Глава 4. Методы исследования математических моделей

1. Вариационные методы решения краевых задач и определения собственных значений

1. Принцип Дирихле

Рассмотрим функционал:

$$D\big(u(x,y)\big) = \int_D F(x, y, u, u_x, u_y)\,dxdy. \qquad (1)$$

Уравнение Эйлера:

$$F_u - \frac{\partial}{\partial x}\{F_p\} - \frac{\partial}{\partial y}\{F_q\} = 0\,, \quad p = u_x, q = u_y. \qquad (2)$$

Уравнение Лапласа $\Delta u(x, y) = 0$ **есть уравнение Эйлера задачи на минимум интеграла Дирихле**

$$D(u) = \int_D (u_x{}^2 + u_y{}^2)\,dxdy. \qquad (3)$$

Непрерывные в $\bar{D}$ функции, кусочно-непрерывно дифференцируемые в $\bar{D}$, которые принимают на кривой Г изнутри D заданные непрерывные значения $\phi(x, y)$ и интеграл Дирихле от которых конечен, называются допустимыми функциями.

$$\begin{cases} \Delta u = 0, \quad M \in D, & (4) \\ u(p) = \phi(p), \quad p \in \Gamma. & (5) \end{cases}$$

Первая вариационная задача: среди допустимых функций найти такую, которая доставляет минимум интегралу Дирихле.

Теорема.

Если заданная на кривой Г функция $\phi(p)$ такова, что класс допустимых функций не является пустым, то задача Дирихле (4), (5) и первая вариационная задача эквивалентны.

Доказательство:

1) Пусть $u(x,y)\in C^{(2)}(D)\cap C^{(1)}(\bar{D})$ - решение первой вариационной задачи. Класс допустимых функций ищем в виде:

$$u(x,y)+\varepsilon h(x,y), \qquad (6)$$

где $h(x,y)\in C^{(2)}(D)\cap C^{(1)}(\bar{D})$, интеграл (3) от h конечен и $h(p)=0,\ p\in\Gamma,\ \varepsilon$ - произвольная постоянная.

$$D(u+\varepsilon h)=D(u)+2\varepsilon D(u,h)+\varepsilon^2 D(h)\geq 0, \tag{7}$$

где

$$D(u,h)=\int_D (u_x h_x + u_y h_y)dxdy. \tag{8}$$

Функция *u* доставляет минимум интегралу (3), следовательно

$$\frac{d}{d\varepsilon}D(u+\varepsilon h)\Big|_{\varepsilon=0} = 2D(u,h)=0 \Rightarrow D(u,h)=0.$$

Запишем первую формулу Грина для функций *u* и *h*:

$$\int_D h\Delta u dxdy = \int_\Gamma h\frac{\partial u}{\partial n}d\sigma - \int_D grad\, u\, grad\, h\, dxdy. \tag{9}$$

Из (9) и $h(p)=0,\ p\in\Gamma \Rightarrow \int_D h\,\Delta u\,dxdy = -D(u,h)=0,$ (10)

$u(M)\in C^{(2)}(D)\Rightarrow \Delta u\in C(D),\ h(x,y)$ - **произвольная функция** $=>$

$$\Delta u(x,y)=0. \tag{11}$$

2) Пусть теперь $u(x,y)$**- решение задачи (4),(5),**

а $u(x,y)+\varepsilon h(x,y)$ **- класс допустимых функций, причем для** $u(x,y)$ **и** $h(x,y)$ **имеет место формула**

$$D\left(u,h\right)=\int\limits_{\Gamma} h\frac{\partial u}{\partial n}d\sigma-\int\limits_{D} h\,\Delta u\;dxdy. \qquad (12)$$

Из (12) и того, что $h(p)=0,\ p\in\Gamma$ **и** $u(M)$ **гармоническая функция следует, что** $D\left(u,h\right)=0$**. Поэтому из (7) =>**

$$D(u)\le D\left(u+\varepsilon h\right), \qquad (13)$$

то есть функция ***u*** **минимизирует интеграл Дирихле и является решением первой вариационной задачи.**

Замечание

Существуют и другие краевые задачи для уравнения Лапласа, которые имеют эквивалентные им вариационные задачи для интеграла Дирихле, например задача Неймана.

Метод сведения краевых задач для уравнения Лапласа к эквивалентным им вариационным задачам носит название принципа Дирихле.

2. Задача о собственных значениях

$$\begin{cases} \Delta u + \lambda u = 0,\ M \in D, & (14) \\ u(P) = 0,\ p \in \Gamma, & (15) \\ u(M) \in C^{(2)}(D) \cap C^{(1)}(\bar{D}). & (16) \end{cases}$$

Вторая вариационная задача: среди допустимых функций, удовлетворяющих условию (15), найти ту, для которой функционал

$$J(u) = \frac{D(u)}{H(u)}, \tag{17}$$

где

$$H(u) = \int_D u^2(x, y)\,dxdy, \tag{18}$$

принимает наименьшее значение.

Теорема 1.

Если *u(x,y)* - решение второй вариационной задачи, то *u(x,y)* является решением задачи (14) - (16).

Доказательство:

Пусть *u* - решение второй вариационной задачи, причем наименьшее значение *J(u)* удовлетворяет условию:

$$J(u) = \frac{D(u)}{H(u)} = \lambda > 0. \qquad (19)$$

Для класса допустимых функций $u(x, y) + \varepsilon h(x, y)$,

где ε**- произвольная постоянная, *h(x,y)* - произвольная допустимая функция,** $h(p) = 0,\ p \in \Gamma$**, имеем**

$$F(\varepsilon) = \frac{D(u+\varepsilon h)}{H(u+\varepsilon h)} = \frac{D(u) + 2\varepsilon D(u,h) + \varepsilon^2 D(h)}{H(u) + 2\varepsilon H(u,h) + \varepsilon^2 H(h)} \geq \lambda, \qquad (20)$$

где

$$H(u,h) = \int_D u\, h\, dxdy. \qquad (21)$$

Так как *F(ε)* при *ε*=0 имеет минимум, то

$$F'(0) = 2\frac{H(u)D(u,h) - D(u)H(u,h)}{H^2(u)} = 0. \qquad (22)$$

(19) =>

$$D(u) = \lambda H(u). \qquad (23)$$

Подставляем (23) в (22):

$$H(u)D(u,h) - D(u)H(u,h) = H(u)\left\{D(u,h) - \lambda H(u,h)\right\} = 0. \qquad (24)$$

Так как $H(u) \neq 0$ **, то из (24) =>**

$$D(u,h) - \lambda\, H(u,h) = 0. \tag{25}$$

Если $u, h \in C^{(2)}(D) \cap C^{(1)}(\bar{D})$ **и контур Г достаточно гладкий, то**

$$D(u,h) = \int_D (u_x h_x + u_y h_y)\,dxdy = -\int_D \Delta u\ h\ dxdy \tag{26}$$

и из (26) получаем:

$$\begin{aligned} D(u,h) - \lambda\, H(u,h) &= -\int_D \left(\Delta u + \lambda u\right) h dxdy = \\ &= -H\left(\Delta u + \lambda u, h\right) = 0. \end{aligned} \tag{27}$$

Поскольку $u \in C^{(2)}(D) => \Delta u \in C(D)$ **и** h **- произвольная функция, то из (27) следует, что**

$$\Delta u + \lambda u = 0. \tag{28}$$

Теорема 2.

Среди собственных значений задачи (14) – (16) найденное собственное значение является минимальным.

Доказательство:

Пусть $\tilde{\lambda} \neq \lambda$ **некоторое собственное значение и** $\tilde{u}(x, y)$ – **соответствующая собственная функция. Тогда:**

$$H(\Delta\tilde{u} + \tilde{\lambda}\,\tilde{u}, \tilde{u}) = \int_D \Delta\tilde{u} \cdot \tilde{u}dxdy + \tilde{\lambda}\int_D \tilde{u}^2 dxdy =$$

$$= \int_\Gamma \frac{\partial\tilde{u}}{\partial n}\tilde{u}\, d\sigma - \int_D (\tilde{u}_x^2 + \tilde{u}_y^2)dxdy + \tilde{\lambda}H(\tilde{u}) = \qquad (29)$$

$$= -D(\tilde{u}) + \tilde{\lambda}H(\tilde{u}) = 0$$

и из (29) получаем, что

$$\tilde{\lambda} = J(\tilde{u}) = \frac{D(\tilde{u})}{H(\tilde{u})}, \tag{30}$$

а так как

$$\lambda = \min_{u \in \mathrm{P}} J(u), \tag{31}$$

где P – класс допустимых функций, то

$$\lambda < \tilde{\lambda}. \tag{32}$$

2. Некоторые алгоритмы проекционного метода

1. Общая схема алгоритмов

Рассмотрим уравнение:

$$Lu = Au + Bu = f, \quad f \in H, \tag{1}$$

где *A*, *B* – линейные операторы в гильбертовом пространстве *H*; *D(A)*, *D(B)* – области их определения, *D(A)* - плотно в *H*.

Введем оператор *K*: $D(K) \supset D(A)$**.**

Введём систему базисных (координатных) функций:.

$$\left\{\phi_i^{(N)}\right\}, \quad \phi_i^{(N)} \in D(A),\, i = 1, 2, \ldots N;\; N = 1, 2, \ldots,$$

H_N – **линейная оболочка** $\phi_i^{(N)}, i = 1, 2, \ldots N, \left\{\phi_i^{(N)}\right\}$ **- базис в** H_N.

Предположим, что:

1) **при любом** N **функции** $\phi_i^{(N)}, i = 1, 2, \ldots N$ **линейно независимы;**

2) **последовательность подпространств** $\left\{H_N\right\}$ **предельно полна в H:**

для $\forall u \in H \quad \exists\, \tilde{u}_N \in H_N,\; N = 1, 2, \ldots,$ **что**

$$\left\| u - \tilde{u}_N \right\| = \inf_{\omega} \left\| u - \omega \right\| \le \varepsilon(u, N),\, \omega \in H_N, \tag{2}$$

$\varepsilon(u, N) \to 0,\, N \to \infty,\, \varepsilon(u, N)$ **- оценка погрешности аппроксимации.**

Введем также базисные функции: $\left\{\psi_i\right\}, \;\psi_i \in D(K)$.

Ищем приближенное решение (1) в виде

$$u_N = \sum_{i=1}^{N} a_i \phi_i, \qquad (3)$$

где $a_i,\ i = 1, 2, ..., N$ **определяется из системы уравнений:**

$$(Au_N + Bu_N - f, K\psi_j) = 0, \quad j = 1, 2, ..., N, \qquad (4)$$

где (u, v) **- скалярное произведение в *H* и** $\|u\| = (u, u)^{1/2}$.

Разрешимость системы (4) и сходимость u_N **к** u **при** $N \to \infty$ **зависят от свойств операторов *A*, *B* и выбора оператора *K* и систем функций** $\{\phi_i\}, \{\psi_j\}$ **.**

2. Метод Ритца

Пусть дано уравнение: $Au = f, \quad f \in H,$ (5)

где $(Au, \upsilon) = (u, A\upsilon), \quad (Au, u) \geq \gamma^2 \|u\|^2, \quad \gamma > 0, \quad u, \upsilon \in D(A).$

1) **Выбирается базис** $\{\phi_i\}, \ \phi_i \in D(A), \ i = 1, 2, \ldots N;.$

2) **Приближенное решение ищется в виде (3).**

3) **Коэффициенты** $\boldsymbol{a_i}$ **находятся из системы уравнений**

$$(Au_N, \phi_j) = (f, \phi_j), \ j = 1, 2, \ldots, N$$

(6)

$$\hat{A}a = b, \ \ \hat{A} = \{A_{i,j}\}, \ \ A_{i,j} = (A\phi_i, \phi_j),$$

$$a = (a_1, \ldots, a_N)^T, \ \ b = (f_1, \ldots, f_N)^T, \ \ f_i = (f, \phi_i), \quad (7)$$

$$i, j = 1, 2, \ldots, N.$$

Вариационная трактовка метода Ритца

Рассмотрим квадратичный функционал энергии

$$F(u) = (Au, u) - 2(u, f). \qquad (8)$$

Теорема. **Для того, чтобы некоторый элемент** $u_0 \in D(A)$ **сообщал минимальное значение функционалу энергии** ***F(u)*****, необходимо и достаточно , чтобы этот элемент удовлетворял уравнению (5). Такой элемент единственный.**

Вариационная задача: найти функцию u_N такую, что

$$F(u_N) = \min_{\upsilon} F(\upsilon), \quad u_N, \upsilon \in H_N. \tag{9}$$

Так как $\upsilon = \sum_{i=1}^{N} a_i \phi_i$ **, то** $F(u_N) = \min_{a_i} F(\upsilon)$ **,**

где $F(\upsilon) = F(a_1, ..., a_N) = \sum_{i,j=1}^{N} a_i a_j (A\phi_i, \phi_j) - 2\sum_{i=1}^{N} a_i (f, \phi_i) \quad \Rightarrow$

$$\frac{\partial F(\upsilon)}{\partial a_i} = 0, i = 1, 2, ..., N. \tag{10}$$

Из (10) следует (7) и (6).

Теорема.

Если для любой функции $u \in D(A)$ **можно построить такую последовательность элементов** $\tilde{u}_N = \sum_{i=1}^{N} \tilde{a}_i \phi_i \in H_N$, $N = 1, 2, \dots$,**что** $\left\| A(u - \tilde{u}_N) \right\| \to 0$ **при** $N \to \infty$ **, то приближенные решения** $\boldsymbol{u_N}$ **сходятся к точному решению** $\boldsymbol{u_0}$ **уравнения (5) при** $N \to \infty$ **и имеет место оценка**

$$\left\| u_0 - u_N \right\| \le C \min_{\tilde{a}_i} \left\| A(u_0 - \tilde{u}_N) \right\|,$$

где $C > 0$ **не зависит от** $\boldsymbol{u_0}$ **и** $\tilde{u}_N$ **.**

Доказательство:

Пусть $\upsilon \in D(A)$ **- произвольная функция.**

$$(A(u_0-\upsilon),u_0-\upsilon)=(Au_0,u_0)+(A\upsilon,\upsilon)-2(Au_0,\upsilon)=$$

$$=(Au_0,u_0)+(A\upsilon,\upsilon)-2(f,\upsilon)=F(\upsilon)+(Au_0,u_0)-$$

$$-F(u_0)+F(u_0)=F(\upsilon)-F(u_0)+2(Au_0,u_0)-2(u_0,f)=$$

$$=F(\upsilon)-F(u_0)\ \Rightarrow\ (A(u_0-\upsilon),u_0-\upsilon)=F(\upsilon)-F(u_0)$$

Поскольку $\boldsymbol{u_0}$ **минимизирует функционал** $\boldsymbol{F(\upsilon)}$ **на** $\boldsymbol{D(A)}$**, а** $\boldsymbol{u_N}$ **минимизирует** $\boldsymbol{F(\upsilon)}$ **на** $\boldsymbol{H_N}$ **, то**

$$(A(u_0-u_N),u_0-u_N)=F(u_N)-F(u_0)\le$$

$$\le F(\upsilon_N)-F(u_0)=(A(u_0-\upsilon_N),u_0-\upsilon_N)$$

при произвольной функции $\upsilon_N=\sum_{i=1}^{N}C_i\phi_i$ **из** $H_N \Rightarrow$

$$\gamma^2 \left\| u_0 - u_N \right\|^2 \le (A(u_0 - u_N), u_0 - u_N) \le$$

$$\le (A(u_0 - \upsilon_N), u_0 - \upsilon_N) \le \left\| A^{-1} \right\| \cdot \left\| A(u_0 - \upsilon_N) \right\|^2 \Rightarrow$$

$$\left\| u_0 - u_N \right\|^2 \le \frac{\left\| A^{-1} \right\|}{\gamma^2} \left\| A(u_0 - \upsilon_N) \right\|^2$$

В силу произвольности выбора коэффициентов C_i, i=1,2,…,N в разложении υ_N, положив $\upsilon_N = \tilde{u}_N$ получаем утверждение теоремы.

Замечание. **При рассмотрении классической формулировки метода Ритца решение вариационной задачи может не существовать.**

Введем в *D(A)* энергетическое скалярное произведение и норму

$$[u,\ \upsilon] = (Au,\ \upsilon),\ \ [u] = [u,u]^{1/2} \tag{11}$$

и пополним *D(A)* по энергетической норме.

Получим энергетическое пространство H_A, порождаемое оператором *A*. В H_A могут появиться предельные элементы не принадлежащие *D(A)*.

Расширим функционал энергии на H_A:

$$F(u) = [u,u] - 2(f,u) \tag{12}$$

и будем искать его минимум на H_A:

Пусть минимум достигается на $u_0 \in H_A$ **. Если** $u_0 \notin D(A)$ **, то** $\boldsymbol{u_0}$ **- обобщенное решение (5), если** $u_0 \in D(A)$ **, то** $\boldsymbol{u_0}$ **- классическое решение (5).**

Краевые условия которым удовлетворяют элементы из *D(A)*, называются естественными для оператора *A*, а краевые условия, которым удовлетворяют как элементы из *D(A)*, так и элементы из H_A, называются главными. Базисные функции выбираются из H_A.

3. Метод Галёркина

Основной недостаток метода Ритца: он применим только для самосопряженных положительно определенных операторов.

Метод Бубнова - Галёркина

$$Lu = Au + Bu = f,\ f \in H, \tag{13}$$

1) **Выбирается базис** $\{\phi_i\},\ \phi_i \in D(A),\ i = 1, 2, \ldots N$ **;**
2) **Приближенное решение ищется в виде (3);**
3) **Коэффициенты a_i находятся из условия ортогональности невязки** $Lu_N - f$ **к** $\phi_1, \phi_2, \ldots, \phi_N$ **:**

$$(Lu_N - f, \phi_i) = 0,\ i = 1, 2, \ldots, N \tag{14}$$

Замечание. Если коэффициенты a_i определяются из условия

$$(Lu_N - f, \psi_i) = 0,\ i = 1, 2, \ldots, N \tag{15}$$

где $\{\psi_i\} \subset H$, **некоторый базис , то метод называется методом Галёркина - Петрова.**

Если $(Au,\upsilon)=(u,A\upsilon),\ (Au,u)\ge\gamma^2\|u\|^2,\ \ \gamma>0,\ \ u,\upsilon\in D(A)$ **, то можно ввести энергетическое пространство *H_A*. Тогда (14) =>**

$$[u,\upsilon]+(Bu,\upsilon)=(f,\upsilon),\quad \upsilon\in H_A. \tag{16}$$

1) Выбираем базис $\{\phi_i\},\ \ \phi_i\in D(A),\ \ i=1,2,\ldots N$ **;**

2) Приближенное решение ищется в виде (3);

3) Коэффициенты *a_i* определяются из системы уравнений:

$$[u_N,\phi_i]+(Bu_N,\phi_i)=(f,\phi_i)\qquad i=1,2,\ldots,N \tag{17}$$

$$\hat{L}a=b,\ \ \hat{L}=\left(L_{i,j}\right),\ \ L_{i,j}=\left[\phi_i,\phi_j\right]+(B\phi_i,\phi_j), \tag{18}$$

$$a=(a_1,\ldots,a_N)^T,\ \ b=(f_1,\ldots,f_N)^T,\ \ f_i=(f,\phi_i),\qquad i,j=1,2,\ldots,N$$

4. Обобщенный метод моментов

$$Au + Bu = f, \quad f \in H,$$

$$(Au, Ku) \geq \gamma^2 \|u\|^2, \ (Au, Ku) \geq \beta^2 \|Ku\|^2; \ \gamma, \beta > 0,$$

$u \in D(A) \subset D(K)$ **- оператор *A* является *К*–положительно определенным.**

1) **Выбирается базис** $\{\phi_i\}, \ \phi_i \in D(A), \ i = 1, 2, \ldots N$**;**

2) **Приближенное решение ищется в виде (3);**

3) **Коэффициенты** a_i **определяются из системы уравнений:**

$$(Au_N + Bu_N - f, K\phi_i) = 0, \qquad i = 1, 2, ..., N. \tag{19}$$

Замечание. **Метод моментов широко используется для решения интегральных уравнений:**

$$u(x) - \lambda \int_a^b K(x,\xi)u(\xi)d\xi = f(x), \quad a \le x \le b \qquad (20)$$

Ищем приближенное решение в виде разложения по полной системе функций $\{\phi_i(x)\}$ **:**

$$u_N(x) = f(x) + \lambda \sum_{i=1}^{N} a_i \phi_i(x). \qquad (21)$$

Коэффициенты a_i **определяются из ортогональности невязки ко всем функциям** $\{\phi_i(x)\}$ **:**

$$\int_a^b \left\{ u_N(x) - \lambda \int_a^b K(x,\xi)u_N(\xi)d\xi - f(x) \right\} \phi_j(x)dx = 0, \quad j = 1, 2, ..., N, \qquad (22)$$

что приводит к системе:

$$\hat{A}a = b, \quad \hat{A} = \left(A_{i,j}\right), \quad i, j = 1, 2, ..., N,$$

$$A_{i,j} = \int_a^b \phi_i(x)\phi_j(x)dx - \int_a^b\int_a^b K(x,\xi)\phi_i(\xi)\phi_j(x)dxd\xi, \quad (23)$$

$$a = (a_1, ..., a_N)^T, \quad b = (b_1, ..., b_N)^T,$$

$$b_i = \int_a^b\int_a^b K(x,\xi)\phi_i(x) f(\xi)dxd\xi, \quad i = 1, 2, ..., N.$$

Замечание. Если система $\left\{\phi_i(x)\right\}$ **ортогональная, то метод моментов эквивалентен замене ядра на специальное вырожденное ядро:**

$$\tilde{K}(x,\xi) = \sum_{i=1}^{N} \phi_i(x)\Phi_i(\xi); \quad \Phi_i(\xi) = \int_a^b K(x,\xi)\phi_i(x)dx \quad (24)$$

5. Метод наименьших квадратов

Пусть оператор в (5) имеет ограниченный обратный A^{-1}.

1) Выбирается базис $\{\phi_i\}$, $\phi_i \in D(A)$, $i = 1, 2, \ldots N$;

2) Приближенное решение ищется в виде (3);

3) Коэффициенты a_i определяются из системы уравнений:

$$(Au_N, A\phi_i) = (f, A\phi_i), \quad i = 1, 2, \ldots, N, \tag{25}$$

$$\hat{A}a = b, \ \hat{A} = \left(A_{i,j}\right), \ A_{i,j} = (A\phi_i, A\phi_j), \ \hat{A} = \hat{A}^T,$$

$$i, j = 1, 2, \ldots, N, \ a = (a_1, \ldots, a_N)^T, \ b = (f_1, \ldots, f_N)^T,$$

$$f_i = (f, A\phi_i), \quad i = 1, 2, \ldots, N.$$

Замечание. Соотношения (25) можно получить из условия минимизации функционала невязки $J(u) = \|Au - f\|^2$ на H_N.