

Вопрос №13

- Задача минимизации в линейной регрессии в одномерном и многомерном случаях. Нормальное уравнение (формулировка)

Одномерный случай:

$$\min_{a_0, a_1} E_2(a_0, a_1) =$$

$$\min_{a_0, a_1} \sum_{i=1}^n [y_i - f(x_i)]^2$$

$E_2(a_0, a_1)$ представляет собой суммы квадратов отклонений.

Аппроксимация функцией вида $f(x) = a_0 + a_1 x$ подставим в

$$E_2(a_0, a_1) \text{ и получаем}$$

$$\min_{a_0, a_1} E_2(a_0, a_1) =$$

$$\min_{a_0, a_1} \sum_{i=1}^n [y_i - a_0 - a_1 x_i]^2$$

Многомерный случай:

Линейная регрессия производится с помощью гиперплоскостей. Для перехода к многомерному случаю нам необходимо воспользоваться векторной формой записи задачи оптимизации

$$\min_a E_2(a) = \min_a (y - Xa)^T (y - Xa), \text{ где } x \in \mathbb{R}^{m \times (n+1)}, a \in \mathbb{R}^{n+1} \text{ и } y \in \mathbb{R}^m.$$

Здесь m обозначает число дискретных данных, а n – размерность пространства данных. Строка матрицы X хранит в себе одну многомерную точку данных, при этом первый столбец является столбцом единиц, что и дает размерность матрицы $m \times (n + 1)$. В общем виде матрица X может быть представлена следующим образом:

$$X = \begin{bmatrix} 1 & x_1^{(1)} & \dots & x_n^{(1)} \\ 1 & x_1^{(2)} & \dots & x_n^{(2)} \\ \vdots & \vdots & \dots & \vdots \\ 1 & x_1^{(m)} & \dots & x_n^{(m)} \end{bmatrix},$$

где нижние индексы обозначают координаты многомерной точки данных, а верхние индексы номера точек.

Нормальное уравнение: $a = (X^T X)^{-1} X^T y$

- Вывод нормального уравнения.

$$E_2(a) = y^T y - a^T X^T y - y^T X a + a^T X^T X a = y^T y - 2a^T X^T y + a^T X^T X a \Rightarrow \frac{\partial E_2}{\partial a} = -2X^T y + 2X^T X a$$

В справедливости подобного «векторного» дифференцирования можно убедиться, перейдя от матричной формы к суммированию:

$$E_2(a) = \sum_{j=1}^m y_j^2 - 2 \sum_{j=1}^m \sum_{i=0}^n a_i X_{ji} y_j + \sum_{j=1}^m \left(\sum_{i=0}^n a_i X_{ji} \right)^2.$$

Тогда дифференцирование относительно произвольного a_k дает

Вопрос №20

- Описание метода Гаусса. Алгоритмическая сложность метода Гаусса. Вычислительная неустойчивость метода Гаусса и способы ее устранения.

Прямой ход:

Метод Гаусса состоит в приведении матрицы A к треугольному виду с помощью элементарных преобразований, после чего решение x находится сравнительно легко.

$$\tilde{A} = [A, b] = \left[\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} & b_n \end{array} \right].$$

Для того, чтобы привести матрицу A к треугольному виду, необходимо последовательно обнулять элементы, находящиеся под главной диагональю.

$$\tilde{A} = \left[\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ 0 & a_{22}^{(1)} & \dots & a_{2n}^{(1)} & b_2^{(1)} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & a_{nn}^{(n-1)} & b_n^{(n-1)} \end{array} \right].$$

Обратный ход:

(нахождение решения x с помощью рекурсивной подстановки)

$$x_n = \frac{b_n^{(n-1)}}{a_{nn}^{(n-1)}},$$

$$x_{n-1} = \frac{b_{n-1}^{(n-2)} - a_{n-1,n}^{(n-2)} x_n}{a_{n-1,n-1}^{(n-2)}},$$

$$\dots$$

$$x_1 = \frac{b_1 - \sum_{i=2}^n a_{1i} x_i}{a_{11}}.$$

Алгоритмическая сложность метода Гаусса:

(а) Прямой ход $O(n^3)$

(б) Обратный ход $O(n^2)$

Метод Гаусса в общем случае является неустойчивым (погрешности возникают, когда диагональный элемент много меньше хотя бы одного элемента в этой же строке)

Способы устранения:

(а) Частичный выбор главного элемента (перестановки строк)

$$|a_{kk}^{(k)}| = \max_{k \leq i \leq n} |a_{ik}^{(k)}|.$$

– $O(n^2)$ сравнений

– используется по умолчанию почти во всех алгоритмах

(б) Полный выбор главного элемента (перестановки строк и столбцов)

$$|a_{kk}^{(k)}| = \max_{k \leq i, j \leq n} |a_{ij}^{(k)}|.$$

– $O(n^3)$ сравнений

– используется только в тех задачах, где особая чувствительность к погрешностям

Вопрос №3

- Приближение, интерполяция и экстраполяция. Наилучшее приближение (задачи минимизации для минимакса и метода наименьших квадратов)

Задача приближения или аппроксимации (эти термины синонимичны) состоит в представлении недоступной функции $f(x)$ в виде более простой аналитической функции по n значениям $f(x_i)$, где $x_i \in [a; b]$ называются узлами. Для решения этой задачи мы задаемся некоторой параметризованной формой для $\tilde{f}(x; c)$, где c – вектор параметров, а затем подбираем c так, что отклонение $\rho[f(x) - \tilde{f}(x; c)]$ минимизировано, где ρ – некоторая метрика, определяющая оценку погрешности приближения. Частными случаями аппроксимации являются интерполяция и экстраполяция.

Интерполяцией называется приближение, при котором требуется, чтобы $\tilde{f}(x)$ проходила через заданные узлы $(x_i, f(x_i))$ внутри отрезка $x \in [a; b]$.

Экстраполяцией называется приближение, при котором требуется, чтобы $\tilde{f}(x)$ по заданным узлы $(x_i, f(x_i))$ предсказывала значение $f(x)$ вне отрезка $[a; b]$.

Наилучшее приближение предполагает приближение полинома с ЗАРАНЕЕ выбранной степенью путем «подгонки» его параметров. ИДЕТ УЧЕТ ПОГРЕШНОСТИ. Для примера рассмотрим набор дискретных данных $D = \{(x_i, y_i)\}_{i=1}^n$. В случае линейной регрессии нам необходимо найти такую функцию $f(x) = a_0 + a_1 x$, что она будет приближаться к данным D наилучшим образом. Самый очевидный подход состоит в решении следующей оптимизационной задачи:

$$\min_{a_0, a_1} E_\infty(a_0, a_1) = \min_{a_0, a_1} \max_{a_0, a_1} |y_i -$$

$f(x_i)|$ известной под названием минимакс (минимизация максимальной ошибки), и результатом которой является наилучший аппроксимационный полином. (наилучшее решение; минимизируем максимальные отклонения; аналитического решение получить трудно даже для линейной функции)

Метод наименьших квадратов представляет собой суммы квадратов отклонений.

$$\min_{a_0, a_1} E_2(a_0, a_1) = \min_{a_0, a_1} \sum_{i=1}^n [y_i -$$

$$f(x_i)]^2$$

(Наиболее комфортный способ; минимизирует суммы квадратов отклонений)

- Приближение тригонометрическим рядом к непрерывной функции (решение оптимизационной задачи метода наименьших квадратов)

Рассмотрим приближение к непрерывной функции $y: \mathbb{R} \rightarrow \mathbb{R}$ с помощью тригонометрического ряда:

$$\frac{\partial E_2}{\partial a_k} = -2 \sum_{j=1}^m X_{jk} y_j + 2 \sum_{j=1}^m X_{jk} \left(\sum_{i=0}^n a_i X_{ji} \right),$$

что при записи в матричном виде дает. Оптимальное значение вектора a находится с помощью приравнивания производной к нулю:

$$-2X^T y + 2X^T X a = 0$$

$$\Rightarrow a = (X^T X)^{-1} X^T y.$$

2. Вывод количества операций в методе Гаусса.

В первую очередь заметим, что на k -й итерации из всего $(n-1)$ итераций прямого хода требуется $(n-k)$ делений для вычисления множителя $m_{ik} = \frac{a_{ik}^{(k)}}{a_{kk}^{(k)}}$. Затем требуется для $(n-k)$ строк произвести $(n-k+1)$ перемножений, т.е. суммарно $(n-k)(n-k+1)$. Таким образом, на k -й итерации мы имеем $n-k+(n-k)(n-k+1)=(n-k)(n-k+2)$ умножений и делений. Суммируя по $k=1, \dots, n-1$, получаем общее число умножений и делений в прямом ходе метода Гаусса:

$$\sum_{k=1}^{n-1} (n-k)(n-k+2) = \sum_{k=1}^{n-1} (n-k)^2 + 2 \sum_{k=1}^{n-1} (n-k) = \sum_{k=1}^{n-1} k^2 + 2 \sum_{k=1}^{n-1} k$$

По арифметической прогрессии,

$$S_n = \frac{(k_1 + k_n)}{2} (n-1) \text{ получаем:}$$

$$\frac{(n-1)n(2n-1)}{6} + 2 \frac{(n-1)n}{2} = \frac{2n^3 + 3n^2 - 5n}{6}$$

В обратном ходе метода Гаусса для нахождения x_k требуется $n-k$ перемножений и одно деление, что в сумме дает:

$$\sum_{k=1}^n (n-k+1) = \sum_{k=1}^n k = \frac{n(n+1)}{2} = \frac{n^2 + n}{2}.$$

Наконец, суммарно требуется следующее число операций для нахождения решения методом Гаусса:

$$\frac{2n^3 + 3n^2 - 5n}{6} + \frac{n^2 + n}{2} = O(n^3)$$

$$f(x) = S_n(x) = a_0 + \sum_{k=1}^n (a_k \cos kx + b_k \sin kx)$$

$$= \sum_{k=0}^n (a_k \cos kx + b_k \sin kx),$$

где $a_k, b_k \in \mathbb{R}$ для $k=0, \dots, n$ и b_0 может быть произвольным, так как $\sin 0 = 0$. Оптимизационная задача метода наименьших квадратов в таком случае имеет вид:

$$\min_a E_2(a) =$$

$$\min_a \int_{-\pi}^{\pi} \omega(x) [y(x) - \sum_{k=0}^n (a_k * \cos kx + b_k \sin kx)]^2$$

система тригонометрических функций $\{\cos kx, \sin kx\}_{k=0}^n$ является

ортогональной на $[-\pi; \pi]$ с весом $\omega(x)=1$. Тогда пользуясь формулой нахождения

оптимальных значений $a_k =$

$$\frac{\langle y(x), \phi_k(x) \rangle_{\omega}}{\langle \phi_k(x), \phi_k(x) \rangle_{\omega}}, \text{ (где ортогональные } \phi_k(x) = \cos kx, \sin kx \text{)}$$

для оптимальных коэффициентов метода наименьших квадратов, можно легко вывести следующие выражения для оптимальных a_k и b_k :

$$a_0 = \frac{1}{2\pi} \int_{-\pi}^{\pi} y(x) dx,$$

$$a_k = \frac{1}{\pi} \int_{-\pi}^{\pi} y(x) \cos kx dx,$$

$$b_k = \frac{1}{\pi} \int_{-\pi}^{\pi} y(x) \sin kx dx.$$

Вопрос №21

- Положительно определенные матрицы (определение). Особенности решения СЛАУ для положительно определенных матриц. Разложения для положительно определенных матриц.

Матрица A называется положительно определенной, если она симметричная, и верным является неравенство $x^T A x > 0$ для любого вектора $x \neq 0$ подходящей размерности.

$$x^T A x = \sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i x_j.$$

Пусть A - положительно определенная матрица. Тогда верно следующее:

- существует A^{-1} ;
- $a_{ii} > 0, i = 1, \dots, n$;
- $\max_{1 \leq i \leq n} |a_{ii}| \geq \max_{1 \leq k, j \leq n} |a_{kj}|$;
- $a_{ii} a_{jj} > a_{ij}^2, i \neq j$

Особенностью положительно определенных матриц является тот факт, что метод Гаусса для них является устойчивым с вычислительной точки зрения и не требует перестановки строк. Более того, для положительно определенных матриц существует два специальных вида разложений - LDL^T и LL^T разложения. Разложение вида LDL^T существует также для симметричных матриц.

Матрица A является положительно определенной тогда и только тогда, когда существует разложение $A = LDL^T$, где L - нижняя треугольная матрица с единицами на диагонали, D - диагональная матрица с положительными элементами. В качестве примера LDL^T -разложения рассмотрим положительно определенную матрицу размерности 3×3 :

$$\begin{aligned} & \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ l_{21} & 1 & 0 \\ l_{31} & l_{32} & 1 \end{bmatrix} \begin{bmatrix} d_1 & 0 & 0 \\ 0 & d_2 & 0 \\ 0 & 0 & d_3 \end{bmatrix} \begin{bmatrix} 1 & l_{21} & l_{31} \\ 0 & 1 & l_{32} \\ 0 & 0 & 1 \end{bmatrix} \\ & = \begin{bmatrix} 1 & 0 & 0 \\ l_{21} & 1 & 0 \\ l_{31} & l_{32} & 1 \end{bmatrix} \begin{bmatrix} d_1 & d_1 l_{21} & d_1 l_{31} \\ 0 & d_2 & d_2 l_{32} \\ 0 & 0 & d_3 \end{bmatrix} \\ & = \begin{bmatrix} d_1 & d_1 l_{21} & d_1 l_{31} \\ d_1 l_{21} & d_1 l_{21}^2 + d_2 & d_1 l_{21} l_{31} + d_2 l_{32} \\ d_1 l_{31} & d_1 l_{21} l_{31} + d_2 l_{32} & d_1 l_{31}^2 + d_2 l_{32}^2 + d_3 \end{bmatrix} \end{aligned}$$

Тогда неизвестные элементы матрицы вычисляются следующим образом:

Вопрос №14

- Тригонометрические ряды. Запись тригонометрических рядов в экспоненциальной форме.

Тригонометрический ряд:

$$\sum_{k=0}^n (a_k \cos kx + b_k \sin kx)$$

Экспоненциальная форма:

$$\sum_{k=-n}^n \hat{a}_k e^{ikx}$$

Доказательство:

Рассмотрим формулу Эйлера:

$$e^{ikx} = \cos kx + i \sin kx \quad | \cdot \hat{a}_k \in \mathbb{C}$$

$$\hat{a}_k e^{ikx} = \hat{a}_k \cos kx + i \hat{a}_k \sin kx =$$

$$\operatorname{Re}(\hat{a}_k) \cos kx + i \operatorname{Im}(\hat{a}_k) \cos kx -$$

$$\operatorname{Im}(\hat{a}_k) \sin kx + i \operatorname{Re}(\hat{a}_k) \sin kx, \text{ где}$$

$\operatorname{Re}(\hat{a}_k)$ и $\operatorname{Im}(\hat{a}_k)$ - вещественная и мнимая части \hat{a}_k соответственно.

Добавим к обоим частям равенства комплексно сопряженное (та же формула, но там где $+i$ добавляется $-i$)

$$\hat{a}_k e^{ikx} + \hat{a}_k^* e^{-ikx} =$$

$$2 \operatorname{Re}(\hat{a}_k) \cos kx - 2 \operatorname{Im}(\hat{a}_k) \sin kx \text{ где}$$

комплексно сопряженное число обозначено через $*$. Тогда

суммирование обеих частей

равенства по $k=0, \dots, n$ приводит к

выражению:

$$\sum_{k=0}^n (\hat{a}_k e^{ikx} + \hat{a}_k^* e^{-ikx}) =$$

$$\sum_{k=-n}^n \hat{a}_k e^{ikx} =$$

$$\sum_{k=0}^n [2 \operatorname{Re}(\hat{a}_k) \cos kx -$$

$$2 \operatorname{Im}(\hat{a}_k) \sin kx], \text{ где } \hat{a}_k^* = \hat{a}_{-k}$$

Таким образом мы получили

экспоненциальную форму

тригонометрического ряда, где

коэффициенты связаны друг с

другом следующим образом:

$$a_k = 2 \operatorname{Re}(\hat{a}_k)$$

$$b_k = -2 \operatorname{Im}(\hat{a}_k), \text{ где } k = 0, \dots, n$$

2. Вывод

тригонометрического

полинома из

тригонометрического ряда.

Для того, чтобы перейти к форме

полинома, обозначим $z = e^{ix}$. Тогда

имеем:

$$\sum_{k=-n}^n \hat{a}_k e^{ikx} = \sum_{k=-n}^n \hat{a}_k z^k \text{ что при}$$

домножении на z^n дает

комплекснозначный полином

степени $2n$, по конвенции

называемый тригонометрическим

полиномом n -й степени:

$$Q_n(z) = z^n \sum_{k=-n}^n \hat{a}_k z^k$$

Вопрос №22

- Ленточные матрицы (определение). Трехдиагональные матрицы (определение). Метод решения СЛАУ с трехдиагональной матрицей коэффициентов и его сложность.

Квадратная матрица $A \in \mathbb{R}^{n \times n}$ называется ленточной, если существуют такие $p, q \in \mathbb{Z}$, где $1 < p, q < n$, что $j - i \geq p \Rightarrow a_{ij} = 0$ и $i - j \geq q \Rightarrow a_{ij} = 0$. Ширина ленты при этом определяется как $w = p + q - 1$.

Где p обозначает число ненулевых диагоналей над главной диагональю (включая ее), в то время как q обозначает число ненулевых диагоналей под главной диагональю (включая ее)

Частным случаем ленточной матрицы является трехдиагональная матрица, возникающая, например, при выводе разрешающих уравнений кубического сплайна:

$$A = \begin{bmatrix} a_{11} & a_{12} & 0 & \dots & \dots & \dots & 0 \\ a_{21} & a_{22} & a_{23} & \ddots & & & \vdots \\ 0 & a_{32} & a_{33} & a_{34} & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & & a_{n-1,n-2} & a_{n-1,n-1} & a_{n-1,n} \\ 0 & \dots & \dots & \dots & 0 & a_{nn-1} & a_{nn} \end{bmatrix}$$

Рассмотрим специальный метод решения СЛАУ, имеющих трехдиагональную матрицу - метод прогонки. Пусть мы имеем следующее матричное уравнение:

$$\begin{bmatrix} a_{11} & a_{12} & 0 & \dots & \dots & \dots & 0 \\ a_{21} & a_{22} & a_{23} & \ddots & & & \vdots \\ 0 & a_{32} & a_{33} & a_{34} & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & & a_{n-1,n-2} & a_{n-1,n-1} & a_{n-1,n} \\ 0 & \dots & \dots & \dots & 0 & a_{nn-1} & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_{n-1} \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_{n-1} \\ b_n \end{bmatrix}$$

где каждая строка эквивалентна рекуррентному соотношению вида:

$$a_{i,i-1}x_{i-1} + a_{ii}x_i + a_{i,i+1}x_{i+1} = b_i. \quad (1)$$

Очевидно, что методом последовательного исключения можно привести такую СЛАУ к верхней треугольной форме:

$$\begin{bmatrix} \tilde{a}_{11} & \tilde{a}_{12} & 0 & \dots & \dots & 0 \\ 0 & \tilde{a}_{22} & \tilde{a}_{23} & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & & \tilde{a}_{n-1,n-1} & \tilde{a}_{n-1,n} \\ 0 & \dots & \dots & \dots & 0 & \tilde{a}_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n-1} \\ x_n \end{bmatrix} = \begin{bmatrix} \tilde{b}_1 \\ \tilde{b}_2 \\ \vdots \\ \tilde{b}_{n-1} \\ \tilde{b}_n \end{bmatrix}$$

Тогда решение для x_{i-1} может быть выражено через x_i :

$$x_{i-1} = \frac{\tilde{b}_{i-1} - \tilde{a}_{i-1,i}x_i}{\tilde{a}_{i-1,i-1}},$$

что задает выражение для обратного хода метода Гаусса. Вместо явного вывода неизвестных коэффициентов, мы построим рекуррентное соотношение, позволяющее рекурсивно найти нужные коэффициенты. Для упрощения записи переопределим коэффициенты:

$$x_{i-1} = \gamma_i x_i + \beta_i \quad (2)$$

Тогда подстановка в (1) дает:

$$\begin{aligned}
d_1 &= a_{11}, \\
l_{21} &= \frac{a_{21}}{d_1}, \\
l_{31} &= \frac{a_{31}}{d_1}, \\
d_2 &= a_{22} - d_1 l_{21}^2, \\
l_{32} &= \frac{1}{d_2} (a_{32} - d_1 l_{21} l_{31}), \\
d_3 &= a_{33} - d_1 l_{31}^2 - d_2 l_{32}^2.
\end{aligned}$$

Матрица является положительно определенной тогда и только тогда, когда существует разложение $A = LDL^T$, называемое разложением Холецкого, где L - нижняя треугольная матрица с ненулевыми элементами на диагонали. Из LDL^T -разложения можно легко получить разложение Холецкого, если разложить D как $D = \sqrt{D}\sqrt{D}$. Тогда имеем:

$$\begin{aligned}
LDL^T &= L\sqrt{D}\sqrt{D}L^T = \\
(L\sqrt{D})(L\sqrt{D})^T &= \tilde{L}\tilde{L}^T
\end{aligned}$$

2. Вывод формул для коэффициентов матриц LDL-разложения и разложения Холецкого на примере разложения произвольной матрицы размерности 3×3 .

В очередной раз воспользуемся положительно определенной матрицей размерности 3×3 для демонстрации разложения Холецкого:

$$\begin{aligned}
&\begin{matrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{matrix} = \\
&= \begin{bmatrix} l_{11} & 0 & 0 \\ l_{21} & l_{22} & 0 \\ l_{31} & l_{32} & l_{33} \end{bmatrix} \begin{bmatrix} l_{11} & l_{21} & l_{31} \\ 0 & l_{22} & l_{32} \\ 0 & 0 & l_{33} \end{bmatrix} \\
&= \begin{bmatrix} l_{11}^2 & l_{11}l_{21} & l_{11}l_{31} \\ l_{11}l_{21} & l_{21}^2 + l_{22}^2 & l_{21}l_{31} + l_{22}l_{32} \\ l_{11}l_{31} & l_{21}l_{31} + l_{22}l_{32} & l_{31}^2 + l_{32}^2 + l_{33}^2 \end{bmatrix}
\end{aligned}$$

Неизвестные элементы матрицы L тогда вычисляются следующим образом:

$$\begin{aligned}
l_{11} &= \sqrt{a_{11}}, \\
l_{21} &= \frac{a_{21}}{l_{11}}, \\
l_{31} &= \frac{a_{31}}{l_{11}}, \\
l_{22} &= \sqrt{a_{22} - l_{21}^2}, \\
l_{32} &= \frac{1}{l_{22}} (a_{32} - l_{21}l_{31}), \\
l_{33} &= \sqrt{a_{33} - l_{31}^2 - l_{32}^2}.
\end{aligned}$$

$$\begin{aligned}
a_{i,i-1}(\gamma_i x_i + \beta_i) + a_{ii}x_i + a_{i,i+1}x_{i+1} &= b_i \\
x_i &= \frac{-a_{i,i+1}}{a_{i,i-1}\gamma_i + a_{ii}}x_{i+1} + \frac{b_i - a_{i,i-1}\beta_i}{a_{i,i-1}\gamma_i + a_{ii}}.
\end{aligned}$$

(3)

Сравнив полученное выражение с (2), получаем рекуррентные соотношения для определения γ_i и β_i :

$$\begin{aligned}
\gamma_{i+1} &= \frac{-a_{i,i+1}}{a_{i,i-1}\gamma_i + a_{ii}}, \\
\beta_{i+1} &= \frac{b_i - a_{i,i-1}\beta_i}{a_{i,i-1}\gamma_i + a_{ii}}.
\end{aligned}$$

(4) (5)

Последовательно вычислив все коэффициенты γ_i и β_i , решение СЛАУ находится с помощью обратного хода метода Гаусса по формуле (2). Для завершения построения метода прогонки достаточно найти выражения для коэффициентов γ_1, β_1 и неизвестной x_n . Рассмотрим уравнение для x_1 :

$$\begin{aligned}
a_{11}x_1 + a_{12}x_2 &= b_1 \\
x_1 &= -\frac{a_{12}}{a_{11}}x_2 + \frac{b_1}{a_{11}}.
\end{aligned}$$

Тогда сравнение с (3) дает:

$$\begin{aligned}
-\frac{a_{12}}{a_{10}\gamma_1 + a_{11}} &= -\frac{a_{12}}{a_{11}}, \\
\frac{b_1 - a_{10}\beta_1}{a_{10}\gamma_1 + a_{11}} &= \frac{b_1}{a_{11}},
\end{aligned}$$

из чего следует $\gamma_1 = \beta_1 = 0$ (или, эквивалентно, $a_{10} = 0$). Для определения x_n рассмотрим уравнение, соответствующее последней строке матрицы:

$$\begin{aligned}
a_{n,n-1}x_{n-1} + a_{nn}x_n &= b_n \\
x_n &= -\frac{a_{n,n-1}}{a_{nn}}x_{n-1} + \frac{b_n}{a_{nn}}.
\end{aligned}$$

Подстановка в уравнение (2), записанное для $i = n$ дает:

$$\begin{aligned}
x_{n-1} &= \gamma_n x_n + \beta_n \\
x_n &= \frac{b_n - a_{n,n-1}\beta_n}{a_{nn} + a_{n,n-1}\gamma_n}.
\end{aligned}$$

(6) (7)

Таким образом, используя рекуррентные соотношения (2), (4) и (5), можно найти решение СЛАУ с помощью $O(n)$ операций, что делает метод прогонки предпочтительным перед методом последовательного исключения, требующим $O(n^3)$ операций. Более того, если исходная матрица обладает свойством строгого диагонального преобладания, то метод прогонки является вычислительно устойчивым.

Пример (столбец-> строка; нумерация 0,1,2):

$$\begin{bmatrix} 3 & 1 & 0 \\ -1 & 4 & 2 \\ 0 & 2 & 3 \end{bmatrix} = \begin{bmatrix} 1 \\ -3 \\ 5 \end{bmatrix} \text{ Используем (4) и (5)}$$

$$\gamma_1 = \frac{-1}{3} \quad \gamma_2 = \frac{-2}{-1(\frac{-1}{3})+4} = -\frac{6}{13} \quad \beta_1 = \frac{1}{3}$$

$$\beta_2 = \frac{-3+1*\frac{1}{3}}{-1*(\frac{-1}{3})+4} = -\frac{8}{13} \text{ Найдем } x \text{ по (6) (7)}$$

$$x_2 = \frac{b_2 - a_{21}\beta_2}{a_{22} + a_{21}\gamma_2} = \frac{5 - 2*(-\frac{8}{13})}{3 + 2*(-\frac{6}{13})} = 3 \quad x_1 = \gamma_2 x_2 +$$

$$\beta_2 = -\frac{6}{13} * 3 - \frac{8}{13} \quad x_0 = -\frac{1}{3} * (-2) + \frac{1}{3} = 1$$

Тогда решение (1;-2;3)

Вопрос №15

1. Быстрое преобразование Фурье (постановка задачи). Сравнение алгоритмических сложностей дискретного преобразования Фурье и быстрого преобразования Фурье.

Быстрое преобразование Фурье предполагает уменьшение вычислительной сложности дискретного преобразования Фурье. Основная идея – быстрое вычисление суммы (когда высчитываем \widehat{a}^k)

Алгоритм Кули-Тьюки, его сложность $O(m \log_2 m)$, в то время как для ДПФ сложность $O(m^2)$.

2. Алгоритм Кули-Тьюки

Для начала запишем коэффициент \widehat{a}^k в следующей форме:

$$\widehat{a}_k = \frac{1}{2m} \sum_{j=0}^{2m-1} y_j e^{ik\pi} e^{-i \frac{kj}{m} \pi} = \frac{(-1)^k}{2m} \sum_{j=0}^{2m-1} y_j e^{-i \frac{kj}{m} \pi}.$$

Алгоритм Кули-Тьюки вычисляет значение суммы, которую мы обозначим как:

$$A_k = \sum_{j=0}^{2m-1} y_j e^{-i \frac{kj}{m} \pi}, \quad k = 0, \dots, 2m-1.$$

Разделим A_k на две части с четными и нечетными индексами соответственно:

$$\begin{aligned} A_k &= \sum_{j=0}^{m-1} y_{2j} e^{-i \frac{2j k \pi}{m}} + \sum_{j=0}^{m-1} y_{2j+1} e^{-i \frac{(2j+1) k \pi}{m}} \\ &= \sum_{j=0}^{m-1} y_{2j} e^{-i \frac{2j k \pi}{m}} + e^{-i \frac{k \pi}{m}} \sum_{j=0}^{m-1} y_{2j+1} e^{-i \frac{2j k \pi}{m}}. \end{aligned}$$

Обозначим полученные суммы E_k и O_k :

$$\begin{aligned} E_k &= \sum_{j=0}^{m-1} y_{2j} e^{-i \frac{2j k \pi}{m}}, \\ O_k &= \sum_{j=0}^{m-1} y_{2j+1} e^{-i \frac{2j k \pi}{m}}. \end{aligned}$$

Заметим, что E_k и O_k являются периодическими относительно k и периода m :

$$\begin{aligned} E_{k \pm m} &= \sum_{j=0}^{m-1} y_{2j} e^{-i \frac{2j(k \pm m) \pi}{m}} = e^{\mp 2j\pi} \sum_{j=0}^{m-1} y_{2j} e^{-i \frac{2j k \pi}{m}} = E_k, \\ O_{k \pm m} &= \sum_{j=0}^{m-1} y_{2j+1} e^{-i \frac{2j(k \pm m) \pi}{m}} = e^{\mp 2j\pi} \sum_{j=0}^{m-1} y_{2j+1} e^{-i \frac{2j k \pi}{m}} = O_k. \end{aligned}$$

Это свойство позволяет посчитать E_k и O_k только для $k = 0, \dots, m$, так как E_k и O_k для $k = m, \dots, 2m-1$ будут иметь те же значения. Тогда A_k для $k = 0, \dots, 2m$ высчитывается следующим образом:

$$A_k = E_k + e^{-i \frac{k \pi}{m}} O_k, \quad k = 0, \dots, m$$

$$A_{k+m} = E_k + e^{-i \frac{(k+m) \pi}{m}} O_k = E_k - e^{-i \frac{k \pi}{m}} O_k, \quad k = 0, \dots, m.$$

Так как E_k и O_k так же могут быть восприняты как БПФ на более грубых сетках, этот алгоритм рекурсивно применяется и к ним, что и замыкает в конечном итоге алгоритм Кули-Тьюки. Ясно, что в

Вопрос №23

1. Прямые и итерационные методы решения СЛАУ (определение и области использования). Метод Гаусса и его связь с LU-разложением.

Методы решения СЛАУ:

- Прямые
 - находят точное решение x
 - используются для плотных матриц малой размерности
- Итерационные
 - рассчитывают такое x_k , что при $k \rightarrow \infty \Rightarrow Ax_k - b \rightarrow 0$
 - используются для разреженных матриц большой размерности.

LU разложение:

$A = LU$, где L и U – нижняя и верхняя треугольные матрицы. Получение матрицы L соответствует прямому ходу метода Гаусса.

Метод Гаусса состоит в приведении матрицы A к треугольному виду с помощью элементарных преобразований, после чего решение x находится сравнительно легко.

$$\tilde{A} = [A, b] = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} & b_n \end{bmatrix}.$$

Для того, чтобы привести матрицу A к треугольному виду, необходимо последовательно обнулять элементы, находящиеся под главной диагональю.

$$\tilde{A} = \begin{bmatrix} a_{11} & a_{12} & \dots & \dots & a_{1n} & b_1 \\ 0 & a_{22}^{(1)} & \dots & \dots & a_{2n}^{(1)} & b_2^{(1)} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & a_{n-1,n-1}^{(n-2)} & a_{n-1,n}^{(n-2)} & b_n^{(n-2)} \\ 0 & 0 & \dots & 0 & a_{nn}^{(n-1)} & b_n^{(n-1)} \end{bmatrix}.$$

2. Вывод LU-разложения. LU-разложение с учетом перестановок.

В таком случае система линейных уравнений $Ax = b$ решается в два шага. Обозначив вектор y как $y = Ux$, мы находим его как решение уравнения: $Ly = b$. Вектор x тогда является решением уравнения: $Ux = y$. Каждый из шагов требует $O(n^2)$ операций. Таким образом, один раз разложив матрицу A в нижнюю и верхнюю треугольные матрицы, мы можем находить решение x для различных b в $O(n^2)$ шагов. LU-разложение связано с прямым ходом метода Гаусса и соответственно требует $O(n^3)$ операций.

Получим явные выражения для матриц L и U . Пусть перед k -й итерацией прямого хода метода Гаусса матрица A находится в форме $A^{(k)}$. Тогда k -я итерация эквивалентна домножению матрицы $A^{(k)}$ и вектора $b^{(k)}$ слева на матрицу $M^{(k)}$, имеющую вид:

Вопрос №9

1. Квадратурные формулы (определение). Метод вывода формул Ньютона-Котеса. Примеры формул Ньютона-Котеса. Вычислительная устойчивость/неустойчивость численного интегрирования.

Аппроксимацию интеграла называют квадратурой: $\int_a^b f(x) dx \approx \sum_{i=1}^n a_i f(x_i)$, $a_i \in \mathbb{R}$

Рассмотрим функцию $f(x) \in C^n[a; b]$ и n различных узлов x_1, \dots, x_n . Тогда, раскладывая $f(x)$ в базисные многочлены Лагранжа, мы получаем

$$f(x) = \sum_{i=1}^n f(x_i) l_i(x) + \frac{\prod_{i=1}^n (x - x_i)}{n!} f^{(n)}(\xi(x)),$$

, где $\xi(x) \in (a; b)$. Тогда интегрирование $f(x)$ на интервале $[a; b]$ дает:

$$\begin{aligned} \int_a^b f(x) dx &= \int_a^b \sum_{i=1}^n [f(x_i) l_i(x)] dx + \int_a^b \frac{\prod_{i=1}^n (x - x_i)}{n!} f^{(n)}(\xi(x)) dx \\ &= \sum_{i=1}^n \left[f(x_i) \int_a^b l_i(x) dx \right] + \int_a^b \frac{\prod_{i=1}^n (x - x_i)}{n!} f^{(n)}(\xi(x)) dx, \end{aligned}$$

что дает выражение для коэффициентов квадратуры a_i : $\int_a^b l_i(x) dx$. В случае, когда интерполяционные узлы x_1, \dots, x_n распределены равномерно, мы получаем формулы Ньютона-Котеса для численного интегрирования.

Формула трапеций ($n = 2$):

$$\int_a^b f(x) dx = \frac{h}{2} [f(x_1) + f(x_2)] - \frac{h^3}{12} f''(\xi), \quad \text{где}$$

$$x_1 = a, x_2 = b \text{ и } h = b - a.$$

Формула Симпсона ($n = 3$):

$$\int_a^b f(x) dx = \frac{h}{3} [f(x_1) + 4f(x_2) + f(x_3)] - \frac{h^5}{90} f^{(4)}(\xi),$$

$$\text{где } x_1 = a, x_2 = x_1 + h \text{ и } x_3 = x_1 + 2h = b, \quad h = \frac{b-a}{2}$$

По сравнению с дифференцированием, операция интегрирования способна к стабилизации вычислительной погрешности. Интуитивное объяснение этого эффект заключается в том, что интегрирование предполагает суммирование близких значений, в то время как дифференцирование вычисляет их разность. Численное интегрирование устойчиво с вычислительной точки зрения, и полная погрешность полностью соответствует остаточному члену вплоть до тех пор, пока E не достигнет машинного эпсилон. Полная погрешность интегрирования падает до тех пор, пока она не достигнет значения, сравнимого с машинным эпсилон, после чего уменьшение погрешности становится невозможным, и она стабилизируется на уровне машинного эпсилон.

2. Вывод составной формулы трапеций.

Рассмотрим интеграл $\int_a^b f(x) dx$. Разделим отрезок $[a; b]$ на четное число подотрезков n и применим формулу трапеций на каждом из них. Тогда для $f(x) \in C^2[a; b]$, $h = \frac{b-a}{n}$ и $x_i = a + (i-1)h$, где $i = 1, \dots, n+1$ мы имеем:

приведенном изложении алгоритма предполагается, что $m = 2^{\tilde{m}}, m \in \mathbb{N}$

$$M^{(k)} = \begin{bmatrix} 1 & 0 & \dots & \dots & \dots & \dots & 0 \\ 0 & \ddots & & & & & \vdots \\ & \ddots & \ddots & & & & \vdots \\ & & 0 & \ddots & & & \vdots \\ & & \vdots & \ddots & \ddots & & \vdots \\ & & \vdots & & -m_{k+1,k} & \ddots & \vdots \\ & & \vdots & & \vdots & \ddots & 0 \\ & & \vdots & & \vdots & & -m_{nk} & \ddots & \vdots \\ 0 & \dots & 0 & -m_{nk} & 0 & \dots & 0 & 1 \end{bmatrix}.$$

Верхняя треугольная матрица, получающаяся в результате обратного хода метода Гаусса, тогда выражается следующим образом: $A^{(n)} = M^{(n-1)} \dots M^{(1)} A$. В

контексте LU-разложения мы таким образом получаем матрицу $U = A^{(n)}$
 $U = M^{(n-1)} \dots M^{(1)} A$

Матрица L в таком случае:

$$\begin{aligned} A &= (M^{(n-1)} \dots M^{(1)})^{-1} U \\ \Rightarrow A &= (M^{(1)})^{-1} \dots (M^{(n-1)})^{-1} U \\ \Rightarrow L &= (M^{(1)})^{-1} \dots (M^{(n-1)})^{-1}. \end{aligned}$$

Явные выражения для обратных матриц могут быть получены, если мы заметим, что операция, обратная $(i) \rightarrow (i) - \frac{a_{ik}}{a_{kk}}(k), i = k+1, \dots, n$,

есть операция $(i) \rightarrow (i) + \frac{a_{ik}}{a_{kk}}(k), i = k+1, \dots, n$. Тогда обратная матрица $(M^{(k)})^{-1}$ имеет вид:

$$(M^{(k)})^{-1} = L^{(k)} = \begin{bmatrix} 1 & 0 & \dots & \dots & \dots & \dots & 0 \\ 0 & \ddots & & & & & \vdots \\ & \ddots & \ddots & & & & \vdots \\ & & 0 & \ddots & & & \vdots \\ & & \vdots & \ddots & \ddots & & \vdots \\ & & \vdots & & m_{k+1,k} & \ddots & \vdots \\ & & \vdots & & \vdots & \ddots & 0 \\ & & \vdots & & \vdots & & -m_{nk} & \ddots & \vdots \\ 0 & \dots & 0 & m_{nk} & 0 & \dots & 0 & 1 \end{bmatrix}$$

Легко убедиться, что перемножение таких матриц дает:

$$L = L^{(1)} \dots L^{(n-1)} = \begin{bmatrix} 1 & 0 & \dots & 0 \\ m_{21} & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ m_{n1} & \dots & m_{n,n-1} & 1 \end{bmatrix}.$$

В данном выводе LU-разложения мы не предполагали перестановок строк, которые зачастую являются необходимыми. Подобные перестановки можно учесть, используя матрицу перестановки. Матрица перестановки P образуется с помощью перестановки строк единичной матрицы E. Несложно проверить, что при умножении такой матрицы слева на матрицу A происходит та же перестановка строк, что использовалась при построении матрицы P из единичной матрицы. При умножении матрицы перестановки справа на матрицу A происходит перестановка столбцов. $PA = LU$ или $PAx = Pb$. (решение матрич-го ур-ия)

$\int_a^b f(x) dx = \sum_{i=1}^{n-1} \int_{x_i}^{x_{i+1}} f(x) dx =$
 $\frac{h}{2} \sum_{i=1}^{n-1} [f(x_i) + f(x_{i+1})] - \frac{h^3}{12} \sum_{i=1}^{n-1} f''(\xi_i)$
 где $\xi_i \in (x_i; x_{i+1})$. Заметим, что все узлы за исключением x_1 и x_{n+1} дважды повторяются в сумме, что позволяет упростить выражение:

$$\int_a^b f(x) dx = \frac{h}{2} [f(x_1) + 2 \sum_{i=2}^n f(x_i) + f(x_{n+1})] - \frac{h^3}{12} \sum_{i=2}^n f''(\xi_i)$$

Для упрощения выражения для остаточного члена заметим, что так как $f(x) \in C^2[a; b]$, мы имеем:

$$\begin{aligned} \min_{x \in [a; b]} f''(x) &\leq f''(\xi_i) \leq \max_{x \in [a; b]} f''(x) = > \\ n \min_{x \in [a; b]} f''(x) &\leq \sum_{i=2}^n f''(\xi_i) \leq \\ n \max_{x \in [a; b]} f''(x) &=> \\ \min_{x \in [a; b]} f''(x) &\leq \frac{1}{n} \sum_{i=2}^n f''(\xi_i) \leq \max_{x \in [a; b]} f''(x) \end{aligned}$$

$f''(\xi) = \frac{1}{n} \sum_{i=2}^n f''(\xi_i)$, где $\xi \in (a; b)$. Остаточный член в таком случае принимает следующую форму:

$$R = -\frac{h^3}{12} \sum_{i=2}^n f''(\xi_i) = -\frac{nh^3}{12} f''(\xi) = -\frac{(b-a)h^2}{12} f''(\xi).$$

$$\int_a^b f(x) dx = \frac{h}{2} [f(x_1) + 2 \sum_{i=2}^n f(x_i) + f(x_{n+1})] - \frac{(b-a)h^2}{12} f''(\xi)$$

3. Вывод формулы Симпсона

Тогда для $f(x) \in C^4[a; b]$, $h = \frac{b-a}{n}$ и $x_i = a + (i-1)h$, где $i = 1, \dots, n+1$ мы имеем:

$$\begin{aligned} \int_a^b f(x) dx &= \sum_{i=1}^{\frac{n}{2}} \int_{x_{2i-1}}^{x_{2i+1}} f(x) dx = \\ &= \frac{h}{3} \sum_{i=1}^{\frac{n}{2}} [f(x_{2i-1}) + 4f(x_{2i}) + f(x_{2i+1})] - \\ &= \frac{h^5}{90} \sum_{i=1}^{\frac{n}{2}} f^{(5)}(\xi_i) \end{aligned}$$

где $\xi_i \in (x_{2i-1}; x_{2i+1})$. Заметим, что все нечетные узлы за исключением x_1 и x_{n+1} дважды повторяются в сумме, что позволяет упростить выражение:

$$\int_a^b f(x) dx = \frac{h}{3} \left[f(x_1) + 2 \sum_{i=1}^{\frac{n}{2}-1} f(x_{2i+1}) + 4 \sum_{i=1}^{\frac{n}{2}} f(x_{2i}) + f(x_{n+1}) \right] - \frac{h^5}{90} \sum_{i=1}^{\frac{n}{2}} f^{(4)}(\xi_i)$$

Для упрощения выражения для остаточного члена заметим, что так как $f(x) \in C^4[a; b]$, мы имеем:

$$\begin{aligned} \min_{x \in [a; b]} f^{(4)}(x) &\leq f^{(4)}(\xi_i) \leq \max_{x \in [a; b]} f^{(4)}(x) = > \\ \frac{n}{2} \min_{x \in [a; b]} f^{(4)}(x) &\leq \sum_{i=1}^{\frac{n}{2}} f^{(4)}(\xi_i) \leq \\ \frac{n}{2} \max_{x \in [a; b]} f^{(4)}(x) &=> \end{aligned}$$

$$\begin{aligned} \min_{x \in [a; b]} f^{(4)}(x) &\leq \frac{2}{n} \sum_{i=1}^{\frac{n}{2}} f^{(4)}(\xi_i) \leq \\ \max_{x \in [a; b]} f^{(4)}(x) & \end{aligned}$$

$$f^{(4)}(\xi) = \frac{2}{n} \sum_{i=1}^{\frac{n}{2}} f^{(4)}(\xi_i), \text{ где } \xi \in (a; b).$$

$$R = -\frac{h^5}{90} \sum_{i=1}^{\frac{n}{2}} f^{(4)}(\xi_i) = -\frac{(b-a)h^4}{180} f^{(4)}(\xi).$$

Тогда составная формула Симпсона:

$$\int_a^b f(x) dx = \frac{h}{3} \left[f(x_1) + 2 \sum_{i=1}^{\frac{n}{2}-1} f(x_{2i+1}) + 4 \sum_{i=1}^{\frac{n}{2}} f(x_{2i}) + f(x_{n+1}) \right] - \frac{(b-a)h^4}{180} f^{(4)}(\xi)$$

Вопрос №24

1. Примеры векторных норм. Примеры матричных норм. Спектральный радиус матрицы

Норма вектора/матрицы – некое обобщение метрики

Примеры векторных норм:

$$x = [x_1, x_2, \dots, x_n] \in \mathbb{R}$$

- Среднеквадратичная векторная норма $\| \cdot \|_2$

$$\|x\|_2 = \left(\sum_{i=1}^n x_i^2 \right)^{\frac{1}{2}}$$

- Векторная супремум-норма $\| \cdot \|_\infty$
 $\|x\|_\infty = \max_{i \in [1, n]} |x_i|$

Геометрическим местом точек x , таких, что $\|x\|_2 = 1$ и $\|x\|_\infty = 1$, являются $(n-1)$ -мерная сфера и $(n-1)$ -мерный куб соответственно.

Пусть $\| \cdot \|$ – векторная норма в пространстве \mathbb{R}^n . Тогда следующий функционал является нормой матрицы:

$$\|A\| = \max_{\|x\|=1} \|Ax\|$$

и называется ИНДУЦИРОВАННОЙ НОРМОЙ.

- Среднеквадратичная матричная норма $\| \cdot \|_2$

$$\|A\|_2 = \max_{\|x\|_2=1} \|Ax\|_2 = \sqrt{\rho(A^T A)}$$

- Матричная супремум-норма $\| \cdot \|_\infty$

$$\|A\|_\infty = \max_{\|x\|_2=1} \|Ax\|_2 = \max_{i \in [1, n]} \sum_{j=1}^n a_{ij}$$

где $\rho(A^T A)$ – спектральный радиус матрицы $A^T A$.

Спектральным радиусом матрицы A называется число $\rho(A) \in \mathbb{R}$ такое, что:

$$\rho(A) = \max_{i \in [1, m]} |\lambda_i|, \text{ где } \lambda_i - \text{одно из } m \text{ собственных чисел матрицы } A.$$

2. Теорема о спектральном радиусе сходящихся матриц

Матрица $A \in \mathbb{R}^{n \times n}$ является сходящейся тогда и только тогда, когда $\rho(A) < 1$.

Докажем утверждение в обе стороны относительно схождения норм:

$$\lim_{k \rightarrow \infty} \|A^k\| = 0$$

В первую очередь докажем $\lim_{k \rightarrow \infty} \|A^k\| = 0 \Rightarrow \rho(A) < 1$. Для

начала заметим, что $\rho(A^k) = \rho^k(A)$. Действительно, пусть λ собственное число матрицы A , ассоциированное с собственным вектором x . Предположим, что A^k имеет собственное число $\tilde{\lambda}$, ассоциированное с тем же собственным вектором. Тогда имеем:

Вопрос №28

1. Методы простой итерации для решения СЛАУ (общая формулировка). Теорема о сходимости метода простой итерации (формулировка).

Итерационные методы ищут такое $x^{(k)}$, что $Ax^{(k)} - b \rightarrow 0$ при $k \rightarrow \infty$. Итерационные методы решения СЛАУ $Ax = b$ имеют форму: $x^{(k)} = Tx^{(k-1)} + c$, $k = 1, 2, \dots$ – номер итерации, где T и c появляются в результате разложения матрицы A . Начальное приближение $x^{(0)}$ предполагается случайным. Решением такого рекуррентного соотношения является неподвижная точка x : $x = Tx + c$. Матрица T должна обладать определенным свойством, чтобы итерации сходились к неподвижной точке, т.е. $x^{(k)} \rightarrow x$ при $k \rightarrow \infty$. Теорема о сходимости:

Последовательность $\{x^{(k)}\}_{k=0}^\infty$, сгенерированная итерацией $x^{(k)} = Tx^{(k-1)} + c$, сходится к единственному решению уравнения $x = Tx + c$, т.е. неподвижной точке x , для любого $x^{(0)} \in \mathbb{R}^n$ тогда и только тогда, когда $\rho(T) < 1$.

2. Теорема о сходимости метода простой итерации (доказательство)

В первую очередь докажем обратную часть теоремы, т.е. предположим, что $\rho(T) < 1$. Тогда, развернув все итерации, имеем:

$$\begin{aligned} x^{(k)} &= Tx^{(k-1)} + c \\ &= T(Tx^{(k-2)} + c) + c \\ &= T^2x^{(k-2)} + (T + E)c \\ &= \dots \\ &= T^kx^{(0)} + (T^{k-1} + \dots + T + E)c. \end{aligned}$$

Рассмотрим предел $\lim_{k \rightarrow \infty} x^{(k)}$. Тогда, воспользовавшись фактом $\rho(T) < 1$

$$\begin{aligned} x &= \lim_{k \rightarrow \infty} x^{(k)} = \lim_{k \rightarrow \infty} [T^kx^{(0)} + (T^{k-1} + \dots + T + E)c] \\ &= \lim_{k \rightarrow \infty} (T^{k-1} + \dots + T + E)c \\ &= (E - T)^{-1}c, \end{aligned}$$

из чего следует $x = Tx + c$. Теперь рассмотрим прямой случай, т.е. предположим, что $x = \lim_{k \rightarrow \infty} x^{(k)}$ есть единственное решение уравнения $x = Tx + c$. Так как $\rho(T) < 1 \Leftrightarrow \lim_{k \rightarrow \infty} T^k z = 0$ для любого z , докажем вторую часть утверждения. Пусть $x^{(0)} = x - z$. Заметим, что:

$$\begin{aligned} x - x^{(k)} &= Tx + c - (Tx^{(k-1)} + c) \\ &= T(x - x^{(k-1)}) \\ &= T^2(x - x^{(k-2)}) \\ &= \dots \\ &= T^k(x - x^{(0)}) \\ &= T^k z. \end{aligned}$$

Однако по условию теоремы $\lim_{k \rightarrow \infty} x - x^{(k)} = 0$

Вопрос №11

1. Степень точности квадратуры (определение). Разница между степенью и порядком точности квадратуры. Квадратуры с наибольшей степенью точности.

Степенью точности квадратуры называется такое наибольшее $n \in \mathbb{N}$, что формула квадратуры дает точный результат (остаточный член равен нулю) для всех $x_i, i = 0, \dots, n$.

Порядок точности – оценка зависимости остаточного члена от шага интегрирования h . Степень точности – оценка наибольшей степени полинома, который интегрируется квадратурой точно.

Задачу максимизации степени полинома, интегрирование которого с помощью данной квадратуры дает точный результат при заданном количестве узлов, решают квадратуры Гаусса, суть которых сводится к нахождению таких x_1, \dots, x_n и c_1, \dots, c_n , что приближение: $\int_a^b f(x) dx \approx \sum_{i=1}^n c_i f(x_i)$ максимизирует степень точности.

2. Теорема о корнях многочленов Лежандра (доказательство). Алгоритм построения квадратур Гаусса произвольной $2n-1$ степени точности.

Пусть x_1, \dots, x_n являются корнями полинома Лежандра n -ой степени $\varphi_n(x)$, и пусть коэффициенты c_1, \dots, c_n определены следующим образом:

$$c_i = \int_{-1}^1 l_i(x) dx = \int_{-1}^1 \prod_{j=1, j \neq i}^n \frac{x - x_j}{x_i - x_j} dx.$$

Тогда, если $P_m(x)$ является полиномом степени $m < 2n$, то верным является следующее равенство:

$$\int_{-1}^1 P_m(x) dx = \sum_{i=1}^n c_i P_m(x_i).$$

Для начала рассмотрим случай $m < n$. Тогда полином $P_m(x)$ может быть переписан в виде многочлена Лагранжа с нулевым остаточным членом, так как n -ая производная от $P_m(x)$ будет равна нулю.

$$P_m(x) = \sum_{i=1}^n P_m(x_i) l_i(x) = \sum_{i=1}^n P_m(x_i) \prod_{j=1, j \neq i}^n \frac{x - x_j}{x_i - x_j}.$$

В таком случае интеграл от $P_m(x)$ имеет вид:

$$\begin{aligned} \int_{-1}^1 P_m(x) dx &= \sum_{i=1}^n P_m(x_i) \left[\int_{-1}^1 \prod_{j=1, j \neq i}^n \frac{x - x_j}{x_i - x_j} dx \right] \\ &= \sum_{i=1}^n c_i P_m(x_i). \end{aligned}$$

Теперь рассмотрим случай $n \leq m < 2n$. Разделим многочлен $P_m(x)$ на многочлен Лежандра $\varphi_n(x)$. Тогда классическое деление многочленов столбиком дает:

$$\begin{aligned} \frac{P_m(x)}{\varphi_n(x)} &= Q(x) + \frac{R(x)}{\varphi_n(x)} \\ \Rightarrow P_m(x) &= Q(x) \varphi_n(x) + R(x), \end{aligned}$$

где $Q(x)$ – полином степени $m-n$ и $R(x)$ – полином степени обязательно меньшей, чем $\varphi_n(x)$. Тогда, учитывая, что $m < 2n$ и соответственно $m-n < n$ мы имеем: $\int_{-1}^1 Q(x) \varphi_n(x) dx = 0$

$$A^k x = A \cdots A A x = \tilde{\lambda} x$$

$$\Rightarrow A \cdots A \lambda x = \tilde{\lambda} x$$

$$\Rightarrow \lambda^k x = \tilde{\lambda} x$$

$$\Rightarrow \tilde{\lambda} = \lambda^k.$$

Очевидным следствием из этого является равенство $\rho(A^k) = \rho^k(A)$. Теперь, воспользовавшись теоремой ($\rho(A) \leq \|A\|$), получаем:

$$\rho(A^k) \leq \|A^k\|$$

$$\Rightarrow \rho^k(A) \leq \|A^k\|,$$

что приводит нас к утверждению: $\lim_{k \rightarrow \infty} \|A^k\| = 0 \Rightarrow \lim_{k \rightarrow \infty} \rho^k(A) = 0 \Leftrightarrow \rho(A) < 1$

Теперь докажем утверждение в обратную сторону: $\rho(A) < 1 \Rightarrow \lim_{k \rightarrow \infty} \|A^k\| = 0$. По теореме ($\rho(A) \leq \|A\| \leq \rho(A) + \epsilon$). мы можем найти такую норму $\|\cdot\|$ и $\epsilon > 0$, что:

$$\|A\| \leq \rho(A) + \epsilon < 1$$

Тогда пользуясь одной из аксиом матричных норм, имеем:

$$\|A^k\| \leq \|A\|^k \leq (\rho(A) + \epsilon)^k \Rightarrow \lim_{k \rightarrow \infty} \|A^k\| = 0.$$

$x^{(k)}$, из чего следует: $\lim_{k \rightarrow \infty} T^k z = 0$

Так как z был выбран произвольно, это эквивалентно утверждению $\rho(T) < 1$.

С другой стороны, так как степень многочлена $R(x)$ меньше n , интеграл от него попадает под первый случай этой теоремы, из чего следует: $\int_{-1}^1 R(x) dx = \sum_{i=1}^n c_i R(x_i)$

В таком случае интеграл от $P_m(x)$ принимает форму: $\int_{-1}^1 P_m(x) dx = \sum_{i=1}^n c_i R(x_i)$

Однако вследствие того, что $\varphi_n(x_i) = 0$, мы получаем $P_m(x_i) = R(x_i)$. Тогда финальная форма интеграла имеет вид: $\int_{-1}^1 P_m(x) dx = \sum_{i=1}^n c_i P_m(x_i)$. Интегрирование на произвольном интервале $[a; b]$ с помощью квадратуры Гаусса реализуется с помощью замены переменных: $x = \frac{1}{2}[(b-a)t + a + b]$, где $x \in [a; b]$ и $t \in [-1; 1]$. Тогда квадр-ра Гаусса вычисляется след. образом:

$$\int_a^b f(x) dx = \int_{-1}^1 f\left(\frac{(b-a)t + a + b}{2}\right) \frac{b-a}{2} dt.$$

Так как всего мы имеем $2n$ оптимизируемых параметров, логично предположить, что полином $2n-1$ степени, имеющий так же $2n$ параметров, может быть интегрирован точно при правильно подобранных параметрах. Очевидно, что если квадратура дает точный результат для любого полинома этой степени, то все полиномы низших степеней автоматически интегрируются точно как частные случаи. Рассмотрим случай $n = 2$ и интервал интегрирования $[-1; 1]$. Тогда квадратура принимает вид: $\int_a^b f(x) dx \approx c_1 f(x_1) + c_2 f(x_2)$. Мы ожидаем, что эта квадратура дает точный результат при интегрировании полинома третьей степени:

$$\begin{aligned} f(x) &= a_0 + a_1 x + a_2 x^2 + a_3 x^3 \\ \Rightarrow \int_{-1}^1 f(x) dx &= a_0 \int_{-1}^1 dx + a_1 \int_{-1}^1 x dx + a_2 \int_{-1}^1 x^2 dx + a_3 \int_{-1}^1 x^3 dx, \end{aligned}$$

где $a_0, a_1, a_2, a_3 \in \mathbb{R}$ – произвольные константы. Можно заметить, что квадратура будет точно вычислять интеграл этого полинома тогда, когда точно будут вычисляться интегралы от функций $f(x) = 1, f(x) = x, f(x) = x^2, f(x) = x^3$. В таком случае мы получаем систему уравнений:

$$\begin{cases} c_1 + c_2 = \int_{-1}^1 dx = 2, \\ c_1 x_1 + c_2 x_2 = \int_{-1}^1 x dx = 0, \\ c_1 x_1^2 + c_2 x_2^2 = \int_{-1}^1 x^2 dx = \frac{2}{3}, \\ c_1 x_1^3 + c_2 x_2^3 = \int_{-1}^1 x^3 dx = 0. \end{cases}$$

Из первых двух уравнений мы получаем:

$$c_1 = \frac{2x_2}{x_2 - x_1} \quad c_2 = \frac{-2x_1}{x_2 - x_1} \quad \text{что при подстановке}$$

в третье дает: $x_1 x_2 = -\frac{1}{3}$. Тогда четвертое уравнение становится: $x_1 + x_2 = 0$

что в результате дает следующее решение:

$$c_1 = 1,$$

$$c_2 = 1,$$

$$x_1 = -\frac{1}{\sqrt{3}},$$

$$x_2 = \frac{1}{\sqrt{3}}.$$

Таким образом квадратура Гаусса со степенью точности 3 имеет следующий вид: $\int_{-1}^1 f(x) dx \approx f\left(-\frac{1}{\sqrt{3}}\right) + f\left(\frac{1}{\sqrt{3}}\right)$

Подобным образом мы можем построить квадратуры Гаусса произвольной степени точности

Вопрос №29

1. Методы простой итерации для решения СЛАУ (общая формулировка). Метод Якоби (формулировка). Метод Гаусса–Зейделя (формулировка).

Итерационные методы ищут такое $x^{(k)}$, что $Ax^{(k)} - b \rightarrow 0$ при $k \rightarrow \infty$. Итерационные методы решения СЛАУ $Ax = b$ имеют форму: $x^{(k)} = Tx^{(k-1)} + c$, $k = 1, 2, \dots$ – номер итерации, где T и c появляются в результате разложения матрицы A . Начальное приближение $x^{(0)}$ предполагается случайным. Решением такого рекуррентного соотношения является неподвижная точка x : $x = Tx + c$. Матрица T должна обладать определенным свойством, чтобы итерации сходились к неподвижной точке, т.е. $x^{(k)} \rightarrow x$ при $k \rightarrow \infty$.

Метод Якоби строит итерационную процедуру, напрямую решая i -ое уравнение СЛАУ

$$x_i = \frac{1}{a_{ii}} \left(- \sum_{j=1, j \neq i}^n a_{ij} x_j + b_i \right)$$

и находя решение на шаге k из решения на шаге $k-1$:

$$x_i^{(k)} = \frac{1}{a_{ii}} \left(- \sum_{j=1, j \neq i}^n a_{ij} x_j^{(k-1)} + b_i \right)$$

В матричном виде метод Якоби предполагает разложение матрицы A на диагональную, нижнюю треугольную и верхнюю треугольную составляющие:

$$\begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{bmatrix} = \begin{bmatrix} a_{11} & 0 & \dots & 0 \\ 0 & a_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & a_{nn} \end{bmatrix} - \begin{bmatrix} 0 & \dots & 0 \\ -a_{21} & \dots & 0 \\ \vdots & \ddots & \vdots \\ -a_{n1} & \dots & -a_{nn-1} \end{bmatrix} + \begin{bmatrix} 0 & -a_{12} & \dots & -a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & -a_{nn} \end{bmatrix}$$

или кратко: $A = D - L - U$. Подстановка в СЛАУ дает: $(D - L - U)x = b$ или $Dx = (L + U)x + b$. При этом, если D имеет обратную (т.е. все диагональные элементы ненулевые), то можно записать: $x = D^{-1}(L + U)x + D^{-1}b$, что легко превращается в следующее рекуррентное соотношение: $x^{(k)} = D^{-1}(L + U)x^{(k-1)} + D^{-1}b$ из которого мы получаем матрицу T и вектор c в контексте метода простой итерации: $T_j = D^{-1}(L + U)$ $c_j = D^{-1}b$ методу Гаусса–Зейделя:

$$x_i^{(k)} = \frac{1}{a_{ii}} \left(- \sum_{j=1}^{i-1} a_{ij} x_j^{(k)} - \sum_{j=i+1}^n a_{ij} x_j^{(k-1)} + b_i \right),$$

В матричном виде метод Гаусса–Зейделя можно записать в следующем виде: $(D - L)x^{(k)} = Ux^{(k-1)} + b \Rightarrow x^{(k)} = (D - L)^{-1}Ux^{(k-1)} + (D - L)^{-1}b$, где матрица $(D - L)^{-1}$ существует тогда и только тогда, когда все

Вопрос №38

1. Приближение, интерполяция и экстраполяция. Аппроксимационная теорема Вейерштрасса (формулировка)

Задача приближения или аппроксимации (эти термины синонимичны) состоит в представлении недоступной функции $f(x)$ в виде более простой аналитической функции по n значениям $f(x_i)$, где $x_i \in [a; b]$ называются узлами. Для решения этой задачи мы задаемся некоторой параметризованной формой для $\tilde{f}(x; c)$, где c – вектор параметров, а затем подбираем c так, что отклонение $\rho[f(x) - \tilde{f}(x; c)]$ минимизировано, где ρ – некоторая метрика, определяющая оценку погрешности приближения. Частными случаями аппроксимации являются интерполяция и экстраполяция.

Интерполяцией называется приближение, при котором требуется, чтобы $\tilde{f}(x)$ проходила через заданные узлы $(x_i, f(x_i))$ внутри отрезка $x \in [a; b]$.

Экстраполяцией называется приближение, при котором требуется, чтобы $\tilde{f}(x)$ по заданным узлы $(x_i, f(x_i))$ предсказывала значение $f(x)$ вне отрезка $[a; b]$. Пусть $f(x) \in C[a; b]$. Тогда для любого $\epsilon > 0$ \exists такой вещественный многочлен $P(x)$, что для $x \in [a; b]$ $\|f(x) - P(x)\|_\infty < \epsilon$. Следствием вышеописанной теоремы является аналогичная для тригонометрических полиномов с периодом равным интервалу.

2. Теорема о полиноме с наименьшими по модулю экстремумами (доказательство). Решение задачи о минимизации ошибки интерполяции Лагранжа.

Пусть $\tilde{T}_k(x) = \frac{1}{2^{k-1}} T_k(x)$, где $k \geq 1$ и $T_k(x)$ является полиномом Чебышева k -й степени, и пусть \tilde{P}_k – множество всех нормированных полиномов степени k . Тогда верным является следующее утверждение:

$$\frac{1}{2^{k-1}} = \max_{x \in [-1; 1]} |\tilde{T}_k(x)| \leq \max_{x \in [-1; 1]} |P_k(x)| \quad \forall P_k(x) \in \tilde{P}_k.$$

Более того, равенство верно только при $P_k(x) = \tilde{T}_k(x)$.

Рассмотрим доказательство от обратного. Пусть верно

$$\max_{x \in [-1; 1]} |P_k(x)| \leq \max_{x \in [-1; 1]} |\tilde{T}_k(x)| =$$

$\frac{1}{2^{k-1}}$ (1) Введем полином $(k-1)$ -й

Вопрос №30

1. Число обусловленности (определение). Связь числа обусловленности с устойчивостью решения СЛАУ к малым изменениям. Связь числа обусловленности с погрешностью округления в контексте арифметики с t значащими цифрами.

Соотношение между относительной погрешностью вычислений $\frac{\|x - \tilde{x}\|}{\|x\|}$ и нормой вектора невязки $\|r\|$ регулируется числом обусловленности.

Теорема (о связи между относительной погрешностью и вектором невязки) Пусть $\tilde{x} \in \mathbb{R}$ является приближением к точному решению СЛАУ $Ax = b$, A – невырожденная матрица и r – вектор невязки. Тогда для любой индуцированной матричной нормы верно: $\|x - \tilde{x}\| \leq \|r\| * \|A^{-1}\|$ и при $x, b \neq 0$: $\frac{\|x - \tilde{x}\|}{\|x\|} \leq K(A) \frac{\|r\|}{\|b\|}$, где $K(A) = \|A\| \cdot \|A^{-1}\|$ называется числом обусловленности.

Устойчивость решения СЛАУ к малым изменениям: Предположим, что приближенное решение \tilde{x} является точным решением модифицированной СЛАУ, полученной с помощью малых возмущений матрицы A :

$$(A + \delta A)\tilde{x} = b \Rightarrow r = b - A\tilde{x} = \delta A\tilde{x}.$$

После подстановки вектора невязки во второе неравенство теоремы:

$$\frac{\|x - \tilde{x}\|}{\|x\|} \leq K(A) \frac{\|\delta A\tilde{x}\|}{\|b\|} \leq K(A) \|\delta A\| \frac{\|\tilde{x}\|}{\|b\|}.$$

Можно заметить, что относительная погрешность вычислений будет тем меньше, чем меньше будет число обусловленности $K(A)$ (при этом легко увидеть, что $K(A) \geq 1$). Слишком большое число обусловленности, в свою очередь, приводит к дестабилизации вычислительных погрешностей (чем больше $K(A)$, тем сильнее влияние погрешностей округления δA на относительную погрешность). Оценим число обусловленности без явного вычисления обратной матрицы, что само по себе было бы источником погрешностей, и найдем связь между числом обусловленности и точностью полученного решения в контексте арифметики с t значащими цифрами.

Можно доказать, что если приближенное решение \tilde{x} было получено с помощью метода Гаусса и арифметики с t значащими цифрами, то верно следующее приближение для нормы вектора невязки: $\|r\| \approx 10^{-t} \|A\| \cdot \|\tilde{x}\|$. Приближение для числа обусловленности теперь можно получить, рассмотрев следующую СЛАУ: $Ay = r$. Т.к. предполагалось, что решение \tilde{x} было получено с помощью метода Гаусса, в наличии гарантированно имеются множители, необходимые для LU-разложения матрицы A , что означает, что решение y

диагональные элементы являются ненулевыми. $T_G = (D - L)^{-1}U$ $c_G = (D - L)^{-1}b$ Методы Якоби и Гаусса–Зейделя являются безусловно сходящимися для матриц со строго диагональным преобладанием.

2. Оценки погрешности методов простой итерации (вывод двух оценок).

$x - x^{(k)} = T^k(x - x^{(0)})$ и оценим его

$$\|x - x^{(k)}\| = \|T^k(x - x^{(0)})\|$$

$$\Rightarrow \|x - x^{(k)}\| \leq \|T^k\| \cdot \|x - x^{(0)}\|,$$

из чего следует оценка погрешности метода: $\|x - x^{(k)}\| \leq \|T\|^k \|x - x^{(0)}\|$ Для вывода второй оценки рассмотрим следующее выражение:

$$x^{(k+1)} - x^{(k)} = T x^{(k)} + c - T x^{(k-1)} - c$$

$$= T(x^{(k)} - x^{(k-1)})$$

$$= \dots$$

$$= T^k(x^{(1)} - x^{(0)}).$$

Тогда для $m > k \geq 1$ справедливо следующее:

$$\|x^{(m)} - x^{(k)}\| = \|x^{(m)} - x^{(m-1)} + x^{(m-1)} - \dots + x^{(k+1)} - x^{(k)}\|$$

$$\leq \|x^{(m)} - x^{(m-1)}\| + \|x^{(m-1)} - x^{(m-2)}\| + \dots + \|x^{(k+1)} - x^{(k)}\|$$

$$\leq (\|T\|^{m-1} + \|T\|^{m-2} + \dots + \|T\|^k) \|x^{(1)} - x^{(0)}\|$$

$$= \|T\|^k (1 + \|T\| + \|T\|^2 + \dots + \|T\|^{m-k-1}) \|x^{(1)} - x^{(0)}\|.$$

В пределе $m \rightarrow \infty$ мы имеем:

$$\|x - x^{(k)}\| = \lim_{m \rightarrow \infty} \|x^{(m)} - x^{(k)}\|$$

$$= \|T\|^k \sum_{j=0}^{\infty} \|T\|^j \cdot \|x^{(1)} - x^{(0)}\|.$$

Сумма представляет собой сумму геометрической прогрессии, и так как $\rho(T) < 1 \Rightarrow \|T\| < 1$ мы имеем:

$$\|x - x^{(k)}\| = \frac{\|T\|^k}{1 - \|T\|} \|x^{(1)} - x^{(0)}\|.$$

Таким образом скорость сходимости методов простой итерации зависит от $\|T\|$, и вследствие $\rho(T) \approx \|T\|$ она, в сущности, зависит от спектрального радиуса матрицы T $\|x - x^{(k)}\| \approx \rho(T)^k \|x - x^{(0)}\|$.

Очевидно, что мы хотели бы найти итерационный метод с минимально возможным $\rho(T)$.

степени $Q(x) = \tilde{T}_k(x) - P_k(x)$. Для $k + 1$ точки \hat{x}_m мы имеем:

$$Q(\hat{x}_m) = \frac{(-1)^m}{2^{k-1}} - P_k(\hat{x}_m).$$

Однако из (1) следует $|P_k(\hat{x}_m)| \leq \frac{1}{2^{k-1}}$. Тогда мы получаем:

$$Q(\hat{x}_m) \geq 0 \quad \text{для } m = 0, 2, \dots$$

$$Q(\hat{x}_m) \leq 0 \quad \text{для } m = 1, 3, \dots$$

Таким образом, вследствие непрерывности $Q(x)$ полином имеет корень в каждом подотрезке $[x_m; x_{m+1}]$, где $m=0, \dots, k-1$, что суммарно дает k различных корней. Однако $Q(x)$ является полиномом $(k-1)$ -й степени. Это возможно только в случае $Q(x) = 0$, что дает $P_k(x) = \tilde{T}_k(x)$.

$$f(x) - L_{n-1}(x) = \frac{f^{(n)}(\xi)}{n!} \prod_{i=1}^n (x - x_i),$$

где $\xi \in (-1; 1)$ и x_1, \dots, x_n – узлы интерполяции. Оптимальная интерполяция соответствует случаю, при котором значение остаточного члена минимизировано. Так как мы не имеем контроля над $f(x)$, минимизация возможна только относительно значений интерполяционных узлов. Таким образом, мы получаем оптимизационную задачу следующего вида:

$$\min_{x_1, \dots, x_n} \prod_{i=1}^n (x - x_i),$$

при условии $x_i \in [-1; 1], \quad i = 1, \dots, n$.

Целевая функция является нормированным полиномом степени n с n различными корнями, равными x_1, \dots, x_n . По теореме ее максимальное значение на $[-1; 1]$ минимально тогда, когда $x_i = \bar{x}_i = \cos\left(\frac{2i-1}{2n}\pi\right), \quad i = 1, \dots, n$.

Более того, теорема позволяет оценить остаточный член подобной оптимальной интерполяции:

$$\max |f(x) - L_{n-1}(x)| \leq \frac{1}{2^n n!} \max |f^{(n)}(x)|,$$

для любой функции $f(x) \in C^n[-1; 1]$. Обобщение до произвольного отрезка $[a; b]$ реализуется с помощью непрерывной замены переменной

$$\tilde{x} = \frac{1}{2} [(b-a)x + a + b], \quad \text{где } \tilde{x} \in [a; b]$$

и $x \in [-1; 1]$. С помощью подобной замены значения \tilde{x}_i пересчитываются на отрезке $[a; b]$.

Интерполяция Лагранжа для любой достаточно гладкой функции $f(x)$ является оптимальной тогда, когда узлы интерполяции распределены в соответствии с чебышевскими узлами.

может быть быстро найдено. Приближение \tilde{y} удовлетворяет:

$$\tilde{y} \approx A^{-1}r = A^{-1}(b - A\tilde{x}) = A^{-1}b - \tilde{x} = x - \tilde{x}.$$

Таким образом вектор \tilde{y} является оценкой вычислительной погрешности, возникшей при нахождении решения \tilde{x} . Оценим его норму:

$$\|\tilde{y}\| \approx \|A^{-1}r\| \leq \|A^{-1}\| \cdot \|r\| \approx \|A^{-1}\| \cdot 10^{-t} \|A\| \cdot \|\tilde{x}\| = 10^{-t} K(A) \cdot \|\tilde{x}\|$$

$$\Rightarrow \frac{\|x - \tilde{x}\|}{\|\tilde{x}\|} \approx 10^{-t} K(A)$$

2. Оценка относительной погрешности на основе числа обусловленности. Итерационное уточнение приближенного решения.

Если число обусловленности имеет порядок $O(1)$, то вычислительная погрешность приближенного решения \tilde{x} имеет оценочный порядок $O(10^{-t})$, в том время как при $K(A) \propto O(10^t)$ ее оценочный порядок $O(1)$. Иными словами, число обусловленности позволяет оценить, сколько значимых цифр теряется в процессе вычислений.

$$K(A) \approx \frac{\|\tilde{y}\|}{\|\tilde{x}\|} 10^t \approx \frac{\|x - \tilde{x}\|}{\|\tilde{x}\|} 10^t$$

$$\Rightarrow \frac{\|x - \tilde{x}\|}{\|\tilde{x}\|} \approx 10^{-t} K(A). \quad (1)$$

На основании этого результата может быть построена техника итерационного уточнения решения \tilde{x} . Действительно, если $\tilde{y} \approx x - \tilde{x}$ то логично предположить, что для достаточно обусловленных матриц (а именно $K(A) < 10^t$) вектор $\tilde{x} + \tilde{y}$ будет давать более точное решение. Обозначив $\tilde{x}^{(1)} = \tilde{x}^{(0)} + \tilde{y}^{(0)}$ и продолжив итерацию уже относительно нового приближенного значения $\tilde{x}^{(1)}$, мы получаем метода итерационного уточнения на основе метода Гаусса. Из (1) очевидно, что при $K(A) \propto 10^q$ после k итераций ожидается $k(t - q)$ корректных значащих цифр в решении. Для хорошо обусловленных матриц будет достаточно одной-двух итераций.

Вопрос №7

1. Глобальная и локальная интерполяция. Кусочная интерполяция. Интерполяция кубическими сплайнами (определение без вывода разрешающих уравнений).

Глобальная интерполяция – интерполяция одной аппроксимирующей функцией по всему отрезку $[a; b]$.

Локальная интерполяция, где мы делим отрезок $[a; b]$ на маленькие подотрезки и используем интерполяцию полиномом малой степени на каждом из этих подотрезков, после чего “склеиваем” полученное множество полиномов в единую функцию, заданную на всем отрезке $[a; b]$. Такую интерполяцию называют кусочной интерполяцией. Пусть функция $f(x)$ задана в n интерполяционных узлах $a = x_1, x_2, \dots, x_n = b$ на отрезке $[a; b]$. Тогда кубическим сплайном для функции $f(x)$ называется функция $S(x)$, для которой верно:

- $S(x)$ кусочно задана кубическими многочленами $S_i(x)$ на каждом отрезке $[x_i; x_{i+1}]$ $i = 1 \dots n - 1$;
 - $S_i(x_i) = f(x_i)$ и $S_i(x_{i+1}) = f(x_{i+1})$, $i = 1 \dots n - 1$;
 - значения смежных многочленов совпадают в общих узлах: $S_i(x_{i+1}) = S_{i+1}(x_{i+1})$, $i = 1, \dots, n - 2$;
 - значения первых производных смежных многочленов совпадают в общих узлах: $S'_i(x_{i+1}) = S'_{i+1}(x_{i+1})$, $i = 1, \dots, n - 2$;
 - значения вторых производных смежных многочленов совпадают в общих узлах: $S''_i(x_{i+1}) = S''_{i+1}(x_{i+1})$, $i = 1, \dots, n - 2$;
 - заданы граничные условия:
 - естественные граничные условия: $S(x_1) = S(x_n) = 0$;
 - граничные условия на касательную: $S(x_1) = f'(x_1)$ и $S(x_n) = f'(x_n)$
2. Вывод разрешающих уравнений кубического сплайна.

Так как кубический многочлен задается 4 константами, для задания кубического сплайна нам необходимо определить $4(n-1)$ констант. Сделаем это в общем виде. Запишем кубический многочлен $S_i(x)$ на отрезке $[x_i; x_{i+1}]$ в форме: $S_i(x) = a_i + b_i(x - x_i) + c_i(x - x_i)^2 + d_i(x - x_i)^3$ что

Вопрос №26

1. Сингулярные числа и вектора (определение). Теорема о сингулярном разложении (формулировка).

Сингулярными числами $\sigma_1, \dots, \sigma_r$ матрицы $A \in \mathbb{R}^{m \times n}$ называются неотрицательные вещественные числа $\sigma_i = \sqrt{\lambda_i}$, где λ_i ненулевые собственные числа соответствующей матрицы Грама $K = A^T A$. Ассоциированные собственные вектора матрицы Грама K называются сингулярными векторами.

Матрица $A \in \mathbb{R}^{m \times n}$ ранга $r > 0$ может быть представлена в виде разложения $A = P \Sigma Q^T$, где $P \in \mathbb{R}^{m \times r}$ и $Q \in \mathbb{R}^{n \times r}$ – ортонормальные матрицы ($P^T P = Q^T Q = E$), а $\Sigma \in \mathbb{R}^{m \times r}$ – диагональная матрица, на диагонали которой расположены сингулярными числами $\sigma_1, \dots, \sigma_r$.

2. Теорема о решении задачи метода наименьших квадратов для СЛАУ

$\min_{a_0, a_1} E_2(a_0, a_1) = \min_{a_0, a_1} \sum_{i=1}^n [y_i - f(x_i)]^2$ Подстановка функции $f(x) = a_0 + a_1 x$ в $E_2(a_0, a_1)$ дает: $E_2(a_0, a_1) = \sum_{i=1}^n [y_i - a_0 - a_1 x_i]^2$ Функция $E_2(a_0, a_1)$ принимает экстремальное значение при таких a_0, a_1 , что:

$$\frac{\partial E_2}{\partial a_0} = 0 \implies -2 \sum_{i=1}^n (y_i - a_0 - a_1 x_i) = 0,$$

$$\frac{\partial E_2}{\partial a_1} = 0 \implies -2 \sum_{i=1}^n x_i (y_i - a_0 - a_1 x_i) = 0,$$

что дает систему уравнений относительно a_0 и a_1 :

$$\begin{cases} \sum_{i=1}^n y_i - n a_0 - a_1 \sum_{i=1}^n x_i = 0, \\ \sum_{i=1}^n (x_i y_i) - a_0 \sum_{i=1}^n x_i - a_1 \sum_{i=1}^n x_i^2 = 0. \end{cases}$$

Решением системы являются следующие выражения для a_0 и a_1 :

$$a_0^{(opt)} = \frac{\sum_{i=1}^n x_i^2 \sum_{i=1}^n y_i - \sum_{i=1}^n x_i \sum_{i=1}^n (x_i y_i)}{n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2},$$

$$a_1^{(opt)} = \frac{n \sum_{i=1}^n (x_i y_i) - \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2}.$$

Тогда прямая $f(x) = a_0^{(opt)} + a_1^{(opt)} x$ наилучшим образом приближается к дискретным данным $\{(x_i, y_i)\}_{i=1}^n = 1$ в среднеквадратичном смысле.

Вопрос №4

1. Интерполяция Лагранжа. Интерполяционный многочлен Лагранжа. Интерполяция Эрмита.

Интерполяция Лагранжа – глобальная интерполяция, предполагающая интерполирование функции по n узлам полиномом $n - 1$ степени.

Пусть функция $f(x)$ задана в n интерполяционных узлах $a = x_1, x_2, \dots, x_n = b$ на отрезке $[a; b]$. Тогда интерполяционным многочленом для функции $f(x)$ и соответствующих интерполяционных узлов называется многочлен следующего вида:

$$L_{n-1}(x) = \sum_{i=1}^n f(x_i) l_i(x) = \sum_{i=1}^n f(x_i) \prod_{j \neq i} \frac{x - x_j}{x_i - x_j},$$

где $l_i = \prod_{j \neq i} \frac{x - x_j}{x_i - x_j}$ – базисный многочлен Лагранжа $n - 1$ степени.

Интерполяция Эрмита – глобальная интерполяция предполагающая согласования не только значений функции и интерполяционного полинома в узлах, но и их производных (многочлены Эрмита предполагают согласование ПЕРВОЙ производной).

Пусть $x_1, \dots, x_n \in [a; b]$ – интерполяционные узлы и $f(x) \in C^1[a; b]$. Тогда единственный многочлен наименьшей степени согласующийся с $f(x_i)$ и $f'(x_i)$, $i = 1, \dots, n$ является многочленом Эрмита степени (максимум) $2n - 1$, заданный выражением

$$H_{2n-1}(x) = \sum_{i=1}^n f(x_i) h_i(x) + \sum_{i=1}^n f'(x_i) \hat{h}_i(x),$$

где $h_i(x)$ и $\hat{h}_i(x)$ заданы как

$$h_i(x) = [1 - 2(x - x_i) l'_i(x_i)] l_i^2(x),$$

$$\hat{h}_i(x) = (x - x_i) l_i^2(x),$$

где $l_i(x)$ – базисные полиномы Лагранжа $n-1$ степени.

2. Теорема об оценке остаточного члена многочлена Лагранжа (доказательство)

Пусть $x_1, \dots, x_n \in [a; b]$ – интерполяционные узлы и $f(x) \in C^n[a; b]$. Тогда $\forall x \in [a; b] \exists \xi \in (a; b)$ такое, что

$$f(x) - L_{n-1}(x) = \frac{f^{(n)}(\xi)}{n!} \prod_{i=1}^n (x - x_i)$$

Случай $x = x_i$ тривиален, так что рассмотрим $x \neq x_i$, $i = 1, \dots, n$. Введем функцию $g(t)$:

$$g(t) = f(t) - L_{n-1}(t) - [f(x) - L_{n-1}(x)] \prod_{i=1}^n \frac{t - x_i}{x - x_i}.$$

Несложно проверить, что $g(x_i) = 0$, $i = 1, \dots, n$ и $g(x) = 0$. Таким образом, функция $g(t)$ имеет $n + 1$ корней. Теорема Ролля гласит, что в этом случае $g'(t)$ имеет как минимум n корней. Обобщая теорему Ролля на производные высшего порядка, получаем, что $g^{(n)}(t)$ имеет как минимум один корень в точке $\xi \in (a; b)$. Тогда, учитывая, что $L_n(t)$ является полиномом $n-1$ степени, имеем:

автоматически дает $S_i(x_i) = a_i = f(x_i)$. Тогда из условия равенства значений смежных многочленов в общих узлах имеем:

$a_{i+1} = S_{i+1}(x_{i+1}) = S_i(x_{i+1}) = a_i + b_i h_i + c_i h_i^2 + d_i h_i^3$, где $h_i = x_{i+1} - x_i$. Так как $S'_i(x_i) = b_i$, из условия равенства значений первых производных в общих узлах смежных многочленов получаем $b_{i+1} = b_i + 2c_i h_i + 3d_i h_i^2$. Наконец, из условия для второй производной имеем:

$c_{i+1} = c_i + 3d_i h_i$ где $c_i = \frac{S''_i(x_i)}{2}$. Из последнего уравнения выразим d_i через c_i : $d_i = \frac{c_{i+1} - c_i}{3h_i}$

Подставив его в выражение для a_{i+1} , имеем: $a_{i+1} = a_i + b_i h_i + \frac{h_i^2}{3}(c_{i+1} + 2c_i)$ (1)

Аналогично получаем для b_{i+1} : $b_{i+1} = b_i + h_i(c_{i+1} + c_i)$ (2)

Чтобы получить замыкание для c_i выраженное через $a_i = f(x_i)$, мы подставляем b_i и b_{i+1} , выраженное из (1)

$$b_i = \frac{1}{h_i}(a_{i+1} - a_i) - \frac{h_i}{3}(c_{i+1} + 2c_i),$$

$$\Rightarrow b_{i-1} = \frac{1}{h_{i-1}}(a_i - a_{i-1}) - \frac{h_{i-1}}{3}(c_i + 2c_{i-1}),$$

в уравнение (2), записанное для b_i , что дает

$$h_{i-1}c_{i-1} + 2(h_i + h_{i-1})c_i + h_i c_i + 1 = \frac{3}{h_i}(a_{i+1} - a_i) -$$

$\frac{3}{h_{i-1}}(a_i - a_{i-1})$ Теперь мы можем составить систему уравнений относительно c_i , решив которую можно вычислить недостающие коэффициенты d_i и b_i по формулам

$$g^{(n)}(\xi) = 0 = f^{(n)}(\xi) - L_{n-1}^{(n)}(\xi) - [f(x) - L_{n-1}(x)] \frac{d^n}{dt^n} \prod_{i=1}^n \left[\frac{t - x_i}{x - x_i} \right]_{t=\xi}$$

$$= f^{(n)}(t) - [f(x) - L_{n-1}(x)] \frac{d^n}{dt^n} \prod_{i=1}^n \left[\frac{t - x_i}{x - x_i} \right]_{t=\xi}.$$

Заметим, что

$$\prod_{i=1}^n \frac{t - x_i}{x - x_i} = t^n \prod_{i=1}^n \frac{1}{x - x_i} + O(t^{n-1}),$$

и тогда мы получаем

$$0 = f^{(n)}(\xi) - [f(x) - L_{n-1}(x)] \frac{n!}{\prod_{i=1}^n (x - x_i)},$$

из чего следуем искомое

$$f(x) - L_{n-1}(x) = \frac{f^{(n)}(\xi)}{n!} \prod_{i=1}^n (x - x_i).$$

Вопрос №39

1. Методы простой итерации для решения СЛАУ (общая формулировка). Метод Якоби (формулировка). Метод Гаусса–Зейделя (формулировка)

Итерационные методы ищут такое $x^{(k)}$, что $Ax^{(k)} - b \rightarrow 0$ при $k \rightarrow \infty$. Итерационные методы решения СЛАУ $Ax = b$ имеют форму: $x^{(k)} = Tx^{(k-1)} + c$, $k = 1, 2, \dots$ – номер итерации, где T и c появляются в результате разложения матрицы A . Начальное приближение $x^{(0)}$ предполагается случайным. Решением такого рекуррентного соотношения является неподвижная точка x : $x = Tx + c$. Матрица T должна обладать определенным свойством, чтобы итерации сходились к неподвижной точке, т.е. $x^{(k)} \rightarrow x$ при $k \rightarrow \infty$.

Метод Якоби строит итерационную процедуру, напрямую решая i -ое уравнение СЛАУ

$$x_i = \frac{1}{a_{ii}} \left(- \sum_{j=1, j \neq i}^n a_{ij} x_j + b_i \right)$$

и находя решение на шаге k из решения на шаге $k-1$:

$$x_i^{(k)} = \frac{1}{a_{ii}} \left(- \sum_{j=1, j \neq i}^n a_{ij} x_j^{(k-1)} + b_i \right)$$

В матричном виде метод Якоби предполагает разложение матрицы A на диагональную, нижнюю треугольную и верхнюю треугольную составляющие:

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} = \begin{bmatrix} a_{11} & 0 & \dots & 0 \\ 0 & a_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_{nn} \end{bmatrix} + \begin{bmatrix} 0 & a_{12} & \dots & a_{1n} \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & \dots & 0 \\ a_{21} & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & 0 \end{bmatrix}$$

или кратко: $A = D - L - U$. Подстановка в СЛАУ дает: $(D - L - U)x = b$ или $Dx = (L + U)x + b$. При этом, если D имеет обратную (т.е. все диагональные элементы ненулевые), то можно записать: $x = D^{-1}(L + U)x + D^{-1}b$, что легко превращается в следующее рекуррентное соотношение: $x^{(k)} = D^{-1}(L + U)x^{(k-1)} + D^{-1}b$ из которого мы получаем матрицу T и вектор c в контексте метода простой итерации: $T_j = D^{-1}(L + U)$ $c_j = D^{-1}b$ методу Гаусса–Зейделя:

$$x_i^{(k)} = \frac{1}{a_{ii}} \left(- \sum_{j=1}^{i-1} a_{ij} x_j^{(k)} - \sum_{j=i+1}^n a_{ij} x_j^{(k-1)} + b_i \right),$$

В матричном виде метод Гаусса–Зейделя можно записать в следующем виде: $(D - L)x^{(k)} = Ux^{(k-1)} + b \Rightarrow x^{(k)} = (D - L)^{-1}Ux^{(k-1)} + (D - L)^{-1}b$, где матрица $(D - L)^{-1}$ существует тогда и только тогда, когда все диагональные элементы являются ненулевыми. $T_G = (D - L)^{-1}U$ $c_G =$

Вопрос №40

1. Степень точности квадратуры (определение). Разница между степенью и порядком точности квадратуры. Квадратуры с наибольшей степенью точности

Степенью точности квадратуры называется такое наибольшее $n \in \mathbb{N}$, что формула квадратуры дает точный результат (остаточный член равен нулю) для всех x_i , $i = 0, \dots, n$. Порядок точности – оценка зависимости остаточного члена от шага интегрирования h . Степень точности – оценка наибольшей степени полинома, который интегрируется квадратурой точно. Задачу максимизации степени полинома, интегрирование которого с помощью данной квадратуры дает точный результат при заданном количестве узлов, решают квадратуры Гаусса, суть которых сводится к нахождению таких x_1, \dots, x_n и c_1, \dots, c_n , что приближение: $\int_a^b f(x) dx \approx \sum_{i=1}^n c_i f(x_i)$ максимизирует степень точности.

2. Вывод квадратуры Гаусса со степенью точности 3.

Рассмотрим случай $n = 2$ и интервал интегрирования $[-1; 1]$. Тогда квадратура принимает вид: $\int_a^b f(x) dx \approx c_1 f(x_1) + c_2 f(x_2)$ Мы ожидаем, что эта квадратура дает точный результат при интегрировании полинома третьей степени:

$$f(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3 \\ \Rightarrow \int_{-1}^1 f(x) dx = a_0 \int_{-1}^1 dx + a_1 \int_{-1}^1 x dx + a_2 \int_{-1}^1 x^2 dx + a_3 \int_{-1}^1 x^3 dx,$$

где $a_0, a_1, a_2, a_3 \in \mathbb{R}$ – произвольные константы. Можно заметить, что квадратура будет точно вычислять интеграл этого полинома тогда, когда точно будут вычисляться интегралы от функций $f(x) = 1, f(x) = x, f(x) = x^2, f(x) = x^3$. В таком случае мы получаем систему уравнений:

$$\begin{cases} c_1 + c_2 = \int_{-1}^1 dx = 2, \\ c_1 x_1 + c_2 x_2 = \int_{-1}^1 x dx = 0, \\ c_1 x_1^2 + c_2 x_2^2 = \int_{-1}^1 x^2 dx = \frac{2}{3}, \\ c_1 x_1^3 + c_2 x_2^3 = \int_{-1}^1 x^3 dx = 0. \end{cases}$$

Из первых двух уравнений мы получаем:

$$c_1 = \frac{2x_2}{x_2 - x_1} \quad c_2 = \frac{-2x_1}{x_2 - x_1} \quad \text{что при подстановке в третье дает: } x_1 x_2 = -\frac{1}{3} \quad \text{Тогда четвертое уравнение становится: } x_1 + x_2 = 0 \quad \text{что в результате дает следующее решение:}$$

Вопрос №41

1. Число обусловленности (определение). Связь числа обусловленности с устойчивостью решения СЛАУ к малым изменениям. Связь числа обусловленности с погрешностью округления в контексте арифметики с t значащими цифрами.

Соотношение между относительной погрешностью вычислений $\frac{\|x - \tilde{x}\|}{\|x\|}$ и нормой вектора невязки $\|r\|$ регулируется числом обусловленности.

Теорема (о связи между относительной погрешностью и вектором невязки) Пусть $\tilde{x} \in \mathbb{R}$ является приближением к точному решению СЛАУ $Ax = b$, A – невырожденная матрица и r – вектор невязки. Тогда для любой индуцированной матричной нормы верно: $\|x - \tilde{x}\| \leq \|r\| * \|A^{-1}\|$ и при $x, b \neq 0$: $\frac{\|x - \tilde{x}\|}{\|x\|} \leq K(A) \frac{\|r\|}{\|b\|}$, где $K(A) = \|A\| \cdot \|A^{-1}\|$ называется числом обусловленности.

Устойчивость решения СЛАУ к малым изменениям: Предположим, что приближенное решение \tilde{x} является точным решением модифицированной СЛАУ, полученной с помощью малых возмущений матрицы A :

$$(A + \delta A)\tilde{x} = b \Rightarrow r = b - A\tilde{x} = \delta A\tilde{x}.$$

После подстановки вектора невязки во второе неравенство теоремы:

$$\frac{\|x - \tilde{x}\|}{\|x\|} \leq K(A) \frac{\|\delta A\tilde{x}\|}{\|b\|} \leq K(A) \|\delta A\| \frac{\|\tilde{x}\|}{\|b\|}.$$

Можно заметить, что относительная погрешность вычислений будет тем меньше, чем меньше будет число обусловленности $K(A)$ (при этом легко увидеть, что $K(A) \geq 1$). Слишком большое число обусловленности, в свою очередь, приводит к дестабилизации вычислительных погрешностей (чем больше $K(A)$, тем сильнее влияние погрешностей округления δA на относительную погрешность). Оценим число обусловленности без явного вычисления обратной матрицы, что само по себе было бы источником погрешностей, и найдем связь между числом обусловленности и точностью полученного решения в контексте арифметики с t значащими цифрами.

Можно доказать, что если приближенное решение \tilde{x} было получено с помощью метода Гаусса и арифметики с t значащими цифрами, то верно следующее приближение для нормы вектора невязки: $\|r\| \approx 10^{-t} \|A\| \cdot \|\tilde{x}\|$. Приближение для числа обусловленности теперь можно получить, рассмотрев следующую СЛАУ: $Ay = r$. Т.к. предполагалось, что решение \tilde{x} было получено с помощью метода Гаусса, в наличии гарантированно имеются множители, необходимые для LU-разложения матрицы A , что означает, что решение у

$(D - L)^{-1}b$ Методы Якоби и Гаусса-Зейделя являются безусловно сходящимися для матриц со строго диагональным преобладанием.

2. Теорема о методе сопряженных направлений (доказательство)

Пусть система векторов $\{v^{(1)}, \dots, v^{(n)}\}$ является А-ортогональной, а матрица A положительно определенной. Тогда $Ax^{(n)} = b$, где $x^{(n)}$ определяется с помощью итерационного алгоритма ниже:

$$t_k = \frac{\langle v^{(k)}, b - Ax^{(k-1)} \rangle}{\langle v^{(k)}, Av^{(k)} \rangle},$$

$$x^{(k)} = x^{(k-1)} + t_k v^{(k)}.$$

Раскроем итерации в выражении $Ax^{(n)}$ вплоть до начального приближения $x^{(0)}$:

$$\begin{aligned} Ax^{(n)} &= A(x^{(n-1)} + t_n v^{(n)}), \\ &= A(x^{(n-2)} + t_{n-1} v^{(n-1)} + t_n v^{(n)}), \\ &= \dots \\ &= A(x^{(0)} + t_1 v^{(1)} + \dots + t_n v^{(n)}), \end{aligned}$$

из чего следует:

$Ax^{(n)} - b = Ax^{(0)} - b + A(t_1 v^{(1)} + \dots + t_n v^{(n)})$ Посчитаем скалярное произведение обеих сторон уравнения, домножив его на $v^{(k)}$:

$$\langle Ax^{(n)} - b, v^{(k)} \rangle = \langle Ax^{(0)} - b, v^{(k)} \rangle + t_k \langle v^{(k)}, Av^{(k)} \rangle.$$

(1) Из определения t_k можно получить выражение:

$$\begin{aligned} t_k \langle v^{(k)}, Av^{(k)} \rangle &= \langle v^{(k)}, b - Ax^{(k-1)} \rangle \\ &= \langle v^{(k)}, b - Ax^{(0)} + Ax^{(1)} - Ax^{(1)} + Ax^{(2)} - \dots - Ax^{(k-2)} + Ax^{(k-2)} - Ax^{(k-1)} \rangle \\ &= \langle v^{(k)}, b - Ax^{(0)} \rangle + \langle v^{(k)}, Ax^{(1)} - Ax^{(1)} \rangle + \dots + \langle v^{(k)}, Ax^{(k-2)} - Ax^{(k-1)} \rangle. \end{aligned}$$

Заметим, что из $x^{(i)} = x^{(i-1)} + t_i v^{(i)}$ следует $Ax^{(i-1)} - Ax^{(i)} = -A t_i v^{(i)}$. Тогда подстановка в выражение выше дает:

$$t_k \langle v^{(k)}, Av^{(k)} \rangle = \langle v^{(k)}, b - Ax^{(0)} \rangle,$$

из чего после подстановки в (1) мы получаем:

$$\begin{aligned} \langle Ax^{(n)} - b, v^{(k)} \rangle &= \langle Ax^{(0)} - b, v^{(k)} \rangle + \langle v^{(k)}, b - Ax^{(0)} \rangle \\ \Rightarrow \langle Ax^{(n)} - b, v^{(k)} \rangle &= 0 \quad \forall k = 1, \dots, n. \end{aligned}$$

Последнее выражение говорит о том, что вектор невязки на n -й итерации ортогонален всем векторам $\{v^{(1)}, \dots, v^{(n)}\}$, которые, будучи линейно независимыми, задают базис пространства. Это возможно тогда и только тогда, когда $Ax^{(n)} - b = 0$, что означает, что $x^{(n)}$ является решением исходной СЛАУ.

$$c_1 = 1,$$

$$c_2 = 1,$$

$$x_1 = -\frac{1}{\sqrt{3}},$$

$$x_2 = \frac{1}{\sqrt{3}}.$$

Таким образом квадратура Гаусса со степенью точности 3 имеет следующий вид: $\int_{-1}^1 f(x) dx \approx$

$$f\left(-\frac{1}{\sqrt{3}}\right) + f\left(\frac{1}{\sqrt{3}}\right)$$

Подобным образом мы можем построить квадратуры Гаусса произвольной степени точности

может быть быстро найдено. Приближение \tilde{y} удовлетворяет:

$$\tilde{y} \approx A^{-1}r = A^{-1}(b - A\tilde{x}) = A^{-1}b - \tilde{x} = x - \tilde{x}.$$

Таким образом вектор \tilde{y} является оценкой вычислительной погрешности, возникшей при нахождении решения \tilde{x} . Оценим его норму:

$$\begin{aligned} \|\tilde{y}\| &\approx \|A^{-1}r\| \leq \|A^{-1}\| \cdot \|r\| \approx \|A^{-1}\| \cdot 10^{-t} \|A\| \cdot \|\tilde{x}\| = 10^{-t} K(A) \cdot \|\tilde{x}\| \\ &\Rightarrow \frac{\|x - \tilde{x}\|}{\|\tilde{x}\|} \approx 10^{-t} K(A) \end{aligned}$$

2. Теорема о связи между относительной погрешностью и вектором невязки

Выше сказано

Вопрос № 8

- Методы вывода формул численного дифференцирования. Формула численного дифференцирования первого порядка (вывод через разложение в ряд Тейлора). Вычислительная устойчивость/неустойчивость численного дифференцирования.

Так или иначе, все формулы численного дифференцирования предполагают аппроксимацию известной функции по значениям в узлах.

Основные методы:

- Определение производной;
- Интерполяция;
- Разложение в ряд Тейлора

Пусть нам известны значения функции $f(x)$ в точках x_1 и $x_1 + h$. Разложим данную функцию в ряд Тейлора относительно точки x_1 : $f(x) = f(x_1) + f'(x_1)(x - x_1) + \frac{f''(\xi)}{2}(x - x_1)^2, \xi \in (x_1; x)$.

Найдем значение ряда в точке $x_1 + h$: $f(x_1 + h) = f(x_1) + f'(x_1)(x_1 + h - x_1) + \frac{f''(\xi)}{2}(x_1 + h - x_1)^2$
 $f(x_1 + h) = f(x_1) + f'(x_1)h + \frac{f''(\xi)}{2}h^2$
 $\frac{f(x_1 + h) - f(x_1)}{h} = \frac{f(x_1)}{h} + f'(x_1) + \frac{f''(\xi)}{2}h$

В предположении, что h достаточно мало, отбрасываем член порядка малости $O(h)$ и получаем формулу численного дифференцирования первого порядка:

$$f'(x_1) \approx \frac{f(x_1 + h) - f(x_1)}{h} - \frac{h}{2}f''(\xi)$$

Численное дифференцирование является вычислительно НЕустойчивым методом. (там из-за вычитания близких значений накапливается погрешность)

- Вывод центральной формулы численного дифференцирования 2-го порядка методом разложения в ряд Тейлора

Метод разложения в ряд Тейлора становится удобен в случае, когда требуется построить формулу для вычисления высших производных. В качестве примера построим формулу для второй производной функции $f(x)$. Для этого допустим, что нам известны ее значения в точках $x_1 - h$, x_1 и $x_1 + h$, и разложим ее в ряд Тейлора в точке x_1 :

$$f(x) = f(x_1) + f'(x_1)(x - x_1) + \frac{f''(x_1)}{2}(x - x_1)^2 + \frac{f'''(x_1)}{6}(x - x_1)^3 + \frac{f^{(4)}(\xi)}{24}(x - x_1)^4, \text{ где } \xi \in (x_1; x).$$

Вопрос №31

- Метод сопряженных градиентов: задача минимизации, направления поиска.

Рассмотрим задачу минимизации вектора невязки. Пусть x^* является решением СЛАУ $Ax^* = b$, где $A \in \mathbb{R}^{n \times n}$. Тогда: $x^* = \operatorname{argmin}_x r^T r = \operatorname{argmin}_x (b - Ax)^T (b - Ax)$.

Рассмотрим положительно определенную и, следовательно, симметричную матрицу A . Для симметричной матрицы A верно: $\langle Ax, y \rangle = \langle Ay, x \rangle$, где $\langle \cdot, \cdot \rangle$ - скалярное произведение векторов; $x, y \in \mathbb{R}^n$ - произвольные вектора. Тогда задача минимизации может быть записана следующим образом:

$$\operatorname{argmin}_x [(b, b) - 2\langle Ax, b \rangle + \langle Ax, Ax \rangle] = \operatorname{argmin}_x [\langle Ax, Ax \rangle - 2\langle Ax, b \rangle] = \operatorname{argmin}_x [\langle x, Ax \rangle - 2\langle x, b \rangle].$$

Обозначим $g(x) = \langle x, Ax \rangle - 2\langle x, b \rangle$, конечная форма задачи минимизации:

$$x^* = \operatorname{argmin}_x g(x).$$

Для нахождения минимума, начав с некоторого приближения x , необходимо определить направление поиска v и шаг t .

Рассмотрим подобный поиск в контексте функции $g(x)$:

$$\begin{aligned} g(x + tv) &= (x + tv, A(x + tv)) - 2\langle x + tv, b \rangle \\ &= (x, Ax) + t\langle x, Av \rangle + t\langle v, Ax \rangle + t^2\langle v, Av \rangle - 2\langle x, b \rangle - 2t\langle v, b \rangle \\ &= (x, Ax) - 2\langle x, b \rangle + 2t\langle v, Ax \rangle + t^2\langle v, Av \rangle - 2t\langle v, b \rangle \\ &= g(x) + t^2\langle v, Av \rangle - 2t\langle v, b - Ax \rangle. \end{aligned}$$

В первую очередь найдем оптимальный шаг $t^{(opt)}$, минимизирующий функцию $g(x + tv)$:

$$\begin{aligned} \frac{\partial g}{\partial t} &= 2t\langle v, Av \rangle - 2\langle v, b - Ax \rangle = 0 \\ \Rightarrow t^{(opt)} &= \frac{\langle v, b - Ax \rangle}{\langle v, Av \rangle} \\ \Rightarrow g(x + t^{(opt)}v) &= g(x) - \frac{\langle v, b - Ax \rangle^2}{\langle v, Av \rangle}. \end{aligned}$$

Из последнего выражения очевидно, что решение x^* минимизирует $g(x)$ для любого направления $v \neq 0$. Задавшись начальным приближением $x^{(0)}$ и начальным направлением поиска $v^{(1)}$, мы можем построить итерационный алгоритм:

$$t_k = \frac{\langle v^{(k)}, b - Ax^{(k-1)} \rangle}{\langle v^{(k)}, Av^{(k)} \rangle},$$

$$x^{(k)} = x^{(k-1)} + t_k v^{(k)}.$$

Следующим шагом является генерация таких направлений поиска $v^{(k)}$, что метод сходится достаточно быстро.

- Направление наискорейшего спуска

$$-\frac{\partial g}{\partial x} = -2(Ax - b) = 2r \quad \text{Тогда}$$

направление поиска определяется как $v^{(k+1)} = r^{(k)}$ - метод

Вопрос №16

- Тригонометрический интерполяционный полином (постановка задачи интерполяции). Связь с дискретным преобразованием Фурье.

Тригонометрический полином:

$$\sum_{k=-n}^n \widehat{a}_k e^{ikx} = 2 \sum_{k=0}^n [\operatorname{Re}(\widehat{a}_k) \cos kx - \operatorname{Im}(\widehat{a}_k) \sin kx], \text{ где } \widehat{a}_k^* = \widehat{a}_{-k}$$

Тригонометрический ряд:

$$\sum_{k=0}^n (a_k \cos kx + b_k \sin kx)$$

Экспоненциальная форма:

$$\sum_{k=-n}^n \widehat{a}_k e^{ikx}$$

Доказательство:

Рассмотрим формулу Эйлера:

$$\begin{aligned} e^{ikx} &= \cos kx + i \sin kx \quad | \quad \widehat{a}_k \in \mathbb{C} \\ \widehat{a}_k e^{ikx} &= \widehat{a}_k \cos kx + i \widehat{a}_k \sin kx = \\ &= \operatorname{Re}(\widehat{a}_k) \cos kx + i \operatorname{Im}(\widehat{a}_k) \cos kx - \\ &- \operatorname{Im}(\widehat{a}_k) \sin kx + i \operatorname{Re}(\widehat{a}_k) \sin kx, \text{ где } \operatorname{Re}(\widehat{a}_k) \text{ и } \operatorname{Im}(\widehat{a}_k) - \text{ вещественная и мнимая части } \widehat{a}_k \text{ соответственно.} \end{aligned}$$

Добавим к обоим частям равенства комплексно сопряженное (та же формула, но там где $+i$ добавляется $-i$)

$$\widehat{a}_k e^{ikx} + \widehat{a}_k^* e^{-ikx} = 2 \operatorname{Re}(\widehat{a}_k) \cos kx - 2 \operatorname{Im}(\widehat{a}_k) \sin kx$$

где комплексно сопряженное число обозначено через $*$. Тогда суммирование обеих частей равенства по $k=0, \dots, n$ приводит к выражению:

$$\sum_{k=0}^n (\widehat{a}_k e^{ikx} + \widehat{a}_k^* e^{-ikx}) = \sum_{k=-n}^n \widehat{a}_k e^{ikx} = \sum_{k=0}^n [2 \operatorname{Re}(\widehat{a}_k) \cos kx - 2 \operatorname{Im}(\widehat{a}_k) \sin kx], \text{ где } \widehat{a}_k^* = \widehat{a}_{-k}$$

Таким образом мы получили экспоненциальную форму

тригонометрического ряда, где коэффициенты связаны друг с другом следующим образом:

$$a_k = 2 \operatorname{Re}(\widehat{a}_k), b_k = -2 \operatorname{Im}(\widehat{a}_k), \text{ где } k = 0, \dots, n$$

Для того, чтобы перейти к форме полинома, обозначим $z = e^{ix}$. Тогда имеем:

$$\sum_{k=-n}^n \widehat{a}_k e^{ikx} = \sum_{k=-n}^n \widehat{a}_k z^k \quad \text{что при} \quad \text{домножении на } z^n \text{ дает комплекснозначный полином степени } 2n, \text{ по конвенции называемый тригонометрическим полиномом } n\text{-й степени:}$$

$$Q_n(z) = z^n \sum_{k=-n}^n \widehat{a}_k z^k$$

Дискретное преобразование Фурье - переход из физического пространства ($y_i = f(x_i)$) в пространство Фурье (\widehat{a}_k).

Тригонометрический интерполяционный полином (дискретное преобразование Фурье):

$$f(x) = \sum_{k=-m+1}^{m-1} \widehat{a}_k e^{ikx} + \widehat{a}_m \cos(mx),$$

$$2\widehat{a}_k = \frac{(-1)^k}{2m} \sum_{j=0}^{2m-1} y_j e^{-\frac{ijk\pi}{2}}, \quad \widehat{a}_m = \frac{(-1)^k}{2m} \sum_{j=0}^{2m-1} y_j (-1)^j.$$

- Вывод тригонометрического интерполяционного полинома.

В первом вопросе

Тогда значения ряда в точках $x_1 - h$ и $x_1 + h$ будут равны:

$$f(x_1 - h) = f(x_1) - hf'(x_1) + \frac{h^2}{2}f''(x_1) - \frac{h^3}{6}f'''(x_1) + \frac{h^4}{24}f^{(4)}(\xi_1),$$

$$f(x_1 + h) = f(x_1) + hf'(x_1) + \frac{h^2}{2}f''(x_1) + \frac{h^3}{6}f'''(x_1) + \frac{h^4}{24}f^{(4)}(\xi_2),$$

где $\xi_1 \in (x_1 - h; x_1)$ и $\xi_2 \in (x_1; x_1 + h)$. Сложив два равенства, получаем (1-ое ур-е, а потом следствие и 2-е):

$$f(x_1 - h) + f(x_1 + h) = 2f(x_1) + \frac{h^4}{24}[f^{(4)}(\xi_1) + f^{(4)}(\xi_2)]$$

$$f''(x_1) = \frac{f(x_1 - h) - 2f(x_1) + f(x_1 + h)}{h^2} - \frac{h^2}{24}[f^{(4)}(\xi_1) + f^{(4)}(\xi_2)].$$

Предположим, что $f(x) \in C^4[x_1 - h; x_1 + h]$. Тогда по теореме о промежуточном значении существует такое $\xi \in (x_1 - h; x_1 + h)$, что $f^{(4)}(\xi) = \frac{1}{2}[f^{(4)}(\xi_1) + f^{(4)}(\xi_2)]$

что в результате дает формулу численного дифференцирования второго порядка для второй производной:

$$f''(x_1) = \frac{f(x_1 - h) - 2f(x_1) + f(x_1 + h)}{h^2} - \frac{h^2}{12}f^{(4)}(\xi),$$

, где $\xi \in (x_1 - h; x_1 + h)$

градиентного спуска. Метод имеет медленную сходимость

- Сопряженные направления - такие направления $v^{(i)}$, что $\langle v^{(i)}, Av^{(j)} \rangle = 0, i \neq j$. Полученная система векторов $v^{(1)}, \dots, v^{(n)}$ называется A-ортогональной. Эта система задает базис пространства, так как является линейно независимой. Метод сопряженных градиентов выбирает такие $\{v^{(i)}\}$, что система, состоящая из векторов невязки $\{r^{(i)}\}$, является ортогональной, т.е. $\langle r^{(i)}, r^{(j)} \rangle = 0$ для $i \neq j$.

2. Итерационный алгоритм метода сопряженных градиентов в общем виде. Вывод оптимального шага в методе сопряженных градиентов.

В первом вопросе

Вопрос №34

1. Метод Ньютона (формулировка для многомерного случая). Геометрическая интерпретация метода Ньютона (одномерный случай).

Метод Ньютона для систем нелинейных алгебраических уравнений:

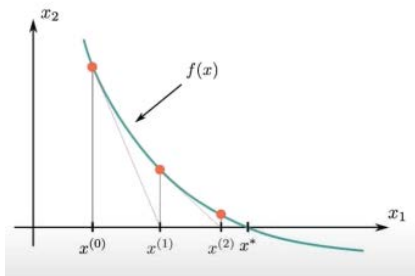
$$x^{(k)} = g(x^{(k)}) = x^{(k-1)} - J^{-1}(x^{(k-1)})f(x^{(k-1)}).$$

На практике обратная от матрицы Якоби не вычисляется, вместо чего используется следующая двухшаговая процедура:

$$J(x^{(k-1)})y^{(k-1)} = f(x^{(k-1)}),$$

$$x^{(k)} = x^{(k-1)} - y^{(k-1)},$$

где $y^{(k-1)}$ находится через решение первого уравнения. (В такой форме метод Ньютона требует на каждой итерации n вычислений значений векторной функции $f(x^{(k-1)})$, n^2 вычислений значений матрицы Якоби $J(x^{(k-1)})$ и $O(n^3)$ операций для нахождения $y^{(k-1)}$. Подобная алгоритмическая сложность не позволяет использовать метод Ньютона в его стандартной форме для систем большой размерности.)



На рисунке показана геометрическая интерпретация: ищется корень функции $f(x) - x^*$. Начальное приближение — $x^{(0)}$, ищем ноль касательной к функции в точке $x^{(0)}$ — это точка $x^{(1)}$, далее ищем ноль касательной к функции в точке $x^{(1)}$ — точка $x^{(2)}$ и т.д., в процессе $\{x^{(k)}\}_{k=0}^{\infty}$ сходится к x^* .

2. Вывод метода Ньютона (многомерный случай).

$g(x) = x - \Omega(x)f(x)$ (возьмем по координатам)

$$g_i(x) = x_i - \sum_{j=1}^n \omega_{ij}(x) f_j(x).$$

Для построения метода простой итерации с квадратичной сходимостью нам необходимо обнулить производные функций $g_i(x)$ в точке x^* . Найдем для начала соответствующие производные:

$$\frac{\partial g_i(x)}{\partial x_j} = \begin{cases} 1 - \sum_{j=1}^n \frac{\partial \omega_{ij}(x)}{\partial x_j} f_j(x) - \sum_{j=1}^n \omega_{ij}(x) \frac{\partial f_j(x)}{\partial x_j}, & i = j, \\ - \sum_{j=1}^n \frac{\partial \omega_{ij}(x)}{\partial x_j} f_j(x) - \sum_{j=1}^n \omega_{ij}(x) \frac{\partial f_j(x)}{\partial x_j}, & i \neq j. \end{cases}$$

Вопрос №42

1. Методы вывода формул численного дифференцирования. Формула численного дифференцирования первого порядка (вывод через разложение в ряд Тейлора). Вычислительная устойчивость/неустойчивость численного дифференцирования.

Так или иначе, все формулы численного дифференцирования предполагают аппроксимацию известной функции по значениям в узлах.

Основные методы:

- Определение производной;
- Интерполяция;
- Разложение в ряд Тейлора

Пусть нам известны значения функции $f(x)$ в точках x_1 и $x_1 + h$. Разложим данную функцию в ряд Тейлора относительно точки x_1 : $f(x) = f(x_1) + f'(x_1)(x - x_1) + \frac{f''(\xi)}{2} (x - x_1)^2, \xi \in (x_1; x)$.

Найдем значение ряда в точке $x_1 + h$: $f(x_1 + h) = f(x_1) + f'(x_1)(x_1 + h - x_1) + \frac{f''(\xi)}{2} (x_1 + h - x_1)^2$
 $f(x_1 + h) = f(x_1) + f'(x_1)h + \frac{f''(\xi)}{2} h^2$

$$\frac{f(x_1 + h) - f(x_1)}{h} = \frac{f(x_1)}{h} + f'(x_1) + \frac{f''(\xi)}{2} h$$

В предположении, что h достаточно мало, отбрасываем член порядка малости $O(h)$ и получаем формулу численного дифференцирования первого порядка:

$$f'(x_1) \approx \frac{f(x_1 + h) - f(x_1)}{h} - \frac{h}{2} f''(\xi)$$

Численное дифференцирование является вычислительно НЕустойчивым методом. (там из-за вычитания близких значений накапливается погрешность)

2. Вывод центральной формулы численного дифференцирования 2-го порядка методом дифференцирования многочлена Лагранжа.

Для этого нам необходимо использовать интерполяцию $f(x)$ в трех точках x_1, x_2, x_3 :

$$f'(x_j) = f(x_1) \frac{d}{dx} \left[\frac{(x - x_2)(x - x_3)}{(x_1 - x_2)(x_1 - x_3)} \right]_{x=x_j} + f(x_2) \frac{d}{dx} \left[\frac{(x - x_1)(x - x_3)}{(x_2 - x_1)(x_2 - x_3)} \right]_{x=x_j} + f(x_3) \frac{d}{dx} \left[\frac{(x - x_1)(x - x_2)}{(x_3 - x_1)(x_3 - x_2)} \right]_{x=x_j} + \frac{1}{6} \prod_{i=1, i \neq j}^3 (x_j - x_i) f^{(3)}(\xi_j),$$

из чего следует:

$$f'(x_j) = f(x_1) \frac{2x_j - x_2 - x_3}{(x_1 - x_2)(x_1 - x_3)} + f(x_2) \frac{2x_j - x_1 - x_3}{(x_2 - x_1)(x_2 - x_3)} + f(x_3) \frac{2x_j - x_1 - x_2}{(x_3 - x_1)(x_3 - x_2)} + \frac{1}{6} \prod_{i=1, i \neq j}^3 (x_j - x_i) f^{(3)}(\xi_j). \quad (3.11)$$

Вопрос №32

1. Системы нелинейных алгебраических уравнений (различные формы представления). Неподвижная точка (определение). Сведение задачи о корнях системы нелинейных алгебраических уравнений к задаче о неподвижной точке.

В общем виде система нелинейных алгебраических уравнений имеет вид:

$$f_1(x_1, \dots, x_n) = 0,$$

$$f_2(x_1, \dots, x_n) = 0,$$

...

$$f_n(x_1, \dots, x_n) = 0.$$

Для векторной функции $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$:

$$f(x) : \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \mapsto \begin{bmatrix} f_1(x_1, \dots, x_n) \\ \vdots \\ f_n(x_1, \dots, x_n) \end{bmatrix}$$

система принимает вид: $f(x) = 0$.

Вектор $x^* \in \mathbb{R}^n$ называется неподвижной точкой векторной функции $f(x)$, если $f(x^*) = x^*$.

Систему вида $f(x) = 0$ можно свести к задаче о поиске неподвижной точки, рассмотрев следующую функцию: $g(x) = x - \Omega(x)f(x)$, где матрица $\Omega(x): \mathbb{R}^n \rightarrow \mathbb{R}^{n \times n}$ является невырожденной (аналог параметра релаксации в линейных системах). В таком случае множество неподвижных точек функции $g(x)$ совпадает с множеством корней функции $f(x)$.

2. Теорема об окрестности сходимости метода Ньютона (доказательство).

Пусть $f(x) \in C^2[a; b]$ и существует такое $x^* \in (a; b)$, что $f(x^*) = 0$ и $f'(x^*) \neq 0$. Тогда существует такое $\delta > 0$, что последовательность $\{x^{(k)}\}_{k=0}^{\infty}$, генерируемая методом Ньютона, сходится к x^* для любого $x^{(0)} \in [x^* - \delta; x^* + \delta]$.

Док-во: Рассмотрим метода Ньютона как метод простой итерации в соответствии с $x^{(k)} = g(x^{(k-1)})$. Тогда необходимо найти интервал $D = [x^* - \delta; x^* + \delta]$, который функция $g(x)$ отображает в себя, и в котором $|g'(x)| \leq \gamma$, где $\gamma \in (0; 1)$. Так как f' является непрерывной и $f'(x^*) \neq 0$, существует такое $\delta_1 > 0$, что $f'(x) \neq 0$ в замкнутой δ_1 -окрестности точки x^* , т.е. для $x \in [x^* - \delta_1; x^* + \delta_1] \subset [a; b]$. Тогда непрерывная производная $g'(x) \in C^1[a; b]$ также будет существовать в этом интервале:

$$g'(x) = 1 - \frac{(f'(x))^2 - f(x)f''(x)}{(f'(x))^2} = \frac{f(x)f''(x)}{(f'(x))^2}.$$

Найдем значение $g'(x)$ в точке x^* :

$$g'(x^*) = \frac{f(x^*)f''(x^*)}{(f'(x^*))^2} = 0.$$

Тогда существует такое $\delta \in (0; \delta_1)$, что $|g'(x)| \leq \gamma$ для $\gamma \in (0; 1)$ в замкнутой δ -окрестности точки x^* , т.е. для $x \in [x^* - \delta; x^* + \delta] \subset [a; b]$. Теперь докажем, что полученный интервал $D = [x^* - \delta; x^* + \delta]$ отображается функцией $g(x)$ сам в себя. По теореме

Тогда обнуление производных дает уравнения:

$$\sum_{j=1}^n \omega_{ij}(\mathbf{x}^*) \frac{\partial f_j(\mathbf{x}^*)}{\partial x_j} = 1, \quad i = j,$$

$$\sum_{j=1}^n \omega_{ij}(\mathbf{x}^*) \frac{\partial f_j(\mathbf{x}^*)}{\partial x_j} = 0, \quad i \neq j.$$

где мы использовали факт $g_i(\mathbf{x}^*) = 0$. Легко убедиться, что в матричном виде это эквивалентно следующему выражению:

$$\Omega(\mathbf{x}^*) \mathbf{J}(\mathbf{x}^*) = \mathbf{E}$$

где матрица \mathbf{J} называется матрицей Якоби и имеет следующую форму:

$$\mathbf{J}(\mathbf{x}) = \begin{bmatrix} \frac{\partial f_1(\mathbf{x})}{\partial x_1} & \dots & \frac{\partial f_1(\mathbf{x})}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_n(\mathbf{x})}{\partial x_1} & \dots & \frac{\partial f_n(\mathbf{x})}{\partial x_n} \end{bmatrix}$$

Таким образом, для квадратичной сходимости в качестве $\Omega(\mathbf{x})$ мы можем выбрать обратную матрицу Якоби: $\Omega(\mathbf{x}) = \mathbf{J}^{-1}(\mathbf{x})$, что дает формулировку метода Ньютона для систем нелинейных алгебраических уравнений:

$$\mathbf{x}^{(k)} = \mathbf{g}(\mathbf{x}^{(k)}) = \mathbf{x}^{(k-1)} - \mathbf{J}^{-1}(\mathbf{x}^{(k-1)}) \mathbf{f}(\mathbf{x}^{(k-1)}).$$

Теперь предположим, что узлы распределены равномерно, т.е. $x_2 = x_1 + h$ и $x_3 = x_1 + 2h$.

Тогда выражение (3.11) принимает вид:

$$f'(x_j) = f(x_1) \frac{2x_j - 2x_1 - 3h}{-h \cdot (-2h)} + f(x_1 + h) \frac{2x_j - 2x_1 - 2h}{h \cdot (-h)} + f(x_1 + 2h) \frac{2x_j - 2x_1 - h}{2h \cdot h} +$$

$$+ \frac{1}{6} \prod_{i=1, i \neq j}^3 (x_j - x_i) f^{(3)}(\xi(x_j)). \quad (3.12)$$

Записав последнее выражение для x_1 , $x_1 + h$ и $x_1 + 2h$, мы получаем три формулы для численного дифференцирования второго порядка точности:

$$f'(x_1) = \frac{-3f(x_1) + 4f(x_1 + h) - f(x_1 + 2h)}{2h} + \frac{h^2}{3} f^{(3)}(\xi),$$

$$f'(x_1 + h) = \frac{f(x_1 + 2h) - f(x_1)}{2h} - \frac{h^2}{6} f^{(3)}(\xi),$$

$$f'(x_1 + 2h) = \frac{f(x_1) - 2f(x_1 + h) + 3f(x_1 + 2h)}{2h} + \frac{h^2}{3} f^{(3)}(\xi).$$

Лагранжа о среднем значении для $x \in D$ существует такое $\xi \in (x; x^*)$ или $\xi \in (x^*; x)$, что:

$$|g'(\xi)| = \frac{|g(x) - g(x^*)|}{|x - x^*|} = \frac{|g(x) - x^*|}{|x - x^*|}$$

$$\implies |g(x) - x^*| = |g'(\xi)| \cdot |x - x^*| \leq \gamma |x - x^*| < |x - x^*|.$$

Так как $|x - x^*| \leq \delta$, мы получаем $|g(x) - x^*| \leq \delta$, из чего следует, что $g(x)$ отображает D в себя. Тогда последовательность $\{x^{(k)}\}_{k=0}^{\infty}$ сходится к x^* для любого $x(0) \in [x^* - \delta; x^* + \delta]$.

Вопрос №10

1. Квадратурные формулы (определение). Составные формулы численного интегрирования (алгоритм их построения). Зависимость порядка точности от гладкости.

Аппроксимацию интеграла называют квадратурой: $\int_a^b f(x)dx \approx \sum_{i=1}^n a_i f(x_i)$, $a_i \in \mathbb{R}$
Алгоритм построения составных формул:

- (а) Рассмотрим интеграл $\int_a^b f(x)dx$, где $f(x) \in C^k[a, b]$ (k зависит от выбранной формулы численного интегрирования).
- (б) Разделим отрезок $[a; b]$ на некоторое число подотрезков n , тогда $h = \frac{b-a}{n}$ – шаг интегрирования.
- (с) Применим одну из формул численного интегрирования на каждой паре подотрезков и получим составную формулу численного интегрирования. Порядок точности квадратуры зависит от гладкости, если функция недостаточно гладкая, погрешность увеличивается.

2. Вывод составной формулы Симпсона.

Рассмотрим интеграл $\int_a^b f(x)dx$. Разделим отрезок $[a; b]$ на четное число подотрезков n и применим формулу трапеций на каждом из них.

Тогда для $f(x) \in C^4[a; b]$, $h = \frac{b-a}{n}$ и $x_i = a + (i-1)h$, где $i = 1, \dots, n+1$ мы имеем: $\int_a^b f(x)dx = \sum_{i=1}^{\frac{n}{2}} \int_{x_{2i-1}}^{x_{2i+1}} f(x)dx =$

$$\frac{h}{3} \sum_{i=1}^{\frac{n}{2}} [f(x_{2i-1}) + 4f(x_{2i}) + f(x_{2i+1})] - \frac{h^5}{90} \sum_{i=1}^{\frac{n}{2}} f^{(5)}(\xi_i)$$

где $\xi_i \in (x_{2i-1}; x_{2i+1})$. Заметим, что все нечетные узлы за исключением x_1 и x_{n+1} дважды повторяются в сумме, что позволяет упростить выражение:

$$\int_a^b f(x)dx = \frac{h}{3} \left[f(x_1) + 2 \sum_{i=1}^{\frac{n}{2}-1} f(x_{2i+1}) + 4 \sum_{i=1}^{\frac{n}{2}} f(x_{2i}) + f(x_{n+1}) \right] - \frac{h^5}{90} \sum_{i=1}^{\frac{n}{2}} f^{(4)}(\xi_i)$$

Для упрощения выражения для остаточного члена заметим, что так как $f(x) \in C^4[a; b]$, мы имеем:

$$\min_{x \in [a; b]} f^{(4)}(x) \leq f^{(4)}(\xi_i) \leq \max_{x \in [a; b]} f^{(4)}(x) \Rightarrow \frac{n}{2} \min_{x \in [a; b]} f^{(4)}(x) \leq \sum_{i=1}^{\frac{n}{2}} f^{(4)}(\xi_i) \leq \frac{n}{2} \max_{x \in [a; b]} f^{(4)}(x) \Rightarrow$$

Вопрос №27

1. Метод главных компонент. Матрица центрированных данных и ее ковариационная матрица. Теорема о главных компонентах (формулировка).

– предполагает формирование такого базиса, где выборочные дисперсии вдоль осей будут максимальны (один из способов понижения размерности)
Матрица центрированных данных A : – приводим набор данных к такому виду, чтобы матожидание = 0.

$$A = X - \bar{e} \bar{x} = (E - \frac{1}{m} \bar{e} \bar{e}^T) X,$$

где $\bar{x} = \frac{1}{m} \bar{e}^T X$ – вектор выборочных средних показаний измерений, m – количество данных для каждого измерения (количество строк), \bar{e} – единичный вектор ($\bar{e} \bar{e}^T$ – матрица единиц), строки матрицы X – показания, а столбцы – измерения. Ковариационная матрица K : $K = \nu A^T A$, где коэффициент ν зависит от выбора оценки ковариации (обычно $\nu = \frac{1}{m-1}$), а элемент k_{ij} является ковариацией i -ого и j -ого показания. Теорема:

Главными компонентами матрицы центрированных данных A являются ее сингулярные векторы, при этом j -ая главная компонента соответствует j -ому сингулярному вектору q_j и стандартному отклонению $\sqrt{\nu} \sigma_j$, где σ_j является j -ым сингулярным числом.

2. Вывод первой главной компоненты в контексте метода главных компонент и связанной с ней выборочной дисперсией.

Главные компоненты формируют ортонормальный базис, состоящий из векторов, ассоциированных с небольшими выборочными дисперсиями, а также в котором межкоординатные коэффициенты корреляции равны нулю. Первая компонента является направлением, вдоль которого выборочная дисперсия максимальная, вторая главная компонента является направлением, ортогональным первой, вдоль которого выборочная дисперсия максимальна и т.д.

$$\bar{p} = A \bar{u} \Rightarrow \sigma_p^2 = \nu \|\bar{p}\|_2^2 = \nu \|A \bar{u}\|_2^2 = \nu (A \bar{u})^T (A \bar{u}) = \nu \bar{u}^T A^T A \bar{u} = \nu \bar{u}^T K \bar{u}$$

Задача максимизации $\max_{\bar{u}, \|\bar{u}\|=1} \nu \bar{u}^T K \bar{u}$. Пот т. о собственных числах и векторах симметричной матрицы: $\nu \lambda_1 = \max_{\bar{u}, \|\bar{u}\|=1} \nu \bar{u}^T K \bar{u} \Rightarrow \lambda_1 = \sigma_1^2$, где σ_1 – максимальное сингулярное число $A \Rightarrow \bar{u}$ – соответственный

Вопрос №2

1. Числа с плавающей и фиксированной запятой (битовая структура). Машинный epsilon и его связь с относительной погрешностью.

Вычислительная погрешность появляется в результате того, что алгебраические структуры, которые определены для поля вещественных чисел \mathbb{R} , неявно заменяются на алгебраические структуры, определенные для поля рациональных чисел \mathbb{Q} .

Числа с плавающей запятой: стандарт IEEE 754

- binary16 - половинная точность;
- binary32 (float в Си) - одинарная точность;
- binary64 (double в Си) - двойная точность;
- Расширенная точность (80 бит);
- binary128 - четверная точность.

Представление вещественного числа состоит из последовательно расположенных бита знака s , бит экспоненты (порядка) e и бит мантиисы m , само значение вычисляется как

$$(-1)^s 2^{e-e_{\text{bias}}} (1 + \sum_{j=1}^p m_j 2^{-j}), \text{ где } e_{\text{bias}} - \text{смещение порядка (для binary32 - 127, для binary64 - 1023), } p - \text{число бит мантиисы (для binary32 - 23, для binary64 - 52).}$$



Представление чисел с плавающей запятой. Числа с плавающей запятой: bfloat16

- Веден в Google TPU для ускорения машинного обучения;
- Тот же динамический диапазон (разница между наименьшим и наибольшим значением), что и у binary32, но меньшая мантииса.

Числа с фиксированной запятой:

- Целая и дробная части описываются целыми числами;
- Отрицательные числа представляются с помощью дополнительного кода;

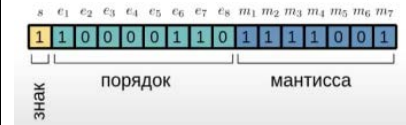


Рис. 2: Представление чисел bfloat16.



Представление чисел с фиксированной запятой.

- Могут использоваться там, где доступна только целочисленная арифметика;
- Более быстрые арифметические операции по сравнению с числами с плавающей запятой;

$\min_{x \in [a; b]} f^{(4)}(x) \leq \frac{2}{n} \sum_{i=1}^{\frac{n}{2}} f^{(4)}(\xi_i) \leq \max_{x \in [a; b]} f^{(4)}(x)$ $f^{(4)}(\xi) = \frac{2}{n} \sum_{i=2}^{\frac{n}{2}} f^{(4)}(\xi_i), \text{ где } \xi \in (a; b).$ $R = -\frac{h^5}{90} \sum_{i=1}^{\frac{n}{2}} f^{(4)}(\xi_i) = -\frac{(b-a)h^4}{180} f^{(4)}(\xi).$ <p>Тогда составная формула Симпсона:</p> $\int_a^b f(x) dx = \frac{h}{3} \left[f(x_1) + 2 \sum_{i=1}^{\frac{n}{2}-1} f(x_{2i+1}) + 4 \sum_{i=1}^{\frac{n}{2}} f(x_{2i}) + f(x_{n+1}) \right] - \frac{(b-a)h^4}{180} f^{(4)}(\xi)$	<p>сингулярный вектор A. Обозначим первую главную компоненту \bar{q}_1. Для \bar{q}_2: $\max_{\bar{u}, \ \bar{u}\ =1, (\bar{u}, \bar{q}_1)=0} \nu \bar{u}^T K \bar{u} \Rightarrow \bar{u}$ – второй сингулярный вектор \bar{q}_2. И т.д. для $\bar{q}_3, \bar{q}_4 \dots$</p>	<ul style="list-style-type: none"> • Меньший динамический диапазон и точность по сравнению с числами с плавающей запятой. <p>Машинный эпсилон:</p> <ul style="list-style-type: none"> • $\epsilon = 2^{-p}$, p - число бит мантиссы; • минимальное ϵ, для которого $1 + \epsilon > 1$ при машинных расчетах с числами данного типа; • верхняя граница относительной погрешности округления; • числовое значение, меньше которого невозможно задавать относительную точность для любого алгоритма, возвращающего вещественные числа. <p>Абсолютное значение «машинного эпсилон» зависит от разрядности сетки применяемой ЭВМ, типа используемых при расчетах чисел, и от принятой в конкретном трансляторе структуры представления чисел (количества бит, отводимых на мантиссу и на порядок).</p> <p>2. L_1 и L_2-регуляризация. Использование в контексте полиномиальной регрессии. Вывод нормального уравнения для задачи метода наименьших квадратов с L_2 регуляризацией.</p> <p>L_1: модификация задачи МНК: $\min_{\bar{a}} E_2(\bar{a}) = \min_{\bar{a}} [(\bar{y} - X\bar{a})^T (\bar{y} - X\bar{a}) + \lambda \ \bar{a}\ _1] = \min_{\bar{a}} [(\bar{y} - X\bar{a})^T (\bar{y} - X\bar{a}) + \lambda \sum_{i=0}^n a_i];$</p> <p>решается с помощью численных методов оптимизации, приводит к разреженному вектору \bar{a}</p> <p>L_2: модификация задачи МНК: $\min_{\bar{a}} E_2(\bar{a}) = \min_{\bar{a}} [(\bar{y} - X\bar{a})^T (\bar{y} - X\bar{a}) + \lambda \ \bar{a}\ _2^2] = \min_{\bar{a}} [(\bar{y} - X\bar{a})^T (\bar{y} - X\bar{a}) + \lambda \sum_{i=0}^n \bar{a}_i^2]$</p> $E_2(\bar{a}) = \sum_{j=1}^m \bar{y}_j^2 - 2 \sum_{j=1}^m \sum_{i=0}^n \bar{a}_i X_{ji} \bar{y}_j + \sum_{j=1}^m \left(\sum_{i=0}^n \bar{a}_i X_{ji} \right)^2 + \lambda \sum_{i=0}^n \bar{a}_i^2$ <p>Тогда дифференцирование относительно произвольного a_k дает</p> $\frac{\partial E_2}{\partial \bar{a}_k} = -2 \sum_{j=1}^m X_{jk} \bar{y}_j + 2 \sum_{j=1}^m X_{jk} \left(\sum_{i=0}^n \bar{a}_i X_{ji} \right) + 2\lambda \sum_{i=0}^n \bar{a}_i$ <p>что при записи в матричном виде дает.</p> $\frac{\partial E_2}{\partial \bar{a}_k} = -2X^T \bar{y} + 2X^T X \bar{a} + 2\lambda \bar{a}$ <p>Оптимальное значение вектора a находится с помощью приравнивания производной к нулю:</p> $-2X^T \bar{y} + 2X^T X \bar{a} + 2\lambda \bar{a} = 0$ $\bar{a} = (X^T X + \lambda E)^{-1} X^T \bar{y}$
---	---	---

<p>Вопрос №43</p> <p>1. Системы нелинейных алгебраических уравнений (различные формы представления). Неподвижная точка (определение). Сведение задачи о корнях системы нелинейных алгебраических уравнений к задаче о неподвижной точке.</p> <p>В общем виде система нелинейных алгебраических уравнений имеет вид:</p> $\begin{aligned} f_1(x_1, \dots, x_n) &= 0, \\ f_2(x_1, \dots, x_n) &= 0, \\ &\dots \\ f_n(x_1, \dots, x_n) &= 0. \end{aligned}$ <p>Для векторной функции $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$:</p> $f(x): \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \mapsto \begin{bmatrix} f_1(x_1, \dots, x_n) \\ \vdots \\ f_n(x_1, \dots, x_n) \end{bmatrix}$ <p>система принимает вид: $f(x) = 0$. Вектор $x^* \in \mathbb{R}^n$ называется неподвижной точкой векторной функции $f(x)$, если $f(x^*) = x^*$. Систему вида $f(x) = 0$ можно свести к задаче о поиске неподвижной точки, рассмотрев следующую функцию: $g(x) = x - \Omega(x)f(x)$, где матрица $\Omega(x): \mathbb{R}^n \rightarrow \mathbb{R}^{n \times n}$ является невырожденной (аналог параметра релаксации в линейных системах). В таком случае множество неподвижных точек функции $g(x)$ совпадает с множеством корней функции $f(x)$.</p> <p>2. Теорема о единственности неподвижной точки (доказательство для одномерного случая)</p> <p>Пусть $D = [a; b]$. Если $g \in C(D)$ и $g(D) \subset D$, то g имеет хотя бы одну неподвижную точку в D. Если при этом существует производная $g'(x)$ в $(a; b)$ и $g'(x) < 1$ для любого $x \in (a; b)$, то эта неподвижная точка является единственной.</p> <p>Доказательство: Случаи $g(a) = a$, $g(b) = b$ очевидны, так что, с учетом $g(D) \subset D$, мы рассмотрим $g(a) > a$, $g(b) < b$. Определим вспомогательную функцию $h(x)$ следующим образом: $h(x) = g(x) - x$. Для нее является справедливым следующее: $h(a) = g(a) - a > 0$, $h(b) = g(b) - b < 0$. Следовательно, по теореме о промежуточном значении можно найти такое $x^* \in D$, что $h(x^*) = 0$, то есть $g(x^*) = x^*$. Докажем вторую часть теоремы от обратного. Предположим, что существуют две неподвижные точки</p>	<p>Вопрос №44</p> <p>1. Глобальная и локальная интерполяция. Кусочная интерполяция. Интерполяция кубическими сплайнами (определение без вывода разрешающих уравнений).</p> <p>Глобальная интерполяция – интерполяция одной аппроксимирующей функцией по всему отрезку $[a; b]$. Локальная интерполяция, где мы делим отрезок $[a; b]$ на маленькие подотрезки и используем интерполяцию полиномом малой степени на каждом из этих подотрезков, после чего “склеиваем” полученное множество полиномов в единую функцию, заданную на всем отрезке $[a; b]$. Такую интерполяцию называют кусочной интерполяцией. Пусть функция $f(x)$ задана в n интерполяционных узлах $a = x_1, x_2, \dots, x_n = b$ на отрезке $[a; b]$. Тогда кубическим сплайном для функции $f(x)$ называется функция $S(x)$, для которой верно:</p> <ul style="list-style-type: none"> • $S(x)$ кусочно задана кубическими многочленами $S_i(x)$ на каждом отрезке $[x_i; x_{i+1}]$, $i = 1 \dots n - 1$; • $S_i(x_i) = f(x_i)$ и $S_i(x_{i+1}) = f(x_{i+1})$, $i = 1 \dots n - 1$; • значения смежных многочленов совпадают в общих узлах: $S_i(x_{i+1}) = S_{i+1}(x_{i+1})$, $i = 1, \dots, n - 2$; • значения первых производных смежных многочленов совпадают в общих узлах: $S'_i(x_{i+1}) = S'_{i+1}(x_{i+1})$, $i = 1, \dots, n - 2$; • значения вторых производных смежных многочленов совпадают в общих узлах: $S''_i(x_{i+1}) = S''_{i+1}(x_{i+1})$, $i = 1, \dots, n - 2$; • заданы граничные условия: <ul style="list-style-type: none"> - естественные граничные условия: $S(x_1) = S(x_n) = 0$; - граничные условия на касательную: $S(x_1) = f'(x_1)$ и $S(x_n) = f'(x_n)$ <p>2. Теорема о единственности естественного кубического сплайна</p> <p>Пусть функция $f(x)$ задана в n интерполяционных узлах $a = x_1, x_2, \dots, x_n = b$ на отрезке $[a; b]$. Тогда функция $f(x)$ имеет уникальный естественный кубический сплайн $S(x)$, т.е. удовлетворяющий граничным условиям $S''(a) = 0$ и $S''(b) = 0$.</p>	<p>Вопрос №33</p> <p>1. Метод простой итерации для нелинейных алгебраических систем. Теорема о сходимости метода простой итерации (формулировка для одномерного случая).</p> <p>Метод простых итераций: $x^{(k)} = g(x^{(k-1)})$, дает последовательность точек $\{x^{(k)}\}_{k=0}^\infty$. Пусть $D = [a; b]$, $g \in C(D)$, $g(D) \subset D$ и существует производная $g'(x)$ для $x \in (a; b)$ с таким $\gamma \in (0; 1)$, что $g'(x) \leq \gamma$ для любого $x \in (a; b)$. Тогда для любого $x^{(0)} \in D$ последовательность $\{x^{(k)}\}_{k=0}^\infty$, полученная итерацией $x^{(k)} = g(x^{(k-1)})$, сходится к единственной неподвижной точке $x^* \in D$.</p> <p>2. Теорема о сходимости метода простой итерации для нелинейных алгебраических систем (доказательство для одномерного случая).</p> <p>Так как $g'(x) < 1$ для любого $x \in (a; b)$, по теореме существует единственная неподвижная точка в D. По теореме Лагранжа о среднем значении для любого $k > 1$ существует такое $\xi(k) \in D$, что:</p> $g'(\xi(k)) = \frac{g(x^{(k-1)}) - g(x^*)}{x^{(k-1)} - x^*} = \frac{x^{(k)} - x^*}{x^{(k-1)} - x^*}$ $\Rightarrow x^{(k)} - x^* = g'(\xi(k)) \cdot x^{(k-1)} - x^* $ $\Rightarrow x^{(k)} - x^* \leq \gamma x^{(k-1)} - x^* .$ <p>Применив тот же подход для нахождения верхней границы для $x^{(k-1)} - x^*$, $x^{(k-2)} - x^*$ и т.д., получаем:</p> $\begin{aligned} x^{(k)} - x^* &\leq \gamma x^{(k-1)} - x^* \\ &\leq \gamma^2 x^{(k-2)} - x^* \\ &\dots \\ &\leq \gamma^k x^{(0)} - x^* . \end{aligned}$ <p>Тогда предел $\lim_{k \rightarrow \infty} x^{(k)} - x^*$ принимает вид:</p> $\lim_{k \rightarrow \infty} x^{(k)} - x^* \leq \lim_{k \rightarrow \infty} \gamma^k x^{(0)} - x^* = 0,$ <p>что эквивалентно сходимости метода простой итерации.</p>
--	--	--

$x^* \neq y^*$. Тогда по теореме Лагранжа о среднем значении существует такое $\xi \in [x^*; y^*] \subset D$, что:

$$\begin{aligned} g'(\xi) &= \frac{g(x^*) - g(y^*)}{x^* - y^*} \\ &= \frac{x^* - y^*}{x^* - y^*} \\ &= 1, \end{aligned}$$

что противоречит установке $|g'(x)| < 1$. Следовательно, существующая неподвижная точка является единственной.

Граничное условие $S''(a) = 0$ соответствует следующему условию на кубический многочлен $S_1(x)$:

$$\begin{aligned} S_1''(x)|_{x=a} &= [2c_1 + 6d_1(x - x_1)]|_{x=a} \\ &= 2c_1 \\ &= 0, \end{aligned}$$

т.е. $c_1 = 0$. Аналогично граничное условие $S''(b) = 0$ трансформируется в $c_n = 0$. Исходя из полученных ограничений на c_1 и c_n и уравнения $h_{i-1}c_{i-1} + 2(h_i + h_{i-1})c_i + h_ic_i + 1 = \frac{3}{h_i}(a_{i+1} - a_i) - \frac{3}{h_{i-1}}(a_i - a_{i-1})$, запишем матричное уравнение $Ac = b$, где $c = [c_1, \dots, c_n]^T$:

$$\begin{bmatrix} 1 & 0 & \dots & \dots & \dots & 0 \\ h_1 & 2(h_2 + h_1) & h_2 & 0 & \dots & 0 \\ 0 & h_2 & 2(h_3 + h_2) & h_3 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & h_{n-2} & 2(h_{n-1} + h_{n-2}) & h_{n-1} \\ 0 & \dots & \dots & \dots & \dots & 1 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ \vdots \\ c_{n-1} \\ c_n \end{bmatrix} = \begin{bmatrix} 0 \\ \frac{3}{h_2}(a_3 - a_2) - \frac{3}{h_1}(a_2 - a_1) \\ \frac{3}{h_3}(a_4 - a_3) - \frac{3}{h_2}(a_3 - a_2) \\ \vdots \\ \frac{3}{h_{n-1}}(a_n - a_{n-1}) - \frac{3}{h_{n-2}}(a_{n-1} - a_{n-2}) \\ 0 \end{bmatrix}$$

Можно заметить, что матрица A является матрицей со строгим диагональным преобладанием, т.е. $|A_{ii}| > \sum_{j \neq i} |A_{ij}|$, $i = 1, \dots, n$. Действительно:

$$2(h_i + h_{i-1}) > h_i + h_{i-1}$$

Теорема о СЛАУ с матрицами со строгим диагональным преобладанием гласит, что такая СЛАУ имеет единственное решение для c

Вопрос №45

1. Методы вывода формулы численного дифференцирования (пример для формулы первого порядка). Вычислительная устойчивость численного дифференцирования.

Так или иначе, все формулы численного дифференцирования предполагают аппроксимацию известной функции по значениям в узлах.

Основные методы:

- Определение производной;
- Интерполяция;
- Разложение в ряд Тейлора

Пусть нам известны значения функции $f(x)$ в точках x_1 и $x_1 + h$. Разложим данную функцию в ряд Тейлора относительно точки x_1 : $f(x) = f(x_1) + f'(x_1)(x - x_1) + \frac{f''(\xi)}{2}(x - x_1)^2, \xi \in (x_1; x)$.

Найдем значение ряда в точке $x_1 + h$: $f(x_1 + h) = f(x_1) + f'(x_1)(x_1 + h - x_1) + \frac{f''(\xi)}{2}(x_1 + h - x_1)^2$
 $f(x_1 + h) = f(x_1) + f'(x_1)h + \frac{f''(\xi)}{2}h^2$
 $\frac{f(x_1 + h) - f(x_1)}{h} = f'(x_1) + \frac{f''(\xi)}{2}h$

В предположении, что h достаточно мало, отбрасываем член порядка малости $O(h)$ и получаем формулу численного дифференцирования первого порядка:

$$f'(x_1) \approx \frac{f(x_1 + h) - f(x_1)}{h} - \frac{h}{2}f''(\xi)$$

Численное дифференцирование является вычислительно НЕустойчивым методом. (там из-за вычитания близких значений накапливается погрешность)

2. Вывод центральной формулы численного дифференцирования 2-го порядка для второй производной.

Метод разложения в ряд Тейлора становится удобен в случае, когда требуется построить формулу для вычисления высших производных. В качестве примера построим формулу для второй производной функции $f(x)$. Для этого допустим, что нам известны ее значения в точках $x_1 - h$, x_1 и $x_1 + h$, и разложим ее в ряд Тейлора в точке x_1 :

$$f(x) = f(x_1) + f'(x_1)(x - x_1) + \frac{f''(x_1)}{2}(x - x_1)^2 + \frac{f'''(x_1)}{6}(x - x_1)^3 + \frac{f^{(4)}(\xi)}{24}(x - x_1)^4, \text{ где } \xi \in (x_1; x).$$

Тогда значения ряда в точках $x_1 - h$ и $x_1 + h$ будут равны:

Вопрос №35

1. Метод Ньютона (вывод для одномерного случая). Особенности сходимости метода Ньютона.

Пусть нам дана функция $f(x) \in C^2[a; b]$ и точка $x^{(0)} \in [a; b]$ является приближением к корню x^* функции $f(x)$, т.е. $f(x^*) = 0$. Разложим функцию $f(x)$ в ряд Тейлора в точке $x^{(0)}$:

$$f(x) = f(x^{(0)}) + f'(x^{(0)})(x - x^{(0)}) + \frac{f''(\xi)}{2}(x - x^{(0)})^2,$$

где $\xi \in (x, x^{(0)})$ (либо $\xi \in (x^{(0)}, x)$). Метод Ньютона основан на предположении, что $|x^* - x^{(0)}|$ достаточно мало. Тогда, вычислив ряд в точке x^* , мы можем отбросить квадратичный член:

$$f(x^*) \approx f(x^{(0)}) + f'(x^{(0)})(x^* - x^{(0)}).$$

Так как $f(x^*) = 0$, мы получаем оценку для x^* , т.е. более точное приближение:

$$0 \approx f(x^{(0)}) + f'(x^{(0)})(x^* - x^{(0)})$$

$$\Rightarrow x^* \approx x^{(0)} - \frac{f(x^{(0)})}{f'(x^{(0)})}.$$

Таким образом, можно сформулировать итерационный метод, называемый методом Ньютона:

$$x^{(k)} = x^{(k-1)} - \frac{f(x^{(k-1)})}{f'(x^{(k-1)})}.$$

Особенности метода Ньютона:

- Начальное приближение $x^{(0)}$ должно быть достаточно близким к корню x^* (существуют ограничения на сходимость для произвольных $x^{(0)}$);
- Градиент $f'(x^{(k)})$ должен быть отличен от нуля для любого k . Более того, метод Ньютона тем эффективнее, чем дальше $f'(x^*)$ от нуля. Существует теорема, доказывающая существование некоторой окрестности x^* , внутри которой сходимость метода Ньютона является безусловной;
- (Метод Ньютона является “оптимальным” методом простой итерации и имеет квадратичную сходимость.)

2. Теорема о квадратичной сходимости метода простой итерации для нелинейных алгебраических систем

Пусть $x^* \in (a; b)$ является неподвижной точкой функции $g(x)$, т.е. $x^* = g(x^*)$. Пусть также $g'(x^*) = 0$ и $|g''(x)| < M$ для $x \in (a; b)$. Тогда существует такое $\delta > 0$, что последовательность $\{x^{(k)}\}_{k=0}^{\infty}$, генерируемая итерацией $x^{(k)} =$

Вопрос №46

1. Квадратурные формулы (определение). Метод вывода формул Ньютона–Котеса. Примеры формул Ньютона–Котеса. Вычислительная устойчивость/неустойчивость численного интегрирования.

Аппроксимацию интеграла называют квадратурой: $\int_a^b f(x)dx \approx \sum_{i=1}^n a_i f(x_i), a_i \in \mathbb{R}$

Рассмотрим функцию $f(x) \in C^n[a; b]$ и n различных узлов x_1, \dots, x_n . Тогда, раскладывая $f(x)$ в базисные многочлены Лагранжа, мы получаем

$$f(x) = \sum_{i=1}^n f(x_i)l_i(x) + \frac{\prod_{i=1}^n (x - x_i)}{n!} f^{(n)}(\xi(x)),$$

, где $\xi(x) \in (a; b)$. Тогда интегрирование $f(x)$ на интервале $[a; b]$ дает:

$$\int_a^b f(x)dx = \int_a^b \sum_{i=1}^n [f(x_i)l_i(x)]dx + \int_a^b \frac{\prod_{i=1}^n (x - x_i)}{n!} f^{(n)}(\xi(x))dx = \sum_{i=1}^n \left[f(x_i) \int_a^b l_i(x)dx \right] + \int_a^b \frac{\prod_{i=1}^n (x - x_i)}{n!} f^{(n)}(\xi(x))dx,$$

что дает выражение для коэффициентов квадратуры $a_i: \int_a^b l_i(x)dx$. В случае, когда интерполяционные узлы x_1, \dots, x_n распределены равномерно, мы получаем формулы Ньютона-Котеса для численного интегрирования.

Формула трапеций ($n = 2$):

$$\int_a^b f(x)dx = \frac{h}{2} [f(x_1) + f(x_2)] - \frac{h^3}{12} f''(\xi), \text{ где}$$

$$x_1 = a, x_2 = b \text{ и } h = b - a.$$

Формула Симпсона ($n = 3$):

$$\int_a^b f(x)dx = \frac{h}{3} [f(x_1) + 4f(x_2) + f(x_3)] - \frac{h^5}{90} f^{(4)}(\xi),$$

где $x_1 = a, x_2 = x_1 + h$ и $x_3 = x_1 + 2h = b$, $h = \frac{b-a}{2}$

По сравнению с дифференцированием, операция интегрирования способна к стабилизации вычислительной погрешности. Интуитивное объяснение этого эффект заключается в том, что интегрирование предполагает суммирование близких значений, в то время как дифференцирование вычисляет их разность. Численное интегрирование устойчиво с вычислительной точки зрения, и полная погрешность полностью соответствует остаточному члену вплоть до тех пор, пока ϵ не достигнет машинного эпсилон. Полная погрешность интегрирования падает до тех пор, пока она не достигнет значения, сравнимого с машинным эпсилон, после чего уменьшение погрешности становится невозможным, и она стабилизируется на уровне машинного эпсилон.

2. Вывод формулы трапеций. Вывод формулы средних. Сравнение формулы средних с формулой трапеций.

Так, для $n = 2$ выражение принимает вид:

$$\int_a^b f(x)dx = f(x_1) \int_{x_1}^{x_2} \frac{(x - x_2)}{(x_1 - x_2)} dx + f(x_2) \int_{x_1}^{x_2} \frac{(x - x_1)}{(x_2 - x_1)} dx + \int_{x_1}^{x_2} \frac{(x - x_1)(x - x_2)}{2} f''(\xi(x))dx,$$

$$f(x_1 - h) = f(x_1) - hf'(x_1) + \frac{h^2}{2}f''(x_1) - \frac{h^3}{6}f'''(x_1) + \frac{h^4}{24}f^{(4)}(\xi_1),$$

$$f(x_1 + h) = f(x_1) + hf'(x_1) + \frac{h^2}{2}f''(x_1) + \frac{h^3}{6}f'''(x_1) + \frac{h^4}{24}f^{(4)}(\xi_2),$$

где $\xi_1 \in (x_1 - h; x_1)$ и $\xi_2 \in (x_1; x_1 + h)$. Сложив два равенства, получаем (1-ое ур-е, а потом следствие и 2-е):

$$f(x_1 - h) - 2f(x_1) + f(x_1 + h) = h^2 f''(x_1) + \frac{h^4}{24} [f^{(4)}(\xi_1) + f^{(4)}(\xi_2)]$$

$$f''(x_1) = \frac{f(x_1 - h) - 2f(x_1) + f(x_1 + h)}{h^2} - \frac{h^2}{24} [f^{(4)}(\xi_1) + f^{(4)}(\xi_2)].$$

Предположим, что $f(x) \in C^4[x_1 - h; x_1 + h]$. Тогда по теореме о промежуточном значении существует такое $\xi \in (x_1 - h; x_1 + h)$, что $f^{(4)}(\xi) = \frac{1}{2} [f^{(4)}(\xi_1) + f^{(4)}(\xi_2)]$

что в результате дает формулу численного дифференцирования второго порядка для второй производной:

$$f''(x_1) = \frac{f(x_1 - h) - 2f(x_1) + f(x_1 + h)}{h^2} - \frac{h^2}{12} f^{(4)}(\xi),$$

, где $\xi \in (x_1 - h; x_1 + h)$

$g(x^{(k-1)})$, $k > 1$, сходится к x^* в области $[x^* - \delta; x^* + \delta]$ как минимум квадратически. Более того, для асимптотически больших k , верно следующее неравенство:

$$|x^{(k+1)} - x^*| < \frac{M}{2} |x^{(k)} - x^*|^2.$$

Выберем такое $\delta > 0$ и $\gamma \in (0; 1)$, что $|g'(x)| \leq \gamma$ для $x \in [x^* - \delta; x^* + \delta] \subset [a; b]$. Последовательность $\{x^{(k)}\}_{k=0}^\infty$ для $x^{(0)} \in [x^* - \delta; x^* + \delta]$ будем так же содержать в $[x^* - \delta; x^* + \delta]$. Разложим $g(x)$ в ряд Тейлора в т. x^* :

$$g(x) = g(x^*) + g'(x^*)(x - x^*) + \frac{g''(\xi)}{2}(x - x^*)^2$$

где $x \in [x^* - \delta; x^* + \delta]$ и, следовательно, $\xi \in [x^*; x]$ или $\xi \in [x; x^*]$. Используя установки теоремы, а именно $g(x^*) = x^*$ и $(g'(x^*) = 0$, мы имеем:

$$g(x) = x^* + \frac{g''(\xi)}{2}(x - x^*)^2$$

Тогда для $x = x^{(k)}$ мы получаем:

$$g(x^{(k)}) = x^* + \frac{g''(\xi_k)}{2}(x^{(k)} - x^*)^2$$

$$\Rightarrow x^{(k+1)} - x^* = \frac{g''(\xi_k)}{2}(x^{(k)} - x^*)^2.$$

Так как $|g'(x)| \leq \gamma$ для $x \in [x^* - \delta; x^* + \delta]$, последовательность $\{x^{(k)}\}_{k=0}^\infty$ сходится к $x^* \Rightarrow \{\xi^{(k)}\}_{k=0}^\infty$ сходится к $x^* \Rightarrow \lim_{k \rightarrow \infty} g''(\xi_k) = g''(x^*)$. Тогда

имеем:

$$\lim_{k \rightarrow \infty} g''(\xi_k) = g''(x^*)$$

$$\Rightarrow \lim_{k \rightarrow \infty} \frac{x^{(k+1)} - x^*}{(x^{(k)} - x^*)^2} = \frac{g''(x^*)}{2}$$

$$\Rightarrow \lim_{k \rightarrow \infty} \frac{|x^{(k+1)} - x^*|}{|x^{(k)} - x^*|^2} = \frac{|g''(x^*)|}{2},$$

что доказывает квадратичную сходимость.

что после интегрирования линейных полиномов дает:

$$\int_a^b f(x)dx = \frac{h}{2} [f(x_1) + f(x_2)] + \int_{x_1}^{x_2} \frac{(x - x_1)(x - x_2)}{2} f''(\xi(x))dx.$$

Так как полином внутри подынтегрального выражения не меняет знак для $x \in [a; b]$, первая теорема о среднем значении позволяет записать выражение

$$\int_a^b f(x)dx = \frac{h}{2} [f(x_1) + f(x_2)] + f''(\xi) \int_{x_1}^{x_2} \frac{(x - x_1)(x - x_2)}{2} dx,$$

где $\xi \in (a; b)$. Таким образом мы получаем выражение для формулы трапеций:

$$\int_a^b f(x)dx = \frac{h}{2} [f(x_1) + f(x_2)] - \frac{h^3}{12} f''(\xi),$$

где $x_1 = a, x_2 = b$ и $h = b - a$.

Рассмотрим случай, когда на отрезке $[a; b]$ мы имеем только один узел, расположенный в центре отрезка, т.е. $x_1 = \frac{(a+b)}{2}$. Разложим функцию $f(x) \in C^2[a; b]$ в ряд Тейлора в этом узле

$$f(x) = f(x_1) + f'(x_1)(x - x_1) + \frac{f''(\xi(x))}{2}(x - x_1)^2.$$

Тогда интеграл от $f(x)$ вычисляется следующим образом:

$$\int_a^b f(x)dx = \left[f(x_1)(x - x_1) + \frac{1}{2} f'(x_1)(x - x_1)^2 \right]_a^b + \int_a^b \frac{f''(\xi(x))}{2}(x - x_1)^2 dx.$$

Так как выражение $(x - x_1)^2$ всюду неотрицательно, мы можем воспользоваться первой теоремой о среднем значении, что дает:

$$\int_a^b f(x)dx = f(x_1)(b - a) + \frac{f''(\xi)}{2} \int_a^b (x - x_1)^2 dx,$$

что после интегрирования приводит к формуле средних:

$$\int_a^b f(x)dx = f(x_1)(b - a) + \frac{f''(\xi)}{24}(b - a)^3.$$

Заметим, что остаточный член в формуле средних пропорционален $O(h^3)$, где $h = b - a$, т.е. имеет тот же порядок погрешности, что и формула трапеций. Однако формула средних использует только один узел и, соответственно, только одно значение функции $f(x)$, что делает ее предпочтительной с вычислительной точки зрения.

Вопрос №47

1. Метод Ньютона (вывод для одномерного случая). Особенности сходимости метода Ньютона.

Пусть нам дана функция $f(x) \in C^2[a; b]$ и точка $x^{(0)} \in [a; b]$ является приближением к корню x^* функции $f(x)$, т.е. $f(x^*) = 0$. Разложим функцию $f(x)$ в ряд Тейлора в точке $x^{(0)}$:

$$f(x) = f(x^{(0)}) + f'(x^{(0)})(x - x^{(0)}) + \frac{f''(\xi)}{2}(x - x^{(0)})^2,$$

где $\xi \in (x, x^{(0)})$ (либо $\xi \in (x^{(0)}, x)$). Метод Ньютона основан на предположении, что $|x^* - x^{(0)}|$ достаточно мало. Тогда, вычислив ряд в точке x^* , мы можем отбросить квадратичный член:

$$f(x^*) \approx f(x^{(0)}) + f'(x^{(0)})(x^* - x^{(0)}).$$

Так как $f(x^*) = 0$, мы получаем оценку для x^* , т.е. более точное приближение:

$$0 \approx f(x^{(0)}) + f'(x^{(0)})(x^* - x^{(0)})$$

$$\Rightarrow x^* \approx x^{(0)} - \frac{f(x^{(0)})}{f'(x^{(0)})}.$$

Таким образом, можно сформулировать итерационный метод, называемый методом Ньютона:

$$x^{(k)} = x^{(k-1)} - \frac{f(x^{(k-1)})}{f'(x^{(k-1)})}.$$

Особенности метода Ньютона:

- Начальное приближение $x^{(0)}$ должно быть достаточно близким к корню x^* (существуют ограничения на сходимость для произвольных $x^{(0)}$);
 - Градиент $f'(x^{(k)})$ должен быть отличен от нуля для любого k . Более того, метод Ньютона тем эффективнее, чем дальше $f'(x^*)$ от нуля. Существует теорема, доказывающая существование некоторой окрестности x^* , внутри которой сходимость метода Ньютона является безусловной;
 - (Метод Ньютона является “оптимальным” методом простой итерации и имеет квадратичную сходимость.)
2. Вывод метода Ньютона как метод простой итерации с квадратичной сходимостью для многомерного случая.

Метод Ньютона для систем нелинейных алгебраических уравнений:

$$x^{(k)} = g(x^{(k)}) = x^{(k-1)} - J^{-1}(x^{(k-1)})f(x^{(k-1)}).$$

На практике обратная от матрицы Якоби не вычисляется, вместо чего

Вопрос №48

1. Квадратурные формулы (определение). Составные формулы численного интегрирования (алгоритм их построения). Зависимость порядка точности от гладкости.

Аппроксимацию интеграла называют квадратурой: $\int_a^b f(x)dx \approx \sum_{i=1}^n a_i f(x_i)$, $a_i \in \mathbb{R}$

Алгоритм построения составных формул:

(а) Рассмотрим интеграл $\int_a^b f(x)dx$, где $f(x) \in C^k[a, b]$ (k зависит от выбранной формулы численного интегрирования).

(б) Разделим отрезок $[a; b]$ на некоторое число подотрезков n , тогда $h = \frac{b-a}{n}$ – шаг интегрирования.

(с) Применим одну из формул численного интегрирования на каждой паре подотрезков и получим составную формулу численного интегрирования.

Порядок точности квадратуры зависит от гладкости, если функция недостаточно гладкая, погрешность увеличивается.

2. Вывод составной формулы Симпсона. Доказательство вычислительной устойчивости/неустойчивости интегрирования на примере составной формулы Симпсона.

Рассмотрим интеграл $\int_a^b f(x)dx$. Разделим отрезок $[a; b]$ на четное число подотрезков n и применим формулу трапеций на каждом из них. Тогда для $f(x) \in C^4[a; b]$, $h = \frac{b-a}{n}$ и $x_i = a + (i-1)h$, где $i = 1, \dots, n+1$ мы имеем: $\int_a^b f(x)dx =$

$$\sum_{i=1}^{\frac{n}{2}} \int_{x_{2i-1}}^{x_{2i+1}} f(x)dx = \frac{h}{3} \sum_{i=1}^{\frac{n}{2}} [f(x_{2i-1}) + 4f(x_{2i}) + f(x_{2i+1})] - \frac{h^5}{90} \sum_{i=1}^{\frac{n}{2}} f^{(5)}(\xi_i)$$

где $\xi_i \in (x_{2i-1}; x_{2i+1})$. Заметим, что все нечетные узлы за исключением x_1 и x_{n+1} дважды повторяются в сумме, что позволяет упростить выражение:

$$\int_a^b f(x)dx = \frac{h}{3} \left[f(x_1) + 2 \sum_{i=1}^{\frac{n}{2}-1} f(x_{2i+1}) + 4 \sum_{i=1}^{\frac{n}{2}} f(x_{2i}) + f(x_{n+1}) \right] - \frac{h^5}{90} \sum_{i=1}^{\frac{n}{2}} f^{(4)}(\xi_i)$$

Для упрощения выражения для остаточного члена заметим, что так как $f(x) \in C^4[a; b]$, мы имеем:

$$\min_{x \in [a; b]} f^{(4)}(x) \leq f^{(4)}(\xi_i) \leq \max_{x \in [a; b]} f^{(4)}(x) \Rightarrow$$

Вопрос №36

1. Сходимость метода (определение). Скорость сходимости метода Ньютона. Скорость сходимости методов простой итерации.

Пусть последовательность $\{x^{(k)}\}_{k=0}^{\infty}$ сходится к x^* . Тогда если существуют такие $\lambda, \alpha \in \mathbb{R}$, что:

$$\lim_{k \rightarrow \infty} \frac{|x^{(k+1)} - x^*|}{|x^{(k)} - x^*|^\alpha} = \lambda,$$

то метод, генерирующий данную последовательность, обладает сходимостью степени α . При $\alpha = 1, \lambda \in (0; 1)$ сходимость называется линейной, при $\alpha = 1, \lambda = 0$ сверхлинейной и при $\alpha = 2$ квадратичной.

Пусть $D = [a; b]$, $g \in C(D)$, $g(D) \subset D$ и существует производная $g'(x)$ для $x \in (a; b)$ с таким $\gamma \in (0; 1)$, что $|g'(x)| \leq \gamma$ для любого $x \in (a; b)$. Если $g'(x^*) \neq 0$, то для любого $x^{(0)} \in D$ последовательность $\{x^{(k)}\}_{k=0}^{\infty}$, сгенерированная итерацией $x^{(k)} = g(x^{(k-1)})$, сходится линейно к единственной неподвижной точке $x^* \in [a; b]$. \Rightarrow Сходящийся метод простой итерации в общем случае сходится только линейно.

Рассмотренная теорема предполагает, что сходимость быстрее линейной возможна только при $g'(x^*) = 0$. Тогда для $g(x) = x - \Omega(x)f(x)$:

$$1 - \Omega'(x^*)f(x^*) - \Omega(x^*)f'(x^*) = 0$$

$$1 - \Omega(x^*)f'(x^*) = 0$$

$$\Omega(x^*) = \frac{1}{f'(x^*)}.$$

Выбрав $\Omega(x^*) = \frac{1}{f'(x^*)}$ получим метод Ньютона:

$$x^{(k)} = g(x^{(k)}) = x^{(k-1)} - \frac{f(x^{(k-1)})}{f'(x^{(k-1)})}.$$

\Rightarrow Метод Ньютона является “оптимальным” методом простой итерации и имеет квадратичную сходимость.

2. Теорема о линейной сходимости метода простой итерации для нелинейных алгебраических систем (доказательство).

Так как $|g'(x)| < 1$ для любого $x \in (a; b)$, по теореме существует единственная неподвижная точка в D . По теореме Лагранжа о среднем значении для любого $k > 1$ существует такое $\xi(k) \in D$, что:

$$g'(\xi(k)) = \frac{g(x^{(k-1)}) - g(x^*)}{x^{(k-1)} - x^*} = \frac{x^{(k)} - x^*}{x^{(k-1)} - x^*}$$

$$\Rightarrow |x^{(k)} - x^*| = |g'(\xi(k))| \cdot |x^{(k-1)} - x^*|$$

$$\Rightarrow |x^{(k)} - x^*| \leq \gamma |x^{(k-1)} - x^*|.$$

Применив тот же подход для нахождения верхней границы для $|x^{(k-1)} - x^*|, |x^{(k-2)} - x^*|$ и т.д., получаем:

$$|x^{(k)} - x^*| \leq \gamma |x^{(k-1)} - x^*| \leq \gamma^2 |x^{(k-2)} - x^*|$$

...

$$\leq \gamma^k |x^{(0)} - x^*|.$$

Тогда предел $\lim_{k \rightarrow \infty} |x^{(k)} - x^*|$ принимает вид:

$$\lim_{k \rightarrow \infty} |x^{(k)} - x^*| \leq \lim_{k \rightarrow \infty} \gamma^k |x^{(0)} - x^*| = 0,$$

используется следующая двухшаговая процедура:

$$J(x^{(k-1)})y^{(k-1)} = f(x^{(k-1)}),$$

$$x^{(k)} = x^{(k-1)} - y^{(k-1)},$$

где $y^{(k-1)}$ находится через решение первого уравнения.

$g(x) = x - \Omega(x)f(x)$ (возьмем покомпонентно)

$$g_i(x) = x_i - \sum_{j=1}^n \omega_{ij}(x) f_j(x).$$

Для построения метода простой итерации с квадратичной сходимостью нам необходимо обнулить производные функций $g_i(x)$ в точке x^* . Найдем для начала соответствующие производные:

$$\frac{\partial g_i(x)}{\partial x_j} = \begin{cases} 1 - \sum_{j=1}^n \frac{\partial \omega_{ij}(x)}{\partial x_j} f_j(x) - \sum_{j=1}^n \omega_{ij}(x) \frac{\partial f_j(x)}{\partial x_j}, & i = j, \\ - \sum_{j=1}^n \frac{\partial \omega_{ij}(x)}{\partial x_j} f_j(x) - \sum_{j=1}^n \omega_{ij}(x) \frac{\partial f_j(x)}{\partial x_j}, & i \neq j. \end{cases}$$

Тогда обнуление производных дает уравнения:

$$\sum_{j=1}^n \omega_{ij}(x^*) \frac{\partial f_j(x^*)}{\partial x_j} = 1, \quad i = j,$$

$$\sum_{j=1}^n \omega_{ij}(x^*) \frac{\partial f_j(x^*)}{\partial x_j} = 0, \quad i \neq j.$$

где мы использовали факт $g_i(x^*) = 0$. Легко убедиться, что в матричном виде это эквивалентно следующему выражению:

$$\Omega(x^*)J(x^*) = E$$

где матрица J называется матрицей Якоби и имеет следующую форму:

$$J(x) = \begin{bmatrix} \frac{\partial f_1(x)}{\partial x_1} & \dots & \frac{\partial f_1(x)}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_n(x)}{\partial x_1} & \dots & \frac{\partial f_n(x)}{\partial x_n} \end{bmatrix}$$

Таким образом, для квадратичной сходимости в качестве $\Omega(x)$ мы можем выбрать обратную матрицу Якоби: $\Omega(x) = J^{-1}(x)$,

что дает формулировку метода Ньютона для систем нелинейных алгебраических уравнений:

$$x^{(k)} = g(x^{(k)}) = x^{(k-1)} - J^{-1}(x^{(k-1)})f(x^{(k-1)}).$$

$$\frac{n}{2} \min_{x \in [a;b]} f^{(4)}(x) \leq \sum_{i=1}^{\frac{n}{2}} f^{(4)}(\xi_i) \leq \frac{n}{2} \max_{x \in [a;b]} f^{(4)}(x) = >$$

$$\min_{x \in [a;b]} f^{(4)}(x) \leq \frac{2}{n} \sum_{i=1}^{\frac{n}{2}} f^{(4)}(\xi_i) \leq \max_{x \in [a;b]} f^{(4)}(x)$$

$$f^{(4)}(\xi) = \frac{2}{n} \sum_{i=2}^{\frac{n}{2}} f^{(4)}(\xi_i), \text{ где } \xi \in (a; b).$$

$$R = -\frac{h^5}{90} \sum_{i=1}^{\frac{n}{2}} f^{(4)}(\xi_i) = -\frac{(b-a)h^4}{180} f^{(4)}(\xi).$$

Тогда составная формула Симпсона:

$$\int_a^b f(x) dx = \frac{h}{3} \left[f(x_1) + 2 \sum_{i=1}^{\frac{n}{2}-1} f(x_{2i+1}) + 4 \sum_{i=1}^{\frac{n}{2}} f(x_{2i}) + f(x_{n+1}) \right] - \frac{(b-a)h^4}{180} f^{(4)}(\xi)$$

По сравнению с дифференцированием, операция интегрирования, как мы вскоре покажем, способна к стабилизации вычислительной погрешности. Интуитивное объяснение этого эффект заключается в том, что интегрирование предполагает суммирование близких значений, в то время как дифференцирование вычисляет их разность.

Рассмотрим составную формулу Симпсона и предположим, что значение $f(x)$ в точке x_i вычисляется с погрешностью округления e_i :

$$f(x_i) = \tilde{f}(x_i) + e_i, \quad i = 1, \dots, n+1.$$

Тогда полная погрешность округления, накапливаемая составной формулой Симпсона, может быть оценена следующим образом:

$$e(h) = \frac{h}{3} \left[e_1 + 2 \sum_{i=1}^{n/2-1} e_{2i+1} + 4 \sum_{i=1}^{n/2} e_{2i} + e_{n+1} \right] \leq \frac{h}{3} \left[|e_1| + 2 \sum_{i=1}^{n/2-1} |e_{2i+1}| + 4 \sum_{i=1}^{n/2} |e_{2i}| + |e_{n+1}| \right].$$

Предположим, что погрешность округления ограничена, например, машинным эпсилон: $|e_i| \leq \epsilon, i = 1, \dots, n+1$. Тогда полная погрешность оценивается как:

$$e(h) \leq \frac{h\epsilon}{3} \left[1 + 2 \left(\frac{n}{2} - 1 \right) + 4 \frac{n}{2} + 1 \right] = nh\epsilon = (b-a)\epsilon.$$

Этот результат показывает, что верхняя грань для накопленной погрешности округления не зависит от n или h , что означает, что увеличение числа подотрезков не приводит к дестабилизации полной погрешности. Полная погрешность интегрирования падает до тех пор, пока она не достигнет значения, сравнимого с машинным эпсилон, после чего уменьшение погрешности становится невозможным, и она стабилизируется на уровне машинного эпсилон.

что эквивалентно сходимости метода простой итерации.

Вопрос №37

- Задача минимизации в нелинейной регрессии. Полиномиальная регрессия как частный случай.

Рассмотрим аппроксимирующую функцию $f: \mathbb{R}^n \rightarrow \mathbb{R}$, заданную следующим образом:

$f(a; x) = \sum_{i=0}^n a_i \phi_i(x)$, где $x \in \mathbb{R}^l$, $\phi_0(x) = 1$ и $\phi_n(x)$, $i = 1, \dots, n$ в общем случае нелинейные функции. Несложно заметить, что в таком случае задача минимизации может быть записана в матричном виде, где матрица $X \in \mathbb{R}^{m \times (n+1)}$ определена как:

$$X = \begin{bmatrix} 1 & \phi_1(x^{(1)}) & \dots & \phi_n(x^{(1)}) \\ 1 & \phi_1(x^{(2)}) & \dots & \phi_n(x^{(2)}) \\ \vdots & \vdots & \dots & \vdots \\ 1 & \phi_1(x^{(m)}) & \dots & \phi_n(x^{(m)}) \end{bmatrix}$$

где $x^{(i)}$, $i = 1, \dots, m$ – дискретные многомерные данные, к которым приближается функция $f(a; x)$. Тогда оптимальные значения параметров a определяются с помощью нормального уравнения:

$$a = (X^T X)^{-1} X^T y$$

Подобный метод может также интерпретироваться как линейная многомерная регрессия в пространстве с нелинейно трансформированными координатами.

Заметим, что повторное использование нормального уравнения возможно благодаря линейности $f(a; x)$ относительно вектора параметров a , несмотря на то, что функция нелинейна относительно x . В случае, когда $f(a; x)$ нелинейна относительно a , требуется пересчет соответствующих производных, что в общем случае будет приводить к более сложным и, вероятно, нелинейным уравнениям для определения оптимальных a .

В качестве примера рассмотрим полиномиальную регрессию в одномерном пространстве, где $f(a; x) = \sum_{i=0}^n a_i x^{(i)}$. Тогда матрица X принимает вид:

$$X = \begin{bmatrix} 1 & x_{(1)} & \dots & x_{(1)}^n \\ 1 & x_{(2)} & \dots & x_{(2)}^n \\ \vdots & \vdots & \dots & \vdots \\ 1 & x_{(m)} & \dots & x_{(m)}^n \end{bmatrix},$$

Так как X является матрицей Вандермонда, ее определитель всегда отличен от нуля, если среди дискретных данных нет повторяющихся. Это доказывает, что при удовлетворении этого условия нормальное уравнение (4.18) всегда будет иметь единственное и нетривиальное решение.

Вопрос №25

- Сходящиеся матрицы (определение). Теорема о спектральном радиусе сходящихся матриц (формулировка).

Матрица $A \in \mathbb{R}^{n \times n}$ называется сходящейся, если для нее верно: $\lim_{k \rightarrow \infty} A^k = \bar{O}$, где \bar{O} – нулевая матрица.

Теорема. Матрица $A \in \mathbb{R}^{n \times n}$ является сходящейся тогда и только тогда, когда $\rho(A) < 1$.

- Теорема о связи спектрального радиуса с нормами матрицы (доказательство).

Пусть $A \in \mathbb{R}^{n \times n}$. Тогда $\rho(A) \leq \|A\|$ для любой индуцированной матричной нормы $\|\cdot\|$.

Доказательство. Рассмотрим собственное число λ с собственным вектором x . Тогда имеем:

$$Ax = \lambda x$$

$$\Rightarrow \|Ax\| = \|\lambda x\|$$

$$\Rightarrow \|Ax\| = |\lambda| \cdot \|x\|.$$

Так как $\|Ax\| \leq \|A\| \cdot \|x\|$, мы получаем:

$$|\lambda| \cdot \|x\| \leq \|A\| \cdot \|x\|$$

$$\Rightarrow |\lambda| \leq \|A\|,$$

что приводит к $\rho(A) = \max |\lambda| \leq \|A\|$

Вопрос №6

- Особенности интерполяции Лагранжа по равномерно распределенным узлам. Оптимальная интерполяция Лагранжа.

Базисные полиномы при увеличении их степени (т.е. при увеличении количества узлов) имеют тенденцию к росту амплитуды ближе к граничным узлам отрезка. Чем выше степень базисного полинома, тем более заметным становится этот эффект. Подобное поведение может привести к появлению нежелательных, паразитных осцилляций у граничных узлов. Интуитивно можно заключить, что использование неравномерно распределенных узлов, концентрирующихся у границ отрезка, потенциально могло бы решить проблему. Такой выбор узлов действительно является оптимальным с точки зрения минимизации ошибки интерполирования.

- Вывод интерполяционного многочлена Лагранжа.

Иными словами мы хотим построить такую функцию $\varphi_i(x)$, что она равна 1 только в узле x_i и обращается в ноль во всех остальных узлах. Тогда искомый интерполяционный многочлен будет вычисляться как $\tilde{f}_n(x) = \sum_{i=1}^n f(x_i) \varphi_i(x)$. Действительно, легко убедиться, что в интерполяционных узлах значения функций совпадают:

$$\tilde{f}_n(x_j) = \sum_{i=1}^n f(x_i) \varphi_i(x_j) = \sum_{i=1}^n f(x_i) \delta_{ij} = f(x_j).$$

Зная, что $\varphi_i(x_i) = 1$ при $i \neq j$, следующий полином удовлетворяет указанному требованию:

$$\varphi_i(x) = C_i \prod_{i \neq j} (x - x_j),$$

где C_i – некоторая константа. Определить неизвестную C_i можно из условия

$$\varphi_i(x_i) = 1$$

$$\Rightarrow C_i \prod_{i \neq j} (x_i - x_j) = 1$$

$$\Rightarrow C_i = \frac{1}{\prod_{i \neq j} (x_i - x_j)},$$

что в результате дает многочлен, известный как интерполяционный многочлен Лагранжа:

$$\tilde{f}_n(x_j) = L_{n-1}(x) = \sum_{i=1}^n f(x_i) \prod_{i \neq j} \frac{x - x_j}{x_i - x_j}.$$

2. Метод наименьших квадратов для приближения к непрерывной функции. Решение задачи минимизации для случая приближения с помощью ортогональных функций.

Пусть аппроксимирующая функция задана так же, как и в случае нелинейной регрессии: $f(a; x) = \sum_{i=0}^n a_i \phi_i(x)$. Метод наименьших квадратов в случае приближения к непрерывной функции трансформируется в интеграл от квадрата отклонения:

$$\begin{aligned} \min_a E_2(a) &= \min_a \int_a^b [y(x) - f(x)]^2 dx \\ &= \min_a \int_a^b \left[y(x) - \sum_{i=0}^n a_i \phi_i(x) \right]^2 dx. \end{aligned}$$

В качестве системы функций $\{\phi_i(x)\}_{i=0}^n$ удобно выбрать ортогональную систему. . Так как многие системы функций являются ортогональными только с весом, активно используется модифицированный метод, называемый взвешенным методом наименьших квадратов:

$$\min_a E_2(a) = \min_a \int_a^b \omega(x) \left[y(x) - \sum_{i=0}^n a_i \phi_i(x) \right]^2 dx.$$

Предположим, что $\{\phi_i(x)\}_{i=0}^n$ является ортогональной системой на $[a; b]$ с весом $\omega(x)$. Для нахождения наименьшего значения функции $E_2(a)$ необходимо найти нули производной:

$$\begin{aligned} \frac{\partial E_2}{\partial a_k} &= -2 \int_a^b \omega(x) \phi_k(x) \left[y(x) - \sum_{i=0}^n a_i \phi_i(x) \right] dx \\ \Rightarrow \int_a^b \omega(x) \phi_k(x) \left[y(x) - \sum_{i=0}^n a_i \phi_i(x) \right] dx &= 0. \end{aligned}$$

Пользуясь свойством ортогональности функций, мы получаем:

$$\int_a^b [\omega(x) y(x) \phi_k(x) - a_k \omega(x) \phi_k^2(x)] dx = 0.$$

Тогда оптимальные значения a_k находятся с помощью выражения:

$$a_k = \frac{\langle y(x), \phi_k(x) \rangle_\omega}{\langle \phi_k(x), \phi_k(x) \rangle_\omega}.$$

Вопрос №17

- Нормальные ОДУ n -го порядка. Сведение нормального ОДУ n -го порядка к системе ОДУ 1-го порядка. Задача Коши для системы ОДУ 1-го порядка.

Нормальные ОДУ n -го порядка:
 $y^{(n)}(t) = f(t, y, y', y'', \dots, y^{(n-1)})$

Система ОДУ 1-го порядка:

$$\frac{d}{dt} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_{n-1} \\ y_n \end{bmatrix} = \begin{bmatrix} y_2 \\ y_3 \\ \vdots \\ y_n \\ f(t, y_1, y_2, y_3, \dots, y_n) \end{bmatrix},$$

где $y_1 = y, y_2 = y', y^{(n)} = y^{(n-1)}$.

Задача Коши для системы ОДУ 1-го порядка: – задача Коши предполагает знание начальных условий (своего рода предсказание будущего)

$$\frac{d}{dt} \begin{bmatrix} y_1(t) \\ y_2(t) \\ \vdots \\ y_{n-1}(t) \\ y_n(t) \end{bmatrix} = \begin{bmatrix} f_1(t, y_1, \dots, y_n) \\ f_2(t, y_1, \dots, y_n) \\ \vdots \\ f_{n-1}(t, y_1, \dots, y_n) \\ f_n(t, y_1, \dots, y_n) \end{bmatrix},$$

где $y_1(a) = \alpha_1, \dots, y_n(a) = \alpha_n, t \in [a; b]$

- Теорема о верхней границе глобальной погрешности метода Эйлера (доказательство).

Пусть функция $f(t, y)$ является липшиц-непрерывной в $D = \{(t, y) \mid t \in [a; b], y \in \mathbb{R}\}$ с константой Липшица L . Пусть существует такое $M > 0$, что $|y''(t)| < M$ для любого $t \in [a; b]$, где $y(t)$ является единственным решением задачи Коши $y' = f(t, y), y(a) = \alpha$. Тогда для последовательности $\{w_i\}_{i=0}^m$, сгенерированной методом Эйлера, верно следующее неравенство:

$$|y(t_i) - w_i| \leq \frac{hM}{2L} (e^{L(t_i-a)} - 1), \quad i = 0, 1, \dots, m,$$

где $h = \frac{b-a}{m}$.

Доказательство:

Из вывода метода Эйлера мы имеем:

$$y_{i+1} = y_i + hf(t_i, y_i) + \frac{h^2}{2} y''(\xi_i),$$

где $y_i = y(t_i)$. Тогда выражение для $y_{i+1} - w_{i+1}$ имеет вид:

$$y_{i+1} - w_{i+1} = y_i - w_i + h[f(t_i, y_i) - f(t_i, w_i)] + \frac{h^2}{2} y''(\xi_i), \quad \text{из чего следует неравенство:}$$

$$|y_{i+1} - w_{i+1}| \leq |y_i - w_i| + h|f(t_i, y_i) - f(t_i, w_i)| +$$

$$\frac{h^2}{2} |y''(\xi_i)|. \quad \text{Так как функция } f(t, y) \text{ является липшиц-непрерывной, мы имеем:}$$

$$|y_{i+1} - w_{i+1}| \leq (1 + hL)|y_i - w_i| + \frac{h^2}{2} |y''(\xi_i)|.$$

Более того, в силу ограниченности $y''(t)$ мы получаем:

Вопрос №49

- Методы вывода формул численного дифференцирования. Формула численного дифференцирования первого порядка (вывод через разложение в ряд Тейлора). Вычислительная устойчивость/неустойчивость численного дифференцирования.

Так или иначе, все формулы численного дифференцирования предполагают аппроксимацию известной функции по значениям в узлах.

Основные методы:

- Определение производной;
- Интерполяция;
- Разложение в ряд Тейлора

Пусть нам известны значения функции $f(x)$ в точках x_1 и $x_1 + h$. Разложим данную функцию в ряд Тейлора относительно точки x_1 :
 $f(x) = f(x_1) + f'(x_1)(x - x_1) + \frac{f''(\xi)}{2} (x - x_1)^2, \xi \in (x_1; x)$.

Найдем значение ряда в точке $x_1 + h$:

$$f(x_1 + h) = f(x_1) + f'(x_1)(x_1 + h - x_1) + \frac{f''(\xi)}{2} (x_1 + h - x_1)^2$$

$$f(x_1 + h) = f(x_1) + f'(x_1)h + \frac{f''(\xi)}{2} h^2$$

$$\frac{f(x_1 + h) - f(x_1)}{h} = f'(x_1) + \frac{f''(\xi)}{2} h$$

В предположении, что h достаточно мало, отбрасываем член порядка малости $O(h)$ и получаем формулу численного дифференцирования первого порядка:

$$f'(x_1) \approx \frac{f(x_1 + h) - f(x_1)}{h} - \frac{h}{2} f''(\xi)$$

Численное дифференцирование является вычислительно НЕустойчивым методом. (там из-за вычитания близких значений накапливается погрешность)

- Вывод центральной формулы численного дифференцирования 2-го порядка для второй производной. Вывод оптимального шага дифференцирования для центральной формулы численного дифференцирования 2-го порядка.

Метод разложения в ряд Тейлора становится удобен в случае, когда требуется построить формулу для вычисления высших производных. В качестве примера построим формулу для второй производной функции $f(x)$. Для этого допустим, что нам известны ее значения в точках $x_1 -$

Вопрос №18

- Метод Эйлера (вывод). Локальная и глобальная погрешности метода (определения). Локальная и глобальная погрешность метода Эйлера.

– базовый численный метод решения задачи Коши Рассматривается ОДУ следующего вида:

$$\frac{dy}{dt} = f(t, y), \quad \text{где } t \in [a; b] \text{ и } y(a) = \alpha.$$

Дискретизируем координату t в сетку вида $t_i = a + h_i, i = 0, \dots, m - 1, h$ – шаг, что автоматически дискретизирует решение. Разложим функцию $y(x) \in C^2[a; b]$ в ряд Тейлора относительно точки t_i :

$$y(t) = y(t_i) + y'(t_i)(t - t_i) + \frac{y''(\xi)}{2} (t - t_i)^2, \xi \in (x_i; x).$$

Найдем значение ряда в точке t_{i+1} :

$$y(t_{i+1}) = y(t_i) + y'(t_i)(t_{i+1} - t_i) + \frac{y''(\xi)}{2} (t_{i+1} - t_i)^2$$

$$y(t_{i+1}) = y(t_i) + y'(t_i)h + \frac{y''(\xi)}{2} h^2$$

В предположении, что h достаточно мало, отбрасываем член порядка малости $O(h^2)$ и получаем формулу численного дифференцирования первого порядка:

$$y(t_{i+1}) \approx y(t_i) + y'(t_i)h \Rightarrow$$

$$\Rightarrow y(t_{i+1}) \approx y(t_i) + f(t_i, y(t_i))h$$

Пусть $w_i \approx y(t_i)$, тогда метод Эйлера может быть сформулирован следующим образом:

$$w_0 = \alpha$$

$$w_{i+1} = w_i + hf(t_i, w_i), \quad i = 0, \dots, m - 1.$$

Погрешности метода:

- Локальная – погрешность на одном шаге (для метода Эйлера $O(h^2)$)
 - Глобальная – погрешность, накопленная на всем интервале $[a; b]$ (для метода Эйлера $O(h)$)
- Вывод метода Рунге–Кутты 2-го порядка.

Нам необходимо найти аппроксимацию для $T^{(2)}(t, y)$ точную до члена порядка $O(h^2)$. Рассмотрим следующую аппроксимирующую функцию:

$\varphi(t, y) = \alpha_1 f(t + \alpha_1, y + \beta_1)$. с неопределенными коэффициентами $\alpha_1, \alpha_1, \beta_1$. Раскроем выражение для $T^{(2)}(t, y)$:

$$\begin{aligned} T^{(2)}(t, y) &= f(t, y) + \frac{h}{2} f'(t, y) \\ &= f(t, y) + \frac{h}{2} \left[\frac{\partial f}{\partial t}(t, y) + \frac{\partial f}{\partial y}(t, y) y'(t, y) \right] \\ &= f(t, y) + \frac{h}{2} \frac{\partial f}{\partial t}(t, y) + \frac{h}{2} \frac{\partial f}{\partial y}(t, y) f(t, y). \end{aligned}$$

1

Теперь разложим функцию $f(t, y)$ в ряд Тейлора в точке (t, y) и вычислим значение ряда, соответствующее $\varphi(t, y) = \alpha_1 f(t + \alpha_1, y + \beta_1)$:

$$\begin{aligned} \alpha_1 f(t + \alpha_1, y + \beta_1) &= \alpha_1 f(t, y) + \\ &+ \alpha_1 \alpha_1 \frac{\partial f}{\partial t}(t, y) + \alpha_1 \beta_1 \frac{\partial f}{\partial y}(t, y) + \alpha_1 R_1(t + \alpha_1, y + \beta_1). \end{aligned} \quad (2)$$

$$|y_{i+1} - w_{i+1}| \leq (1 + hL)|y_i - w_i| + \frac{h^2 M}{2}. \quad \text{Тогда применение}$$

$$\text{леммы вместе с } |y_0 - w_0| = 0 \text{ и } (i + 1)h = t_{i+1} - a \text{ дает:}$$

$$|y_{i+1} - w_{i+1}| \leq e^{(i+1)hL} \left(|y_0 - w_0| + \frac{h^2 M}{2hL} \right) - \frac{h^2 M}{2hL}$$

$$= \frac{hM}{2L} \left(e^{(i+1)hL} - 1 \right)$$

$$= \frac{hM}{2L} \left(e^{(t_{i+1} - a)L} - 1 \right).$$

h , x_1 и $x_1 + h$, и разложим ее в ряд Тейлора в точке x_1 :

$$f(x) = f(x_1) + f'(x_1)(x - x_1) + \frac{f''(x_1)}{2}(x - x_1)^2 + \frac{f'''(x_1)}{6}(x - x_1)^3 + \frac{f^{(4)}(\xi)}{24}(x - x_1)^4.$$

где $\xi \in (x_1; x)$. Тогда значения ряда в точках $x_1 - h$ и $x_1 + h$ будут равны:

$$f(x_1 - h) = f(x_1) - hf'(x_1) + \frac{h^2}{2}f''(x_1) - \frac{h^3}{6}f'''(x_1) + \frac{h^4}{24}f^{(4)}(\xi_1),$$

$$f(x_1 + h) = f(x_1) + hf'(x_1) + \frac{h^2}{2}f''(x_1) + \frac{h^3}{6}f'''(x_1) + \frac{h^4}{24}f^{(4)}(\xi_2).$$

где $\xi_1 \in (x_1 - h; x_1)$ и $\xi_2 \in (x_1; x_1 + h)$. Сложив два равенства, получаем (1-ое ур-е, а потом следствие и 2-е):

$$f(x_1 - h) - 2f(x_1) + f(x_1 + h) = h^2 f''(x_1) + \frac{h^4}{24} [f^{(4)}(\xi_1) + f^{(4)}(\xi_2)]$$

$$f''(x_1) = \frac{f(x_1 - h) - 2f(x_1) + f(x_1 + h)}{h^2} - \frac{h^2}{24} [f^{(4)}(\xi_1) + f^{(4)}(\xi_2)].$$

Предположим, что $f(x) \in C^4[x_1 - h; x_1 + h]$. Тогда по теореме о промежуточном значении существует такое $\xi \in (x_1 - h; x_1 + h)$, что $f^{(4)}(\xi) = \frac{1}{2} [f^{(4)}(\xi_1) + f^{(4)}(\xi_2)]$

что в результате дает формулу численного дифференцирования второго порядка для второй производной:

$$f''(x_1) = \frac{f(x_1 - h) - 2f(x_1) + f(x_1 + h)}{h^2} - \frac{h^2}{12} f^{(4)}(\xi),$$

, где $\xi \in (x_1 - h; x_1 + h)$

Рассмотрим формулы численного дифференцирования второго порядка:

$$f'(x_1 + h) = \frac{f(x_1 + 2h) - f(x_1)}{2h} - \frac{h^2}{6} f^{(3)}(\xi),$$

и предположим, что при округлении значений $f(x_1 + 2h)$ и $f(x_1)$ вычислительная погрешность равна $e(x_1 + 2h)$ и $e(x_1)$, то есть

$$f(x_1 + 2h) = \tilde{f}(x_1 + 2h) + e(x_1 + 2h),$$

$$f(x_1) = \tilde{f}(x_1) + e(x_1).$$

Тогда полная погрешность, включающая погрешность метода и вычислительную погрешность, вычисляется следующим образом:

$$f'(x_1 + h) - \frac{\tilde{f}(x_1 + 2h) - \tilde{f}(x_1)}{2h} = \frac{e(x_1 + 2h) - e(x_1)}{2h} - \frac{h^2}{6} f^{(3)}(\xi).$$

Пусть вычислительная погрешность ограничена ϵ (например, машинным эпсилон) и пусть $f^{(3)}$ ограничена M . Тогда верным является:

$$\left| f'(x_1 + h) - \frac{\tilde{f}(x_1 + 2h) - \tilde{f}(x_1)}{2h} \right| \leq \frac{\epsilon}{h} + \frac{h^2}{6} M.$$

Можно заметить, что при $h \rightarrow 0$ погрешность будет стремиться к бесконечности, что и обуславливает вычислительную неустойчивость численного дифференцирования. Несложно проверить, что выражение справа имеет минимум в этой точке

$h_{opt} = \left(\frac{3\epsilon}{M} \right)^{\frac{1}{3}}$ что задает оптимальный шаг дифференцирования для данной формулы

Приравнявая коэффициенты при $f(t, y)$, $\frac{\partial f}{\partial t}(t, y)$ и $\frac{\partial f}{\partial y}(t, y)$ в выражениях (1) и (2), получаем:

$$\alpha_1 = 1,$$

$$\alpha_1 = \frac{h}{2},$$

$$\alpha_2 = \frac{h}{2} f(t, y),$$

из чего следует:

$$T^{(2)}(t, y) = f\left(t + \frac{h}{2}, y + \frac{h}{2} f(t, y)\right) - R_1\left(t + \frac{h}{2}, y + \frac{h}{2} f(t, y)\right).$$

Мы ожидаем, что функция R_1 будет иметь порядок $O(h^2)$. Убедимся в этом:

$$R_1\left(t + \frac{h}{2}, y + \frac{h}{2} f(t, y)\right) = \frac{h^2}{8} \frac{\partial^2 f}{\partial t^2}(\xi, \mu) + \frac{h^2}{4} f(t, y) \frac{\partial^2 f}{\partial t \partial y}(\xi, \mu) + \frac{h^2}{8} f^2(t, y) \frac{\partial^2 f}{\partial y^2}(\xi, \mu) = O(h^2).$$

Таким образом, мы получаем формулировку метода Рунге–Кутты 2-го порядка, также называемого модифицированным методом Эйлера:

$$w_0 = \alpha,$$

$$w_{i+1} = w_i + hf(t_i + \frac{h}{2}, w_i + \frac{h}{2} f(t_i, w_i)), \quad i = 0, 1, \dots, m - 1.$$

Функция $\phi(t, y)$ была подобрана таким образом, что соответствующее разложение в ряд Тейлора включало в себя составляющие функции $T^{(2)}(t, y)$, т.е. $f(t, y)$, $\frac{\partial f}{\partial t}(t, y)$ и $\frac{\partial f}{\partial y}(t, y)$.

Вопрос №5

1. Тригонометрический интерполяционный полином (постановка задачи интерполяции). Связь с дискретным преобразованием Фурье.

Тригонометрический полином:

$$\sum_{k=-n}^n \widehat{a}_k e^{ikx} =$$

$$2 \sum_{k=0}^n [\operatorname{Re}(\widehat{a}_k) \cos kx -$$

$$\operatorname{Im}(\widehat{a}_k) \sin kx], \text{ где } \widehat{a}_k^* = \widehat{a}_{-k}$$

Тригонометрический ряд:

$$\sum_{k=0}^n (a_k \cos kx + b_k \sin kx)$$

Экспоненциальная форма:

$$\sum_{k=-n}^n \widehat{a}_k e^{ikx}$$

Доказательство:

Рассмотрим формулу Эйлера:

$$e^{ikx} = \cos kx + i \sin kx \mid * \widehat{a}_k \in \mathbb{C}$$

$$\widehat{a}_k e^{ikx} = \widehat{a}_k \cos kx + i \widehat{a}_k \sin kx =$$

$$\operatorname{Re}(\widehat{a}_k) \cos kx + i \operatorname{Im}(\widehat{a}_k) \cos kx -$$

$$\operatorname{Im}(\widehat{a}_k) \sin kx + i \operatorname{Re}(\widehat{a}_k) \sin kx, \text{ где}$$

$\operatorname{Re}(\widehat{a}_k)$ и $\operatorname{Im}(\widehat{a}_k)$ – вещественная и мнимая части \widehat{a}_k соответственно.

Добавим к обоим частям равенства комплексно сопряженное (та же формула, но там где $+i$ добавляется $-i$)

$$\widehat{a}_k e^{ikx} + \widehat{a}_k^* e^{-ikx} =$$

$$2 \operatorname{Re}(\widehat{a}_k) \cos kx - 2i \operatorname{Im}(\widehat{a}_k) \sin kx \text{ где}$$

комплексно сопряженное число обозначено через $*$. Тогда суммирование обеих частей равенства по $k=0, \dots, n$ приводит к выражению:

$$\sum_{k=0}^n (\widehat{a}_k e^{ikx} + \widehat{a}_k^* e^{-ikx}) =$$

$$\sum_{k=-n}^n \widehat{a}_k e^{ikx} =$$

$$\sum_{k=0}^n [2 \operatorname{Re}(\widehat{a}_k) \cos kx -$$

$$2i \operatorname{Im}(\widehat{a}_k) \sin kx], \text{ где } \widehat{a}_k^* = \widehat{a}_{-k}$$

Таким образом мы получили экспоненциальную форму тригонометрического ряда, где коэффициенты связаны друг с другом следующим образом:

$$a_k = 2 \operatorname{Re}(\widehat{a}_k)$$

$$b_k = -2i \operatorname{Im}(\widehat{a}_k), \text{ где } k = 0, \dots, n$$

Для того, чтобы перейти к форме полинома, обозначим $z = e^{ix}$. Тогда имеем:

$\sum_{k=-n}^n \widehat{a}_k e^{ikx} = \sum_{k=-n}^n \widehat{a}_k z^k$ что при домножении на z^n дает комплекснозначный полином степени $2n$, по конвенции называемый тригонометрическим полиномом n -й степени:

$$Q_n(z) = z^n \sum_{k=-n}^n \widehat{a}_k z^k$$

Дискретное преобразование Фурье – переход из физического пространства ($y_i = f(x_i)$) в пространсво Фурье (\widehat{a}_k).

Тригонометрический интерполяционный полином (дискретное преобразование Фурье):

Вопрос №12

1. Сингулярные числа и вектора (определение). Теорема о сингулярном разложении (формулировка).

Сингулярными числами $\sigma_1, \dots, \sigma_r$ матрицы $A \in \mathbb{R}^{m \times n}$ называются неотрицательные вещественные числа $\sigma_i = \sqrt{\lambda_i}$, где λ_i ненулевые собственные числа соответствующей матрицы Грама $K = A^T A$. Ассоциированные собственные вектора матрицы Грама K называются сингулярными векторами.

Матрица $A \in \mathbb{R}^{m \times n}$ ранга $r > 0$ может быть представлена в виде разложения $A = P \Sigma Q^T$, где $P \in \mathbb{R}^{m \times r}$ и $Q \in \mathbb{R}^{n \times r}$ – ортонормальные матрицы ($P^T P = Q^T Q = E$), а $\Sigma \in \mathbb{R}^{m \times r}$ – диагональная матрица, на диагонали которой расположены сингулярными числами $\sigma_1, \dots, \sigma_r$.

2. Псевдообратная матрица (определение). Лемма об обратной Мура–Пенроуза для матрицы $A \in \mathbb{R}^{m \times n}$ с рангом n (доказательство).

A^+ называется псевдообратной матрицей для матрицы A , если она удовлетворяет следующим критериям

$$1. AA^+A = A;$$

$$2. A^+AA^+ = A^+ \text{ (} A^+ \text{ является слабым обращением в мультипликативной полугруппе);}$$

$$3. (AA^+)^* = AA^+ \text{ (это означает, что } AA^+ \text{ — эрмитова матрица);}$$

$$4. (A^+A)^* = A^+A \text{ (} A^+A \text{ — тоже эрмитова матрица).}$$

$$A^+ = Q \Sigma^{-1} P^T$$

Лемма: Пусть матрица $A \in \mathbb{R}^{m \times n}$ имеет ранг n . Тогда $A^+ = (A^T A)^{-1} A^T$

Доказательство:

$$A^T A = (P \Sigma Q^T)^T (P \Sigma Q) =$$

$$Q \Sigma^T P^T P \Sigma Q^T = Q \Sigma^2 Q^T [rank(A) =$$

$$n \Rightarrow Q \in \mathbb{R}^{n \times n} \Rightarrow Q^{-1} = Q^T] \Rightarrow$$

$$(A^T A)^{-1} A^T =$$

$$(Q \Sigma^2 Q^T)^{-1} (P \Sigma Q^T)^T =$$

$$(Q^T)^{-1} \Sigma^{-2} Q^{-1} Q \Sigma P^T = Q \Sigma^{-1} P^T =$$

$$A^+$$

Вопрос №1

1. Этапы моделирования в прикладной математике. Источники погрешностей. Абсолютная и относительная погрешности.

Этапы моделирования в прикладной математике: Объект исследования \rightarrow Математическая модель \rightarrow Дискретная модель \rightarrow Компьютерная модель
Упрощения, осуществляемые на каждом шаге моделирования, вносят погрешность. Принято выделять четыре вида погрешностей:

- Неустраняемая погрешность – неточность при задании исходных данных.
- Погрешность математической модели – неточность при описании реального объекта математическими понятиями.
- Погрешность метода – неточность при замене математической модели приближенной.
- Вычислительная погрешность – погрешность математических операций, производимых компьютером.

Определение 1. Абсолютной погрешностью приближенного значения a^* называют величину $\Delta(a^*)$, которая определена как $\Delta(a^*) = |a - a^*|$, где a – точное значение. Число a записывают с учетом абсолютной погрешности в следующей форме: $a = a^* \pm \Delta(a^*)$.

Определение 2. Относительной погрешностью приближенного значения a^* называют величину $\delta(a^*)$, которая определена как

$$\delta(a^*) = \left| \frac{a^* - a}{a} \right|,$$

где a – точное значение. Число a записывают с учетом относительной погрешности в следующей форме:

$$a = a^* (1 \pm \delta(a^*))$$

2. Теорема о полиноме с наименьшими по модулю экстремумами (доказательство). Решение задачи о минимизации ошибки интерполяции Лагранжа.

Пусть $\widetilde{T}_k(x) = \frac{1}{2^{k-1}} T_k(x)$, где $k \geq 1$ и $T_k(x)$ является полиномом Чебышева k -й степени, и пусть \widetilde{P}_k – множество всех нормированных полиномов степени k . Тогда верным является следующее утверждение:

$$\frac{1}{2^{k-1}} = \max_{x \in [-1; 1]} |\widetilde{T}_k(x)| \leq \max_{x \in [-1; 1]} |P_k(x)| \quad \forall P_k(x) \in \widetilde{P}_k.$$

Более того, равенство верно только при $P_k(x) = \widetilde{T}_k(x)$.

Рассмотрим доказательство от обратного. Пусть верно

$$\max_{x \in [-1; 1]} |P_k(x)| \leq \max_{x \in [-1; 1]} |\widetilde{T}_k(x)| = \frac{1}{2^{k-1}} (1)$$

Введем полином $(k-1)$ -й степени $Q(x) = \widetilde{T}_k(x) - P_k(x)$. Для $k+1$ точки \widehat{x}_m мы имеем:

$$f(x) = \sum_{k=-m+1}^{m-1} \hat{a}_k e^{ikx} + \hat{a}_m \cos(mx),$$

$$2\partial \hat{a}_k = \frac{(-1)^k}{2m} \sum_{j=0}^{2m-1} y_j e^{-\frac{ijk\pi}{2}}, \quad \hat{a}_m = \frac{(-1)^k}{2m} \sum_{j=0}^{2m-1} y_j (-1)^j.$$

$$2\partial \hat{a}_k = \frac{(-1)^k}{2m} \sum_{j=0}^{2m-1} y_j e^{-\frac{ijk\pi}{2}}, \quad \hat{a}_m = \frac{(-1)^k}{2m} \sum_{j=0}^{2m-1} y_j (-1)^j.$$

2. Вывод

тригонометрического
интерполяционного
полинома. Алгоритм Кули–
Тьюки.

Для начала запишем коэффициент \hat{a}^k в следующей форме:

$$\hat{a}_k = \frac{1}{2m} \sum_{j=0}^{2m-1} y_j e^{ik\pi} e^{-i\frac{kj\pi}{m}} = \frac{(-1)^k}{2m} \sum_{j=0}^{2m-1} y_j e^{-\frac{ijk\pi}{m}}.$$

Алгоритм Кули–Тьюки вычисляет значение суммы, которую мы обозначим как:

$$A_k = \sum_{j=0}^{2m-1} y_j e^{-\frac{ijk\pi}{m}}, \quad k = 0, \dots, 2m-1.$$

Разделим A_k на две части с четными и нечетными индексами соответственно:

$$\begin{aligned} A_k &= \sum_{j=0}^{m-1} y_{2j} e^{-\frac{2j ik \pi}{m}} + \sum_{j=0}^{m-1} y_{2j+1} e^{-\frac{(2j+1) ik \pi}{m}} \\ &= \sum_{j=0}^{m-1} y_{2j} e^{-\frac{2j ik \pi}{m}} + e^{-\frac{ik \pi}{m}} \sum_{j=0}^{m-1} y_{2j+1} e^{-\frac{2j ik \pi}{m}}. \end{aligned}$$

Обозначим полученные суммы E_k и O_k :

$$E_k = \sum_{j=0}^{m-1} y_{2j} e^{-\frac{2j ik \pi}{m}},$$

$$O_k = \sum_{j=0}^{m-1} y_{2j+1} e^{-\frac{2j ik \pi}{m}}.$$

Заметим, что E_k и O_k являются периодическими относительно k и периода m :

$$E_{k \pm m} = \sum_{j=0}^{m-1} y_{2j} e^{-\frac{2j i(k \pm m) \pi}{m}} = e^{\mp 2j i \pi} \sum_{j=0}^{m-1} y_{2j} e^{-\frac{2j ik \pi}{m}} = E_k,$$

$$O_{k \pm m} = \sum_{j=0}^{m-1} y_{2j+1} e^{-\frac{2j i(k \pm m) \pi}{m}} = e^{\mp 2j i \pi} \sum_{j=0}^{m-1} y_{2j+1} e^{-\frac{2j ik \pi}{m}} = O_k.$$

Это свойство позволяет посчитать E_k и O_k только для $k = 0, \dots, m$, так как E_k и O_k для $k = m, \dots, 2m-1$ будут иметь те же значения. Тогда A_k для $k = 0, \dots, 2m$ высчитывается следующим образом:

$$A_k = E_k + e^{-\frac{ik \pi}{m}} O_k, \quad k = 0, \dots, m$$

$$A_{k+m} = E_k + e^{-\frac{i(k+m)\pi}{m}} O_k = E_k - e^{-\frac{ik\pi}{m}} O_k, \quad k = 0, \dots, m.$$

Так как E_k и O_k так же могут быть восприняты как БПФ на более грубой сетке, этот алгоритм рекурсивно применяется и к ним, что и замыкает в конечном итоге алгоритм Кули–Тьюки. Ясно, что в приведенном изложении алгоритма предполагается, что $m = 2^{\tilde{m}}, m \in \mathbb{N}$

$$Q(\hat{x}_m) = \frac{(-1)^m}{2^{k-1}} - P_k(\hat{x}_m).$$

Однако из (1)

следует $|P_k(\hat{x}_m)| \leq \frac{1}{2^{k-1}}$. Тогда мы получаем:

$$Q(\hat{x}_m) \geq 0 \quad \text{для } m = 0, 2, \dots$$

$$Q(\hat{x}_m) \leq 0 \quad \text{для } m = 1, 3, \dots$$

Таким образом, вследствие непрерывности $Q(x)$ полином имеет корень в каждом подотрезке $[x_m; x_{m+1}]$, где $m=0, \dots, k-1$, что суммарно дает k различных корней. Однако $Q(x)$ является полиномом $(k-1)$ -й степени. Это возможно только в случае $Q(x) = 0$, что дает $P_k(x) = \widetilde{T}_k(x)$.

$$f(x) - L_{n-1}(x) = \frac{f^{(n)}(\xi)}{n!} \prod_{i=1}^n (x - x_i),$$

где $\xi \in$

$(-1; 1)$ и x_1, \dots, x_n – узлы интерполяции. Оптимальная интерполяция соответствует случаю, при котором значение остаточного члена минимизировано. Так как мы не имеем контроля над $f(x)$, минимизация возможна только относительно значений интерполяционных узлов. Таким образом, мы получаем оптимизационную задачу следующего вида:

$$\min_{x_1, \dots, x_n} \prod_{i=1}^n (x - x_i),$$

при условии $x_i \in [-1; 1], \quad i = 1, \dots, n$.

Целевая функция является нормированным полиномом степени n с n различными корнями, равными x_1, \dots, x_n . По теореме ее максимальное значение на $[-1; 1]$ минимально тогда, когда

$$x_i = \bar{x}_i = \cos\left(\frac{2i-1}{2n}\pi\right), \quad i = 1, \dots, n.$$

Более того, теорема позволяет оценить остаточный член подобной оптимальной интерполяции:

$$\max |f(x) - L_{n-1}(x)| \leq \frac{1}{2^n n!} \max |f^{(n)}(x)|,$$

для любой функции $f(x) \in C^n[-1; 1]$. Обобщение до произвольного отрезка $[a; b]$ реализуется с помощью непрерывной замены переменной

$\tilde{x} = \frac{1}{2}[(b-a)x + a + b]$, где $\tilde{x} \in [a; b]$ и $x \in [-1; 1]$. С помощью подобной замены значения \tilde{x}_i пересчитываются на отрезке $[a; b]$.

Интерполяция Лагранжа для любой достаточно гладкой функции $f(x)$ является оптимальной тогда, когда узлы интерполяции распределены в соответствии с чебышевскими узлами.

Вопрос №19

1. Многошаговые методы. Схема вывода методов Адамса–Башфорта и Адамса–Моултона

- Заметим, что в одношаговых методах использовалась информация в пределах $t \in [t_i, t_{i+1}]$
- Увеличить точность можно за счет уже посчитанных значений $y(t)$ на предыдущих шагах t_{i-1}, t_{i-2}, \dots

Общая форма p -шаговой формулы численного решения задачи Коши:

$$w_{i+1} = a_{p-1}w_i + a_{p-2}w_{i-1} + \dots + a_0w_{i+1-p} + h[b_p f(t_{i+1}, w_{i+1}) + b_{p-1}f(t_i, w_i) + \dots + b_0 f(t_{i+1-p}, w_{i+1-p})]$$

При $b_p \neq 0$ формула называется неявной. Схема вывода метода Адамса–Башфорта. Для расчета значений $w_{i+1}; i < p - 1$ использовать одношаговый метод.

Рассмотрим следующее ОДУ $\frac{dy}{dt} = f(t, y)$,

Где $t \in [a, b]$ и $y(a) = \alpha$. Проинтегрируем обе части уравнения на интервале $[t_i, t_{i+1}]$

$$\int_{t_i}^{t_{i+1}} \frac{dy}{dt} dt = \int_{t_i}^{t_{i+1}} f(t, y(t)) dt \\ \Rightarrow y(t_{i+1}) = y(t_i) + \int_{t_i}^{t_{i+1}} f(t, y(t)) dt.$$

$f(t, y(t))$ по сути является функцией одной переменной t . Для интегрирования требуется найти приближение $f(t, y(t)) \approx P(t)$ воспользуемся интерполяцией Лагранжа для узлов:

$$(t_i, f(t_i, y(t_i))), (t_{i-1}, f(t_{i-1}, y(t_{i-1}))), \dots, (t_{i+1-p}, f(t_{i+1-p}, y(t_{i+1-p}))).$$

Мы получаем:

$$f(t, y(t)) = L_{p-1}(t) + \frac{f^{(p)}(\xi_i, y(\xi_i))}{p!} \prod_{j=1}^p (t - t_{i-j+1}) \\ = \sum_{j=1}^p f(t_{i-j+1}, y(t_{i-j+1})) \prod_{j \neq k} \frac{t - t_{i-k+1}}{t_{i-j+1} - t_{i-k+1}} + \frac{f^{(p)}(\xi_i, y(\xi_i))}{p!} \prod_{j=1}^p (t - t_{i-j+1}) \\ = \sum_{j=1}^p f(t_{i-j+1}, y(t_{i-j+1})) \prod_{j \neq k} \frac{t - t_{i-k+1}}{h(k-j)} + \frac{f^{(p)}(\xi_i, y(\xi_i))}{p!} \prod_{j=1}^p (t - t_{i-j+1}),$$

Где $\xi_i = \xi_i(t) \in [t_i - p + 1, t_i]$. Тогда интеграл от $f(t, y(t))$ принимает вид:

$$\int_{t_i}^{t_{i+1}} f(t, y(t)) dt = \sum_{j=1}^p f(t_{i-j+1}, y(t_{i-j+1})) \int_{t_i}^{t_{i+1}} \prod_{j \neq k} \frac{t - t_{i-k+1}}{h(k-j)} dt + \\ + \int_{t_i}^{t_{i+1}} \frac{f^{(p)}(\xi_i, y(\xi_i))}{p!} \prod_{j=1}^p (t - t_{i-j+1}) dt.$$

Для вывода зависимости остаточного члена от h , проведем замену $t = t_i + sh$ во втором интеграле:

$$\int_{t_i}^{t_{i+1}} \frac{f^{(p)}(\xi_i, y(\xi_i))}{p!} \prod_{j=1}^p (t - t_{i-j+1}) dt = \frac{h}{p!} \int_0^1 f^{(p)}(\xi_i, y(\xi_i)) \prod_{j=1}^p (t_i + sh - t_{i-j+1}) ds \\ = \frac{h^{p+1}}{p!} \int_0^1 f^{(p)}(\xi_i, y(\xi_i)) \prod_{j=1}^p (s + j - 1) ds \\ = \frac{h^{p+1}}{p!} f^{(p)}(\mu_i, y(\mu_i)) \int_0^1 \prod_{j=1}^p (s + j - 1) ds,$$

где в последнем шаге мы воспользовались теоремой о среднем значении ($\prod_{j=1}^p (s + j - 1)$ не меняет знак для $s \in [0; 1]$) и $\mu_i \in (t_i; t_{i+1})$. Полученное выражение говорит о том, что локальная погрешность метода будет иметь порядок $O(h^{p+1})$. Проведем аналогичную замену для первого интеграла:

$$\int_{t_i}^{t_{i+1}} f(t, y(t)) dt = h \sum_{j=1}^p f(t_{i-j+1}, y(t_{i-j+1})) \int_0^1 \prod_{j \neq k} \frac{s + k - 1}{k - j} ds.$$

Таким образом мы получаем следующую обобщенную формулировку явного p -

		<p>шагового метода, известного как метод Адамса–Башфорта p-го порядка:</p> $w_0 = \alpha_0, w_1 = \alpha_1, \dots, w_{p-1} = \alpha_{p-1},$ $w_{i+1} = w_i + h \sum_{j=1}^p a_j f(t_{i-j+1}, w_{i-j+1}), \quad i = p-1, p, \dots, m-1,$ <p>Где коэффициенты a_j имеют вид:</p> $a_j = \int_0^1 \prod_{j \neq k} \frac{s+k-1}{k-j} ds, j = 1, \dots, p$ <p>Схема вывода метода Адамса–Моултона. Аналогичным образом можно получить формулировку для неявного метода. В таком случае мы имеет следующие p точек интерполяции: $(t_{i+1}, f(t_{i+1}, y(t_{i+1}))) ; (t_i, f(t_i, y(t_i))) ; \dots ; (t_{i+2-p}, f(t_{i+2-p}, y(t_{i+2-p})))$ Вывод будет идентичен выводу явного метода. В результате мы получаем $(p-1)$-шаговый метод, известный как метод Адамса–Моултона:</p> $w_0 = a_0, w_1 = a_1, \dots, w_{p-2} = a_{p-2}$ $w_{i+1} = w_i + h \sum_{j=1}^p a_j f(t_{i-j+2}, y(t_{i-j+2})), i = p-2, p-1, \dots, m-1$ <p>Где коэффициенты a_j имеют вид:</p> $a_j = \int_0^1 \prod_{j \neq k} \frac{s+k-2}{k-j} ds, j = 1, \dots, p$ <p>и остаточный член, формирующийся из-за интерполяции функции $f(t, y(t))$, имеет форму:</p> $\frac{h^{p+2}}{p!} f^{(p)}(\mu_i, y(\mu_i)) \int_0^1 \prod_{j=1}^p (s+j-2) ds$ <p>Таким образом, $(p-1)$-шаговый метод Адамса–Моултона имеет порядок точности $O(h^p)$, что делает его на один порядок точным, чем $(p-1)$шаговый метод Адамса–Башфорта</p>
--	--	---