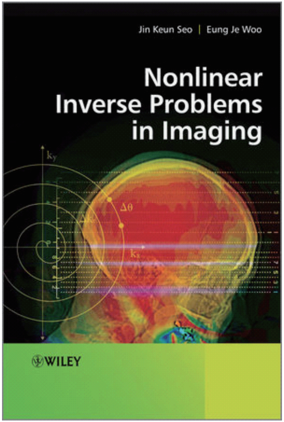
Partial Differential Equations



Part III-1: Energy functional & Sobolev Space

Homepage

http://www.seojinkeun.com/♯!pde-course-/c1laj

Jin Keun Seo
CSE, Yonsei University

Elliptic PDE for inhomogeneous material

Consider the following PDE:

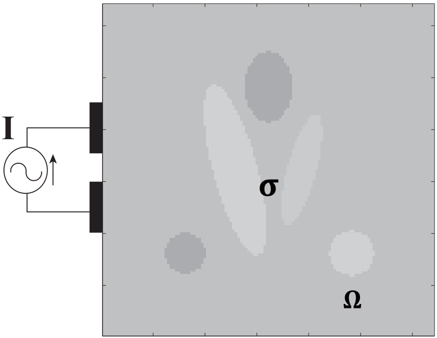
$$\begin{cases} -\nabla \cdot (\sigma(\mathbf{r})\nabla u(\mathbf{r})) = 0 & \text{in } \Omega \\ u|_{\partial\Omega} = f \end{cases}$$

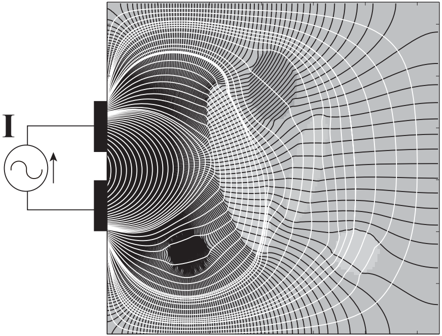
- This PDE makes sense in classical sense when $\sigma \in C^1(\overline{\Omega})$ and $u \in C^2(\overline{\Omega})$.
- In practice, the material property σ may change abruptly. For example, a conductivity distribution σ inside the human body Ω may have a jump along the boundary of two different organs. Along such a boundary, the electrical field $\mathbf{E} = -\nabla u$ may be **discontinuous** due to interface conditions of the electric field (like the refractive condition of Snell's law). In this case, there exists no solution $u \in C^2(\Omega)$ in the classical sense.

The difficulty regarding the refraction contained in PDE can be removed by the use of the **variational framework**:

$$\int_\Omega \sigma \nabla u \cdot \nabla \phi = 0, \quad \forall \phi \in H_0^1(\Omega)$$

$$\begin{cases} -\nabla \cdot (\sigma(\mathbf{r})\nabla u(\mathbf{r})) = 0 & \text{in } \Omega \\ \mathbf{n} \cdot (\sigma \nabla u)|_{\partial\Omega} \approx I\chi_{E_1} - I\chi_{E_2} \end{cases}$$





Variational framework

Consider the following one-dimensional Dirichlet problem:

$$\begin{cases} -\frac{d}{dx}\left(\sigma(x)\frac{d}{dx}u(x)\right) = 0 & \text{in } (-1,1) \\ u(-1) = -2, \quad u(1) = 1 \end{cases} \qquad \text{where } \sigma(x) = \begin{cases} 2 & \text{if } x \geq 0 \\ 1 & \text{if } x < 0. \end{cases}$$

- u satisfies $u''(x) = 0$ in $(-1,1) \setminus \{0\}$ with $u(-1) = 2, \quad u(1) = 1$.
- $u \neq v$, the solution of $v''(x) = 0$ in $(-1,1)$ with $v(-1) = 2, \quad v(1) = 1$.
- The classical derivative u' does not exist at $x = 0$:

$$u(x) = \begin{cases} x & \text{if } \; 0 \leq x < 1 \\ 2x & \text{if } \; -1 < x < 0 \end{cases} \qquad \& \qquad u'(x) = \begin{cases} 1 & \text{if } x > 0 \\ \nexists & \text{if } x = 0 \\ 2 & \text{if } x < 0 \end{cases}$$

- The difficulty can be removed by the use of the **variational framework**:

$$\int_{-1}^{1} \sigma(x)u'(x)\phi'(x)dx = 0 \qquad \forall \phi \in C_0^1(-1,1).$$

Energy functional

$$\begin{cases} -\frac{d}{dx}\left(\sigma(x)\frac{d}{dx}u(x)\right) = 0 & \text{in } (-1,1) \\ u(-1) = -2, \quad u(1) = 1 \end{cases} \qquad \text{where } \sigma(x) = \begin{cases} 2 & \text{if } x \geq 0 \\ 1 & \text{if } x < 0. \end{cases}$$

- Show that u satisfies the transmission conditions

$$u(0^+) = u(0^-) \quad \& \quad 2u'(0^+) = u'(0^-).$$

- A practically meaningful solution u should have a finite energy

$$\Phi(u) := \frac{1}{2}\int_{-1}^{1} \sigma |u'|^2 dx < \infty.$$

- Indeed, $u = \arg\min_{v \in \mathcal{A}} \Phi(v)$, a minimizer of $\Phi(v)$ within the admissible set

$$\mathcal{A} := \{v \; : \; \Phi(v) < \infty, \quad v(-1) = -1, \quad v(1) = 2\}.$$

Elliptic PDE

$$\begin{cases} -\nabla \cdot (\sigma(\mathbf{r})\nabla u(\mathbf{r})) = 0 & \text{in } \Omega \\ u|_{\partial\Omega} = f \end{cases}$$

- The physically meaningful solution u must have a finite energy:

$$\Phi(v) = \int_\Omega \sigma(x)|\nabla v(\mathbf{r})|^2 dx < \infty.$$

- Hence, u should be contained in the set

$$\mathcal{A}_\Phi := \{v \in L^2(\Omega) : \Phi(v) < \infty\}.$$

- Assuming $0 < \inf_\Omega \sigma < \sup_\Omega \sigma < \infty$,

$$\mathcal{A}_\Phi \ = \ H^1(\Omega) := \ \{v \ : \ \|v\|_{H^1(\Omega)} := \sqrt{\int_\Omega |\nabla u|^2 + |u|^2 d\mathbf{x}} < \infty\}$$

This space is called the Sobolev space H^1.

Weak solution

Assume $\sigma \notin C(\bar{\Omega})$ and $0 < \sigma < \infty$. Consider

$$\begin{cases} -\nabla \cdot (\sigma(\mathbf{r})\nabla u(\mathbf{r})) = 0 & \text{in } \Omega \\ u|_{\partial\Omega} = f \end{cases}$$

[Assume $\sigma \notin C(\bar{\Omega})$]

- When $u \in C^2(\bar{\Omega})$, there is no difference between the classical and variational problems.
- However, if $\sigma \notin C(\bar{\Omega})$, then the minimization problem has no solution in the class $C^2(\bar{\Omega})$. Obviously, the classical problem does not have a solution.
- We can construct a minimizing sequence $\{u_n\}$ in $C^2(\bar{\Omega})$ which is a Cauchy sequence with respect to the norm $\|u\|_{H^1(\Omega)}$.
- Although the Cauchy sequence $\{u_n\}$ does not converge within $C^2(\bar{\Omega})$, it converges in the Sobolev space $H^1(\Omega)$, the completion of $C^2(\bar{\Omega})$ with respect to the norm $\|u\|_{H^1(\Omega)}$. This means that we can solve the minimization and variational problem within the Sobolev space $H^1(\Omega)$.

Generalized derivative

The generalized derivative can be explained by means of the integrating by parts formula:

$$\int_\Omega u\, \partial_{x_i}\, \phi\, dx = -\int_\Omega \partial_{x_i} u\, \phi\, d\mathbf{x} \quad (\forall \phi \in C_0^\infty(\Omega)).$$

In general,

$$\int_\Omega u\, \partial^\alpha\, \phi\, d\mathbf{x} = (-1)^{|\alpha|} \int_\Omega \partial^\alpha u\, \phi\, d\mathbf{x} \quad (\forall \phi \in C_0^\infty(\Omega))$$

where the notions ∂^α and $|\alpha|$ are understood in the following way:

- $\alpha = (\alpha_1, \cdots, \alpha_n) \in \mathbb{N}_0^n$, $\mathbb{N} = \{1, 2, \cdots\}$, and $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$.
- $\partial^\alpha u = \partial_{x_1}^{\alpha_1} \cdots \partial_{x_n}^{\alpha_n} u$; for example, $\partial^{(2,0,3)} u = \partial_{x_1}^2 \partial_{x_3}^3 u$.
- $|\alpha| = \sum_{k=1}^n \alpha_k$.

A function v_i satisfying the following equality behaves like the classical derivative $\partial_{x_i} u$:

$$\int_\Omega u\, \partial_{x_i} \phi\, dx = -\int_\Omega \underbrace{v_i}_{\partial_{x_i} u}\, \phi\, dx, \quad \forall \phi \in C_0^1(\Omega).$$

Sobolev space

We are now ready to introduce the Sobolev space $W^{k,p}(\Omega)$ and $W_0^{k,p}(\Omega)$ where Ω be a domain in $\mathbb{R}^n$ with its boundary $\partial\Omega$ and $1 \leq p < \infty$:

- $W^{1,p}(\Omega)$ and $W_0^{1,p}(\Omega)$ are the completion (or closure) of $C^\infty(\bar{\Omega})$ and $C_0^\infty(\Omega)$, respectively, with respect to the norm

$$\|u\|_{W^{1,p}} := \left(\int_\Omega |u|^p + |\nabla u|^p dx \right)^{1/p}.$$

  In other words,

$$W_0^{1,p}(\Omega) = \{u \ : \ \exists u_k \in C_0^\infty(\Omega) \ s.t. \ \lim_{k\to\infty} \|u_k - u\|_{W^{1,p}(\Omega)} = 0\}.$$

- $W^{2,p}(\Omega)$ and $W_0^{2,p}(\Omega)$ are the completion (or closure) of $C^\infty(\bar{\Omega})$ and $C_0^\infty(\Omega)$, respectively, with respect to the norm

$$\|u\|_{W^{2,p}} := \left(\int_\Omega |u|^p + |\nabla u|^p + |\nabla\nabla u|^p dx \right)^{1/p}.$$

- We denote $H^1(\Omega) = W^{1,2}(\Omega)$, $H_0^1(\Omega) = W_0^{1,2}(\Omega)$, $H^2(\Omega) = W^{2,2}(\Omega)$ and $H_0^1(\Omega) = W_0^{2,2}(\Omega)$.

Two examples

- Let $\Omega = (0,1)$.

$$u(x) = x(1-x) \in H_0^1(\Omega) \cap H^2(\Omega) \quad \text{but} \quad u \notin H_0^2(\Omega).$$

- There exists a function $u \in H^1(\mathbb{R}^3)$ which is not bounded on every non-empty open set in $\mathbb{R}^3$. Denoting the set of rational numbers by $\mathbb{Q}$, there exists a sequence $\{\mathbf{q}_m\}_{m=1}^{\infty} = \mathbb{Q}^3$.

$$u(\mathbf{r}) := \frac{1}{1+|\mathbf{r}|^2} \sum_{m=1}^{\infty} 2^{-m} \log|\mathbf{r} - \mathbf{q}_m| \in H^1(\mathbb{R}^3)$$

but it is unbounded on every non-empty open set.

Practically meaningful solution

Example

Let $\Omega := \{(r\cos\theta, r\sin\theta) \ : \ 0 < r < 1, \ \ 0 < \theta < \frac{3\pi}{2}\}$. Let

$$u = (r^{-2/3} - r^{2/3})\sin(\frac{2}{3}\theta)$$

with $r = \sqrt{x^2 + y^2}$ and $\tan\theta = \frac{y}{x}$.

Clearly, u satisfies

$$\nabla^2 u = 0 \quad \text{in} \quad \Omega \ , \qquad\qquad u|_{\partial\Omega} = 0.$$

According to the maximum principle,

$$u = 0 \quad \text{in} \quad \Omega \qquad ???.$$

What is wrong with this conclusion? We should note that $u \notin H^1(\Omega)$ and hence u is not a practically meaningful solution. Because of $u \notin H^1(\Omega)$, we cannot apply the maximum principle.

Inaccurate statement

For a given $f \in C(\partial\Omega)$, let u be a function in $\mathcal{A}_f := \{v \in C^2(\Omega) \cap C(\overline{\Omega}) : \ v|_{\partial\Omega} = f\}$ which minimizes the integral $\Phi(v) = \int_\Omega |\nabla v|^2$.

Is the following statement correct? **NO!**

- Any function u which satisfies $\Delta u = 0$ in Ω with $u|_{\partial\Omega} = f \in C(\partial\Omega)$ is a minimizer of $\Phi(v) = \int_\Omega |\nabla v|^2$. .

The space $C^2(\Omega)$ is not the right space in which to set the admissible set for the existence of a minimizer due to the lack of completeness with respect to the distance $f(u_n - u_m)$ in terms of energy between two functions u_n and u_m. In practice, the existence of a minimizer is obvious provided that the given system is physically existing or concerns observable quantities. Hence, it requires to find the correct space of physically meaningful functions with a proper distance concept. The correct space for solutions of the Laplace equation is the Sobolev space $\{v \in H^1(\Omega) \ : \ v|_{\partial\Omega} = \phi\}$ equipped with the norm $\|v\| = \sqrt{\int_\Omega |v|^2 + |\nabla v|^2}$, which measures both the size and regularity of a function. Before the twentieth century, the Hilbert space $H^1(\Omega)$ and measure theory had not been introduced; thus, there was insufficient knowledge to validate Dirichlet's principle in a rigorous way. This issue of ascertaining the existence of a minimizer might have been a possible motivation for the development of the notion of compactness.

Hadamard's example

Is the following statement correct? **NO!**

- Any function u which satisfies $\Delta u = 0$ in Ω with $u|_{\partial\Omega} = f \in C(\partial\Omega)$ is a minimizer of $\Phi(v) = \int_\Omega |\nabla v|^2$. .

Example

Hadamard gave the following counter example of a solution of Dirichlet's problem in the unit disk Ω whose energy blows up:

$$u(\mathbf{x}) = \sum_{n=1}^{\infty} 2^{-n} |\mathbf{x}|^{2^{2n}} \cos(2^{2n}\theta) \qquad (\tan\theta = x_2/x_1).$$

In this example, u has the boundary data $\phi = \sum_{n=1}^{\infty} 2^{-n} \cos(2^{2n}\theta)$ which is continuous but not differentiable almost everywhere.

Theorem (Simplified version of Poincaré inequalities)

Let $\Omega = \{(x, y) \ : 0 < x, y < a\}$. A simplified version of the Poincaré inequality is

$$\sup_{u \in C_0^1(\Omega)} \frac{\|u\|_{L^2(\Omega)}}{\|\nabla u\|_{L^2(\Omega)}} \leq C$$

where C is a positive constant depending only on Ω.

The Poincaré inequality uses the special property of $u|_{\partial\Omega} = 0$ to get

$$|u(x, y)|^2 \leq \left| \int_0^a |\frac{\partial}{\partial x} u(x', y)| dx' \right|^2 \leq a \int_0^a |\nabla u(x', y)|^2 dx'$$

and hence

$$\int_\Omega |u|^2 \leq a^2 \int_\Omega |\nabla u|^2,$$

Theorem (Simplified version of trace inequalities)

Let $\Omega = \{(x, y) \ : 0 < x, y < a\}$. A simplified version of the trace inequality is

$$\sup_{u \in C^1(\overline{\Omega})} \frac{\|u\|_{L^2(\partial\Omega)}}{\|u\|_{H^1(\Omega)}} \leq C$$

where C is a positive constant depending only on Ω.

From the fundamental theorem of calculus,

$$\int_0^a |u(0,y)|^2 dy = \int_0^a \left| u(x,y) - \int_0^x \frac{\partial}{\partial x} u(x', y) dx' \right|^2 dy$$

and therefore

$$\begin{aligned}
\int_0^a |u(0,y)|^2 dy &\leq \int_0^a \left[|u(x,y)| + \int_0^a |\nabla u(x', y)| dx' \right]^2 dy \\
&\leq \int_0^a \left[2|u(x,y)|^2 + a \int_0^a |\nabla u(x', y)|^2 dx' \right] dy \\
&\leq 2 \int_0^a |u(x,y)|^2 dy + 2a \int_\Omega |\nabla u|^2 d\mathbf{r} \qquad (d\mathbf{r} = dxdy).
\end{aligned}$$

Theorem (Sobolev's inequality)

Let $u \in H^1(\mathbb{R}^n)$. The following inequality holds:

- *For $n \geq 3$,*
$$\|u\|^2_{L^{2n/(n-2)}} \leq C_n \|\nabla u\|^2_{L^2(\mathbb{R}^n)}$$
where $C_n = \frac{4}{n(n-2)} 2^{-2/n} \pi^{-1-1/n} [\Gamma(\frac{n+1}{2})]^{2/n}$.

- *For $n = 1$,* $\|u\|_{L^\infty(\mathbb{R})} \leq \frac{1}{2}\|u\|^2_{H^1(\mathbb{R})}$, $\quad \sup_{x,y} \frac{|u(x)-u(y)|}{|x-y|} \leq \|u'\|_{L^2}$.

- *For $n = 2$,* $\|u\|_{L^q(\mathbb{R}^2)} \leq C_q \|u\|^2_{H^1(\mathbb{R})}$, $\quad (\forall 2 \leq q < \infty)$ *where*
$C_q \leq \left[q^{1-2/q}(q-1)^{-1+1/q}((q-2)/8\pi)^{1/2-1/q}\right]^2$.

- *Let Ω be a $C^{0,1}$-domain and $1 \leq p \leq q$, $m \geq 1$ and $k \leq m$. Then,*
$$\begin{aligned} \|u\|_{L^{np/(n-kp)}(\Omega)} &\leq C\|u\|_{W^{k,p}(\Omega)} && \text{if } kp < n, \\ \|u\|_{C^m(\Omega)} &\leq C\,\|u\|_{W^{k+m,p}(\Omega)} && \text{if } kp > n \end{aligned}$$

  where C is independent of u.

For $n \geq 3$, the proof of the Sobolev inequality is based on the identity

$$u(\mathbf{x}) = \int_{\mathbb{R}^n} \nabla\left(\frac{c}{|\mathbf{x}-\mathbf{y}|^{n-2}}\right) \cdot \nabla u(\mathbf{y}) d\mathbf{y}.$$

where $c/|\mathbf{x}-\mathbf{y}|^{n-2}$ is the fundamental solution of Laplacian.

Helmholtz Decomposition

The Helmholtz decomposition states that any smooth vector field **F** in a smooth bounded domain Ω can be resolved into the sum of a divergence-free (solenoidal) vector field and a curl-free (irrotational) vector field.

Theorem

Every vector field $\mathbf{F}(\mathbf{r}) = (F_1(\mathbf{r}), F_2(\mathbf{r}), F_3(\mathbf{r})) \in [L^2(\Omega)]^3$ *can be decomposed into*

$$\mathbf{F}(\mathbf{r}) = -\nabla u(\mathbf{r}) + \nabla \times \mathbf{A}(\mathbf{r}) + \mathit{Harmonic} \quad \text{in } \Omega$$

where u *is a scalar function,* $\nabla \cdot \mathbf{A} = 0$ *and Harmonic is a vector field whose Laplacian is zero in* Ω*. Moreover,* u *and* $\mathbf{A}$ *are solutions of* $\nabla^2 u = \nabla \cdot \mathbf{F}$ *and* $\nabla^2 \mathbf{A} = \nabla \times \mathbf{F}$ *with appropriate boundary conditions. Hence, these can be uniquely determined up to harmonic functions;*

$$u(\mathbf{r}) = -\int_\Omega \frac{\nabla \cdot \mathbf{F}(\mathbf{r}')}{4\pi|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' + \mathit{Harmonic}$$

and

$$\mathbf{A}(\mathbf{r}) = \int_\Omega \frac{\nabla \times \mathbf{F}(\mathbf{r}')}{4\pi|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' + \mathit{Harmonic}.$$

Helmholtz's decomposition

$$\mathbf{F}(\mathbf{r}) = -\nabla \int_\Omega \frac{\nabla \cdot \mathbf{F}(\mathbf{r}')}{4\pi|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}' + \nabla \times \int_\Omega \frac{\nabla \times \mathbf{F}(\mathbf{r}')}{4\pi|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}' + Harmonic$$

- Write

$$\mathbf{F}(\mathbf{r}) = \int_\Omega \delta(\mathbf{r}-\mathbf{r}')\mathbf{F}(\mathbf{r}')d\mathbf{r}' = -\int_\Omega \nabla^2 \left(\frac{1}{4\pi|\mathbf{r}-\mathbf{r}'|}\right) \mathbf{F}(\mathbf{r}')d\mathbf{r}'.$$

- Integrating by parts

$$\mathbf{F}(\mathbf{r}) = -\int_\Omega \frac{1}{4\pi|\mathbf{r}-\mathbf{r}'|} \nabla^2 \mathbf{F}(\mathbf{r}')d\mathbf{r}' + Harmaonic.$$

- Using the vector identity $-\nabla^2 \mathbf{F} = \nabla \times (\nabla \times \mathbf{F}) - \nabla(\nabla \cdot \mathbf{F})$,

$$\mathbf{F}(\mathbf{r}) = \int_\Omega \frac{1}{4\pi|\mathbf{r}-\mathbf{r}'|} \left[\nabla \times (\nabla \times \mathbf{F}) - \nabla(\nabla \cdot \mathbf{F})\right] dV + Harmonic, \qquad \mathbf{r} \in \Omega.$$

Integrating by parts again, we have the result.

Variational Problems

& FEM

Example: Poisson's equation

Define the map $a : H_0^1(\Omega) \times H_0^1(\Omega) \to \mathbb{R}$ by

$$a(u,\phi) = \int_\Omega \sigma \nabla u \cdot \nabla \phi \qquad (0 < c_0 < \sigma(\mathbf{r}) < c_1 < \infty)$$

and define $b : H_0^1(\Omega) \to \mathbb{R}$ by

$$b(\phi) = \int_\Omega f \phi dx.$$

The solvability problem of $\begin{cases} -\nabla \cdot (\sigma \nabla u) = f & \text{in } \Omega \\ u|_{\partial\Omega} = 0 \end{cases}$ is equivalent to the uniqueness and existence question of finding $u \in X = H_0^1(\Omega)$ satisfying

$$a(u,\phi) = b(\phi), \quad \forall \phi \in X.$$

Note that the map $a : H_0^1(\Omega) \times H_0^1(\Omega) \to \mathbb{R}$ satisfies the followings:

- both $a(u,\cdot) : X \to \mathbb{R}$ and $a(\cdot, w) : X \to \mathbb{R}$ are linear for all $u, w \in X$,
- $|a(u,v)| < c_1 \|u\| \|v\|$ and $c_0 \|u\|^2 \leq a(u,u)$.

Example: Finite Element Method (FEM) for solving Poisson's Equation

Assuming that the Hilbert space $H_0^1(\Omega)$ has a basis $\{\phi_k \ : \ k = 1, 2, \cdots\}$, the variational problem [$a(u, \phi) = b(\phi), \forall \phi \in X$] is to determine the coefficient $\{u_k : k = 1, 2, \cdots\}$ of $u = \sum u_k \phi_k$ satisfying

$$\underbrace{a(\sum_k u_k \phi_k, \phi_j)}_{\int_\Omega \sigma \nabla u \cdot \nabla \phi_j dx} = \underbrace{b(\phi_j)}_{\int_\Omega f \phi_j dx}, \quad \forall j = 1, 2, \cdots .$$

Taking advantage of the linearity of $a(\cdot, \cdot)$ and $b(\cdot)$, the problem is equivalent to solve

$$\underbrace{\begin{pmatrix} a(\phi_1, \phi_1) & a(\phi_1, \phi_2) & \cdots & \\ a(\phi_2, \phi_1) & a(\phi_2, \phi_2) & \cdots & \\ a(\phi_3, \phi_1) & a(\phi_3, \phi_2) & \cdots & \\ & & & \ddots \end{pmatrix}}_{A} \underbrace{\begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ \vdots \end{pmatrix}}_{\mathbf{u}} = \underbrace{\begin{pmatrix} b(\phi_1) \\ b(\phi_2) \\ b(\phi_3) \\ \vdots \end{pmatrix}}_{\mathbf{b}}$$

From the fact that $c_0 \|u\|^2_{H^1(\Omega)} \le a(u, u)$, A is positive definite and invertible. Therefore, for a given data $\mathbf{b}$, there exists a unique $\mathbf{u}$ satisfying $A\mathbf{u} = \mathbf{b}$.

Theorem (Lax-Milgram theorem)

Suppose the bounded bilinear map $a : X \times X \to \mathbb{R}$ *is symmetric and*

$$a(u,u) \geq c\|u\|^2, \quad c > 0.$$

Suppose $b(\cdot) : X \to \mathbb{R}$ *is linear.*

- *There exists a unique solution* $u \in X$ *of the minimization problem*

$$\text{Minimize} \quad \Phi(u) := \frac{1}{2} a(u,u) - b(u), \quad u \in X.$$

- *The solution* u *of the minimization problem is the solution of variational problem:*

$$a(u,\phi) = b(\phi), \quad \forall \phi \in X.$$

Example: Consider the solution $u \in X = H_0^1(\Omega)$ of Poisson's equation $-\nabla^2 u = f$ in Ω. This u is a minimizer of the problem:

$$\min \quad \Phi(u) := \underbrace{\frac{1}{2}\int_\Omega |\nabla v(x)|^2 dx}_{a(u,u)} - \underbrace{\int_\Omega f(x)u(x)dx}_{b(u)}, \qquad u \in X = H_0^1(\Omega).$$

Proof of Lax-Milgram: Existence

Existence of the minimizer can be proven by showing

$$\exists\, u \in X \quad s.t.\ \ \Phi(u) = \alpha := \inf_{v \in X} \Phi(v).$$

Proof. Let $\{u_n\}$ be a sequence s.t. $\Phi(u_n) \to \alpha$.

- $a(u_n, u_n) + a(u_m, u_m) = \frac{1}{2} a(u_n - u_m, u_n - u_m) + \frac{1}{2} a(u_n + u_m, u_n + u_m)$.
- $2\Phi(u_n) + 2\Phi(u_m) = \frac{1}{2} a(u_n - u_m, u_n - u_m) + 4\Phi(\frac{u_n + u_m}{2})$.
- Setting $n, m \to \infty$, we have

$$\begin{aligned} 4\alpha \longleftarrow 2\Phi(u_n) + 2\Phi(u_m) &= \frac{1}{2} a(u_n - u_m, u_n - u_m) + 4\Phi(\frac{u_n + u_m}{2}) \\ &\geq \frac{c}{2} \|u_n - u_m\|^2 + 4\alpha. \end{aligned}$$

- Hence $\{u_n\}$ is a Cauchy sequence and $\exists u \in X$ so that $u = \lim_{n\to\infty} u_n$. Since $\Phi(\cdot)$ is continuous due to the assumption that $a(\cdot,\cdot)$ and $b(\cdot)$ are bounded and linear, we prove existence of the minimizer u:

$$\Phi(u) = \lim_{n\to\infty} \Phi(u_n) = \alpha.$$

Proof of Lax-Milgram: Uniqueness

- If u is a minimizer,

$$\frac{d}{dt}\Phi(u + t\phi)|_{t=0} = 0 \quad (\forall \phi \in X) \qquad \Longrightarrow \qquad a(u, \phi) = b(\phi) \quad (\forall \phi \in X).$$

- Hence, if u and v are minimizers, then

$$a(u, \phi) = b(\phi) = a(v, \phi) \qquad (\forall \phi \in X)$$

which is equivalent to

$$a(u - v, \phi) = 0 \qquad (\forall \phi \in X).$$

- In particular, $a(u - v, u - v) = 0$. Then, it follows from **the condition of $a(\cdot, \cdot)$** that

$$c\|u - v\|^2 \leq a(u - v, u - v) = 0,$$

which gives $u = v$.

Ritz Approach

Taking account of the finite element method, we assume the followings:

- X is a real Hilbert space with a norm $\|\cdot\|$.
- $\{X_n\}$ is a sequence of a finite dimensional subspace of X such that

$$X_n \subset X_{n+1} \qquad \& \quad \overline{\cup_{n=1}^{\infty} X_n} = X.$$

- $\{\phi_j^n : j = 1, \cdots, N_n\}$ is a basis of X_n.
- $a(\cdot,\cdot) : X \times X \to \mathbb{R}$ be a bounded, symmetric, strongly positive, bilinear map and

$$|a(u,v)| \le M\|u\|\|v\|, \quad c\|u\|^2 \le a(u,u), \quad \forall u, v \in X.$$

- $b \in X^*$ where X^* is the set of linear functional on X.

Consider the minimization problem:

$$\min \quad \Phi(u_n) := \frac{1}{2} a(u_n, u_n) - b(u_n), \quad u_n \in X_n.$$

The variational problem is as follows:

$$\text{Find} \quad u_n \in X_n \quad s.t. \quad a(u_n, \phi_n) = b(\phi_n), \quad \forall \phi_n \in X_n.$$

Theorem (Projection theorem)

The space X equipped with the new inner product and norm

$$\langle u, v\rangle_a := a(u,v), \quad \|u\|_a := \sqrt{a(u,u)} \qquad (\forall\ u, v \in X)$$

is also a Hilbert space. For any closed subspace $V = \overline{V} \subset X$ and for each $u \in X$,

$$\exists\ u_V \in \quad s.t. \quad \|u - u_V\|_a = \min_{v \in V} \|u - v\|_a.$$

Denoting $V^{\perp_a} := \{w \in X\ :\ a(w,v) = 0.\ \forall\ v \in V\}$, if u is decomposed as

$$u = u_1 + u_2, \quad (u_1 \in V,\ \ u_2 \in V^{\perp_a}),$$

then $u_1 = u_V$ and $u_2 = u - u_V$.

Example: Let $X = \mathbb{R}^3$ with its new inner product given by

$$\langle u, v\rangle_a := \langle \begin{pmatrix} 2 & 1 & 0 \\ 1 & 3 & 0 \\ 0 & 0 & 4 \end{pmatrix} u\ ,\ v\ \rangle \qquad (\forall\ u, v \in X = \mathbb{R}^3)$$

Let $V = \text{span}\{(1,1,0),(0,0,1)\} = \{v \in X\ :\ (-1,1,0)\cdot v = 0\}$. Then, $V^{\perp_a} := \text{span}\{(-4,3,0)\}$. For any $u \in X$, $u_{V^{\perp_a}} = \alpha(-4,3,0)$ for some α. This is the key of Riesz Representation:

$$\exists\ \text{unique}\ u \in X\ \ s.t.\ \langle u, v\rangle_a = \underbrace{(-1,1,0)\cdot v}_{\text{linear functional}} \qquad (\forall v \in X)$$

Proof of $\exists\, u_V \in V$ s.t. $\|u - u_V\|_a = \min_{v\in V} \|u - v\|_a$

- Let u be fixed. Since

$$\|u - v\|_a^2 = a(v, v) - 2a(u, v) + a(u, u),$$

the problem is equivalent to solve the following minimization problem:

$$\min_{v\in V} \quad \tilde{\Phi}(v) := \frac{1}{2} a(v, v) - b(v) \qquad b(v) = a(u, v)$$

- According to the Lax-Milgram theorem, $\exists u_V \in V$ of the minimization problem.
- Next, we will show $u - u_V \in V^\perp$.

$$\begin{aligned} & \|u - u_V\|_a^2 \le \|u - (u_V + tv)\|_a^2 \quad (\forall\, v \in V,\ \ \forall\, t \in \mathbb{R}) \\ \Rightarrow \quad & 0 \le -2ta(u - u_V, v) + t^2 a(v, v) \quad (\forall\, v \in V,\ \ \forall\, t \in \mathbb{R}) \\ \Rightarrow \quad & -t\, a(v, v) \le 2a(u - u_V, v) \le t\, a(v, v), \quad (\forall\, v \in V,\ \ \forall\, t > 0). \end{aligned}$$

- Letting $t \to 0$, we get $a(u - u_V, v) = 0 \quad (\forall\, v \in V) \quad \Rightarrow \quad u - u_V \in V^{\perp_a}$.
- Uniqueness of the decomposition follows from the fact that

$$0 = u - u = \underbrace{(u_1 - u_V)}_{\in\ V} + \underbrace{(u_2 - u + u_V)}_{\in\ V^{\perp_a}} \quad \Rightarrow \quad u_1 = u_V \ \ \&\ \ u_2 = u - u_V.$$

Theorem (Riesz representation theorem)

For each linear functional $b \in X^*$, there is a unique $u_* \in X$ such that

$$b(v) = a(u_*, v), \quad (\forall v \in X).$$

Moreover, $\|b\| = \sqrt{a(u_*, u_*)}$.

Assume $b \neq 0$. Let $V := \{w \in X : b(w) = 0\}$. Claim: $\dim V^{\perp_a} = 1$.

- Since $V^{\perp_a} \neq \emptyset$, $\exists u \in V^{\perp_a}$ s.t. $b(u) = 1$.
- Since $b(w - b(w)u) = b(w) - b(w)b(u) = 0$ for all $w \in X$,

$$w = \underbrace{w - b(w)u}_{\in\ V} + \underbrace{b(w)u}_{\in\ V^{\perp_a}} \qquad (\forall w \in X).$$

Projection theorem $\Rightarrow$ the decomposition is unique & $\dim V^{\perp_a} = 1$.

For all $w \in X$, we have

$$a(u, w) = a(u, \underbrace{w - b(w)u}_{\in V} + b(w)u) = a(u, b(w)u) = a(u, u)b(w).$$

Hence, $u_* = u/a(u, u)$ is the solution!
Since $|b(w)| \leq \|u_*\|_a \|w\|_a$ and $|b(u_*)| = \|u_*\|_a^2$, we have $\|b\| = \|u_*\|_a$.

Theorem (Ritz Error Estimate)

Assume that $u_n \in X_n$ is a unique solution of the minimization problem

$$\min \quad \Phi(u_n) := \frac{1}{2}a(u_n, u_n) - b(u_n), \quad u_n \in X_n.$$

Let u be a solution of the minimization problem with X_n replaced by X. Then,

$$(i) \quad \lim_{n\to\infty} \|u_n - u\| = 0,$$

$$(ii) \quad \|u - u_n\| \leq \frac{M}{c} \min_{v\in X_n} \|u - v\|,$$

$$(iii) \quad \frac{c}{2}\|u - u_n\|^2 \leq \Phi(u_n) - \Phi(u).$$

- $a(u, v) = b(v) = a(u_n, v)$ for all $v \in X_n$ implies $a(u - u_n, v) = 0 \quad (\forall v \in X_n)$.
- For all $v \in X_n$, we have

$$\|u - u_n\|^2 \leq \frac{1}{c}a(u - u_n, u - u_n) = \frac{1}{c}a(u - u_n, u - v) \leq \frac{M}{c}\|u - u_n\| \, \|u - v\|$$

  which gives (ii). (i) follows from the assumption that $X_n \to X$ as $n \to \infty$.
- For all $v \in X$,

$$\Phi(u + v) = \frac{1}{2}a(u + v, u + v) - b(u + v) = \frac{1}{2}a(v, v) + \underbrace{(a(u, v) - b(v))}_{=0} + \Phi(u)$$