Section 04. FEM and Algebraic Representations

Weak Form of Poisson's Equation

The linear operator $\mathcal{A} : \mathscr{V} \mapsto \mathscr{V}'$ (for example, $-\Delta$) is defined by

$$(\mathcal{A}u, v) := a[u, v] = \int_\Omega \nabla u \cdot \nabla v \, dx, \quad \forall v \in \mathscr{V}$$

and $f \in \mathscr{V}'$ is a function or distribution. Suppose that $\mathcal{A}$ is bounded, i.e.,

$$a[u, v] \le C_a \|u\|_{\mathscr{V}} \|v\|_{\mathscr{V}}, \quad \forall\, u, v \in \mathscr{V}$$

and coercive, i.e.,

$$a[v, v] \ge \alpha \|v\|_{\mathscr{V}}^2, \quad \forall v \in \mathscr{V}.$$

We would like to find $u \in \mathscr{V}$ such that $\mathcal{A}u = f$ or in the weak form

$$a[u, v] = \langle f, v \rangle\,, \quad \forall\, v \in \mathscr{V} \quad \Longrightarrow \quad \text{Well-posed} \tag{22}$$

Hence, $\kappa(\mathcal{A}) = \|\mathcal{A}\|_{\mathscr{L}(\mathscr{V};\mathscr{V}')} \|\mathcal{A}^{-1}\|_{\mathscr{L}(\mathscr{V}';\mathscr{V})} \le C_a/\alpha \quad \Longrightarrow$ Well-conditioned

> However, the problem here lies in that we are working on two different spaces $\mathscr{V}$ and $\mathscr{V}'$.
> Question: If we consider $-\Delta : L^2(\Omega) \mapsto L^2(\Omega)$ instead, do we lost boundedness?

Linear Lagrange Finite Element Method

The weak formulation of the model equation can be written as: Find $u \in H_0^1(\Omega)$, such that

$$\int_\Omega \nabla u \cdot \nabla v dx = \int_\Omega f v \, dx, \quad \forall v \in H_0^1(\Omega).$$

Let $\mathcal{P}_k(\tau)$ be the space of all polynomials of degree less than or equal to k on τ. Let

$$V = V_h := \big\{ v \in C(\overline{\Omega}) \, : \, v \in \mathcal{P}_1(x_{i-1}, x_i), \; v(0) = v(1) = 0 \big\}.$$

Now we can write the discrete variational problem as: Find $u_h \in V_h$, such that

$$a[u_h, v_h] = (f, v_h), \quad \forall v_h \in V_h.$$

Furthermore, we use nodal basis functions $\phi_i \in V_h$, i.e. $\phi_i(x_j) = \delta_{i,j}$. In this way, we can express a given function $u_h \in V_h$ as $u_h(x) = \sum_{j=1}^N u_j \phi_j(x)$. Hence we arrive at the following equation:

$$\sum_{j=1}^N a[\phi_j, \phi_i] u_j = (f, \phi_i) \quad \text{or} \quad \sum_{j=1}^N A_{i,j} u_j = f_i, \qquad i = 1, \ldots, N.$$

This is a system of algebraic linear equations

$$A\vec{u} = \vec{f}, \tag{23}$$

with $(A)_{i,j} = a_{i,j} := a[\phi_i, \phi_j]$, $\vec{u} := (u_i)_{i=1}^N$, and $\vec{f} = (f_i)_{i=1}^N := \big(\langle f, \phi_i \rangle \big)_{i=1}^N$.

Galerkin Approximation

Approximate space: Replaces the underlying function space by appropriate finite dimensional subspaces. We choose a finite dimensional space $V = V_N$ with finite dimension $\dim(V_N) = N$ to approximate $\mathscr{V}$.

Galerkin method: Then we arrive at the Galerkin discretization:

$$\text{Find } u_N \in V: \quad a[u_N, v_N] = \langle f, v_N \rangle, \quad \forall\, v_N \in V. \tag{24}$$

Equation (24) yields the so-called Galerkin discretization. If the bilinear form $a[\cdot,\cdot]$ is symmetric and coercive, it is called the Ritz–Galerkin discretization.

Conforming discretizations: If the bilinear form $a[\cdot,\cdot]$ is coercive, then we have

$$a[v_N, v_N] \geq \alpha_N \|v_N\|_{\mathscr{V}}^2, \quad \forall\, v_N \in V \subset \mathscr{V}.$$

Since coercivity is inherited from $\mathscr{V}$ to its subspace V, we can see that the constant α_N is bounded from below, i.e., $\alpha_N \geq \alpha$, $\forall N$. The bilinear form $a[\cdot,\cdot]$ is well-defined on $V \times V$.

Finite Elements

Definition (Finite element)

A triple $(K, \mathcal{P}, \mathcal{N})$ is called a finite element if and only if

(i) Element domain: $K \subseteq \mathbb{R}^d$ be a bounded closed set with nonempty interior and piecewise smooth boundary;

(ii) Shape functions: $\mathcal{P}$ be a finite-dimensional space of functions on K;

(iii) Set of nodal variables: $\mathcal{N} = \{\mathcal{N}_1, \ldots, \mathcal{N}_k\}$ be a basis of $\mathcal{P}'$.

Problem (Finite element discretization)

Let $V_h \subset \mathscr{V}$ be the space of continuous piecewise polynomials over a quasi-uniform conforming mesh $\mathcal{M}_h$, i.e., $V_h := \left\{v \in C(\overline{\Omega}) \,:\, v|_\tau \in \mathcal{P}_\tau,\ \text{for all } \tau \in \mathcal{M}_h\right\} \bigcap \mathscr{V}$. Find $u_h \in V_h$ such that

$$a[u_h, v_h] = \langle f, v_h \rangle, \quad \forall\, v_h \in V_h, \tag{25}$$

or, equivalently,

$$\mathcal{A}_h u_h = f_h. \tag{26}$$

Properties of Finite Element Space

Proposition (Interpolation error)

Let $\mathcal{M}_h$ be a uniform mesh and V_h be a C^α $(\alpha \geq 0)$ finite element space on $\mathcal{M}_h$. The interpolant $\mathcal{J}_h : W_p^m(\Omega) \mapsto V_h$ satisfies

$$\|v - \mathcal{J}_h v\|_{W_p^k(\Omega)} \lesssim h^{m-k} \|v\|_{W_p^m(\Omega)}, \quad \forall\, v \in W_p^m(\Omega),\ 0 \leq k \leq \min\{m, \alpha+1\}.$$

Proposition (Inverse estimate)

Let $\mathcal{M}_h$ be a uniform mesh and $\mathcal{P} \subseteq W_p^k(K) \bigcap W_q^m(K)$ and $0 \leq m \leq k$. If V_h is a finite element space for $(K, \mathcal{P}, \mathcal{N})$ on $\mathcal{M}_h$, then we have

$$\Big(\sum_{\tau \in \mathcal{M}_h} \|v\|_{W_p^k(\tau)}^p\Big)^{\frac{1}{p}} \lesssim h^{m-k+\min\{0, \frac{d}{p}-\frac{d}{q}\}} \Big(\sum_{\tau \in \mathcal{M}_h} \|v\|_{W_q^m(\tau)}^q\Big)^{\frac{1}{q}}, \quad \forall\, v \in V_h.$$

Some Important Inequalities

NCMIS

Proposition (Some useful inverse estimates)

For any $v \in V_h$, we have

$$\begin{cases} \|v\|_{L^\infty(\Omega)} \lesssim h^{-\frac{d}{p}} \|v\|_{L^p(\Omega)}, & p \in [1, \infty); \\ \|v\|_{H^s(\Omega)} \lesssim h^{-s} \|v\|_{L^2(\Omega)}, & s \in [0, 1]; \\ \|v\|_{H^{1+\alpha}(\Omega)} \lesssim h^{-\alpha} \|v\|_{H^1(\Omega)}, & \alpha \in (0, \frac{1}{2}). \end{cases}$$

Proposition (Discrete Sobolev inequality)

The following inequality holds

$$\|v\|_{L^\infty(\Omega)} \lesssim C_d(h) \|v\|_{H^1(\Omega)}, \quad \forall\, v \in V_h,$$

where $C_1(h) \equiv 1$, $C_2(h) = |\log h|^{1/2}$, and $C_3(h) = h^{-\frac{1}{2}}$.

Simultaneous Estimate

Proposition (Weighted estimate for L^2 projection)

Define $\mathcal{Q}_h : L^2(\Omega) \mapsto V_h$ by, for any $v \in L^2(\Omega)$, it holds that

$$(\mathcal{Q}_h v, w) = (v, w), \quad \forall\, w \in V_h.$$

Then we have the following weighted L^2-estimate

$$\|v - \mathcal{Q}_h v\|_0 + h\|\mathcal{Q}_h v\|_1 \lesssim h\|v\|_1, \quad \forall\, v \in H_0^1(\Omega).$$

Remark (Simultaneous estimate)

From the above weighted L^2-estimate, we can easily show the so-called simultaneous estimate

$$\inf_{w \in V_h} \Big(\|v - w\|_0 + h\|v - w\|_1 \Big) \lesssim h\|v\|_1, \quad \forall\, v \in H_0^1(\Omega).$$

Condition of FE Matrices

NCMIS

Remark (Spectral radius and condition number of $\mathcal{A}_h$)

Suppose that we have a uniform partition with meshsize h. It is clear, from the Poincaré inequality and the inverse inequality, that

$$\|v\|_0^2 \lesssim \|\nabla v\|_0^2 = (\mathcal{A}_h v, v) \leq \|v\|_1^2 \lesssim h^{-2}\|v\|_0^2, \quad \forall v \in V_h.$$

In fact, we have $\rho(\mathcal{A}_h) \cong h^{-2}$ and $\kappa(\mathcal{A}_h) \cong h^{-2}$.

We lost boundedness! The condition number of the discrete operator is not uniformly bounded.

- The key point is the norm used in the above estimates.
- This means the discrete finite element problems are difficult to solve if the mesh size is small.
- Other factors, like diffusion coefficients, will also contribute to ill-condition.
- Reminder: We want to construct a SPD matrix $\mathcal{B}_h$ such that

$$\mu_0\big(\mathcal{A}_h v, v\big) \leq \big(\mathcal{B}_h^{-1} v, v\big) \leq \mu_1\big(\mathcal{A}_h v, v\big), \quad \forall v \in V.$$

Vector Representations

Assume that $\{\phi_i\}_{i=1,\dots,N}$ is a basis of V. Any function $v \in V$ can be represented as

$$v = \sum_{i=1}^{N} \underline{v}_i \phi_i$$

and the vector representation (coefficient vector) of v is defined as

$$\text{primal representation} \quad \underline{v} := \begin{pmatrix} \underline{v}_1 \\ \vdots \\ \underline{v}_N \end{pmatrix} \in \mathbb{R}^N. \tag{27}$$

There is another natural and easier-to-compute vector representation

$$\text{dual representation} \quad \vec{v} := \begin{pmatrix} (v, \phi_1) \\ \vdots \\ (v, \phi_N) \end{pmatrix} \in \mathbb{R}^N \quad \text{and} \quad \vec{v} = M\underline{v}, \tag{28}$$

where $M \in \mathbb{R}^{N\times N}$ with $M_{i,j} := (\phi_i, \phi_j)$ is the mass matrix. Apparently, we have

$$(\underline{u}, \vec{v}) \equiv (\underline{u}, \vec{v})_{l^2} = \underline{u}^T M \underline{v} = (u, v)_V.$$

Matrix Representations

Let W be a linear space with basis $\{\psi_i\}_{i=1,\dots,N'}$. For any $\mathcal{A} : V \mapsto W$, we give its primal representation, $\underline{\mathcal{A}} \in \mathbb{R}^{N' \times N}$, s.t. $\sum_{i=1}^{N'} \left(\underline{\mathcal{A}}\right)_{i,j} \psi_i = \mathcal{A}\phi_j \;\; (j = 1, \dots, N)$:

$$(\psi_1, \dots, \psi_{N'})\underline{\mathcal{A}} = \mathcal{A}(\phi_1, \dots, \phi_N) = (\mathcal{A}\phi_1, \dots, \mathcal{A}\phi_N). \tag{29}$$

The dual representation for $\mathcal{A} : V \mapsto V$ is denoted by $\left(\hat{\mathcal{A}}\right)_{i,j} := (\mathcal{A}\phi_j, \phi_i)$. $\Longleftarrow$ the stiffness matrix

Lemma (Simple relations between matrix representations)

If $\mathcal{A}, \mathcal{B} : V \mapsto V$ and $v, u \in V$, we have the following results:

1. $\underline{\mathcal{A}\mathcal{B}} = \underline{\mathcal{A}}\,\underline{\mathcal{B}}$;
2. $\underline{\mathcal{A}v} = \underline{\mathcal{A}}\,\underline{v}$;
3. $\sigma(\mathcal{A}) = \sigma(\underline{\mathcal{A}}), \quad \kappa(\mathcal{A}) = \kappa(\underline{\mathcal{A}})$;
4. $\vec{v} = M\,\underline{v}, \quad \overrightarrow{\mathcal{A}v} = \hat{\mathcal{A}}\,\underline{v}$;
5. $\hat{\mathcal{A}} = M\,\underline{\mathcal{A}}$;
6. $(u, v) = (M\underline{u}, \underline{v})$.

Example: Finite Element Matrices

NCMIS

Setting for finite element methods:

- $V = V_h$ is the piecewise linear finite element space.
- $\{\phi_i\}_{i=1,\dots,N}$ are the basis functions for FEM.
- Let A be the coefficient matrix $(A)_{i,j} = a_{i,j} := a[\phi_i, \phi_j]$.
- By definition, $A = \hat{\mathcal{A}} \in \mathbb{R}^{N\times N}$ is the stiffness matrix corresponding to $\mathcal{A}$.
- Let $\underline{u} = (u_i)_{i=1}^{N} \in \mathbb{R}^N$ be the vector of coefficients of u_h, namely $\underline{u_h}$.
- Let $\vec{f} = (f_i)_{i=1}^{N} := \{\langle f, \phi_i\rangle\}_{i=1}^{N}$.

Linear algebraic system for finite element discretization of $\mathcal{A}u = f$:

$$\hat{\mathcal{A}}\underline{u} = \vec{f} \quad \text{or} \quad A\underline{u} = \vec{f}. \qquad \text{(We have been abusing notation!)}$$

Upon solving this linear system, we obtain a discrete approximation

$$u_h = \sum_{i=1}^{N} \underline{u}_i \phi_i.$$

Condition of FE Matrices, Revisited

Spectral radius and condition number of $\mathcal{A}_h$: Suppose that mesh is uniform with meshsize h. From the Poincaré inequality and the inverse inequality, that

$$\|v\|_0^2 \lesssim \|\nabla v\|_0^2 = (\mathcal{A}_h v, v) \le \|v\|_1^2 \lesssim h^{-2}\|v\|_0^2, \quad \forall v \in V_h.$$

Hence, we have $\rho(\mathcal{A}_h) \cong h^{-2}$ and $\kappa(\mathcal{A}_h) \cong h^{-2}$.

Spectrum of mass matrix: The mass matrix $M \in \mathbb{R}^{N\times N}$, $M_{i,j} = (\phi_i, \phi_j)$, satisfies

$$(M\underline{v}, \underline{v}) = \sum_{i,j} \underline{v}_i\, \underline{v}_j\, (\phi_i, \phi_j) = (v, v) = \int_\Omega v^2(x)\, dx \cong h^d \sum_i \underline{v}_i^2 \cong h^d(\underline{v}, \underline{v}).$$

It is well-known that the mass matrix is SPD and well-conditioned, i.e.,

$$h^d\|\xi\|_0^2 \lesssim \xi^T M\, \xi \lesssim h^d\|\xi\|_0^2, \quad \forall\, \xi \in \mathbb{R}^N.$$

Spectrum of stiffness matrix: The stiffness matrix A is SPD and

$$h^d\|\xi\|_0^2 \lesssim \xi^T A\, \xi \lesssim h^{d-2}\|\xi\|_0^2, \quad \forall\, \xi \in \mathbb{R}^N.$$

Hence the spectral radius $\rho(A) \cong h^{d-2}$ and the condition number $\kappa(A) \cong h^{-2}$.

Richardson Method, Revisited

Consider solving the $\mathcal{P}_1$-Lagrange finite element system for the Poisson's equation:

$$\hat{A}\underline{u} = \vec{f}.$$

The simplest iterative solver is the well-known Richardson method:

$$\underline{u}^{\text{new}} = \underline{u}^{\text{old}} + \omega\left(\vec{f} - \hat{A}\,\underline{u}^{\text{old}}\right).$$

It is equivalent to

$$\underline{u}^{\text{new}} = \underline{u}^{\text{old}} + \omega\left(M\underline{f} - M\underline{\mathcal{A}}\,\underline{u}^{\text{old}}\right) = \underline{u}^{\text{old}} + \omega M\left(\underline{f} - \underline{\mathcal{A}}\,\underline{u}^{\text{old}}\right).$$

That is to say, the Richardson method, can be written in the operator form as

$$u^{\text{new}} = u^{\text{old}} + \mathcal{B}_\omega\left(f - \mathcal{A}\,u^{\text{old}}\right)$$

with an iterator $\mathcal{B}_\omega$, whose matrix representation is $\underline{\mathcal{B}_\omega} = \omega M$.

The operator form of the Richardson method is

$$\mathcal{B}_\omega v := \omega \sum_{i=1}^{N} (v, \phi_i)\phi_i, \quad \forall v \in V.$$