Math 574
class #24
11/20/2025

We have been working on the Stokes Problem:

$$-\Delta \vec{u} + \nabla p = \vec{f} \quad \text{in } \Omega$$

$$\nabla \cdot \vec{u} = 0 \quad \text{in } \Omega$$

$$\vec{u} = 0 \quad \text{on } \partial\Omega$$

Last time, using the inf-sup condition, we were able to establish the following fundamental stability:

$$|\vec{u}|_{H^1} + \|p\|_{L^2} \leq \left(1 + \frac{2}{\beta}\right) \|\vec{f}\|_{H^{-1}}.$$

Now, let us investigate finite element approximations.

Mixed FEM for Stokes Problem

Let

$$X_h \subset X := [H_0^1(\Omega)]^d, \quad \dim(X_h) = M$$

$$Q_h \subset Q := L_0^2(\Omega), \quad \dim(Q_h) = N.$$

Typically $M > N$. Suppose that we have the respective bases

$$B_X = \{\vec{\phi}_1, \dots, \vec{\phi}_M\} \subset X_h$$

$$B_Q = \{\psi_1, \dots, \psi_N\} \subset Q_h$$

The <u>Galerkin</u> <u>approximation</u> is given as usual:

Find $\vec{u}_h \in X_h$ and $p_h \in Q_h$ such that

$$
(24.1) \quad
\begin{aligned}
a(\vec{u}_h, \vec{v}_h) - b(\vec{v}_h, p_h) &= (\vec{f}, \vec{v}_h), \quad \forall \vec{v}_h \in X_h, \\
-b(\vec{u}_h, q_h) &= 0, \quad \forall q_h \in Q_h.
\end{aligned}
$$

where

$$a(\vec{u}, \vec{v}) = (\nabla \vec{u}, \nabla \vec{v})_{L^2}, \quad \forall \vec{u}, \vec{v} \in X$$

$$b(\vec{v}, q) = (q, \nabla \cdot \vec{v})_{L^2}, \quad \forall \vec{v} \in X, q \in Q.$$

Now, set

$$\vec{u}_h = \sum_{j=1}^{M} u_j \vec{\phi}_j,$$

$$p_h = \sum_{j=1}^{N} p_j \psi_j.$$

Define

$$A = [a_{ij}] \in \mathbb{R}^{M \times M}, \quad a_{ij} = a(\vec{\phi}_j, \vec{\phi}_i),$$

$$B = [b_{ij}] \in \mathbb{R}^{M \times N}, \quad b_{ij} = b(\vec{\phi}_i, \psi_j).$$

Define

$$\vec{u} = [u_i], \quad i = 1, \cdots, M,$$

$$\vec{p} = [p_i], \quad i = 1, \cdots, N.$$

Finally, define

$$\vec{f} = [f_i] \in \mathbb{R}^M, \qquad f_i := (\vec{f}, \vec{\phi}_i), \quad i = 1, \cdots, M.$$

Then,

$$\begin{aligned} A\vec{u} - B\vec{p} &= \vec{f} \\ -B^T\vec{u} &= \vec{0} \end{aligned} \tag{24.2}$$

Typically, as we have indicated, $M > N$. We will show this shortly. Therefore, in block form,

$$\underbrace{\overbrace{\begin{bmatrix} A & -B \\ -B^T & 0 \end{bmatrix}}^{M+N}}_{M+N} \begin{bmatrix} | \\ \vec{u} \\ | \\ \hline \vec{p} \end{bmatrix} = \begin{bmatrix} | \\ \vec{f} \\ | \\ \hline \vec{0} \end{bmatrix} \in \mathbb{R}^{M+N}$$

The coefficient matrix

$$C = \begin{bmatrix} A & -B \\ -B^T & 0 \end{bmatrix} \in \mathbb{R}^{(M+N)\times(M+N)}$$

is symmetric, but not SPD. Why?

We will show, using the inf-sup condition, that the coefficient matrix C is invertible under certain structural conditions.

Definition (24.1): Let X_h and Q_h be finite dimensional subspaces of

$$X := [H_0^1(\Omega)]^d \quad \text{and} \quad Q := L_0^2(\Omega),$$

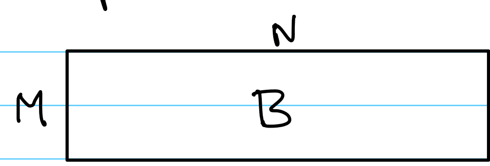
respectively. We say that (X_h, Q_h) form an *inf-sup stable pair* iff they satisfy the following condition

(24.3) $$\sup_{\substack{\vec{v}_h \in X_h \\ \vec{v}_h \neq 0}} \frac{(q, \nabla \cdot \vec{v})}{|\vec{v}|_{H^1}} \geq \beta \|q\|_{L^2}, \ \forall\, q \in Q_h.$$

Lemma (24.2): Suppose that X_h and Q_h form and inf-sup stable pair. Then

$$\dim(X_h) =: M \geq N := \dim(Q_h).$$

Proof: Consider the matrix $B \in \mathbb{R}^{M \times N}$ defined above. To get a contradiction, suppose that $M < N$. Thus, B has the shape



This implies that the columns of B are linearly dependent. There exists an $\vec{\alpha} \in \mathbb{R}^N$, $\vec{\alpha} \neq \vec{0}$, such that $B\vec{\alpha} = \vec{0}$, ie, $\vec{\alpha} \in \mathrm{Ker}(B)$.

$\forall i$, $$0 = \sum_{j=1}^{N} b_{i,j}\, \alpha_j = \sum_{j=1}^{N} b(\vec{\phi}_i, \psi_j)\, \alpha_j$$

$$= b\Big(\vec{\phi}_i, \sum_{j=1}^{N} \alpha_j \psi_j\Big).$$

Define $q_* = \sum_{j=1}^{N} \alpha_j \psi_j \in Q_h$. Then

$$b(\vec{\phi}_i, q_*) = 0, \quad i = 1, \cdots, M.$$

This implies that

$$b(\vec{v}, q_*) = 0, \quad \forall \vec{v} \in X_h,$$

which implies that

$$0 = \sup_{\vec{v} \in X_h \setminus \{\vec{0}\}} \frac{b(\vec{v}, q_*)}{|\vec{v}|_{H^1}} \overset{(24.3)}{\geq} \beta \|q_*\|_{L^2}$$

This implies $q_* = 0$, which is a contradiction.
Thus, $M \geq N$. ///

Theorem (24.3): Suppose that X_h and Q_h form an inf-sup stable pair. Then there is a unique solution $(\vec{u}_h, p_h) \in X_h \times Q_h$ to (24.1), the Galerkin approximation of the Stokes Problem.

Proof: For each fixed $q \in Q_h$, define $l : X_h \to \mathbb{R}$ via

$$l(\vec{v}) := b(\vec{v}, q) = (q, \nabla \cdot \vec{v}).$$

l is a bounded linear functional. The norm of l can be estimated via

$$\frac{|l(\vec{v})|}{\|\nabla \vec{v}\|_{L^2}} \overset{C.S.}{\leq} \frac{\|q\|_{L^2} \|\nabla \cdot \vec{v}\|_{L^2}}{\|\nabla \vec{v}\|_{L^2}} \leq \frac{d \|q\|_{L^2} \cancel{\|\nabla \vec{v}\|_{L^2}}}{\cancel{\|\nabla \vec{v}\|_{L^2}}} = d \|q\|_{L^2}.$$

By the RRT, $\exists! \vec{z}_h \in X_h$ such that

$$a(\vec{z}_h, \vec{v}_h) = \ell(\vec{v}_h) = (q, \nabla \cdot \vec{v}_h),$$

for all $\vec{v}_h \in X_h$. Remember q is fixed. This defines a linear† mapping,

† linearity is left as an exercise

$$T : Q_h \longrightarrow X_h \qquad (\text{Range}(T) \subseteq X_h)$$

such that

(24.4) $$T(q) = \vec{z}_h$$

To see that T is 1-1, suppose that

$$T(q_1) = T(q_2), \qquad q_1, q_2 \in Q_h.$$

This implies that

$$T(q_1 - q_2) = \vec{0} \in X_h$$

$$0 = \sup_{\substack{\vec{v} \in X_h \\ \vec{v} \neq \vec{0}}} \frac{(\nabla T(q_1 - q_2), \nabla \vec{v})}{\|\nabla \vec{v}\|_{L^2}}$$

$$\overset{(24.4)}{=} \sup_{\substack{\vec{v} \in X_h \\ \vec{v} \neq \vec{0}}} \frac{(q_1 - q_2, \nabla \cdot \vec{v})}{\|\nabla \vec{v}\|_{L^2}}$$

$$\overset{(24.3)}{\geq} \beta \|q_1 - q_2\|_{L^2}.$$

Hence $q_1 = q_2$ in Q_h. T is 1-1.

Now define

$$Z_h := T(Q_h) \qquad (Z_h \text{ is range of } T)$$

$$\dim(Z_h) = \dim(Q_h) = N.$$

$$\underset{\dim M}{X_h} = \underset{\dim N}{Z_h} \oplus \underset{\dim M-N}{Z_h^{\perp}}$$

We see that Z_h is a vector subspace of X_h, which has dimension M.

Define
$$Z_h^{\perp} := \{ \vec{z} \in X_h \mid a(\vec{z}, \vec{v}) = 0, \ \forall \vec{v} \in Z_h \}.$$

$Z_h^{\perp}$ is the orthogonal complement of Z_h with respect to $a(\cdot,\cdot)$.

<u>Construction of $\vec{u}_h$:</u> Define $\vec{u}_h \in Z_h^{\perp}$ to be the unique solution to

$$a(\vec{u}_h, \vec{v}_h) = (\vec{f}, \vec{v}_h), \quad \forall \vec{v}_h \in Z_h^{\perp}, \tag{24.5}$$

which is guaranteed by RRT.

<u>Construction of p_h:</u> Define $p_h \in Q_h$ to be the unique solution to

$$(p_h, \nabla \cdot \vec{w}_h) = -(\vec{f}, \vec{w}_h), \quad \forall \vec{w}_h \in Z_h. \tag{24.6}$$

How do we know a unique solution exists?

Suppose that $\vec{z}_h \in Z_h$ is the unique solution to

$$a(\vec{z}_h, \vec{w}_h) \overset{\text{RRT}}{=} -(\vec{f}, \vec{w}_h) \quad \forall \vec{w}_h \in Z_h. \tag{24.7}$$

Now, let $p_h \in Q_h$ be the unique solution of

$$T(p_h) = \vec{z}_h \in Z_h \qquad \left(T : Q_h \xrightarrow[\text{onto}]{1-1} Z_h\right)$$

In other words,

$$(p_h, \nabla \cdot \vec{w}_h) \overset{T}{=} a(\vec{z}_h, \vec{w}_h) \overset{(24.7)}{=} -(\vec{f}, \vec{w}_h)$$

Thus, there is a unique solution to (24.6).

Solution to Galerkin Approx: Let

$$\vec{v}_h = \vec{w}_h^{(1)} + \vec{w}_h^{(2)} \in X_h = Z_h \oplus Z_h^{\perp}$$

where

$$\vec{w}_h^{(1)} \in Z_h, \quad \vec{w}_h^{(2)} \in Z_h^{\perp},$$

be arbitrary. Then,

$$\begin{aligned}
a(\vec{u}_h, \vec{v}_h) &= a(\vec{u}_h, \vec{w}_h^{(1)} + \vec{w}_h^{(2)}) \\
&\overset{\vec{u}_h \in Z_h^{\perp}}{=} a(\vec{u}_h, \vec{w}_h^{(2)}) \\
&\overset{(24.5)}{=} (\vec{f}, \vec{w}_h^{(2)}) \\
&= (\vec{f}, \vec{v}_h) - (\vec{f}, \vec{w}_h^{(1)}) \\
&\overset{(24.6)}{=} (\vec{f}, \vec{v}_h) + (p_h, \nabla \cdot \vec{w}_h^{(1)}) \\
&\overset{T \text{ defn}}{=} (\vec{f}, \vec{v}_h) + (p_h, \nabla \cdot \vec{w}_h^{(1)}) \\
&\qquad + (p_h, \nabla \cdot \vec{w}_h^{(2)}) \\
&= (f, \vec{v}_h) + (p_h, \nabla \cdot \vec{v}_h)
\end{aligned}$$

for all $\vec{v}_h \in X_h$. Furthermore,

$$(q_h, \nabla \cdot \vec{u}_h) = 0 \qquad \forall\, q_h \in Q_h.$$

This also follows from the definition of T, since $\vec{u}_h \in Z_h^{\perp}$.

In fact, we have the following: for any $q_h \in Q_h$, and any $\vec{v}_h \in Z_h^{\perp}$,

$$(q_h, \nabla \cdot \vec{v}_h) = 0.$$

This is a simple exercise. ///