Тема : Метод Ритца. Метод Галеркина. Метод конечных элементов

$\mathbf{1^0}$. Вариационная постановка краевой задачи для обыкновенного дифференциального уравнения на отрезке. $\mathbf{2^0}$. Формулировка метода Ритца. $\mathbf{3^0}$. Расчетные формулы метода Ритца. $\mathbf{4^0}$. Проекционная постановка краевой задачи. Сопутствующие интегральные тождества. $\mathbf{5^0}$. Метод Галёркина. Специфика матриц сопутствующих систем линейных уравнений. $\mathbf{6^0}$. Метод конечных элементов как разновидность проекционных методов. Реализация в одномерном случае.

$1^0 - 2^0$. На конечном отрезке числовой оси задано дифференциальное уравнение второго порядка

$$-(p(x)u')' + q(x)u = f(x), \qquad a < x < b. \qquad (DE_s)$$

Функции $p(x)$, $q(x)$ и $f(x)$ здесь известны, а функция $u = u(x)$ — искомая, ее требуется найти. При этом выполнены неравенства

$$p(x) \geq p_0 > 0, \quad q(x) \geq 0 \quad \text{для} \quad a \leq x \leq b.$$

Предполагается также, что $p(x)$, $q(x)$ и $f(x)$ достаточно гладкие для выполнения с ними всех требуемых аналитических операций. К уравнению (DE_s) добавляются краевые условия первого рода на краях отрезка $[a, b]$.

Для приближенного решения краевой задачи методом Ритца совместно с ней рассматривается **вариационная** задача на отыскание минимума сопутствующего дифференциальному уравнению функционала

$$J(w) = \frac{1}{2}\int\limits_a^b \Big[p(x)(w')^2 + q(x)w^2\Big]dx - \int\limits_a^b f(x)w(x)dx.$$

Точнее, среди всевозможных функций $w = w(x)$ из класса $C^{(1)}[a,b]$, удовлетворяющих краевым условиям

$$w(a) = u_a, \quad w(b) = u_b,$$

требуется найти ту, которая доставляет минимум функционалу $J(w)$ на указанном классе:

$$w_0 = \arg\min_{w} J(w). \qquad (VP)$$

Решение $w_0 = w_0(x)$ этой экстремальной задачи существует и единственно среди функций сопутствующего линейного пространства

$$\mathbb{U} = \left\{ w(x) \in C^{(1)}[a, b] \mid w(a) = u_a,\ w(b) = u_b \right\}.$$

Решение w_0, как доказывается в вариационном исчислении, совпадает с единственным же решением $u(x)$ дифференциальной краевой задачи, то есть

$$w_0(x) = u(x) \quad \text{при} \quad a \leq x \leq b.$$

Это существенное математическое наблюдение позволяет искать функцию $u(x)$ как решение вариационной задачи (VP).

Преимущество вариационной задачи $(\boldsymbol{VP})$ перед дифференциальной постановкой состоит в том, что изначально не требуется, чтобы искомая функция $u(x)$ имела вторые непрерывные производные. Более того, даже первые производные функции $u(x)$ могут быть лишь кусочно-непрерывными, а значение минимизируемого функционала $J(w)$ при этом все равно будет определено.

В методе Ритца вместо минимизации функционала $J(w)$ на всем пространстве $\mathbb{U}$ его минимум ищется на конечномерном подпространстве $\mathbb{U}^N$. Точнее, находится такая функция $u_N(x)$, для которой выполнены условия

$$\begin{cases} J(u_N) = \min\limits_{v \in \mathbb{U}^N} J(v), \\ \text{где} \quad u_N(x) = \sum\limits_{j=0}^{N} \alpha_j \varphi_j(x). \end{cases} \qquad (VP_N)$$

Здесь линейно независимые функции

$$\varphi_0, \quad \varphi_1, \quad \dots, \quad \varphi_N, \quad \varphi_j \in \mathbb{U}, \qquad (B)$$

образуют базис конечномерного подпространства $\mathbb{U}^N$, представляющего собой их линейную оболочку:

$$\mathbb{U}^N = \mathbf{span}\,\{\varphi_0, \varphi_1, \varphi_2, \dots, \varphi_N\} \subset \mathbb{U}.$$

С целью решить задачу (VP_N), предъявим более детальные требования к исходному базису ($\boldsymbol{B}$). Точнее предположим, что выполняются следующие условия:

$$\begin{cases} \varphi_0(a) = 1, \\ \varphi_j(a) = 0 \qquad \text{для} \qquad j = 1, 2 \ldots, N, \\ \varphi_N(b) = 1, \\ \varphi_j(b) = 0 \qquad \text{для} \qquad j = 0, 1, 2, \ldots, N - 1. \end{cases}$$

В этом случае сумма $\sum\limits_{j=0}^{N} \alpha_j \varphi_j(x)$ удовлетворяет краевым условиям дифференциальной постановки в том и только том случае, если $\alpha_0 = u_a$ и $\alpha_N = u_b$. В этом случае имеем разложение

$$u_N(x) = u_a\varphi_0(x) + \sum_{j=1}^{N-1} \alpha_j\varphi_j(x) + u_b\varphi_N(x). \quad (E_N)$$

Определение. *Функция $u_N(x)$, полученная по формуле (E_N) как решение вариационной задачи (VP_N), называется* ***приближением по Ритцу*** *к $u(x)$.*

3^0. Получим расчетные формулы, пригодные к отысканию приближений по Ритцу для искомого решения $u = u(x)$, исходной краевой задачи.

Пусть функция $v = v(x)$ имеет вид

$$v = u_a\varphi_0(x) + \sum_{j=1}^{N-1} \alpha_j\varphi_j(x) + u_b\varphi_N(x).$$

Тогда $J(v) = J\Big(\sum\limits_{j=0}^{N} \alpha_j \varphi_j\Big)$ представляет собой функцию переменных $(\alpha_1, \alpha_2, \ldots, \alpha_{N-1})$ из пространства $\mathbb{R}^{N-1}$.

Необходимым условием экстремума этой функции многих переменных на пространстве $\mathbb{R}^{N-1}$ является обращение в нуль ее производных по указанным переменным:

$$\frac{\partial}{\partial \alpha_i} J\Big(\sum_{j=0}^{N} \alpha_j \varphi_j\Big) = 0, \qquad i = 1, 2, \ldots, N-1.$$

Запишем указанные равенства в виде линейных уравнений, которым должны удовлетворять искомые переменные $(\alpha_1, \alpha_2, \ldots, \alpha_{N-1})$.

Учтем, что $\alpha_0 = u_a$, $\alpha_N = u_b$, а значение функционала $J\big(\sum\limits_{j=0}^{N} \alpha_j \varphi_j\big)$ задается следующим аналитическим выражением:

$$\frac{1}{2}\int\limits_a^b \Big[p(x)\Big(\sum_{j=0}^{N}\alpha_j\varphi_j'\Big)^2 + q(x)\Big(\sum_{j=0}^{N}\alpha_j\varphi_j\Big)^2\Big]dx -$$

$$-\int\limits_a^b f(x)\sum_{j=0}^{N}\alpha_j\varphi_j(x)dx.$$

Дифференцируя это выражение по переменной α_i, получаем значение производной

$$\frac{\partial}{\partial\alpha_i}J\Big(\sum_{j=0}^{N}\alpha_j\varphi_j\Big)$$

в виде суммы интегралов

$$\int\limits_a^b\Big[p(x)\Big(\sum_{j=0}^{N}\alpha_j\varphi_j'\Big)\varphi_i'+q(x)\Big(\sum_{j=0}^{N}\alpha_j\varphi_j\Big)\varphi_i\Big]dx-\int\limits_a^b f(x)\varphi_i(x)dx.$$

Полагая здесь $i = 1, 2, \ldots, N-1$, приходим к искомой системе линейных алгебраических уравнений

$$\sum_{j=0}^{N}\Big[\int_a^b \big(p(x)\varphi_j'\varphi_i' + q(x)\varphi_j\varphi_i\big)dx\Big]\alpha_j = \int_a^b f(x)\varphi_i(x)dx.$$

Неизвестными здесь выступают коэффициенты α_j, $j = 1, 2, \ldots, N-1$, искомой линейной комбинации. При этом $\alpha_0 = u_a$ и $\alpha_N = u_b$ известны из краевых условий.

Введем вспомогательные обозначения

$$\begin{cases} a_{ij} = \int\limits_a^b \Big[p(x)\varphi'_i\varphi'_j + q(x)\varphi_i\varphi_j\Big]dx, \\ \\ b_i = \int\limits_a^b f(x)\varphi_i(x)dx, \end{cases}$$

где $i = 1, 2, \ldots, N-1$ и $j = 1, 2, \ldots, N-1$. Затем образуем матрицу $A = (a_{ij})$ и вектор

$$\overrightarrow{b} =\uparrow (b_1, b_2, \ldots, b_{N-1}).$$

Тогда полученные выше условия стационарности запишутся в виде следующей системы линейных уравнений:

$$\begin{cases} \sum\limits_{j=0}^{N} a_{ij}\alpha_j = b_i, \\ i = 1, 2, \ldots, N-1, \\ \alpha_0 = u_a, \quad \alpha_N = u_b. \end{cases} \qquad (RM_N)$$

Исключив из первых $(N-1)$ уравнений переменные α_0 и α_N, приведем систему к виду

$$A\overrightarrow{\alpha} = \overrightarrow{d}, \qquad (LR_N)$$

где вектор $\overrightarrow{d} =\uparrow (d_1, d_2, \ldots, d_{N-1})$ имеет координаты

$$d_i = b_i - a_{i0}u_a - a_{iN}u_b, \qquad i = 1, 2, \ldots, N-1.$$

По определению матрица A — симметричная, то есть $a_{ij} = a_{ji}$.

Вектор $\overrightarrow{\alpha} =\uparrow (\alpha_1, \alpha_2, \ldots, \alpha_{N-1})$, являющийся решением СЛАУ (RM_N), имеет в качестве координат коэффициенты искомой линейной комбинации

$$u_N^*(x) = u_a\varphi_0(x) + \sum_{j=1}^{N-1} \alpha_j \varphi_j(x) + u_b\varphi_N(x).$$

Функция $u_N^*(x)$ доставляет функционалу $J(v)$ минимум на подпространстве $\mathbb{U}^N$.

Для полного обоснования изложенного метода Ритца требуется еще доказать:

во-первых, что построенная последовательность функций u_N^* является при $N \to \infty$ фундаментальной в какой-либо норме линейного пространства $\mathbb{U}$.

Во-вторых, нужно также установить, что предел последовательности функций u_N^* является дважды непрерывно дифференцируемой функцией.

4^0. Рассмотрим теперь следующую краевую задачу:

$$\begin{cases} -(p(x)u')' + v(x)u' + q(x)u = f(x), \\ u(a) = u_a, \\ u(b) = u_b, \end{cases} \qquad (E_V)$$

где $a < x < b$ и функция $v(x) \not\equiv 0$.

Задача (E_V) не формулируется в вариационном виде и поэтому применить к ее решению метод Ритца невозможно. Вместо этого дифференциальная формулировка (E_V) заменяется на эквивалентную задачу в **проекционной постановке**, формулировка которой дается ниже.

Назовем **пробной** любую непрерывную и кусочно дифференцируемую на отрезке $[a, b]$

функцию $\varphi = \varphi(x)$, принимающую на краях a и b отрезка нулевые значения. Совокупность всех пробных на $[a, b]$ функций обозначим через Φ.

Возьмем любую функцию φ из Φ и умножим на нее обе части дифференциального уравнения (E_V).

Получившееся в результате поточечное равенство проинтегрируем по отрезку $[a, b]$.

Тогда придем к следующему интегральному тождеству:

$$\int\limits_a^b \Big[-(p(x)u')' + v(x)u' + q(x)u\Big]\varphi(x)dx =$$

$$= \int\limits_a^b f(x)\varphi(x)dx. \qquad (IT)$$

Функция φ здесь пробная, то есть принадлежит Φ, а в остальном произвольна.

Таким образом, если функция $\boldsymbol{u = u(x)}$ — это решение дифференциальной задачи $(\boldsymbol{E_V})$, то эта же функция обязана удовлетворять и семейству интегральных равенств $(\boldsymbol{IT})$.

С другой стороны, если интегральное тождество $(\boldsymbol{IT})$ выполнено для произвольной пробной функции φ из пространства Φ, то, как доказывается в вариационном исчислении, функция $u(x)$ с необходимостью является решением дифференциального уравнения

$$-(p(x)u')' + v(x)u' + q(x)u = f(x), \qquad a < x < b.$$

Таким образом, исходное дифференциальное уравнение допускает эквивалентную переформулировку в виде семейства интегральных тождеств (IT).

Задача, сформулированная в $(\boldsymbol{IT})$-виде, называется **проекционной**. Приведем ее более подробную формулировку.

Требуется найти такую функцию $\boldsymbol{u = u(x)}$, для которой

$$\boldsymbol{u(a) = u_a, \quad u(b) = u_b}$$

и при этом выполняются соотношения

$$\int_a^b [-(p(x)u')' + v(x)u' + q(x)u]\varphi(x)dx =$$

$$= \int_a^b f(x)\varphi(x)dx. \qquad (IT')$$

Здесь функция φ пробная, то есть принадлежит Φ, а в остальном произвольна.

Интегральному тождеству в проекционной постановке придают несколько иную форму с помощью интегрирования по частям. Для любой пробной функции φ из Φ имеет место равенство

$$-\int\limits_a^b (p(x)u')'\varphi(x)dx = -p(x)u'\varphi(x)\Big|_a^b + \int\limits_a^b p(x)u'\varphi' dx.$$

По условию $\varphi(a) = \varphi(b) = 0$ и поэтому

$$-\int_a^b (p(x)u')'\varphi(x)dx = \int_a^b p(x)u'\varphi' dx.$$

Подставляя это равенство в (IT'), получаем

$$\int_a^b [p(x)u'\varphi' + v(x)u'\varphi + q(x)u\varphi]dx = \int_a^b f(x)\varphi(x)dx. \tag{IT''}$$

Здесь функция φ пробная, то есть принадлежит Φ, а в остальном произвольна.

С помощью замены исходного дифференциального уравнения семейством интегральных тождеств (IT'') получена возможность рассматривать в качестве решений исходной краевой задачи функции, обладающие лишь **первыми** кусочно непрерывными производными.

Этот переход важен: он позволяет перейти от классических гладких решений дифференциальных уравнений к их обобщенным решениям.

Проекционная постановка оказывается также весьма удобной при рассмотрении уравнений с кусочно непрерывными коэффициентами $p = p(x)$ и $q = q(x)$.

5^0. Найдем решение интегрального тождества (IT'') в виде линейной комбинации

$$u_N(x) = \sum_{j=0}^{N} \alpha_j \varphi_j(x).$$

Определение. *Пусть функция $u_N(x)$ удовлетворяет интегральным соотношениям*

$$\int_a^b \left[p(x)u_N'\varphi' + v(x)u_N'\varphi + q(x)u_N\varphi\right]dx = \int_a^b f(x)\varphi(x)dx,$$

$$\forall\, \varphi \in \mathrm{span}\,\{\varphi_0, \varphi_1, \ldots, \varphi_N\} = \mathbb{U}^N.$$

Тогда $u_{\mathbf{N}} = u_{\mathbf{N}}(x)$ *называется* ***приближением по Галеркину*** *решения* $u(x)$ *исходной краевой задачи.*

Расчетные формулы для отыскания коэффициентов приближенного по Галеркину решения $u_N(x)$ имеют следующий вид:

$$
\begin{cases}
\int\limits_a^b \Big[p(x)(\sum\limits_{j=0}^{N} \alpha_j \varphi'_j)\varphi'_i + v(x)(\sum\limits_{j=0}^{N} \alpha_j \varphi'_j)\varphi_i\Big] dx + \\
+ \int\limits_a^b q(x)(\sum\limits_{j=0}^{N} \alpha_j \varphi_j)\varphi_i dx = \int\limits_a^b f(x)\varphi_i(x)dx, \\
i = 1, 2, \ldots, N-1, \\
\alpha_0 = u_a, \quad \alpha_N = u_b.
\end{cases}
$$

Если функция $v(x)$ тождественно равна нулю, то полученная СЛАУ в точности такая же, как и для приближения по методу Ритца. Таким образом, в этом случае оба метода совпадают.

Отметим, что методы Ритца и Галеркина дают приближение к решению $u(x)$ как функцию, значения которой в любой точке из $[a, b]$

вычисляются с помощью подстановки найденных коэффициентов линейной комбинации в формулу общего вида соответствующего приближения.

Таким образом, в случае метода Галеркина изначально указан естественный интерполянт для исходной функции. В этом важное отличие методов Ритца и Галеркина от разностных методов, в которых интерполянт еще требуется сконструировать.

В общем случае применение методов Ритца и Галеркина к решению краевых задач приводит к системам линейных уравнений $\boldsymbol{A}\overrightarrow{\boldsymbol{x}} = \overrightarrow{\boldsymbol{b}}$ с заполненными матрицами $\boldsymbol{A}$, что отличает эти методы от разностных, в которых матрица $\boldsymbol{A}$ основной системы разрежена.

При больших значениях N соответствующая матрица $A = A_N$ оказывается плохо обусловленной, причем с ростом N обусловленность только ухудшается.

6^0. Метод конечных элементов (МКЭ) основан на специальном выборе базисных функций с последующим использованием идеи проекционного метода.

МКЭ был предложен Р. Курантом в 1943 г. В начале 50-х годов XX века инженеры-специалисты по строительной механике разработали новый подход к численному решению задач теории упругости.

Характерной чертой этого нового подхода служило разбиение расчетной области сложной геометрии на конечное число подобластей простой геометрии, в каждой из которых решение искалось аналитически.

Эти элементарные подобласти получили название **конечных элементов**, а сам подход назвали МКЭ. С той поры теория МКЭ бурно развивалась, проникая во все новые области естествознания. К настоящему времени МКЭ получил самое широкое распространение в области вычислительной математики. На его основе создано огромное число пакетов прикладных программ для решения научных задач.

Кроме разбиения расчетной области на конечное число элементарных подмножеств стандартной и по возможности единообразной формы, в МКЭ используются базисные функции φ_j достаточно простого вида.

Локально базисная функция $\varphi_j(x)$ задается чаще всего как полином. При этом φ_j отличается от нуля лишь в нескольких соседних конечных элементах.

Рассмотрим МКЭ в одномерном случае в применении к решению краевой задачи

$$\begin{cases} -(p(x)u')' + q(x)u = f(x), \qquad a < x < b, \\ u(a) = u_a, \\ u(b) = u_b, \end{cases}$$

Отрезок $[a, b]$ разобьем узловыми точками

$$x_0 = a < x_1 < \ldots < x_{N-1} < b = x_N$$

на N элементарных отрезков $[x_{i-1}, x_i]$ длины $h_i = x_i - x_{i-1}$. Каждый из этих малых отрезков выступает в роли **конечного элемента**.

Базисную функцию $\varphi_j(x)$ введем как кусочно линейную, равную единице в узле x_j и нулю во всех остальных x_i.

Точнее, $\varphi_j(x)$ при $j = 1, 2, \ldots, N-1$ зададим следующим образом:

$$\varphi_j(x) = \begin{cases} \frac{x - x_{j-1}}{h_j} & \text{при } x_{j-1} \le x \le x_j; \\ \frac{x_{j+1} - x}{h_{j+1}} & \text{при } x_j \le x \le x_{j+1}; \\ 0 & \text{для } x \notin [x_{j-1}, x_{j+1}]. \end{cases}$$

График такой функции — это треугольник с основанием $[x_{j-1}, x_{j+1}]$ на оси абцисс и высотой 1. При увеличении j от 1 до $N-1$ этот график — треугольник перемещается вдоль оси Ox в положительном направлении.

График функции $\varphi_0(x)$ — это прямоугольный треугольник с левой вершиной $(x_0, 0)$ в точке с координатами $(a, 0)$, с правой вершиной $(x_1, 0)$ в точке $(a, a+h_1)$, и с третьей вершиной в точке $(0, y_1) = (0, 1)$.

Аналогично, график функции $\varphi_N(x)$ — это прямоугольный треугольник с вершинами в точках $(b - h_N, 0)$, $(b, 0)$, $(0, 1)$.

Реализуем метод Ритца с таким образом выбранным базисом

$$\{\varphi_0(x),\ \varphi_1(x),\ \varphi_2(x),\ \ldots,\ \varphi_N(x)\}.$$

Полагая $u_N(x) = \sum\limits_{j=0}^{N} \alpha_j \varphi_j(x)$, возьмем здесь

$x = x_j$. Тогда получим, что значения приближения по Ритцу $u_N(x)$ в узлах совпадают с коэффициентами α_j:

$$u_N(x_j) \equiv u_j^h = \alpha_j, \quad j = 0, 1, 2, \ldots, N.$$

Коэффициенты $\{\alpha_j\}$ при этом находим как решение СЛАУ (RM_N), то есть

$$\begin{cases} \sum\limits_{j=0}^{N} a_{ij}\alpha_j = b_i, \\ i = 1, 2, \ldots, N-1, \\ \alpha_0 = u_a, \quad \alpha_N = u_b. \end{cases}$$

Здесь коэффициент a_{ij} задается равенством

$$a_{ij} = \int\limits_a^b \Big[p(x)\varphi_i'(x)\varphi_j'(x) + q(x)\varphi_i(x)\varphi_j(x)\Big]dx.$$

Но произведения $\varphi'_i(x)\varphi'_j(x)$ и $\varphi_i(x)\varphi_j(x)$ могут быть отличны от нуля лишь при условии, что $|i-j| \le 1$, то есть при $j = i-1$, $j = i$ или же $j = i+1$. Для всех остальных номеров i, j имеем $a_{ij} = 0$.

Следовательно, система линейных уравнений относительно неизвестных $\alpha_j = u_j^h$ имеет матрицу трехдиагонального вида:

$$\begin{cases} a_{i,j-1}u_{i-1}^h + a_{i,i}u_i^h + a_{i,i+1}u_{i+1}^h = b_i, \\ 1 \le i \le N-1, \\ u_0^h = u_a, \quad u_N^h = u_b. \end{cases} \tag{FEM_N}$$

Коэффициенты этой разностной схемы вычисляются по следующим расчетным формулам:

$$a_{i,j-1} = -\frac{1}{h_i}p_{i-\frac{1}{2}}^h + h_{i+\frac{1}{2}}q_{i-\frac{1}{2}}^h;$$

$$a_{i,i+1} = -\frac{1}{h_{i+1}} p^h_{i+\frac{1}{2}} + h_{i+\frac{1}{2}} q^h_{i+\frac{1}{2}};$$

$$a_{i,i} = \frac{1}{h_i} p^h_{i-\frac{1}{2}} + \frac{1}{h_{i+1}} k^h_{i+\frac{1}{2}} + h_{i+\frac{1}{2}} q^h_i;$$

$$b_i = h_{i+\frac{1}{2}} f^h_i, \qquad h_{i+\frac{1}{2}} = \frac{h_i + h_{i+1}}{2}.$$

Числовые параметры в правой части этих формул определяются равенствами

$$p^h_{i-\frac{1}{2}} = \frac{1}{h_i} \int\limits_{x_{i-1}}^{x_i} p(x)dx;$$

$$q^h_{i\pm\frac{1}{2}} = \frac{1}{h_{i+1/2}} \int\limits_{x_{i-1}}^{x_{i+1}} q(x)\varphi_i(x)\varphi_{i\pm1}(x)dx;$$

$$q^h_i = \frac{1}{h_{i+1/2}} \int\limits_{x_{i-1}}^{x_{i+1}} q(x)\varphi^2_i(x)dx;$$

$$f^h_i = \frac{1}{h_{i+1/2}} \int\limits_{x_{i-1}}^{x_{i+1}} f(x)\varphi_i(x)\,dx.$$

Систему $(\boldsymbol{FEM}_{\boldsymbol{N}})$ называют **системой метода конечных элементов**, или проекционно-разностной схемой. Также эту СЛАУ называют **проекционно-сеточным методом**, или же **вариационно-разностной схемой**. Система $(\boldsymbol{FEM}_{\boldsymbol{N}})$ имеет единственное решение, которое можно отыскать методом прогонки.

При некоторых условиях на сетку разбиения и коэффициенты $p(x)$, $q(x)$, $f(x)$ проекционно-разностная схема имеет второй порядок точности.

При решении системы $(\boldsymbol{FEM}_{\boldsymbol{N}})$ возникает необходимость вычисления большого количества интегралов, через которые определяются коэффициенты системы и ее правая часть. В случае, когда функции $p(x)$, $q(x)$ и $f(x)$ гладкие, указанные интегралы вычисляются с помощью квадратурных формул.

Тема : Метод наименьших квадратов

$\mathbf{1^0}$. Постановка линейной задачи метода наименьших квадратов. Связь с задачей интерполяции функции обобщенным полиномом конечной длины в большом числе узлов. Среднеквадратичное уклонение. Полином наилучшего среднеквадратичного приближения. $\mathbf{2^0}$. Нормальная система метода наименьших квадратов. Теорема о существовании единственного полинома наилучшего среднеквадратичного приближения. Случай алгебраических полиномов. Пример. $\mathbf{3^0}$. Вычислительные аспекты задачи метода наименьших квадратов: симметричность и положительная определенность матрицы нормальной системы, ее обусловленность. Зависимость обусловленности от выбора полиномиального базиса.

1^0. Пусть функция $y = f(x)$ задана некоторой таблицей приближенных значений

$$y_i \approx f(x_i), \quad i = 0, 1, 2, \ldots, n.$$

Величины y_i известны с ошибками $\varepsilon_i = y_i^0 - y_i$, где $y_i^0 = f(x_i)$. Если значения y_i получены экспериментально, то ошибки носят случайный характер. Уровень погрешности (“шумов”) при этом бывает весьма существенным.

Предположим, что для восстановления $f(x)$ используются линейные комбинации вида

$$\Phi_m(x) = a_0\varphi_0(x) + a_1\varphi_1(x) + \ldots + a_m\varphi_m(x).$$

Здесь $\varphi_0(x)$, $\varphi_1(x)$, ..., $\varphi_m(x)$ — заданные вещественнозначные базисные функции. Коэффициенты a_0, a_1, ..., a_m — это параметры модели, которые требуется определить по исходной таблице.

Функция $\Phi_m(x)$ называется **обобщенным полиномом**. Часто в качестве базисных выбираются степени независимой переменной. В этом случае вместо обобщенного используется обычный полином

$$P_m(x) = a_0 + a_1 x + a_2 x^2 + \ldots + a_m x^m.$$

Если уровень неопределенности исходных данных высок, неестественно требовать от модели совпадения значений обобщенного полинома $\Phi_m(x)$ в узловых точках x_i таблицы с величинами y_i.

Иными словами, не имеет смысла использовать в этом случае обычную интерполяцию. При таком интерполировании происходит **повторение** исходных больших ошибок наблюдений, в то время как обработка экспериментальных данных требует **сглаживания ошибок** (избавления от шумов).

Отказ от требования выполнения в точках x_i точных равенств $\Phi_m(x_i) = y_i$ все же необходимо компенсировать требованием приближенных равенств $\Phi_m(x_i) \approx y_i$.

Система приближенных равенств при этом

записывается в покомпонентном виде:

$$\begin{cases} a_0\varphi_0(x_0)+a_1\varphi_1(x_0)+\ldots+a_m\varphi_m(x_0)\approx y_0, \\ a_0\varphi_0(x_1)+a_1\varphi_1(x_1)+\ldots+a_m\varphi_m(x_1)\approx y_1, \\ a_0\varphi_0(x_2)+a_1\varphi_1(x_2)+\ldots+a_m\varphi_m(x_2)\approx y_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots, \\ a_0\varphi_0(x_n)+a_1\varphi_1(x_n)+\ldots+a_m\varphi_m(x_n)\approx y_n. \end{cases}$$

Можно предложить различные критерии, позволяющие выбрать параметры a_0, a_1, ..., a_m так, чтобы приближенные равенства выполнялись наилучшим в каком-либо смысле образом.

Один из таких критериев, используемый чаще всего, называется **методом наименьших квадратов**. В этом методе минимизируется **среднеквадратичное уклонение**, определяемое равенством

$$\delta(\Phi_m, \overrightarrow{y}) = \left[\frac{1}{n+1}\sum_{i=0}^{n}(\Phi_m(x_i) - y_i)^2\right]^{1/2}.$$

Величина $\delta(\Phi_m, \overrightarrow{y})$ является мерой отклонения обобщенного полинома Φ_m от вектора $\overrightarrow{y} = (y_0, y_1, \ldots, y_n)$ заданных табличных значений.

Минимум среднеквадратичного отклонения достигается при тех же значениях коэффициентов a_0, a_1, ..., a_m, что и минимум функции, задаваемой следующим равенством

$$S(\overrightarrow{a}, \overrightarrow{y}) = \sum_{i=0}^{n} \left\{ \sum_{j=0}^{m} a_j \varphi_j(x_i) - y_i \right\}^2 = \|P\overrightarrow{a} - \overrightarrow{y}\|_2^2.$$

Здесь P — прямоугольная матрица размеров $n \times m$, определяемая по значениям базисных функций $\varphi_j(\cdot)$ в узлах интерполяции следующим образом:

$$P = \begin{pmatrix} \varphi_0(x_0) & \varphi_1(x_0) & \cdots & \varphi_m(x_0) \\ \varphi_0(x_1) & \varphi_1(x_1) & \cdots & \varphi_m(x_1) \\ \varphi_0(x_2) & \varphi_1(x_2) & \cdots & \varphi_m(x_2) \\ \vdots & \vdots & \vdots & \vdots \\ \varphi_0(x_n) & \varphi_1(x_n) & \cdots & \varphi_m(x_n) \end{pmatrix}. \qquad (P_M)$$

Функции $\delta(\Phi_m, \overrightarrow{y})$ и $S(\overrightarrow{a}, \overrightarrow{y})$ связаны между собой равенством

$$[\delta(\Phi_m, \overrightarrow{y})]^2 = \frac{1}{n+1} S(\overrightarrow{a}, \overrightarrow{y}).$$