Homework 1

Finite Element Methods 1

Due date: 6th November 2023

Submit a PDF/scan of the answers to the following questions before the deadline via the *Study Group Roster* (*Záznamník učitele*) in SIS, or hand-in directly at the practical class on the 6th November 2023.

1. (2 points) Consider the boundary value problem

$$-\sum_{i,j=1}^{n} \frac{\partial}{\partial x_i}\left(a_{ij}\frac{\partial u}{\partial x_j}\right) + cu = f \qquad \text{in } \Omega,$$

$$\sum_{i,j=1}^{n} n_i a_{ij}\frac{\partial u}{\partial x_j} + hu = g \qquad \text{on } \partial\Omega,$$

where $\Omega \subset \mathbb{R}^n$ is a bounded domain with a Lipschitz continuous boundary, $a_{ij} \in L^\infty(\Omega)$, $c \in L^\infty(\Omega)$, $f \in L^2(\Omega)$, $h \in L^\infty(\partial\Omega)$, and $g \in L^2(\partial\Omega)$. We assume the matrix $(a_{ij})_{i,j=1}^{n}$ is uniformly positive definite a.e. in Ω, $c \geq 0$ a.e. in Ω, and $h \geq h_0$ on $\partial\Omega$ where h_0 is a positive constant.

Derive the variational formulation for the above boundary value problem, using the test space $V = H^1(\Omega)$, and prove a unique solution exists.

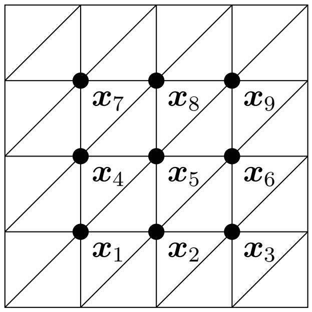
Solution:

Multiplying the PDE by a smooth test function v, integrating over Ω and applying *Gauss' integration theorem* we obtain

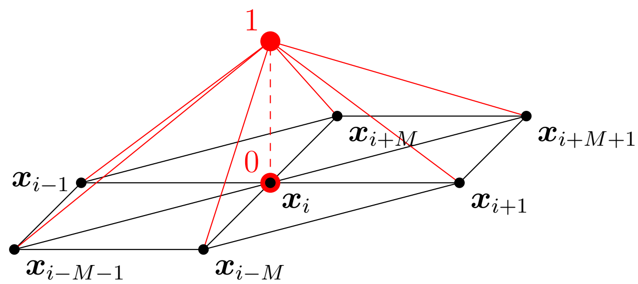
$$\sum_{i,j=1}^{n} \int_\Omega a_{ij}\frac{\partial u}{\partial x_j}\frac{\partial v}{\partial x_i}\,\mathrm{d}\boldsymbol{x} - \sum_{i,j=1}^{n} \int_{\partial\Omega} n_i a_{ij}\frac{\partial u}{\partial x_j} v\,\mathrm{d}s + \int_\Omega cuv\,\mathrm{d}\boldsymbol{x} = \int_\Omega fv\,\mathrm{d}\boldsymbol{x}.$$

By applying the boundary condition we obtain a variational formulation which gives the following weak formulation: find $u \in V$ such that

$$a(u, v) = \langle F, v\rangle \qquad \forall v \in V,$$



(a) Example of 4×4 triangular mesh



(b) Nodal linear basis function φ_i

Figure 1: Question 3

where $V = H^1(\Omega)$ and

$$a(u,v) = \sum_{i,j=1}^{n} \int_\Omega a_{ij} \frac{\partial u}{\partial x_j} \frac{\partial v}{\partial x_i} \,\mathrm{d}\boldsymbol{x} + \int_\Omega cuv \,\mathrm{d}\boldsymbol{x} + \int_{\partial\Omega} huv \,\mathrm{d}s,$$
$$\langle F, v\rangle = \int_\Omega fv \,\mathrm{d}\boldsymbol{x} + \int_{\partial\Omega} gv \,\mathrm{d}s.$$

Then, V is a real Hilbert space and, for any $u, v \in V$ we have that

$$\begin{aligned}
|a(u,v)| &\le \left(\sum_{i,j=1}^{n} \|a_{ij}\|_{0,\infty,\Omega}^2\right)^{1/2} |u|_{1,\Omega}|v|_{1,\Omega} + \|c\|_{0,\infty,\Omega}\|u\|_{0,\Omega}\|v\|_{0,\Omega} \\
&\quad + \|h\|_{0,\infty,\partial\Omega}\|u\|_{0,\partial\Omega}\|v\|_{0,\partial\Omega} \\
&\le C\|u\|_{1,\Omega}\|v\|_{1,\Omega}, && \text{(by \textit{trace theorem})}, \\
a(u,u) &\ge \alpha|u|_{1,\Omega}^2 + h_0\|u\|_{0,\partial\Omega}^2 \ge C\|u\|_{1,\Omega}^2, && \text{(by \textit{Friedrichs})}, \\
|\langle F, v\rangle| &\le \|f\|_{0,\Omega}\|v\|_{0,\Omega} + \|g\|_{0,\partial\Omega}\|v\|_{0,\partial\Omega} \le C\|v\|_{1,\Omega}, && \text{(by \textit{trace theorem})}.
\end{aligned}$$

Therefore, a is a continuous V-elliptic bilinear form and F is a continuous linear functional; hence, a unique solution exists by Lax-Milgram.

2. (2 points) Consider the Poisson equation on the unit square with homogeneous boundary conditions:

$$\begin{aligned} -\Delta u &= f \quad \text{in } \Omega := (0,1)^2 \\ u &= 0 \quad \text{on } \partial\Omega, \end{aligned} \tag{1}$$

where f is a constant.

We define the finite element method for this problem as: Find $u_h \in V_h$ such that

$$a(u_h, v_h) = \langle F, v_h\rangle \qquad \text{for all } v_h \in V_h, \tag{2}$$

where

$$a(u_h, v_h) = \int_\Omega \nabla u_h \cdot \nabla v_h \,\mathrm{d}\boldsymbol{x}, \qquad \langle F, v_h\rangle = \int_\Omega f v_h \,\mathrm{d}\boldsymbol{x},$$

and V_h is finite-dimensional subspace of $H_0^1(\Omega)$. Let $\varphi_1, \ldots, \varphi_N$ be the basis functions of V_h; then, the solution u_h of the finite element discretization (2) can be written in the form $u_h = \sum_{j=1}^{N} u_j \varphi_j$. Hence, the discretization (2) is equivalent to solving the following linear system of N unknown coefficients $u_1, \ldots, u_N$:

$$\sum_{j=1}^{N} a(\varphi_j, \varphi_i) u_j = \langle F, \varphi_i \rangle \qquad \text{for } i = 1, \ldots, N. \tag{3}$$

We denote by $\mathcal{T}_h$ the triangulation of Ω into triangles in the following manner:

1. subdivide the domain into $(M+1) \times (M+1)$ squares of equal size,
2. divide each square into two triangles by splitting from the bottom left to top-right corner of the square;

see Figure 1a for an example when $M = 3$. We define the width and height of each square as $h = {}^1\!/_{(M+1)}$. Let

$$V_h = \{ v_h \in H_0^1(\Omega) : v_h|_T \in P_1(T) \ \forall T \in \mathcal{T}_h \};$$

i.e. the space of continuous piecewise linear functions vanishing on the boundary of Ω. To the interior vertices $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N$ of $\mathcal{T}_h$, where $N = M^2$, (see Figure 1a for one possible numbering of the vertices) we assign a basis function of V_h such that

$$\varphi_i(x_j) = \delta_{ij} \qquad \text{for } i, j = 1, \ldots, N.$$

The support of the basis function φ_i consists of the six triangles sharing the vertex $\boldsymbol{x}_i$, see Figure 1b. This implies that every row of the matrix for the linear system (3) contains at most seven non-zero entries.

Compute the entries for the matrix and right-hand side vector for the linear system (3) and compare these entries to a discretization using the finite difference scheme on a uniform square mesh.

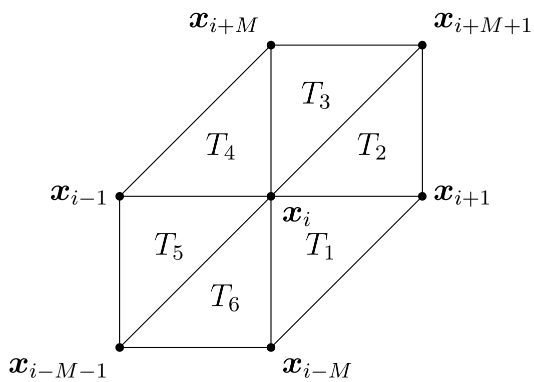
Hint. Computation of these entries is fairly trivial. Consider, for example, the calculation of $a(\varphi_j, \varphi_i)$, where $j = i + 1$. The nodes $\boldsymbol{x}_j$ and $\boldsymbol{x}_{i+1}$ are connected by an edge and only two triangles share this edge; see Figure 1b. We denote these two triangles as T_1 and T_2, and note that $\operatorname{supp} \varphi_j \cap \operatorname{supp} \varphi_i = T_1 \cup T_2$. Note, also, that $\nabla \varphi_j$ and $\nabla \varphi_i$ are constant on each triangle; therefore,

$$a(\varphi_j, \varphi_i) = \int_{T_1 \cup T_2} \nabla \varphi_j \cdot \nabla \varphi_i \, \mathrm{d}\boldsymbol{x} = |T_1| \, (\nabla \varphi_j)|_{T_1} \cdot (\nabla \varphi_i)|_{T_1} + |T_2| \, (\nabla \varphi_j)|_{T_2} \cdot (\nabla \varphi_i)|_{T_2}.$$

The derivatives of φ_j and φ_i with respect to x and y can be computed on the horizontal and vertical edges, respectively, of the triangles T_1 and T_2.

Solution:

For simplicity we shall only deal with all vertices on the interior of the mesh, as when a vertex is on the boundary the value is zero. As can be seen from Figure 1b, or the diagram below, for fixed $i = 1, \ldots, M \times M$ we only need to compute $a(\varphi_j, \varphi_i)$ for $j = i - M - 1, i - M, i - 1, i, i + 1, i + M, i + M + 1$.



From the above diagram we note that $|T_1| = |T_2| = |T_3| = |T_4| = |T_5| = |T_6| = {}^{h^2}/_2$.

$\boldsymbol{j = i + 1}$ From the above diagram we can see the intersections of the supports of φ_i and φ_{i+1} is the triangles T_1 and T_2, and using the fact that the gradients of the basis functions are constant on each triangle:

$$\begin{aligned} a(\varphi_{i+1}, \varphi_i) &= \int_{T_1 \cup T_2} \nabla\varphi_{i+1} \cdot \nabla\varphi_i \,\mathrm{d}\boldsymbol{x} \\ &= |T_1| \, (\nabla\varphi_{i+1})|_{T_1} \cdot (\nabla\varphi_i)|_{T_1} + |T_2| \, (\nabla\varphi_{i+1})|_{T_2} \cdot (\nabla\varphi_i)|_{T_2} \\ &= \frac{h^2}{2} \begin{pmatrix} 1/h \\ 0 \end{pmatrix} \cdot \begin{pmatrix} -1/h \\ -1/h \end{pmatrix} + \frac{h^2}{2} \begin{pmatrix} 1/h \\ -1/h \end{pmatrix} \cdot \begin{pmatrix} -1/h \\ 0 \end{pmatrix} \\ &= -1. \end{aligned}$$

$\boldsymbol{j = i + M + 1}$ Intersection of supports is T_2 and T_3:

$$a(\varphi_{i+M+1}, \varphi_i) = |T_2| \, (\nabla\varphi_{i+M+1})|_{T_2} \cdot (\nabla\varphi_i)|_{T_2} + |T_3| \, (\nabla\varphi_{i+M+1})|_{T_3} \cdot (\nabla\varphi_i)|_{T_3} = 0.$$

$\boldsymbol{j = i + M}$ Intersection of supports is T_3 and T_4:

$$a(\varphi_{i+M}, \varphi_i) = |T_3| \, (\nabla\varphi_{i+M})|_{T_3} \cdot (\nabla\varphi_i)|_{T_3} + |T_4| \, (\nabla\varphi_{i+M})|_{T_4} \cdot (\nabla\varphi_i)|_{T_4} = -1.$$

$\boldsymbol{j = i - 1}$ Intersection of supports is T_4 and T_5:

$$a(\varphi_{i-1}, \varphi_i) = |T_4| \, (\nabla\varphi_{i-1})|_{T_4} \cdot (\nabla\varphi_i)|_{T_4} + |T_5| \, (\nabla\varphi_{i-1})|_{T_5} \cdot (\nabla\varphi_i)|_{T_5} = -1.$$

$\boldsymbol{j = i - M - 1}$ Intersection of supports is T_5 and T_6:

$$a(\varphi_{i-M-1}, \varphi_i) = |T_5| \, (\nabla\varphi_{i-M-1})|_{T_5} \cdot (\nabla\varphi_i)|_{T_5} + |T_6| \, (\nabla\varphi_{i-M-1})|_{T_6} \cdot (\nabla\varphi_i)|_{T_6} = 0.$$

$\boldsymbol{j = i - M}$ Intersection of supports is T_6 and T_1:

$$a(\varphi_{i-M}, \varphi_i) = |T_6| \, (\nabla\varphi_{i-M-1})|_{T_6} \cdot (\nabla\varphi_i)|_{T_6} + |T_1| \, (\nabla\varphi_{i-M-1})|_{T_1} \cdot (\nabla\varphi_i)|_{T_1} = -1.$$

$\boldsymbol{j = i}$ Intersection of supports is T_1, T_2, T_3, T_4, T_5, and T_6:

$$\begin{aligned} a(\varphi_i, \varphi_i) &= \sum_{k=1}^{6} |T_k| \left|(\nabla\varphi_i)|_{T_k}\right|^2 \\ &= \frac{h^2}{2}\left|\begin{pmatrix} -1/h \\ -1/h \end{pmatrix}\right|^2 + \frac{h^2}{2}\left|\begin{pmatrix} -1/h \\ 0 \end{pmatrix}\right|^2 + \frac{h^2}{2}\left|\begin{pmatrix} 0 \\ -1/h \end{pmatrix}\right|^2 \\ &\quad + \frac{h^2}{2}\left|\begin{pmatrix} 1/h \\ -1/h \end{pmatrix}\right|^2 + \frac{h^2}{2}\left|\begin{pmatrix} 1/h \\ 0 \end{pmatrix}\right|^2 + \frac{h^2}{2}\left|\begin{pmatrix} 0 \\ -1/h \end{pmatrix}\right|^2 \\ &= 4. \end{aligned}$$

Furthermore, by formula for the volume of a tetrahedron $\int_{T_k} \varphi_i \, \mathrm{d}\boldsymbol{x} = |T_k|/3$, and hence

$$\langle F, \varphi_i \rangle = \sum_{k=1}^{6} \int_{T_k} f\varphi_i \, \mathrm{d}\boldsymbol{x} = \sum_{k=1}^{6} f\frac{|T_k|}{3} = fh^2.$$

Thus, we have the linear system

$$4u_i - u_{i-1} - u_{i+1} - u_{i-M} - u_{i+M} = fh^2,$$

which is identical to the finite difference scheme.

3. (2 points) Let T be an n-simplex, let $\{a_i\}_{i=1}^n$, $\{a_{iij}\}_{i \neq j}$, $\{a_{ijk}\}_{i<j<k}$ be the points of $L_3(T)$ and let $\{p_i\}_{i=1}^n$, $\{p_{iij}\}_{i \neq j}$, $\{p_{ijk}\}_{i<j<k}$ be the corresponding basis functions of $P_3(T)$. For $i < j < k$ define the linear functionals

$$\Phi_{ijk}(p) = 12\, p(a_{ijk}) + 2 \sum_{\ell \in \{i,j,k\}} p(a_\ell) - 3 \sum_{\substack{\ell,m \in \{i,j,k\} \\ \ell \neq m}} p(a_{llm})$$

and the space

$$P_3'(T) = \{p \in P_3(T) : \Psi_{ijk}(p) = 0, 1 \leq i < j < k \leq n+1\}.$$

Prove that any function from the space $P_3'(T)$ is uniquely determined by its values at the points $\{a_i\}_{i=1}^n \cup \{a_{iij}\}_{i \neq j}$ and derive basis functions such that each basis function equals 1 at one of these points and vanishes at the rest.

Hint. The basis functions $\{p_i\}_{i=1}^n$, $\{p_{iij}\}_{i\neq j}$ can be modified by adding linear combinations of the functions $\{p_{ijk}\}_{i<j<k}$ in such a way that the resulting functions are in $P_3'(T)$. Show that these function form a basis of $P_3'(T)$ and find formulas for these basis functions.

Solution:

Let $\{p_i\}_{i=1}^{n+1}$, $\{p_{iij}\}_{i\neq j}$, and $\{p_{ijk}\}_{i<j<k}$ be the basis of $P_3(T)$. Then, for any $r<s<t$,

$$\Phi_{rst}(p_i) = 2\sum_{\ell\in\{r,s,t\}} p_i(a_\ell) = \begin{cases} 2 & \text{if } i\in\{r,s,t\}, \\ 0 & \text{otherwise,} \end{cases}$$
$$\Phi_{rst}(p_{iij}) = -3\sum_{\substack{\ell,m\in\{r,s,t\}\\ \ell\neq m}} p_{iij}(a_{\ell\ell m}) = \begin{cases} -3 & \text{if } i,j\in\{r,s,t\}, \\ 0 & \text{otherwise,} \end{cases}$$
$$\Phi_{rst}(p_{ijk}) = 12p_{ijk}(a_{rst}) = \begin{cases} 12 & \text{if } i=r, j=s, k=t, \\ 0 & \text{otherwise.} \end{cases}$$

Define the functions

$$p_i' \coloneqq p_i + \sum_{k<\ell<m} \alpha_{k\ell m} p_{k\ell m}, \tag{4}$$
$$p_{iij}' \coloneqq p_{iij} + \sum_{k<l<m} \beta_{k\ell m} p_{k\ell m} \tag{5}$$

as linear combinations of p_i or p_{iij} with $p_{k\ell m}$. Now we investigate if it is possible to select $\alpha_{k\ell m}, \beta_{k\ell m}\in\mathbb{R}$ such that $p_i', p_{iij}'\in P_3(T)$. For any $r<s<t$ we have that

$$\Phi_{rst}(p_i') = \Phi_{rst}(p_i) + \sum_{k<\ell<m} \alpha_{k\ell m}\Phi_{rst}(p_{k\ell m}) = \Phi_{rst}(p_i) + 12\alpha_{rst},$$
$$\Phi_{rst}(p_{iij}') = \Phi_{rst}(p_{iij}) + \sum_{k<\ell<m} \beta_{k\ell m}\Phi_{rst}(p_{k\ell m}) = \Phi_{rst}(p_{iij}) + 12\beta_{rst};$$

hence, as we require that $\Phi_{rst}(p_i')=0$ and $\Phi_{rst}(p_{iij}')=0$, for $1\leq r<s<t\leq n+1$, we can set

$$\alpha_{rst} = -\frac{1}{12}\Phi_{rst}(p_i) = \begin{cases} -\frac{1}{6} & \text{if } i\in\{r,s,t\} \\ 0 & \text{otherwise,} \end{cases} \tag{6}$$
$$\beta_{rst} = -\frac{1}{12}\Phi_{rst}(p_{iij}) = \begin{cases} \frac{1}{4} & \text{if } i,j\in\{r,s,t\} \\ 0 & \text{otherwise,} \end{cases} \tag{7}$$

We note the functions $\{p_i'\}_{i=1}^{n+1}$, $\{p_{iij}'\}_{i\neq j}$, and $\{p_{ijk}\}_{ijk}$ form a basis of P_3. Furthermore, as $P_3'(T)\subset P_3(T)$ and $P_3'(T)\cap\operatorname{span}\{p_{ijk}\}_{i<j<k} = \{0\}$ then $\{p_i'\}_{i=1}^{n+1}$, $\{p_{iij}'\}_{i\neq j}$

form a basis of $P_3'(T)$. For all $i,j,k,\ell = 1,\dots n+1$, $i \neq k$, $j \neq \ell$ we have that

$$p_i'(a_k) = \delta_{ik}, \qquad p_i'(a_{kk\ell}) = 0,$$
$$p_{iij}'(a_k) = 0, \qquad p_{iij}'(a_{kk\ell}) = \begin{cases} 1 & \text{if } i = k, j = \ell \\ 0 & \text{otherwise,} \end{cases}$$

which implies that any $p \in P_3'(T)$ is uniquely determined by its values at the points $\{a_i\}_{i=1}^{n+1} \cup \{a_{iij}\}_{i \neq j}$.

From (4), (6), and the definitions of the basis for $P_3(T)$ we have that

$$p_i' = \frac{1}{2}\lambda_i(3\lambda_i - 1)(3\lambda_i - 2) - \frac{27}{6} \sum_{\substack{k<\ell<m \\ i \in \{k,\ell,m\}}} \lambda_k \lambda_\ell \lambda_m = \frac{1}{2}\lambda_i(3\lambda_i - 1)(3\lambda_i - 2) - \frac{9}{2} \sum_{\substack{j<k \\ j,k \neq i}} \lambda_i \lambda_j \lambda_k,$$

as

$$\begin{aligned} \sum_{\substack{k<\ell<m \\ i \in \{k,\ell,m\}}} \lambda_k \lambda_\ell \lambda_m &= \sum_{\substack{k<\ell<m \\ i=k}} \lambda_k \lambda_\ell \lambda_m + \sum_{\substack{k<\ell<m \\ i=\ell}} \lambda_k \lambda_\ell \lambda_m + \sum_{\substack{k<\ell<m \\ i=m}} \lambda_k \lambda_\ell \lambda_m \\ &= \sum_{i<j<k} \lambda_i \lambda_j \lambda_k + \sum_{j<i<k} \lambda_j \lambda_i \lambda_k + \sum_{j<k<i} \lambda_j \lambda_k \lambda_i \\ &= \sum_{\substack{j<k \\ j,k \neq i}} \lambda_i \lambda_j \lambda_k. \end{aligned}$$

Similarly, from (5) and (7) we have that

$$p_{iij}' = \frac{9}{2}\lambda_i \lambda_j (3\lambda_i - 1) + \frac{27}{4} \sum_{\substack{k<\ell<m \\ i,j \in \{k,\ell,m\}}} \lambda_k \lambda_\ell \lambda_m = \frac{9}{2}\lambda_i \lambda_j (3\lambda_i - 1) + \frac{27}{4} \sum_{k \neq i,j} \lambda_i \lambda_j \lambda_k.$$