omega)} &= \int_\Omega A_1 (\frac{du}{dx})^2 dx \\ \|u-u^N\|^2_{A_1(\Omega)} &= \int_\Omega A_1 (\frac{d(u-u^N)}{dx})^2 dx \\ &= A_1 \int_\Omega (\frac{du}{dx} - \frac{du_N}{dx})^2 dx \\ &= A_1 \sum_{e=1}^{N_e} \int_{x_e}^{x_{e+1}} (\frac{du}{dx} - \frac{du_N}{dx})^2 dx \\ &= A_1 \sum_{e=1}^{N_e} \int_{-1}^{1} (\frac{du}{d\xi} - \frac{du_N}{d\xi})^2 J d\xi \end{aligned} \tag{31}$$

6 Postprocessing

The postprocessing is more complicated for higher order element. In order to plot the result, we calculate the value of the numerical solution for 10 points in each

P	Ne_{opt}	$N_{opt} = Ne_{Opt} * P + 1$
1	1348	1349
2	94	189
3	34	103

Table 1: Minimum number of nodes needed for the error criteria

elements. For any given ξ, the corresponding value for the solution is calculated, as well as the global coordinate x for the given ξ:

$$\begin{aligned} u(x) &= u(x(\xi)) = \sum_{i=1}^{P+1} a_i^e \hat{\phi}_i \\ x(\xi) &= \sum x^e \hat{\phi}(\xi) \end{aligned} \tag{32}$$

For each element, we can also calculate the derivative of u(x):

$$\frac{du}{dx} = \frac{d}{dx} \sum_{i=1}^{P+1} a_i \phi_i = (\frac{d}{d\xi} \sum_{i=1}^{P+1} a_i \hat{\phi}_i) \frac{d\xi}{dx} \tag{33}$$

7 RESULTS

Minimum number of elements and number of nodes to achieve to error criteria for different order of elements are tabulated bellow. N_{Opt} drops dramatically with the order of the elements.

The numerical solution and the analytical solution are compared in figure **??**.

7.1 RELATIONSHIP BETWEEN THE ERROR AND THE ELEMENT SIZE

Error estimate for FEM is:

$$e^N \leq Ch^{\gamma = min(r-1,P)} \tag{34}$$

r is smoothness, and P is the polynomial order.

From the error plotting, we can observe the linear relation between $log(e^N)$ and $log(h)$,

$$\gamma = 0.98, P = 1$$
$$\gamma = 1.98, P = 2$$
$$\gamma = 2.93, P = 3 \tag{35}$$

Overall, γ is similar to P.

8 CONCLUSION

This homework extends the previous one to higher order of element. Having higher order element shape functions can reduce the number of degree of freedom needed for solving the problem by better capturing the fluctuations in the solution.