AMATH 585 Homework 4

Cade Ballew

February 14, 2022

1 Problem 1

1.1 Part a

Considering the problem

$$-\frac{d}{dx}\left((1+x^2)\frac{du}{dx}\right) = f(x), \quad 0 \le x \le 1,$$

$$u(0) = 0, \quad u(1) = 0,$$

we derive the FEM system needed to solve it numerically. Considering $\varphi \in S$ where S is the space of continuous piecewise linear functions that satisfy the above boundary conditions. Letting

$$\mathcal{L} = -\frac{d}{dx}\left((1+x^2)\frac{d}{dx}\right)$$

be our differential operator, we note that a u that satisfies our BVP also satisfies

$$\langle \mathcal{L}u, \varphi \rangle = -\int_0^1 \frac{d}{dx}\left((1+x^2)\frac{du}{dx}\right)\varphi(x)dx = \int_0^1 f(x)\varphi(x)dx = \langle f, \varphi \rangle.$$

Integrating by parts,

$$\int_0^1 \frac{d}{dx}\left((1+x^2)\frac{du}{dx}\right)\varphi(x)dx = \underbrace{\left[(1+x^2)\frac{du}{dx}\varphi(x)\right]_0^1}_{=0} - \int_0^1 (1+x^2)\frac{du}{dx}\varphi'(x)dx,$$

so we can rewrite the weak form as

$$\int_0^1 (1+x^2)u'(x)\varphi'(x)dx = \int_0^1 f(x)\varphi(x)dx.$$

Now, consider the hat functions $\varphi_1, \dots, \varphi_m$ defined in the standard way to be a basis for S and consider $\hat{u} \in S$ so that

$$\hat{u} = \sum_{j=1}^{m} c_j \varphi_j.$$

Then, we want to choose coefficients $c_1, \ldots, c_m$ such that $\langle \mathcal{L}\hat{u}, \varphi_j \rangle = \langle f, \varphi_j \rangle$ for $j = 1, \ldots, m$. We can rewrite this as

$$\sum_{i=1}^{m} c_i \langle \mathcal{L}\varphi_i, \varphi_j \rangle = \langle f, \varphi_j \rangle.$$

This gives way to a matrix problem $Ac = f$ where c is the vector of coefficients that we wish to find and

$$A_{i,j} = \langle \mathcal{L}\varphi_i, \varphi_j \rangle = \int_0^1 (1 + x^2)\varphi_i'(x)\varphi_j'(x)dx,$$

$$f_i = \langle f, \varphi_i \rangle = \int_0^1 (1 + x^2) f(x)\varphi_i(x)dx.$$

Noting that

$$\varphi_j'(x) = \begin{cases} \frac{1}{x_j - x_{j-1}}, & x \in [x_{j-1}, x_j] \\ \frac{-1}{x_{j+1} - x_j}, & x \in [x_j, x_{j=1}] \\ 0, & \text{otherwise,} \end{cases}$$

we first consider that when $j = i$,

$$\begin{aligned} A_{i,j} &= \int_{x_{i-1}}^{x_i} (1 + x^2) \left(\frac{1}{x_i - x_{i-1}} \right)^2 dx + \int_{x_i}^{x_{i+1}} (1 + x^2) \left(\frac{-1}{x_{i+1} - x_i} \right)^2 dx \\ &= \frac{1}{(x_i - x_{i-1})^2} \left[x + x^3/3 \right]_{x_{i-1}}^{x_i} + \frac{1}{(x_{i+1} - x_i)^2} \left[x + x^3/3 \right]_{x_i}^{x_{i+1}} \\ &= \frac{1}{(x_i - x_{i-1})^2}((x_i + x_i^3/3) - (x_{i-1} + x_{i-1}^3/3))\frac{1}{(x_{i+1} - x_i)^2}((x_{i+1} + x_{i+1}^3/3) - (x_i + x_i^3/3)). \end{aligned}$$

When $j = i + 1$,

$$\begin{aligned} A_{i,j} &= \int_{x_i}^{x_{i=1}} (1 + x^2) \frac{-1}{x_{i+1} - x_i} \frac{1}{x_{i+1} - x_i} dx = -\frac{1}{(x_{i+1} - x_i)^2} \left[x + x^3/3 \right]_{x_i}^{x_{i+1}} \\ &= -\frac{1}{(x_{i+1} - x_i)^2}((x_{i+1} + x_{i+1}^3/3) - (x_i + x_i^3/3)), \end{aligned}$$

and when $j = i - 1$,

$$\begin{aligned} A_{i,j} &= \int_{x_{i-1}}^{x_i} (1 + x^2) \frac{1}{x_i - x_{i-1}} \frac{-1}{x_i - x_{i-1}} dx = -\frac{1}{(x_i - x_{i-1})^2} \left[x + x^3/3 \right]_{x_{i-1}}^{x_i} \\ &= -\frac{1}{(x_i - x_{i-1})^2}((x_i + x_i^3/3) - (x_{i-1} + x_{i-1}^3/3)). \end{aligned}$$

When $|i - j| > 1$, we will clearly have that $A_{i,j} = 0$. Since the function $f(x)$ depends on our specific problem, we cannot evaluate the inner products involving it exactly (we could in theory do so if $f(x)$ and $xf(x)$ have closed form antiderivatives, but this is not necessarily the case), so we instead find the entries of f via a midpoint quadrature rule approximation. Namely, if we let $x_{i-1/2} = (x_{i-1} + x_i)/2$ and $x_{i+1/2} = (x_i + x_{i+1})/2$, then

$$
\begin{aligned}
f_i &= \int_{x_{i-1}}^{x_i} f(x)\frac{x - x_{i-1}}{x_i - x_{i-1}}dx + \int_{x_i}^{x_{i+1}} f(x)\frac{x_{i+1} - x}{x_{i+1} - x_i}dx \\
&\approx (x_i - x_{i-1})f(x_{i-1/2})\frac{x_{i-1/2} - x_{i-1}}{x_i - x_{i-1}} + (x_{i+1} - x_i)f(x_{i+1/2})\frac{x_{i+1} - x_{i+1/2}}{x_{i+1} - x_i} \\
&= f(x_{i-1/2})(x_{i-1/2} - x_{i-1}) + f(x_{i+1/2})(x_{i+1} - x_{i+1/2})
\end{aligned}
$$

where we have multiplied the width of each integral by the integrand evaluated at the midpoint. The midpoint rule is known to be second order accurate, so this should not degrade our approximation if our FEM approximation is second order. Using the defintions above for $i, j = 1, \ldots, m$ and letting

$$
c = \begin{pmatrix} c_1 \\ \vdots \\ c_m \end{pmatrix},
$$

we have the system $Ac = f$ needed to solve this problem numerically.

1.2 Part b

Letting $u(x) = x(1 - x)$ so that $f(x) = 2(3x^2 - x + 1)$, we use the following MATLAB code to solve the system of equations in part a for various values of $h = x_i - x_{i-1}$ on a uniform mesh and on a nonuniform mesh where $x_i = (i/(m + 1))^2$, $i = 0, 1, \ldots, m + 1$. when varying the value of m.

```
ufunc = @(x) x.*(1-x);
f = @(x) 2*(3*x.^2-x+1);
P = @(x) x+x.^3/3; %antiderivative of p(x)=1+x^2

fprintf(['  h           error                   error/h^2                 h_max    ' ...
    '       error                 error/h^2\n'])
for m = [9 99 999 9999]
    x = linspace(0,1,m+2)';
    h = x(2)-x(1);
    x2 = (((0:m+1)./(m+1)).^2)'; %nonuniform grid
    hmax = max(x2(2:m+2)-x2(1:m+1));
    x = x(2:m+1); x2 = x2(2:m+1);
    [u,errvec] = solve_fem(x,P,f,ufunc);
    [u2,errvec2] = solve_fem(x2,P,f,ufunc);
    err = norm(errvec,"inf"); err2 = norm(errvec2,"inf");
    fprintf('%.0e %.16d %.16d     %.4e %.16d %.16d\n',h,err,err/h^2,hmax,err2,err2/h^2)
end

function [u,errvec] = solve_fem(x,P,func,utrue)
```

```
m=length(x);
h = [x;1]-[0;x]; %ith entry of h denotes x_i-x_{i-1}

A=zeros(m);
A(1,1:2) = [(P(x(1))-P(0))/h(1)^2+(P(x(2))-P(x(1)))/h(2)^2, -(P(x(2))-P(x(1)))/h(2)^2];
for i = 2:m-1
    A(i,i-1:i+1) = [-(P(x(i))-P(x(i-1)))/h(i)^2, (P(x(i))-P(x(i-1)))/h(i)^2+...
        (P(x(i+1))-P(x(i)))/h(i+1)^2, -(P(x(i+1))-P(x(i)))/h(i+1)^2];
end
A(m,m-1:m) = [-(P(x(m))-P(x(m-1)))/h(m)^2, (P(x(m))-P(x(m-1)))/h(m)^2+...
    (P(1)-P(x(m)))/h(m+1)^2];

xmid = [x(1)/2; (x(1:m-1)+x(2:m))/2; (x(m)+1)/2]; %midpoints of subintervals
f = zeros(m,1);
f(1) = func(xmid(1))*(xmid(1)-0)+func(xmid(2))*(x(2)-xmid(2));
for i = 2:m-1
    f(i) = func(xmid(i))*(xmid(i)-x(i-1))+func(xmid(i+1))*(x(i+1)-xmid(i+1));
end
f(m) = func(xmid(m))*(xmid(m)-x(m-1))+func(xmid(m+1))*(1-xmid(m+1));

u=A\f;
errvec = u-utrue(x);
end
```

1.3 Part c

The results for the uniform case are displayed in the following table.

h	$\|\|u-\hat{u}\|\|_\infty$	$\|\|u-\hat{u}\|\|_\infty/h^2$
1e-01	7.8120188134384039e-04	7.8120188134384025e-02
1e-02	7.8686163202390524e-06	7.8686163202390524e-02
1e-03	7.8686027193963781e-08	7.8686027193963781e-02
1e-04	7.8205966702604712e-10	7.8205966702604712e-02

We can see that our error in the infinity norm appears to be second order as it is nearly constant when divided by h^2.

1.4 Part d

The results for the nonuniform case are displayed in the following table

$\max_i(x_{i+1}-x_i)$	$\|\|u-\hat{u}\|\|_\infty$	$\|\|u-\hat{u}\|\|_\infty/h^2_{\max}$
1.9000e-01	3.2693499644595170e-03	3.2693499644595164e-01
1.9900e-02	3.3145534150957889e-05	3.3145534150957889e-01
1.9990e-03	3.3148672473615193e-07	3.3148672473615193e-01
1.9999e-04	3.3147659339594071e-09	3.3147659339594071e-01

As before, our error in the infinity norm appears to be second order as it is nearly constant when divided by $h^2_{\max}$.

1.5 Part e

If we instead consider the boundary conditions $u(0)=a$, $u(1)=b$, we must modify our set of hat functions as it no longer forms a basis for the space of continuous piecewise linear functions that satisfy the above boundary conditions. To amend this, we introduce functions

$$\varphi_0(x)=\begin{cases}\frac{x_1-x}{x_1}, & x\in[0,x_1]\\ 0, & \text{otherwise}\end{cases}$$

and

$$\varphi_{m+1}(x)=\begin{cases}\frac{x-x_m}{1-x_m}, & x\in[x_m,1]\\ 0, & \text{otherwise}\end{cases}$$

where we also adopt the convention that $x_0=0$, $x_{m+1}=1$. Now, we let

$$\hat{u}=\sum_{j=0}^{m+1}c_j\varphi_j.$$

In order for $\hat{u}$ to satisfy the boundary conditions, we need that $c_0=a$, $c_{m+1}=b$ as the other hat functions are zero on the boundary. Since the values of these coefficients are known, we do not add rows to our matrix A; however we do need to account for its impact on the elements of A. Since φ_0 is only nonzero where φ_1 is nonzero, it only impacts our first equation which becomes

$$\sum_{i=0}^{m+1}c_i\langle\mathcal{L}\varphi_i,\varphi_1\rangle=a\langle\mathcal{L}\varphi_0,\varphi_1\rangle+\sum_{i=1}^{m}c_i\langle\mathcal{L}\varphi_i,\varphi_1\rangle=\langle f,\varphi_1\rangle.$$

Similarly, φ_{m+1} is only nonzero where φ_m is nonzero, so

$$\sum_{i=0}^{m+1}c_i\langle\mathcal{L}\varphi_i,\varphi_m\rangle=b\langle\mathcal{L}\varphi_{m+1},\varphi_m\rangle+\sum_{i=1}^{m}c_i\langle\mathcal{L}\varphi_i,\varphi_m\rangle=\langle f,\varphi_m\rangle.$$

This means that we can keep our matrix A unmodified if we subtract $a\langle\mathcal{L}\varphi_0,\varphi_1\rangle$ and $b\langle\mathcal{L}\varphi_{m+1},\varphi_m\rangle$ from the appropriate elements of our RHS vector. Noting that

$$\varphi_0'(x)=\begin{cases}\frac{-1}{x_1}, & x\in[0,x_1]\\ 0, & \text{otherwise}\end{cases}$$

and

$$\varphi'_{m+1}(x) = \begin{cases} \frac{1}{1-x_m}, & x \in [x_m, 1] \\ 0, & \text{otherwise} \end{cases}$$

we compute

$$\langle \mathcal{L}\varphi_0, \varphi_1 \rangle = \int_0^{x_1} (1+x^2) \frac{-1}{x_1} \frac{1}{x_1} dx = -\frac{1}{x_1^2} \left[x + x^3/3 \right]_0^{x_1} = -\frac{x_1 + x_1^3/3}{x_1^2}$$

and

$$\langle \mathcal{L}\varphi_{m+1}, \varphi_m \rangle = \int_{x_m}^1 (1+x^2) \frac{1}{1-x_m} \frac{-1}{1-x_m} dx = -\frac{1}{(1-x_m)^2} \left[x + x^3/3 \right]_{x_m}^1 = -\frac{4/3 - (x_m + x_m^3/3)}{(1-x_m)^2}.$$

With this, we now let

$$f_1 = f(x_{1/2})(x_{1/2}) + f(x_{3/2})(x_2 - x_{3/2}) + a\frac{x_1 + x_1^3/3}{x_1^2}$$

and

$$f_m = f(x_{m-1/2})(x_{m-1/2} - x_{m-1}) + f(x_{m+1/2})(1 - x_{m+1/2}) + b\frac{4/3 - (x_m + x_m^3/3)}{(1-x_m)^2}$$

and leave the rest of our system unmodified.

2 Problem 2

We use the following code to solve the BVP from problem 1 using chebfun.

```
d = domain(0,1);
x = chebfun('x',d);
utrue = x*(1-x);
f = 2*(3*x^2-x+1);
L = chebop(@(x,u) -diff((1+x^2)*diff(u)),d,0,0);
u = L\f;
err_2 = norm(u-utrue,2)
err_inf = norm(u-utrue,'inf')
```

With this, we compute the L_2 norm of the error as 2.7779×10^{-15} and the infinity norm of the error as 3.8309×10^{-15}.