Непрерывные математические модели

А. В. Лапин

Казанский Федеральный Университет

Оглавление

Глава 1

Решение линейных одномерных эллиптических краевых задач

1.1 Конечно-разностная аппроксимация двухточечной краевой задачи

Рассмотрим так называмую двухточечную краевую задачу с краевыми условиями первого рода (задачу Дирихле для одномерного уравнения):

$$-u''(x) = f(x) \text{ при } 0 < x < 1; \ \ u(0) = u(1) = 0. \tag{1.1}$$

Далее будем обозначать через $C^k([0,1])$ для целого $k \geqslant 0$ пространство функций, непрерывных на отрезке $[0.1]$ вместе со своими производными до k-ого порядка включительно.

Для $f \in C^0([0,1])$ существует единственное решение $u \in C^2([0,1])$ задачи (1.1). Известно свойство монотонности: если $f(x)$ неотрицательна, то решение $u(x)$ также неотрицательно. Другое свойство, называемое принципом максимума, гласит, что для $f \in C^0([0,1])$ справедливо неравенство

$$\|u\|_\infty \leqslant \frac{1}{8}\|f\|_\infty, \text{ где } \|u\|_\infty = \max_{0 \leqslant x \leqslant 1} |u(x)|. \tag{1.2}$$

Построим на отрезке $[0,1]$ равномерную сетку $\omega = \{0 = x_0 < x_1 < \dots < x_{n-1} < x_n = 1\}, x_j - x_{j-1} = h$. При построении конечно-разностной аппроксимации будем использовать разностные отношения, аппроксимирующие производные первого порядка:

$$u'(x_j) \approx \frac{u(x_{j+1}) - u(x_j)}{h} \equiv \partial u(x_j) \text{ (правая разность)},$$

$$u'(x_j) \approx \frac{u(x_j) - u(x_{j-1})}{h} \equiv \bar{\partial} u(x_j) \text{ (левая разность)}$$

и симметричное разностное отношение, аппроксимирующее производную второго порядка

$$u''(x_j) \approx \frac{u(x_{j+1}) - 2u(x_j) + u(x_{j-1})}{h^2} \equiv \bar{\partial}\partial u(x_j) = \partial\bar{\partial} u(x_j).$$

Пусть

$$V_h = \{y_h = (y_0, y_1, \ldots, y_n) \in \mathbb{R}^{n+1} : y_0 = y_n = 0\}$$

— пространство сеточных функций, определенных на сетке ω и обращающихся в 0 на границе. Определим сеточный оператор $L_h : V_h \to V_h$ равенством

$$L_h y_h = -\frac{y_{j+1} - 2y_j + y_{j-1}}{h^2}.$$

Тогда конечно-разностной аппроксимацией краевой задачи (1.1) является система уравнений

$$L_h y_h = f_h$$

с $f_h(x_j) = f(x_j)$, или, в явном виде,

$$\begin{cases} -\dfrac{y_{j+1} - 2y_j + y_{j-1}}{h^2} = f_j \text{ для } j = 1, 2, \ldots, n-1, \\ y_0 = y_n = 0. \end{cases} \tag{1.3}$$

Сеточная функция y_h с значениями y_j в узлах x_j сетки ω (иными словами, с узловыми параметрами y_j) аппроксимирует точное решение u.

Для разностной схемы (1.3) справедливо свойство монотонности:

$$f_j \geqslant 0 \; \forall j \Rightarrow y_j \geqslant 0 \; \forall j,$$

и дискретный аналог принципа максимума (1.2):

$$\|y_h\|_\infty = \max_{1 \leqslant j \leqslant n-1} |y_j| \leqslant \frac{1}{8}\|f_h\|_\infty. \tag{1.4}$$

Из неравенства (1.4) следует существование единственного решения и устойчивость в дискретной максимум-норме для конечно-разностного решения y_h.

Пусть $f \in C^0([0,1])$, тогда решение краевой задачи (1.1) $u \in C^2([0,1])$. Определим сеточную функи погрешности аппроксимации ψ_h равенствами

$$\psi_h(x_j) = (L_h u)(x_j) - f(x_j), j = 1, 2, \ldots, n-1.$$

Легко проверить, что
$$\lim_{h\to+0}\|\psi_h\|_\infty = 0.$$
Более того, если $f \in C^2([0,1])$, так что $u \in C^4([0,1])$, то
$$\|\psi_h\|_\infty \leqslant \frac{\|u^{(iv)}\|_\infty}{12}h^2 = \frac{\|f''\|_\infty}{12}h^2. \tag{1.5}$$
Это неравенство показывает, что конечно-разностная схема (1.3) имеет второй порядок аппроксимации на гладком решении.

Обозначим через $z_h = u - y_h = (z_0, z_1, \ldots, z_n) \in \mathbb{R}^{n+1} : z_0 = z_n = 0$ сеточную функцию погрешности с компонентами $z_j = u(x_j) - y_j$. Тогда
$$L_h z_h = \psi_h$$
и аналогично (1.4) получим
$$\|z_h\|_\infty \leqslant \frac{1}{8}\|\psi_h\|_\infty. \tag{1.6}$$
теперь из (1.5) и (1.6) следует
$$\|z_h\|_\infty \leqslant Mh^2,\ M = \frac{\|f''\|_\infty}{96}. \tag{1.7}$$

Обозначим через $y = (y_1, y_2, \ldots, y_{n-1})^T$ и $f = (f_1, f_2, \ldots, f_{n-1})^T$ векторы узловых параметров соответствующих сеточных функций. Тогда (1.3) можно записать в ви де системы линейных алгебраических уравнений
$$Ay = f, \tag{1.8}$$
где A — $(n-1)\times(n-1)$ матрица:
$$A = h^{-2}\mathrm{tridiagn}(-1,2,-1) = h^{-2}\begin{pmatrix} 2 & -1 & 0 & 0 & 0 & 0 \\ -1 & 2 & -1 & 0 & 0 & 0 \\ 0 & -1 & 2 & -1 & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & 0 & -1 & 2 \end{pmatrix}. \tag{1.9}$$
Трехдиагональная матрица A симметрична и положительно определена. Кроме того, она обладает свойством диагонального преобладания. Это свойство в общем случае матрицы $A = \{a_{ij}\}$ означает, что
$$|a_{ii}| \geqslant \sum_{j\neq i}|a_{ij}| \text{ для всех } i \text{ и } \exists i^* : |a_{i^*i^*}| > \sum_{j\neq i^*}|a_{i^*j}|.$$

1.2 Вариационный метод построения конечно-разностной схемы

Далее $L_2(0,1)$ – пространство Лебега функций, интегрируемых с квадратом на $[0,1]$: $u \in L_2(0,1) \Leftrightarrow \int\limits_0^1 u^2 dx < \infty$. Здесь интеграл Лебега, однако для простоты можно считать, что это "обычный"интеграл Римана. Через $H^1(0,1) = \{u \in L_2(0,1) : u' \in L_2(0,1)\}$ обозначим пространство Соболева. Для нас в дальнейшем важно знать, что это пространство содержит непрерывные на $[0,1]$ функции, имеющие производные первого порядка, интегрируемые с квадратом: $\int\limits_0^1 (u')^2 dx < \infty$. В частности, непрерывные функции с кусочно непрерывной на $[0,1]$ производной принадлежат $H^1(0,1)$. Через на $H_0^1(0,1) = \{u \in H^1(0,1) : u(0) = u(1) = 0\}$ будем обозначать подпространство $H^1(0,1)$ функций, обращающихся в нуль в граничных точках отрезка $[0,1]$.

Умножим уравнение (1.1) на функцию $v \in C^1([0,1]), v(0) = v(1) = 0$, которую будем называть тестовой, и затем проинтегрируем по $[0,1]$. Использовав формулу интегрирования по частям, получим

$$\int\limits_0^1 u'v'dx = \int\limits_0^1 fvdx.$$

Назовем обобщенным решением краевой задачи (1.1) функцию u, удовлетворяющую следующему **интегральному тождеству:**

$$u \in H_0^1(0,1) : \int\limits_0^1 u'v'dx = \int\limits_0^1 fvdx \;\forall v \in H_0^1(0,1). \tag{1.10}$$

Если функция u является классическим решением краевой задачи (1.1), т.е. $u \in C^2(0,1) \cap C([0,1])$, удовлетворяет уравнению (1.1) и обращается в ноль на границе области $x = 0$ и $x = 1$, то u есть также решение интегрального тождества (1.10). Обратно, если u — решение интегрального тождества (1.10) и обладает дополнительной гладкостью: $u \in C^2(0,1) \cap C([0,1])$, то u — классическое решение (1.1).

Мы построим конечно-разностную схему для (1.1), аппроксимируя интегральное тождество (1.10). Для этого мы будем использовать следующие квадратурные формулы ($\varphi(x)$ — непрерывная функция):

$$\begin{cases} \int\limits_0^1 \varphi(x)dx \approx h \sum\limits_{i=0}^{n-1} \varphi(x_i) \equiv [\varphi, 1) \text{ (составная формула левых прямоугольников)}, \\ \int\limits_0^1 \varphi(x)dx \approx h \sum\limits_{i=1}^{n} \varphi(x_i) \equiv (\varphi, 1] \text{ (составная формула правых прямоугольников)}, \\ \int\limits_0^1 \varphi(x)dx \approx \dfrac{h}{2}\varphi(x_0) + h \sum\limits_{i=1}^{n-1} \varphi(x_i) + \dfrac{h}{2}\varphi(x_n) \equiv [\varphi, 1] = \\ = \dfrac{1}{2}\big([\varphi, 1) + (\varphi, 1]\big) \text{ (составная формула трапеций)}. \end{cases} \tag{1.11}$$

Пусть как и ранее $V_h = \{y_h = (y_0, y_1, \dots, y_n) \in \mathbb{R}^{n+1} : y_0 = y_n = 0\}$ — пространство сеточных функций, определенных на сетке $\omega = \{0 = x_0 < x_1 < \dots < x_{n-1} < x_n = 1\}, x_j - x_{j-1} = h$ и обращающихся в 0 на границе.

Аппроксимируем интегральное тождество (1.10) так называемым сумматорным тождеством

$$y_h \in V_h : \ \frac{1}{2}(\bar{\partial} y_h, \bar{\partial} v_h] + \frac{1}{2}[\partial y_h, \partial v_h) = [f, v_h] \ \forall v_h \in V_h. \tag{1.12}$$

Выбирая в качестве тестовой функции $v_h = (0, 0, \dots, 1, 0, \dots, 0)$ с $v_j = 1$ последовательно для $j = 1, 2, \dots, n-1$, получим из (1.12) линейную систему (1.3). Таким образом, сумматорное тождество (1.12) представляет собой неявную форму записи разностной схемы (1.3) для задачи (1.1).

Очевидно, положив $y = (y_1, y_2, \dots, y_{n-1})^T$ и $f = (f_1, f_2, \dots, f_{n-1})^T$, мы снова получим систему линейных алгебраических уравнений с матрицей (1.9).

Результаты секции 1.1 об устойчивости и сходимости разностной схемы в максимум-норме остаются справедливыми. Более того, используя запись схемы в виде (1.12), можно получить оценки в сеточной норме — аналоге нормы пространства Соболева H_0^1.

Положим

$$\|y_h\|_0 = (y_h, y_h)^{1/2} = \big(h \sum_{j=1}^{n-1} y_j^2\big)^{1/2},$$

$$\|y_h\|_1 = \big(\frac{1}{2}(\bar{\partial} y, \bar{\partial} y] + \frac{1}{2}[\partial y, \partial y)\big)^{1/2} = \big(h \sum_{j=1}^{n} (\bar{\partial} y)_j^2\big)^{1/2}.$$

Это нормы в пространстве V_h, являющиеся сеточными аналогами норм L_2 и H_0^1. Очевидно, что как нормы в конечномерном пространстве они эквивалентны. В общем случае константы эквивалентности зависят от размерности пространства, или, иначе говоря, от шага сетки h. Но для норм $\|.\|_0$, $\|.\|_\infty$ и $\|.\|_1$ в V_h справедливы неравенства:

$$\|y_h\|_0 \leqslant \|y_h\|_1. \tag{1.13}$$

$$\|y_h\|_\infty \leqslant \|y_h\|_1. \tag{1.14}$$

Для разностной схемы (1.3) справедливо неравенство устойчивости:

$$\|y_h\|_1 \leqslant \|f_h\|_0. \tag{1.15}$$

Из (1.15), (1.5) и неравенства $\|f_h\|_0 \leqslant \|f_h\|_\infty$ следует оценка точности

$$\|z_h\|_1 \leqslant Mh^2,\ M = \frac{\|f''\|_\infty}{12}. \tag{1.16}$$

В силу (1.14) норма $\|.\|_1$ "сильнее" чем максимум-норма $\|.\|_\infty$, поэтому оценка (1.16) лучше, чем (1.7).

Однако основное достоинство метода аппроксимации с помощью сумматорного тождества состоит в том, что он является универсальным инструментом построения аппроксимаций краевых задач с переменными коэффициентами и различными краевыми условиями. Эти задачи кратко рассмотрены ниже.

Конечно-разностная схема для задачи с переменными коэффициентами

Рассмотрим краевую задачу Дирихле для уравнения с переменными коэффициентами:

$$-\big(p(x)u'(x)\big)' + q(x)u(x) = f(x) \text{ for } 0 < x < 1;\ \ u(0) = u(1) = 0, \tag{1.17}$$

где $p(x)$, $q(x)$ и $f(x)$ непрерывны и

$$p(x) \geqslant p_0, p_0 = const > 0,\ q(x) \geqslant 0 \text{ для всех } x \in \bar{\Omega}.$$

Умножив уравнение (1.17) на тестовую функцию $v \in C^1([0,1]), v(0) = v(1) = 0$, и проинтегрировав по частям, получим интегральное тождество

$$\int\limits_0^1 p(x)u'v'dx + \int\limits_0^1 q(x)uvdx = \int\limits_0^1 fvdx. \tag{1.18}$$

Функция $u \in H_0^1(0,1)$, удовлетворяющая (1.18) при всех $v \in H_0^1(0,1)$, является обобщенным решением задачи (1.17).

Используя введенные выше обозначения, аппоксимируем интегральное тождество сумматорным тождеством

$$y_h \in V_h: \ \frac{1}{2}(p\,\bar{\partial} y_h, \bar{\partial} v_h] + \frac{1}{2}[p\,\partial y_h, \partial v_h) + (q\,y_h, v_h) = (f_h, v_h) \ \forall v_h \in V_h. \quad (1.19)$$

Явная форма соответствующей разностной схемы:

$$\begin{cases} -\frac{1}{2}\partial(p\,\bar{\partial} y_h)(x_j) - \frac{1}{2}\bar{\partial}(p\,\partial y_h)(x_j) + q y_h(x_j) = f(x_j), \ 1 \leqslant j \leqslant n-1, \\ y_0 = y_n = 0. \end{cases} \quad (1.20)$$

В силу предположений относительно коэффициентов и неравенства (1.13) справедливо неравенство устойчивости подобное (1.15):

$$\|y_h\|_1 \leqslant c\,\|f_h\|_0, \ c = const.$$

Конечно-разностная схема для задачи Неймана

Рассмотрим задачу Неймана (вторую краевую задачу) для одномерного уравнения:

$$-\big(p(x)u'(x)\big)' + q(x)u(x) = f(x) \text{ for } 0 < x < 1; \ \ u'(0) = u'(1) = 0. \quad (1.21)$$

Здесь коэффициенты $p(x)$, $q(x)$ и правая часть $f(x)$ непрерывны и удовлетворяют условиям

$$p(x) \geqslant p_0, p_0 = const > 0, \ q(x) \geqslant q_0 > 0 \text{ для всех } x \in \bar{\Omega}.$$

Умножая уравнение (1.21) на функцию $v \in C^1([0,1])$ и интегрируя по частям, получаем интегральное тождество

$$\int\limits_0^1 p(x)u'v'dx + \int\limits_0^1 q(x)uvdx = \int\limits_0^1 fvdx. \quad (1.22)$$

Функция $u \in H^1(0,1)$, удовлетворяющая (1.22) для всех $v \in H^1(0,1)$, является обобщенным решением задачи (1.21).

Пусть $V_h = \{y_h = (y_0, y_1, \dots, y_n) \in \mathbb{R}^{n+1}\}$ — пространство сеточных функций, определенных на сетке $\omega = \{0 = x_0 < x_1 < \dots < x_{n-1} < x_n = 1\}, x_j - x_{j-1} = h$ (в данном случае сеточные функции не обращаются в 0 на границе). Сумматорное тождество, аппроксимирующее интегральное тождество (1.22), имеет вид

$$y_h \in V_h: \ \frac{1}{2}(p\,\bar{\partial} y_h, \bar{\partial} v_h] + \frac{1}{2}[p\,\partial y_h, \partial v_h) + [q\,y_h, v_h] = [f_h, v_h] \ \forall v_h \in V_h. \quad (1.23)$$

Отметим, что теперь выражения $[q\,y_h, v_h]$ и $[f, v_h]$ использованы для аппроксимации соответствующих интегралов, так как сеточные функции не обращаются в ноль на границе.

Явная форма конечно-разностной схемы:

$$\begin{cases} -\dfrac{1}{2}\partial(p\,\bar{\partial}y_h)(x_j) - \dfrac{1}{2}\bar{\partial}(p\,\partial y_h)(x_j) + qy_h(x_j) = f(x_j) \text{ for } j = 1, 2, \dots, n-1, \\ -\dfrac{p(0)+p(h)}{2}\,\partial y_h(0) + \dfrac{h}{2}qy_h(0) = \dfrac{h}{2}f_h(0), \\ \dfrac{p(1)+p(1-h)}{2}\,\bar{\partial}y_h(1) + \dfrac{h}{2}qy_h(1) = \dfrac{h}{2}f_h(1). \end{cases}$$

Оснастив пространство V_h нормой (сеточный аналог нормы пространства $H^1(0,1)$)

$$\|y_h\|_1 = \big((\bar{\partial}y_h, \bar{\partial}y_h] + [\partial y_h, \partial y_h) + [y_h, y_h]\big)^{1/2},$$

можно получиить оценку устойчивости

$$\|y_h\|_1 \leqslant c\,\|f_h\|_0 = c\,[f,f]^{1/2},\ c = const.$$

Конечно-разностная схема для задачи Робина

Рассмотрим задачу Робина (третью краевую задачу) для одномерного уравнения:

$$\begin{cases} -\big(p(x)u'(x)\big)' + q(x)u(x) = f(x) \text{ for } 0 < x < 1, \\ -p(0)u'(0) + ru(0) = \phi, \\ p(1)u'(1) + ru(1) = \psi, \end{cases} \tag{1.24}$$

где $p(x)$, $q(x)$ и $f(x)$ непрерывны и

$$p(x) \geqslant p_0, p_0 = const > 0,\ q(x) \geqslant 0 \text{ для всех } x \in \bar{\Omega}, r > 0.$$

В очередной раз умножая уравнение на функцию $v \in C^1([0,1])$ и интегрируя по частям, получаем интегральное тождество

$$\int\limits_0^1 p(x)u'v'dx + \int\limits_0^1 q(x)uvdx + ru(0)v(0) + ru(1)v(1) = \int\limits_0^1 fvdx + \phi v(0) + \psi v(0), \tag{1.25}$$

которое определяет обобщенное решение задачи (1.24) из постранства $H^1(0,1)$.

Пусть $V_h = \{y_h = (y_0, y_1, \dots, y_n) \in \mathbb{R}^{n+1}\}$ — пространство сеточных функций, определенных на равномерной сетке $\omega = \{0 = x_0 < x_1 < \dots < x_{n-1} < x_n = 1\}, x_j - x_{j-1} = h$. Сумматорное тождество, аппроксимирующее интегральное тождество (1.25), имеет вид

$$\frac{1}{2}(p\,\bar\partial y_h, \bar\partial v_h] + \frac{1}{2}[p\,\partial y_h, \partial v_h) + [q\,y_h, v_h] + ry_h(0)v_h(0) + ry_h(1)v_h(1) =$$

$$= [f_h, v_h] + \phi v_h(0) + \psi v_h(0). \tag{1.26}$$

Явная форма конечно-разностной схемы:

$$\begin{cases} -\dfrac{1}{2}\partial(p\,\bar\partial y_h)(x_j) - \dfrac{1}{2}\bar\partial(p\,\partial y_h)(x_j) + qy_h(x_j) = f(x_j) \text{ for } j = 1, 2, \dots, n-1, \\ -\dfrac{p(0)+p(h)}{2}\,\partial y_h(0) + ry_h(0) + \dfrac{h}{2}qy_h(0) = \phi + \dfrac{h}{2}f_h(0), \\ \dfrac{p(1)+p(1-h)}{2}\,\bar\partial y_h(1) + ry_h(1) + \dfrac{h}{2}qy_h(1) = \psi + \dfrac{h}{2}f_h(1). \end{cases}$$

Оснастив пространство V_h нормой (сеточный аналог одной из эквивалентых норм пространства $H^1(0,1)$)

$$\|y_h\|_1 = \left((\bar\partial y_h, \bar\partial y_h] + [\partial y_h, \partial y_h) + y_h^2(0) + y_h^2(1)\right)^{1/2},$$

можно получиить оценку устойчивости

$$\|y_h\|_1 \leqslant c\,(\|f_h\|_0 + \phi + \psi),\ c = const.$$

1.3 Аппроксимация по методу конечных элементов

Схема МКЭ для задачи (1.1)

Для построение схемы метода конечных элементов для задачи (1.1) снова воспользуемся ее обощенной формулировкой (1.10):

$$u \in H_0^1(0,1): \int_0^1 u'v'dx = \int_0^1 fvdx \quad \forall v \in H_0^1(0,1).$$

Пусть отрезок разбит точками $\{x_j\}_{j=0}^n$ и пусть V_h — пространство непрерывных функций, лиейных на каждом интервале (x_j, x_{j+1}), и равных нулю на концах отрезка $x = 0$ и $x = 1$. Тогда V_h — это пространство размерности $\dim V_h = n-1$, и $V_h \subset V = H_0^1(0,1)$. Обозначая через $u_h, v_h, ...$

функции из V_h, определим схему МКЭ с помощью вариационного уравнения

$$u_h \in V_h: \int_0^1 u_h' v_h' dx = \int_0^1 f v_h dx \ \ \forall v_h \in V_h. \tag{1.27}$$

(1.27) — это неявная форма схемы МКЭ. Для построения явного вида схемы требуется выбрать базис $\{\varphi_i(x)\}_{i=1}^{n-1}$ в пространстве V_h и написать соответствующую систему линейных алгебраических уравнений для коэффициентов Фурье y_i решения

$$u_h = \sum_{i=1}^{n-1} y_i\, \varphi_i(x).$$

Далее для простоты рассмотрим лишь равномерную сетку: $x_j - x_{j-1} = h \ \forall j$. Традиционным базисом в V_h является так называемый базис Куранта:

$$\varphi_i(x) = \begin{cases} \dfrac{x - x_{i-1}}{h} \text{ for } x \in (x_{i-1}, x_i) \\ \dfrac{x_{i+1} - x}{h} \text{ for } x \in (x_{i-1}, x_i) \\ 0 \text{ для других точек в } [0,1]. \end{cases}$$

Подставляя $u_h = \sum\limits_{i=1}^{n-1} y_i\, \varphi_i(x)$, $v_h = \varphi_i(x)$ в уравнение (1.27), получим

$$\begin{cases} -\dfrac{y_{i-1} - 2y_i + y_{i+1}}{h^2} = f_i \text{ for } i = 1, 2, \ldots, n-1, \\ y_0 = y_n = 0. \end{cases} \tag{1.28}$$

Матрица системы (1.28) — та же, что и в разностной схеме (1.3), тогда как вектор правых частей имеет вид:

$$f_i = \frac{1}{h} \int_{x_{i-1}}^{x_{i+1}} f(x) \varphi_i(x) dx$$

Как и для конечно-разностной схемы, для схемы МКЭ справедливо неравенство устойчивости

$$\|u_h\|_\infty = \max_{1 \leqslant i \leqslant n-1} |y_i| \leqslant \frac{1}{8} \max_{1 \leqslant i \leqslant n-1} |f_i|,$$

из которого, в частности, следует существование единственного решения (1.28). Более того, справедлива оценка точности

$$\|u - u_h\|_\infty \leqslant M_2 h^2, \ M_2 = const.$$

Схема МКЭ для уравнения с переменными коэффициентами

Рассмотрим задачу (1.17), считая функции $p(x)$, $q(x)$ и $f(x)$ непрерывными и такими, что

$$p(x) \geqslant p_0, p_0 = const > 0,\ q(x) \geqslant 0 \text{ для всех } x \in \bar{\Omega}.$$

Как мы знаем, обобщенное решение этой задачи удовлетворяет инткгральному тождеству (1.18):

$$u \in H_0^1(0,1):\ \int\limits_0^1 p(x)u'v'dx + \int\limits_0^1 q(x)uvdx = \int\limits_0^1 fvdx\ \forall v \in H_0^1(0,1)$$

Для построения схемы МКЭ для (1.18) мы снова построим равноменую сетку шага h на отрезке $[0,1]$ и введем в рассмотрение пространство кусочно линейных и непрерывных функций V_h, обращающихся в ноль при $x = 0$ и $x = 1$. Схема МКЭ в неявном виде для задачи (1.18) — это интегральное тождество

$$u_h \in V_h:\ \int\limits_0^1 p(x)u_h'v_h'dx + \int\limits_0^1 q(x)u_h v_h dx = \int\limits_0^1 f v_h dx\ \ \forall v_h \in V_h. \tag{1.29}$$

Выбирая базис Куранта в пространстве V_h и подставляя $u_h = \sum\limits_{i=1}^{n-1} y_i\, \varphi_i(x)$, $v_h = \varphi_i(x)$ в уравнение (1.29), получим явный вид схемы:

$$\begin{cases} a_i\, y_{i-1} + b_i\, y_i + c_i\, y_{i+1} = f_i \text{ для } i = 1, 2, \ldots, n-1, \\ y_0 = y_n = 0. \end{cases} \tag{1.30}$$

Здесь

$$a_i = \int\limits_0^1 \big(p(x)\varphi_{i-1}'\varphi_i'dx + q(x)\varphi_{i-1}\varphi_i\big)\, dx =$$

$$= -\frac{1}{h^2}\int\limits_{x_{i-1}}^{x_i} p(x)\, dx - \frac{1}{h^2}\int\limits_{x_{i-1}}^{x_i} q(x)(x - x_{i-1})(x - x_i)dx, \tag{1.31}$$

$$c_i = \int\limits_0^1 \left(p(x)\varphi_i'\varphi_{i+1}'dx + q(x)\varphi_i\varphi_{i+1}\right) dx =$$

$$= -\frac{1}{h^2}\int\limits_{x_i}^{x_{i+1}} p(x)\,dx - \frac{1}{h^2}\int\limits_{x_i}^{x_{i+1}} q(x)(x-x_i)(x-x_{i+1})dx, \quad (1.32)$$

$$b_i = \int\limits_0^1 \left(p(x)(\varphi_i')^2 dx + q(x)\varphi_i^2\right) dx =$$

$$= \frac{1}{h^2}\int\limits_{x_{i-1}}^{x_{i+1}} p(x)\,dx + \frac{1}{h^2}\int\limits_{x_{i-1}}^{x_i} q(x)(x-x_{i-1})^2 dx + \frac{1}{h^2}\int\limits_{x_i}^{x_{i+1}} q(x)(x-x_{i+1})^2 dx, \quad (1.33)$$

$$f_i = \frac{1}{h}\int\limits_{x_{i-1}}^{x_{i+1}} f(x)\varphi_i(x)dx.$$

Матрица системы (1.30) имеет диагональное преобладание, т.е.

$$b_i \geqslant |a_i| + |c_i| \text{ для всех } i = 1, 2, \ldots, n-1,$$

где a_i, c_i и b_i определены в (1.31)- (1.33).

Глава 2

Основы теории линейных краевых задач для эллиптических уравнений

2.1 Функциональные пространства

Пространства Банаха и Гильберта

Пусть X – вещественное линейное пространство. Норма $||.||$ на X – это отображение из X в $\mathbb{R}$ такое, что

1. $\|u\| \geqslant 0$ для каждого $u \in X$ и $||u|| = 0 \Rightarrow u = 0$,
2. $\|u\| = |c|\|u\|$ для $c \in \mathbb{R}$ и $u \in X$,
3. $\|u + v\|) \leqslant \|u\| + \|v\|$ для $u, v \in X$.

Мы будем часто использовать обозначение $(X, ||.||_X)$ для пространства X с нормой $||.||_X$.

Скалярное произведение $(.,.)$ в X – это билинейное отображение из $X \times X$ в $\mathbb{R}$ такое, что

1. $(u, v) = (v, u)$ для $u, v \in X$,
2. $(u, u) \geqslant 0,\ \forall u \in X$,
3. $(u, u) = 0 \Leftrightarrow u = 0$.

Скалярное произведение определяет норму $||u|| = (u, u)^{1/2}$.

Линейное пространство X со скалярным произведением (соответственно, с нормой) называется предгильбертовым (соответственно, нормированным) пространством. Линейное нормированное пространство X называется полным, если любая последовательность Коши имеет в X предел, т.е.

$$\{u_k\} \subset X, ||u_k - u_l|| \to 0 \text{ при } k, l \to \infty \Rightarrow \exists u \in X : ||u_k - u|| \to 0 \text{ при } k \to +\infty.$$

Полное линейное нормированное пространство — пространство Банаха (банахово). Соответственно, полное предгильбертово пространство – пространство Гильберта (гильбертово).

Две нормы $\|\cdot\|$ и $\|\cdot\|_1$ на линейном пространстве X называются эквивалентными (используется обозначение $\|.\| \sim \|.\|_1$), если существуют постоянные c_1 и c_0 такие, что

$$c_0||u|| \leqslant ||u||_1 \leqslant c_1||u|| \quad \forall u \in X.$$

Если $X = (X, ||.||_X)$ и $Y = (Y, ||.||_Y)$ – два линейных нормированных пространства, то $\mathbb{L}(X; Y)$ – линейное пространство линейных непрерывных операторов из X в Y. Оснащенное нормой

$$||A|| = \sup_{u \in X,\, u \neq 0} \frac{||Au||_Y}{||u||_X} \quad \text{для } A \in \mathbb{L}(X; Y),$$

пространство $\mathbb{L}(X; Y)$ становится линейным нормированным пространством. Если Y банахово, то $\mathbb{L}(X; Y)$ – также банахово.

Пространство $\mathbb{L}(X; \mathbb{R})$ называется сопряженным пространством к X и обозначается X^*. Сопряженное к линейному нормированному пространству является банаховым с нормой (двойственной к $||.||_X$)

$$\|f\|_* = \sup_{u \in X,\, u \neq 0} \frac{|f(u)|}{||u||_X} \quad \text{для } f \in X^*.$$

Пусть X – гильбертово пространство. Тогда согласно теореме Рисса (о представлении линейного функционала) для любого $f \in X^*$ существует единственный элемент $u_f \in X$ такой, что

$$f(v) = (u_f, v) \quad \forall v \in X,$$

при этом $\|f\|_* = \|u_f\|_X$.

Для линейного нормированного пространства X определим второе сопряженное пространство $X^{**} \equiv (X^*)^*$ и линейный оператор изометрического вложения $\pi : X \to X^{**}$, $\pi(X) \subset X^{**}$, $||\pi|| = 1$. Теперь X может быть отождествлено с подмножеством X^{**} (далее пишем $X \subset X^{**}$).

Если X^* – сопряженное к X, то обозначаем через $(x, f) \equiv (f, x)$ отношение двойственности между $x \in X$ и $f \in X^*$, то есть значение функционала $f \in X^*$ на элементе $x \in X$, или, что равносильно, значение функционала $x \in X \subset X^{**}$ на элементе $f \in X^*$.

Линейное нормированное пространство X называется рефлексивным, если $X^{**} = X$. Гильбертово пространство является примером рефлексивного банахова пространства.

Множество $M \subset X$ называется предкомпактным, если из любой последовательности элементов в M можно выделить сильно сходящуюся в X подпоследовательность. Если, кроме того, предел этой подпоследовательности принадлежит M, то M компактно.

Оператор A из X в Y называется компактным, если он переводит любое ограниченное в X множество в предкомпактное множество в Y. Непрерывный и компактный оператор – вполне непрерывный оператор.

Пусть $X = (X, \|\cdot\|_X)$ и $Y = (Y, \|\cdot\|_Y)$ – два линейных нормированных пространства. Говорят, что X непрерывно вложено в Y $(X \subset Y)$, если существует линейный непрерывный оператор P, переводящий любой элемент $u \in X$ в тот же самый элемент $u \in Y$ так, что

$$\|u\|_Y \leqslant c\,\|u\|_X \quad \forall u \in X$$

с постоянной c, не зависящей от $u \in X$.

Если оператор P к тому же компактный, то вложение X в Y непрерывно и компактно (то есть вполне непрерывно).

Пространства гладких функций

Далее используются следующие обозначения: пусть α_i – целые неотрицательные числа и вектор $\alpha = (\alpha_1, \ldots, \alpha_n)$ – так называемый мультииндекс длины $|\alpha| = \alpha_1 + \cdots + \alpha_n$. Тогда $D^\alpha u = \dfrac{\partial^{|\alpha|} u}{\partial x_1^{\alpha_1} \cdots \partial x_n^{\alpha_n}}$ – частная производная порядка $|\alpha|$.

Пусть $\Omega \subset \mathbb{R}^n$ – область с кусочно гладкой границей $\partial\Omega$ и $\overline{\Omega} = \Omega \cup \partial\Omega$ – замыкание Ω. Для целого $k \geqslant 0$ обозначим через $C^k(\Omega)$ линейное пространство функций $u = u(x)$, непрерывных в Ω вместе со всеми производными $D^\alpha u$ вплоть до порядка k, а через $C^k(\overline{\Omega})$ – линейное пространство сужений на $\overline{\Omega}$ функций из $C^k(\mathbb{R}^n)$. Оснащенное нормой

$$\|u\|_{C^k(\overline{\Omega})} = \max_{|\alpha| \leqslant k} \sup_{x \in \Omega} |D^\alpha u(x)|,$$

пространство $C^k(\overline{\Omega})$ становится банаховым. $C^\infty(\overline{\Omega}) = \bigcap\limits_{k=0}^{\infty} C^k(\overline{\Omega})$.

Через $C_0^\infty(\Omega)$ обозначим пространство бесконечно дифференцируемых функций $u(x)$ с компактным в Ω носителем $\operatorname{supp} u = \overline{\{x \in \Omega : u(x) \neq 0\}}$ – замыканием множества точек, в которых функция отлична от нуля. Иными словами, для каждой функции $u \in C_0^\infty(\Omega)$ существует окрестность границы $\partial\Omega$ в Ω, где $u(x)$ тождественно равна нулю.

Пространства Лебега

Пусть M – измеримое по Лебегу множество. Пространством Лебега $L_p(M)$, $1 \leqslant p < \infty$, называется линейное нормированное пространство измеримых и p-интегрируемых функций (т. е. $\int\limits_M |u|^p dx < \infty$) с нормой

$$\|u\|_{L_p} = \Big(\int\limits_M |u|^p dx\Big)^{1/p}.$$

Пространство $L_\infty(M)$ – это пространство измеримых и почти всюду (п. вс.) ограниченных функций с нормой

$$||u||_{L_\infty} = \operatorname*{ess\,sup}_{x\in M} |u(x)| \equiv \inf_{\mathrm{M_0\subset M,\ meas M_0=0}} \sup_{\mathrm{x\in M\setminus M_0}} |u(x)|.$$

Пространства $L_p(M)$ являются банаховыми для всех $1 \leqslant p \leqslant \infty$. Пространство $L_2(\Omega)$ – гильбертово со скалярным произведением $(u,v)_{L_2} = \int\limits_M u(x)\,v(x)\,dx$.

Далее используются пространства $L_p(\Omega)$ и $L_p(\Gamma)$, где Γ – часть границы $\partial\Omega$.

Множество $C_0^\infty(\Omega)$ всюду плотно в $L_p(\Omega)$ для всех $1 \leqslant p < \infty$, т.е. для любой функции $u \in L_p(\Omega)$ найдется последовательность $\{u_n\}$ из $C_0^\infty(\Omega)$: $\|u_n - u\|_{L_p} \to 0$ при $n \to \infty$.

В свою очередь, множество $C^\infty(\bar\Omega)$ всюду плотно в $L_p(\Omega)$ для всех $1 \leqslant p \leqslant \infty$.

Если $1 \leqslant p < \infty$, то сопряженным к $L_p(\Omega)$ является $L_q(\Omega)$, $1/p+1/q = 1$ (и $q = \infty$ при $p = 1$).

Пространства Соболева

Функция $u(x) \in L_p(\Omega)$ имеет обобщенную (или слабую) производную $\dfrac{\partial u}{\partial x_i} \in L_p(\Omega)$, если существует $u^*(x) \in L_p(\Omega)$ такая, что

$$\int\limits_\Omega u(x)\frac{\partial v}{\partial x_i}dx = -\int\limits_\Omega u^*(x) v dx \quad \forall v \in C_0^\infty(\Omega).$$

Функция u^* обозначается как $\dfrac{\partial u}{\partial x_i}$ и в случае существования "обычной" (классической) частной производной $\dfrac{\partial u}{\partial x_i}$ в точках Ω эти производные совпадают. Если u имеет в Ω обобщенную производную $\dfrac{\partial u}{\partial x_i} \in L_p(\Omega)$ и классическую производную в $\Omega^* \subset \Omega$, то в Ω^* они совпадают.

Обобщенные производные высших порядков определяются аналогично. Именно, производная $D^\alpha u = \dfrac{\partial^{|\alpha|} u}{\partial x_1^{\alpha_1} \partial x_2^{\alpha_2} \dots \partial x_n^{\alpha_n}} \in L_p(\Omega)$ определена равенством

$$\int\limits_\Omega u(x) D^\alpha v(x) dx = (-1)^{|\alpha|} \int\limits_\Omega D^\alpha u(x) v(x) dx \quad \forall v \in C_0^\infty(\Omega).$$

Обобщенная производная $D^\alpha u \in L_p(\Omega)$ определена однозначно.

Если $D^\alpha u$ – обобщенная производная от u из $L_p(\Omega)$, $\Omega^* \subset \Omega$ и $u(x)$ имеет классическую частную производную $D^\alpha u$ в точках Ω^*, то обобщенная производная и классическая производная совпадают в Ω^*.

Пусть $1 \leqslant p \leqslant \infty$ и $l \in \mathbb{N}$. Пространством Соболева $W_p^l(\Omega)$ называется линейное нормированное пространство функций $u \in L_p(\Omega)$, у которых существуют все обобщенные производные $D^\alpha u = \dfrac{\partial^{|\alpha|} u}{\partial x_1^{\alpha_1} \cdots \partial x_n^{\alpha_n}} \in L_p(\Omega)$ порядка $|\alpha| \leqslant l$. Норма в $W_p^l(\Omega)$ определена равенством

$$\begin{cases} \|u\|_{W_p^l} \equiv ||u||_{l,p,\Omega} = \left(\sum\limits_{|\alpha| \leqslant l} ||D^\alpha u||_{L_p}^p \right)^{1/p} & \text{для } 1 \leqslant p < \infty; \\ \|u\|_{W_\infty^l} \equiv \|u\|_{l,\infty,\Omega} = \max\limits_{|\alpha| \leqslant l} \|D^\alpha u\|_{L_\infty} & \text{для } p = +\infty. \end{cases} \tag{2.1}$$

Пространство $W_p^l(\Omega)$ является банаховым при всех $1 \leqslant p \leqslant \infty$. Более того, $W_2^l(\Omega) = H^l(\Omega)$ – гильбертово пространство со скалярным произведением

$$(u, v)_{H^l} = \int\limits_\Omega \sum_{|\alpha| \leqslant l} D^\alpha u \, D^\alpha v \, dx.$$

Множество $C^\infty(\bar{\Omega})$ всюду плотно в $W_p^l(\Omega)$. т.е. пространство $W_p^l(\Omega)$ можно определить как замыкание множества $C^\infty(\bar{\Omega})$ по норме пространства $W_p^l(\Omega)$, определенной равенством (2.1).

Замыкание множества $C_0^\infty(\Omega)$ по норме пространства $W_p^l(\Omega)$ приводит к пространству $\overset{\circ}{W}{}_p^l(\Omega)$, которое является подпространством $W_p^l(\Omega)$.

Функция из пространства Соболева $u \in W_p^{(l)}(\Omega)$ в действительности – это класс функций, различающихся на множестве нулевой меры. В связи с этим мы не можем в общем случае определить сужение (след) функции из пространства Соболева на многообразии меньшей размерности, чем размерность Ω, в частности на $\partial\Omega$. Тем не менее, при некоторых дополнительных предположениях относительно параметров p, l можно дать корректное определение следа функции из $W_p^{(l)}(\Omega)$, которое обобщает обычное понятие сужения непрерывной в $\overline{\Omega}$ функции на любую ее часть. Мы не будем обсуждать эти вопросы, однако отметим, что если $u \in W_p^l(\Omega), l \geqslant 1$, то существуют следы на границе $\partial\Omega$ всех производных до порядка $l-1$ из пространства $L_p(\partial\Omega$ и

$$||D^\alpha u||_{L_p(\partial\Omega)} \leqslant M||u||_{W_p^l(\Omega)} \quad \forall u \in W_p^l(\Omega),\ \forall \alpha :\ |\alpha| \leqslant l-1. \tag{2.2}$$

Это позволяет определить пространства $\overset{\circ}{W}{}_p^l(\Omega)$ для рассматриваемых областей с гладкими и кусрчно гладкими границами следующим образом: $\overset{\circ}{W}{}_p^l(\Omega) = \{u \in W_p^l(\Omega) : D^\alpha u = 0$ на $\partial\Omega$ для $|\alpha| \leqslant l-1\}$.

Далее используем обозначение $H_0^l(\Omega)$ для гильбертова пространства $\overset{\circ}{W}{}_2^l(\Omega)$.

Эквивалентные нормировки пространств Соболева

Напомним, что две нормы $\|\cdot\|$ и $\|\cdot\|_1$ на линейном пространстве X называются эквивалентными , если существуют постоянные c_1 и c_0 такие, что

$$c_0||u|| \leqslant ||u||_1 \leqslant c_1||u|| \ \ \forall u \in X.$$

В пространстве $\overset{\circ}{W}{}_p^1(\Omega), 1 \leqslant p < \infty$, исходной норме эквивалентна норма $\Big(\sum\limits_{|\alpha|=l} ||D^\alpha u||_{L_p}^p\Big)^{1/p}$. Скалярное произведение в $H_0^l(\Omega) \equiv W_0^{l,2}(\Omega)$ можно определить равенством $(u,v)_{H_0^l} = \sum\limits_{|\alpha|=l}(D^\alpha u, D^\alpha v)_{L_2}$.

Пусть $\Gamma_0 \subset \partial\Omega,\ \operatorname{meas}\Gamma_0 \neq 0,\ 1 \leqslant p < \infty$. В пространстве $V = \{u \in W_p^1(\Omega) : u = 0$ п. вс. на $\Gamma_0\}$ норма $\|u\| = \Big(\int\limits_\Omega |\nabla u|^p dx\Big)^{1/p}$ эквивалентна исходной норме (2.1).

В частности, в пространстве $H_0^1(\Omega)$ норма $\|.\|_{1,\Omega}$ эквивалентна $\|u\| =$

$\left(\int\limits_\Omega |\nabla u|^2 dx\right)^{1/2}$. Это следует из т.н. неравенства Фридрихса:

$$\left(\int\limits_\Omega u^2\, dx\right)^{1/2} \leqslant c_\Omega \left(\int\limits_\Omega |\nabla u|^2\, dx\right)^{1/2} \ \forall u \in H_0^1(\Omega) \tag{2.3}$$

с постоянной c_Ω, зависящей лишь от области.

Вложение пространств

Напомним что два линейное нормированное пространство X непрерывно вложено в Y $(X \subset Y)$, если существует линейный непрерывный оператор P, переводящий любой элемент $u \in X$ в тот же самый элемент $u \in Y$ так, что

$$\|u\|_Y \leq c\, \|u\|_X \quad \forall u \in X$$

с постоянной c, не зависящей от $u \in X$. Если при этом оператор P – компактный, то вложение X в Y непрерывно и компактно (то есть вполне непрерывно).

Справедливы, в частности, следующие теоремы вложения для пространств Соболева ($\Omega \subset \mathbb{R}^n$, $l \geqslant 0$ – целое, $1 \leqslant p \leqslant \infty$):

1. Если $lp < n$, то пространство $W_p^l(\Omega)$ непрерывно вложено в $L_q(\Omega)$ для любого $q \leqslant q_0 \equiv np/(n - lp)$ и компактно для $q < q_0$.

2. Если $lp = n$, то пространство $W_p^l(\Omega)$ вполне непрерывно вложено в $L_q(\Omega)$ для любого $q < \infty$.

3. Если $lp > n$, то пространство $W_p^l(\Omega)$ вполне непрерывно вложено в $C(\overline{\Omega})$.

4. Если $lp < n$, то пространство $W_p^l(\Omega)$ непрерывно вложено в $L_q(\partial\Omega)$ для любого $q \leqslant q_1 \equiv (n-1)p/(n - lp)$ и компактно для $q < q_1$.

5. Если $lp = n$ то пространство $W_p^l(\Omega)$ вполне непрерывно вложено в $L_q(\partial\Omega)$ для любого $q < \infty$.

2.2 Линейные уравнения в пространстве Гильберта

Пусть H — вещественное пространство Гильберта со скалярным произведением $(\cdot,\cdot)$ и нормой $\|\cdot\|$. Обозначим через $a(u,v)$ непрерывную билинейную форму: $a(\cdot,\cdot) : H \times H \to \mathbb{R}$,

$$|a(u,v)| \leqslant M\|u\|\|v\| \tag{2.4}$$

$f : H \to \mathbb{R}$ — непрерывная линейная форма (линейный функционал). Вследствие теоремы Рисса об общем виде линейного ограниченного функционала форма $a(\cdot,\cdot)$ порождает линейный ограниченный оператор A, определяемый соотношением

$$(Au,v) = a(u,v) \quad \forall u,v \in H. \tag{2.5}$$

Функционалу f будем сопоставлять элемент $f \in H$, определяемый по теореме Рисса (единственным образом) тождеством

$$f(u) = (f,u) \quad \forall u \in H. \tag{2.6}$$

В дальнейшем мы часто будем рассматривать уравнения вида

$$Au = f, \tag{2.7}$$

где $f \in H$ — заданный элемент.

Эквивалентная формулировка: найти такой элемент $u \in H$, что

$$a(u,v) = f\,(v) \quad \forall v \in H. \tag{2.8}$$

Форма $a(\cdot,\cdot)$ называется положительно определенной (эллиптичной), если существует такая положительная постоянная m, что

$$a(u,u) \geqslant m\|u\|^2 \quad \forall u \in H. \tag{2.9}$$

Соответствующий оператор A также будем называть положительно определенным.

Теорема 2.1. *(Теорема Лакса-Мильграма) Пусть оператор A ограничен и положительно определен. Тогда задача (2.7) при любом $f \in H$ имеет единственное решение.*

Ограниченная билинейная форма называется симметричной, если

$$a(u, v) = a(v, u) \quad \forall u, v \in H.$$

Линейный оператор A, порожденный симметричной формой, называется самосопряженным. Ясно, что

$$(Au, v) = (Av, u) \quad \forall u, v \in H.$$

Пусть оператор A самосопряжен. Свяжем с уравнением (2.7) задачу минимизации квадратичного функционала: найти такой элемент $u \in H$, что

$$F(u) = \min_{v \in H} F(v), \quad F(v) = \frac{1}{2}a(v, v) - f(v). \tag{2.10}$$

Теорема 2.2. *Пусть оператор A самосопряжен и положительно определен. Тогда задачи (2.7), (2.10) эквивалентны.*

2.3 Обобщенные решения эллиптических краевых задач 2-ого порядка

Задача о равновесии мембраны

Понятие обобщенного решения естественным образом возникает при изучении вариационных постановок задач математической физики.

Рассмотрим в качестве примера задачу о равновесии мембраны, жестко закрепленной по контуру и находящейся под действием внешней силы f. Эта задача может быть сформулирована как задача об отыскании функции перемещений точек мембраны $u(x)$, доставляющей минимальное значение интегралу (потенциальной энергии мембраны)

$$F(u) = \frac{1}{2}\int\limits_{\Omega} T|\nabla u|^2 dx - \int\limits_{\Omega} fu\, dx \tag{2.11}$$

на множестве функций, равных нулю на границе $\partial\Omega$ области $\Omega \subset \mathbb{R}^2$. Здесь $T = T(x)$ — заданная непрерывная функция — натяжение мембраны.

Получим уравнение равновесия мембраны как необходимое условие минимума (уравнение Эйлера) функционала (2.11). Пусть функция u доставляет минимум $F(u)$. Будем считать, что u – дважды непрерывно дифференцируемая, функция T — непрерывно дифференцируемая в Ω. Возьмем произвольную, дважды непрерывно дифференцируемую на Ω

функцию η, равную нулю на $\partial\Omega$. Числовая функция $\varphi(t) = F(u + t\eta)$ достигает минимального значения при $t = 0$, поэтому $\varphi'(0) = 0$. Легко подсчитать, что $\varphi'(0) = \int\limits_\Omega T\nabla u \cdot \nabla\eta \; dx - \int\limits_\Omega f\eta \; dx$, поэтому, если u — решение задачи (2.15), то

$$\int\limits_\Omega T\nabla u \cdot \nabla\eta \, dx = \int\limits_\Omega f\eta \, dx \quad \forall\eta.$$

При помощи формулы интегрирования по частям из этого равенства получим

$$\int\limits_\Omega (\operatorname{div}(T\nabla u) - f)\,\eta \; dx = 0. \tag{2.12}$$

Покажем, что $\operatorname{div}(T\nabla u) - f = 0$ всюду в области Ω. Если предположить противное, то найдется точка $x \in \Omega$, такая, что $(\operatorname{div}(T\nabla u) - f)(x) \neq 0$. Не ограничивая общности, можно считать, что

$$(\operatorname{div}(T\nabla u) - f)(x) > 0.$$

В силу непрерывности функции $\operatorname{div}(T\nabla u) - f$ это неравенство сохраняется и для некоторой окрестности точки x, которую можно считать лежащей внутри Ω. Выберем теперь функцию η так, чтобы она была положительной внутри указанной окрестности и тождественно равной нулю вне этой окрестности. Тогда интеграл в левой части равенства (2.12) будет положительным, что по доказанному невозможно ни при какой функции η, следовательно, $\operatorname{div}(T\nabla u) - f = 0$ всюду в области Ω Итак, уравнение равновесия мембраны имеет вид

$$-\operatorname{div}(T\nabla u) = f, \quad x \in \Omega. \tag{2.13}$$

Присоединяя к уравнению (2.13) условия жесткого закрепления

$$u(x) = 0, \quad x \in \partial\Omega, \tag{2.14}$$

получим граничную задачу Дирихле для отыскания функции u.

Под классическим решением этой задачи понимают функцию u, непрерывную в области $\overline{\Omega}$, дважды непрерывно дифференцируемую в открытой области Ω и удовлетворяющую уравнению (2.13) и граничному условию (2.14).

Заметим, что функционал (2.11) определен на гораздо более широком множестве функций. Именно, он, очевидно, имеет смысл для любой функции $u \in H^1(\Omega)$. Учтем граничные условия (2.14), полагая, что

$u \in H_0^1(\Omega)$. Таким образом, возникает задача об отыскании минимума функционала F на пространстве $H_0^1(\Omega)$

$$F(u) = \min_{v \in H_0^1(\Omega)} F(v). \tag{2.15}$$

Решение этой задачи естественно назвать обобщенным решением задачи Дирихле (2.13), (2.14).

Обобщенной постановке задачи (2.13), (2.14) можно придать и другую, эквивалентную, форму, используя результат теоремы 2.2. Именно, u — решение задачи (2.15) тогда и только тогда, когда выполнено интегральное тождество

$$\int\limits_{\Omega} T\nabla u \cdot \nabla \eta \, dx = \int\limits_{\Omega} f\eta \, dx \quad \forall \eta \in H_0^1(\Omega). \tag{2.16}$$

К интегральному тождеству (2.16) можно придти и исходя из поточечной записи задачи в виде (2.13), (2.14). Именно, пусть u – классическое решение задачи (2.13), (2.14). Умножим уравнение (2.13) на некоторую дважды непрерывно дифференцируемую на Ω функцию η, равную нулю на $\partial\Omega$ и проинтегрируем полученное равенство по области Ω. Получим

$$-\int\limits_{\Omega} \operatorname{div}(T\nabla u)\, \eta \, dx = \int\limits_{\Omega} f\eta \, dx. \tag{2.17}$$

Преобразуем интеграл в левой части равенства (2.17) при помощи формулы интегрирования по частям, учитывая при этом граничное условие для функции η. В результате придем к тождеству вида (2.16).

Таким образом, если u – классическое решение задачи (2.13), (2.14), то оно удовлетворяет интегральному тождеству (2.16), т.е. является и обобщенным решением. Обратно, если обобщенное решение u задачи (2.13), (2.14) – дважды непрерывно дифференцируемая в Ω функция, а функция T непрерывно дифференцируема в Ω, то u – классическое решение задачи (2.13), (2.14).

Краевые задачи для линейного эллиптического уравнения общего вида

Задача Дирихле

Пусть Ω — ограниченная область с липшицевой границей $\partial\Omega$. Будем искать функцию $u \in C^2(\Omega) \cap C\left(\overline{\Omega}\right)$, удовлетворяющую уравнению

$$-\sum_{i,j=1}^{n} \frac{\partial}{\partial x_i}\left(a_{ij}\frac{\partial u}{\partial x_j}\right) + \sum_{i=1}^{n} a_i \frac{\partial u}{\partial x_i} + a_0 u = f, \quad x \in \Omega, \tag{2.18}$$

и граничному условию

$$u(x) = 0, \quad x \in \partial\Omega. \tag{2.19}$$

Задача (2.18), (2.19) называется первой краевой задачей или задачей Дирихле.

Относительно коэффициентов и правой части уравнения (2.18) будем предполагать выполненными следующие условия

$$a_{ij} \in C^1(\Omega)\ \forall i,j, \quad a_i \in C^0(\Omega)\ \forall i, \quad f \in C^0(\Omega). \tag{2.20}$$

Вводя в рассмотрение матрицу $A(x) = \{a_{ij}\}_{i,j=1}^n$ и вектор $a = (a_1, \ldots, a_n)^T$, уравнению (2.18) можно придать более компактный вид

$$-\mathrm{div}\,(A\nabla u) + a \cdot \nabla u + a_0 u = f, \quad x \in \Omega.$$

В дальнейшем предполагается выполненным условие эллиптичности. Это означает, что матрица A равномерно по $x \in \overline{\Omega}$ положительно определена:

$$A(x)\xi \cdot \xi \geqslant c_0 |\xi|^2 \quad \forall \xi \in \mathbb{R}^n,\ x \in \overline{\Omega} \quad c_0 = \mathrm{const} > 0. \tag{2.21}$$

В силу ограниченности коэффициентов справедливо, очевидно, и обратное неравенство

$$A(x)\xi \cdot \xi \leqslant c_1 |\xi|^2 \quad \forall \xi \in \mathbb{R}^n, x \in \overline{\Omega}. \tag{2.22}$$

Сформулируем понятие обобщенного решения задачи (2.18), (2.19). Умножим уравнение на гладкую функцию η, удовлетворяющую условию (2.19), и преобразуем интеграл, содержащий старшие производные, при помощи формулы интегрирования по частям. Получим

$$\int\limits_{\Omega} \left((A\nabla u \cdot \nabla)\,\eta\, dx + (a \cdot \nabla u)\,\eta + a_0 u \eta \right) dx = \int\limits_{\Omega} f\eta\, dx. \tag{2.23}$$

Равенство (2.23) – интегральное тождество, соответствующее задаче (2.18), (2.19). Его естественно положить в основу определения обобщенного решения. Понятно, что (2.23) сохраняет смысл для любых функций u, $\eta \in H_0^1(\Omega)$. Условия на коэффициенты и правую часть уравнения также могут быть ослаблены. Именно, достаточно считать, что

$$a_{ij},\ a_i \in L_\infty(\Omega), \quad f \in L_2(\Omega). \tag{2.24}$$

Функция $u \in H_0^1(\Omega)$ называется обобщенным решением задачи (2.18), (2.19), если для любой функции $\eta \in H_0^1(\Omega)$ выполнено интегральное тождество (2.23).

Исследуем существование и единственность обобщенного решения задачи (2.18), (2.19) в частном случае, когда

$$a_i = 0 \ \text{ для } i = 1, 2, ..n, \ \ a_0(x) \geqslant 0. \tag{2.25}$$

Далее будем использовать следующие обозначения для норм:

$$\|u\|_0^2 = \int\limits_\Omega u^2 dx, \ \|u\|_1^2 = \int\limits_\Omega |\nabla u|^2 dx.$$

Определим билинейную форму $a(\cdot,\cdot) : H_0^1(\Omega) \times H_0^1(\Omega) \to \mathbb{R}$ при помощи соотношения

$$a(u, \eta) = \int\limits_\Omega (A\nabla u \cdot \nabla \eta + a_0 u \eta)\, dx. \tag{2.26}$$

При выполнении условий (2.24) форма $a(\cdot,\cdot)$ непрерывна. Действительно для любых $u, v \in H_0^1(\Omega)$ имеем

$$|a(u, v)| \leqslant \int\limits_\Omega (c_1 |\nabla u|\, |\nabla v| + c_3 |u||v|)\, dx.$$

Используя неравенство Фридрихса (2.3), получим

$$|a(u, v)| \leqslant c_1 \left|u\right|_1 \left|v\right|_1 + c_3 \left|u\right|_0 \left|v\right|_0 \leqslant c \left|u\right|_1 \left|v\right|_1 ,$$

где $c = c_1 + c_3 c_\Omega$, c_Ω — постоянная из неравенства Фридрихса.

Форма $a(\cdot,\cdot)$ положительно определена. Для доказательства положим $u = v$ в равенстве (2.26), получим

$$a(u, u) \geqslant c_0 \int\limits_\Omega |\nabla u|^2 dx = c_0 |u|_1^2.$$

Из свойств $a(\cdot,\cdot)$ и теоремы 2.1 Лакса-Мильграма непосредственно вытекает существование единственного решения задачи (2.18), (2.19) при любой правой части.

Особо остановимся на важном для приложений случае, когда форма $a(\cdot,\cdot)$, соответствующая задаче (2.18), (2.19) при выполнении (2.25) симметрична. Для этого достаточно потребовать выполнения условий

$$A = A^T.$$

В этом случае задача отыскания обобщенного решения задачи (2.18), (2.19) эквивалентна задаче минимизации

$$F(u) = \min_{v \in H_0^1(\Omega)} \{F(v) = \frac{1}{2} \int\limits_\Omega \left(A|\nabla u|^2 + a_0 u^2\right) dx - \int\limits_\Omega f u\; dx\}. \tag{2.27}$$

Функционал (2.27) принято называть энергетическим функционалом.

Задача Робина (3-я краевая задача)

Возможна постановка и других краевых условий для уравнения (2.18). Рассмотрим так называемую третью краевую задачу. Под классическим решением этой задачи понимается функция $u \in C^1\left(\overline{\Omega}\right) \cap C^2(\Omega)$, удовлетворяющая уравнению (2.18) и граничному условию

$$A\nabla u \cdot \nu + \sigma u = g, \quad x \in \partial\Omega. \tag{2.28}$$

Здесь $\sigma = \sigma(x), g = g(x)$ – заданные функции, ν – единичный вектор внешней нормали. Предполагаются выполненными условия гладкости (2.20), функции σ, g считаются непрерывными.

Замечание 2.1. *Выражение $A\nabla u \cdot \nu$ называется конормальной производной функции u. В простейшем случае, когда $A = E$ — единичная матрица, то есть старшие производные уравнения (2.18) образуют оператор Лапласа*

$$\Delta u = \frac{\partial^2 u}{\partial x_1} + \frac{\partial^2 u}{\partial x_2} + \dots + \frac{\partial^2 u}{\partial x_n},$$

это выражение совпадает с нормальной производной функции u.

Интегральное тождество для задачи (2.18), (2.28). имеет вид

$$\int\limits_{\Omega} \left(A\nabla u \cdot \nabla \eta + (a \cdot \nabla u)\,\eta + a_0 u\eta\right) dx + \int\limits_{\partial\Omega} \sigma u\eta\, dx = \int\limits_{\Omega} f\eta\, dx + \int\limits_{\partial\Omega} g\eta\, dx. \tag{2.29}$$

Обобщенным решением задачи (2.18), (2.28) называется функция $u \in H^1(\Omega)$, удовлетворяющая интегральному тождеству (2.29) при любой функции $\eta \in H^1(\Omega)$. Как и выше, можно доказать существование единственного обобщенного решения задачи (2.18), (2.28).

Глава 3

Элементы выпуклого анализа и теории вариационных неравенств

3.1 Выпуклые функции и субдифференциалы

Функция $F : \mathbb{R}^n \to \overline{\mathbb{R}}$ называется:

выпуклой,

если $F(tu + (1-t)v) \leqslant tF(u) + (1-t)F(v)\ \ \forall u, v \in \mathbb{R}^n,\ \ \forall t \in (0,1)$;

строго выпуклой,

если $F(tu + (1-t)v) < tF(u) + (1-t)F(v)\ \forall u \neq v \in \mathbb{R}^n,\ \forall t \in (0,1)$;

собственной,

если $F(u) > -\infty\ \forall u \in \mathbb{R}^n$ и эффективное множество $\operatorname{dom} F \equiv \{u \in \mathbb{R}^n : F(u) < +\infty\}$ функции F не пусто;

полунепрерывной снизу,

если $u^k \to u \Rightarrow \liminf\limits_{k\to\infty} F(u^k) \geqslant F(u)$.

Множество $\operatorname{epi} F = \{[u, c] : u \in \mathbb{R}^n,\ c \in \mathbb{R},\ F(u) \leqslant c\} \subset \mathbb{R}^{n+1}$ называется надграфиком функции F. Ортогональная в $\mathbb{R}^{n+1}$ проекция надграфика $\operatorname{epi} F$ на $\mathbb{R}^n$ совпадает с эффективным множеством $\operatorname{dom} F$.

Лемма 3.1. *Для выпуклости функции F необходимо и достаточно, чтобы надграфик* $\operatorname{epi} F$ *был выпуклым множеством.*

Лемма 3.2. *Для полунепрерывности снизу функции F необходимо достаточно, чтобы надграфик* $\operatorname{epi} F$ *был замкнутым множеством.*

Пусть функция $F : \mathbb{R}^n \to \overline{\mathbb{R}}$ — выпуклая, собственная и полунепрерывная снизу. Вектор $\mu \in \mathbb{R}^n$ - это субградиент функции F в точке u, если

$$F(v) - F(u) \geqslant (\mu, v - u) \ \ \forall v \in \mathbb{R}^n.$$

Множество субградиентов в точке u образует субдифференциал $\partial F(u)$. Многозначный оператор ∂F имеет область определения $D(\partial F) \subseteq \operatorname{dom} F$ и область значений в $\mathbb{R}^n$. Отметим также, что множества внутренних точек $\operatorname{int} D(\partial F)$ и $\operatorname{int} \operatorname{dom} F$ совпадают. Чтобы подчеркнуть многозначность оператора ∂F, часто пишут $\partial F : \mathbb{R}^n \to 2^{\mathbb{R}^n}$.

Из определения субдифференциала можно вывести следующие свойства:

- Если выпуклая функция F дифференцируема в точке $u \in \mathbb{R}^n$, то ее субдифференциал $\partial F(u)$ содержит единственный элемент — это градиент функции:

$$\partial F(u) = \{\nabla F(u)\}, \ \ \nabla F(u) = \left(\frac{\partial F}{\partial u_1}(u), \frac{\partial F}{\partial u_2}(u), \ldots, \frac{\partial F}{\partial u_n}(u)\right)^T.$$

- Задача отыскания минимума функции F равносильна решению включения, а именно,

$$u = \arg\min_{v \in \mathbb{R}^n} F(v) \ \Leftrightarrow \ 0 \in \partial F(u). \tag{3.1}$$

- Субдифференциал ∂F — монотонный оператор:

$$(\mu_1 - \mu_2, u_1 - u_2) \geqslant 0 \ \ \forall u_1, u_2 \in D(\partial F), \ \forall \mu_i \in \partial F(u_i).$$

  В дальнейшем мы будем для сокращения записи писать

$$(\partial F(u_1) - \partial F(u_2), u_1 - u_2) \geqslant 0 \ \ \forall u_1, u_2 \in D(\partial F).$$

- субдифференциал $\partial F(u)$ в точке $u \in D(\partial F)$ — это выпуклое и замкнутое множество в $\mathbb{R}^n$.

Элементы субдифференциального исчисления

Лемма 3.3. *Пусть $\varphi : \mathbb{R}^n \to \overline{\mathbb{R}}$ — выпуклая, собственная и полунепрерывная снизу функция. Тогда*

- *если $F(u) = \varphi(u + u_0)$, то $\partial F(u) = \partial\varphi(u + u_0)$;*

- *если* $F(u) = \lambda\varphi(u)$, *где* $\lambda > 0$, *то* $\partial F(u) = \lambda\partial\varphi(u)$;
- *если* $F(u) = \varphi(\lambda u)$, *где* $\lambda > 0$, *то* $\partial F(u) = \lambda\partial\varphi(\lambda u)$.

Лемма 3.4. *Пусть* $F : \mathbb{R}^n \to \overline{\mathbb{R}}$ *и* $G : \mathbb{R}^n \to \overline{\mathbb{R}}$ *— выпуклые, полунепрерывные снизу и собственные функции и* $\operatorname{int}\operatorname{dom} F \cap \operatorname{dom} G \neq \emptyset$. *Тогда*

$$\partial(F + G) = \partial F + \partial G.$$

Следствие 3.1. *Если* $G : \mathbb{R}^n \to \overline{\mathbb{R}}$ *— выпуклая, полунепрерывная снизу и собственная функция, а* $F : \mathbb{R}^n \to \mathbb{R}$ *— выпуклая и дифференцируемая функция то*

$$\partial(F + G) = \nabla F + \partial G.$$

Лемма 3.5. *Пусть* $\Lambda \in \mathbb{R}^{n\times m}$ *—* $n \times m$ *матрица,* $F : \mathbb{R}^n \to \overline{\mathbb{R}}$ *— выпуклая, полунепрерывная снизу, собственная функция и существует вектор* $\Lambda\bar{u} \in \operatorname{int}\operatorname{dom} F$. *Тогда сложная функция* $F \circ \Lambda : \mathbb{R}^m \to \overline{\mathbb{R}}$ *— выпуклая, полунепрерывная снизу и собственная. При этом справедливо равенство*

$$\partial[F(\Lambda u)] = \Lambda^T \partial F(\Lambda u).$$

Примеры выпуклых, полунепрерывных снизу функций и их субдифференциалов.

Пример 3.1. Субдифференциал скалярной функции $F(u) = u^+ = \max\{0, u\}$ — это многозначная функция Хевисайда:

$$\partial F(u) \equiv H(u) = \begin{cases} 0 \text{ при } u < 0; \\ [0,1] \text{ при } u = 0; \\ 1 \text{ при } u > 0 \end{cases}$$

с областью определения $D(\partial F) = \operatorname{dom} F = \mathbb{R}$.

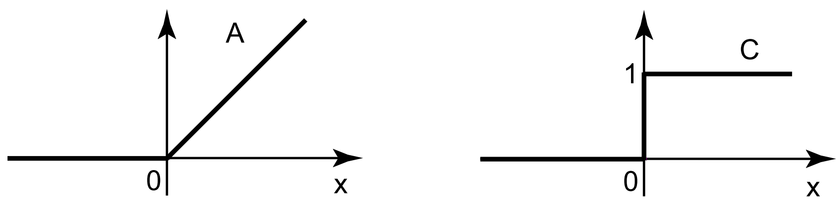

Рис. 3.1: Функция $F(u) = u^+$ и ее субдифференциал

Пример 3.2. Индикаторная функция $I_{[a,b]}$ отрезка $[a,b]$ и ее субдифференциал

$$I_{[a,b]}(u) = \begin{cases} 0, \ u \in [a,b] \\ +\infty, \ u \notin [a,b], \end{cases} \qquad \partial I_{[a,b]}(u) = \begin{cases} (-\infty, 0] \text{ для } u = a, \\ 0 \text{ для } u \in [a,b], \\ [0, +\infty) \text{ для } u = b. \end{cases}$$

Область определения оператора $\partial I_{[a,b]}$ совпадает с эффективным множеством функции $I_{[a,b]}$: $D(\partial I_{[a,b]}) = \operatorname{dom} I_{[a,b]} = [a,b]$.

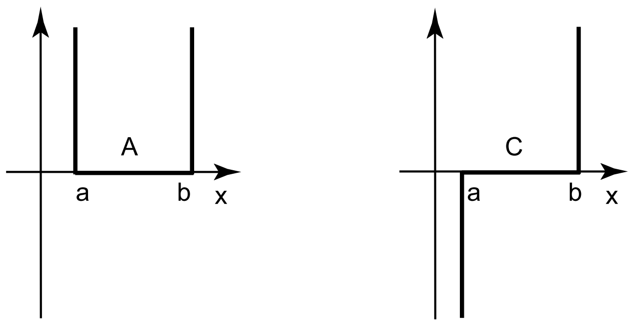

Рис. 3.2: Функция $F(x) = I_{[a,b]}(x)$ и ее субдифференциал

Пример 3.3. Пусть $F(u) = \|u\|$ — евклидова норма вектора $u \in \mathbb{R}^n$. Тогда

$$\partial F(u) = \begin{cases} \nabla F(u) = \dfrac{u}{\|u\|} \text{ для } u \neq 0, \\ \{v \in \mathbb{R}^n : \ \|v\| \leqslant 1\} - \text{ замкнутый единичный шар для } u = 0, \end{cases}$$

$D(\partial F) = \operatorname{dom} F = \mathbb{R}^n$.

Пример 3.4. Индикаторная функция выпуклого замкнутого множества $K \in \mathbb{R}^n$ ее субдифференциал

$$I_K(u) = \begin{cases} 0, \ u \in K \\ +\infty, \ u \notin K, \end{cases} \qquad \partial I_K(u) = \{\mu \in \mathbb{R}^n : (\mu, v - u) \leqslant 0 \ \forall v \in K\},$$

$D(\partial I_K) = \operatorname{dom} I_K = K$. Из определения следует, что $\partial I_K(u) = \{0\}$, если u —- внутренняя точка множества K. Если u принадлежит границе ∂K множества K, то $\partial I_K(u)$ — это так называемый нормальный конус в точке $u \in \partial K$.

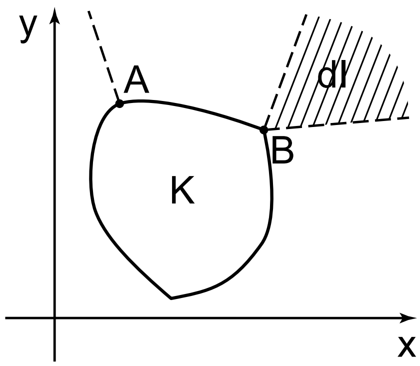

Рис. 3.3: Выпуклое замкнутое множество K в двумерном случае и нормальные конусы в точках A и B

Функция $F : \mathbb{R}^n \to \overline{\mathbb{R}}$ называется сепарабельной, если

$$F(u) = \sum_{i=1}^{n} F_i(u_i), \quad F_i : \mathbb{R} \to \overline{\mathbb{R}}.$$

Сепарабельная функция F является выпуклой, собственной и полунепрерывной снизу тогда и только тогда, когда теми же свойствами обладают все функции F_i. При этом $\operatorname{dom} F = \operatorname{dom} F_1 \times \operatorname{dom} F_2 \times \cdots \times \operatorname{dom} F_n$. Субдифференциал сепарабельной функции — это диагональный оператор

$$\partial F = \operatorname{diag}(\partial F_1, \partial F_2, \dots, \partial F_n),$$

т.е. для любого $u \in \operatorname{dom} F \subset \mathbb{R}^n$ вектор $\partial F(u)$ имеет вид $\partial F(u) = (\partial F_1(u_1), \partial F_2(u_2), \dots, \partial F_n(u_n))^T$.

Пример 3.5. Частный случай сепарабельной функции — индикаторная функция множества $K = \prod_{i=1}^{n} [a_i, b_i]$, где $-\infty \leqslant a_i < b_i \leqslant +\infty$, а обозначение $[a_i, b_i]$ использовано как для конечного отрезка, так и для бесконечного или полубесконечного промежутка:

$$I_K(u) = \sum_{i=1}^{n} I_{[a_i,b_i]}(u_i), \quad \partial I_K = \operatorname{diag}(\partial I_{[a_1,b_1]}, \dots, \partial I_{[a_n,b_n]}).$$

3.2 Примеры включений с многозначными операторами, допускающих прямое решение

Рассмотрим включение

$$Bu + \partial\varphi(u) \ni g \tag{3.2}$$

с известным вектором g и некоторой невырожденной матрицей B и приведем примеры, когда это включение может быть легко решено.

Пример 3.6. Пусть $B = \mathrm{diag}(b_{11}, b_{22}, \ldots, b_{nn})$ — диагональная $n \times n$ матрица с положительными элементами b_{ii} и $\partial\varphi = \mathrm{diag}(\partial\varphi_1, \partial\varphi_2, \ldots, \partial\varphi_n) : \mathbb{R}^n \to 2^{\mathbb{R}^n}$ — диагональный оператор (т. е. субдифференциал сепарабельной функции $\varphi(u) = \sum\limits_{i=1}^{n} \varphi_i(u_i)$). Тогда включение $Bu + \partial\varphi(u) \ni g$ расщепляется на n скалярных включений

$$b_{ii}u_i + \partial\varphi_i(u_i) \ni g_i, \tag{3.3}$$

или, эквивалентно, на n задач минимизации скалярных, строго выпуклых и полунепрерывных снизу функций

$$\frac{1}{2}b_{ii}u_i^2 + \varphi_i(u_i) - g_i u_i.$$

Эти задачи могут быть решены методами одномерной минимизации. Более того, в случае кусочно линейного графика $\partial\varphi_i$ задачу (3.3) можно решить прямыми вычислениями.

Пример 3.7. Включение (3.2) с треугольной невырожденной матрицей B и диагональным $\partial\varphi$ решается рекуррентно. Так, если B — левая треугольная матрица, т. е. $b_{ij} = 0$ для $j > i$, $b_{ii} > 0$, то последовательно решаются одномерные задачи

$$b_{ii}u_i + \partial\varphi_i(u_i) \ni g_i - \sum_{j<i} b_{ij}u_j,\ i = 1, 2, \ldots, n.$$

Пример 3.8. Включение

$$u + \tau\partial I_K(u) \ni g$$

с некоторым параметром $\tau > 0$ и индикаторной функцией I_K выпуклого и замкнутого множества K в $\mathbb{R}^n$ эквивалентно задаче минимизации функции

$$F(u) = \frac{1}{2}\|u\|^2 + \tau I_K(u) - (g, u),$$

которая, в свою очередь, равносильна задаче

$$\min_{u \in K}\{F_0(u) = \frac{1}{2}\|u\|^2 - (g, u)\}. \tag{3.4}$$

Поскольку

$$F_0(u) = \frac{1}{2}\|u - g\|^2 - \frac{1}{2}\|g\|^2,$$

то (3.4) — это задача минимизации евклидова расстояния $\|u - g\|$ от вектора g до множества K. Это означает, что решением рассматриваемого включения является ортогональная проекция $\mathrm{Pr}_K g$ в $\mathbb{R}^n$ вектора g на множество K.

В частности, если

$$K = \prod_{i=1}^{N} K_i,\ K_i = \{u_i : a_i \leqslant u_i \leqslant b_i\},\ \ -\infty < a_i < b_i < +\infty,$$

то

$$u_i = (\mathrm{Pr}_K g)_i = \begin{cases} a_i \text{ при } g_i \leqslant a_i \\ g_i \text{ при } a_i < g_i < b_i, \\ b_i \text{ при } g_i \geqslant b_i. \end{cases}$$

Если же, например, $K = \{u \in \mathbb{R}^n : \|u - u_0\| \leqslant \rho\}$ — шар с центром в точке u_0, то

$$\mathrm{Pr}_K g = \begin{cases} u_0 + \dfrac{\rho}{\|g - u_0\|}(g - u_0), & \text{если } \|g - u_0\| > \rho, \\ g, \quad \text{если } \|g - u_0\| \leqslant \rho. \end{cases}$$

3.3 Экстремальные задачи, вариационные неравенства и включения с положительно определенными операторами

Теорема 3.1. *Пусть $A \in \mathbb{R}^{n \times n}$ — симметричная и положительно определенная матрица, $\varphi : \mathbb{R}^n \to \overline{\mathbb{R}}$ — выпуклая, собственная и полунепрерывная снизу функция и $f \in \mathbb{R}^n$ — заданный вектор. Тогда вектор*

$u^* \in \mathbb{R}^n$ является решением задачи минимизации

$$u^* = \arg\min_{u\in\mathbb{R}^n}\{F(u) = \frac{1}{2}(Au,u) - (f,u) + \varphi(u)\} \tag{3.5}$$

в том и только том случае, если u^* удовлетворяет вариационному неравенству

$$(Au, v-u) + \varphi(v) - \varphi(u) \geqslant (f, v-u) \ \ \forall v \in \mathbb{R}^n, \tag{3.6}$$

или, эквивалентно, u^* является решением включения

$$Au + \partial\varphi(u) \ni f. \tag{3.7}$$

Следствие 3.2. *Пусть K — выпуклое и замкнутое множество в пространстве $\mathbb{R}^n$, $A \in \mathbb{R}^{n\times n}$ — симметричная и положительно определенная матрица и $f \in \mathbb{R}^n$ — заданный вектор. Тогда вектор $u^* \in K$ является решением задачи минимизации квадратичной функции*

$$u^* = \arg\min_{u\in K}\{F(u) = \frac{1}{2}(Au,u) - (f,u)\}$$

в том и только том случае, если u^* удовлетворяет вариационному неравенству

$$(Au, v-u) \geqslant (f, v-u) \ \ \forall v \in K. \tag{3.8}$$

Подчеркнем, что вариационное неравенство (3.8) может быть записано в форме вариационного неравенства (3.6) с $\varphi(u) = I_K(u)$, поэтому (3.6) является общей формой записи вариационного неравенства, и его эквивалентная запись — включение (3.7).

Теорема 3.2. *Пусть $\varphi : \mathbb{R}^n \to \overline{\mathbb{R}}$ — выпуклая, собственная и полунепрерывная снизу функция и $A \in \mathbb{R}^{n\times n}$ — положительно определенная матрица:*

$$(Au,u) \geqslant m\|u\|^2 \ \forall u \in \mathbb{R}^n,\ m > 0.$$

Тогда для любой правой части f существует единственное решение вариационного неравенства (3.6). Если u_1 и u_2 — решения вариационного неравенства (3.6) с правыми частями f_1 и f_2, соответственно, то

$$\|u_1 - u_2\| \leqslant \frac{1}{m}\,\|f_1 - f_2\|.$$

Иными словами, оператор $A + \partial\varphi$ имеет левый обратный $(A+\partial\varphi)^{-1}$, удовлетворяющий условию Липшица

$$\|(A+\partial\varphi)^{-1}f_1 - (A+\partial\varphi)^{-1}f_2\| \leqslant \frac{1}{m}\,\|f_1 - f_2\|.$$

Следствие 3.3. *Пусть* $\varphi : \mathbb{R}^n \to \overline{\mathbb{R}}$ *— выпуклая, собственная и полунепрерывная снизу функция и* $A \in \mathbb{R}^{n \times n}$ *— симметричная и положительно определенная матрица. Тогда существует единственное решение задачи минимизации* (3.5).

Глава 4

Решение вариационных неравенств с поточечными ограничениями на решение

4.1 Примеры вариационных неравенств

Пример 1. Задача о препятствии

Пусть $K = \{u \in H_0^1(\Omega) : u(x) \geqslant 0$ почти всюду в $\Omega\}$– выпуклое и замкнутое множество в $H_0^1(\Omega)$. Задачей о препятствии называется следующее вариационное неравенство:

$$u \in K : \int\limits_\Omega \nabla u \cdot \nabla(v-u)dx \geqslant \int\limits_\Omega f(v-u)dx \quad \forall v \in K \tag{4.1}$$

для заданной правой части $f(x) \in L_2(\Omega)$.

Вариационное неравенство (4.1) определяет, например, решение задачи о прогибах $u(x)$ мембраны, закрепленной по контуру $\partial\Omega$, под действием силы $f(x)$ при наличии снизу жесткой поверхности (препятствия) $u = 0$ (таким образом, решение задачи (4.1) не может быть отрицательным).

Эквивалентной формулировкой задачи о препятствии является следующая задача минимизации квадратичного функционала:

$$u = \arg\min_{v\in K} J(v), \quad J(v) = \frac{1}{2}\int\limits_\Omega |\nabla v|^2 dx - \int\limits_\Omega f v dx.$$

Из вариационного неравенства (4.1) можно вывести поточечную формулировку задачи о препятствии в случае, когда $f(x)$ и $\Delta u(x)$ принадлежат

$L_2(\Omega)$:

$$(-\Delta u - f)(x) \geqslant 0,\ u(x) \geqslant 0,\ (\Delta u + f)(x)u(x) = 0 \text{ для почти всех } x \in \Omega.$$

Область Ω разбивается на два множества: $\Omega_0 = \{x \in \Omega : u(x) = 0\}$ (т. н. коинцидентное множество) и $\Omega_+ = \{x \in \Omega : u(x) > 0\}$. Если решение $u(x)$ непрерывное, то во внутренних точках Ω_+ выполняется уравнение $-\Delta u(x) = f(x)$. Неизвестная граница $\partial\Omega_0 \setminus \partial\Omega$ называется свободной границей.

Пример $1'$. Задача о двухстороннем препятствии

Пусть $\phi(x), \psi(x) \in C(\overline{\Omega})$ и $\phi(x) \leqslant \psi(x)$ для всех $x \in \Omega$; $\phi(x) \leqslant 0 \leqslant \psi(x)$для всех $x \in \partial\Omega$. Определим выпуклое множество

$$K = \{u \in H_0^1(\Omega) : \phi(x) \leqslant u(x) \leqslant \psi(x) \text{ п.вс. в } \Omega\}$$

и рассмотрим вариационное неравенство

$$\int\limits_\Omega \nabla u \cdot \nabla(v-u)dx + \int\limits_\Omega \bar{a}\nabla u(v-u)dx \geqslant \int\limits_\Omega f(v-u)dx \quad \forall v \in K\ u \in K \quad (4.2)$$

с заданным вектором $\bar{a}$.

Если решение $u(x)$ задачи (4.2) достаточно гладкое, то область Ω представима в виде объединения подобластей

$$\Omega_- = \{x \in \Omega:\ u(x) = \phi(x)\},\ \ \Omega_+ = \{x \in \Omega:\ u(x) = \psi(x)\},$$

$$\Omega_{eq} = \{x \in \Omega:\ \phi(x) < u(x) < \psi(x)\}.$$

В Ω_{eq} справедливо уравнение $-\Delta u(x) + \bar{a}\nabla u(x) = f(x)$, в то время как в остальных подобластях решение удовлетворяет неравенствам:

$$-\Delta u(x) + \bar{a}\nabla u(x) - f(x) \geqslant 0 \text{ в } \Omega_-,\ \ -\Delta u(x) + \bar{a}\nabla u(x) - f(x) \leqslant 0 \text{ в } \Omega_+.$$

Вариационное неравенство (4.2) **не эквивалентно задаче на минимум**, так как дифференциальный оператор $-\Delta + \bar{a}\cdot\nabla$ не симметричен.

Пример 2.

Рассмотрим вариационное неравенство: найти функцию $u \in H_0^1(\Omega)$, такую что

$$\int\limits_\Omega \nabla u \cdot \nabla(v-u)dx + \int\limits_\Omega (|v| - |u|)dx \geqslant \int\limits_\Omega f(v-u)dx\ \forall v \in H_0^1(\Omega). \quad (4.3)$$

В отличие от (4.1) оно не содержит множества ограничений, но благодаря наличию недифференцируемого функционала $J(u) = \int\limits_{\Omega} |u(x)|dx$ не может быть записано в виде интегрального тождества (вариационного равенства). Задача (4.3) эквивалентна следующей: найти $u \in H_0^1(\Omega)$, доставляющую минимум выпуклому функционалу

$$J(u) = \frac{1}{2}\int\limits_{\Omega} |\nabla u|^2 dx + \int\limits_{\Omega} |u|dx - \int\limits_{\Omega} fudx.$$

Как и в предыдущих примерах область Ω может быть разбита на подмножества

$$\Omega_+ = \{x \in \Omega : u(x) > 0\}; \;\; \Omega_- = \{x \in \Omega : u(x) < 0\},$$

$$\Omega_0 = \{x \in \Omega : u(x) = 0\}.$$

Если решение $u(x)$ задачи (4.3) достаточно гладкое, то множество Ω_0 "разделяет" Ω_+ и Ω_-, функция u удовлетворяет уравнениям

$$-\Delta u(x) = f(x) - 1 \;\text{ при } x \in \Omega_+, \;\; -\Delta u(x) = f(x) + 1 \;\text{ при } x \in \Omega_-,$$

а часть границы $\partial\Omega_0 \setminus \partial\Omega$ множества Ω_0 образует свободную границу.

Пример 3. Задача Синьорини

Пусть теперь граница области Ω и состоит из двух частей: $\partial\Omega = \bar{\Gamma}_D \cup \bar{\Gamma}_C$, и пусть $V = \{u \in H^1(\Omega) : u(x) = 0$ п. в. на $\Gamma_D\}$. Определим множество

$$K = \{u : u \in V \, u(x) \geqslant 0 \text{ п. в. на } \Gamma_C\}$$

и рассмотрим вариационное неравенство: найти такое $u \in K$, что

$$\int\limits_{\Omega} \nabla u \nabla(v - u)dx \geqslant \int\limits_{\Omega} f(v - u)dx \quad \forall v \in K, \tag{4.4}$$

где $f \in L_2(\Omega)$.

Неравенство (4.4) служит моделью физической задачи об установившемся течении жидкости в области, в которой часть границы Γ_C представляет собой полупроницаемую мембрану, позволяющую жидкости, поступающей в Ω, проходить свободно, но препятствующую ее вытеканию.

Эквивалентной формулировкой (4.4) является следующая задача на минимум функционала: найти $u \in K$, доставляющую минимум выпуклому функционалу

$$J(u) = \frac{1}{2} \int\limits_{\Omega} |\nabla u|^2 dx - \int\limits_{\Omega} f u dx.$$

При дополнительном предположении о гладкости решения задача может быть сформулирована в поточечном виде:

$$\begin{cases} -\Delta u = f & \text{в } \Omega, \\ u = 0 & \text{на } \Gamma_D, \\ u \geqslant 0, \quad \dfrac{\partial u}{\partial \mathbf{n}} \geqslant 0, \quad u \dfrac{\partial u}{\partial \mathbf{n}} = 0 & \text{на } \Gamma_C, \end{cases}$$

где $\mathbf{n}$ — единичный вектор внешней нормали к $\partial\Omega$.

4.2 Аппроксимация вариационных неравенств

4.2.1 Аппроксимация с помощью метода конечных элементов

Далее будем предполагать, что область $\Omega \subset \mathbb{R}^2$ и является многоугольником. Пусть $\mathcal{T}_h = \{\delta_i\}$ — множество треугольных или прямоугольных конечных элементов конформной триангуляции области $\bar{\Omega}$, h — максимальный диаметр элементов.

Введем следующие обозначения для пространств полиномов: пусть $\mathbb{P}_1$ — пространство линейных функций, $\mathbb{Q}_1$ — пространство билинейных, т. е. линейных по каждой переменной, функций:

$$p \in \mathbb{P}_1 \Leftrightarrow p = c_{00} + c_{10} x_1 + c_{01} x_2, \qquad p \in \mathbb{Q}_1 \Leftrightarrow p = c_{00} + c_{10} x_1 + c_{01} x_2 + c_{11} x_1 x_2,$$

где $c_{i,j}$ — некоторые числа. Если используются треугольные элементы, то для аппроксимации $H^1(\Omega)$ выберем пространство непрерывных функций, линейных на каждом треугольнике:

$$V_{h,1} = \{u_h \in H^1(\Omega) \cap C(\bar{\Omega}) : u_h \in \mathbb{P}_1 \quad \forall \delta \in \mathcal{T}_h\},$$

а в случае прямоугольных элементов будем использовать пространство непрерывных функций, билинейных на каждом элементе:

$$V_{h,2} = \{u_h \in H^1(\Omega) \cap C(\bar{\Omega}) : u_h \in \mathbb{Q}_1 \quad \forall \delta \in \mathcal{T}_h\}.$$

Мы построим аппроксимации лишь для нескольких из рассмотренных выше вариационных неравенств. Аппроксимация остальных задач проводится аналогично.

Аппроксимация (4.1)

Аппроксимацию вариационного неравенства (4.1) получим, заменив в нем пространство V пространством V_h, а множество K — множеством K_h:

$$V_h = \{u_h \in H_0^1(\Omega) \cap C(\bar{\Omega}) | \; u_h \in \mathbb{P}_1(\text{или } \mathbb{Q}_1) \quad \forall \delta \in \mathcal{T}_h\},$$
$$K_h = \{u_h \in V_h | \; u_h(x) \geqslant 0 \quad \forall x \in \Omega\}.$$

В результате аппроксимация задачи (4.1) по методу конечных элементов примет вид:

$$u_h \in K_h : \int\limits_\Omega \nabla u_h \nabla (v_h - u_h) dx \geqslant \int\limits_\Omega f(v_h - u_h) dx \quad \forall v_h \in K_h. \tag{4.5}$$

Для удобства численной реализации можно интеграл в правой части (4.5) аппроксимировать с помощью составных квадратурных формул. В случае треугольного элемента δ используем квадратуру

$$\int\limits_\delta v(x) dx \approx S_\delta(v) = \frac{|\delta|}{3} \sum_{i=1}^{3} v(a_i),$$

где a_i — вершины δ, $|\delta| = \operatorname{meas} \delta$, а $v(x)$ — непрерывная функция. Соответственно, в случае прямоугольного элемента δ используем квадратуру

$$\int\limits_\delta v(x) dx \approx S_\delta(v) = \frac{|\delta|}{4} \sum_{i=1}^{4} v(a_i), \quad a_i - \text{ вершины } \delta.$$

Аппроксимация (4.4)

Заменим пространство V и множество K в задаче (4.4) аналогично предыдущему случаю:

$$V_h = \{u_h \in V \cap C(\bar{\Omega}) | \; u_h \in \mathbb{P}_1(\text{или } \mathbb{Q}_1) \quad \forall \delta \in \mathcal{T}_h\},$$
$$K_h = \{u_h \in V_h | \; u_h(x) \geqslant 0 \quad \forall x \in \Gamma_C\},$$

и получим конечномерную задачу

$$u_h \in K_h : \int\limits_\Omega \nabla u_h \nabla (v_h - u_h) dx \geqslant \int\limits_\Omega f(v_h - u_h) dx \, \forall v_h \in K_h. \tag{4.6}$$

Алгебраические формулировки сеточных вариационных неравенств. Запись в виде включений

Вернемся к аппроксимации задачи о препятствии (4.5) и запишем ее в алгебраической форме.

Поставим в соответствие $v_h \in V_h$ вектор $v \in \mathbb{R}^N$ с координатами $v_i = v_h(x_i), x_i \in \omega$, и будем использовать для этого взаимно однозначного соответствия обозначение $v \Leftrightarrow v_h$.

Пусть далее $(\cdot\,,\cdot)$ — скалярное произведение в $\mathbb{R}^N$. Положим

$$(Au,v) = \int\limits_{\Omega} \nabla u_h(x) \nabla v_h(x) dx, \quad (f,v) = \int\limits_{\Omega} f(x) v_h(x) dx,\ u \Leftrightarrow u_h,\ v \Leftrightarrow v_h.$$

и пусть $\{\varphi_i(x)\}$ — ***базис Куранта*** пространства V_h, т. е. множество функций таких, что

$$\varphi_i \in V_h, \quad \varphi_i(x_i) = 1, \quad \varphi_i(x_j) = 0, \quad i \neq j.$$

Используя разложение функций по базису: $u_h = \sum_i \varphi_i(x) u_i$ и $v_h = \sum_j \varphi_j(x) q_j$, получим, что элементы матрицы A вычисляются по формулам

$$a_{ij} = \int\limits_{\Omega} \nabla \varphi_i(x) \nabla \varphi_j(x) dx. \tag{4.7}$$

Из определения следует, что A — симметричная положительно определенная матрица:

$$(Au,u) = \int\limits_{\Omega} |\nabla u_h|^2 dx > 0 \quad u_h \in V_h, \quad u_h \not\equiv 0 \Leftrightarrow u \not\equiv 0.$$

Заметим, что как в случае $\mathbb{P}_1$-интерполяции, так и в случае $\mathbb{Q}_1$-интерполяции, ограничение $u_h(x) \geqslant 0 \quad \forall x \in \Omega$ эквивалентно ограничению на значения узловых параметров u_h: $u_h(x_i, y_j) \geqslant 0$ для всех внутренних узлов. Далее используем обозначение K для следующего множества: $K = \{u \in \mathbb{R}^N : u_i \geqslant 0\}$, тогда вариационное неравенство (4.5) можно записать как соответствующее неравенство в $\mathbb{R}^N$:

$$u \in K : \ (Au, v-u) \geqslant (f, v-u) \quad \forall\, v \in K. \tag{4.8}$$

Пусть I_K — индикаторная функция множества K:

$$I_K(u) = \begin{cases} 0, & u \in K, \\ +\infty, & u \notin K, \end{cases}$$

тогда задача (4.8) может быть записана следующим образом:

$$u \in \mathbb{R}^N : \ (Au, v-u) + I_K(v) - I_K(u) \geqslant (f, v-u) \quad \forall v \in \mathbb{R}^N. \tag{4.9}$$

Вариационное неравенство (4.9) эквивалентно включению

$$Au + P(u) \ni f, \quad P = \partial I_K. \tag{4.10}$$

Из определения множества K следует, что I_K — сепарабельный функционал, поэтому P — диагональный максимально монотонный оператор, причем в данном случае все компоненты этого оператора одинаковы: $P = \mathrm{diag}(p, \ldots, p)$, где $p(u)$ — многозначные функции

$$p(u) = \begin{cases} (-\infty, 0], & u = 0, \\ 0, & u > 0 \end{cases} \tag{4.11}$$

Оператор P в задаче Синьорини (4.6) и (**??**), при соответствующей нумерации узлов, можно представить в виде $P = \mathrm{diag}(\underbrace{p, \ldots, p}_{N_C}, 0, \ldots, 0)$, где N_C — число узлов, принадлежащих Γ_C.

4.2.2 Конечно-разностные аппроксимации вариационных неравенств

Выше были рассмотрены конечно-элементные аппроксимации вариационных задач с использованием треугольных $\mathbb{P}_1$-элементов и четырехугольных $\mathbb{Q}_1$-элементов. Если ограничиться случаем квадратных $\mathbb{Q}_1$-элементов и использовать составные квадратурные формулы трапеций для вычисления интегралов (ср. с 1.2), то в результате получим конечно-разностные аппроксимации соответствующих задач.

Именно, пусть $\bar{\omega} = \{x = (ih, jh) : 0 \leqslant i, j \leqslant m+1, \ (m+1)h = 1\}$ — равномерная сетка на области $\Omega = (0,1) \times (0,1)$. Множество внутренних узлов сетки обозначим через ω, множество граничных узлов — через $\partial\omega$. В случае, если граница области $\partial\Omega$ состоит из частей Γ_D и Γ_C, обозначим через Γ_D и Γ_C — множества узлов сетки на $\bar{\Gamma}_D$ (замыкание Γ_D) и Γ_C.

Обозначим через Δ_h конечно-разностный оператор Лапласа: $\Delta_h = \bar{\partial}_1 \partial_1 + \bar{\partial}_2 \partial_2$. Здесь

$$\bar{\partial}_1 u_i = \frac{u_{ij} - u_{i-1,j}}{h}, \ \partial_1 u_i = \frac{u_{i+1,j} - u_{ij}}{h}$$

и разностные производные $\overline{\partial}_2 u_i, \ \partial_2 u_i$ определены аналогично.

Для аппроксимации интегралов по Ω от непрерывных функций, равных нулю на границе области, используем составные квадратурные формулы трапеций

$$\int\limits_{\Omega} g(x)dx \approx S_{\Omega}(g) = h^2 \sum_{x \in \omega} g(x).$$

В результате получим следующие конечно-разностные схемы.

Задача о препятствии.

$$\begin{gathered} -\partial_1\bar{\partial}_1 u - \partial_2\bar{\partial}_2 u + p(u) \ni f_{ij}, \quad \text{в } \omega \\ u = 0, \quad \text{на } \partial\omega. \end{gathered} \tag{4.12}$$

Задача Синьорини.

$$\begin{gathered} -\partial_1\bar{\partial}_1 u - \partial_2\bar{\partial}_2 u = f_{ij}, \quad \text{в } \omega \\ -\frac{1}{2}\partial_1\bar{\partial}_1 u - \frac{1}{h}\partial_2 u + \frac{1}{h}p(u) \ni \frac{1}{2}f_{ij}, \quad \text{на } \Gamma_C \\ u = 0, \quad \text{на } \Gamma_D. \end{gathered} \tag{4.13}$$

здесь нелинейный многозначный оператор $p(u)$ определен формулой (4.11).

4.3 Итерационные методы для сеточных вариационных неравенств с положительно определенными матрицами

Аппроксимации всех рассмотренных в предыдущей главе задач записаны в виде конечномерного включения

$$Au + P(u) \ni f, \tag{4.14}$$

где $A = A^T$ — симметричная и положительно определенная матрица, P – диагональный и монотонный (многозначный) оператор. В этом разделе приведены итерационные методы для решения (4.14), результаты о их сходимости и описаны алгоритмы реализации эти методов.

4.3.1 Метод простой итерации

Применим для решения (4.14) предобусловленный метод простой итерации

$$B\frac{u^{n+1}-u^n}{\tau} + Au^n + P(u^{n+1}) \ni f, \quad n = 0, 1, 2, \ldots, \tag{4.15}$$

где предобуслвливатель B — симметричная и положительно определенная матрица, $\tau > 0$ — итерационный параметр.

Теорема 4.1. *Пусть $A = A^T > 0$, $B = B^T > 0$ и $mB \leqslant A \leqslant MB$, $m > 0$. Тогда метод (4.15) сходится при $\tau \in (0, 2/M)$ с любого начального приближения u^0. При оптимальном параметре $\tau = \tau_0 = 2/(m+M)$ справедлива оценка скорости сходимости:*

$$\|u^{n+1} - u^*\|_B \leqslant \frac{M-m}{M+m}\|u^n - u^*\|_B \quad \forall n.$$

4.3.2 Методы релаксации: Якоби, Гаусса-Зейделя, SOR, SSOR

Далее рассмотрим методы релаксационного типа, где будет использоваться представление матрицы в виде $A = D - L - L^T$, где L — строго нижняя треугольная матрица, D — диагональная матрица.

Метод Якоби

В матрично-векторном виде метод Якоби записывается следующим образом:

$$Du^{n+1} + P(u^{n+1}) \ni (L + L^T)u^n + f, \quad n = 0, 1, \ldots \tag{4.16}$$

Так как оператор P является диагональным, i-я строка включения (4.22) имеет вид

$$a_{ii}u_i^{n+1} + p_i(u_i^{n+1}) \ni \sum_{j \neq i} a_{ij}u_j^n + f_i. \tag{4.17}$$

Следовательно, одна итерация метода Якоби состоит в последовательном нахождении компонент вектора u^{n+1} из скалярного включения (4.17).

Подробно опишем, как осуществляется решение этого одномерного включения. Разделим обе части включения (4.17) на a_{ii}:

$$u_i^{n+1} + \frac{1}{a_{ii}}p_i(u_i^{n+1}) \ni F_i^n \equiv \frac{1}{a_{ii}}\left(\sum_{j \neq i} a_{ij}u_j^n + f_i\right). \tag{4.18}$$

В задаче о препятствии p_i задаются формулой (4.11). Скалярный многозначный оператор $q_i(t) = t + \dfrac{1}{a_{ii}} p_i(t)$ имеет однозначный обратный оператор. Таким образом, из включения (4.17) единственным образом находится решение u_i^{n+1}, которое определяется по формуле:

$$u_i = \begin{cases} 0, & F_i \leqslant 0, \\ F_i, & F_i > 0, \end{cases} \tag{4.19}$$

то есть решение свелось к нахождению проекции числа F_i на множество неотрицательных чисел.

Для задачи Синьорини оператор P не равен нулю только в точках границы с условием Синьорини, следовательно нахождение решения в этих точках осуществляется по формуле (4.19), в остальных точках u находится из соответствующего уравнения.

Значения $p(u_i)$ на точном решении определяются однозначно и могут быть найдены по формулам

$$p(u_i) = \begin{cases} 0, & u_i \geqslant 0, \\ a_{ii} F_i, & u_i = 0, \end{cases}$$

в задачах о препятствии (4.5) и Синьорини (4.6).

Метод Гаусса – Зейделя

Матрично-векторная запись метода Гаусса-Зейделя для решения включения (4.10) имеет вид

$$(D - L)\, u^{n+1} + P(u^{n+1}) \ni L^T u^n + f, \quad n = 0, 1, \ldots, \tag{4.20}$$

На каждой итерации метода (4.20) нужно решить включение с нижней треугольной матрицей $(D - L)$, т. е. реализация одной итерации сводится к последовательному решению одномерных включений

$$a_{ii} u_i^{n+1} + p_i(u_i^{n+1}) \ni \sum_{j>i} a_{ij} u_j^{n+1} + \sum_{j<i} a_{ij} u_j^n + f_i. \tag{4.21}$$

Здесь, в отличие от метода Якоби, на каждой итерации при вычислениях используются уже найденные ранее значения вектора u^{n+1}. С вычислительной точки зрения отличие будет лишь в определении правой части F скалярного включения (4.23), теперь i-я компонента вектора вычисляется как

$$F_i^n \equiv \frac{1}{a_{ii}} \left(\sum_{j>i} a_{ij} u_j^{n+1} + \sum_{j<i} a_{ij} u_j^n + f_i \right).$$

Метод верхней релаксации (SOR)

Поточечный метод верхней релаксации (называемый далее SOR от английского “successive over relaxation”) может быть записан как

$$\left(\frac{1}{\sigma}D - L\right)u^{n+1} + P(u^{n+1}) \ni \left(\frac{1}{\sigma} - 1\right)Du^n + L^T u^n + f, \quad n = 0, 1, \ldots \tag{4.22}$$

Отметим, что метод Гаусса-Зейделя является частным случаем (4.22) при $\sigma = 1$.

Разделим i-ю строку включения (4.22) на a_{ii} и перенесем известные величины в правую часть:

$$u_i^{n+1} + \frac{\sigma}{a_{ii}} p_i(u_i^{n+1}) \ni (1-\sigma)u_i^n + \frac{\sigma}{a_{ii}}\left(\sum_{j>i} a_{ij}u_j^{n+1} + \sum_{j<i} a_{ij}u_j^n + f_i\right) \equiv F_i^n. \tag{4.23}$$

В задачах (4.5) и (4.6), т. е. в случае, когда c_i задаются формулой (4.11), решение включения (4.23) сводится к нахождению проекции точки на множество неотрицательных чисел: $u_i = \max\{0, F_i\}$, а решение (**??**) в точках Γ_C определяется формулой

$$u_i = \begin{cases} F_i - \sigma/a_{ii}, & \text{если} \quad F_i \geqslant \sigma/a_{ii}, \\ 0, & \text{если} \quad -\sigma/a_{ii} \leqslant F_i \leqslant \sigma/a_{ii}, \\ F_i + \sigma/a_{ii}, & \text{если} \quad F_i \leqslant -\sigma/a_{ii}. \end{cases} \tag{4.24}$$

Сходимость метода (4.22) вытекает из следующей теоремы.

Теорема 4.2. *Пусть $A = A^T > 0$. Тогда итерации метода верхней релаксации (4.22) сходятся с любого начального приближения u_0 к решению включения (4.14) при $\sigma \in (0,\, 2)$.*

Замечания об алгоритмах реализации итерационных методов для конечно-разностной аппроксимации задачи о препятствии (4.12)

Метод Якоби

При написании программного кода понадобятся два массива $\hat{u}$, в котором хранятся значения текущей итерации, и u, в котором хранятся значения, полученные на предыдущей итерации.

1. задаем начальное приближение u
2. для $i = \overline{2, n-1}$, $j = \overline{2, n-1}$ вычисляем компоненты вектора $\hat{u}$ следующим образом

2.1. находим $\hat{u}_{i,j} = \dfrac{h^2}{4}\left(\dfrac{1}{h^2}(u_{i-1,j} + u_{i+1,j} + u_{i,j-1} + u_{i,j+1}) + f_{i,j}\right)$

2.2. если $\hat{u}_{i,j} < 0$, полагаем $\hat{u}_{i,j} = 0$.

3. если не выполнен критерий остановки, то полагаем $u := \hat{u}$ и переход к п. 2.

Метод Гаусса-Зейделя

В этом методе, как и в методе SOR, достаточно одного массива u. При нахождении u_{ij}, используются обновленные значения $u_{i-1,j}$ и $u_{i,j-1}$.

1. задаем начальное приближение u
2. для $i = \overline{2, n-1}$, $j = \overline{2, n-1}$ вычисляем компоненты вектора u:

   2.1. находим $u_{i,j} = \dfrac{h^2}{4}\left(\dfrac{1}{h^2}(u_{i-1,j} + u_{i+1,j} + u_{i,j-1} + u_{i,j+1}) + f_{i,j}\right)$

   2.2. если $u_{i,j} < 0$, полагаем $u_{i,j} = 0$.
3. если не выполнен критерий остановки, то переход к п. 2.

SOR-метод

1. задаем начальное приближение u
2. для $i = \overline{2, n-1}$, $j = \overline{2, n-1}$ вычисляем компоненты вектора u:

   2.1. находим $u_{i,j} = (\sigma - 1)u_i + \sigma\dfrac{h^2}{4}\left(\dfrac{1}{h^2}(u_{i-1,j} + u_{i+1,j} + u_{i,j-1} + u_{i,j+1}) + f_{i,j}\right)$

   2.2. если $u_{i,j} < 0$, полагаем $u_{i,j} = 0$.
3. если не выполнен критерий остановки, то переход к п. 2.

4.3.3 О контроле точности и критерии окончания итераций

Решая задачу итерационным методом мы хотим получить решение, близкое к точному. Возникает вопрос, когда прекратить вычисления, или, другими словами, как оценить точность приближенного решения, располагая доступными в результате расчетов данными.

Выход по погрешности

Используется в тестовых примерах, когда известно точное решение сеточной задачи. Пусть u^n — решение полученное на n-й итерации любым из рассмотренных численных методов, u — точное решение сеточной задачи. На каждой итерации вычисляется норма погрешности решения $\|u^n - u\|$, малость которой используется в качестве критерия остановки.

Как правило, используются следующие сеточные нормы (формулы соответствуют сеточным схемам на равномерной сетке шага h для двумерных задач):

$$\|u\|_C = \max_{i,j} |u_{ij}|, \quad \|u\|_{L_2} = (\sum_{i,j=1}^{n} h^2 u_{i,j}^2)^{1/2}.$$

Выход по числу итераций

Используется, когда имеются теоретические оценки скорости сходимости метода.

Пусть известна оценка нормы оператора перехода в итерационном методе, т.е. известно число q:

$$\|u^n - u\| \leqslant q\|u^{n-1} - u\|, \tag{4.25}$$

Имеем:

$$\|u^n - u\| \leqslant q\|u^{n-1} - u^*\| \leqslant q^2\|u^{n-2} - u\| \leqslant q\|u^{n-1} - u\| \leqslant \ldots \leqslant q^n\|u^0 - u\|.$$

Откуда можно найти число итераций, необходимых для достижения относительной погрешностью величины ε:

$$\varepsilon = \frac{\|u^{n-1} - u^*\|}{\|u^0 - u\|} \geqslant q^n,$$

следовательно

$$n \geqslant \frac{\ln q}{\ln \varepsilon}. \tag{4.26}$$

Формула (4.26) гарантирует, что необходимая точность будет достигнута через n итераций.

Выход по близости итераций

В качестве критерия остановки иногда используется норма разности решений, полученных на соседних итерациях. Однако такой критерий **мало подходит в случае медленно сходящихся методов**. Именно, если справедлива оценка (4.25), то можно оценить разность решений на соседних итерациях:

$$\|u^{n+1} - u^n\| \geqslant \|u^n - u^*\| - \|u^{n+1} - u^*\| \geqslant \|u^n - u^*\| - q\|u^n - u^*\| \geqslant (1-q)\|u^n - u^*\|.$$

Отсюда следует
$$\|u^n - u\| \le \frac{1}{1-q}\|u^{n+1} - u^n\|.$$
Как видно, если множитель $q < 1$ близок к единице, т. е. итерационный метод медленно сходится, то малая величина $\|u^{n+1} - u^n\|$ не означает близости u^{n+1} к точному решению. При знании значения q или оценки для него имеет смысл контролировать величину $\dfrac{1}{1-q}\|u^{n+1} - u^n\|$.

Выход по невязке

Универсальный метод.

При решении любым итерационным методом включения $Au + P(u) \ni f$ мы находим на n-ой итерации не только вектор u^n, но и единственную селекцию $\partial\Omega^n \in P(u^n)$ из множества $P(u^n)$. Вектор невязки определяется равенством
$$r^n = Au^n + \partial\Omega^n - f, \quad \partial\Omega^n \in P(u^n).$$
Если u — точное решение, то
$$Au + \partial\Omega - f = 0, \quad \partial\Omega \in P(u).$$
Представим невязку в виде $r^n = Au^n + \partial\Omega^n - f - Au + \partial\Omega + f = A(u^n - u) + \partial\Omega^n - \partial\Omega$. Имеем
$$(r^n, u^n - u) = (A(u^n - u), u^n - u) + (\partial\Omega^n - \partial\Omega, u^n - u).$$
В силу максимальной монотонности оператора P последнее слагаемое неотрицательно, следовательно
$$(r^n, u^n - u) \geqslant (A(u^n - u), u^n - u) \geqslant m\|u^n - u\|^2,$$
где m — минимальное собственное число оператора A, поэтому из оценки нормы невязки $\|r^n\|$ следует и оценка для погрешности:
$$\|u^u - u\| \leqslant m^{-1}\|r^n\|.$$

Глава 5

Решение сеточных седловых задач с ограничениями

5.1 Конечномерные седловые задачи с ограничениями.

Рассмотрим задачу

$$\begin{pmatrix} A & -B^T \\ -B & 0 \end{pmatrix} \begin{pmatrix} u \\ \lambda \end{pmatrix} + \begin{pmatrix} \partial\varphi(u) \\ 0 \end{pmatrix} \ni \begin{pmatrix} f \\ -g \end{pmatrix}, \tag{5.1}$$

где

$$A \in \mathbb{R}^{n\times n} - \text{положительно определенная матрица},$$

$$B \in \mathbb{R}^{s\times n} - \text{ матрица полного столбцового ранга: } \operatorname{rank} B = s \leqslant n, \tag{5.2}$$

$$\varphi \text{ - выпуклая, собственная и полунепрерывная снизу функция}, \tag{5.3}$$

$$f \in \mathbb{R}^n \text{ и } g \in \mathbb{R}^s - \text{ заданные векторы.}$$

Теорема 5.1. *Пусть выполнены условия* (5.2), (5.3) *и*

$$K \cap \operatorname{int}\operatorname{dom}\varphi \neq \emptyset, \quad K = \{u \in \mathbb{R}^n : Bu = g\}. \tag{5.4}$$

Тогда задача (5.1)*имеет решение* (u, λ), *и его первая компонента* u *определяется однозначно.*

Замечание 5.1. Вторая компонента решения λ задачи (5.1) может быть как единственной, так и не единственной.

Установим теперь связь решения седловой задачи (5.1) в случае симметричной матрицы A с седловой точкой соответствующей функции Лагранжа. Сначала приведем некоторые сведения о функциях Лагранжа.

Пусть $K \in \mathbb{R}^n$ и $M \in \mathbb{R}^s$ – выпуклые множества и функция $\mathcal{L} : K \times M \to \mathbb{R}$ удовлетворяет следующим предположениям:

$$\begin{aligned} \mathcal{L}(u,\lambda) &\text{ – выпуклая и полунепрерывная снизу по } u \in K \ \forall \lambda \in M; \\ -\mathcal{L}(u,\lambda) &\text{ – выпуклая и полунепрерывная снизу по } \lambda \in M \ \forall u \in K. \end{aligned} \tag{5.5}$$

$\mathcal{L}$ называется функцией Лагранжа. Точка $(u^*, \lambda^*) \in K \times M$ — её седловая точка, если

$$\mathcal{L}(u^*,\lambda) \leqslant \mathcal{L}(u^*,\lambda^*) \leqslant \mathcal{L}(u,\lambda^*) \ \ \forall u \in K, \ \forall \lambda \in M. \tag{5.6}$$

Теорема 5.2. *Пусть φ — выпуклая, собственная и полунепрерывная снизу функция, а $A \in \mathbb{R}^{n\times n}$ — симметричная и положительно полуопределенная матрица: $(Au,u) \geqslant 0 \ \forall u \in \mathbb{R}^n$. Тогда задача* (5.1) *равносильна задаче отыскания седловой точки функции Лагранжа*

$$\mathcal{L}(u,\lambda) = \frac{1}{2}(Au,u) + \varphi(u) - (Bu,\lambda) - (f,u) + (g,\lambda),$$

на множестве $\operatorname{dom}\varphi \times \mathbb{R}^s$.

5.2 Примеры сеточных седловых задач с ограничениями

5.2.1 Конечно-разностная аппроксимация одномерной задачи с ограничением на производную от решения

Рассмотрим задачу минимизации функционала

$$J(u) = \frac{1}{2}\int\limits_0^1 u'^2(x)dx - \int\limits_0^1 f(x)\,u(x)dx, \ f(x) \in C[0,1], \tag{5.7}$$

на выпуклом и замкнутом множестве $U_{ad} = \{u(x) \in H_0^1(0,1) : |u'(x)| \leqslant 1 \text{ для п. вс. } x \in (0,1)\}$ пространства $H_0^1(0,1) = \{u(x) \in H^1(0,1) : u(0) = u(1) = 0\}$.

Пусть $\{x_i = i\,h,\ i = 0, \ldots, n+1;\ (n+1)\,h = 1\}$ – равномерная сетка шага $h > 0$ на отрезке $[0,1]$, $u_i = u(x_i)$ $(u_0 = u_{n+1} = 0)$ и $f_i = f(x_i)$. Будем использовать следующие обозначения для разностных производных:

$$\partial u_i = \frac{1}{h}(u_{i+1} - u_i),\ \overline{\partial} u_i = \frac{1}{h}(u_i - u_{i-1}).$$

Конечно-разностные аппроксимации функционала цели и множества ограничений возьмем в виде:

$$F(u) = \frac{1}{2}\sum_{i=1}^{n+1}(\overline{\partial} u_i)^2 - \sum_{i=1}^{n} f_i u_i = \frac{1}{2}\left(\frac{u_1^2}{h^2} + \sum_{i=2}^{n}\left(\frac{u_i - u_{i-1}}{h}\right)^2 + \frac{u_n^2}{h^2}\right) - \sum_{i=1}^{n} f_i u_i,$$

$$K = \left\{u \in \mathbb{R}^n :\ |\overline{\partial} u_i| \leqslant 1\ \forall i = 1, \ldots, n+1\right\} =$$
$$= \{u \in \mathbb{R}^n :\ |u_i - u_{i-1}| \leqslant h\ \forall i = 2, \ldots, n,\ |u_1| \leqslant h,\ |u_n| \leqslant h\}.$$

Пусть матрица $L \in \mathbb{R}^{(n+1)\times n}$ определена равенством

$$L = h^{-1}\begin{pmatrix} 1 & 0 & 0 & \ldots & 0 & 0 \\ -1 & 1 & 0 & \ldots & 0 & 0 \\ 0 & -1 & 1 & \ldots & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & -1 & 1 \\ 0 & 0 & 0 & \ldots & 0 & -1 \end{pmatrix}.$$

L — это матрица сеточного оператора $\overline{\partial}$ на равномерной сетке шага h на отрезке $[0,1]$ с нулевыми граничными условиями:

$$(Lu)_i = \overline{\partial} u_i,\ i = 1, 2, \ldots, n+1, \text{с учетом равенств } u_0 = u_{n+1} = 0.$$

Пусть $\varphi_p(p)$ – индикаторная функция множества $\{p \in \mathbb{R}^{n+1} : |p_i| \leqslant 1\ \forall i\}$. Тогда конечно-разностная аппроксимация исходной задачи минимизации есть

$$\min_{u\in\mathbb{R}^n}\left(\frac{1}{2}(L^T L u, u) - (f, u) + \varphi_p(Lu)\right). \tag{5.8}$$

Матрица $L^T L$ — это матрица сеточного оператора $-\partial\overline{\partial}$ на равномерной сетке шага h на отрезке $[0,1]$ с нулевыми граничными условиями. Она симметрична и положительно определена. Ясно, что функция $\varphi_p \circ L$ — выпуклая, собственная и полунепрерывная снизу. Применив следствие 3.3 к теореме 3.2, получим результат о существовании единственного решения задачи минимизации (5.8).

Поскольку нулевой вектор принадлежит $\operatorname{int}\operatorname{dom}\varphi_p = \{p \in \mathbb{R}^{n+1} : |p_i| < 1 \ \forall i = 1, 2, \ldots, n+1\}$, то из леммы 3.5 следует $\partial[\varphi_p(Lu)] = L^T \partial\varphi_p(Lu)$, и согласно теореме 3.1 задача (5.8) эквивалентна включению

$$L^T L u + L^T \partial\varphi_p(Lu) \ni f. \tag{5.9}$$

Введя векторы $p = Lu$ и $\lambda \in p + \partial\varphi_p(p)$, преобразуем это включение к седловой задаче

$$\begin{pmatrix} 0 & 0 & L^T \\ 0 & E & -E \\ L & -E & 0 \end{pmatrix} \begin{pmatrix} u \\ p \\ \lambda \end{pmatrix} + \begin{pmatrix} 0 \\ \partial\varphi_p(p) \\ 0 \end{pmatrix} \ni \begin{pmatrix} f \\ 0 \\ 0 \end{pmatrix}, \tag{5.10}$$

где E – единичная $(n+1) \times (n+1)$ матрица.

Замечание 5.2. *В силу симметрии матрицы L (а значит, эквивалентности включения соответствующей задаче минимизации) можно получить седловую задачу* (5.10) *и другим способом, используя функцию Лагранжа. Именно, заменим задачу минимизации* (5.8) *эквивалентной задачей*

$$\min_{Lu=p} \left(\frac{1}{2}\|p\|^2 - (f, u) + \varphi_p(p) \right)$$

и определим функцию Лагранжа равенством

$$\mathcal{L}(u, p, \lambda) = \frac{1}{2}\|p\|^2 - (f, u) + \varphi_p(p) + (Lu - p, \lambda).$$

Тогда седловая точка функции $\mathcal{L}$ удовлетворяет системе (5.10).

Задача (5.10) — это частный случай (5.1) с $A = \begin{pmatrix} 0 & 0 \\ 0 & E \end{pmatrix}$, $B = \begin{pmatrix} -L & E \end{pmatrix}$. Матрица B имеет полный столбцовый ранг, т. к. $\det E = 1$. Вектор $(u, p) = (0, 0)$ удовлетворяет условию (5.4). Однако матрица A лишь положительно полуопределена. Для применения теоремы 5.1 о существования решения проведем эквивалентное преобразование системы (5.10).

Эквивалентное преобразование седловой задачи (5.10)

Прибавим к первому уравнению третье, умноженное слева на матрицу rL^T с положительной постоянной r, а к второму включению прибавим третье уравнение, умноженное на $-r$. В результате получим эквивалентную (5.10) систему

$$\begin{pmatrix} rL^T L & -rL^T & L^T \\ -rL & (1+r)E & -E \\ L & -E & 0 \end{pmatrix} \begin{pmatrix} u \\ p \\ \lambda \end{pmatrix} + \begin{pmatrix} 0 \\ \partial\varphi_p(p) \\ 0 \end{pmatrix} \ni \begin{pmatrix} f \\ 0 \\ 0 \end{pmatrix}. \tag{5.11}$$

Отметим, что седловая задача (5.11) характеризует стационарную точку расширенной функции Лагранжа

$$\mathcal{L}_r(u,p,\lambda)=\frac{1}{2}\|p\|^2-(f,u)+\varphi_p(p)+(Lu-p,\lambda)+\frac{r}{2}\|Lu-p\|^2,\ \ r>0.$$

При любом $r>0$ матрица $A=\begin{pmatrix} rL^TL & -rL^T \\ -rL & (1+r)E \end{pmatrix}$ в задаче (5.11) положительно определена. Более того, справедлива

Лемма 5.1. *Матрица A спектрально эквивалентна блочно диагональной и положительно определенной матрице* $A^0=\begin{pmatrix} L^TL & 0 \\ 0 & E \end{pmatrix}$:

$$m(r)(A^0x,x)\leqslant(Ax,x)\leqslant M(r)(A^0x,x)\quad\forall x=(u,p)^T.$$

Здесь $0<m(r)<M(r)$ *— минимальное и максимальное собственные числа матрицы* $\begin{pmatrix} r & -r \\ -r & 1+r \end{pmatrix}$.

Проведем теперь еще одно эквивалентное преобразование (5.10) к системе с положительно определенной и блочно треугольной матрицей A. С этой целью прибавим к первому уравнению (5.10) третье, умноженное слева на матрицу rL^T. Преобразованная седловая задача выглядит следующим образом:

$$\begin{pmatrix} rL^TL & -rL^T & L^T \\ 0 & E & -E \\ L & -E & 0 \end{pmatrix}\begin{pmatrix} u \\ p \\ \lambda \end{pmatrix}+\begin{pmatrix} 0 \\ \partial\varphi_p(p) \\ 0 \end{pmatrix}\ni\begin{pmatrix} f \\ 0 \\ 0 \end{pmatrix}. \qquad (5.12)$$

Матрица $A=\begin{pmatrix} rL^TL & -rL^T \\ 0 & E \end{pmatrix}$ в задаче (5.12) положительно определена при дополнительном условии на параметр r. Именно, справедливо утверждение:

Лемма 5.2. *Пусть $0<r<4$. Тогда матрица A энергетически эквивалентна* $A^0=\begin{pmatrix} L^TL & 0 \\ 0 & E \end{pmatrix}$:

$$m(r)(A^0x,x)\leqslant(Ax,x)\leqslant M(r)(A^0x,x)\quad\forall x=(u,p)^T,$$

где постоянные $0<m(r)<M(r)$ *— это минимальное и максимальное собственные числа матрицы* $\begin{pmatrix} r & -r/2 \\ -r/2 & 1 \end{pmatrix}$.

Преимущество задачи (5.12) по сравнению с (5.11) состоит в том, что матрица A в этой задаче имеет блочно треугольную структуру, при этом включение $p+\partial\varphi_p(p)\ni F$ легко решить.

5.2.2 Конечно-разностные аппроксимации вариационных неравенств с ограничением на градиент решения

Пусть область $\Omega = (0,1)\times(0,1)$, $V = H_0^1(\Omega)$ и $K = \{u \in V : |\nabla u(x)| \leqslant 1 \text{ в } \Omega\}$ – выпуклое и замкнутое множество в пространстве V. Определим на $V \times V$ билинейную форму

$$a(u,v) = \int\limits_{\Omega} \nabla u \cdot \nabla v \, dx.$$

Пусть функция $k(x)$ непрерывна в $\overline{\Omega}$ и $k(x) \geqslant 0$.

Далее будем строить аппроксимации двух вариационных неравенств:

$$u \in K:\ a(u, v-u) \geqslant \int\limits_{\Omega} f(v-u)\, dx \quad \forall v \in K, \tag{5.13}$$

$$u \in V:\ a(u,v-u) + \int\limits_{\Omega} |\nabla v|\, dx - \int\limits_{\Omega} |\nabla u|\, dx \geqslant \int\limits_{\Omega} f(v-u)\, dx \quad \forall v \in V. \tag{5.14}$$

Вариационные неравенства (5.13) и (5.14) однозначно разрешимы.

Построим конечно-разностные аппроксимации вариационных неравенств (5.13) и (5.14) на равномерной сетке

$$\bar{\omega} = \{x = (ih, jh) : 0 \leqslant i,j \leqslant m+1,\ (m+1)h = 1\}.$$

Множество внутренних узлов сетки обозначим через ω, множество граничных узлов – через $\partial\omega$. Далее используем также следующие сеточные множества:

$$\omega_1^+ = \{x = (ih, jh) : 1 \leqslant i \leqslant m+1,\ 1 \leqslant j \leqslant m\},$$

$$\omega_2^+ = \{x = (ih, jh) : 1 \leqslant i \leqslant m,\ 1 \leqslant j \leqslant m+1\},\ \omega^+ = \omega_1^+ \cup \omega_2^+.$$

Пусть V_h – это пространство сеточных функций, определенных в узлах сетки $\bar{\omega}$ и равных нулю в граничных узлах $\partial\omega$, а W_{ih} для $i = 1,2,$ – пространства сеточных функций, определенных в узлах сетки ω_i^+. Примем следующие обозначения для разностных производных по x_1 от $u_h \in V_h$ в узле сетки $x = (ih, jh)$:

$$\overline{\partial}_1 u_h(x) = h^{-1}(u_h(ih,jh) - u_h((i-1)h, jh)),$$

$$\partial_1 u_h(x) = h^{-1}(u_h((i+1)h, jh) - u_h(ih, jh))$$

и аналогично для разностных производных по x_2.

Определим конечно-разностные аппроксимации билинейной формы $a(u,v)$, множества K и функционала $I(u) = \int\limits_\Omega |\nabla u|\,dx$:

$$a_h(u_h, v_h) = h^2 \sum_{x\in\omega^+} \sum_{i=1}^{2} \overline{\partial}_i u_h(x) \overline{\partial}_i v_h(x),$$

$$K_h = \left\{ u_h \in V_h :\ (\overline{\partial}_1 u_h(x))^2 + (\overline{\partial}_2 u_h(x))^2 \leqslant 1 \quad \forall x \in \omega^+ \right\},$$

$$I_h = h^2 \sum_{x\in\omega^+} \sqrt{|\overline{\partial}_1 u_h(x)|^2 + |\overline{\partial}_2 u_h(x)|^2}.$$

Считая $f(x)$ непрерывной функцией, аппроксимируем функционал $\int\limits_\Omega f\,v\,dx$ выражением $f_h(v_h) = h^2 \sum\limits_{x\in\omega} f(x) v_h(x)$.

Конечно-разностные аппроксимации вариационных неравенств (5.13) и (5.14) имеют вид:

$$u_h \in K_h :\ a_h(u_h, v_h - u_h) \geqslant f_h(v_h - u_h)\ \forall v_h \in K_h,$$

$$u_h \in V_h :\ \ a_h(u_h, v_h - u_h) + I_h(v_h) - I_h(u_h) \geqslant f_h(v_h - u_h)\ \forall v_h \in V_h.$$

Поставим во взаимно однозначное соответствие функции $v_h \in V_h$ вектор $v \in \mathbb{R}^n,\ n = m^2$, ее узловых параметров при какой-либо нумерации узлов сетки и будем использовать обозначение $v \Leftrightarrow v_h$. Аналогично, сеточной функции $w_h \in W_{ih}, i = 1, 2$, соответствует вектор $w \in \mathbb{R}^{n+m}$ при выбранной нумерации узлов. Определим прямоугольные $(n+m)\times n$ матрицы L_1, L_2 и вектор $f \in \mathbb{R}^n$ равенствами

$$(L_1 u, w_1) = \sum_{x\in\omega^+} \overline{\partial}_1 u_h(x) w_{1h}(x),\ (L_2 u, w_2) = \sum_{x\in\omega^+} \overline{\partial}_2 u_h(x) w_{2h}(x),$$

$$(f, v) = \sum_{x\in\omega} f(x) v_h(x).$$

Здесь u и v – векторы узловых параметров сеточных функций u_h и v_h из V_h, а w_i – векторы узловых параметров сеточных функций $w_{ih} \in W_{ih}$. Скобки $(,)$ в левой части равенств означают евклидово скалярное произведение в пространстве векторов соответствующей размерности.

Пусть $L^T = \begin{pmatrix} L_1^T & L_2^T \end{pmatrix}$, тогда $L^T L = L_1^T L_1 + L_2^T L_2$ – это матрица сеточного оператора $-\Delta_h = -\overline{\partial}_1 \partial_1 - \overline{\partial}_2 \partial_2$ в пространстве V_h. Она симметрична и положительно определена, ее минимальное собственное число равно $\mu_{\min} = \dfrac{8}{h^2} \sin^2 \dfrac{\pi h}{2}$.

Множество ограничений для векторов векторов узловых параметров $u \Leftrightarrow u_h$ сеточных функций $u_h \in K_h$ – это множество

$$\mathcal{K} = \{u \in \mathbb{R}^n : \ (L_1 u)_j^2 + (L_2 u)_j^2 \leqslant 1 \ \forall j = 1, 2, \ldots, n + m.\}$$

Индикаторная функция множества $\mathcal{K}$ определятся равенством

$$\varphi(Lu) = \sum_{j=1}^{n+m} I_1((L_1 u)_j, ((L_2 u)_j),$$

где $I_1(x_1, x_2)$ – индикаторная функция единичного круга $x_1^2 + x_2^2 \leqslant 1$ в $\mathbb{R}^2$. Функция, соответствующая сеточному функционалу I_h, есть

$$\varphi(Lu) = \sum_{j=1}^{n+m} \sqrt{(L_1 u)_j^2 + (L_2 u)_j^2}.$$

В результате оба сеточных вариационных неравенства могут быть записаны в виде

$$u \in \mathbb{R}^n : \ (L^T L u, v - u) + \varphi(Lv) - \varphi(Lu) \geqslant (f, v - u) \ \ \forall v \in \mathbb{R}^n, \tag{5.15}$$

где функция $\varphi \circ L$ – выпуклая, собственная и полунепрерывная снизу (непрерывная в случае второго вариационного неравенства).

Матрица $L^T L$ положительно определена, поэтому вариационное неравенство (5.15) имеет единственное решение.

Вариационное неравенство (5.15) может быть записано в форме включения

$$L^T L u + L^T \partial\varphi(Lu) \ni f. \tag{5.16}$$

Введем в рассмотрение векторы $p = Lu$ и $\lambda \in p + \partial\varphi(p)$. Тогда тройка (u, p, λ) удовлетворяет системе

$$L^T \lambda = f, \quad \lambda \in p + \partial\varphi(p), \quad p = Lu.$$

Иными словами, это седловая задача

$$\begin{pmatrix} 0 & 0 & L^T \\ 0 & E & -E \\ L & -E & 0 \end{pmatrix} \begin{pmatrix} u \\ p \\ \lambda \end{pmatrix} + \begin{pmatrix} 0 \\ \partial\varphi(p) \\ 0 \end{pmatrix} \ni \begin{pmatrix} f \\ 0 \\ 0 \end{pmatrix}. \tag{5.17}$$

Проведем эквивалентное преобразование седловой задачи (5.17), прибавляя к первому уравнению третье, умноженное на rL^T, $r > 0$. Получим:

$$\begin{pmatrix} rL^TL & -rL^T & L^T \\ 0 & E & -E \\ L & -E & 0 \end{pmatrix} \begin{pmatrix} u \\ p \\ \lambda \end{pmatrix} + \begin{pmatrix} 0 \\ \partial\varphi(p) \\ 0 \end{pmatrix} \ni \begin{pmatrix} f \\ 0 \\ 0 \end{pmatrix}. \tag{5.18}$$

Лемма 5.3. *Пусть* $0 < r < 4$. *Тогда матрица* A *энергетически эквивалентна* $A^0 = \begin{pmatrix} L^TL & 0 \\ 0 & E \end{pmatrix}$:

$$m(r)(A^0x, x) \leqslant (Ax, x) \leqslant M(r)(A^0x, x) \quad \forall x = (u, p)^T,$$

где постоянные $0 < m(r) < M(r)$ *— это минимальное и максимальное собственные числа матрицы* $\begin{pmatrix} r & -r/2 \\ -r/2 & 1 \end{pmatrix}$.

Из леммы 5.3 и из свойств $\partial\varphi$ следует существование решения у седловой задачи (5.18) при $0 < r < 4$.

5.2.3 Сеточные аппроксимации задач оптимального управления в правой части эллиптического уравнения

Задача с распределенным в области наблюдением

Пусть функция $u(x)$ является решением задачи Дирихле для уравнения Пуассона:

$$-\Delta u = p \text{ для } x \in \Omega, \quad u = 0 \text{ для } x \in \partial\Omega,$$

или, в обобщенной постановке,

$$u \in H_0^1(\Omega): \int\limits_\Omega \nabla u \cdot \nabla z \, dx = \int\limits_\Omega p \, z \, dx \;\; \forall z \in H_0^1(\Omega). \tag{5.19}$$

Определим целевой функционал

$$J(u, p) = \frac{1}{2} \int\limits_\Omega (u - u_d)^2 \, dx + \frac{\alpha}{2} \int\limits_\Omega p^2 \, dx, \;\; \alpha = \text{const} > 0,$$

с заданной функцией $u_d(x) \in H_0^1(\Omega)$, а также множества ограничений

$$\begin{array}{ll} \tilde{P}_{ad} = & \{p \in L_2(\Omega) : \, |p(x)| \leqslant p_0 \; \forall x \in \Omega\}, \quad p_0 = const > 0; \\ \tilde{U}_{ad} = & \{u \in H_0^1(\Omega) : \, u(x) \geqslant 0 \; \forall x \in \Omega\}. \end{array} \tag{5.20}$$

Будем решать задачу оптимального управления

$$\min_{(u,p) \in K} J(u,p), \quad K = \{(u,p) \in \tilde{U}_{ad} \times \tilde{P}_{ad} : \text{ выполнено (5.19)}\}. \tag{5.21}$$

Построим конечно-разностную аппроксимацию сформулированной задачи на равномерной сетке $\bar{\omega} = \{x = (ih, jh) : 0 \leqslant i, j \leqslant m+1, \ (m+1)h = 1\}$ в случае $\Omega = (0,1) \times (0,1)$. Множество внутренних узлов сетки обозначим через ω, множество граничных узлов — через $\partial\omega$. Будем считать для простоты, что функции $p(x)$ и $u_d(x)$ непрерывны в $\overline{\Omega}$ и равны нулю на $\partial\Omega$. Обозначим через Δ_h конечно-разностный оператор Лапласа: $\Delta_h = \bar{\partial}_1 \partial_1 + \bar{\partial}_2 \partial_2$. Здесь

$$\Delta_h = \bar{\partial}_1 \partial_1 + \bar{\partial}_2 \partial_2, \text{ где } \bar{\partial}_1 u_i = \frac{u_{ij} - u_{i-1,j}}{h}, \ \partial_1 u_i = \frac{u_{i+1,j} - u_{ij}}{h}$$

и разностные производные $\overline{\partial}_2 u_i, \ \partial_2 u_i$ определены аналогично. Обозначения u_h, p_h будем применять для сеточных функций, заданных в узлах сетки $\bar{\omega}$. Аппроксимируем уравнение состояния (5.19) конечно-разностной схемой:

$$-\Delta_h u_h(x) = p_h(x) \text{ для } x \in \omega; \ u_h(x) = 0 \text{ для } x \in \partial\omega.$$

Далее $u, \, p \in \mathbb{R}^n, \ n = m^2,$ — это векторы узловых параметров сеточных функций $u_h, \, p_h$, определенных в узлах сетки $\bar{\omega}$ и равных нулю в граничных узлах. Конечно-разностная схема может быть записана в виде системы уравнений

$$Lu = p,$$

где матрица $L = L^T > 0$ соответствует конечно-разностному оператору Лапласа $-\Delta_h$ с однородными граничными условиями Дирихле. Для аппроксимации интегралов по Ω от непрерывных функций, равных нулю на границе области, используем простейшие квадратурные формулы

$$\int\limits_\Omega g(x) dx \approx S_\Omega(g) = h^2 \sum_{x \in \omega} g(x).$$

Целевой функционал аппроксимируем функцией $\dfrac{1}{2} S_\Omega((u - u_d)^2) + \dfrac{\alpha}{2} S_\Omega(p^2)$, которая после сокращения на h^2 может быть записана в виде

$$J(u,p) = \frac{1}{2}\|u - u_d\|^2 + \frac{\alpha}{2}\|p\|^2, \quad \text{где } \|.\| - \text{ евклидова норма в } \mathbb{R}^n.$$

Наконец, пусть $\varphi_u = I_{U_{ad}}$ и $\varphi_p = I_{P_{ad}}$ — индикаторные функции множеств ограничений на векторы u и p

$$P_{ad} = \{p \in \mathbb{R}^n : \, |p_i| \leqslant p_0 \ \ \forall i\}, \quad \text{и} \quad U_{ad} = \{u \in \mathbb{R}^n : \, u_i \geqslant 0 \, \forall i\}.$$

Теперь конечномерная задача оптимального управления, аппроксимирующая исходную задачу (5.21), приобретает вид

$$\text{найти} \quad \min_{Lu=p} \left\{ J(u,p) = \frac{1}{2}\|u-u_d\|^2 + \frac{\alpha}{2}\|p\|^2 + \varphi_u(u) + \varphi_p(p) \right\}. \qquad (5.22)$$

Возьмем функцию Лагранжа для задачи (5.22) в виде

$$\mathcal{L}(u,p,\lambda) = \frac{1}{2}\|u-u_d\|^2 + \frac{\alpha}{2}\|p\|^2 + \varphi_u(u) + \varphi_p(p) + (Lu-p,\lambda).$$

В силу теоремы 5.2 ее седловая точка на $\mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n$ удовлетворяет системе

$$\begin{pmatrix} E & 0 & L^T \\ 0 & \alpha E & -E \\ L & -E & 0 \end{pmatrix} \begin{pmatrix} u \\ p \\ \lambda \end{pmatrix} + \begin{pmatrix} \partial\varphi_u(u) \\ \partial\varphi_p(p) \\ 0 \end{pmatrix} \ni \begin{pmatrix} u_d \\ 0 \\ 0 \end{pmatrix}. \qquad (5.23)$$

В очередной раз задача (5.23) — это частный случай (5.1), при этом

$$A = \begin{pmatrix} E & 0 \\ 0 & \alpha E \end{pmatrix}, \; B = \begin{pmatrix} -L & E \end{pmatrix}, \; \text{где } E \in \mathbb{R}^{n\times n} - \; \text{единичная матрица.}$$

Матрица $A \in \mathbb{R}^{2n\times 2n}$ симметрична и положительно определена, $B \in \mathbb{R}^{n\times 2n}$ имеет полный ранг: $\operatorname{rank} B = n$. Так как ранее установлено, что и условие (5.4) выполнено, то из теоремы 5.1 следует существование решения (u,p,λ) системы (5.23) и единственность пары векторов (u,p), которые совпадают с решением (5.22).

Задача с наблюдением в части области

Пусть уравнение состояние и множества ограничений — такие же, как и в предыдущем случае. т. е. (5.19) и (5.20). В качестве целевого функционала возьмем

$$G(u,p) = \frac{1}{2}\int\limits_{\Omega_1} (u-u_d)^2 dx + \frac{\alpha}{2}\int\limits_{\Omega} p^2 dx, \; \alpha = \text{const} > 0,$$

где Ω_1 — подобласть Ω и будем решать задачу оптимального управления

$$\min_{(u,p)\in K} G(u,p), \quad K = \{(u,p) \in \tilde{U}_{ad} \times \tilde{P}_{ad} : \text{ выполнено (5.19)}\}.$$

Построим сеточную аппроксимацию сформулированной задачи в случае, когда $\Omega = (0,1) \times (0,1)$, а подобласть Ω_1 — прямоугольник. Будем считать, что сетка такова, что границы Ω_1 совпадают с линиями сетки. Аппроксимируем целевой функционал функцией

$$\frac{1}{2} S_{\Omega_1}((u-u_d)^2) + \frac{\alpha}{2} S_{\Omega}(p^2), \text{ где } S_{\Omega_1}(g) = h^2 \sum_{x \in \omega_1} g(x),$$

а ω_1 — множество внутренних узлов сетки $\bar{\omega}_1$. Проведя все построения как в предыдущем случае, придем к следующей конечномерной задаче:

$$\min_{Lu=p} \left\{ \frac{1}{2}(M_u(u-u_d), u-u_d) + \frac{\alpha}{2}\|p\|^2 + \varphi_u(u) + \varphi_p(p) \right\}. \tag{5.24}$$

Здесь $\varphi_u = I_{U_{ad}}$ и $\varphi_p = I_{P_{ad}}$ — индикаторные функции множеств ограничений на векторы u и p, а матрица M_u — диагональная и положительно полуопределенная (ее положительные элементы, равные 1, соответствуют узлам сетки $x \in \omega_1$).

Соответствующие задаче минимизации (5.24) функция Лагранжа и седловая задача приобретают вид

$$\mathcal{L}(u,p,\lambda) = \frac{1}{2}(M_u(u-u_d), u-u_d) + \frac{\alpha}{2}\|p\|^2 + \varphi_u(u) + \varphi_p(p) + (Lu-p, \lambda), \tag{5.25}$$

$$\begin{pmatrix} M_u & 0 & L \\ 0 & \alpha E & -E \\ L & -E & 0 \end{pmatrix} \begin{pmatrix} u \\ p \\ \lambda \end{pmatrix} + \begin{pmatrix} \partial\varphi_u(u) \\ \partial\varphi_p(p) \\ 0 \end{pmatrix} \ni \begin{pmatrix} M_u u_d \\ 0 \\ 0 \end{pmatrix}. \tag{5.26}$$

В данном случае матрица $A = \begin{pmatrix} M_u & 0 \\ 0 & \alpha E \end{pmatrix}$ – положительно полуопределенная, поэтому проведем эквивалентное преобразование седловой задачи.

Умножим уравнение $Lu - p = 0$ в системе (5.26) на $r\,L^{-1}$ и прибавим к первому включению системы. Получим следующую задачу:

$$\begin{pmatrix} M_u + rE & -r\,L^{-1} & L \\ 0 & \alpha E & -E \\ L & -E & 0 \end{pmatrix} \begin{pmatrix} u \\ p \\ \lambda \end{pmatrix} + \begin{pmatrix} \partial\varphi_u(u) \\ \partial\varphi_p(p) \\ 0 \end{pmatrix} \ni \begin{pmatrix} M_u u_d \\ 0 \\ 0 \end{pmatrix}. \tag{5.27}$$

Лемма 5.4. *Матрица* $A = \begin{pmatrix} M_u + rE & -r\,L^{-1} \\ 0 & \alpha E \end{pmatrix}$ *положительно определена при* $0 < r < 4\alpha\mu^2$, *и энергетически эквивалентна единичной матрице:*

$$m(r)\|x\|^2 \leqslant (A_2 x, x) \leqslant M(r)\|x\|^2 \quad \forall x = (u,p)^T,\ m_2(r) > 0.$$

Замечание 5.3. *В качестве постоянных спектральной эквивалентности можно взять*

$$M(r) \text{ — максимальное собственное число матрицы } \begin{pmatrix} 1+r & \dfrac{r}{2\,\mu} \\ \dfrac{r}{2\,\mu} & \alpha \end{pmatrix};$$

$$m(r) > 0 \text{ — минимальное собственное число матрицы } \begin{pmatrix} r & -\dfrac{r}{2\,\mu} \\ -\dfrac{r}{2\,\mu} & \alpha \end{pmatrix}.$$

5.3 Итерационные методы для седловых задач

5.3.1 Обобщенный метод Узавы

Будем строить итерационноый метод для задачи

$$\begin{pmatrix} A & -B^T \\ -B & 0 \end{pmatrix} \begin{pmatrix} x \\ \lambda \end{pmatrix} + \begin{pmatrix} \partial\varphi(x) \\ 0 \end{pmatrix} \ni \begin{pmatrix} f \\ -g \end{pmatrix}, \tag{5.28}$$

считая, что матрица $A \in \mathbb{R}^{n\times n}$ положительно определена,

$$B \in \mathbb{R}^{s\times n} - \text{ матрица полного ранга: } \operatorname{rank} B = s \leqslant n, \tag{5.29}$$

$$\varphi \text{ - выпуклая, собственная и полунепрерывная снизу функции.} \tag{5.30}$$

Пусть $D \in \mathbb{R}^{s\times s}$ – симметричная и положительно определенная матрица (предобусловливатель в итерационном методе). Предобусловленный итерационный метод Узавы имеет вид:

$$\begin{aligned} &Ax^{k+1} + \partial\varphi(x^{k+1}) \ni B^T\lambda^k + f, \\ &D\lambda^{k+1} = D\lambda^k + \tau(g - Bx^{k+1}), \end{aligned} \tag{5.31}$$

где $\lambda^0 \in \mathbb{R}^s$ – начальное приближение.

Теорема 5.3. *Пусть выполнены условия* (5.29), (5.30), *матрица A положительно определена и задача* (5.28) *имеет непустое множество решений $X = \{(x,\lambda)\}$. Пусть, также, выполнено одно из следующих условий:*

$$D > \frac{\tau}{2} B A_s^{-1} B^T, \tag{5.32}$$

$$A_s > \frac{\tau}{2} B^T D^{-1} B, \tag{5.33}$$

где $A_s = 0.5(A + A^T)$ – симметричная часть матрицы A. Тогда итерационный метод (5.31) *сходится с любого начального приближения λ^0: $(x^k,\lambda^k) \to (x^*,\lambda^*) \in X$.*

5.3.2 Применение к вариационным неравенствам

Применим итерационный метод (5.31) для решения задачи (5.18)

$$\begin{pmatrix} rL^TL & -rL^T & L^T \\ 0 & E & -E \\ L & -E & 0 \end{pmatrix} \begin{pmatrix} u \\ p \\ \lambda \end{pmatrix} + \begin{pmatrix} 0 \\ \partial\varphi(p) \\ 0 \end{pmatrix} \ni \begin{pmatrix} f \\ 0 \\ 0 \end{pmatrix}.$$

В рассматриваемом случае $B = \begin{pmatrix} L & -E \end{pmatrix}$ и $A = \begin{pmatrix} rLL^T & -rL^T \\ 0 & E \end{pmatrix}$. Поскольку матрица A энергетически эквивалентна блочно диагональной матрице $A^0 = \begin{pmatrix} L^TL & 0 \\ 0 & E \end{pmatrix}$, то можно строить предобусловливатель D в методе Узавы энергетически эквивалентным матрице $B(A^0)^{-1}B^T$. Прямыми вычислениями получим $B(A^0)^{-1}B^T = L(LL^T)^{-1}L^T + E$. Спектр симметричной матрицы $L(LL^T)^{-1}L^T$ состоит из двух точек: 0 и 1. Поэтому $E \leqslant B(A^0)^{-1}B^T \leqslant 2E$ и в качестве D можно взять единичную матрицу. Выберем для определенности $r = 2$. В этом случае метод принимает вид:

$$\begin{aligned} p^{k+1} &= (E + \partial\varphi)^{-1}(\lambda^k); \\ 2L^TLu^{k+1} &= 2L^Tp^{k+1} - L^T\lambda^k + f; \\ \lambda^{k+1} &= \lambda^k + \tau(p^{k+1} - Lu^{k+1}). \end{aligned} \tag{5.34}$$

Теорема 5.4. *Итерационный метод* (5.34) *сходится с любого начального приближения* λ^0 *при* $\tau < 2$.

5.3.3 Применение к задачам оптимального управления

Сеточная схема, аппроксимирующая задачу управления с наблюдением во всей области и распределенным в области управлением

Применим метод Узавы к седловой задаче (5.23):

$$\begin{pmatrix} E & 0 & L^T \\ 0 & \alpha E & -E \\ L & -E & 0 \end{pmatrix} \begin{pmatrix} u \\ p \\ \lambda \end{pmatrix} + \begin{pmatrix} \partial\varphi_u(u) \\ \partial\varphi_p(p) \\ 0 \end{pmatrix} \ni \begin{pmatrix} u_d \\ 0 \\ 0 \end{pmatrix}.$$

Поскольку в рассматриваемом случае $B = \begin{pmatrix} -L & E \end{pmatrix}$ и $A = \begin{pmatrix} E & 0 \\ 0 & \alpha E \end{pmatrix}$. то $BA^{-1}B^T = L^2 + \dfrac{1}{\alpha}E$. Однако для удобства численной реализации метода возьмем в качестве предобусловливателя факторизованную матрицу

$D = L^2$, спектрально эквивалентную матрице $L^2 + \dfrac{1}{\alpha}E$ с постоянными эквивалентности, не зависящими от параметра сетки. Тогда алгоритм реализации итерационного метода для задачи (5.23) состоит из следующих шагов:

для известного λ^k находим

$$\begin{gathered} u^{k+1} = (E + \partial\varphi_u)^{-1}(-L\lambda^k + u_d); \\ p^{k+1} = (\alpha E + \partial\varphi_p)^{-1}(\lambda^k); \\ y^{k+1} = L^{-1}p^{k+1}; \\ \lambda^{k+1} = \lambda^k + \tau L^{-1}(u^{k+1} - y^{k+1}). \end{gathered} \tag{5.35}$$

Теорема 5.5. *Пусть* $\mu_{\min} > 0$ *– это минимальное собственное число матрицы* L*, тогда итерационный метод* (5.35) *сходится при условии*

$$0 < \tau < \frac{2\alpha\mu_{\min}^2}{1 + \alpha\mu_{\min}^2}. \tag{5.36}$$

Сеточная схема, аппроксимирующая задачу управления с наблюдением в части области

Соответствующая седловая задача (5.27) имеет вид

$$\begin{pmatrix} M_u + rE & -r\,L^{-1} & L \\ 0 & \alpha E & -E \\ L & -E & 0 \end{pmatrix} \begin{pmatrix} u \\ p \\ \lambda \end{pmatrix} + \begin{pmatrix} \partial\varphi_u(u) \\ \partial\varphi_p(p) \\ 0 \end{pmatrix} \ni \begin{pmatrix} M_u u_d \\ 0 \\ 0 \end{pmatrix}.$$

Матрица $A = \begin{pmatrix} M_u + rE & -r\,L^{-1} \\ 0 & \alpha E \end{pmatrix}$ положительно определена при $0 < r < 4\alpha\mu^2$, $\mu = \|L^{-1}\|^{-1}$, и энергетически эквивалентна единичной матрице. Как и в предыдущем случае в качестве предобусловливателя D можно взять матрицу L^2. Тогда алгоритм реализации итерационного метода Узавы для задачи (5.27) состоит из следующих шагов:

для известного λ^k последовательно находим

$$\begin{gathered} p^{k+1} = (\alpha E + \partial\varphi_p)^{-1}(\lambda^k); \\ y^{k+1} = L^{-1}p^{k+1}; \\ u^{k+1} = (M_u + rE + \partial\varphi_u)^{-1}(-L^T\lambda^k + ry^{k+1} + M_u u_d); \\ \lambda^{k+1} = \lambda^k + \tau L^{-1}(u^{k+1} - y^{k+1}). \end{gathered} \tag{5.37}$$

Теорема 5.6. *Пусть* $0 < r < 4\alpha\mu^2$*, где* $\mu^2 > 0$ *– минимальное собственное число матрицы* L^2*, и* $m(r) > 0$ *– минимальное собственное число* *матрицы* $\begin{pmatrix} r & -\dfrac{r}{2\,\mu} \\ -\dfrac{r}{2\,\mu} & \alpha \end{pmatrix}$. *Тогда итерационный метод* (5.37) *сходится при условии*

$$0 < \tau < \frac{2m(r)\mu^2}{1+\mu^2}. \tag{5.38}$$