TMA026/MMA430

Exercises and solutions

Per Ljung
perlj@chalmers.se

April 24, 2020

Exercise session 3: 5.3, 5.6, 5.11.

Exercise 5.3

Use the basis $\{\Phi_i\}_{i=1}^{M_h}$ to show that

$$(P_h v - v, \chi) = 0, \quad \forall \chi \in S_h, \quad v \in L_2$$

can be written in matrix form as $BV = b$, where the matrix B (the so-called mass matrix) is symmetric, positive definite, and sparse if M_h is large.

<u>Solution:</u> Since $P_h v \in S_h = \text{span}(\{\Phi_i\})$, we can write it as the linear combination

$$P_h v = \sum_{j=1}^{M_h} \alpha_j \Phi_j.$$

Then

$$(P_h v, \chi) = (v, \chi) \implies \sum_{i=j}^{M_h} \alpha_j (\Phi_j, \chi) = (v, \chi).$$

This equality should hold for all test functions $\chi \in S_h$, so we test against all basis functions $\chi = \Phi_i$ for $i = 1, 2, ..., M_h$. For each i it must hold that

$$\sum_{j=1}^{M_h} \alpha_j (\Phi_j, \Phi_i) = (v, \Phi_i).$$

This corresponds to the i'th row in a $M_h \times M_h$ matrix system written as

$$\underbrace{\begin{bmatrix} (\Phi_1, \Phi_1) & (\Phi_2, \Phi_1) & \cdot \ \cdot \ \cdot & (\Phi_{M_h}, \Phi_1) \\ \cdot & \cdot & \cdot \ \cdot \ \cdot & \cdot \\ \cdot & \cdot & \cdot \ \cdot \ \cdot & \cdot \\ (\Phi_1, \Phi_{M_h}) & (\Phi_2, \Phi_{M_h}) & \cdot \ \cdot \ \cdot & (\Phi_{M_h}, \Phi_{M_h}) \end{bmatrix}}_{=:B} \underbrace{\begin{bmatrix} \alpha_1 \\ \cdot \\ \cdot \\ \alpha_{M_h} \end{bmatrix}}_{=:V} = \underbrace{\begin{bmatrix} (v, \Phi_1) \\ \cdot \\ \cdot \\ (v, \Phi_{M_h}) \end{bmatrix}}_{=:b},$$

and so we have our matrix system $BV = b$. Now we want to show the stated properties of B. Obviously B is symmetric since the product $(\cdot, \cdot)$ is symmetric. To show the positive definiteness, we recall that a matrix B is positive definite if it holds that

$$v^T B v > 0, \quad \forall v \neq 0.$$

Let $v = \begin{bmatrix} v_1 & \cdot & \cdot & \cdot & v_{M_h} \end{bmatrix}^T$ be a non-zero column vector. Then

$$v^T B v = \begin{bmatrix} v_1 & \cdot & \cdot & \cdot & v_{M_h} \end{bmatrix} \begin{bmatrix} \sum_j v_j (\Phi_j, \Phi_1) \\ \cdot \\ \cdot \\ \sum_j v_j (\Phi_j, \Phi_{M_h}) \end{bmatrix} = \sum_i v_i \sum_j v_j (\Phi_j, \Phi_i)$$

$$= \left(\sum_j v_j \Phi_j, \sum_i v_i \Phi_i \right) = \left\| \sum_j v_j \Phi_j \right\|^2 > 0.$$

It remains to show that B is sparse for large M_h. This follows since the basis functions fulfill the property that

$$\operatorname{supp}(\Phi_i) \cap \operatorname{supp}(\Phi_j) = \emptyset$$

if the nodes i and j are not neighbors. Consequently, for all such i and j we have that

$$(\Phi_j, \Phi_i) = 0.$$

Hence, if M_h is large, there will be a large amount of nodes not neighboring one another, and consequently there will be many zero elements in the matrix, so B is sparse.

Exercise 5.6

Let $a(\cdot,\cdot)$ and $L(\cdot)$ satisfy the assumptions of the Lax-Milgram lemma, i.e.,

$$\begin{aligned}
|a(v,w)| &\le C_1\|v\|_V\|w\|_V, && \forall v,w \in V,\\
a(v,v) &\ge C_2\|v\|_V^2, && \forall v \in V,\\
|L(v)| &\le C_3\|v\|_V, && \forall v \in V.
\end{aligned}$$

Let $u \in V$ be the solution of

$$a(u,v) = L(v), \quad \forall v \in V. \tag{1}$$

Let $\tilde{V} \subset V$ be the finite-dimensional subspace and let $\tilde{u} \in \tilde{V}$ be determined by Galerkin's method:

$$a(\tilde{u},v) = L(v), \quad \forall v \in \tilde{V}. \tag{2}$$

Prove that (note that $a(\cdot,\cdot)$ may be non-symmetric)

$$\|\tilde{u}-u\|_V \le \frac{C_1}{C_2}\min_{\chi\in\tilde{V}}\|\chi-u\|_V. \tag{3}$$

Prove that, if $a(\cdot,\cdot)$ is symmetric and $\|v\|_a = a(v,v)^{1/2}$, then

$$\begin{aligned}
\|\tilde{u}-u\|_a &= \min_{\chi\in\tilde{V}}\|\chi-u\|_a,\\
\|\tilde{u}-u\|_V &\le \sqrt{\frac{C_1}{C_2}}\min_{\chi\in\tilde{V}}\|\chi-u\|_V.
\end{aligned}$$

<u>Solution:</u> At first we recall that (1) and (2) gives the Galerkin orthogonality

$$a(\tilde{u}-u,\tilde{v}) = 0, \quad \forall \tilde{v} \in \tilde{V}.$$

We first show (3) by

$$\begin{aligned}
\|\tilde{u}-u\|_V^2 &\le \frac{1}{C_2}a(\tilde{u}-u,\tilde{u}-u)\\
&= \frac{1}{C_2}\big[a(\tilde{u}-u,\tilde{u}-\tilde{v}) + a(\tilde{u}-u,\tilde{v}-u)\big]\\
&= \frac{1}{C_2}a(\tilde{u}-u,\tilde{v}-u)\\
&\le \frac{C_1}{C_2}\|\tilde{u}-u\|_V\|\tilde{v}-u\|_V.
\end{aligned}$$

Here we first used the coercivity assumption on $a(\cdot,\cdot)$. After this we added and subtracted an arbitrary $\tilde{V}$-function so that we could apply the Galerkin orthogonality on one of the terms. The final step then followed from the boundedness assumption on $a(\cdot,\cdot)$. We can now cancel one factor of $\|\tilde{u}-u\|_V$ on each side to obtain

$$\|\tilde{u}-u\|_V \le \frac{C_1}{C_2}\min_{\chi\in\tilde{V}}\|\chi-u\|_V,$$

since the $\tilde{V}$-function was chosen arbitrarily.

Now we assume that the bilinear form $a(\cdot,\cdot)$ is symmetric. We make the same trick of adding and subtracting a function $\tilde{V}$ to apply Galerkin orthogonality, but continue by applying this once more. We get

$$\begin{aligned}
\|\tilde{u}-u\|_a^2 &= a(\tilde{u}-u,\tilde{u}-u) && \text{(definition)}\\
&= a(\tilde{u}-u,\tilde{u}-\tilde{v}) + a(\tilde{u}-u,\tilde{v}-u) && (\pm\tilde{v} \text{ in right})\\
&= a(\tilde{u}-u,\tilde{v}-u) && \text{(G.O.)}\\
&= a(\tilde{u}-\tilde{v},\tilde{v}-u) + a(\tilde{v}-u,\tilde{v}-u) && (\pm\tilde{v} \text{ in left})\\
&= a(\tilde{u}-\tilde{v},\tilde{v}-u) + \|\tilde{v}-u\|_a^2 && \text{(definition)}\\
&= a(\tilde{u}-\tilde{v},\tilde{v}-\tilde{u}) + a(\tilde{u}-\tilde{v},\tilde{u}-u) + \|\tilde{v}-u\|_a^2 && (\pm\tilde{u} \text{ in right})\\
&= -a(\tilde{u}-\tilde{v},\tilde{u}-\tilde{v}) + a(\tilde{u}-u,\tilde{u}-\tilde{v}) + \|\tilde{v}-u\|_a^2 && (a \text{ symmetric})\\
&= -\|\tilde{u}-\tilde{v}\|_a^2 + 0 + \|\tilde{v}-u\|_a^2 && \text{(G.O.)}\\
&\leq \|\tilde{v}-u\|_a^2. && (\|\cdot\| \geq 0)
\end{aligned}$$

This shows that

$$\|\tilde{u}-u\|_a \leq \min_{\chi\in\tilde{V}} \|\chi-u\|_a, \tag{4}$$

as we chose $\tilde{v}\in\tilde{V}$ arbitrarily. However, since $\tilde{u}\in\tilde{V}$, it follows that

$$\|\tilde{u}-u\|_a = \min_{\chi\in\tilde{V}} \|\chi-u\|_a.$$

To show the last inequality, we begin by noting that for arbitrary $\tilde{v}\in\tilde{V}$

$$\begin{aligned}
\|\tilde{u}-u\|_V^2 &\leq \frac{1}{C_2} a(\tilde{u}-u,\tilde{u}-u) && \text{(Coercivity)}\\
&= \frac{1}{C_2}\|\tilde{u}-u\|_a^2 && \text{(Definition)}\\
&\leq \frac{1}{C_2}\|\tilde{v}-u\|_a^2 && \text{(Inequality (4))}\\
&= \frac{1}{C_2} a(\tilde{v}-u,\tilde{v}-u) && \text{(Definition)}\\
&\leq \frac{C_1}{C_2}\|\tilde{v}-u\|_V\|\tilde{v}-u\|_V. && \text{(Boundedness)}
\end{aligned}$$

Take square root on each side and we find that, since $\tilde{v}\in\tilde{V}$ was arbitrarily chosen,

$$\|\tilde{u}-u\|_V \leq \sqrt{\frac{C_1}{C_2}} \min_{\chi\in\tilde{V}} \|\chi-u\|_V.$$

Exercise 5.11

Formulate a finite element problem corresponding to the Robin problem in Problem 3.6, i.e.

$$\begin{aligned} -\Delta u &= f, \quad \text{in } \Omega, \\ \frac{\partial u}{\partial n} + u &= g, \quad \text{on } \Gamma. \end{aligned}$$

Prove error estimates.

<u>Solution:</u> By standard procedure, we multiply by test function $v \in H^1(\Omega)$ and integrate over the domain to get

$$\int_\Omega -\Delta u v \, \mathrm{d}x = \int_\Omega f v \, \mathrm{d}x.$$

Next, we apply Green's formula on the left hand side and use the Robin boundary condition to find

$$\begin{aligned} -\int_\Omega \Delta u v \, \mathrm{d}x &= \int_\Omega \nabla u \cdot \nabla v \, \mathrm{d}x - \int_\Gamma \frac{\partial u}{\partial n} v \, \mathrm{d}s \\ &= \int_\Omega \nabla u \cdot \nabla v \, \mathrm{d}x + \int_\Gamma uv \, \mathrm{d}s - \int_\Gamma gv \, \mathrm{d}s. \end{aligned}$$

We find the variational formulation by passing everything that does not depend on the solution to the right hand side, i.e., we seek $u \in H^1(\Omega)$ such that

$$\underbrace{\int_\Omega \nabla u \cdot \nabla v \, \mathrm{d}x + \int_\Gamma uv \, \mathrm{d}s}_{=:a(u,v)} = \underbrace{\int_\Omega fv \, \mathrm{d}x + \int_\Gamma gv \, \mathrm{d}s}_{=:L(v)}, \quad \forall v \in H^1(\Omega).$$

Denote the finite element space consisting of continuous piecewise linear functions by S_h. Then the finite element problem is to find $u_h \in S_h$ such that

$$a(u_h, v) = L(v), \quad \forall v \in S_h.$$

Now, since Problem 3.6 consists of showing there is a unique weak solution to this problem, I will assume from here on that $a(\cdot,\cdot)$ and $L(\cdot)$ fulfills the assumption for the Lax-Milgram lemma, i.e.

$$\begin{aligned} |a(v,w)| &\le C_1 \|v\|_{H^1} \|w\|_{H^1}, && \forall v, w \in H^1(\Omega), \\ a(v,v) &\ge C_2 \|v\|_{H^1}^2, && \forall v \in H^1(\Omega), \\ |L(v)| &\le C_3 \|v\|_{H^1}, && \forall v \in H^1(\Omega). \end{aligned}$$

We prove a priori error estimates in $L_2(\Omega)$-norm and $H^1(\Omega)$-norm respectively. The estimates are proven using the Lagrange nodal interpolant I_h and known estimates for it. Moreover, we make the regularity assumption

$$\|u\|_2 \le C \|f\|,$$

so that the error estimates will be stated in terms of given data. For the $H^1(\Omega)$-error we note that

$$
\begin{aligned}
\|u-u_h\|_{H^1}^2 &\le \frac{1}{C_2} a(u-u_h, u-u_h) && \text{(Coercivity)}\\
&= \frac{1}{C_2}\big[a(u-u_h, u-I_h u) + a(u-u_h, I_h u - u_h)\big] && (\pm I_h u \text{ in right})\\
&= \frac{1}{C_2} a(u-u_h, u-I_h u) && \text{(G.O.)}\\
&\le \frac{C_1}{C_2}\|u-u_h\|_{H^1}\|u-I_h u\|_{H^1} && \text{(Boundedness)}\\
&\le \frac{C_1}{C_2}\|u-u_h\|_{H^1} Ch\|u\|_2. && (I_h\text{-estimate})
\end{aligned}
$$

Here we used the fact that $I_h u - u \in V_h$ (where V_h denotes the finite element space) so that the Galerkin orthogonality could be applied. Canceling one factor on each side now yields

$$
\|u-u_h\|_{H^1} \le C' h\|u\|_2 \le Ch\|f\|, \tag{5}
$$

where we applied the regularity assumption to get the error estimate in terms of given data. For the $L_2(\Omega)$-error, we introduce an auxiliary problem. Denote $e = u - u_h$, and define our problem as

$$
\begin{aligned}
-\Delta\phi &= e, \text{ in } \Omega,\\
\frac{\partial\phi}{\partial n} + \phi &= g, \text{ on } \Gamma.
\end{aligned}
$$

Corresponding variational formulation is to find $\phi \in H^1(\Omega)$ such that

$$
a(w,\phi) = (w,e), \quad \forall w \in H^1(\Omega).
$$

Note that our bilinear form is symmetric, so defining it with the test functions in the left lane is not necessary, but we use this convention as it is necessary in the non-symmetric case. We get

$$
\begin{aligned}
\|e\|^2 &= (e,e) && \text{(Definition)}\\
&= a(e,\phi) && \text{(Weak form)}\\
&= a(e,\phi - I_h\phi) && (a(e, I_h\phi) = 0)\\
&\le C_1\|e\|_{H^1}\|\phi - I_h\phi\|_{H^1} && \text{(Boundedness)}\\
&\le C_1 C_2 h\|e\|_{H^1}\|\phi\|_2 && (I_h\text{-estimate})\\
&\le C_1 C_2 C_3 h^2\|u\|_2\|\phi\|_2 && \text{(Estimate (5))}\\
&\le Ch^2\|f\|\|e\| && \text{(Assumption).}
\end{aligned}
$$

Now cancel one $\|e\|$ on each side and we have shown that

$$
\|u-u_h\| \le Ch^2\|f\|.
$$

In conclusion, we have found that the finite element method yields quadratic convergence in L_2-norm and linear convergence in H^1-norm.