17.1. Вариационный подход Ритца

Рассмотрим две задачи:

$$\hat{L}_1\, u(x) \equiv -(k(x)\, u'_x(x))'_x + p(x)\, u(x) = f(x), \tag{17.1}$$

$$u(0) = a; \quad u(X) = b;$$

$$k(x) \geqslant k_0 > 0; \quad p(x) \geqslant 0.$$

$$\hat{L}_2\, u(x,\ y) \equiv -\mathrm{div}\ (k(x,\ y)\ \mathrm{grad}\, u(x,\ y)) + p(x,\ y)\, u(x,\ y) = g(x,\ y), \tag{17.2}$$

$$u\ |_{\partial\Omega}\ = \psi(l);$$

$$k(x,\ y) \geqslant k_0 > 0; p(x,\ y) \geqslant 0.$$

Эти задачи похожи: (17.1) является одномерным случаем более общей задачи (17.2). Уравнения (17.1) и (17.2) записаны в самосопряженной форме. Поставим задачам (17.1) и (17.2) в соответствие функционалы

$$I_1(u) = \int\limits_0^X (k(u'_x)^2 + pu^2 - 2fu)\, dx \tag{17.3}$$

и

$$I_2(u) = \iint\limits_\Omega (k(\nabla u, \nabla u) + pu^2 - 2gu) dx dy. \tag{17.4}$$

Будем рассматривать пространство функций $w \in W_2^1$ (пространство Соболева) с нормой

$$||w||^2_{W_2^1} = \int\limits_0^X (w^2 + (w'_x)^2)\,dx \quad \text{для одномерного случая,}$$

$$||w||^2_{W_2^1} = \iint\limits_\Omega w^2 dxdy + \iint\limits_\Omega \left(\frac{\partial w}{\partial x}\right)^2 + \left(\frac{\partial w}{\partial y}\right)^2 dxdy \quad \text{для двумерного случая.}$$

Это — функции с ограниченным интегралом.

Теорема 5. *Среди всех функций $w \in W_2^1$, удовлетворяющих граничным условиям, решение задачи* (17.1) *придает наименьшее значение функционалу* (17.3)*, а решение* (17.2) *— функционалу* (17.4).

Доказательство.

Докажем это утверждение для одномерного случая, а доказательство для уравнений (17.2, 17.4) оставим в качестве упражнений.

Введем $\xi(x) \equiv w(x) - u(x)$. Поскольку $w(x) \in W_2^1$, а $u(x)$ — дважды непрерывно дифференцируемая функция, то $\xi(x) \in W_2^1$ и $\xi(0) = \xi(X) = 0$.

$$I_1(w) = I_1(u(x) + \xi(x)) =$$

$$= I_1(u) + \int\limits_0^X (k(\xi'_x)^2 + p\xi^2 - 2f\xi)\,dx + \int\limits_0^X 2(ku'_x\xi'_x + p\xi u)\,dx =$$

$$= I_1(u) + \int\limits_0^X (k(\xi'_x)^2 + p\xi^2)\,dx + \int\limits_0^X 2\xi(pu - f)\,dx + 2\int\limits_0^X ku'_x\xi'_x\,dx =$$

$$= I_1(u) + J(\xi) + 2ku'_x\xi \Big|_0^X + \int\limits_0^X 2\xi(-(ku'_x)'_x + pu - f)\,dx,$$

$$\text{где} J(\xi) \equiv \int\limits_0^X (k(\xi'_x)^2 + p\xi^2)\,dx \geqslant 0. \tag{17.5}$$

Третье слагаемое в (17.5) равно нулю в силу граничных условий для функции ξ; последнее слагаемое равно нулю, так как u — решение (17.1); второе слагаемое — неотрицательное. Следовательно, минимум функционала $I_1(w)$ достигается, когда $J(\xi) = 0$, т. е. $\xi \equiv 0$ или, что то же самое, $w(x) = u(x)$. ■

Чуть сложнее эта теорема доказывается для двумерного случая, где надо воспользоваться теоремой Остроградского–Гаусса. Таким образом, решение соответствующей задачи в частных производных (17.2) или краевой задачи для ОДУ (17.1) сводится к задаче минимизации некоторого функционала.

В том случае, если функционал (17.3) или (17.4) ограничен снизу, то экстремаль функционала — минимум, и численный метод, который будет построен ниже, носит название метода Ритца. Чаще, когда нет необходимости тщательно исследовать постановку задачи, говорят об экстремальной точке, стационарной точке функционала и т.д.

17.2. Общая схема метода Ритца

Решение задачи (17.1) ищут в виде

$$u^N(x) = \psi_0^N(x) + \sum_{k=1}^{N} C_k \psi_k^N(x), \tag{17.6}$$

где $\psi_0^N(x), \ldots, \psi_N^N(x)$ — базисные функции в W_2^1; $\psi_0^N(x)$ удовлетворяет граничным условиям, а $\psi_k^N(x)$ при $k \geqslant 1$ такие, что $\psi_k^N(0) = \psi_k^N(X) = 0$. Если суммирование в (17.6) происходит до бесконечности, то эта формула дает точное решение задачи (17.1). Так как рассматривается конечное число базисных функций, то получаем лишь приближенное решение. Примером базисных функций для метода Ритца может служить тригонометрический базис, а в качестве приближенного решения получим конечный отрезок ряда Фурье.

Подставив (17.6) в (17.3), получаем

$$I_1(u^N) = \sum_{p=1}^{N} \sum_{q=1}^{N} C_p C_q \cdot \left\{ \int\limits_0^X \left(k(x) \, \frac{\partial \, \psi_p^N}{\partial \, x} \, \frac{\partial \, \psi_q^N}{\partial \, x} + p(x) \psi_p^N \psi_q^N \right) dx \right\} -$$

$$-2 \sum_{r=1}^{N} \left\{ C_r \int\limits_0^X \left(f(x) \psi_r^N - p(x) \psi_0^N \psi_r^N - k(x) \frac{\partial \psi_0^N}{\partial x} \frac{\partial \psi_r^N}{\partial x} \right) dx \right\} + I_1(\psi_0^N). \tag{17.7}$$

Находим минимум функционала (17.7) из условия $\frac{\partial \, I_1(u^N)}{\partial \, C_i} = 0$, получаем систему из N линейных уравнений для определения коэффициентов C_k. Затем объявляем (17.6) решением задачи.

Точно так же поступаем и для функционала (17.4). Число уравнений в системе для определения коэффициентов тоже будет N (у базисной функции только один индекс!). Вид функционала будет аналогичен (17.7), но вместо интегралов по отрезку будут стоять двойные интегралы по рассматриваемой области пространства Ω, а вместо производных — градиенты.

Первая проблема, которая возникает в методе Ритца — выбор подходящего базиса. Как от набора функций $\psi_0^N(x), \ldots, \psi_N^N(x)$ зависит решение? Как оценить ошибки?

Существуют два типа базиса: *глобальный базис* для метода Ритца и *базис из функций с финитным носителем*.

Для того чтобы решения по методу Ритца сходились к точному, необходимо и достаточно, чтобы $\forall\ g \in W_2^1$ и $\forall\ \varepsilon > 0$ существовала линейная комбинация

$$g^N(x) \equiv \psi_0^N(x) + \sum_{j=1}^{N} C_j \psi_j^N(x), \text{такая, что} ||g^N - g||_{W_2^1} \leqslant \varepsilon,$$

если вычисления проводятся точно.

Допустимый базис для применения в методе Ритца $\sin\left(\frac{\pi q x}{X}\right)$, $q = 1, \ldots, N$.

На отрезке $[0, 1]$ допустимые базисы: $\psi_j^N = x(1 - x)T_j(2x - 1)$, где $T_j(x)$ — j-й полином Чебышева; $\psi_j^N = x^j(1 - x)$.

Матрица системы линейных уравнений для определения коэффициентов разложения по базису метода Ритца получается заполненной. В случае использования «неудачных» базисов ее число обусловленности достаточно велико.

Технологичность метода Ритца заключается в следующем. Матрица соответствующей системы является самосопряженной с диагональным преобладанием при правильном выборе базиса. Можно решать систему быстро сходящимися итерационными методами.

Рассмотрим простейший вариант метода Ритца с использованием базиса функций с финитным носителем. Напомним, что носитель функции — множество точек x, для которых $\psi_j^N(x) \neq 0$. Введем разбиение отрезка $[0,\ \ X]$ точками x_j (сетку): $0 = x_0 < x_1 < \ldots < x_N = X$. Строим базисные функции:

$$\psi_0^N = \begin{cases} a \cdot \frac{x - x_1}{x_0 - x_1}, & 0 \leqslant x \leqslant x_1; \\ 0, & x_1 \leqslant x \leqslant x_{N-1}; \\ b \cdot \frac{x - x_{N-1}}{x_N - x_{N-1}}, & x_{N-1} \leqslant x \leqslant x_N; \end{cases}$$

$$\psi_j^N = \begin{cases} \frac{x - x_{j-1}}{x_j - x_{j-1}}, & x_{j-1} \leqslant x \leqslant x_j; \\ 0, & x > x_{j+1}, \quad x < x_{j-1}; \\ \frac{x_{j+1} - x}{x_{j+1} - x_j}, & x_j < x \leqslant x_{j+1}. \end{cases}$$

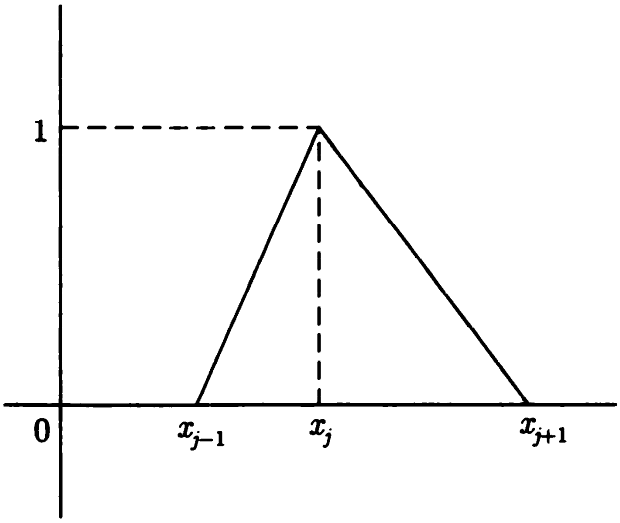

Рис. 17.2

Можно проверить, что $\psi_j^N \in W_2^1[0, \quad X]$. Интегралы и производные определяются в смысле обобщенных функций — недостаток базиса! Достоинством этого базиса является то, что базисные функции почти ортогональны.

Пусть

$$(\psi_j^N, \psi_k^N) = \int\limits_0^X \psi_j^N(x)\, \psi_k^N(x)\, dx,$$

тогда

$$(\psi_j^N, \psi_j^N) = \int\limits_{x_{j-1}}^{x_j} \frac{(x - x_{j-1})^2}{(x_j - x_{j-1})^2}\, dx + \int\limits_{x_j}^{x_{j+1}} \frac{(x - x_{j+1})^2}{(x_{j+1} - x_j)^2}\, dx =$$

$$= (x_j - x_{j-1}) \int\limits_0^1 t^2 dt + (x_{j+1} - x_j) \int\limits_0^1 t^2 dt = \frac{(x_j - x_{j-1})}{3}\, t^3 \Big|_0^1 + \ldots =$$

$$= \frac{x_j - x_{j-1}}{3} + \frac{x_{j+1} - x_j}{3} = \frac{x_{j+1} - x_{j-1}}{3}, (\psi_j^N, \psi_{j+1}^N) =$$

$$= \int\limits_{x_j}^{x_{j+1}} \frac{x - x_j}{x_{j+1} - x_j}\, \frac{x_{j+1} - x}{x_{j+1} - x_j}\, dx = \frac{x_{j+1} - x_j}{6}, (\psi_{j-1}^N, \psi_j^N) = \frac{x_j - x_{j-1}}{6},$$

а все остальные скалярные произведения равны нулю.

Также достаточно легко берутся интегралы, включающие в себя производные базисных функций. Носитель каждой такой базисной функции называется конечным элементом, а метод Ритца с использованием такого базиса — первый метод из семейства МКЭ. Иногда конечным элементом также называют саму базисную функцию с финитным носителем.

17.3. Формулировка проекционного метода Галеркина

По-прежнему рассматриваем задачи (17.1) и (17.2).

В дальнейшем будет рассмотрен класс дифференциальных операторов. Главный недостаток метода Ритца — применимость лишь к дифференциальным задачам, допускающим вариационную формулировку, т. е. в линейном случае $\hat{L}$ — самосопряженный положительно определенный оператор (все собственные числа $\hat{L}$ положительны).

Наряду с формулировкой (17.1) и (17.2) будем использовать запись, определяющую *слабое* (обобщенное) решение:

$$(\hat{L}u,\ v) - (f,\ v) = 0, \tag{17.8}$$

где v — *любая* функция из рассмотренного ранее функционального пространства W_2^1, а скалярное произведение определено как

$$(u,\ v) = \int_0^X u(x)\ v(x)\ dx \quad \text{в одномерном случае;}$$

$$(u, v) = \iint_\Omega u(x,y)v(x,y)dxdy \quad \text{в двумерном случае.}$$

Равенство (17.8) определяет обобщенное решение задачи. Известно, что если u — классическое решение задачи, то оно является обобщенным решением в смысле (17.8). Обратное, по понятным причинам, неверно — в W_2^1 «больше» функций, чем в C^1 или C^2. У задачи может существовать обобщенное решение, но не существовать классического.

Рассмотрим конечномерное подпространство пространства W_2^1 с введенным базисом:

$$u^N = \psi_0^N + \sum_{k=1}^N C_k \psi_k^N,$$

ψ_k^N — базисные функции в W_2^1; они обязаны обладать теми же свойствами, что и базисные функции для метода Ритца. Рассмотрим теперь для (17.8) *конечную* систему *весовых* функций из W_2^1 : $v_1^N, \ldots, v_N^N$. Вместо (17.8) рассмотрим конечную систему проекций на весовые функции.