



# Note For Finite Element Methods

Zhejiang University

作者: Shuang Hu

组织: Zhejiang University

时间: Sept 14, 2022

版本: 1.0

简介: 2022 秋冬学期“有限元方法”课程笔记



# 目录

<b>第 1 章 一维有限元方法</b>	<b>1</b>
1.1 为什么需要有限元方法?	1
1.2 从一维边值问题说起	1
1.3 有限元思想的导出	2
1.3.1 Galerkin 近似	2
1.3.2 Ritz 方法	2
1.4 有限元方法	3
1.4.1 线性有限元空间	3
1.4.2 刚度矩阵与负载	3
1.5 误差分析	4
1.6 一阶非齐次 BVP	8
1.7 其他边值问题	9
1.7.1 Neumann 边值问题	9
1.7.2 Robin 边值问题	9
1.8 本章总结	9
<b>第 2 章 变分原理</b>	<b>10</b>
2.1 变分问题以及等价形式	10
2.1.1 预备知识	10
2.1.2 存在唯一性	11
2.1.3 变分问题的等价形式	12
2.2 Lax-Milgram 引理	13
2.3 具体实例	14
2.3.1 二阶椭圆方程	14
2.3.2 四阶双调和方程	15
2.4 本章总结	16
<b>第 3 章 Sobolev 空间</b>	<b>17</b>
3.1 $L^p$ 空间内容回顾	17
3.2 广义导数	18
3.3 磨光算子以及相关应用	19
3.3.1 磨光算子	19
3.3.2 均值逼近定理	21
3.3.3 单位分解定理	23
3.4 Sobolev 空间	24
3.4.1 相关定义	24
3.4.2 Sobolev 嵌入定理	25
3.4.3 迹算子	27
3.5 本章小结	27
<b>第 4 章 有限元离散化</b>	<b>28</b>
4.1 有限元离散的步骤	28

4.1.1	变分问题	28
4.1.2	有限元子空间	29
4.1.3	离散线性方程组与求解	29
4.2	二维情形	31
4.2.1	重心坐标	31
4.2.2	三角单元内分段多项式的形式	32
4.2.2.1	线性 Lagrange 元	33
4.2.2.2	完全二次 Lagrange 元	33
4.2.2.3	完全三次 Lagrange 元	33
4.2.2.4	不完全三次 Lagrange 元	34
4.2.2.5	完全三次 Hermite 元	34
4.2.3	矩形单元内分段多项式的形式	34
4.2.3.1	双线性元	34
4.2.3.2	双二次元	35
4.2.3.3	奇妙八点元	35
4.3	有限元求解方程示例	35
4.3.1	构造变分形式	36
4.3.2	进行区域剖分	36
4.3.3	建立有限元子空间 $V_h$	37
4.3.4	刚度矩阵与负载向量计算	37
4.4	本章总结	38
<b>第 5 章</b>	<b>协调有限元的误差分析</b>	<b>39</b>
5.1	误差估计的整体流程	39
5.2	Sobolev 空间上的插值误差估计	40
5.2.1	仿射等价元之间的范数关系	40
5.2.2	单元上的插值误差估计	41
5.3	多边形区域上二阶问题的分析	42
5.3.1	解的 $H^1$ 模误差估计	43
5.4	$L^2$ 模与负模的估计	43
5.4.1	$L^2$ 模估计	43
5.4.2	负模估计	45
5.5	逆不等式	45
5.5.1	单元上的逆不等式	45
5.5.2	逆不等式	46
5.6	本章小结	48
<b>第 6 章</b>	<b>抛物型方程的有限元求解与分析</b>	<b>49</b>
6.1	空间离散 (半离散)	49
6.1.1	离散过程	49
6.1.2	半离散解的误差	50
6.2	全离散与误差估计	53
6.2.1	Euler-Galerkin 方法	53

# 第 1 章 一维有限元方法

## 1.1 为什么需要有限元方法？

此前在《微分方程数值解》课程中，我们已经学习了有限差分法和有限体积法。这两种方法有不少优点：首先，比较直观，只要知道如何利用差分近似导数即可得到对应的差分公式；其次，在一些情形下，有限差分法和有限体积法可以实现较高的计算精度。

但是，这两种算法有一些明显的缺陷。

- 算法稳定性的分析比较复杂。
- 处理不规则区域的问题时较为麻烦，需要多次利用插值近似。
- 只是求解离散格点的近似点值/离散网格的近似积分平均值，未能给出函数整体的近似。

为此，基于函数逼近论的**有限元方法**被提出。该算法能弥补有限差分法的一些明显缺陷，目前是最主流的数值算法之一。

## 1.2 从一维边值问题说起

考虑如下例子：

$$\begin{cases} -u'' + u = f(x), x \in (0, 1) \\ u(0) = u(1) = 0. \end{cases} \quad (1.1)$$

类似于“偏微分方程”课程中对弱解的讨论方式，在(1.1)两边同时乘某个函数  $v$  并在  $[0, 1]$  上积分，得到如下形式：

$$\int_0^1 (-u'' + u)v dx = \int_0^1 f v dx. \quad (1.2)$$

定义函数空间  $V$  如下：

$$V := \left\{ v \mid v(0) = v(1) = 0, \int_0^1 ((v')^2 + v^2) dx < \infty \right\}. \quad (1.3)$$

如果函数  $v \in V$ ，利用分部积分法，(1.1)可以转化为以下问题：

**例题 1.1** 记  $a(u, v) = \int_0^1 (u'v' + uv) dx$ ,  $h(v) = \int_0^1 f v dx$ ,  $v \in V$ . 求  $u \in V$ ，使得  $a(u, v) = h(v) \forall v \in V$ .

下面的定理说明了该问题可以转化为一个优化问题：

### 定理 1.1

记泛函  $J(v) := \frac{1}{2}a(v, v) - h(v)$ ，问题 1.1 与最小化  $J(v)$  的优化问题等价。即：如果  $a(u, v) = h(v) \forall v \in V$ ，那么  $J(u) \leq J(v) \forall v \in V$ 。

**证明**  $\Rightarrow$ :

$$\begin{aligned} J(v) - J(u) &= \frac{1}{2}a(v, v) - h(v) - \frac{1}{2}a(u, u) + h(u) \\ &= \frac{1}{2}(a(v, v) - a(u, u) - 2h(v - u)) \\ &= \frac{1}{2}(a(v, v) - a(u, u) - 2a(u, v - u)) \\ &= \frac{1}{2}(a(v, v) + a(u, u) - 2a(v, u)) \\ &= \frac{1}{2}(a(v - u, v - u)) \geq 0. \end{aligned} \quad (1.4)$$

由(1.4)可得，如果  $a(u, v) = h(v)$ ，那么  $J(u) \leq J(v)$ 。



$\Leftarrow: \forall v \in V, t \in \mathbb{R}$ , 有  $J(u+tv) \geq J(u)$ 。我们定义函数  $g(t) := J(u+tv)$ , 根据上面的讨论可知:  $g'(0) = 0$ 。另一方面, 计算  $g(t)$  的表达式, 有:

$$\begin{aligned} g(t) &= J(u+tv) \\ &= \frac{1}{2} \int_0^1 ((u' + tv')^2 + (u+tv)^2) dx - \int_0^1 f(u+tv) dx \end{aligned} \quad (1.5)$$

对(1.5)求一阶导数, 可得:

$$g'(0) = \int_0^1 (u'v' + uv) dx - \int_0^1 f v dx. \quad (1.6)$$

根据  $g$  的一阶条件, 可得  $a(u, v) = h(v)$ 。又由于  $v$  的任意性, 结论得证。

如此, 我们把一个解微分方程的问题, 利用1.1和1.2转化为了一个变分问题。由于  $V$  是一个无穷维空间, 我们不能期望利用算法给出这个变分问题的精确解, 但我们可以考虑对空间  $V$  进行有限维近似, 并在有限维空间上近似求解这个变分问题。

## 1.3 有限元思想的导出

接下来, 根据上一节的思路, 我们继续问题(1.1)的近似求解。根据上面的分析, 我们的数值算法需要解决两个问题:

- 如何对函数空间  $V$  进行有限维近似?
- 在进行有限维近似之后, 如何在有限维空间中对变分问题进行求解?

首先, 我们考虑第二个问题。根据上面的讨论, 近似求解变分问题有两种不同的思路, 分别对应的是 **Galerkin 近似方法**和 **Ritz 近似方法**。

### 1.3.1 Galerkin 近似

假设已经给出有限维子空间  $V_N \leq V$ , Galerkin 近似的目的是求解  $u_N \in V_N$  使得

$$a(u_N, v_N) = h(v_N) \quad (1.7)$$

对所有  $v_N \in V_N$  都成立。

由于  $V_N$  是有限维的空间, 我们可以找到这个空间中的一组基函数  $\{\phi_1, \dots, \phi_N\}$ , 注意到  $a(u, v)$  是对称双线性函数, 设  $u_N = \alpha_1 \phi_1(x) + \dots + \alpha_N \phi_N(x)$ ,  $v_N(x) = \phi_i(x)$ , 代入(1.7), 可得一个线性方程组:

$$\begin{bmatrix} a(\phi_1, \phi_1) & a(\phi_1, \phi_2) & \cdots & a(\phi_1, \phi_N) \\ a(\phi_2, \phi_1) & a(\phi_2, \phi_2) & \cdots & a(\phi_2, \phi_N) \\ \vdots & \vdots & \ddots & \vdots \\ a(\phi_N, \phi_1) & a(\phi_N, \phi_2) & \cdots & a(\phi_N, \phi_N) \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_N \end{bmatrix} = \begin{bmatrix} h(\phi_1) \\ h(\phi_2) \\ \vdots \\ h(\phi_N) \end{bmatrix} \quad (1.8)$$

解这个线性方程组, 得到系数向量, 即可给出该方程的近似解。

### 1.3.2 Ritz 方法

同样, 假设有限维子空间  $V_N \leq V$  已经给出, **Ritz 方法**的思路是求解有关  $J(u)$  的优化问题, 即: 求  $u_N \in V_N$ , 使得

$$J(u_N) \leq J(v) \forall v \in V_N. \quad (1.9)$$

这里  $J(v) := \frac{1}{2}a(v, v) - h(v)$ 。

给出这个问题之后, 我们在有限维空间中, 利用最优化解法求解该问题。

这两个思路都是建立在有限维子空间  $V_N$  已经给出的前提下的。但这个有限维空间如何构造?

一个很容易想到的思路是利用  $v(0) = v(1)$  这一性质，构造三角函数系作为基底。这种选取思路对于问题(1.1)而言当然是极好的，三角函数系的正交性也使得(1.8)中的系数矩阵变得相当简单易求解。但这个方案的扩展性并不强，如果扩展到二维平面上，乃至更高维度的椭圆偏微分方程，就很难找到像这样全局定义的基函数。如果问题区域非规则或是存在不同方程的耦合，则更是如此。

有限元方法由此引出。

## 1.4 有限元方法

在上面两种思路的基础上，我们需要一个方便推广的构建有限维子空间的方法。

多项式函数空间是最容易表示的函数空间，因此这是我们的首选。但全局定义的多项式很难保证其符合边界条件。于是，借助样条插值的思想，我们转为考虑分段多项式空间。

### 1.4.1 线性有限元空间

我们先针对问题(1.1)，考虑最简单的近似形式—分段线性近似。在这种情形下，有限维子空间  $V_h$  由  $(0, 1)$  上的分段线性函数表示。

#### 定义 1.1 (线性有限元空间)

线性有限元空间的定义为：

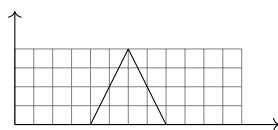
$$V_h := \{v_h \in C(0, 1) : v_h(0) = v_h(1) = 0, v_h|_{[x_i, x_{i+1}]} \in \mathbb{P}_1\}. \quad (1.10)$$

其中  $\{x_i\}_{i=0}^n$  为  $[0, 1]$  上给定的互异节点,  $x_0 = 0$ ,  $x_n = 1$ 。

首先需要讨论的是，空间  $V_h$  的维数和基底。空间(1.10)的形式很容易联想到数值分析课程中学习过的 **B-样条空间**。特别地，一维 B-样条基函数为所谓的“hat-function”，定义如下：

$$\phi_i(x) = \begin{cases} \frac{x - x_{i-1}}{x_i - x_{i-1}}, & x_{i-1} \leq x \leq x_i \\ \frac{x_{i+1} - x}{x_{i+1} - x_i}, & x_i \leq x \leq x_{i+1} \\ 0, & \text{otherwise} \end{cases} \quad (1.11)$$

图 1.1: hat 函数的示意图



容易验证，如此定义的  $(\phi_1, \dots, \phi_{n-1})$  构成了空间  $V_h$  的一组基。从而可得  $\dim(V_h) = n - 1$ 。

### 1.4.2 刚度矩阵与负载

此处我们采用 Galerkin 近似的思路，下面我们需要讨论的是方程组(1.8)的导出与求解。

#### 定义 1.2 (刚度矩阵)

线性有限元空间 1.1 的刚度矩阵定义为：

$$A = (a(\phi_i, \phi_j))_{i,j=1}^{n-1}. \quad (1.12)$$

其中  $\phi_i$  为(1.11)中定义的基函数。

**定义 1.3 (负载)**

方程(1.1)关于线性有限元空间1.1的负载向量定义为:

$$\mathbf{b} := \left( \int_0^1 f(x) \phi_i(x) dx \right)_{i=1}^{n-1}. \quad (1.13)$$

下面讨论(1.1)负载的计算。由于我们是用一次样条多项式近似真实解的,在选取数值积分公式的时候只需要使用一阶代数精度的公式即可不损失计算精度。因此此处采用复化梯形公式进行负载的近似计算。

记  $h_i = x_i - x_{i-1}$ , 负载的近似计算公式如下:

$$b_i = \int_0^1 f(x) \phi_i(x) dx = \int_{x_{i-1}}^{x_{i+1}} f(x) \phi_i(x) \approx \frac{1}{2} (h_i + h_{i+1}) f(x_i). \quad (1.14)$$

对于刚度矩阵的计算,虽然基函数  $\phi_i$  在整个区间  $(0, 1)$  上并不可导,但不可导点为零测集,不会对积分的计算产生影响。将基函数  $\phi_i$  的表达式代入刚度矩阵,计算得:

$$\begin{cases} a_{ij} = 0, |i - j| > 1 \\ a_{ii} = \frac{1}{h_i} + \frac{1}{h_{i+1}} + \frac{1}{3} (h_i + h_{i+1}) \\ a_{i-1,i} = -\frac{1}{h_i} + \frac{1}{6} h_i \\ a_{i,i+1} = -\frac{1}{h_{i+1}} + \frac{1}{6} h_{i+1} \end{cases} \quad (1.15)$$

由(1.15)可知,刚度矩阵是一个三对角矩阵,因此如果采用传统的矩阵存储方式和求解方式可能会造成计算资源的浪费。下面给出一种采用“局部矩阵”的形式存储(1.15)的技巧。

记负载矩阵  $A = \sum_{k=1}^n A^{(k)}$ , 其中  $A^{(k)} := (a_{ij}^{(k)})$ ,  $a_{ij}^{(k)} := \int_{x_{k-1}}^{x_k} [\phi_i'(x) \phi_j'(x) + \phi_i(x) \phi_j(x)] dx$ , 由定义可得,  $a_{ij}^{(k)}$  只在  $i, j \in \{k-1, k\}$  时取非零值。所以每个  $A^{(k)}$  都可以存储为一个  $2 \times 2$  的矩阵,而整体的负载矩阵  $A$  可以视为把所有的  $A^{(k)}$  “装配”起来的结果。在求解最终的线性方程组的时候,由于矩阵  $A$  是一个三对角对称正定矩阵,我们可以采用一系列数值代数算法加速方程组的求解。

## 1.5 误差分析

在本节中我们针对求解(1.1)的线性有限元近似算法进行误差分析。在叙述之前,为方便起见,先给出两个记号。

**注**  $\|f\|_0$ : 表示  $L^2$  范数,  $\|f\|_0 := \sqrt{\int_0^1 f^2 dx}$ .  $\|f\|_1$ : 表示  $W^{1,2}$  范数,  $\|f\|_1 := \sqrt{\int_0^1 ((f')^2 + f^2) dx}$ .

回顾一下有限元逼近问题的提法:

**例题 1.2** 记泛函  $a(u, v) = \int_0^1 (uv + u'v') dx$ ,  $h(v) = \int_0^1 f v dx$ ,  $V = \{v(x) : v(0) = v(1) = 0, \int_0^1 ((v')^2 + v^2) dx < \infty\}$ , 原问题是求解  $u \in V$  使得  $a(u, v) = h(v) \forall v \in V$ , 而在线性有限元空间上的逼近问题则是,在线性有限元空间  $V_h \leq V$  上求解  $u_h \in V$  使得  $a(u_h, v_h) = h(v_h) \forall v_h \in V_h$ .

在进行讨论前,首先给出一个重要的引理。该引理保证了在  $W^{1,2}$  范数的意义下,  $u_h \in V_h$  是  $u \in V$  在子空间  $V_h$  下的最优逼近。

**引理 1.1 (Cea 引理)**

设  $u$  和  $u_h$  是1.2的解, 则有:

$$\|u - u_h\|_1 \leq \inf_{v_h \in V_h} \|u - v_h\|_1. \quad (1.16)$$

**证明** 由于  $u$  满足:  $\forall v \in V, a(u, v) = h(v) \Rightarrow \forall v_h \in V_h, a(u, v_h) = h(v_h)$ . 又有:  $u_h$  满足  $a(u_h, v_h) = h(v_h) \forall v_h \in V_h$ , 两式相减, 得:

$$a(u - u_h, v_h) = 0 \forall v_h \in V_h. \quad (1.17)$$

(1.17)表示了向量  $u - u_h$  和空间  $V_h$  的正交性。下面讨论  $\|u - u_h\|_1^2$  的估计。

$$\begin{aligned}
\|u - u_h\|_1^2 &= a(u - u_h, u - u_h) \\
&= a(u - u_h, u - v_h) \\
&= \int_0^1 ((u' - u'_h)(u' - v'_h) + (u - u_h)(u - v_h)) dx \\
&\leq \left( \int_0^1 (u' - u'_h)^2 dx \right)^{\frac{1}{2}} \left( \int_0^1 (u' - v'_h)^2 dx \right)^{\frac{1}{2}} + \left( \int_0^1 (u - u_h)^2 dx \right)^{\frac{1}{2}} \left( \int_0^1 (u - v_h)^2 dx \right)^{\frac{1}{2}} \\
&\leq \left\{ \int_0^1 ((u' - u'_h)^2 + (u - u_h)^2) dx \right\}^{\frac{1}{2}} \left\{ \int_0^1 ((u' - v'_h)^2 + (u - v_h)^2) dx \right\}^{\frac{1}{2}} \\
&= \|u - u_h\|_1 \|u - v_h\|_1 \Rightarrow \|u - u_h\|_1 \leq \|u - v_h\|_1.
\end{aligned} \tag{1.18}$$

该引理在  $\|\cdot\|$  的意义下, 把有限元解的误差转化为计算解函数的插值误差 (这个插值误差可以作为一个上界)。目前解函数用  $u$  表示, 假设  $u$  在节点  $\{x_i\}_{i=0}^n$  上的分段线性插值多项式记为  $u_I$ , 考虑  $\|u - u_I\|_1$  的估计。关于分段线性插值, 我们有下面的误差估计公式:

### 定理 1.2

$$\begin{cases} \max_{[0,1]} |u(x) - u_I(x)| \leq \frac{1}{8} h^2 \max_{[0,1]} |u''(x)| \\ \max_{[0,1]} |u'(x) - u'_I(x)| \leq h \max_{[0,1]} |u''(x)| \end{cases} \tag{1.19}$$

其中  $h = \max_{1 \leq i \leq n} |x_i - x_{i-1}|$ 。

**证明** 见任何一本数值分析教材。

根据1.2和1.1, 可以针对  $\|u - u_h\|_1$  进行估计。

$$\|u - u_h\|_1^2 \leq \|u - u_I\|_1^2 = \int_0^1 ((u - u_I)^2 + (u' - u'_I)^2) dx \leq \left( \frac{1}{64} h^4 + h^2 \right) \max_{[0,1]} |u''(x)|^2. \tag{1.20}$$

从而, 我们导出了下面的估计定理:

### 定理 1.3

$$\|u - u_h\|_1 \leq 2h \max_{[0,1]} |u''|. \tag{1.21}$$

1.3给出了利用无穷范数控制  $\|\cdot\|_1$  的方法。但由于二阶求导可能会大幅度提升  $u$  的振幅,  $\max_{[0,1]} |u''|$  可能会大的令人难以接受, 所以我们最好能找到更合理的范数估计。事实上, 如果右端项取  $u''$  的  $L^2$ -norm, 我们可以有下面的估计:

### 定理 1.4

$$\|u - u_h\|_1 \leq Ch \|u''\|_0 \leq Ch \|f\|_0. \tag{1.22}$$

**证明** 在介绍1.4的证明前, 我们需要先给出一些命题。



## 命题 1.1

如果  $w \in C^2$  是下面的微分方程的解:

$$\begin{cases} -w'' + w = g(x), x \in (0, 1), \\ w(0) = w(1) = 0. \end{cases} \quad (1.23)$$

则有

$$\int_0^1 (w'')^2 dx \leq \int_0^1 g^2 dx. \quad (1.24)$$

**证明** 在(1.23)的第(1)式左右两边同乘  $w''$ , 并同时在区间  $[0, 1]$  上积分:

$$\begin{aligned} \int_0^1 (-w'' + w)w'' dx &= \int_0^1 g(x)w''(x) dx \\ \Rightarrow \int_0^1 (w'')^2 dx + \int_0^1 (w')^2 dx &= - \int_0^1 g(x)w''(x) dx \\ \Rightarrow \int_0^1 (w'')^2 dx &\leq \int_0^1 g(-w'') dx \leq \|g\|_0 \|w''\|_0 \\ \Rightarrow \|w''\|_0 &\leq \|g\|_0 \end{aligned} \quad (1.25)$$

**注** 该结论的证明思路正是偏微分方程理论中非常常见的“能量模估计”。

## 命题 1.2

记

$$G_i(x, t) = \begin{cases} (t - x_i)(x - x_{i+1}), & x > t, \\ (x - x_i)(t - x_{i+1}), & x \leq t. \end{cases} \quad (1.26)$$

那么在  $[x_i, x_{i+1}]$  上, 有下面这一等式成立:

$$(u(x) - u_I(x))(x_{i+1} - x_i) = \int_{x_i}^{x_{i+1}} G_i(x, t)u''(t)dt. \quad (1.27)$$

**证明** 记符号  $h_i := x_{i+1} - x_i$ , 计算右端项的积分, 有:

$$\begin{aligned} &\int_{x_i}^{x_{i+1}} G_i(x, t)u''(t)dt \\ &= (x - x_{i+1}) \int_{x_i}^x (t - x_i)u''(t)dt + (x - x_i) \int_x^{x_{i+1}} u''(t)(t - x_{i+1})dt \\ &= (x - x_{i+1}) \left[ \int_{x_i}^x (t - x_i)du'(t) \right] + (x - x_i) \left[ \int_x^{x_{i+1}} (t - x_i)du'(t) \right] \\ &= (x - x_{i+1}) \left[ (x - x_i)u'(x) - \int_{x_i}^x u'(t)dt \right] + (x - x_i) \left[ (x_{i+1} - x)u'(x) - \int_x^{x_{i+1}} u'(t)dt \right] \\ &= h_i u(x) + (x - x_{i+1})u(x_i) - (x - x_i)u(x_{i+1}) \\ &= h_i(u(x) - u_I(x)). \end{aligned} \quad (1.28)$$

下面利用 1.1 和 1.2 证明定理 1.4。首先讨论  $\|u - u_1\|_0$  的估计。由 1.2，我们有：

$$\begin{aligned}
 \int_0^1 [u(x) - u_I(x)]^2 dx &= \sum_{i=0}^{n-1} \int_{x_i}^{x_{i+1}} [u(x) - u_I(x)]^2 dx \\
 &= \sum_{i=0}^{n-1} \int_{x_i}^{x_{i+1}} \left[ \frac{1}{x_{i+1} - x_i} \int_{x_i}^{x_{i+1}} G_i(x, t) u''(t) dt \right]^2 dx. \\
 &\leq \sum_{i=0}^{n-1} (x_{i+1} - x_i) \left[ \frac{1}{x_{i+1} - x_i} \int_{x_i}^{x_{i+1}} (t - x_i)(x_{i+1} - t) |u''(t)| dt \right]^2 \\
 &\leq \sum_{i=0}^{n-1} (x_{i+1} - x_i) \left[ \frac{x_{i+1} - x_i}{4} \int_{x_i}^{x_{i+1}} |u''(t)| dt \right]^2 \\
 &\leq \sum_{i=0}^{n-1} \frac{(x_{i+1} - x_i)^3}{16} \left[ \int_{x_i}^{x_{i+1}} |u''(t)| dt \right]^2 \\
 &\leq \sum_{i=0}^{n-1} \frac{(x_{i+1} - x_i)^4}{16} \int_{x_i}^{x_{i+1}} |u''(t)|^2 dt \\
 &\leq \frac{h^4}{16} \int_0^1 |u''(t)|^2 dt.
 \end{aligned} \tag{1.29}$$

(1.29) 中，第三行和第四行的不等式都源于对函数  $G_i(x, t)$  的估计，第六行的不等式则源于柯西不等式。

由 1.2 可知，在  $[x_i, x_{i+1}]$  上，我们有：

$$\begin{aligned}
 |u'(x) - u'_I(x)| &= \left| \frac{1}{x_{i+1} - x_i} \left[ \int_{x_i}^x \frac{(t - x_i) u''(t)}{x_{i+1} - x_i} dt + \int_x^{x_{i+1}} \frac{(t - x_{i+1}) u''(t)}{x_{i+1} - x_i} dt \right] \right| \\
 &\leq C(x_{i+1} - x_i)^{\frac{1}{2}} \left( \int_{x_i}^{x_{i+1}} |u''(t)|^2 dt \right)^{\frac{1}{2}}.
 \end{aligned} \tag{1.30}$$

(1.30) 中， $C$  为常数，该不等式由柯西不等式导出。借助该式，我们有：

$$\begin{aligned}
 \int_0^1 |u'(x) - u'_I(x)|^2 dx &= \sum_{i=0}^{n-1} \int_{x_i}^{x_{i+1}} [u'(x) - u'_I(x)]^2 dx \\
 &\leq \sum_{i=0}^{n-1} C^2 (x_{i+1} - x_i)^2 \int_{x_i}^{x_{i+1}} |u''(t)|^2 dt \\
 &\leq C^2 h^2 \int_0^1 |u''(t)|^2 dt.
 \end{aligned} \tag{1.31}$$

联立 (1.29) 和 (1.31)，结合引理 1.1 和命题 1.1，定理 1.4 得证。

**注** 对 1.4，可以用 Fourier 分析给出一个更优的估计。

**练习 1.1** 补充 (1.31) 的证明细节，并给出常数  $C$  的值。

利用定理 1.4，对于  $\|u - u_h\|_0$ ，我们可以给出更精细的估计：

#### 定理 1.5 (Aubin-Nitsche 技巧)

$$\|u - u_h\|_0 \leq Ch^2 \|u''\|_0. \tag{1.32}$$

**证明** 考虑辅助问题：

$$\begin{cases} -z'' + z = u - u_h, \\ z(0) = z(1) = 0. \end{cases} \tag{1.33}$$

其变分形式为：

$$a(z, v) = (u - u_h, v)_{L^2}, \quad \forall v \in V. \tag{1.34}$$

从而：

$$\begin{aligned}\|u - u_h\|_0 &= a(z, u - u_h) \\ &= a(z - \omega_h, u - u_h) \quad \forall \omega_h \in V_h\end{aligned}\quad (1.35)$$

取  $\omega_h$  为函数  $z$  的分段线性插值  $z_I$ ，那么：

$$a(z - z_I, u - u_h) \leq C\|z - z_I\|_1\|u - u_h\|_1 \leq Ch^2\|z''\|_0\|f\|_0. \quad (1.36)$$

由 (1.24)， $\|z''\|_0 \leq C\|u - u_h\|_0$ ，结论得证。

## 1.6 一阶非齐次 BVP

接下来我们把研究对象改为一般的非齐次两点边值问题。事实上，虽然方程看似复杂了些，但处理方法和之前基本一致。

考虑下面的例子：

$$\begin{cases} -(k(x)u')' + q(x)u = f(x), & x \in (0, 1), \\ u(0) = \alpha, \quad u(1) = \beta. \end{cases} \quad (1.37)$$

其中  $k(x) \in C^1$ ,  $k(x) \geq k_0 > 0$ ,  $q(x) \geq 0$ .

和齐次问题同样的道理，先考虑该问题的变分形式。在 (1.37) 两端同乘函数  $v$  并在  $[0, 1]$  上积分，有：

$$\begin{aligned}& \int_0^1 [-(k(x)u')'v(x) + q(x)u(x)v(x)]dx = \int_0^1 f(x)v(x)dx \\ \Rightarrow & \int_0^1 quvdx - \int_0^1 v(x)d(k(x)u') = \int_0^1 f(x)v(x)dx \\ \Rightarrow & \int_0^1 (quv + ku'v')dx + v(0)k(0)u'(0) - v(1)k(1)u'(1) = \int_0^1 f(x)v(x)dx.\end{aligned}\quad (1.38)$$

要保证 (1.38) 有意义，同样需要确保  $u, v \in V := W^{1,2}$ 。在此基础上，如果  $v(0) = v(1) = 1$ ，令  $a(u, v) := \int_0^1 (ku'v' + quv)dx$ ， $h(v) := \int_0^1 fvdv$ ，则有：

$$a(u, v) = h(v) \quad \forall v \in V, \quad v(0) = v(1) = 0. \quad (1.39)$$

其中  $u \in W := \{v : \int_0^1 ((v')^2 + v^2)dx < \infty, v(0) = \alpha, v(1) = \beta\}$ .

同样地，该问题也有与之相对应的极小化问题：

**例题 1.3** 求  $u \in W$  s.t.  $J(u) \leq J(v) \forall v \in W$ ，此处  $J(v) := \frac{1}{2}a(v, v) - f(v)$ 。

下一步则是构造对应的线性有限元空间。注意到，对于非齐次边值问题，函数集合  $W$  并不构成线性空间，所以我们构建的“线性有限元空间”实际上是一个由分段线性函数构成的子集。但与此同时，这个子集可以写成陪集 (coset) 的形式，所以我们依旧可以利用基底的线性组合来表示这个函数集合。

取  $[0, 1]$  区间的分划  $0 = x_0 < x_1 < \dots < x_{n-1} < x_n = 1$ ，由于  $u(0) = \alpha, u(1) = \beta$ ，记函数集合  $W_h := \{w \in C(0, 1) : w|_{[x_i, x_{i+1}]} \in \mathbb{P}_1, w(0) = \alpha, w(1) = \beta\}$ ，那么我们必定有  $u \in W_h$ 。 $V_h$  的定义同 1.10，设分段线性函数  $g(x)$  满足：

$$g(0) = \alpha, g(1) = \beta, \forall i \in [1, n-1], g(x_i) = 0. \quad (1.40)$$

那么我们有  $W_h = g + V_h$ 。

最后考虑有限元近似求解。此处  $\phi_i$  的定义沿用 (1.11) 的表达式，记  $u_h(x) := g(x) + \sum_{i=1}^{n-1} \alpha_i \phi_i(x)$ ， $v_i(x) := \phi_i(x)$ ，代入 (1.39)，得：

$$a(g, \phi_i) + \sum_{j=1}^{n-1} \alpha_j a(\phi_j, \phi_i) = h(\phi_i). \quad (1.41)$$

由  $a(\cdot, \cdot)$  的对称性，我们把 (1.41) 转化为矩阵形式  $(Ax = b)$ 。其中刚度矩阵  $A := a(\phi_i, \phi_j)_{i,j=1}^{n-1}$ ，负载  $b := (h(\phi_i) - a(g, \phi_i))^t$ 。可以证明，刚度矩阵仍旧为对称正定三对角矩阵。但刚度矩阵和负载向量的计算则需要通过数值积

分完成。当然，在编程计算刚度矩阵时，利用装配技巧同样能节省存储空间并增加程序的可读性。

与齐次方程同样，非齐次方程变分问题的有限元解也有与之相对应的 Cea 引理和各种范数估计，且证明方法近乎雷同，故此处省略不表，留作习题。

**注 (1.37)** 中  $q(x) \geq 0$  的条件似乎应当改为  $q(x) \geq q_0 > 0$ ，否则 Cea 引理的证明会出问题。举个例子，如果  $q(x) \equiv 0$ ，引出的就不再是对  $\|u - u_h\|_1$  的估计，而是  $\|u' - u'_h\|_0$  的估计。

## 1.7 其他边值问题

### 1.7.1 Neumann 边值问题

问题描述：

$$\begin{cases} -(ku')' + qu = f \\ -k(0)u'(0) = \gamma_1, k(1)u'(1) = \gamma_2 \end{cases} \quad (1.42)$$

变分形式：

$$\int_0^1 (ku'v' + quv)dx = \int_0^1 fvdx + \gamma_2v(1) + \gamma_1v(0). \quad (1.43)$$

### 1.7.2 Robin 边值问题

问题描述：

$$\begin{cases} -(ku')' + qu = f \\ -k(0)u'(0) = -\beta_1u(0) + \gamma_1, k(1)u'(1) = -\beta_2u(1) + \gamma_2 \end{cases} \quad (1.44)$$

变分形式：

$$\int_0^1 (ku'v' + quv)dx = \int_0^1 fvdx + \gamma_1v(0) + \gamma_2v(1) - \beta_1u(0)v(0) - \beta_2u(1)v(1). \quad (1.45)$$

## 1.8 本章总结

有限元的基本逻辑：

- 两边相乘试验函数，同时积分。
- 给出变分形式。
- 构造函数空间的有限维子空间。
- 在有限维子空间上求解变分形式。
- 误差估计。

由于有限元方法的输出并非解函数在某些点上的取值或者控制体上的积分平均值，而是整体函数，其误差范数定义和有限差/有限体方法有一些区别。本章中常用的范数是  $L^2$  范数和  $W^{1,2}$  范数。同样，由于有限元方法是在函数空间中求解优化问题或者线性方程，函数逼近论在有限元方法的研究中居于重要地位。

## 第2章 变分原理

上一章中，我们用一个具体的例子说明了有限元的大体思路和操作步骤。本章则会着眼于变分问题的提法，给出一般情形下的变分问题（以及其等价形式），并证明其解的存在唯一性。最后则是给出几个具体的算例以及它们各自的变分形式。

### 2.1 变分问题以及等价形式

在上一章定理1.2中，我们针对一个特殊问题，提出了两种变分问题，并证明了这两种问题的等价性。本节中，我们将着眼于更一般的问题，并对一般的问题尝试说明两者的等价性，以及其解的存在唯一性。在此之前，先要做一些概念和定理上的准备。

#### 2.1.1 预备知识

首先，我们讨论双线性函数。

##### 定义 2.1 (双线性函数)

一个双线性函数  $a(\cdot, \cdot) : V \times V \rightarrow \mathbb{R}$  需要满足下面的条件：

- $a(k_1u + k_2v, w) = k_1a(u, w) + k_2a(v, w)$ .
- $a(u, k_1v + k_2w) = k_1a(u, v) + k_2a(u, w)$ .

对任意  $u, v \in V, k_1, k_2 \in \mathbb{R}$  均成立。特别地，假设  $V$  是以  $\|\cdot\|$  为范数的赋范线性空间，如果对任意  $u, v \in V$ ，存在  $M > 0$  使得：

$$|a(u, v)| \leq M \|u\| \|v\| \quad (2.1)$$

总成立，那么我们称双线性型  $a$  为有界的或连续的。如果  $\forall u, v \in V$ ，均有：

$$a(u, v) = a(v, u), \quad (2.2)$$

那么双线性型  $a$  为对称的。如果满足下面的性质：

$$\exists \alpha > 0, \forall v \in V, \alpha \|v\|^2 \leq a(v, v), \quad (2.3)$$

则称双线性型  $a$  是 **V-椭圆**的。

特别地，第一章所述

$$a(u, v) = \int_0^1 (uv + u'v') dx, \quad (2.4)$$

是一个有界且 V-椭圆的双线性函数。双线性函数的性质在后续对变分问题的讨论中相当重要。

为方便后续内容的展开，本节中将给出一些泛函分析中的重要结论。

##### 命题 2.1 (Riesz 表示定理)

设  $H$  是 Hilbert 空间，则任意  $f \in H^*$ ，存在唯一  $x_f \in H$ ，使得：

$$f(y) = (x_f, y) \quad \forall y \in H, \|f\|_{H^*} = \|x_f\|_H. \quad (2.5)$$

**证明** 存在性：考查闭子空间  $\ker f \leq H$ 。

如果  $\ker f = H$ ，这意味着  $f(x) = 0 \forall x \in H$ ，取  $x_f = 0$  则有  $f(y) = 0 = (x_f, y), \forall y \in H$ ，且  $\|f\|_{H^*} = \|x_f\|_H = 0$ 。

如果  $\ker f \neq H$ ，由于  $f \in H^*$ ，由第一同构定理可知  $\text{codim } \ker f = 1$ 。记  $(\ker f)^\perp$  的一组基为  $\{y_0\}$ ，则由正交分解定理， $\forall y \in H, y = ky_0 + \tilde{y}$ ，其中  $\tilde{y} \in \ker f$ 。



此时,  $f(y) = kf(y_0)$ , 记  $x_f = \lambda y_0$ ,  $(x_f, y) = \lambda k \|y_0\|_H^2$ 。从而: 取  $x_f = \frac{f(y_0)}{\|y_0\|_H^2} y_0$ , 等式(2.5)成立。且有:

$$\|x_f\|_H = \frac{|f(y_0)|}{\|y_0\|_H} = \|f\|_{H^*}. \quad (2.6)$$

唯一性: 如果存在  $w_f$  s.t.  $f(y) = (w_f, y)$ , 那么  $\forall y \in H$ , 我们有

$$(x_f - w_f, y) = 0, \forall y \in H. \quad (2.7)$$

取  $y = x_f - w_f$ , 由内积的定义即得  $x_f - w_f = 0$ , 即  $x_f = w_f$ 。

### 命题 2.2 (闭凸子集投影的存在唯一性)

设  $V$  是一个 Hilbert 空间, 如果  $U$  是  $V$  的闭凸子集, 对于  $v \in V$ , 存在唯一  $u_0 \in U$  使得:

$$\langle v - u_0, v - u_0 \rangle = \min_{u \in U} \langle v - u, v - u \rangle. \quad (2.8)$$

**证明** 记范数  $\|u\| := \langle u, u \rangle^{\frac{1}{2}}$ ,  $U$  是闭凸子集, 由 F-Riesz 定理,  $\|v - u\|_U$  在  $u \in U$  中存在下确界  $d$ 。由此, 存在序列  $\{u_n\} \subset U, \forall \epsilon > 0, \exists N$ , 当  $n > N$  时有:

$$d^2 \leq d_n^2 := \|v - u_n\|^2 < d^2 + \epsilon \quad (2.9)$$

往证序列  $\{u_n\}$  存在极限, 且极限在集合  $U$  内部。由平行四边形公式,  $\forall n, m > N$ , 有:

$$\|u_m + u_n - 2v\|^2 + \|u_m - u_n\|^2 = 2(d_m^2 + d_n^2). \quad (2.10)$$

根据等式(2.10), 我们有:

$$\begin{aligned} \|u_m - u_n\|^2 &= 2(d_m^2 + d_n^2) - \|u_m + u_n - 2v\|^2 \\ &< 4d^2 + 4\epsilon - 4\left(\left\|\frac{u_m + u_n}{2} - v\right\|^2\right) \\ &\leq 4\epsilon. \end{aligned} \quad (2.11)$$

最后一个不等号源于  $U$  为凸集, 从而  $\frac{u_m + u_n}{2} \in U$ , 这意味着  $\left\|\frac{u_m + u_n}{2} - v\right\| \geq d$ 。

(2.11)保证了  $\{u_n\}$  是 Cauchy 列, 即该序列收敛。又  $U$  是闭的, 该序列收敛于  $u_0 \in U$ 。存在性得证。

唯一性证明完全同理, 留作习题。

### 命题 2.3 (压缩映射原理)

设  $V$  是 Banach 空间, 连续映射  $T$  满足:

$$\|Tv_1 - Tv_2\| \leq L\|v_1 - v_2\|, \forall v_1, v_2 \in V, \quad (2.12)$$

其中  $0 < L < 1$  为常数, 则存在唯一的  $u \in V$  使得  $u = Tu$ 。

该定理证明从略, 留作习题。

## 2.1.2 存在唯一性

本节中讨论抽象意义下的优化问题。

### 定义 2.2 (优化问题)

给定赋范线性空间  $V$  以及其上的有界双线性函数  $a$ ,  $f \in V^*$ , 对于  $U \subset V$ , 寻找  $u \in U$ , 使得:

$$J(u) = \inf_{v \in U} J(v), \quad J(v) := \frac{1}{2}a(v, v) - f(v) \quad (2.13)$$

2.2中叙述的优化问题和第一章对应的问题相比, 极大放宽了对  $U$  的要求, 这也使得第一章对该问题的讨论不一定适用于此。为此, 我们需要给出问题2.2解存在唯一的条件。

**定理 2.1**

如果2.2满足下列额外条件:

- $V$  完备,
- $U$  是  $V$  的闭凸子集,
- 双线性函数  $a(u, v)$  是对称且  $V$ -椭圆的,

那么优化问题2.2存在唯一解。



**证明** 由于  $a(u, v)$  是连续对称椭圆双线性型, 可知  $a(u, v)$  构成空间  $V$  上的一个内积, 导出的范数记作  $\|\cdot\|$ 。由 Riesz 表示定理, 可以把  $f$  用  $a$  表示, 即:

$$\forall v \in V, f(v) = a(\sigma_f, v). \quad (2.14)$$

由(2.14), 代入(2.13)可得:

$$\begin{aligned} J(v) &= \frac{1}{2}a(v, v) - a(\sigma_f, v) \\ &= \frac{1}{2}a(v - \sigma_f, v - \sigma_f) - \frac{1}{2}a(\sigma_f, \sigma_f). \end{aligned} \quad (2.15)$$

由此, 该极小化问题转化为在  $U$  上最小化  $\|v - \sigma_f\|$ 。由于  $\sigma_f \in V$ , 而  $U$  是  $V$  的闭凸子集, 借助闭凸子集投影的存在唯一性, 当且仅当  $v$  是  $\sigma_f$  在  $U$  上的投影时, 优化问题2.2取极小值。由此便证明了该问题解的存在唯一性。

**2.1.3 变分问题的等价形式**

在第一章中, 我们说最小化  $J(v)$  的问题等价于求解泛函方程  $a(u, v) = f(v)$ 。本节中, 我们将进一步探究这两个问题之间的关系。

**定理 2.2**

1. 如果  $u$  是问题2.2的解, 当且仅当

$$\forall v \in U, a(u, v - u) \geq f(v - u). \quad (2.16)$$

2. 特别地, 如果  $U$  是以 0 为顶点的凸锥, 那么:

$$\begin{cases} \forall v \in U, a(u, v - u) \geq f(v - u), \\ a(u, u) = f(u). \end{cases} \quad (2.17)$$

3. 特别地, 如果  $U$  是  $V$  的闭子空间, 那么:

$$\forall v \in U, a(u, v) = f(v). \quad (2.18)$$



**证明** 由 Riesz 表示定理,  $f(v - u) = a(\sigma_f, v - u)$ 。

先证明问题 1 的充分性部分。任取  $v \in U$ , 可得:

$$\begin{aligned} a(v - \sigma_f, v - \sigma_f) &= a(v - u + u - \sigma_f, v - u + u - \sigma_f) \\ &= a(v - u, v - u) + 2a(v - u, u - \sigma_f) + a(u - \sigma_f, u - \sigma_f) \\ &\geq a(u - \sigma_f, u - \sigma_f). \end{aligned} \quad (2.19)$$

即, 在(2.16)成立时,  $u$  是2.2的解。

再证明其必要性部分。由于  $U$  是凸集,  $\forall t \in (0, 1)$ ,  $tv + (1 - t)u \in U$ 。如果  $u$  是2.2的解, 那么:

$$a(\sigma_f - tv - (1 - t)u, \sigma_f - tv - (1 - t)u) \geq a(\sigma_f - u, \sigma_f - u). \quad (2.20)$$

化简, 有:

$$-2ta(\sigma_f - u, v - u) + t^2a(v - u, v - u) \geq 0. \quad (2.21)$$

取  $t \rightarrow 0$  即可证(2.16)成立。

对于问题 2, 由凸锥的性质,  $u+v \in U$ , 在(2.16)中取  $v_1 := u+v$ , 可得  $a(u, v) \geq f(v)$ 。又 0 在凸锥顶点, 取  $v=0$ , 可得  $a(u, u) \leq f(u)$ 。由此,  $a(u, u) = f(u)$ 。

对于问题 3, 根据刚刚对于凸锥的推理,  $a(u, v) \geq f(v)$ 。根据闭子空间的性质,  $-v \in U$ , 将  $\tilde{v} := -v$  可得  $a(u, -v) \geq f(-v)$ , 即  $a(u, v) \leq f(v)$ 。由此,  $a(u, v) = f(v)$ 。

## 2.2 Lax-Milgram 引理

### 定理 2.3 (Lax-Milgram 引理)

设  $V$  是一个 Hilbert 空间,  $a(\cdot, \cdot) : V \times V \rightarrow \mathbb{R}$  是一个连续  $V$ -椭圆双线性型,  $f : V \rightarrow \mathbb{R}$  是连续线性泛函, 那么, 存在唯一  $u \in V$ , 使得:

$$a(u, v) = f(v) \quad (2.22)$$

对任意  $v \in V$  成立。

**注** 此处  $a(u, v)$  并没有对称性条件, 故不能诱导  $V$  上的内积。

**证明** 第一步: 将双线性型  $a(u, v)$  转化为  $V \rightarrow V'$  的映射。

定义  $Au(v) := a(u, v)$ , 则  $Au \in V'$ , 且

$$\|Au\| = \sup_{v \in V} \frac{|Au(v)|}{\|v\|} \leq M\|u\|. \quad (2.23)$$

由  $a$  的有界性, 即可直接推出  $Au$  的有界性。由此: 映射  $A : V \rightarrow V', u \mapsto Au$  为连续线性映射。(2.22)转换为  $Au(v) = f(v)$ 。但直接比较  $V'$  上的两个元素并不容易。

第二步: 将  $V'$  中两元素的比较转化为  $V$  中两元素的比较。

设  $\tau$  为  $V' \rightarrow V$  的 Riesz 表示映射, 那么  $Au = f \Leftrightarrow \tau Au = \tau f$ 。下面需要证明满足该等式的  $u$  存在唯一。关于存在唯一性的问题, 我们可以通过构造压缩映射求解。

第三步: 构造压缩映射。

记  $T : V \rightarrow V$ , 其定义为:

$$T(v) = v - \rho(\tau Av - \tau f). \quad (2.24)$$

$\rho$  为一个可以自行选定的参数。 $T$  的不动点即  $v = v - \rho(\tau Av - \tau f)$ , 可得  $\tau Av = \tau f$ 。下证  $T$  在  $\rho$  取适当值的时候是压缩映射。

记  $\varphi = v - w$ ,  $v \in V, w \in V$ , 那么

$$\begin{aligned} & \|Tv - Tw\|^2 \\ &= \|T\varphi\|^2 \\ &= \|\varphi - \rho(\tau A\varphi)\|^2 \\ &= \|\varphi\|^2 - 2\rho \langle \tau A\varphi, \varphi \rangle + \rho^2 \langle \tau A\varphi, \tau A\varphi \rangle \\ &\leq \|\varphi\|^2 - 2\rho\alpha\|\varphi\|^2 + \rho^2 M^2\|\varphi\|^2. \end{aligned} \quad (2.25)$$

最后一个不等号的依据是: 1.  $\tau$  是等距同构。2.  $A$  是有界线性算子。3.  $a(\varphi, \varphi)$  的椭圆性。事实上:

$$\langle \tau A\varphi, \varphi \rangle = (A\varphi)(\varphi) = a(\varphi, \varphi) \geq \alpha\|\varphi\|^2. \quad (2.26)$$

$$\langle \tau A\varphi, \tau A\varphi \rangle = (A\varphi)(\tau A\varphi) = a(\varphi, \tau A\varphi) \leq M\|\varphi\|\|\tau A\varphi\| \leq M^2\|\varphi\|^2. \quad (2.27)$$

由此, 只要  $\rho < \frac{2\alpha}{M^2}$ ,  $T$  即为压缩映射, 这意味着  $\tau Au = \tau f$  存在唯一解。

## 2.3 具体实例

**注** 此处内容可能用到一些 Sobolev 空间的相关知识，由于之后会专题讨论该内容，此处不过多赘述相关知识点。请读者翻阅后续的笔记内容，或参考任何一本泛函分析教材。

### 2.3.1 二阶椭圆方程

#### 定义 2.3 (齐次边界二阶椭圆问题)

设  $\Omega \subset \mathbb{R}^2$  是一个有界单连通区域，其边界为  $\partial\Omega$ ，考虑下面的 Poisson 方程：

$$\begin{cases} -\Delta u = f, x \in \Omega, \\ u = 0, x \in \partial\Omega. \end{cases} \quad (2.28)$$

其中  $f \in C(\Omega)$ 。

在(2.28)的第1式两边同乘测试函数  $v$  并在区域  $\Omega$  上积分，得：

$$-\int_{\Omega} \Delta u v dx = \int_{\Omega} f v dx \Rightarrow \int_{\Omega} \nabla u \cdot \nabla v dx - \int_{\partial\Omega} \frac{\partial u}{\partial n} v dx = \int_{\Omega} f v dx \quad (2.29)$$

取测试函数空间为：

$$v \in H_0^1(\Omega) := \{v | v, \nabla v \in L^2(\Omega), v|_{\partial\Omega} = 0\}. \quad (2.30)$$

那么(2.29)转化为

$$\int_{\Omega} \nabla u \cdot \nabla v dx = \int_{\Omega} f v dx, \forall v \in H_0^1(\Omega). \quad (2.31)$$

(2.31)称为问题(2.28)的弱形式。根据本章定理2.2，该问题等价于如下泛函极小化问题：

$$J(v) := \frac{1}{2} \int_{\Omega} \nabla v \cdot \nabla v dx - \int_{\Omega} f v dx, J(u) := \inf_{v \in V} J(v). \quad (2.32)$$

#### 定理 2.4

(2.28)的弱解存在唯一。

**证明** 取  $a(u, v) := \int_{\Omega} \nabla u \cdot \nabla v$ ,  $f(v) := \int_{\Omega} f v dx$ ，只需验证 Lax-Milgram 引理的两个条件。

$a(u, v)$  的双线性由积分的线性性质即可导出，其连续性则是 Cauchy-Schwarz 不等式的直接推论，下证其  $H_0^1(\Omega)$ -椭圆性。证明该结论前，先不加证明地给出 Friedrichs 不等式的叙述，该不等式的证明可以参考 Evans 的偏微分方程教材。

#### 命题 2.4 (Friedrichs 不等式)

如果  $\Omega$  是  $\mathbb{R}^n$  的有界单连通子区域，其直径为  $d$ ，设  $u \in H_0^1(\Omega)$ ，我们有：

$$\|u\|_{L^2(\Omega)} \leq d \|\nabla u\|_{L^2(\Omega)}. \quad (2.33)$$

由2.4，我们有： $\exists C_1 > 0$  s.t.

$$a(v, v) = \int_{\Omega} \nabla v \cdot \nabla v \geq C_1 \|v\|_{L^2(\Omega)}^2. \quad (2.34)$$

又由 Sobolev 空间范数的定义，

$$\|v\|_{H_0^1(\Omega)}^2 = \int_{\Omega} (v^2 + \nabla v \cdot \nabla v) dx. \quad (2.35)$$

联立上面两等式可得：

$$(1 + C_1) a(v, v) \geq C_1 \|v\|_{H_0^1(\Omega)}^2. \quad (2.36)$$

$a$  的椭圆性即得证。

关于泛函  $f$  的有界性, 证明如下:

$$\frac{|f(v)|}{\|v\|_{H_0^1}} = \frac{\int_{\Omega} f v}{\|v\|_{H_0^1}} \leq \frac{\|f\|_{L^2} \|v\|_{L^2}}{\|v\|_{H_0^1}} \leq \|f\|_{L^2} < \infty. \quad (2.37)$$

由 Lax-Milgram 引理, 可知问题(2.28)的弱解存在唯一。

**练习 2.1** 用同样的方法讨论 Poisson 方程非齐次 Dirichlet 边界问题的弱解。

#### 定义 2.4 (Neumann 边界二阶椭圆方程)

设  $\Omega \subset \mathbb{R}^2$  是一个有界单连通区域, 其边界为  $\Gamma$ ,  $b, f, g \in C(\Omega)$ , 考虑下面的方程:

$$\begin{cases} -\Delta u + bu = f, & x \in \Omega \\ \frac{\partial u}{\partial \nu} = g, & x \in \Gamma \end{cases} \quad (2.38)$$

取测试函数空间  $V = H^1(\Omega)$ , 定义双线性型

$$a(u, v) := \int_{\Omega} (\nabla u \cdot \nabla v + b(x)uv) dx. \quad (2.39)$$

右端的线性泛函

$$f(v) := \int_{\Omega} f v dx + \int_{\Gamma} g v ds. \quad (2.40)$$

这里  $b(x) \geq b_0 > 0$ ,  $f \in L^2(\Omega)$ ,  $g \in L^2(\Omega)$ 。则问题(2.38)的弱形式为:

$$a(u, v) = f(v), \forall v \in V. \quad (2.41)$$

对应的极小化问题同(2.13)。事实上, 该弱形式同样存在唯一解, 直接验证 Lax-Milgram 引理即可。

**练习 2.2** 如果我们把问题(2.38)左端项中  $b(x)$  取为  $b(x) \equiv 0$ , 该问题弱形式的适定性是否还满足? 如果不满足适定性, 是 Lax-Milgram 引理的哪个部分出了问题?

### 2.3.2 四阶双调和方程

#### 定义 2.5 (四阶双调和方程)

设  $\Omega \in \mathbb{R}^n$  是一个有界单连通区域, 其边界为  $\partial\Omega$ , 考虑下面的双调和方程:

$$\begin{cases} -\Delta^2 u = f, & x \in \Omega \\ u = \frac{\partial u}{\partial n} = 0, & x \in \partial\Omega \end{cases} \quad (2.42)$$

该方程称为四阶双调和方程。

对这个方程, 首先我们需要给出其变分问题的具体形式。此处取测试函数空间  $V = H_0^2(\Omega)$ , 即  $v \in H_0^2(\Omega)$ , 在(2.42)左右两边同乘函数  $v$ , 并在  $\Omega$  上作积分, 得:

$$\begin{aligned} \int_{\Omega} -v \Delta^2 u dx &= - \int_{\Omega} [\nabla \cdot (v \nabla \Delta u) - \nabla v \cdot \nabla (\Delta u)] dx \\ &= \int_{\Omega} \nabla v \cdot \nabla (\Delta u) dx \\ &= \int_{\Omega} \Delta u \Delta v dx \end{aligned} \quad (2.43)$$

**注** (2.43)真的没有少一个负号?

由(2.43), 设  $B(u, v) = \int_{\Omega} \Delta u \Delta v dx$ ,  $\langle f, v \rangle = \int_{\Omega} f v dx$ , 变分问题的描述为:

$$B(u, v) = \langle f, v \rangle. \quad (2.44)$$

由于此处  $u, v \in H_0^2(\Omega)$ , 验证 Lax-Milgram 引理条件可得该弱形式具有适定性。下面给出  $B(u, v)$  满足  $H_0^2(\Omega)$ -椭



圆性的证明。

### 引理 2.1

设  $u \in H_0^2(\Omega)$ ,  $|\cdot|$  为 Sobolev 半范数,  $\|\cdot\|$  为 Sobolev 范数, 那么

$$\|\Delta u\|_{0,\Omega}^2 = |u|_{2,\Omega}^2. \quad (2.45)$$

**证明** 根据定义:

$$\|\Delta v\|_{0,\Omega}^2 = \int_{\Omega} \left( \sum_{i=1}^n (\partial_{ii} v)^2 + \sum_{i \neq j} \partial_{ii} v \partial_{jj} v \right) dx. \quad (2.46)$$

$$|v|_{2,\Omega}^2 = \int_{\Omega} \left( \sum_{i=1}^n (\partial_{ii} v)^2 + \sum_{i \neq j} (\partial_{ij} v)^2 \right) dx. \quad (2.47)$$

由分部积分公式 (该公式由 Green 公式导出), 有:

$$\begin{aligned} \int_{\Omega} (\partial_{ij} v)^2 dx &= \int_{\Omega} \partial_{ij} v \partial_j (\partial_i v) dx \\ &= - \int_{\Omega} \partial_i v \partial_{ijj} v dx \\ &= \int_{\Omega} \partial_{ii} v \partial_{jj} v dx. \end{aligned} \quad (2.48)$$

由此即证明了引理 2.1。

### 定理 2.5 (Poincare-Friedrichs)

设集合  $\Omega$  有界,  $v \in H_0^m(\Omega)$ , 那么必定存在一个常数  $C(\Omega)$ , 使得:

$$\|v\|_{0,\Omega} \leq C(\Omega) |v|_{m,\Omega}. \quad (2.49)$$

由 2.1 和 2.5 可知:

$$B(u, u) = \|\Delta u\|_{0,\Omega}^2 = |u|_{2,\Omega}^2 \geq C \|u\|_{2,\Omega}^2. \quad (2.50)$$

即: 算子  $B$  是椭圆算子。

## 2.4 本章总结

本章主要讨论有限元方法的第一步: 将微分方程转化为其弱形式。重点讨论了两个结论的成立条件:

- 方程弱形式和它对应优化问题的等价性。
- 方程弱形式的适定性。

Lax-Milgram 引理是本章的核心内容, 弱解的存在唯一性由该引理保证。后面我们同样讨论了一些具体方程的例子。时间所限, 我没有把所有课本例子记录下来。并且, 这一章用到了很多 Sobolev 空间的相关记号和定理, 这部分将在第三章进行讲述。

## 第3章 Sobolev 空间

本章的内容主要是在 Lebesgue 积分的框架下，简单介绍一下广义函数与 Sobolev 空间，为后续的讨论巩固基础，提供依据。

为叙述方便，先给出一些符号的定义。

### 定义 3.1

$$\text{supp}(u) := \overline{\{\mathbf{x} : \mathbf{x} \in \Omega, u(\mathbf{x}) \neq 0\}}. \quad (3.1)$$

$$C_0^\infty(\Omega) = D(\Omega) := \{u : u \in C^\infty(\Omega), \text{supp}(u) \subset \Omega\}. \quad (3.2)$$

$$L_{loc}^1(\Omega) := \{f : f \in L^1(\Omega_1) \forall \text{ compact set } \Omega_1 \subset \Omega\}. \quad (3.3)$$

### 3.1 $L^p$ 空间内容回顾

$L^p$  空间的内容我们早在“实变函数”课程中已经学过，此处仅作简单回顾，不给出具体证明。如无特殊说明，本章中  $\int_\Omega f(x)dx$  均表示 Lebesgue 积分。

### 定义 3.2 ( $L^p$ 范数)

设区域  $\Omega \in \mathbb{R}^n$  为 Lebesgue 非空可测集， $f$  是  $\Omega$  上的实值函数， $f$  的  $L^p$  范数定义为：

$$\|f\|_{L^p(\Omega)} = \left( \int_\Omega |f(x)|^p dx \right)^{\frac{1}{p}}, \quad 1 \leq p < \infty, \quad (3.4)$$

$$\|f\|_{L^\infty(\Omega)} = \text{ess sup}_{x \in \Omega} |f(x)|. \quad (3.5)$$

定义空间：

$$L^p(\Omega) = \{f : \|f\|_{L^p(\Omega)} < \infty\}, \quad 1 \leq p \leq \infty. \quad (3.6)$$

### 命题 3.1 ( $L^p$ 空间上的一些重要不等式)

1. (Young 不等式) 对于  $a, b \geq 0, 0 \leq p, q \leq +\infty, \frac{1}{p} + \frac{1}{q} = 1$ , 我们有：

$$ab \leq \frac{1}{p}a^p + \frac{1}{q}b^q. \quad (3.7)$$

2. (Holder 不等式)  $1 \leq p, q \leq \infty, f, g \in L^p(\Omega)$ , 则：

$$\|fg\|_{L^1(\Omega)} \leq \|f\|_{L^p(\Omega)} \cdot \|g\|_{L^q(\Omega)}. \quad (3.8)$$

3. (Minkowski 不等式)  $1 \leq p \leq \infty, f, g \in L^p(\Omega)$ , 则：

$$\|f + g\|_{L^p(\Omega)} \leq \|f\|_{L^p(\Omega)} + \|g\|_{L^p(\Omega)}. \quad (3.9)$$

**注** Minkowski 不等式表明了  $L^p$  范数满足三角不等式，结合其正定性和正齐次性，可以说明由(3.4)定义的表达式确实是一个范数。

### 定理 3.1

对  $1 \leq p \leq \infty, L^p(\Omega)$  是一个 Banach 空间。

### 定理 3.2

对于  $1 \leq p < \infty, C_0^\infty(\Omega)$  在  $L^p(\Omega)$  中稠密。其中  $C_0^\infty(\Omega)$  表示  $\Omega$  上所有紧支集光滑函数构成的集合。

上面两个定理的证明可以参考任何一本“实变函数”课程的教材。

## 3.2 广义导数

在数学分析课程中，我们给出的导数定义如下：

$$f'(x) = \lim_{\Delta x \rightarrow 0} \frac{f(x + \Delta x) - f(x)}{\Delta x}. \quad (3.10)$$

但这个形式逐渐无法适应对方程广义解的研究，原因主要有两点：

- 该定义式对函数  $f$  的光滑性要求较高。
- 在广义解的研究中，我们更关注导数的整体的性质而非某点处的取值。但传统导数却是逐点定义的。

这是我们推广导数定义的动机。

把局部定义的导数概念向全局定义推广，重要的突破口是分部积分公式。

### 命题 3.2 (分部积分)

设  $\Omega \in \mathbb{R}^n$ ,  $f(x) \in C^n(\Omega)$ ,  $\alpha := \{\alpha_1, \dots, \alpha_n\}$  为多重指标且  $|\alpha| \leq n$ ,  $\phi(x) \in D(\Omega)$ , 那么：

$$\int_{\Omega} \partial^{\alpha} f(x) \cdot \phi(x) dx = (-1)^{|\alpha|} \int_{\Omega} f(x) \partial^{\alpha} \phi(x) dx. \quad (3.11)$$

可以看到，等式(3.11)右端仅仅要求  $f(x) \in L^1(\Omega)$ 。相比(3.10)，(3.11)降低了对函数正则性的要求，并且也是一个  $\Omega$  上全局定义的函数。已知分部积分公式对  $f \in C^n(\Omega)$  成立，我们不妨利用该公式进行一些推广。设  $g(x)$  满足

$$\int_{\Omega} g(x) \phi(x) dx = (-1)^{|\alpha|} \int_{\Omega} f(x) \partial^{\alpha} \phi(x) dx. \quad (3.12)$$

对任意  $\phi \in D(\Omega)$  均成立，那么在允许相差一个零测集的情形下我们可以近似认为  $g(x) = \partial^{\alpha} f(x)$ 。于是，我们由此给出了广义导数的定义。

### 定义 3.3 (广义导数)

对于  $f(x) \in L^1_{loc}(\Omega)$ ，如果存在  $g(x) \in L^1_{loc}(\Omega)$ ，使得：

$$\int_{\Omega} g(x) \phi(x) dx = (-1)^{|\alpha|} \int_{\Omega} f(x) \partial^{\alpha} \phi(x) dx, \forall \phi \in C(\Omega), \quad (3.13)$$

那么我们称  $g(x)$  为  $f(x)$  的  $|\alpha|$  阶广义导数，记作

$$D^{\alpha} f(x) = g(x). \quad (3.14)$$

**例题 3.1** 设区域  $\Omega = (-1, 1)$ ，求  $f(x) = |x|^t$  的广义导数。

按定义(3.12)，计算下面的积分：

$$\begin{aligned} \int_{-1}^1 f(x) \phi'(x) dx &= \int_{-1}^0 (-x)^t \phi'(x) dx + \int_0^1 x^t \phi'(x) dx \\ &= (-x)^t \phi(x)|_{0-} + \int_{-1}^0 t(-x)^{t-1} \phi(x) dx - x^t \phi(x)|_{0+} - \int_0^1 t x^{t-1} dx. \end{aligned} \quad (3.15)$$

取

$$g(x) = \begin{cases} t|x|^{t-1}, & 0 < x < 1 \\ -t|x|^{t-1}, & -1 < x < 0 \end{cases} \quad (3.16)$$

当  $t < 0$  时，

$$\int_{-1}^1 f(x) \phi'(x) dx = - \int_{-1}^1 g(x) \phi(x) dx, \quad (3.17)$$

且  $g(x) \in L^1_{loc}(\Omega)$ ，于是  $t > 0$  时我们有  $Df = g$ 。而  $t < 0$  时， $f$  的广义导数不存在。

关于广义导数，我们有下面这些结论：

**命题 3.3**

如果  $u \in C^{|\alpha|}(\Omega)$ , 那么它的弱导数  $D^\alpha$  存在, 且该弱导数恰好就是其常义导数。

**证明** 由分部积分公式即可直接得到。

**命题 3.4**

设  $\Omega = \Omega_1 \cup \Omega_2$ ,  $m(\Omega_1 \cap \Omega_2) = 0$ , 设函数  $u$  在  $\bar{\Omega}$  上连续, 分别在  $\Omega_1, \Omega_2$  上连续可微, 那么  $u$  的一阶弱导数总是存在, 并且在  $\Omega_1$  或是  $\Omega_2$  内部与常义的一阶导数相等。

**证明** 设  $v(x) = \frac{\partial u}{\partial x_i}$ , 那么对于任何  $\phi(x) \in C_0^\infty(\Omega)$ , 我们有:

$$\begin{aligned} \int_{\Omega} v(x) \phi(x) dx &= \int_{\Omega_1} \frac{\partial u}{\partial x_i} \phi(x) dx + \int_{\Omega_2} \frac{\partial u}{\partial x_i} \phi(x) dx \\ &= \int_{\Gamma} u \frac{\partial \phi}{\partial x_i} ds - \int_{\Omega_1} u \frac{\partial \phi}{\partial x_i} dx - \int_{\Omega_2} u \frac{\partial \phi}{\partial x_i} ds + \int_{\bar{\Gamma}} u \frac{\partial \phi}{\partial x_i} ds. \end{aligned} \quad (3.18)$$

其中  $\Gamma$  和  $\bar{\Gamma}$  位置相同, 方向相反。由(3.18)可知,

$$\int_{\Omega} v \phi dx = - \int_{\Omega} u \frac{\partial \phi}{\partial x_i} dx. \quad (3.19)$$

从而,  $v$  是  $\phi$  关于  $x_i$  的弱导数。这就说明了弱导数的存在性, 并且  $v$  分别限制在  $\Omega_1$  和  $\Omega_2$  上, 就是常义导数的定义。

**命题 3.5**

$\Omega_1, \Omega_2$  的定义同上面的命题, 函数  $u(x)$  定义为:

$$u = \begin{cases} 1, & x \in \Omega_1, \\ 2, & x \in \Omega_2. \end{cases} \quad (3.20)$$

那么  $u$  至少一个方向的弱偏导数不存在。

**证明** 如果所有方向上弱偏导数  $v(x) = \frac{\partial u}{\partial x_i}$  均存在, 那么  $v$  在区域  $\Omega_1$  和  $\Omega_2$  上均等于其常义导数。由广义导数的定义:

$$- \int_{\Omega} u \frac{\partial \phi}{\partial x_i} dx = \int_{\Omega} v \phi dx = 0. \quad (3.21)$$

与此同时, 由格林公式:

$$- \int_{\Omega} u \frac{\partial \phi}{\partial x_i} dx = \int_{\Gamma} u|_{\Omega_1} \phi n_i ds + \int_{\bar{\Gamma}} u|_{\Omega_2} \phi n_i ds = \int_{\bar{\Gamma}} \phi n_i ds. \quad (3.22)$$

这意味着对任意  $i \in [1, n] \cap \mathbb{N}$ , 均有

$$\int_{\bar{\Gamma}} \phi n_i ds = 0. \quad (3.23)$$

这说明  $n_i \equiv 0$ , 矛盾!

## 3.3 磨光算子以及相关应用

### 3.3.1 磨光算子

**定义 3.4 (磨光算子)**

设  $j(x)$  是  $\mathbb{R}^n$  上的实值函数, 且:

- $j(x) \in C_0^\infty(\mathbb{R}^n)$ ;
- $j(x) \geq 0$  且当  $|x| \geq 1$  时,  $j(x) \equiv 0$ ;

$$\bullet \int_{\mathbb{R}^n} j(x) dx = 1.$$

对于  $u(x) \in L^1(\Omega)$ , 作该函数的简单延拓

$$\tilde{u}(x) = \begin{cases} u(x), & x \in \Omega, \\ 0, & x \notin \Omega. \end{cases} \quad (3.24)$$

那么  $\tilde{u} \in L^1(\mathbb{R}^n)$ 。我们定义磨光算子如下:

$$J_\epsilon u(x) := \epsilon^{-n} \int_{\mathbb{R}^n} j\left(\frac{x-y}{\epsilon}\right) \tilde{u}(y) dy = \epsilon^{-n} \int_{\Omega} j\left(\frac{x-y}{\epsilon}\right) u(y) dy. \quad (3.25)$$

### 注

1. 将  $j(x)$  通过伸缩变换写为  $j_\epsilon(x) := \epsilon^{-n} j(\frac{x}{\epsilon})$ , 变换后的函数依旧满足光滑性和区域积分为 1 的性质, 但此时其支集可以进行收缩。
2. 对  $u(x)$  作用磨光算子是一个“光滑化”的过程, 也就是说, 在尽量少改变函数值的情况下, 提升输入函数  $u(x)$  的光滑性。

下面考虑磨光算子对函数光滑性的影响, 有下面的定理成立:

#### 定理 3.3

若  $u(x) \in L^1(\Omega)$ , 则  $u_\epsilon(x) := J_\epsilon u(x) \in C^\infty(\mathbb{R}^n)$ 。又设  $\bar{A} \subset \Omega$ ,  $\text{dist}(\bar{A}, \partial\Omega) > 0$ , 且  $u$  在  $\Omega \setminus A$  上等于 0, 而  $\text{dist}(A, \partial\Omega) > \epsilon$ , 则  $u_\epsilon(x) \in C_0^\infty(\Omega)$ 。

**证明** 由于  $h_x(y) = \tilde{u}(y)j_\epsilon(x, y)$  关于  $x$  具有一致的紧支集, 且  $j_\epsilon(x, y)$  关于  $x$  无穷次可微, 我们有:

$$D_x^\alpha J_\epsilon u(x) = \int_{\mathbb{R}^n} \tilde{u}(y) D_x^\alpha j_\epsilon(x, y) dy. \quad (3.26)$$

由此即得  $u_\epsilon(x) \in C^\infty(\mathbb{R}^n)$ 。

下面证明第二个结论。首先对(3.25)进行换元, 记  $t := \frac{x-y}{\epsilon}$ , 我们有:

$$J_\epsilon u(x) = \epsilon^{-n} \int_{\mathbb{R}^n} j(t) \tilde{u}(x - \epsilon t) d(y + \epsilon t) = \int_{\mathbb{R}^n} \tilde{u}(x - \epsilon t) j(t) dt. \quad (3.27)$$

作集合

$$A_\epsilon := \{x : x \in \mathbb{R}^n, \text{dist}(x, A) < \epsilon\}, B(x, \epsilon) := \{x - \epsilon y, \|y\| \leq 1\}. \quad (3.28)$$

在  $x \notin A_\epsilon$  时,  $B(x, \epsilon) \cap A = \emptyset$ 。由题目条件, 当  $x \notin A_\epsilon$  时,  $u_\epsilon(x) = 0$ 。而  $\text{dist}(A, \partial\Omega) > \epsilon$ , 此即  $\Omega \setminus A_\epsilon \neq \emptyset$ 。由此,  $u_\epsilon(x) \in C_0^\infty(\Omega)$ 。

下面讨论广义导数和磨光算子的交换性:

#### 定理 3.4

设  $f(x) \in L_{loc}^1(\Omega)$ , 且具有  $|\alpha|$  阶广义导数  $D^\alpha f(x)$ , 那么

$$D^\alpha J_\epsilon f(x) = J_\epsilon D^\alpha f(x). \quad (3.29)$$

**证明** 对  $\delta > 0$ , 作集合

$$\Omega_\delta := \{x : x \in \Omega, \text{dist}(x, \partial\Omega) \geq \delta\}. \quad (3.30)$$

对于  $x \in \Omega_\delta$  ( $\delta > \epsilon$ ), 我们有:

$$J_\epsilon f(x) = \epsilon^{-n} \int_{\Omega} j\left(\frac{x-y}{\epsilon}\right) f(y) dy. \quad (3.31)$$

求导可得:

$$D^\alpha J_\epsilon f(x) = \epsilon^{-n} \int_{\Omega} (-1)^{|\alpha|} f(y) D_y^\alpha j\left(\frac{x-y}{\epsilon}\right) dy. \quad (3.32)$$



由于  $j$  的紧支集包含在  $x$  为球心,  $\epsilon$  为半径的闭球中, 可得  $j(\frac{x-y}{\epsilon}) \in D(\Omega)$ 。由广义导数的定义可得:

$$\int_{\Omega} (-1)^{|\alpha|} f(y) D_y^{\alpha} j(\frac{x-y}{\epsilon}) dy = \int_{\Omega} (-1)^{|\alpha|} (-1)^{|\alpha|} j(\frac{x-y}{\epsilon}) D_y^{\alpha} f(y) dy = \int_{\Omega} j(\frac{x-y}{\epsilon}) D_y^{\alpha} f(y) dy. \quad (3.33)$$

综上所述, 我们可以推知, 在  $\Omega_{\delta}$  内,  $D^{\alpha} J_{\epsilon} = J_{\epsilon} D^{\alpha}$ 。

### 3.3.2 均值逼近定理

根据上面的结论, 磨光算子  $J_{\epsilon}$  可以把一个一般的函数  $u$  转化为一个光滑函数  $J_{\epsilon} u$ , 由此导出了一个很自然的问题: 作用后的  $J_{\epsilon} u$  和函数  $u$  之间具体有什么关系? 这就是本节即将描述的**均值逼近定理**。作为重要推论, 该定理说明了之前讨论的**变分原理**的合理性。

在讨论均值逼近定理之前, 先给出一个引理。

#### 定理 3.5

设  $\Omega$  是  $\mathbb{R}^n$  中的有界可测集,  $u(x) \in L^p(\Omega)$  是有界函数,  $1 \leq p < +\infty$ , 如果在  $\Omega$  外补充定义  $u(x) = 0$ , 那么  $u$  一致连续。

**注** 这里的一致连续定义不同于数学分析中所叙述的一致连续, 而是指在  $L^p$  范数的意义下一致连续。

**证明** 根据定义, 要证一致连续, 我们只需要证明对任何  $\epsilon > 0$ , 存在  $\eta > 0$  使得对任何  $|h| < \eta$  有

$$\|u(x+h) - u(x)\|_{L^p(\Omega)} < \epsilon. \quad (3.34)$$

不妨假设  $\Omega$  是闭长方体, 否则可以将  $\Omega$  置于某个闭长方体内。由有界性,  $|u(x)| \leq M$ , 根据 Borel 定理,  $\forall \epsilon_1 > 0, \delta_1 > 0$ , 存在连续函数  $v(x)$  使得  $|v(x)| \leq M$  且:

$$|u(x) - v(x)| < \epsilon_1, \forall x \in \Omega \setminus E, m(E) < \delta_1. \quad (3.35)$$

其中  $E = \{x : |u(x) - v(x)| \geq \epsilon_1\}$ 。根据三角不等式 (在  $L^p$  空间中表现为 Minkovsky 不等式), 我们有:

$$\|u(x+h) - u(x)\|_{L^p} \leq \|u(x+h) - v(x+h)\|_{L^p} + \|v(x+h) - v(x)\|_{L^p} + \|u(x) - v(x)\|_{L^p}. \quad (3.36)$$

对(3.36)右端的三项内容逐次进行分析。首先, 由(3.35), 我们如下估计:

$$\begin{aligned} \|u(x) - v(x)\|_{L^p(\Omega)}^p &= \int_{\Omega \setminus E} |u(x) - v(x)|^p dx + \int_E |u(x) - v(x)|^p dx \\ &\leq \epsilon_1^p m(\Omega \setminus E) + (2M)^p \delta_1. \end{aligned} \quad (3.37)$$

可以取充分小的  $\epsilon_1$  和  $\delta_1$ , 使得(3.36)的右端项小于  $\frac{\epsilon}{3}$ 。

如果对  $v$  在  $\Omega$  外做零延拓, 我们可得:

$$\begin{aligned} \|u(x+h) - v(x+h)\|_{L^p(\Omega)} &= \left( \int_{\Omega} |u(x+h) - v(x+h)|^p dx \right)^{\frac{1}{p}} \\ &\leq \left( \int_{\mathbb{R}^n} |u(x+h) - v(x+h)|^p dx \right)^{\frac{1}{p}} \\ &= \left( \int_{\Omega} |u(x) - v(x)|^p dx \right)^{\frac{1}{p}} \\ &= \|u - v\|_{L^p(\Omega)} < \frac{\epsilon}{3}. \end{aligned} \quad (3.38)$$

最后, 由  $v(x)$  在  $\Omega$  上的一致连续性 (此处为微积分里定义的一致连续性, 依据是 Cantor 定理), 可以取  $|h|$  充分小, 使得

$$|v(x+h) - v(x)| < \frac{1}{6} (m(\Omega))^{-\frac{1}{p}} \epsilon. \quad (3.39)$$

从而:

$$\begin{aligned}
 \|v(x+h) - v(x)\|_{L^p(\Omega)} &= \left( \int_{\Omega} |v(x+h) - v(x)|^p dx \right)^{\frac{1}{p}} \\
 &\leq \left( \int_{\Omega_i} |v(x+h) - v(x)|^p dx \right)^{\frac{1}{p}} + \left( \int_{\Omega_e} |v(x+h) - v(x)|^p dx \right)^{\frac{1}{p}} \\
 &< \frac{\epsilon}{6} + \left( \int_{\Omega_e} |v(x)|^p dx \right)^{\frac{1}{p}} \\
 &< \frac{\epsilon}{6} + M(m(\Omega_e))^{\frac{1}{p}}.
 \end{aligned} \tag{3.40}$$

其中

$$\Omega_i := \{x : x \in \Omega, x+h \in \Omega\}, \Omega_e := \{x : x \in \Omega, x+h \notin \Omega\}. \tag{3.41}$$

取  $\eta$  足够小即可保证(3.40)小于  $\frac{\epsilon}{3}$ , 从而原定理得证。

**注** 该定理的证明思路在一致连续的证明中非常常见, 即先寻求一个光滑函数  $v$  来逼近已知的函数  $u$ , 然后使用三段估计法来估计  $\|u(x+h) - u(x)\|$ 。

下面介绍均值逼近定理, 该定理表明用  $J_\epsilon u$  来估计  $u$  是合理的。

#### 定理 3.6 (均值逼近定理)

设  $u(x) \in L^p(\Omega)$ ,  $1 \leq p < +\infty$ , 则:

- $\|J_\epsilon u\|_{0,p,\Omega} \leq \|u\|_{0,p,\Omega}$ .
- $\lim_{\epsilon \rightarrow 0} \|J_\epsilon u - u\|_{0,p,\Omega} = 0$ .

**证明** 对  $1 < p < +\infty$ , 由 Holder 不等式, 有:

$$\begin{aligned}
 |u_\epsilon(x)|^p &= \epsilon^{-np} \left| \int_{\Omega} u(y) \left[ j\left(\frac{x-y}{\epsilon}\right)^{\frac{1}{p}} \right] \left[ j\left(\frac{x-y}{\epsilon}\right)^{\frac{1}{q}} \right] dy \right|^p \\
 &\leq \epsilon^{-np} \left[ \int_{\Omega} |u(y)|^p \left| j\left(\frac{x-y}{\epsilon}\right) \right| dy \right] \left[ \int_{\Omega} \left| j\left(\frac{x-y}{\epsilon}\right) \right| dy \right]^{\frac{p}{q}} \\
 &\leq \epsilon^{-np(1-\frac{1}{q})} \int_{\Omega} |u(y)|^p \left| j\left(\frac{x-y}{\epsilon}\right) \right| dy \\
 &= \epsilon^{-n} \int_{\Omega} |u(y)|^p j\left(\frac{x-y}{\epsilon}\right) dy.
 \end{aligned} \tag{3.42}$$

对  $p=1$ , 直接可得(3.42)依旧成立。从而, 我们在(3.42)两端同时对  $x$  进行积分, 有:

$$\int_{\Omega} |u_\epsilon(x)|^p dx \leq \int_{\Omega} \left[ \int_{\Omega} |u(y)|^p \epsilon^{-n} j\left(\frac{x-y}{\epsilon}\right) dx \right] dy = \int_{\Omega} |u(y)|^p dy. \tag{3.43}$$

从而:

$$\|J_\epsilon u\|_{0,p,\Omega} \leq \|u\|_{0,p,\Omega}. \tag{3.44}$$

此即第一个结论。

下面证明第二个结论, 即  $u_\epsilon(x)$  在  $L^p(\Omega)$  中收敛于  $u$ 。为此, 首先对  $u(x)$  做零延拓, 设延拓后的函数为  $\tilde{u}$ , 则:

$$\|u - u_\epsilon\|_{0,p,\Omega}^p \leq \int_{\mathbb{R}^n} |\tilde{u}(x) - u_\epsilon(x)|^p dx. \tag{3.45}$$

令  $\xi = x - y$ , 则有:

$$\begin{aligned}
 \|u - u_\epsilon\|_{0,p,\Omega}^p &\leq \int_{\mathbb{R}^n} \epsilon^{-n} \left| \int_{\mathbb{R}^n} (\tilde{u}(x) - \tilde{u}(y)) j\left(\frac{x-y}{\epsilon}\right) dy \right|^p dx \\
 &\leq \epsilon^{-n} \int_{\mathbb{R}^n} \left[ \int_{\mathbb{R}^n} |\tilde{u}(x) - \tilde{u}(x-\xi)| j\left(\frac{\xi}{\epsilon}\right) d\xi \right]^p dx \\
 &\leq \epsilon^{-n} \int_{\mathbb{R}^n} \left[ \int_{|\xi| \leq \epsilon} |\tilde{u}(x) - \tilde{u}(x-\xi)|^p d\xi \right] \left[ \int_{|\xi| \leq \epsilon} \left| j\left(\frac{\xi}{\epsilon}\right) \right|^q d\xi \right]^{\frac{p}{q}} dx \\
 &\leq C \epsilon^{-n} \int_{\Omega} \left( \int_{|\xi| \leq \epsilon} |u(x) - u(x-\xi)|^p d\xi \right) dx.
 \end{aligned} \tag{3.46}$$

由  $u$  的一致连续性, 可得

$$\lim_{\epsilon \rightarrow 0} \|u_\epsilon - u\|_{L^p(\Omega)} = 0 \tag{3.47}$$

**注** 这个定理事实上阐述了“用光滑函数来逼近  $L^p(\Omega)$  空间上的函数”的行为, 在后面的讨论中, 这个结论非常常用。下面两个非常有用的推论就是一些例子。

#### 推论 3.1

对  $1 \leq p < +\infty$ ,  $\Omega$  有界, 则  $C_0^\infty(\Omega)$  在  $L^p(\Omega)$  中稠密。

**证明** 对  $\delta > 0$ , 基于有界区域  $\Omega$ , 可以作子区域如下:

$$\Omega_\delta := \{x : x \in \Omega, \text{dist}(x, \partial\Omega) \geq \delta\}. \tag{3.48}$$

$\forall u \in L^p(\Omega), \exists \delta$  使得  $\forall \eta > 0$ ,

$$\int_{\Omega \setminus \Omega_\delta} |u(x)|^p dx < \eta^p. \tag{3.49}$$

考虑函数

$$u_\delta(x) = \begin{cases} u(x), & x \in \Omega_\delta, \\ 0, & x \notin \Omega_\delta. \end{cases} \tag{3.50}$$

对于  $0 < \epsilon < \frac{\delta}{2}$ ,  $J_\epsilon u_\delta \in C_0^\infty(\Omega)$ , 并且:

$$\|u - J_\epsilon u_\delta\|_{0,p,\Omega} \leq \|u - u_\delta\|_{0,p,\Omega} + \|u_\delta - J_\epsilon u_\delta\|_{0,p,\Omega} \tag{3.51}$$

取  $\eta, \delta \rightarrow 0$  即得  $\|u - J_\epsilon u_\delta\|_{0,p,\Omega} \rightarrow 0$ 。

#### 推论 3.2 (变分法基本原理)

设  $u(x) \in L^p(\Omega), 1 \leq p < +\infty, \Omega$  有界, 且

$$\int_{\Omega} u(x) \varphi(x) dx = 0, \forall \varphi(x) \in C_0^\infty(\Omega), \tag{3.52}$$

则在  $\Omega$  上有  $u = 0$  a.e.

**证明**  $\forall \delta > 0$ , 按(3.48)式做区域  $\Omega_\delta$ , 取  $0 < \epsilon < \delta$ , 当  $x \in \Omega_\delta$  时  $j(\frac{x-y}{\epsilon}) \in C_0^\infty(\Omega)$ 。由变分法条件:

$$J_\epsilon u(x) = 0. \tag{3.53}$$

由均值逼近定理: 在  $\Omega_\delta$  上,  $u(x) = 0$  a.e. 又由  $\delta$  的任意性,  $u(x) = 0$  在  $\Omega$  上几乎处处成立。

**注** 上面的推论保证了第二章阐述的弱形式和变分原理是处理原方程的合理方案。

### 3.3.3 单位分解定理

在前面几个小节, 我们通过磨光算子研究了  $L^p(\Omega)$  上函数的局部光滑逼近。接下来我们需要着眼于局部性质与整体性质的联系, 建立局部与整体的关系。单位分解定理正是在局部性质和整体性质之间, 构建了一道桥梁。

此处给出单位分解定理的叙述，具体证明详见微分几何的教材。

### 定理 3.7 (有穷单位分解定理)

设  $O_1, \dots, O_n$  是有限个开集， $F \in \mathbb{R}^n$  是一个有界闭集，且  $F \subset \cup_{i=1}^m O_i$ ，那么存在函数  $\phi_i(x)$  满足下面几条性质：

- $0 \leq \phi_i(x) \leq 1, i = 1, 2, \dots, m$ 。
- $\phi_i(x) \in C_0^\infty(\mathbb{R}^n)$ ，且  $\text{supp}(\phi_i) \subset O_i$ 。
- $\sum_{i=1}^m \phi_i(x) = 1, \forall x \in F$ 。



### 定理 3.8 (无穷单位分解定理)

设  $\Omega$  是  $\mathbb{R}^n$  中的任何有界开集，开集族  $\{O_i\}_{i=1}^\infty$  是  $\Omega$  的一个开覆盖，则存在一族函数  $\phi_i(x)$  满足：

- $0 \leq \phi_i(x) \leq 1$ ;
- $\phi_i(x) \in C_0^\infty(\mathbb{R}^n)$ ;
- 对任何  $\phi_i(x)$ ，存在  $O_{n_i}$  使得  $\text{supp}(\phi_i) \subset O_{n_i}$ ;
- $\sum_{i=1}^\infty \phi_i(x) = 1$ 。



## 3.4 Sobolev 空间

### 3.4.1 相关定义

**注** 本节中仅就之后可能用到的 Sobolev 空间相关结论做一简介，并不关注具体细节，如对具体细节感兴趣可以关注“偏微分方程”和“泛函分析”的相关著作。

### 定义 3.5 (Sobolev 空间)

对非负整数  $m$ ，定义空间

$$H^{m,p}(\Omega) := \{u \in L^p(\Omega) : D^\alpha u \in L^p(\Omega), \forall |\alpha| \leq m\}, \quad (3.54)$$

并定义其上的范数为

$$\|u\|_{H^{m,p}(\Omega)} := \left( \sum_{|\alpha| \leq m} \|D^\alpha u\|_{L^p(\Omega)}^p \right)^{\frac{1}{p}}. \quad (3.55)$$

$$\|u\|_{H^{m,\infty}(\Omega)} := \max_{|\alpha| \leq m} \|D^\alpha u\|_{L^\infty(\Omega)}. \quad (3.56)$$

特别地，如果  $p = 2$ ，我们可以把  $H^{m,2}(\Omega)$  简写为  $H^m(\Omega)$ 。



### 引理 3.1

对于  $1 \leq p \leq \infty$ ，Sobolev 空间  $H^{m,p}(\Omega)$  均为 Banach 空间。特别地， $H^m(\Omega)$  是 Hilbert 空间，其上的内积定义为

$$\langle u, v \rangle := \sum_{|\alpha| \leq m} \langle D^\alpha u, D^\alpha v \rangle. \quad (3.57)$$



### 命题 3.6

直接根据定义可以推知，对于不同上标的 Sobolev 空间  $H^{m,p}(\Omega)$ ，有下面的包含关系：

$$C^\infty(\bar{\Omega}) \subset \dots \subset H^{m+1,p}(\Omega) \subset H^{m,p}(\Omega) \subset \dots \subset H^{0,p}(\Omega) = L^p(\Omega). \quad (3.58)$$



Sobolev 空间的指数  $m$  可以小于 0，对于负指数的 Sobolev 空间我们如下定义：

**定义 3.6**

$H^{-m}(\Omega)$  定义为  $H_0^m(\Omega)$  的对偶空间，并赋范数

$$\|f\|_{-m,\Omega} := \sup_{0 \neq v \in H_0^m(\Omega)} \frac{|\langle f, v \rangle|}{|v|_{m,\Omega}}. \quad (3.59)$$

在 Sobolev 空间  $H^{m,p}(\Omega)$  上可以定义半范数如下:

$$|u|_{m,p,\Omega} := \left( \sum_{|\alpha|=m} \|D^\alpha u\|_{0,p,\Omega}^p \right)^{\frac{1}{p}}, \quad 1 \leq p < +\infty, \quad (3.60)$$

$$|u|_{m,\infty,\Omega} := \sup_{|\alpha|=m} \|D^\alpha u\|_{0,\infty,\Omega}. \quad (3.61)$$

事实上，如果  $\Omega$  满足一定条件，Sobolev 半范数和 Sobolev 范数是等价的。该结论由下面的 Poincare-Friedrichs 不等式保证，为我们研究 Sobolev 范数提供了不小便利。

**定理 3.9 (Poincare-Friedrichs 不等式)**

如果  $\Omega$  单连通，且至少在一个方向上有界，那么对任何正整数  $m$ ，存在常数  $C(m)$  使得：

$$\|v\|_{m,\Omega} \leq C(m)|v|_{m,\Omega}, \quad \forall v \in H_0^m(\Omega). \quad (3.62)$$

**3.4.2 Sobolev 嵌入定理**

本节中讨论 Sobolev 嵌入定理，主要思路是将一些比较难以直接研究的 Sobolev 空间，转移到更大的 Sobolev 空间中进行处理。

首先我们需要给出连续嵌入的定义。

**定义 3.7**

如果空间  $X, Y$  满足下面两个条件：

- $X \subset Y$ .
- $\exists C > 0$  s.t.

$$\|x\|_Y \leq C\|x\|_X, \quad \forall x \in X. \quad (3.63)$$

则称空间  $X$  嵌入到  $Y$ ，记作  $X \hookrightarrow Y$ 。特别地，如果  $id: X \rightarrow Y, x \mapsto x$  为紧算子，称  $X$  紧嵌入到  $Y$ ，记作  $X \hookrightarrow\hookrightarrow Y$ 。

对于一个 Sobolev 空间，我们希望找到一个类似连续函数空间的 Banach 空间来方便研究其内部函数的解析性质。为此我们对  $C^{(m)}(\Omega)$  的定义做一些扩展，给出下面的  $(m, \alpha)$  阶 Holder 连续函数空间的定义。

**定义 3.8**

Holder 连续函数空间  $C^{(m,\alpha)}(\Omega)$  由满足下面性质的函数组成：

- $f \in C^m(\Omega)$ .
- $\forall 0 \leq |\beta| \leq m, f$  满足下面的 Holder 条件：

$$|\partial^\beta f(x) - \partial^\beta f(y)| \leq C|x - y|^\alpha. \quad (3.64)$$

在该函数空间中赋予范数

$$\|v\|_{C^{m,\alpha}(\bar{\Omega})} := \|v\|_{m,\infty,\Omega} + \max_{|\beta| \leq m} \sup_{x,y \in \bar{\Omega}, x \neq y} \frac{|\partial^\beta v(x) - \partial^\beta v(y)|}{\|x - y\|^\alpha}, \quad (3.65)$$

可知此时  $C^{m,\alpha}(\bar{\Omega})$  为一个 Banach 空间。

讨论嵌入定理前，我们对区域  $\Omega$  的性质做一些限定，下面介绍两类特殊的区域形式。



**定义 3.9 (有限锥形区域)**

称区域  $\Omega$  为有限锥形区域, 如果它满足下面两个条件:

- $\Omega$  是  $\mathbb{R}^n$  中的有界开集。
- $\Omega$  中任何一点都是有限锥  $C_x (\subset \Omega)$  的顶点, 并且每一个有限锥  $C_x$  全同于过原点的一个有限锥  $C_0$ 。

**定义 3.10 (L 型区域)**

称区域  $\Omega$  为 L 型区域, 如果  $\Omega$  满足下面的条件:

- $\Omega$  是  $\mathbb{R}^n$  中的有界开集, 且  $\Omega$  在其边界  $\partial\Omega$  的固定一侧。
- 存在有限个开集  $O_1, \dots, O_m$  使得  $\Omega$  的边界满足  $\partial\Omega \subset \bigcup_{i=1}^m O_i$ , 且  $\partial\Omega_i := \partial\Omega \cap O_i$  在某一局部坐标  $(\xi_{i_1}, \xi_{i_2}, \dots, \xi_{i_n})$  中可用一个 Lip 函数

$$\xi_{i_n} = f_i(\xi_{i_1}, \dots, \xi_{i_{n-1}}) \quad (3.66)$$

表示。

**注** Lipschitz 区域的几何意义: 在区域  $\Omega$  边界上每一点  $x$  有一个邻域  $U_x$  使得  $\partial\Omega \cap U_x$  是一个 Lipschitz 连续函数的图形。

**注** Lipschitz 区域是一个有限锥形区域, 反之则不一定, 但有限锥形区域是有限个 Lipschitz 区域的并。

在做完上述准备后, 给出一些常用的 Sobolev 嵌入定理的介绍。嵌入定理的目标大多是  $L^p$  空间和  $C^{m,\alpha}$  空间。具体的证明则略去不表。

**定理 3.10**

设  $\Omega \in \mathbb{R}^d$  是一个 Lipschitz 区域, 则下面的结论成立:

- 若  $k < \frac{d}{p}$ , 则  $W^{k,p}(\Omega) \hookrightarrow L^q(\Omega)$  对任何  $q \leq p^*$  成立, 其中  $\frac{1}{p^*} = \frac{1}{p} - \frac{k}{d}$ 。
- 若  $k = \frac{d}{p}$ , 则  $W^{k,p}(\Omega) \hookrightarrow L^q(\Omega)$  对任何  $q < \infty$  成立。
- 若  $k > \frac{d}{p}$ , 则:

$$W^{k,p}(\Omega) \hookrightarrow C^{k-[\frac{d}{p}]-1,\beta}(\Omega), \quad (3.67)$$

如果  $\frac{d}{p}$  不为整数, 那么  $\beta = [\frac{d}{p}] + 1 - \frac{d}{p}$ , 否则  $\beta$  可取任何小于 1 的正数。事实上, 把该定理中的“嵌入”改为“紧嵌入”, 结论依然成立。

下面是几个常用的 Sobolev 嵌入结论:

**命题 3.7**

•

$$H^1(\Omega) \hookrightarrow L^2(\Omega) \quad (3.68)$$

对任何维数的 Lipschitz 空间  $\Omega \subset \mathbb{R}^n$  成立。

•

$$W^{1,p}(\Omega) \hookrightarrow L^p(\Omega), W^{k+1,p}(\Omega) \hookrightarrow W^{k,p}(\Omega). \quad (3.69)$$

- 在  $n = 2$  时,

$$H^2(\Omega) \hookrightarrow C^0(\Omega), H^2(\Omega) \hookrightarrow C^0(\bar{\Omega}); \quad (3.70)$$

但  $H^1(\Omega)$  并不嵌入到  $C^0(\Omega)$ 。

### 3.4.3 迹算子

#### 定理 3.11 (迹定理)

设  $\Omega \subset \mathbb{R}^n$  为一个 Lipschitz 空间, 则存在一个有界线性算子

$$\gamma : W^{1,p}(\Omega) \rightarrow L^p(\partial\Omega) \quad (3.71)$$

满足:

$$\gamma u = u|_{\partial\Omega}. \quad (3.72)$$

对  $u \in W^{1,p}(\Omega) \cap C(\bar{\Omega})$  成立。线性算子  $\gamma$  称为迹算子。



迹定理给出了  $u$  的内部信息和边界信息的关系。关于迹算子, 有下面这些不等式成立:

#### 命题 3.8 (迹不等式)

- 存在  $C > 0$  使得

$$\|\gamma\phi\|_{0,\partial\Omega} \leq C\|\phi\|_{1,\Omega}. \quad (3.73)$$

- 存在  $C > 0$  使得

$$\|\gamma v\|_{0,\partial\Omega} \leq C\|v\|_{0,\Omega}^{\frac{1}{2}} \cdot \|v\|_{1,\Omega}^{\frac{1}{2}}. \quad (3.74)$$

- 对于  $1 \leq p \leq \infty$ , 存在  $C > 0$  使得

$$\|\gamma v\|_{0,p,\partial\Omega} \leq C\|v\|_{0,p,\Omega}^{1-\frac{1}{p}} \cdot \|v\|_{1,p,\Omega}^{\frac{1}{p}}. \quad (3.75)$$



## 3.5 本章小结

本章中对有限元方法中可能用到的一系列 Sobolev 相关的知识内容做了简单的归纳。由于本课程的重点并不在于研究 Sobolev 空间, 故很多结论都只是粗略介绍或是省去了证明过程。如果读者对 Sobolev 空间理论感兴趣, 更建议去阅读相关的专著, 或是与泛函分析/偏微分方程相关的著作。

## 第4章 有限元离散化

经过第二章和第三章的前期准备，本章开始我们正式开始讨论有限元离散化的问题。

### 4.1 有限元离散的步骤

#### 4.1.1 变分问题

考虑抽象的变分问题:

##### 定义 4.1 (变分问题)

设  $V$  是某个无限维 Banach 空间，且  $a(\cdot, \cdot)$  和  $f$  分别为满足 Lax-Milgram 定理条件的双线性型和线性泛函，变分问题的描述为：求  $u \in V$  使得  $\forall v \in V$ ,

$$a(u, v) = f(v). \quad (4.1)$$

根据第二章的 Lax-Milgram 定理，可知(4.1)的解存在唯一。但  $V$  是一个无穷维空间，直接在  $V$  上求解弱形式依旧是一件无法完成的任务。

回想第一章我们对 ode 边值问题的讨论，我们选取了分段线性函数子空间  $V_h \leq V$ ，然后在这个有限维空间上求解了弱形式(4.1)，并把  $V_h$  空间上的弱解作为  $V$  上弱解的一种近似。这种思路被称作 **Galerkin 方法**。同样的，对于一般情形下的问题，我们也可以定义与之相对应的 **Galerkin 方法** 和 **Ritz 方法**。

假设  $V_h \leq V$  是已经给定的有限维子空间，那么：

##### 定义 4.2 (Galerkin)

**Galerkin 方法**的思路是求解  $u_h \in V_h$ ，使得

$$a(u_h, v_h) = f(v_h). \quad (4.2)$$

由 Lax-Milgram 定理，可知 Galerkin 方法求得的近似解也具有唯一性。

如果  $a(u, v)$  是一个对称双线性型，那么(4.2)等价于一个最优化问题，**Ritz 方法**就是着眼于求解这个最优化问题的思路。

##### 定义 4.3 (Ritz)

**Ritz 方法**的思路是求解  $u_h \in V_h$  使得：

$$J(u_h) = \min_{v_h \in V_h} J(v_h). \quad (4.3)$$

其中

$$J(v) = \frac{1}{2}a(v, v) - f(v). \quad (4.4)$$

第二章我们已经证明过求解弱形式和优化问题的等价性。

一般的有限元离散问题，在给出 Ritz 方法和 Galerkin 方法的描述之后，就产生了下面两个主要的问题：

- 如何确定子空间  $V_h \leq V$ ？
- 如何寻找子空间  $V_h$  的一组基使得问题容易求解？

本章将给出这两个问题的回答，并特别针对二维空间的问题，给出一整套可行的有限元计算方案。

### 4.1.2 有限元子空间

对于 ode 边值问题，我们用区间段的形式对整个区间  $[0, 1]$  进行划分，对应的基函数即为数值分析课程中讲授的“hat-function”。在考虑高维区域时，继续使用分片多项式逼近依旧是一个好主意，但此时对区域  $\Omega$  的划分方案就不得不进行一些改变。

#### 定义 4.4 (单形)

一个  $k$ -单形是一个由  $k+1$  个顶点组成的  $k$ -维多面体凸包。设  $x_0, \dots, x_n \in \mathbb{R}^n$  且仿射无关，那么  $n$  维单形定义为：

$$K_n := \left\{ \sum_{i=0}^n \theta_i x_i : \sum_{i=0}^n \theta_i = 1, \theta_j \geq 0 \forall j \in [0, n] \cap \mathbb{Z} \right\}. \quad (4.5)$$

对于  $\mathbb{R}^n$  空间的有界区域  $\Omega$ ，我们使用  $n$ -单形对其进行划分。假设我们对区域  $\Omega$  的单形剖分为  $\mathcal{T}_h$ ， $T$  为  $\mathcal{T}_h$  的单元，那么我们对这样的单形剖分有下面这些要求：

1. 对单元  $T$  的要求：
  - $\forall T \in \mathcal{T}_h$ ， $T$  为闭集，其内部  $\overset{\circ}{T}$  非空且连通。
  - $\partial T$  是 Lipschitz-连续的。
  - $\bar{\Omega} = \cup_{T \in \mathcal{T}_h} T$ 。
  - 对于任何两个不同的  $T_1, T_2 \in \mathcal{T}_h$ ，均有  $\overset{\circ}{T}_1 \cap \overset{\circ}{T}_2 = \emptyset$ 。
  - 对每个  $T \in \mathcal{T}_h$ ， $\partial T$  或者是  $\partial \Omega$  的一部分，或者是相邻单元  $T'$  的边。
2. 对于其上求解的多元分片多项式， $h \rightarrow 0$  时，解收敛到原问题的解。
3. 基函数的支集尽量小，使得计算简单。所得的刚度矩阵应当是稀疏矩阵。

图 4.1: 正确的单形剖分

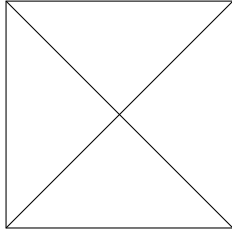
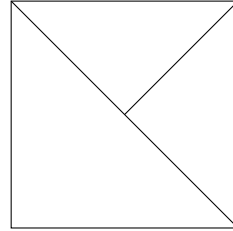


图 4.2: 错误的单形剖分



有了空间划分  $\mathcal{T}_h$  后，定义在该划分空间上的分片多项式即可构成子空间  $V_h \leq V$ 。

### 4.1.3 离散线性方程组与求解

若  $\dim(V_h) = N$ ，设  $\phi_1, \phi_2, \dots, \phi_N$  为  $V_h$  上的一组基函数，设

$$u_h(x) = \sum_{i=1}^N u_i \phi_i(x), v_h(x) = \sum_{i=1}^N v_i \phi_i(x), \quad (4.6)$$

代入(4.2)，我们有：

$$\sum_{i=1}^N a(\phi_i, \phi_j) u_i = f(\phi_j), \quad (4.7)$$

这是一个线性系统，求解该系统即可得到  $u(x)$  的近似值  $u_h(x)$  (事实上是一个分片多项式)。类似对 ode 边值问题的讨论，我们依旧可以称矩阵  $K = (a(\phi_i, \phi_j))_{ij}$  为**刚度矩阵**， $F = (f_j)_{j=1}^N$  为**负载向量**。

如果  $a(\cdot, \cdot)$  是对称的，Galerkin 方法和 Ritz 方法等价。

对于刚度矩阵  $K$ ，我们有下面的结论：

**命题 4.1**

设  $a(\cdot, \cdot) : V \times V \rightarrow \mathbb{R}^n$  是对称且 V-椭圆的双线性形式, 那么对应 Galerkin-Ritz 变分问题近似解的刚度矩阵  $K$  是对称正定的。

**证明** 对称性由  $a(\cdot, \cdot)$  的对称性可即刻得到。对于正定性, 取  $v \in \mathbb{R}^n$ , 记

$$v_h(x) = \sum_{i=1}^N v_i \phi_i(x), \quad (4.8)$$

那么:

$$v^T K v = a(v_h, v_h) \geq a \|v_h\|^2 \geq 0. \quad (4.9)$$

该式取等号当且仅当  $v_h = 0$ , 即  $v = 0$ 。由此, 矩阵  $K$  是正定的。

**注** 上面的性质保证了在变分问题的左侧项  $a(\cdot, \cdot)$  满足一定条件的情况下, 其对应的离散化线性方程组的解存在唯一, 且可以通过针对正定矩阵的算法进行求解。

以上推导均需要满足 **Lax-Milgram 引理**, 因此需要  $V_h$  是  $V$  的子空间, 这种有限元方法被称为**协调有限元方法**。

对于一个具体的问题,  $V = H^1(\Omega)$ , 那么  $V_h$  中的分段多项式也必须在  $H^1(\Omega)$  中。这就要求我们对分段多项式在单元交界处的行为给出一些额外的条件。下面两个定理说明了分片多项式成为  $V_h$  中元素的充要条件。

**定理 4.1**

设  $v \in H^1(T), \forall T \in \mathcal{T}_h$ , 且  $v \in C^0(\bar{\Omega})$ , 则  $v \in H^1(\Omega)$ 。

**证明** 只需要证明  $v$  的广义导数  $D_i v \in L^2(\Omega)$ , 即:

$$-\int_{\Omega} D_i v \cdot \phi dx = \int_{\Omega} v \cdot \partial_i \phi dx, \forall \phi \in D(\Omega). \quad (4.10)$$

由于  $v \in C^0(\bar{\Omega})$ , 且  $v$  在每个单元  $T$  上均为多元多项式, 在每个单元  $T$  上运用 Gauss-Green 公式, 有:

$$\int_{\Omega} v \cdot \partial_i \phi dx = \sum_{T \in \mathcal{T}_h} \int_T v \cdot \partial_i \phi dx = - \sum_T \int_T D_i v \cdot \phi dx + \sum_T \int_{\partial T} v \phi \cdot n_i ds. \quad (4.11)$$

$n_i$  表示单位法向量在第  $i$  个维度下的投影。由(4.11), 要证(4.10), 即证

$$\sum_T \int_{\partial T} v \phi \cdot n_i ds = 0. \quad (4.12)$$

根据我们选取单元  $T$  的第五点要求, 取  $T_0 \in \partial T$ , 有两种情况:

- $T_0$  是区域  $\Omega$  的边界: 由  $\phi \in D(\Omega)$ , 即  $\phi|_{\partial\Omega} = 0$ , 可知  $\int_{T_0} v \phi \cdot n_i ds = 0$ 。
- $T_0$  是两个单元  $T_a$  和  $T_b$  的交线。此时积分式(4.12)中有关线段  $T_0$  的积分表达式为

$$\int_{T_0 \in T_a} v \phi \cdot n_i ds + \int_{T_0 \in T_b} v \phi \cdot n_i ds. \quad (4.13)$$

由于在  $T_a$  和  $T_b$  上, 线段  $T_0$  位置相同, 法向量方向相反, 从而(4.13)的值为 0。故(4.12)得证, 即(4.10)成立。

**定理 4.2**

设  $v|_T \in C^0(T), \forall T \in \mathcal{T}_h$ , 且  $v \in H^1(\Omega)$ , 则  $v \in C^0(\bar{\Omega})$ 。

**证明** 如果  $v$  在两个相邻单元  $e$  和  $e'$  上不连续, 即在其公共边  $\gamma$  上  $v|_e - v|_{e'}$  不恒为 0。不妨其在  $\gamma$  上的某段处为正, 作开集  $U$  使得  $U \subset e \cap e'$ , 且  $v|_e - v|_{e'} > 0$  在  $U \cap \gamma$  上成立, 则作函数  $\phi \in D(U)$  使得在  $\text{supp}(\phi)$  上  $\phi > 0$ 。如果  $v \in H^1(\Omega)$ , 根据  $H^1(\Omega)$  广义函数的定义, 必定有:

$$\int_{U \cap \gamma} (v|_e - v|_{e'}) \cdot \phi \cdot n_i^e ds = 0. \quad (4.14)$$

这与前面的假定相矛盾。

**注** 上面两个定理要求: 我们求得的有限元逼近解必须在整个区域内是连续的。

## 4.2 二维情形

在本节中，我们着眼于  $\Omega \subset \mathbb{R}^2$ ，讨论区域  $\Omega$  的三角划分，以及定义在该划分上的分片多项式函数。首先，我们讨论定义在三角形有界区域上的多项式函数。

### 4.2.1 重心坐标

一个很容易想到的思路是，依据三角形上某些点 (例如三个顶点) 处的函数取值，插值得到一个多元多项式。但如果采用直角坐标，插值多项式的表达式不免会显得非常麻烦。如果能通过一些合理的坐标变换，将一般三角形区域转化为等腰直角三角形区域，则可以大大降低该区域上插值多项式形式的复杂性。这就是引入重心坐标的原因。

图 4.3: 重心坐标变换



#### 定义 4.5 (重心坐标)

考虑平面上的三角形  $\Delta A_1 A_2 A_3$ ，顶点编号为逆时针，设点  $P(x, y)$  落在三角形内部，则定义点  $P$  的重心坐标  $(\lambda_1, \lambda_2, \lambda_3)$  为：

$$\lambda_1 = \frac{m(\Delta P A_2 A_3)}{m(\Delta A_1 A_2 A_3)}, \lambda_2 = \frac{m(\Delta P A_1 A_3)}{m(\Delta A_1 A_2 A_3)}, \lambda_3 = \frac{m(\Delta P A_1 A_2)}{m(\Delta A_1 A_2 A_3)}. \quad (4.15)$$

其中  $m$  表示区域的测度，在这里表示的是三角形的面积。

根据定义，下面的表达式是显然的：

$$\lambda_i(A_j) = \delta_{ij}. \quad (4.16)$$

所以，定义在  $\Delta A_1 A_2 A_3$  的线性函数  $u$  可以表示为：

$$u(x, y) = u(A_1)\lambda_1(x, y) + u(A_2)\lambda_2(x, y) + u(A_3)\lambda_3(x, y). \quad (4.17)$$

分别取  $u(x, y) = x, u(x, y) = y, u(x, y) = 1$ ，代入上式，可得下面的性质：

#### 命题 4.2 (直角坐标与重心坐标的关系)

记  $A_i$  的坐标为  $(x_i, y_i)$ ， $\Delta A_1 A_2 A_3$  上直角坐标和重心坐标的关系式如下：

$$\begin{cases} x = x_1\lambda_1 + x_2\lambda_2 + x_3\lambda_3 \\ y = y_1\lambda_1 + y_2\lambda_2 + y_3\lambda_3 \\ 1 = \lambda_1 + \lambda_2 + \lambda_3. \end{cases} \quad (4.18)$$

重心坐标有许多非常棒的性质，可以大大简化三角形区域上的函数计算。

#### 命题 4.3 (重心坐标的性质)

1. 三角形三顶点  $A_1, A_2, A_3$ ，重心  $G$  的坐标分别为：

$$A_1(1, 0, 0), A_2(0, 1, 0), A_3(0, 0, 1), G\left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right). \quad (4.19)$$

2. 三角形三边的方程:

$$A_1A_2 : \lambda_3 = 0; A_2A_3 : \lambda_1 = 0; A_3A_1 : \lambda_2 = 0. \quad (4.20)$$

3. 平行于三角形某一边 (如  $A_1A_2$ ) 的方程:  $\lambda_3 = c$ 。

4. 任一  $x, y$  的  $k$  次多项式是  $\lambda_1, \lambda_2, \lambda_3$  的  $k$  次齐次多项式, 反之亦然。

5. 由  $\lambda_1 + \lambda_2 + \lambda_3 = 1$  知这三个变量线性相关。如果取独立变量  $\lambda_1, \lambda_2$ , 则:

$$\begin{cases} \frac{\partial}{\partial \lambda_1} = (x_1 - x_3) \frac{\partial}{\partial x} + (y_1 - y_3) \frac{\partial}{\partial y}, \\ \frac{\partial}{\partial \lambda_2} = (x_2 - x_3) \frac{\partial}{\partial x} + (y_2 - y_3) \frac{\partial}{\partial y}. \end{cases} \quad (4.21)$$

6. 重心坐标和直角坐标变换的 Jacobi 行列式为:

$$\det \left( \frac{\partial(x, y)}{\partial(\lambda_1, \lambda_2)} \right) = 2m(\Delta_e). \quad (4.22)$$

$m(\Delta_e)$  为三角形单元的面积。

**练习 4.1** 证明重心坐标的上述性质。

重心坐标对简化单元上的积分计算具有重要的价值, 由于重心坐标把一般的三角形映射到直角边长为 1 的等腰直角三角形, 我们可以把一般积分区域上的求积分问题转化为三角形区域上求积分的问题。记  $e$  为三角单元,  $\hat{e}$  为变换后的等腰直角三角形, 具体推导如下:

$$\begin{aligned} \int \int_e F(x, y) dx dy &= \int \int_{\hat{e}} F(\lambda_1, \lambda_2) \det \frac{\partial(x, y)}{\partial(\lambda_1, \lambda_2)} d\lambda_1 d\lambda_2 \\ &= 2m(\Delta_e) \int_0^1 d\lambda_1 \int_0^{1-\lambda_1} F(\lambda_1, \lambda_2, 1 - \lambda_1 - \lambda_2) d\lambda_2. \end{aligned} \quad (4.23)$$

下面的积分等式在有限元计算中是重要的:

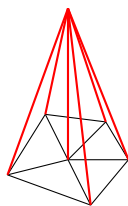
**命题 4.4**

$$\int \int_e \lambda_1^m \lambda_2^n \lambda_3^k dx dy = 2m(\Delta_e) \frac{m!n!k!}{(m+n+k+2)!}. \quad (4.24)$$

**练习 4.2** 证明(4.24)。

以分片线性多项式元为例, 在实际计算中, 插值基函数选取为: 在其中一个顶点取值为 1, 其余顶点取值均为 0 的分片线性函数。其形式与一维情形的 hat-function 比较类似, 示意图如下:

图 4.4: 二维分片线性插值基函数



## 4.2.2 三角单元内分段多项式的形式

本节中, 我们借助重心坐标, 在有合适插值条件的前提下, 写出每个三角单元内部分片多项式的表达式。如无特殊说明, 本节中单元  $e$  指的是三角形  $\Delta A_1A_2A_3$ 。



### 4.2.2.1 线性 Lagrange 元

二元线性多项式的自由度为 3。要在单元  $e$  上确认一个线性二元多项式，需要三个插值条件，考虑在单元  $e$  的三个顶点处给出：

$$u(A_i) = u_i, i = 1, 2, 3. \quad (4.25)$$

由重心坐标的性质，即

$$\lambda_i(A_j) = \delta_{ij}, \quad (4.26)$$

可得：

$$p(x, y) = \sum_{i=1}^3 u_i \lambda_i(x, y). \quad (4.27)$$

这就是重心坐标下线性 Lagrange 元的表达式。

### 4.2.2.2 完全二次 Lagrange 元

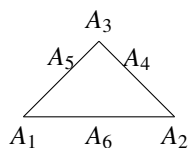
$e$  上的完全二次多项式可以表示为：

$$p(x, y) = a_{0,0} + a_{1,0}x + a_{2,0}x^2 + a_{0,1}y + a_{1,1}xy + a_{0,2}y^2. \quad (4.28)$$

由表达式知共有 6 个自由变量，需要 6 个插值条件。在实际计算中，我们取三角形三个顶点以及三条边的中点作为插值节点，问题转化为求一组基函数使得  $u_i(A_j) = \delta_{ij}$ 。

插值节点如图所示：

图 4.5: 插值节点示意



首先考虑构造二次多项式  $p(\lambda_1, \lambda_2, \lambda_3)$ ，使得在顶点  $A_1$  上的取值为 1，在其他顶点和线段中点处取值均为 0。由于  $\lambda_1(A_1) = 1, \lambda_1(A_i) = 0 (i \neq 1)$ ，可以取  $\lambda_1$  为  $p$  的一个因子。由于  $A_5A_6$  平行于  $A_2A_3$ ，且为中点连线，故该线段的表达式为  $\lambda_1 = \frac{1}{2}$ 。从而  $(2\lambda_1 - 1)$  也是  $p$  的因子。此时取基函数  $p(\lambda_1, \lambda_2, \lambda_3) = \lambda_1(2\lambda_1 - 1)$ ，经检验可知该多项式在  $A_1$  取 1，在  $A_i (i \neq 1)$  均取 0。

下面考虑如何选取多项式使得中点处取值为 1，其余为 0。以  $A_1A_2$  的中点为例，在该点处， $\lambda_1 = \lambda_2 = \frac{1}{2}, \lambda_3 = 0$ 。故一个比较合理的取法是  $p(x, y) = 4\lambda_1\lambda_2$ 。经验证知，该多项式满足基函数的要求。


因此，完全二次 Lagrange 元的表达式可以写为：

$$p_2(x, y) = u_1\lambda_1(2\lambda_1 - 1) + u_2\lambda_2(2\lambda_2 - 1) + u_3\lambda_3(2\lambda_3 - 1) + 4u_6\lambda_1\lambda_2 + 4u_4\lambda_2\lambda_3 + 4u_5\lambda_3\lambda_1. \quad (4.29)$$

### 4.2.2.3 完全三次 Lagrange 元

完全三次 Lagrange 元需要 10 个插值节点，这里取三角形三个顶点，三条边上的三等分点以及三角形的重心。类似完全二次 Lagrange 元的讨论，完全三次 Lagrange 元的表达式可以写成：

$$p(x, y) = \sum_{i=1}^3 u_i \frac{\lambda_i(3\lambda_i - 1)(3\lambda_i - 2)}{2} + \sum_{i \neq j} \frac{9}{2} \lambda_i \lambda_j (3\lambda_i - 1) u_{ij} + 27\lambda_1\lambda_2\lambda_3 u_{123}. \quad (4.30)$$

 **练习 4.3** 证明(4.30)中选取的函数确实构成一组插值基底。

#### 4.2.2.4 不完全三次 Lagrange 元

在介绍不完全 Lagrange 元之前，先给出**升幂形式**的定义。

##### 定义 4.6 (升幂形式)

对于定义在三角元上的三次多项式，按重心坐标我们可以写成如下形式：

$$p_3(x, y) = \alpha_1 \lambda_1^3 + \alpha_2 \lambda_2^3 + \alpha_3 \lambda_3^3 + \alpha_4 \lambda_1^2 \lambda_2 + \alpha_5 \lambda_1^2 \lambda_3 + \alpha_6 \lambda_1 \lambda_2^2 + \alpha_7 \lambda_3 \lambda_2^2 + \alpha_8 \lambda_1 \lambda_3^2 + \alpha_9 \lambda_2 \lambda_3^2 + \alpha_{10} \lambda_1 \lambda_2 \lambda_3. \quad (4.31)$$

这是一个三次齐次多项式。由于  $\lambda_1 + \lambda_2 + \lambda_3 = 1$ ，将(4.31)进行变形，可得下面的升幂形式：

$$p_3(x, y) = a_1 \lambda_1 + a_2 \lambda_2 + a_3 \lambda_3 + a_4 \lambda_1 \lambda_2 + a_5 \lambda_2 \lambda_3 + a_6 \lambda_3 \lambda_1 + a_7 \lambda_1^2 \lambda_2 + a_8 \lambda_2^2 \lambda_3 + a_9 \lambda_3^2 \lambda_1 + a_{10} \lambda_1 \lambda_2 \lambda_3. \quad (4.32)$$

对于**不完全三次 Lagrange 元**，我们给的明确插值条件只有三个顶点和三条边的三分点，一共 9 个方程，对于 10 个未知数的  $p_3(x, y)$  并不是 **poised lattice**，所以需要有一个额外的条件。事实上，我们列这 9 个方程时，对于(4.32)， $a_{10}$  前的系数  $\lambda_1 \lambda_2 \lambda_3$  总是 0。对于这个问题，额外的插值条件可以取为：

##### 定义 4.7 (不完全三次 Lagrange 元)

$p_3(x, y)$  针对二次多项式的插值误差为 0。

也就是说， $a_7 = a_8 = a_9 = 0$  时， $a_{10}$  也应当为 0。这种情况下我们可以取

$$a_{10} = c_1 a_7 + c_2 a_8 + c_3 a_9. \quad (4.33)$$

这种取法满足前提条件。事实上，由于  $c_i$  可以是任意实数，不完全三次 Lagrange 元并不唯一。我们取  $c_1 = c_2 = c_3 = \frac{1}{2}$ ，就给出了线性系统的第十个方程，从而保证不完全三次 Lagrange 元的存在性。

#### 4.2.2.5 完全三次 Hermite 元

TBD

### 4.2.3 矩形单元内分段多项式的形式

相比于三角单元，矩形单元上分段多项式的形式稍微好写一些，但在应用角度上并不是主流。究其原因，主要是对于一般区域，三角划分总是容易的，但矩形划分受限于特殊的几何结构并不容易进行。本节中我们主要讨论矩形单元上的分片多项式形式。由于一般矩形  $[a_1, a_2] \times [b_1, b_2]$  总是能通过仿射变换变为标准区域  $[-1, 1] \times [-1, 1]$ ，我们只需要考虑矩形区域  $C := [-1, 1] \times [-1, 1]$  上的多项式。

#### 4.2.3.1 双线性元

对于双线性元，我们选取插值节点为矩形的四个顶点，即：

$$A_1(1, 1), A_2(-1, 1), A_3(-1, -1), A_4(1, -1). \quad (4.34)$$

若双线性多项式  $p$  满足  $p(A_i) = u_i$ ，我们可以写：

$$p(x, y) = \sum_{i=1}^4 u_i \psi_i(x, y). \quad (4.35)$$

其中  $\psi_i$  为 Lagrange 插值基函数，满足：

$$\psi_i(A_j) = \delta_{ij}. \quad (4.36)$$

根据(4.36)和(4.34)，可以写出  $\psi_i$  的具体表达式如下：

$$\begin{aligned}\psi_1(x, y) &= \frac{1}{4}(1+x)(1+y), \\ \psi_2(x, y) &= \frac{1}{4}(1-x)(1+y), \\ \psi_3(x, y) &= \frac{1}{4}(1-x)(1-y), \\ \psi_4(x, y) &= \frac{1}{4}(1+x)(1-y).\end{aligned}\tag{4.37}$$

据此，我们给出了矩形区域上的双线性插值。

#### 4.2.3.2 双二次元

对矩形区域  $\hat{K} = [-1, 1] \times [-1, 1]$ ，要构造完全双二次型，需要 9 个插值节点。插值节点选取为 4 个顶点，4 边中点以及正方形的中心节点，具体为：

$$\begin{aligned}A_1(1, 1), A_2(-1, 1), A_3(-1, -1), A_4(1, -1) \\ A_5(0, 1), A_6(-1, 0), A_7(0, -1), A_8(1, 0), A_9(0, 0).\end{aligned}\tag{4.38}$$

插值基函数依旧要求满足  $\psi_i(A_j) = \delta_{ij}$ ，形式为：

$$\begin{aligned}\psi_1(x, y) &= \frac{1}{4}x(1+x)y(1+y), \psi_2(x, y) = \frac{1}{4}x(1-x)y(1+y), \psi_3(x, y) = \frac{1}{4}x(1-x)y(1-y), \\ \psi_4(x, y) &= \frac{1}{4}x(1+x)y(1-y), \psi_5(x, y) = \frac{1}{2}(1-x^2)y(1+y), \psi_6(x, y) = -\frac{1}{2}x(1-x)(1-y^2), \\ \psi_7(x, y) &= -\frac{1}{2}(1-x^2)y(1-y), \psi_8(x, y) = \frac{1}{2}x(1+x)(1-y^2), \psi_9(x, y) = (1-x^2)(1-y^2).\end{aligned}\tag{4.39}$$

#### 4.2.3.3 奇妙八点元

这种插值节点取法为四个顶点 + 四条边的中点，与之相应的是，插值多项式也少了一项。奇妙八点元的插值多项式是：

$$q_8(x, y) = \text{span}\{x^i y^j\} \setminus \{x^2 y^2\}, 0 \leq i, j \leq 2.\tag{4.40}$$


相比于双二次元，奇妙八点元的特征是，其表达式为二元三次多项式，并没有包含四次项  $x^2 y^2$ 。

插值节点为：

$$\begin{aligned}A_1(1, 1), A_2(-1, 1), A_3(-1, -1), A_4(1, -1) \\ A_5(0, 1), A_6(-1, 0), A_7(0, -1), A_8(1, 0).\end{aligned}\tag{4.41}$$

插值基函数要求  $\psi_i(A_j) = \delta_{ij}$ ，形式为：

$$\begin{aligned}\psi_1(x, y) &= \frac{1}{4}(x+1)(y+1)(x+y-1), \psi_2(x, y) = \frac{1}{4}(1-x)(1+y)(y-x-1), \\ \psi_3(x, y) &= -\frac{1}{4}(x-1)(y-1)(x+y+1), \psi_4(x, y) = \frac{1}{4}(y-1)(x+1)(y-x+1), \\ \psi_5(x, y) &= \frac{1}{2}(1-x^2)(1+y), \psi_6(x, y) = \frac{1}{2}(1-y^2)(1-x), \\ \psi_7(x, y) &= \frac{1}{2}(1-x^2)(1-y), \psi_8(x, y) = \frac{1}{2}(1-y^2)(1+x).\end{aligned}\tag{4.42}$$

 **练习 4.4** 推导奇妙八点元的表达式。

### 4.3 有限元求解方程示例

本节中，通过一个例子展示有限元方法求解二维 Poisson 方程的过程。

**定义 4.8 (二维 Poisson 方程)**

本节中考虑的 2 维 Poisson 方程形式为:

$$\begin{cases} -\Delta u = f(x), x \in \Omega, \\ u|_{\Gamma_1} = 0, \\ \frac{\partial u}{\partial \nu}|_{\Gamma_2} = 0. \end{cases} \quad (4.43)$$

其中  $\Omega = (0, 1) \times (0, 1)$ ,  $\Gamma_1 := \{(0, y) : 0 \leq y \leq 1\} \cup \{(x, 0) : 0 \leq x \leq 1\}$ ,  $\Gamma_2 := \{(1, y) : 0 \leq y \leq 1\} \cup \{(x, 1) : 0 \leq x \leq 1\}$ ,  $f \in L^2(\Omega)$ .

回忆一维有限元问题的步骤: 变分形式-区域划分-建立有限元空间-刚度矩阵与负载向量-处理边界条件-有限元方程组求解。对于二维 Poisson 方程问题, 我们依旧可以按照上面的步骤进行。

**4.3.1 构造变分形式**

在(4.43)第 1 式左右两边同乘辅助函数  $v$ , 并在区域  $\Omega$  上求积分, 可得:

$$\int_{\Omega} -\Delta u v dx = \int_{\Omega} f v dx. \quad (4.44)$$

对(4.44)的左端项使用格林公式, 可得:

$$-\int_{\Omega} \Delta u v dx = \int_{\Omega} \nabla u \cdot \nabla v dx - \int_{\partial\Omega} v \frac{\partial u}{\partial n} ds(x). \quad (4.45)$$

由于  $\frac{\partial u}{\partial n}|_{\Gamma_2} = 0$ , 取函数空间  $V = \{v : v|_{\Gamma_1} = 0\}$ , 可知  $\forall v \in V$ , 均有

$$\int_{\Omega} \nabla u \cdot \nabla v = \int_{\Omega} f v dx. \quad (4.46)$$

取  $B(u, v) = \int_{\Omega} \nabla u \cdot \nabla v dx$ ,  $\langle f, v \rangle = \int_{\Omega} f v dx$ , 可得变分问题的表示形式为:

$$B(u, v) = \langle f, v \rangle, \forall v \in V. \quad (4.47)$$

**4.3.2 进行区域剖分**

本例中问题区域  $\Omega$  为矩形区域, 因此三角剖分和矩形剖分均可实现有限元离散和求解。在本例中, 我们采用三角剖分来分割矩形区域  $\Omega$ , 具体的离散方案如下图所示:

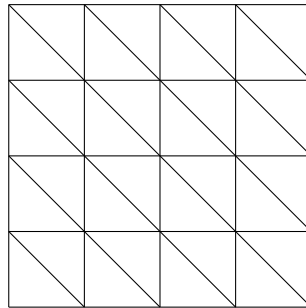


图 4.6: 矩形区域三角剖分

在本例中, 我们采用每个单元上的分段线性函数作有限元近似。根据分段线性函数的性质, 我们只需要确定单元交点处的取值, 即可给出分段线性函数的表达式。

值得注意的是, 并非所有节点处的值都是未知的。在本例中, 由于我们在  $\Gamma_1$  处给出了 Dirichlet 边值条件, 矩形区域左边界和下边界的取值都是确定的, 从而这些点是已知取值的节点。我们关心的是取值未知的节点, 在上图中一共有 16 个点, 这意味着我们需要 16 个离散方程。

在列方程前，我们需要对剖分后的区域和剖分节点进行编号。一个合理的编号方案可以大大降低最后一步矩阵求解的时间复杂度。下图为一种常用的标注方案：

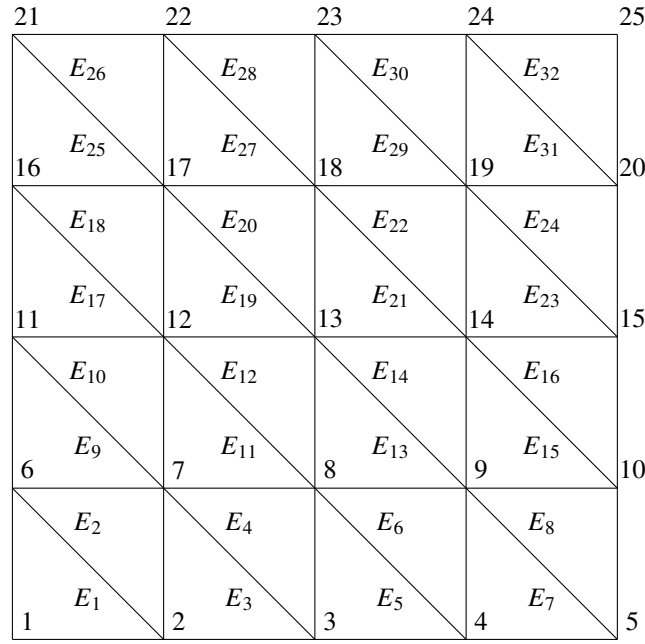


图 4.7: 三角单元符号标注

事实上，在上图中，节点 1, 2, 3, 4, 5, 6, 11, 16, 21 处的值都是已知的，剩余的节点才是需要求解的部分。为后续列方程方便，我们可以记录一下与每个节点相邻的三角单元的编号。

### 4.3.3 建立有限元子空间 $V_h$

一阶有限元子空间由 4.4 张成，记  $\Psi_i(x)$  为满足下列条件的基函数：

1.  $\Psi_i(P_j) = \delta_{ij}$ 。
2.  $\Psi_i$  为线性函数。

根据边界条件，最终的一阶有限元解是节点 7, 8, 9, 10, 12, 13, 14, 15, 17, 18, 19, 20, 22, 23, 24, 25 上定义的  $\Psi_i$  的线性组合。

### 4.3.4 刚度矩阵与负载向量计算

根据(4.47)，在有限元子空间  $V_h$  上，我们需要寻找函数  $v \in V_h$  使得  $B(v, \Psi_i) = \langle f, \Psi_i \rangle \forall i$ 。设

$$v = \sum_{i=1}^N u_i \Psi_i, \quad (4.48)$$

可知

$$\sum_{i=1}^N B(\Psi_i, \Psi_j) u_i = \langle f, \Psi_j \rangle, \forall j \in [1, N]. \quad (4.49)$$

要求解函数  $v$ ，即求解系数  $u_i$ ，需要计算刚度矩阵  $B(\Psi_i, \Psi_j)$  和负载向量  $\langle f, \Psi_j \rangle$ 。

计算刚度矩阵时，我们采用的做法是在每个单元上计算其对刚度矩阵的贡献，最后整体装配到刚度矩阵中。对于一维有限元问题，装配过程的作用有限，因为整个矩阵的形式足够清晰。但在二维甚至更高维度的问题中，装配在写程序的时候就相当必要了。

以图 4.7 的  $E_{11}$  为例，说明一下刚度矩阵装配的过程。

1. 明确  $E_{11}$  落在哪些基函数的支集内。

在图中,  $E_{11}$  落在其三个顶点对应的基函数的支集内部, 即  $\Psi_7, \Psi_8, \Psi_{12}$ 。

2. 设下标集  $I = \{7, 8, 12\}$ , 通过双线性泛函  $B$  的定义计算  $B(\Psi_i, \Psi_j), i, j \in I$ 。
3. 将第二步计算所得子矩阵装配到刚度矩阵  $M$  内, 所谓“装配”即为把计算结果与装配前的刚度矩阵作加法。

第二步中对  $B(\Psi_i, \Psi_j)$  的计算需要使用重心坐标进行简化。

对于负载向量  $\langle f, \Psi_j \rangle$  的计算, 则是在注意  $\Psi_j$  支集的前提下, 设计三角形上的数值积分公式计算即可。

最后一步, 则是求解得到的线性方程, 并给出近似解的表达式。

 **练习 4.5** 实现二维 Poisson 方程的一阶有限元算法。

## 4.4 本章总结

在前三章基础上, 本章开始了本课程的主题—有限元算法的原理与实践。实践中需要特别注意:

- 变分问题的选取。
- 区域的三角划分。
- 插值多项式空间  $V_h$  的形式。
- 刚度矩阵  $B$  的装配。
- 负载向量的计算。
- 计算完成后计算结果与原问题数值解的对应。

至于该方法的误差分析, 详情请见下一个章节。

## 第5章 协调有限元的误差分析

本章中，我们将借助 Sobolev 空间以及其上的范数，对协调有限元的求解误差进行估计。

### 5.1 误差估计的整体流程

经过协调有限元离散后，原微分方程边值问题对应的变分问题为：

#### 定义 5.1 (微分方程弱形式)

求  $u \in V$  使得

$$a(u, v) = f(v), \forall v \in V. \quad (5.1)$$

其中  $V$  是定义在  $\Omega$  上的函数的 Hilbert 空间，双线性型  $a(\cdot, \cdot)$  是连续且  $V$ -椭圆的， $f$  是线性连续泛函。

设  $V_h \leq V$  是有限元子空间，则有限元逼近问题为

#### 定义 5.2 (有限元逼近问题)

求  $u_h \in V_h$  使得

$$a(u_h, v_h) = f(v_h). \quad (5.2)$$

直接讨论有限元逼近问题的解不太容易，讨论对应的插值问题相对容易一些。和一维有限元分析类似，高维情况下也有对应的 Cea 引理。

#### 引理 5.1 (Cea 引理)

设双线性型  $a(\cdot, \cdot)$  是连续且  $V$ -椭圆的，则存在  $C > 0$  使得：

$$\|u - u_h\|_V \leq C \inf_{v_h \in V_h} \|u - v_h\|_V, \quad (5.3)$$

其中  $u, u_h$  分别为准确解和有限元解。

**证明** 由于  $a(\cdot, \cdot)$  连续且  $V$ -椭圆，可知双线性型  $a$  可以在空间  $V$  上诱导一个内积。由问题 5.1 和 5.2 的描述可知： $\forall v_h \in V_h, u \in V$ ，有  $a(u - u_h, v_h) = 0$  成立。即向量  $u - u_h$  与子空间  $V_h$  关于内积  $a(\cdot, \cdot)$  正交。由此，我们有：

$$\|u - u_h\|_V^2 \leq \frac{1}{\alpha} |a(u - u_h, u - u_h)| = \frac{1}{\alpha} |a(u - u_h, u - v_h)| \leq \frac{M}{\alpha} \|u - u_h\|_V \|u - v_h\|_V. \quad (5.4)$$

上式中，第一个不等号由  $V$ -椭圆性导出，第二个等号源于  $u - u_h$  与子空间  $V_h$  的正交性，第三个不等号源于双线性型的有界性。

由此可得：

$$\|u - u_h\|_V \leq \frac{M}{\alpha} \|u - v_h\|_V. \quad (5.5)$$

**练习 5.1** 试证明在 Cea 引理中常数  $C$  可以优化为  $\sqrt{\frac{M}{\alpha}}$ 。

**注** Cea 引理的重要意义在于将有限元的误差估计问题归结于插值误差估计问题。事实上，假定  $\mathcal{T}_h$  为空间  $\Omega$  的有限元划分， $\pi_h u$  代表对函数  $u$  的样条插值， $\pi_T u$  代表对函数  $u$  在单元  $T$  上的插值，那么我们有：

$$\inf_{v_h \in V_h} \|u - v_h\|_V \leq \|u - \pi_h u\|_V = \left( \sum_{T \in \mathcal{T}_h} \|u - \pi_T u\|_V^2 \right)^{\frac{1}{2}}. \quad (5.6)$$

由此，问题归结于估计 (三角) 单元  $T$  上的插值误差  $\|u - \pi_T u\|_{1,T}$ 。

单元  $T$  的任意性可能会使得插值误差不好计算，所以我们需要考虑一个仿射变换  $F_T : \hat{T} \rightarrow T$ ，把标准参考单元  $\hat{T} : \{(x, y) : x \geq 0, y \geq 0, x + y \leq 1\}$  变换为我们需要讨论的三角单元  $T$ ，而后把  $T$  上的误差估计转化到  $\hat{T}$  上完成。



根据  $F_T$ , 我们可以定义下面两个映射:

$$\hat{v}(\hat{x}) := v(F(\hat{x})) = v(F_T(\hat{x})) = v(x). \quad (5.7)$$

$$\hat{\pi}_{\hat{T}} \hat{u}(\hat{x}) := \pi_T u(F(\hat{x})) = \pi_T u(x). \quad (5.8)$$

据此我们可以写出误差估计的全流程:

1. 把有限元解的误差根据 Cea 引理转化为插值误差。
2. 把整体的插值误差估计转化为每个三角单元  $T$  上的插值误差估计。
3. 将三角单元  $T$  上的插值误差转化为标准单元  $\hat{T}$  上的插值误差。
4. 利用等价范数定理建立  $\hat{T}$  上的插值误差估计。
5. 将标准单元  $\hat{T}$  上的范数  $|\hat{u}|_{k+1, \hat{T}}$  转化到一般单元上的范数  $|u|_{k+1, T}$ 。

## 5.2 Sobolev 空间上的插值误差估计

### 5.2.1 仿射等价元之间的范数关系

#### 定义 5.3 (仿射等价)

称  $\mathbb{R}^n$  中的两个开子集  $\Omega, \hat{\Omega}$  是仿射等价的, 如果存在可逆的仿射变换

$$F : F(\hat{x}) = B\hat{x} + b = x \in \Omega, \forall \hat{x} \in \hat{\Omega}. \quad (5.9)$$

使得  $\Omega = F(\hat{\Omega})$ 。

#### 定理 5.1

设  $\Omega$  和  $\hat{\Omega}$  仿射等价, 若  $v \in W^{m,p}(\Omega)$ , 令

$$\hat{v}(\hat{x}) := v(F(\hat{x})) = v(x), \quad (5.10)$$

则  $\hat{v} \in W^{m,p}(\hat{\Omega})$  且:

$$|\hat{v}|_{m,p,\hat{\Omega}} \leq C \|B\|^m |\det B|^{-\frac{1}{p}} |v|_{m,p,\Omega}. \quad (5.11)$$

其中  $C$  为仅与  $m, n$  有关的正常数,  $\|B\|$  为矩阵  $B$  的 Euclid 范数。类似地, 下面这个等式也成立:

$$|v|_{m,p,\Omega} \leq C \|B^{-1}\|^m |\det B|^{\frac{1}{p}} |\hat{v}|_{m,p,\hat{\Omega}}. \quad (5.12)$$

**证明** 先考虑  $v \in C^m$ 。设  $1 \leq p < \infty$ , 对  $v \in C^m(\bar{\Omega})$ , 根据  $\hat{v}$  的定义则有  $\hat{v} \in C^m(\bar{\hat{\Omega}})$ , 且对任何满足  $|\alpha| = m$  的多重指标  $\alpha = (\alpha_1, \dots, \alpha_n)$ , 有:

$$D^\alpha \hat{v}(\hat{x}) = \nabla^m \hat{v}(\hat{x})(\xi_1, \dots, \xi_m). \quad (5.13)$$

这里  $\nabla^m \hat{v}$  是  $m$  阶张量,  $\xi_i$  表示  $\mathbb{R}^m$  内的基向量。因此我们有:

$$|\hat{v}|_{m,p,\hat{\Omega}} \leq C_1(m, n) \left( \int_{\hat{\Omega}} \|\nabla^m \hat{v}(\hat{x})\|^p d\hat{x} \right)^{\frac{1}{p}} \quad (5.14)$$

下面讨论如何将(5.14)中的变量  $\hat{x}$  转化为  $\Omega$  内的变量  $x$ 。利用复合函数的微分法则:

$$\nabla^m \hat{v}(\hat{x})(\xi_1, \dots, \xi_m) = \nabla^m v(x)(B\xi_1, \dots, B\xi_m). \quad (5.15)$$

我们有:

$$\|\nabla^m \hat{v}(\hat{x})\| \leq \|\nabla^m v(x)\| \cdot \|B\|^m = \|\nabla^m v(F(\hat{x}))\| \cdot \|B\|^m. \quad (5.16)$$

将上式代入(5.14)右侧的表达式, 可知:

$$\int_{\hat{\Omega}} \|\nabla^m \hat{v}(\hat{x})\|^p d\hat{x} \leq \int_{\hat{\Omega}} \|\nabla^m v(F(\hat{x}))\|^p \cdot \|B\|^{mp} d\hat{x} = |\det(B^{-1})| \|B\|^{mp} \int_{\Omega} \|\nabla^m v(x)\|^p dx. \quad (5.17)$$

又由于有界性, 可得存在常数  $C_2(m, n)$  使得:

$$\|D^m v(x)\| \leq C_2(m, n) \max_{|\alpha|=m} |D^\alpha v(x)|. \quad (5.18)$$

综合上述, 可得:

$$|\hat{v}|_{m,p,\hat{\Omega}} \leq C(m, n) \|B\|^m |\det B|^{-\frac{1}{p}} |v|_{m,p,\Omega}. \quad (5.19)$$

在  $p < +\infty$  的情况下, 利用  $C^m(\bar{\Omega})$  在  $H^{m,p}(\Omega)$  中的稠密性可知, 上述不等式对  $v \in H^{m,p}(\Omega)$  也成立。

如果  $p = +\infty$ , 对于  $\forall v \in H^{m,\infty}(\Omega)$ , 有  $v \in H^{m,p}(\Omega) \forall p < \infty$ . 根据前面的结论,  $|D^\alpha \hat{v}|_{0,p,\hat{\Omega}}$  的上界与  $p$  无关, 因此函数  $D^\alpha \hat{v} \in L^\infty(\hat{\Omega})$ . 由  $\alpha$  任意性可得  $\hat{v} \in H^{m,\infty}(\hat{\Omega})$ . 再利用嵌入定理即证。

**练习 5.2** 用类似的方法证明等式(5.12)。

**注** 上面的定理说明了如果  $\Omega$  和  $\hat{\Omega}$  仿射等价, 则半范数  $|\cdot|_{m,p,\hat{\Omega}}$  和半范数  $|\cdot|_{m,p,\Omega}$  为等价半范数。

定理 5.1 给出了两个范数之间的等价关系。具体的倍率则需要由  $\|B\|$  和  $|\det(B)|$  来进行控制。而由于矩阵  $B$  刻画了  $\Omega$  和  $\hat{\Omega}$  几何特征的差异。为了估计  $\|B\|$  和  $\det(B)$ , 引入下列记号来描述这两个区域的几何信息:

- $h$ :  $\Omega$  的直径。
- $\hat{h}$ :  $\hat{\Omega}$  的直径。
- $\rho$ :  $\Omega$  最大内接球的直径。
- $\hat{\rho}$ :  $\hat{\Omega}$  最大内接球的直径。

对于  $\|B\|$  和  $\det(B)$  的控制, 有下面这些结论成立:

#### 引理 5.2

$$\|B\| \leq \frac{h}{\hat{\rho}}, \|B^{-1}\| \leq \frac{\hat{h}}{\rho}. \quad (5.20)$$

#### 引理 5.3

$$\det(B) = \frac{|\Omega|}{|\hat{\Omega}|}. \quad (5.21)$$

#### 引理 5.4

设  $\sigma_n$  是  $\mathbb{R}^n$  中单位球的体积,  $\hat{\sigma}_n = C \frac{\sigma_n}{m(\hat{\Omega})}$ , 则:

$$C_1 \hat{\sigma}_n \rho^n \leq |\det(B)| \leq C_2 \hat{\sigma}_n h^n. \quad (5.22)$$

## 5.2.2 单元上的插值误差估计

下面的定理给出的是全空间  $\Omega$  上的插值误差估计。

#### 定理 5.2

设  $W^{k+1,p}(\hat{\Omega})$  和  $W^{m,q}(\hat{\Omega})$  是两个 Sobolev 空间,  $0 \leq m \leq k+1$ , 满足

$$W^{k+1,p}(\hat{\Omega}) \hookrightarrow W^{m,q}(\hat{\Omega}). \quad (5.23)$$

又设  $\pi_h$  是定义在  $W^{k+1,p}(\hat{\Omega})$  上的有界线性算子, 且保持  $k$  次多项式不变。又设  $\Omega$  和  $\hat{\Omega}$  仿射等价, 且

$$\widehat{\pi_h v} = \pi_h \hat{v}. \quad (5.24)$$

则存在正常数  $C$  使得下面的不等式成立:

$$|v - \pi_h v|_{m,q,\Omega} \leq C [m(\Omega)]^{\frac{1}{q} - \frac{1}{p}} \frac{h^{k+1}}{\rho^m} |v|_{k+1,p,\Omega}, v \in W^{k+1,p}(\Omega). \quad (5.25)$$

**证明** 根据定理 5.1:

$$|v - \pi_h v|_{m,q,\Omega} \leq C(m, n) \|B^{-1}\|^m |\det(B)|^{\frac{1}{q}} |\hat{v} - \widehat{\pi_h v}|_{m,q,\hat{\Omega}}. \quad (5.26)$$

由上式, 我们只需要估计  $|\hat{v} - \widehat{(\pi_h v)}|$ 。由题设条件, 我们知道对任意的  $\hat{p}_k \in P_k(\hat{\Omega})$ , 有:

$$\hat{v} - \widehat{(\pi_h v)} = \hat{v} - \hat{\pi}_h \hat{v} = (I - \hat{\pi}_h)(\hat{v} + \hat{p}_k). \quad (5.27)$$

从而根据等价范数定理,

$$|\hat{v} - \widehat{(\pi_h v)}|_{m,q,\hat{\Omega}} \leq C |\hat{v}|_{k+1,p,\hat{\Omega}}. \quad (5.28)$$

再利用前述引理 5.2 和引理 5.4, 可得待证等式成立。

**注** 该定理利用  $v$  在  $W^{k+1,p}(\Omega)$  上的范数, 估计了插值在  $W^{m,q}(\Omega)$  上的误差。

下面我们给出一般单元  $e \in \mathcal{T}_h$  上的插值误差估计。

#### 定理 5.3

如果下面几个条件成立:

- $\hat{v} \in H^{k+1,p}(\hat{e})$
- $\hat{\pi}_{\hat{e}} : H^{k+1,p}(\hat{e}) \rightarrow H^{m,q}(\hat{e})$ , 满足

$$\hat{\pi}_{\hat{e}} \hat{p}_k = \hat{p}_k, \forall \hat{p}_k \in \hat{P}_k(\hat{e}). \quad (5.29)$$

- 对整数  $k \geq 0, m \geq 0$ , 有下列嵌入关系:

$$\begin{aligned} H^{k+1,p}(\hat{e}) &\hookrightarrow \bar{C}^s(\hat{e}), s = 0, 1, 2 \\ H^{k+1,p}(\hat{e}) &\hookrightarrow H^{m,q}(\hat{e}). \end{aligned} \quad (5.30)$$

- 对任意单元  $e$ , 存在可逆变换  $F_e$

$$x = F_e(\hat{x}) = B_e \hat{x} + b. \quad (5.31)$$

使  $e = F(\hat{e})$ .

则:

$$|v - \pi_e v|_{m,q,e} \leq C [m(e)]^{\frac{1}{q} - \frac{1}{p}} \frac{h_e^{k+1}}{\rho_e^m} |v|_{k+1,p,e}, v \in W^{k+1,p}(e). \quad (5.32)$$

**证明** 和定理 5.2 同理。

如果剖分  $\mathcal{T}_h$  是拟正则的, 即存在常数  $\sigma > 0$  使得

$$\frac{h_e}{\rho_e} \leq \sigma, \forall e \in \mathcal{T}_h, h_e > 0, h_e \rightarrow 0. \quad (5.33)$$

那么根据定理 5.3, 存在正常数  $C$  使得下面的插值误差估计成立:

$$|v - \Pi_e v|_{m,q,e} \leq C h_e^{k+1-m+n(\frac{1}{q}-\frac{1}{p})} |v|_{k+1,p,e}. \quad (5.34)$$

根据(5.34)中的局部误差估计, 推广到整体, 可以得到以下推论:

#### 推论 5.1

假设定理 5.3 的条件满足, 且剖分是拟正则的,  $h$  为所有单元的最大直径, 则:

$$|v - \Pi_h v|_{m,p,\Omega} \leq C h^{k+1-m} |v|_{k+1,p,\Omega}, \forall v \in W^{k+1,p}(\Omega). \quad (5.35)$$

**注** 可以把上面的推论 5.1 和一维样条插值的误差估计结果进行一些比对。

## 5.3 多边形区域上二阶问题的分析

下面我们具体讨论二阶问题的误差估计, 即区域  $\Omega \in \mathbb{R}^2$  上的变分问题。设这个二阶问题的解空间  $V$  满足:  $H_0^1(\Omega) \subset V \subset H^1(\Omega)$ 。以下是下面的误差分析成立需要满足的一系列前提条件:

- 区域划分  $\mathcal{T}_h$  是拟正则的。
- 有限元族  $\{e, P_e, \Sigma_e\}_{e \in \mathcal{T}_h, h>0}$  与一个参考元  $(\hat{e}, \hat{P}_{\hat{e}}, \hat{\Sigma}_{\hat{e}})$  仿射等价。

- $\{(e, P_e, \Sigma_e)\}$  属于  $C^0$  类, 即为协调元。

**注** 第一个条件保证了区域划分不会出现一些“过于尖锐的”三角形。第二个条件则是我们利用参考元族研究有限元族的依据。第三个条件则是利用插值估计有限元误差的基本依据。

### 5.3.1 解的 $H^1$ 模误差估计

对于  $H^1$  模的误差估计, 我们可以使用 Cea 引理, 把有限元解的误差分析转化为插值的误差分析。在前提条件满足的情况下, 有下面的定理成立:

#### 定理 5.4

如果  $P_k(\hat{e}) \subset \hat{P}_{\hat{e}} \subset H^l(\hat{e}), H^{k+1}(\hat{e}) \hookrightarrow C^s(\hat{e})$ , 其中  $0 \leq l \leq k$ ,  $s$  是  $\hat{\Sigma}_{\hat{e}}$  中出现的最高阶导数, 则对任意  $v \in H^{k+1}(\Omega) \cap V$ , 有:

$$\begin{aligned} \|v - \Pi_h v\|_{m,\Omega} &\leq Ch^{k+1-m} |v|_{k+1,\Omega}, 0 \leq m \leq \min\{1, l\}, \\ \left( \sum_{e \in \mathcal{T}_h} \|v - \Pi_e v\|_{m,e}^2 \right)^{\frac{1}{2}} &\leq Ch^{k+1-m} |v|_{k+1,\Omega}, 0 \leq m \leq l. \end{aligned} \quad (5.36)$$

其中  $\Pi_h v$  表示  $v$  在  $V_h$  中的分片插值。

**证明** 在表达式(5.34)中, 取  $p = q = 2$  可得:

$$\|v - \Pi_e v\|_{m,e} \leq Ch_e^{k+1-m} |v|_{k+1,e}. \quad (5.37)$$

由于  $h$  是单元直径的最大值, 我们有  $h_e \leq h$ , 由上式求和可得:

$$\left( \sum_{e \in \mathcal{T}_h} \|v - \Pi_e v\|_{m,e}^2 \right)^{\frac{1}{2}} \leq Ch^{k+1-m} \left( \sum_{e \in \mathcal{T}_h} |v|_{k+1,e}^2 \right)^{\frac{1}{2}} = Ch^{k+1-m} |v|_{k+1,\Omega}. \quad (5.38)$$

而当  $0 \leq m \leq \min\{1, l\}$  时, (5.36)的第二式左边可以写成  $\|v - \Pi_h v\|_{m,\Omega}$ , 从而第一式也可得证。

在定理 5.4 中取  $l = 1$ , 再借助 Cea 引理, 可以得到  $H^1$  范数的估计式:

#### 定理 5.5

如果  $v \in H^{k+1}(\Omega) \cap V$  且定理 5.4 的条件成立, 那么我们有下面的估计式成立:

$$\|u - u_h\|_{1,\Omega} \leq Ch^k |u|_{k+1,\Omega}. \quad (5.39)$$

**注** 定理 5.5 说明了子空间  $V$  选取  $p$  次元确实意味着有限元解的  $H^1$  误差能达到  $p$  阶精度。但  $p$  也并非越大越好。一方面高次元的计算复杂度也会大幅度上升, 另一方面, 定理 5.5 的估计式需要真实解  $u$  也满足  $u \in H^{k+1}(\Omega)$ 。

## 5.4 $L^2$ 模与负模的估计

### 5.4.1 $L^2$ 模估计

上面我们利用插值误差和 Cea 引理, 估计了在  $H^1(\Omega)$  空间范数意义下的相对误差阶数。有时候我们需要对有限元解在其他范数意义下的误差进行估计, 比较常用的范数是  $L^2(\Omega)$ 。这时我们不能直接用插值误差导出  $L^2$  误差, 因为没有对应的 Cea 引理, 这时我们必须要建立  $u - u_h$  的  $L^2$  误差与  $H^1$  误差的关系。下面是对应的结论, 这是一维情形下 Aubin-Nitsche 技巧的推广。

**定理 5.6 (Aubin-Nitsche)**

设  $a(\cdot, \cdot)$  是连续且 V-椭圆的双线性型, 变分问题: 求  $u \in V$  使得  $a(u, v) = \langle f, v \rangle \forall v$  在空间  $V_h \leq V$  上的有限元解为  $u_h$ , 那么存在一个正常数  $C$  使得下列不等式成立:

$$\|u - u_h\|_{0,\Omega} \leq C \|u - u_h\|_{1,\Omega} \sup_{g \in L^2(\Omega)} \left\{ \frac{1}{\|g\|_{0,\Omega}} \inf_{v_h \in V_h} \|\Phi - v_h\|_{1,\Omega} \right\}. \quad (5.40)$$

其中  $\Phi$  为共轭问题

$$a(\Phi, v) = \langle g, v \rangle \forall v \in V_h \quad (5.41)$$

的解。

**证明** 由  $L^2(\Omega)$  范数的定义, 我们有

$$\|u - u_h\|_{0,\Omega} = \sup_{g \in L^2(\Omega)} \frac{|\langle u - u_h, g \rangle|}{\|g\|_{0,\Omega}}. \quad (5.42)$$

下面我们需要估计  $|\langle u - u_h, g \rangle|$ 。根据 Lax-Milgram 引理, 设  $V = L^2(\Omega)$ , 存在唯一的  $\Phi \in V$  使得:

$$a(\Phi, u - u_h) = \langle g, u - u_h \rangle. \quad (5.43)$$

由投影定理,  $\forall v_h \in V_h$ ,  $a(v_h, u - u_h) = 0$ 。从而:

$$\langle g, u - u_h \rangle = a(\Phi - v_h, u - u_h). \quad (5.44)$$

综合上述等式, 可得:

$$|\langle u - u_h, g \rangle| = |a(u - u_h, \Phi - v_h)| \leq C \|u - u_h\|_1 \|\Phi - v_h\|_1 \leq C \|u - u_h\|_1 \inf_{v_h \in V_h} \|\Phi - v_h\|_1. \quad (5.45)$$

根据(5.42)和(5.45), Aubin-Nitsche 不等式成立。

如果我们讨论的偏微分方程为一个二阶椭圆问题, 那么其变分形式的解满足下面的正则性估计:

**定理 5.7**

设椭圆方程

$$\mathcal{L}u := - \sum_{i,j=1}^2 \partial_i(a_{ij}(\mathbf{x})\partial_j u) + a_0(\mathbf{x})u = f(\mathbf{x}), \mathbf{x} \in \Omega. \quad (5.46)$$

其中  $\Omega$  是凸多边形区域, 其系数满足一致椭圆性条件。在赋予一定的边值条件后, 该问题的变分形式为:

$$a(u, v) = \langle f, v \rangle \forall v \in V. \quad (5.47)$$

其有限元变分形式为:

$$a(u_h, v) = \langle f, v \rangle, \forall v \in V_h. \quad (5.48)$$

那么, 变分问题的解  $u$  和共轭问题的解  $\Phi$  满足:

$$\|u\|_{2,\Omega} \leq C \|f\|_{0,\Omega}, \|\Phi\|_{2,\Omega} \leq C \|g\|_{0,\Omega}. \quad (5.49)$$

根据定理 5.4, 5.6 和 5.7, 我们有下面的结论成立:

**定理 5.8**

$$\|u - u_h\|_{0,\Omega} \leq Ch \|u - u_h\|_{1,\Omega}. \quad (5.50)$$

**证明** 利用不等式

$$\|\Phi - \Phi_h\|_{0,\Omega} \leq Ch \|\Phi\|_{2,\Omega} \leq Ch \|g\|_{0,\Omega}. \quad (5.51)$$

结合 Aubin-Nitsche 定理即可得证。

### 5.4.2 负模估计

对于负模，类似于 Aubin-Nitsche 定理，我们有如下结论成立：

#### 定理 5.9

设  $u$  和  $u_h$  分别为  $\Omega$  上椭圆方程的广义解和有限元离散解，那么存在一个正常数  $C$ ，使得下列不等式成立：

$$\|u - u_h\|_{-l, \Omega} \leq C \|u - u_h\|_{1, \Omega} \sup_{\psi \in H_0^l(\Omega)} \left\{ \frac{1}{\|\psi\|_{l, \Omega}} \inf_{v_h \in V_h} \|\Psi - v_h\|_{1, \Omega} \right\}. \quad (5.52)$$

其中  $\Psi$  为共轭问题

$$a(v, \Psi) = \langle \psi, v \rangle, \forall v \in V \quad (5.53)$$

的解。

**证明** 根据定义，我们有：

$$\|u - u_h\|_{-l, \Omega} = \sup_{\psi \in H_0^l(\Omega)} \frac{|\langle u - u_h, \psi \rangle|}{\|\psi\|_{l, \Omega}}. \quad (5.54)$$

令  $\Psi_h$  为 (5.53) 的有限元解，根据方程 (5.53) 和  $u - u_h$  对空间  $V_h$  的正交性，我们可以推知：

$$\begin{aligned} \langle u - u_h, \psi \rangle &= a(u - u_h, \Psi) \\ &= a(u - u_h, \Psi - \Psi_h) \\ &\leq M \|u - u_h\|_1 \|\Psi - \Psi_h\|_1 \\ &\leq C \|u - u_h\|_1 \inf_{v_h \in V_h} \|\Psi - v_h\|_1. \end{aligned} \quad (5.55)$$

将 (5.55) 代入 (5.54)，即可得 (5.52) 成立。

对于 (5.53)，有下面的正则性估计成立：

$$\|\Psi\|_{l+2, \Omega} \leq C \|\psi\|_{l, \Omega}. \quad (5.56)$$

**练习 5.3** 利用定理 5.9 和正则性估计证明下面的误差估计式：

$$\|u - u_h\|_{-l, \Omega} \leq Ch^{l+1} \|u - u_h\|_{1, \Omega}. \quad (5.57)$$

## 5.5 逆不等式

下面我们把关注点转向有限元空间  $V_h$  本身。我们知道，对于  $H_0^1(\Omega)$  上的函数，我们可以用高阶模来估计低阶模，即 Poincare-Friedrichs 不等式：

$$\|u\|_{0, \Omega} \leq C \|u\|_{1, \Omega}, \forall u \in H_0^1(\Omega). \quad (5.58)$$

但对于一般的函数，我们不能用低阶模来估计高阶模，例如一致有界的函数列，其导数却完全可以无界。

然而，有限元空间作为一个特殊的函数空间，我们可以在其上用低阶模来估计高阶模。这就是本节中需要讨论的逆不等式。

### 5.5.1 单元上的逆不等式

#### 定理 5.10

设  $\mathcal{T}_h$  是  $\Omega$  的拟正则剖分， $P(e)$  是定义在单元  $e \in \mathcal{T}_h$  上的多项式空间，那么下面的逆不等式成立：

$$\begin{aligned} |p|_{1, e} &\leq Ch_e^{-1} \|p\|_{0, e}, \forall p \in P(e), e \in \mathcal{T}_h, \\ \|p\|_{0, \infty, e} &\leq Ch_e^{-\frac{n}{2}} \|p\|_{0, e}, \forall p \in P(e), e \in \mathcal{T}_h. \end{aligned} \quad (5.59)$$

**证明** 令  $\hat{e}$  是标准三角单元, 并存在可逆仿射变换

$$F: \hat{e} \rightarrow e, F(\hat{x}) = B\hat{x} + b = x \in e, \forall \hat{x} \in \hat{e}. \quad (5.60)$$

那么: 根据定理 5.1, 有:

$$|p|_{1,e} \leq C \|B^{-1}\| |\det(B)|^{\frac{1}{2}} |\hat{p}|_{1,\hat{e}}. \quad (5.61)$$

又, 由于  $\|\hat{p}\|_{0,\hat{e}}$  和  $\|\hat{p}\|_{1,\hat{e}}$  均为  $P(\hat{e})$  上的范数, 且在有限元空间上, 因此这两个范数等价, 即:

$$\|\hat{p}\|_{1,\hat{e}} \leq \hat{C} \|\hat{p}\|_{0,\hat{e}}. \quad (5.62)$$

于是:

$$\begin{aligned} \|p\|_{1,e} &\leq C \|B^{-1}\| |\det(B)|^{\frac{1}{2}} \|\hat{p}\|_{0,\hat{e}} \\ &\leq C \|B^{-1}\| |\det(B)|^{\frac{1}{2}} |\det(B)|^{-\frac{1}{2}} \|p\|_{0,e} \\ &\leq C \rho_e^{-1} \|p\|_{0,e} \\ &\leq C h_e^{-1} \|p\|_{0,e}. \end{aligned} \quad (5.63)$$

对于第二式, 由定理 5.1 可得

$$\|p\|_{0,\infty,e} \leq C \|\hat{p}\|_{0,\infty,\hat{e}}, \quad (5.64)$$

而在  $P(\hat{e})$  上,  $\|\hat{p}\|_{0,\infty,\hat{e}}$  与  $\|\hat{p}\|_{0,\hat{e}}$  等价, 从而:

$$\|\hat{p}\|_{0,\infty,\hat{e}} \leq \hat{C} \|\hat{p}\|_{0,\hat{e}}. \quad (5.65)$$

于是:

$$\|p\|_{0,\infty,e} \leq C \|\hat{p}\|_{0,\hat{e}} \leq C |\det(B)|^{-\frac{1}{2}} \|p\|_{0,e} \leq C h_e^{-\frac{n}{2}} \|p\|_{0,e}. \quad (5.66)$$

## 5.5.2 逆不等式

### 定义 5.4 (拟一致条件)

存在正常数  $\nu$ , 使得:

$$\frac{h}{h_e} \leq \nu, \forall e \in \mathcal{T}_h, h \rightarrow 0. \quad (5.67)$$

其中  $h_e = \text{diam}(e)$ ,  $h = \max_{e \in \mathcal{T}_h} \{h_e\}$ .

下面叙述整体区域下的逆不等式。

### 定理 5.11 (逆不等式)

设  $\Omega \subset \mathbb{R}^n$  为有界凸多边形开区域, 设剖分  $\mathcal{T}_h$  满足正则性条件, 仿射等价条件和拟一致条件, 且:

$$\hat{P}_{\hat{e}} \subset H^{l,r}(\hat{e}) \cap H^{m,q}(\hat{e}), 0 \leq l \leq m, 1 \leq r, q \leq \infty. \quad (5.68)$$

则存在常数使得对任意  $v_h \in V_h$ , 有:

$$\left( \sum_{e \in \mathcal{T}_h} |v_h|_{m,q,e}^q \right)^{\frac{1}{q}} \leq \frac{C}{(h^n)^{\max\{0, \frac{1}{r} - \frac{1}{q}\}} h^{m-l}} \left( \sum_{e \in \mathcal{T}_h} |v_h|_{l,r,e}^r \right)^{\frac{1}{r}}. \quad (5.69)$$

其中分片多项式空间:

$$V_h := \{v_h : v_h|_e \in P_e \leftrightarrow \hat{P}_{\hat{e}} \subset H^{l,r}(\hat{e}) \cap H^{m,q}(\hat{e})\}. \quad (5.70)$$

特别地, 取  $\hat{P}_{\hat{e}} \subset H^1(\hat{e})$ , 且  $V_h \subset C(\bar{\Omega}) \cap H^1(\Omega)$ , 则:

$$\begin{aligned} |v_h|_{0,\infty,\Omega} &\leq C h^{-\frac{n}{2}} |v_h|_{0,\Omega}, \forall v_h \in V_h \\ |v_h|_{1,\Omega} &\leq C h^{-1} |v_h|_{0,\Omega}, \forall v_h \in V_h. \end{aligned} \quad (5.71)$$



证明在本笔记中按下不表。

最后，作为本节的结束，我们利用逆不等式证明两个模估计定理。

#### 定理 5.12 ( $H^s$ 模估计)

设  $u \in H^m \cap H_0^1(\Omega)$  为变分问题的解， $u_h \in V_h$  为有限元逼近问题的解，那么有下面的最佳误差估计：

$$\|u - u_h\|_{s,\Omega} \leq Ch^{m-s} \|u\|_{m,\Omega}, 0 \leq s \leq m-1. \quad (5.72)$$

**证明** 我们只证明  $1 \leq s \leq m-1$  的情形。根据定理 5.11，取  $m = s$ ， $r = q = 2$ ， $l = 1$ ，对任意  $v_h \in V_h$ ，有：

$$|v_h|_{s,\Omega} = \left( \sum_{e \in \mathcal{T}_h} |v_h|_{s,e}^2 \right)^{\frac{1}{2}} \leq Ch^{1-s} \left( \sum_{e \in \mathcal{T}_h} |v_h|_{1,e}^2 \right)^{\frac{1}{2}} \leq Ch^{1-s} |v_h|_{1,\Omega}, s \geq 1. \quad (5.73)$$

于是：

$$\begin{aligned} \|u - u_h\|_{s,\Omega} &\leq \|u - \Pi_h u\|_{s,\Omega} + \|u_h - \Pi_h u\|_{s,\Omega} \\ &\leq \|u - \Pi_h u\|_{s,\Omega} + Ch^{1-s} \|u_h - \Pi_h u\|_{1,\Omega} \\ &\leq \|u - \Pi_h u\|_{s,\Omega} + Ch^{1-s} (\|u - \Pi_h u\|_{1,\Omega} + \|u - u_h\|_{1,\Omega}). \end{aligned} \quad (5.74)$$

利用插值逼近定理，可知

$$\begin{aligned} \|u - \Pi_h u\|_{s,\Omega} &\leq C_1 h^{m-s} \|u\|_{m,\Omega}, \\ \|u - \Pi_h u\|_{1,\Omega} &\leq C_2 h^{m-1} \|u\|_{m,\Omega}. \end{aligned} \quad (5.75)$$

由  $H^1$  模估计可知：

$$\|u - u_h\|_{1,\Omega} \leq C_3 h^{m-1} \|u\|_{m,\Omega}. \quad (5.76)$$

综上所述，

$$\|u - u_h\|_{s,\Omega} \leq Ch^{m-s} \|u\|_{m,\Omega}. \quad (5.77)$$

#### 定理 5.13 (最大模估计)

设  $\Omega$  的剖分是拟正则的， $|\alpha| = s < m-1$ ，有限元解  $u_h$  具有如定理 5.12 所述的估计，则  $u_h$  有下面的最大模估计：

$$\max_{\Omega} |D^\alpha(u - u_h)| \leq Ch^{m-1-s} \|u\|_{m,\Omega}. \quad (5.78)$$

**证明** 设  $\Pi_h u$  是  $u$  的  $m-1$  次插值函数，在单元  $e_0$  上达到  $\max_{\Omega} |D^\alpha(u - u_h)|$ ，则：

$$\max_{\Omega} |D^\alpha(u - u_h)| \leq \max_{\Omega} |D^\alpha(u - \Pi_h u)| + \max_{\Omega} |D^\alpha(\Pi_h u - u_h)|. \quad (5.79)$$

由插值误差分析公式可知：

$$\max_{\Omega} |D^\alpha(u - \Pi_h u)| = \max_{e_0} |D^\alpha(u - \Pi_h u)| \leq Ch^{m-1-s} \|u\|_{m,e_0} \leq Ch^{m-1-s} \|u\|_{m,\Omega}. \quad (5.80)$$

令  $v_h = \Pi_h u - u_h$ ，则由逆不等式，有：

$$|v_h|_{s,\Omega} = \left( \sum_{e \in \mathcal{T}_h} |v_h|_{s,e}^2 \right)^{\frac{1}{2}} \leq Ch^{-s} \left( \sum_{e \in \mathcal{T}_h} |v_h|_{0,e}^2 \right)^{\frac{1}{2}} \leq Ch^{-s} |v_h|_{0,\Omega}. \quad (5.81)$$

再由逆不等式，有

$$\max_{\Omega} |D^\alpha v_h| \leq Ch^{m-1-s} \|u\|_{m,\Omega}. \quad (5.82)$$

综上，最大模估计式得证。

## 5.6 本章小结

第四章在一维情形的基础上，给出了高维情形下求解椭圆方程的有限元算法，而本章则是对第四章所述算法进行了误差分析。

事实上，误差分析的基本流程和一维情形大致相同，都是先把有限元逼近的误差问题利用 Cea 原理转换为插值逼近的误差问题，再分析得到误差范数的最优估计表达式。但高维问题的 Sobolev 范数计算有其复杂性，这导致了对高维问题插值误差估计比一维情形更困难一些。

高维问题中的几何与拓扑比一维问题更复杂，在上一章节中我们利用三角剖分来逼近一个一般区域。问题即转换为在一个三角形区域上的范数估计。为了方便进行积分的运算，我们考虑利用仿射变换把一个一般的三角区域转化为以  $O$  点为中心的等腰直角三角形。

本章中的定理 5.1 是一个比较核心的结论，它体现了高维问题中函数  $v$  与它的插值函数  $\hat{v}$  之间范数的关系。借此即可得到单元上以及全空间上在 Sobolev 范数意义下插值误差的估计。

如果我们想讨论其他模长意义下 (如  $L^2$ ) 的误差估计，由于没有对应的 Cea 引理，我们不能直接利用插值误差来估计  $L^2$  误差，此时我们需要推导高维情形下的 Aubin-Nitsche 不等式。

一般来说，我们不能用低阶模来估计高阶模，但在有限元空间下可以，这就是所谓**逆不等式**。借助逆不等式，我们可以在给出  $H^m(\Omega)$  模下误差估计的情形下，针对  $s < m$ ，给出  $H^s(\Omega)$  模下的误差估计。

## 第6章 抛物型方程的有限元求解与分析

前面的章节中我们主要讨论的是椭圆方程的有限元解法以及其求解误差的分析。在本章中，我们转而关注抛物型方程的有限元求解与分析。在抛物型方程的有限差分算法 (MOL) 中，我们的思路是先对空间进行离散形成一个常微分方程组，再对生成的常微分方程组进行时间积分求解。有限元求解抛物型方程的过程中也可以借鉴这一思路，即先固定时间参数  $t$ ，在空间上给出变分问题的形式，然后转化为一个常微分方程组，最后通过时间积分解决问题。

本章中，如无特殊说明，讨论的微分方程均为如下二维热传导方程：

$$\begin{cases} u_t - \Delta u = f(\mathbf{x}, t), (\mathbf{x}, t) \in \Omega \times [0, +\infty), \\ u(\mathbf{x}, t) = 0, (\mathbf{x}, t) \in \partial\Omega \times [0, +\infty), \\ u(\mathbf{x}, 0) = u_0(\mathbf{x}), \mathbf{x} \in \Omega. \end{cases} \quad (6.1)$$

在方程(6.1)中，要求  $\Omega \subset \mathbb{R}^2$  为有界区域，而  $\Gamma := \partial\Omega$  为  $\Omega$  的光滑边界。

### 6.1 空间离散 (半离散)

本节着重介绍有限元求解抛物型方程的半离散算法，以及半离散解的  $L^2(\Omega)$  范数估计。

#### 6.1.1 离散过程

第四章里我们讨论了有限元求解微分方程的一般流程：

1. 寻求原问题的变分形式。
2. 对区域  $\Omega$  进行剖分。
3. 构造有限元子空间  $V_h$ 。
4. 利用问题的变分形式建立有限元方程组。
5. 对离散后的有限元方程组 (本质上是一个线性方程组) 进行求解。

本节中，我们用同样的流程来分析热传导方程(6.1)。

设  $v \in H_0^1(\Omega)$  为测试函数， $t$  看作一个给定的常数，在(6.1)的第(1)式两边同乘  $v$ ，在  $\Omega$  上积分得：

$$\int_{\Omega} (u_t v - v \Delta u) d\mathbf{x} = \int_{\Omega} f(\mathbf{x}, t) v(\mathbf{x}) d\mathbf{x}. \quad (6.2)$$

由于  $v|_{\partial\Omega} \equiv 0$ ，利用格林公式可得：

$$\int_{\Omega} u_t v d\mathbf{x} + \int_{\Omega} \nabla u \cdot \nabla v d\mathbf{x} = \int_{\Omega} f v d\mathbf{x}. \quad (6.3)$$

我们定义以下符号：

$$B(u, v) := \int_{\Omega} \nabla u \cdot \nabla v d\mathbf{x}, \langle f, v \rangle := \int_{\Omega} f(\mathbf{x}, t) v(\mathbf{x}) d\mathbf{x}. \quad (6.4)$$

由此可得热传导方程的变分形式：

#### 定义 6.1

热传导方程的变分形式为：对每个固定的  $t \in [0, +\infty)$ ，求  $u(t) \in H_0^1(\Omega)$  使得：

$$\begin{cases} \langle u_t, v \rangle + B(u, v) = \langle f, v \rangle, \forall v \in H_0^1(\Omega) \\ u(\mathbf{x}, 0) = u_0(\mathbf{x}) \end{cases} \quad (6.5)$$

对  $\Omega$  的剖分可以参考椭圆方程求解时的剖分方式，即三角剖分  $\mathcal{T}_h$ 。记  $\Omega_h = \cup_{e \in \mathcal{T}_h} e$ ，其中  $h$  为三角剖分的最大直径。我们可以假定  $\mathcal{T}_h$  是拟一致的划分，此时每个单元  $e$  的面积有下界  $Ch^2$ 。

由剖分  $\Omega_h$ , 我们可以构造  $H_0^1(\Omega)$  的有限元子空间  $V_h$  如下:

$$V_h := \{v_h | v_h \in C(\bar{\Omega}), v_h|_e = p_k(\mathbf{x}), v_h(\mathbf{x}) = 0 \text{ if } \mathbf{x} \notin \Omega_h\}. \quad (6.6)$$

根据如上所述的区间划分和有限元子空间  $V_h$  定义, 可得(6.1)的半离散变分问题:

### 定义 6.2

对每个固定的  $t \in [0, +\infty)$ , 求  $u_h(t) \in V_h$  使得:

$$\begin{cases} \langle u_{h,t}, v \rangle + B(u_h, v) = \langle f, v \rangle, \forall v \in H_0^1(\Omega) \\ u_h(\mathbf{x}, 0) = u_{0h}(\mathbf{x}) \end{cases} \quad (6.7)$$

(6.7)的解称为变分形式6.1的**半离散解**。

下面讨论如何给出  $u_h(t)$  满足的常微分方程。设空间  $V_h$  有一组基函数为  $\{\varphi_i(\mathbf{x})\}$ , 对任意  $u_h(t) \in V_h$  进行 Fourier 展开, 得:

$$u_h(t) = \sum_{i=1}^N \alpha_i(t) \varphi_i(\mathbf{x}). \quad (6.8)$$

此时半离散问题(6.7)可以转化为以下常微分方程组:

$$\begin{cases} \sum_{i=1}^N \langle \varphi_i, \varphi_j \rangle \alpha_i'(t) + \sum_{i=1}^N B(\varphi_i, \varphi_j) \alpha_i(t) = \langle f, \varphi_j \rangle. \\ \alpha_i(0) = \gamma_i. \end{cases} \quad (6.9)$$

其中  $\gamma_i$  为初值  $u_{0h}(\mathbf{x})$  关于基底  $\{\varphi_i(\mathbf{x})\}$  展开的 Fourier 系数。如果引入下列矩阵记号:

$$\begin{aligned} \mathbf{A} &:= (\langle \varphi_i, \varphi_j \rangle) \\ \mathbf{B} &:= (B(\varphi_i, \varphi_j)), \alpha(t) := (\alpha_1(t), \dots, \alpha_N(t))^T \\ F(t) &:= (\langle f, \varphi_1 \rangle, \dots, \langle f, \varphi_N \rangle)^T, \gamma := (\gamma_1, \dots, \gamma_N)^T, \end{aligned} \quad (6.10)$$

那么方程组(6.9)可以改写为:

$$\begin{cases} \mathbf{A} \alpha'(t) + \mathbf{B} \alpha(t) = \mathbf{F}, \\ \alpha(0) = \gamma. \end{cases} \quad (6.11)$$

其中  $\mathbf{A}$  被称为**质量矩阵**,  $\mathbf{B}$  被称为**刚度矩阵**。由 ode 理论可知, 该微分方程存在唯一解。

## 6.1.2 半离散解的误差

本节中我们估计上述半离散解的求解误差。在分析时, 我们假定求解常微分方程的过程不引入任何误差, 即(6.9)的解是精确的。

在分析之前, 我们假定有限元空间  $V_h$  有下面的逼近性质:

### 命题 6.1

$$\begin{aligned} \inf_{\chi \in V_h} \{ \|v - \chi\|_{0,\Omega} + h \|\nabla(v - \chi)\|_{0,\Omega} \} &\leq C h^s \|v\|_{s,\Omega}, \\ 1 \leq s \leq k+1, k \geq 1, \forall v &\in H^s(\Omega) \cap H_0^1(\Omega). \end{aligned} \quad (6.12)$$

**注** 上面的性质体现了有限元空间  $V_h$  确实可以较好地逼近函数空间  $H^s(\Omega) \cap H_0^1(\Omega)$ 。

在讨论半离散问题有限元解的  $L^2(\Omega)$ -模估计前, 我们先引入一个**椭圆投影算子**。

**定义 6.3 (椭圆投影算子)**

椭圆投影算子  $P_h : H_0^1(\Omega) \rightarrow V_h$  对任意  $u \in H_0^1(\Omega)$ , 满足:

$$B(P_h u, \chi) = B(u, \chi), \forall \chi \in V_h. \quad (6.13)$$

**注** 椭圆投影算子可以看作在以  $B(\cdot, \cdot)$  为内积的 Hilbert 空间下, 函数空间  $H_0^1(\Omega)$  到空间  $V_h$  的投影。

下面讨论椭圆投影算子的性质。对于一个满足逼近性质的有限元空间  $V_h$ , 有下面的引理成立:

**引理 6.1**

设  $v \in H^s(\Omega) \cap H_0^1(\Omega)$ ,  $1 \leq s \leq k+1$ , 则:

$$\|P_h v - v\|_{0,\Omega} + h \|\nabla(P_h v - v)\|_{0,\Omega} \leq Ch^s \|v\|_{s,\Omega}. \quad (6.14)$$

**注** 上面引理保证了当  $V_h$  满足逼近性质时,  $P_h v$  是对  $v$  的一个好的逼近。

**证明** 证明分两步, 先估计  $\|\nabla(P_h v - v)\|_{0,\Omega}$ , 再估计  $\|P_h v - v\|_{0,\Omega}$ 。根据投影算子的性质, 我们有如下不等式成立:

$$\begin{aligned} \|\nabla(P_h v - v)\|_{0,\Omega}^2 &= \langle \nabla(P_h v - v), \nabla(P_h v - v) \rangle \\ &= \langle \nabla(P_h v - v), \nabla(\chi - v) \rangle \quad \forall \chi \in V_h \\ &\leq \|\nabla(P_h v - v)\|_{0,\Omega} \|\nabla(\chi - v)\|_{0,\Omega} \end{aligned} \quad (6.15)$$

从而我们有  $\|\nabla(P_h v - v)\|_{0,\Omega} \leq \|\nabla(\chi - v)\|_{0,\Omega}$ 。又由逼近性质, 有:

$$\|\nabla(P_h v - v)\|_{0,\Omega} \leq \inf_{\chi \in V_h} \|\nabla(v - \chi)\|_{0,\Omega} \leq Ch^{s-1} \|v\|_{s,\Omega}. \quad (6.16)$$

对于  $\|P_h v - v\|$  的估计, 我们没法直接使用投影的性质, 因此我们需要利用一些技巧来提升  $(P_h v - v)$  的导数阶数。根据第二格林公式, 我们可以试着讨论  $P_h v - v$  与一个标量场  $\Delta w$  的内积, 来得到  $\|P_h v - v\|$  的估计。而要使得  $-\Delta w = P_h v - v$ , 这就导出了一个椭圆方程。

因此我们需要讨论如下辅助问题。设  $g \in L^2(\Omega)$ , 考虑下面的椭圆方程:

$$\begin{cases} -\Delta w = g, \\ w|_{\partial\Omega} = 0. \end{cases} \quad (6.17)$$

根据椭圆方程的适定性和 Poincare-Friedriches 不等式, 该椭圆方程存在唯一解  $w$ , 且满足不等式

$$\|w\|_{2,\Omega} \leq C \|\Delta w\|_{0,\Omega} = C \|g\|_{0,\Omega}. \quad (6.18)$$

于是:

$$\begin{aligned} \langle P_h v - v, g \rangle &= -\langle P_h v - v, \Delta w \rangle \\ &= \langle \nabla(P_h v - v), \nabla w \rangle \\ &= \langle \nabla(P_h v - v), \nabla(w - P_h w) \rangle \\ &\leq \|P_h v - v\|_{0,\Omega} \|\nabla(w - P_h w)\|_{0,\Omega} \\ &\leq Ch^s \|v\|_{s,\Omega} \|w\|_{2,\Omega} \leq Ch^s \|v\|_{s,\Omega} \|g\|_{0,\Omega}. \end{aligned} \quad (6.19)$$

特别地, 取  $g = P_h v - v$ , 可得  $\|P_h v - v\|_{0,\Omega} \leq Ch^s \|v\|_{s,\Omega}$ 。

综上, 引理成立。

下面的定理利用投影算子, 讨论了有限元解  $u_h$  与真实解  $u$  的误差的  $L^2$ -模估计。

**定理 6.1 ( $L^2$ -模估计)**

设  $u(\mathbf{x}, t)$  为(6.1)的解,  $u_h(t)$  为其在  $V_h$  上的有限元近似解, 且  $u(\mathbf{x}, t) \in H^{k+1}(\Omega) \cap H_0^1(\Omega)$ , 则有:

$$\|u_h(t) - u(\mathbf{x}, t)\|_{0,\Omega} \leq \|u_{0h} - u_0\|_{0,\Omega} + Ch^{k+1} \{ \|u_0\|_{k+1,\Omega} + \int_0^t \|u_t\|_{k+1,\Omega} ds \}. \quad (6.20)$$

**注** 抛物方程有限元半离散解的误差由两部分组成，第一部分是估计初值时产生的误差，第二部分是演化过程中的误差。由上面的定理可以看出，如果需要对抛物方程的有限元半离散解达到四阶精度，则至少需要三次多项式单元。

**证明** 证明分三步完成。

第一步：利用投影算子将误差分割为两部分。

$$u_h(t) - u(\mathbf{x}, t) = (u_h(t) - P_h u) + (P_h u - u) := \theta(t) + \rho(t). \quad (6.21)$$

$\rho(t)$  表示将真实解投影到解空间  $V_h$  上产生的误差，而  $\theta(t)$  则表示在解空间上近似  $P_h u$  产生的误差。

第二步：估计  $\rho(t)$ 。

利用引理6.1，取  $s = k + 1$ ，我们有：

$$\begin{aligned} \|\rho(t)\|_{0,\Omega} &= \|P_h u - u\|_{0,\Omega} \\ &\leq Ch^{k+1} \|u\|_{k+1,\Omega} \\ &\leq Ch^{k+1} \|u_0\|_{k+1,\Omega} + \int_0^t \|u_t\|_{k+1,\Omega} dt \\ &\leq Ch^{k+1} \left\{ \|u_0\|_{k+1,\Omega} + \int_0^t \|u_t\|_{k+1,\Omega} dt \right\}. \end{aligned} \quad (6.22)$$

第三步：估计  $\theta(t)$ 。首先我们需要推导  $\theta$  满足的方程。

$\forall \chi \in V_h$ ，我们有：

$$\begin{aligned} \langle \theta_t, \chi \rangle + \langle \nabla \theta, \nabla \chi \rangle &= \langle u_{h,t}, \chi \rangle - \langle P_h u_t, \chi \rangle + \langle \nabla u_h, \nabla \chi \rangle - \langle \nabla P_h u, \nabla \chi \rangle \\ &= \langle f, \chi \rangle - \langle P_h u_t, \chi \rangle - \langle \nabla u, \nabla \chi \rangle \\ &= \langle u_t - P_h u_t, \chi \rangle = \langle -\rho_t, \chi \rangle. \end{aligned} \quad (6.23)$$

根据投影算子的性质， $\theta(t) \in V_h$ ，那么我们可以取  $\chi = \theta$ ，得：

$$\langle \theta_t, \theta \rangle + \|\nabla \theta\|_{0,\Omega}^2 = -\langle \rho_t, \theta \rangle. \quad (6.24)$$

根据上式，可知：

$$\frac{1}{2} \frac{d}{dt} \|\theta\|_{0,\Omega}^2 \leq \|\rho_t\|_{0,\Omega} \|\theta\|_{0,\Omega}. \quad (6.25)$$

从而：

$$\frac{d}{dt} \|\theta\|_{0,\Omega} \leq \|\rho_t\|_{0,\Omega} \quad (6.26)$$

上式两边同时从 0 到  $t$  积分，可得：

$$\|\theta(t)\|_{0,\Omega} \leq \|\theta(0)\|_{0,\Omega} + \int_0^t \|\rho_t\|_{0,\Omega} ds. \quad (6.27)$$

又：

$$\begin{aligned} \|\theta(0)\|_{0,\Omega} &= \|u_h(0) - u(\mathbf{x}, 0) + u(\mathbf{x}, 0) - P_h u(\mathbf{x}, 0)\|_{0,\Omega} \\ &\leq \|u_{0h} - u_0\|_{0,\Omega} + \|u_0 - P_h u_0\|_{0,\Omega} \\ &\leq \|u_{0h} - u_0\|_{0,\Omega} + Ch^{k+1} \|u_0\|_{k+1,\Omega}. \end{aligned} \quad (6.28)$$

并且

$$\|\rho_t\|_{0,\Omega} \leq Ch^{k+1} \|u_t\|_{k+1,\Omega}. \quad (6.29)$$

综合上述，结论得证。

如果初值近似也为  $k + 1$  阶的，那么  $L^2$ -模估计可以简化为：

$$\|u_h(t) - u(\mathbf{x}, t)\|_{0,\Omega} \leq Ch^{k+1} \left\{ \|u_0\|_{k+1,\Omega} + \int_0^t \|u_t\|_{k+1,\Omega} ds \right\}. \quad (6.30)$$

我们可以用类似的技巧证明下面的梯度估计定理：

**定理 6.2 (梯度估计)**

设  $u$  为(6.1)的解,  $u_h$  为其有限元解, 且  $u \in H^{k+1}(\Omega) \cap H_0^1(\Omega)$ , 则:

$$\|\nabla u_h - \nabla u\|_{0,\Omega} \leq \|\nabla u_{0h} - \nabla u_0\|_{0,\Omega} + Ch^k \left\{ \|u_0\|_{k+1,\Omega} + \|u(t)\|_{k+1,\Omega} + \left( \int_0^t \|u_t\|_{k,\Omega}^2 ds \right)^{\frac{1}{2}} \right\}. \quad (6.31)$$



**练习 6.1** 完成定理 6.2 的证明。

## 6.2 全离散与误差估计

上面一节中我们讨论了针对热方程进行半离散思路, 并分析了半离散解的  $L^2$  误差。然而, 半离散后导出的常微分方程(6.9)往往无法直接求得解析解, 而是需要对应的 ode 数值算法求解。这意味着我们同样需要针对时间参数  $t$  对系统进行离散化处理。在本节中, 我们将讨论两种不同的时间积分格式, 对应的是 Euler-Galerkin 全离散方法和 Crank-Nicolson-Galerkin 全离散方法, 并分析这两种算法的误差阶数。

### 6.2.1 Euler-Galerkin 方法

和数值求解常微分方程初值问题的隐式 Euler 方法类似, 在 Euler-Galerkin 算法中, 我们利用向后差商来近似一阶导数, 即:

$$\bar{\partial}_t U^m \approx \frac{U^m - U^{m-1}}{\Delta t}. \quad (6.32)$$

将向后差商的表达式代入半离散方程(6.7), 可得 Euler-Galerkin 全离散格式如下:

$$\begin{cases} \langle U^m, \chi \rangle + \tau B(U^m, \chi) = \langle U^{m-1} + \tau f(t_m), \chi \rangle, \forall \chi \in V_h \\ U^0 = u_{0h} \end{cases} \quad (6.33)$$

**注** 此处的  $U^m$  指  $t = t_m$  时刻解  $u$  的近似, 其中  $U^i \in V_h$ ,  $\tau$  指时间间隔, 即  $t_{i+1} - t_i = \tau$ 。实际计算中, 利用全离散格式, 一定能求解出数值近似解, 因为如果  $U^{m-1}$  已知, 等式的右端项可以计算, 只需要利用待定系数法求解左端分片多项式的系数即可给出  $u(t_m)$  的估计。

**注** 一个疑问: 为什么课本上介绍的时间积分方法都采用隐式格式?

下面讨论该算法的  $L^2$  误差估计。

**定理 6.3**

设  $u(\mathbf{x}, t)$  为(6.1)的解, 而  $U^m$  为全离散方程(6.33)的解, 那么求解误差有如下上界估计:

$$\|U^m - u(t_m)\|_{0,\Omega} \leq \|u_{0h} - u_0\| + Ch^{k+1} \left\{ \|u_0\|_{k+1,\Omega} + \int_0^{t_m} \|u_t\|_{k+1,\Omega} ds \right\} + \tau \int_0^{t_m} \|u_{tt}\|_{0,\Omega} ds. \quad (6.34)$$



**注** 上面的估计式中, 可以分为三个部分: 初值误差, 空间离散误差与时间积分误差。我们可以看到在误差分析过程中空间离散和时间积分是解耦的。

**证明** 首先改写  $U^m - u(t_m)$  为:

$$U^m - u(t_m) = (U^m - P_h u(t_m)) + (P_h u(t_m) - u(t_m)) := \theta^m + \rho^m. \quad (6.35)$$

关于  $\rho^m$  的分析和前面半离散问题完全相同, 下面主要讨论  $\theta^m$  的分析。



首先我们需要给出  $\theta^m$  满足的方程:

$$\begin{aligned}
& \langle \bar{\partial}_t \theta^m, \chi \rangle + \langle \nabla \theta^m, \nabla \chi \rangle \\
&= \langle \bar{\partial}_t U^m, \chi \rangle - \langle \bar{\partial}_t P_h u(t_m), \chi \rangle + \langle \nabla U^m, \nabla \chi \rangle - \langle \nabla P_h u(t_m), \nabla \chi \rangle \\
&= \langle f(t_m), \chi \rangle - \langle \bar{\partial}_t P_h u(t_m), \chi \rangle - \langle \nabla P_h u(t_m), \nabla \chi \rangle \\
&= \langle u_t(t_m), \chi \rangle + \langle \nabla u(t_m), \nabla \chi \rangle - \langle \nabla P_h u(t_m), \nabla \chi \rangle - \langle \bar{\partial}_t P_h u(t_m), \chi \rangle \\
&= \langle u_t - \bar{\partial}_t P_h u(t_m), \chi \rangle := -\langle \omega^m, \chi \rangle, \forall \chi \in V_h.
\end{aligned} \tag{6.36}$$

其中

$$\begin{aligned}
\omega^m &= \bar{\partial}_t P_h u(t_m) - u_t^m \\
&= \bar{\partial}_t P_h u(t_m) - \bar{\partial}_t u(t_m) + \bar{\partial}_t u(t_m) - u_t^m \\
&= (P_h - I) \bar{\partial}_t u(t_m) + \bar{\partial}_t u(t_m) - u_t^m \\
&= \omega_1^m + \omega_2^m.
\end{aligned} \tag{6.37}$$

这里  $\omega_1^m$  可以看作将  $\bar{\partial}_t u(t_m)$  投影到  $V_h$  产生的误差, 而  $\omega_2^m$  即为关于时间进行参数离散产生的截断误差。

在(6.36)中取  $\chi = \theta^m$ , 可得:

$$\langle \bar{\partial}_t \theta^m, \theta^m \rangle \leq -\langle \omega^m, \theta^m \rangle \leq \|\omega^m\|_{0,\Omega} \|\theta^m\|_{0,\Omega}. \tag{6.38}$$

代入  $\bar{\partial}_t \theta^m$  的表达式, 可得:

$$\|\theta^m\|_{0,\Omega}^2 - \langle \theta^{m-1}, \theta^m \rangle \leq \tau \|\omega^m\|_{0,\Omega} \|\theta^m\|_{0,\Omega}. \tag{6.39}$$

不断使用 Cauchy-Schwarz 不等式, 可得:

$$\|\theta^m\|_{0,\Omega} \leq \|\theta^0\|_{0,\Omega} + \tau \left( \sum_{j=1}^m \|\omega^j\|_{0,\Omega} \right) \leq \|\theta^0\|_{0,\Omega} + \tau \sum_{j=1}^m \|\omega_1^j\|_{0,\Omega} + \tau \sum_{j=1}^m \|\omega_2^j\|_{0,\Omega}. \tag{6.40}$$

首先估计  $\|\theta^0\|_{0,\Omega}$ :

$$\|\theta^0\|_{0,\Omega} \leq \|u_{0h} - P_h u_0\|_{0,\Omega} \leq \|u_{0h} - u_0\| + Ch^{k+1} \|u_0\|_{k+1,\Omega}. \tag{6.41}$$

由于

$$\omega_1^j = \tau^{-1} (P_h - I) \int_{t_{j-1}}^{t_j} u_t(s) ds. \tag{6.42}$$

我们有:

$$\tau \sum_{j=1}^m \|\omega_1^j\|_{0,\Omega} \leq \sum_{j=1}^m \int_{t_{j-1}}^{t_j} \|(P_h - I)u_t(s)\|_{0,\Omega} ds \leq Ch^{k+1} \sum_{j=1}^m \int_{t_{j-1}}^{t_j} \|u_t\|_{k+1,\Omega} ds. \tag{6.43}$$

同样的, 由于

$$\omega_2^j = -\tau^{-1} \int_{t_{j-1}}^{t_j} (s - t_{j-1}) u_{tt}(s) ds. \tag{6.44}$$

可得:

$$\tau \sum_{j=1}^m \|\omega_2^j\|_{0,\Omega} \leq \tau \sum_{j=1}^m \int_{t_{j-1}}^{t_j} \|u_{tt}\|_{0,\Omega} ds. \tag{6.45}$$

综合上述, 结论得证。

**注** 上面的推导说明了 Euler-Galerkin 方法是收敛的, 但不够理想, 因为这个算法关于时间仅一阶收敛。要提升关于时间的精度, 我们需要使用更高阶数的时间积分方法。