

Тема : Методы поиска безусловного экстремума функции

1⁰. Постановка задачи, локальный и глобальный минимумы. Целевая функция. Условный минимум. 2⁰. Связь задачи минимизации функций с задачей решения системы нелинейных алгебраических уравнений. 3⁰. Минимизация функций одной переменной: метод перебора. 4⁰. Минимизация функций одной переменной: метод исключения отрезков, метод дихотомии. 5⁰. Минимизация функций одной переменной: метод золотого сечения.

4⁰. Модификацией метода перебора служит метод исключения отрезков. Сначала исходный отрезок $[a, b]$ делится на три части с помощью двух точек u_1 и u_2 , удовлетворяющих условию

$$a < u_1 < u_2 < b.$$

Затем сравниваем значения $\Phi(u_1)$ и $\Phi(u_2)$.

Если $\Phi(u_1) \leq \Phi(u_2)$, то вместо $[a, b]$ рассматриваем далее меньший отрезок $[a, u_2]$.

Если же $\Phi(u_1) > \Phi(u_2)$, то рассматриваем отрезок $[u_1, b]$.

Новый отрезок снова разбиваем на три части, причем в качестве узлов нового разбиения возьмем уже существующий и какой-то новый.

Возьмем в итерационном процессе метода исключения отрезков узлы u_1 и u_2 близки-

ми к середине отрезка, то есть положим

$$u_1 = \frac{b + a - \delta}{2} < u_2 = \frac{b + a + \delta}{2},$$

где число δ положительно и достаточно мало. Тогда получится метод дихотомии.

После выполнения первого шага длина отрезка, на котором ищется минимум, уменьшится либо в $\frac{b-u_1}{b-a}$ раз, либо в $\frac{u_2-a}{b-a}$ раз. Если

δ мало, то оба этих отношения близки к $\frac{1}{2}$, как это следует из равенств

$$\frac{b - u_1}{b - a} = \frac{1}{2} + \frac{\delta}{2(b - a)};$$

$$\frac{u_2 - a}{b - a} = \frac{1}{2} - \frac{\delta}{2(b - a)}.$$

Длина δ_n отрезка Δ_n , получившегося через n итераций метода, вычисляется по формуле

$$\delta_n = \frac{b - a}{2^n} + \left(1 - \frac{1}{2^n}\right)\delta.$$

Искомую точку минимума u^* выбираем внутри отрезка Δ_n , получившегося через n итераций, или же на границе этого отрезка. В качестве приближения к u^* выбирается середина последнего найденного отрезка. Следовательно, точность локализации минимума u^* через n итераций не будет превышать величины $\frac{\delta_n}{2} = \varepsilon_n$.

Условие окончания итерационного процесса имеет вид $\varepsilon_n \leq \varepsilon$, где ε — приемлемая

точность. Учитывая явный вид ε_n , получаем следующую оценку снизу на необходимое число итераций метода:

$$n \geq \log_2 \frac{b - a - \delta}{2\varepsilon - \delta}.$$

Здесь $\delta = u_2 - u_1$. Если $\delta \approx 0$, то точность локализации минимума $\varepsilon_n \approx \frac{b-a}{2^{n+1}}$.

5⁰. Расположим точки u_1 и u_2 , удовлетворяющие условию $a < u_1 < u_2 < b$, таким образом, чтобы одна из них стала пробной и на

новом отрезке, который получается на первой итерации после исключения из исходного отрезка некоторой его части.

Тогда значение $\Phi(u)$ на следующей итерации достаточно будет вычислить только во вновь добавленном узле. Во втором из узлов значение функции $\Phi(\cdot)$ найдено еще на предыдущем шаге. Этот простой прием позволяет существенно уменьшить количество вычислений.

Пусть $a = 0$ и $b = 1$. Предположим, что на первой итерации исключается правая часть исходного отрезка, то есть промежуток $(u_2, 1]$. Укажем, как должны быть расположены в этом случае узлы u_1 и u_2 .

Пусть $\frac{1}{2} \leq \tau < 1$, $u_2 = \tau$. В качестве u_1 возьмем симметричную относительно $\frac{1}{2}$ точку, то есть $u_1 = 1 - \tau$, $0 < u_1 \leq \frac{1}{2}$.

После первой итерации отрезок $[0, 1]$ перейдет в отрезок $[0, \tau]$, а узел u_1 станет узлом $u_2^{(1)} = 1 - \tau$ нового отрезка $[0, \tau]$.

Далее отрезок $[0, 1]$ точка $u_2 = \tau$ делит в отношении $\frac{1}{\tau}$. Отрезок же $[0, \tau]$ узел $u_2^{(1)} = 1 - \tau = u_1$ делит в отношении $\frac{\tau}{1 - \tau}$.

Потребуем, чтобы эти отношения совпада-

ли, тогда получим

$$\frac{1}{\tau} = \frac{\tau}{1 - \tau} \Leftrightarrow \tau^2 + \tau - 1 = 0.$$

Положительный корень этого уравнения

$$\tau = \frac{\sqrt{5} - 1}{2} \approx 0.61803.$$

Таким образом, следует взять

$$u_1 = 1 - \tau = \frac{3 - \sqrt{5}}{2}, \quad u_2 = \tau = \frac{\sqrt{5} - 1}{2}.$$

Соответствующие формулы для отрезка $[a, b]$ имеют вид

$$u_1 = a + \frac{3 - \sqrt{5}}{2}(b - a), \quad u_2 = a + \frac{\sqrt{5} - 1}{2}(b - a).$$

Метод минимизации, основанный на указанном разбиении $a < u_1 < u_2 < b$, называется методом золотого сечения.

Точки u_1 и u_2 обладают следующим свойством: каждая из них делит отрезок $[a, b]$ на

две неравные части. При этом отношение длины всего отрезка к длине его большей части равно отношению длин его большей и меньшей частей. Обладающие этим свойством точки u_1 и u_2 называются точками золотого сечения.

В методе золотого сечения на каждой итерации длина отрезка, на котором продолжается поиск минимума, уменьшается в одном и том же отношении $\tau = \frac{\sqrt{5}-1}{2}$.

После n итераций длина отрезка становится равной $\delta_n = \tau^n(b - a)$.

Таким образом, погрешность в определении точки минимума u^* после n итераций равна

$$\varepsilon_n = \frac{\delta_n}{2} = \frac{1}{2} \left(\frac{\sqrt{5} - 1}{2} \right)^n (b - a).$$

Условие окончания вычислительного процесса имеет вид $\varepsilon_n \leq \varepsilon$, где ε — приемлемая точность.

Тема : Итерационные методы минимизации функций

1⁰. Метод покоординатного спуска. Пример. 2⁰. Метод градиентного спуска. Пример. 3⁰. Метод наискорейшего спуска. Пример. 4⁰. Выпуклые множества и выпуклые функции. Пример. Достаточные условия строгой выпуклости дважды непрерывно дифференцируемой функции. 5⁰. Теорема о минимуме выпуклой (строго выпуклой) функции. Следствие: одноэкстремальность задачи минимизации. 6⁰. Задачи линейного программирования.

1⁰. Среди алгоритмов минимизации функций многих вещественных переменных выделяется особый класс методов, позволяющий перейти из точки начального приближения

$$\vec{u}_0 = (u_1^0, u_2^0, \dots, u_n^0)$$

в некоторую новую точку

$$\vec{u}_1 = (u_1^1, u_2^1, \dots, u_n^1)$$

таким образом, чтобы новое значение $\Phi(\vec{u}_1)$ функции приблизилось к искомому минимальному значению целевой функции.

Алгоритмы, построенные на указанном принципе поиска, получили общее название методов спуска. Характерный представитель указанного класса алгоритмов — это метод покоординатного спуска (МПС). В нем поиск минимума функции многих переменных

сведется к последовательной минимизации функций одной переменной.

Пусть $\Phi(\vec{u})$ — целевая функция и \vec{u}_0 из U — это начальное приближение к точке минимума $\Phi(\vec{u})$.

Зафиксируем значения $(u_2^0, u_3^0, \dots, u_n^0)$ всех координат вектора \vec{u}_0 начиная со второй. Затем рассмотрим функцию одной переменной

u_1 , задаваемую равенством

$$f_1(u_1) = \Phi(u_1, u_2^0, u_3^0, \dots, u_n^0).$$

Найдем минимум $f_1(u_1)$ каким-либо из ранее описанных методов. Полученное значение переменной u_1 , доставляющее минимум

$$\min_{u_1 \in U_1} f_1(u_1) = \min_{\vec{u} \in U} \Phi(u_1, u_2^0, u_3^0, \dots, u_n^0)$$

обозначим как u_1^1 . При этом справедлива следующая оценка:

$$\Phi(u_1^1, u_2^0, u_3^0, \dots, u_n^0) \leq \Phi(u_1^0, u_2^0, u_3^0, \dots, u_n^0).$$

Далее зафиксируем переменные u_1^1 и (u_3^0, \dots, u_n^0) , а затем минимизируем по переменной u_2 функцию

$$f_2(u_2) = \Phi(u_1^1, u_2, u_3^0, \dots, u_n^0).$$

Точку минимума этой функции обозначим как u_2^1 . Тогда справедлива следующая оценка:

$$\Phi(u_1^1, u_2^1, u_3^0, \dots, u_n^0) \leq \Phi(u_1^1, u_2^0, u_3^0, \dots, u_n^0).$$

Аналогично поступаем с третьей и последующими координатами целевой функции. В результате получаем вектор

$$\vec{u}_1 = (u_1^1, u_2^1, u_3^1, \dots, u_n^1),$$

обладающий по построению тем свойством,
что

$$\Phi(\vec{u}_1) \leq \Phi(\vec{u}_0).$$

Затем возьмем вектор \vec{u}_1 в качестве нового начального приближения и повторим процесс. Получим в результате следующий вектор \vec{u}_2 со свойством

$$\Phi(\vec{u}_2) \leq \Phi(\vec{u}_1).$$

Таким образом, получаем последовательность $\vec{u}_1, \vec{u}_2, \vec{u}_3, \dots, \vec{u}_n, \dots$ приближений к искомой точке минимума.

Возникает вопрос: на каком номере n (векторе \vec{u}_n) следует остановиться?

В качестве условия выхода из итерационного процесса часто выбирают оценку следу-

ющего вида:

$$|\Phi(\vec{u}_{n+1}) - \Phi(\vec{u}_n)| \leq \varepsilon,$$

где $\varepsilon > 0$ — наперед заданная (или приемлемая) точность.

Пример. Пусть $\Phi(u_1, u_2) = (u_1^2 + u_2^2)$. Взяв начальное приближение $\vec{u}_0 = (2, 2)$, найти точку минимума функции $\Phi(\vec{u})$ методом покоординатного спуска.

Имеем равенство

$$f_1(u_1) = u_1^2 + 4 \Rightarrow \min_{u_1} f_1(u_1) = 4.$$

Этот минимум достигается в точке $u_1^1 = 0$.

Далее полагаем $\vec{u}_1 = (0, u_2)$ и рассматриваем функцию $f_2(u_2) = \Phi(0, u_2) = u_2^2$. Минимум этой функции достигается при $u_2 = 0 \Rightarrow \vec{u}_2 = (0, 0)$. Таким образом, итерационный процесс сошелся за два шага.

Теорема (достаточное условие сходимости МПС). Пусть существуют вторые производные целевой функции $\Phi(u_1, u_2)$, причем

$$\frac{\partial^2 \Phi}{\partial u_1^2} \geq a_1 > 0, \quad \frac{\partial^2 \Phi}{\partial u_2^2} \geq a_2 > 0, \quad \frac{\partial^2 \Phi}{\partial u_1 \partial u_2} \leq a_3.$$

Если при этом $a_1 a_2 > a_3^2$, то последовательность приближений, построенная по методу покоординатного спуска, сходится к точке минимума функции $\Phi(u_1, u_2)$.

2⁰. Градиент функции $\Phi(u) = \Phi(u_1, u_2, \dots, u_n)$ определяется равенством

$$\nabla \Phi(u) = \left(\frac{\partial \Phi}{\partial u_1}(u), \frac{\partial \Phi}{\partial u_2}(u), \dots, \frac{\partial \Phi}{\partial u_n}(u) \right).$$

Если $\nabla \Phi \neq 0$ в некоторой точке \vec{u}_0 , то определяемый градиентом вектор ортогонален линии уровня функции $\Phi(\vec{u})$, проходящей через точку \vec{u}_0 . Направление $\nabla \Phi(\vec{u}_0)$ совпадает с направлением наибольшего роста функции $\Phi(\vec{u})$ в точке \vec{u}_0 .

Отметим, что в точке минимума функции $\Phi(\vec{u})$ имеет место равенство $\nabla\Phi = 0$.

Зададим начальное приближение \vec{u}_0 и рассмотрим следующий итерационный процесс

$$\vec{u}_{k+1} = \vec{u}_k - \tau \nabla\Phi(\vec{u}_k), \quad k = 0, 1, 2, \dots \quad (G)$$

Здесь τ — это скаляр, называемый итерационным параметром.

Алгоритм, основанный на итерационных приближениях (G), называется методом градиентного спуска (МГС).

Условие выхода из процесса (G) можно записать как оценку

$$\|\nabla\Phi(\vec{u}_{k+1})\| \leq \varepsilon,$$

где ε — приемлемая точность.

Пример. Пусть $\Phi(u_1, u_2) = \frac{u_1^2}{4} + u_2^2$. Тогда

$$\nabla \Phi(u_1, u_2) = \left(\frac{u_1}{2}, 2u_2 \right).$$

Формулы (**G**) при этом принимают вид

$$\begin{cases} u_1^{(k+1)} = u_1^{(k)} - \tau \frac{1}{2} u_1^{(k)}, \\ u_2^{(k+1)} = u_2^{(k)} - \tau 2 u_2^{(k)}, \end{cases} \quad k = 0, 1, 2, \dots \quad (G')$$

Если $\vec{u}_0 = (1, 1)$ и $\tau = 0.1$, то имеем

$$\begin{cases} \vec{u}_1 = (0.95, 0.80), \\ \vec{u}_2 = (0.9025, 0.6400), \\ \vec{u}_3 = (0.8574, 0.5120). \end{cases}$$

При этом $\Phi(\vec{u}_1) = 1.25$, $\Phi(\vec{u}_3) = 0.446$. Если $\tau = 2$, то из (G') получаем

$$\vec{u}_1 = (0, -3); \quad \Phi(\vec{u}_1) = 9.$$

При этом искомый минимум равен нулю и достигается в начале координат.

Таким образом, выбор шага τ оказывает существенное влияние на точность результата МГС.

В этой связи часто используют переменный итерационный параметр, то есть выбирают значение параметра τ в определенной зависимости от номера итерации.

3⁰. В методе наискорейшего спуска на каждой итерации шаг τ выбирается таким образом, чтобы целевая функция $\Phi(\vec{u})$ при переходе от аргумента \vec{u}_k к \vec{u}_{k+1} уменьшала свое значение в максимально возможной степени. Таким образом, итерационный параметр находится из следующего условия

$$\Phi(\vec{u}_{k+1}) = \min_{\tau} \Phi(\vec{u}_k - \tau \nabla \Phi(\vec{u}_k)).$$

Проиллюстрируем на примере работу метода наискорейшего спуска.

Пример. Пусть целевая функция и начальное приближение заданы равенствами

$$\Phi(u_1, u_2) = \frac{u_1^2}{4} + u_2^2 \quad \text{и} \quad \vec{u}_0 = (1, 1).$$

Требуется найти приближение \vec{u}_1 по методу наискорейшего спуска.

Имеем векторные равенства

$$\nabla\Phi(u_1, u_2) = \left(\frac{u_1}{2}, 2u_2\right), \quad \nabla\Phi(\vec{u}_0) = \left(\frac{1}{2}, 2\right).$$

Искомое приближение \vec{u}_1 ищем из условия

$$\Phi(\vec{u}_1) = \min_{\tau} \Phi(\vec{u}_0 - \tau \nabla\Phi(\vec{u}_0)).$$

Сосчитаем значение $\Phi(\vec{u}_0 - \tau \nabla\Phi(\vec{u}_0))$. Имеем

$$\vec{u}_0 - \tau \nabla\Phi(\vec{u}_0) = (u_1^0 - \tau \frac{u_1^0}{2}, u_2^0 - 2\tau u_2^0).$$

Следовательно,

$$\Phi(\vec{u}_0 - \tau \nabla \Phi(\vec{u}_0)) = \frac{1}{4}(u_1^0 - \frac{\tau}{2}u_1^0)^2 + (u_2^0 - 2\tau u_2^0)^2.$$

Подставляя сюда координаты начального вектора $\vec{u}_0 = (1, 1)$, получаем

$$\Phi(\vec{u}_0 - \tau \nabla \Phi(\vec{u}_0)) = \frac{1}{4}\left(1 - \frac{\tau}{2}\right)^2 + (1 - 2\tau)^2.$$

Минимум квадратичной функции

$$f(\tau) = \frac{1}{4}\left(1 - \frac{\tau}{2}\right)^2 + (1 - 2\tau)^2$$

достигается в точке τ^* с условием $f'(\tau^*) = 0$,
то есть при $\tau^* = \frac{34}{65}$.

Искомое первое приближение задается формулой

$$\vec{u}_1 = \vec{u}_0 - \frac{34}{65} \nabla \Phi(\vec{u}_0) = (1, 1) - \frac{34}{65} \left(\frac{1}{2}, 2 \right).$$

На втором шаге выбор параметра τ^* происходит аналогично с учетом найденного вектора (u_1^1, u_2^1) .

4⁰. Решение экстремальных задач в многомерных пространствах сопряжено со значительными трудностями, особенно если искомым экстремумов много. Некоторых из этих трудностей можно избежать, если искать экстремумы выпуклых функций на выпуклых множествах.

Определение. Множество U выпуклое, если для любых $\vec{u}, \vec{v} \in U$ множеству U принадлежит весь отрезок $\lambda \vec{u} + (1 - \lambda) \vec{v}$, где $0 \leq \lambda \leq 1$.

Определение. Функция $\Phi(\vec{u})$, заданная на выпуклом множестве $U \subset \mathbb{R}^n$, называется выпуклой, если для любых двух векторов \vec{u} и \vec{v} из U и любого числа λ из отрезка $[0, 1]$ имеет место неравенство

$$\Phi(\lambda \vec{u} + (1 - \lambda) \vec{v}) \leq \lambda \Phi(\vec{u}) + (1 - \lambda) \Phi(\vec{v}).$$

Пусть для любых двух векторов \vec{u} и \vec{v} из U , $\vec{u} \neq \vec{v}$, и любого числа λ , $0 < \lambda < 1$, справед-

ливо строгое неравенство

$$\Phi(\lambda \vec{u} + (1 - \lambda) \vec{v}) < \lambda \Phi(\vec{u}) + (1 - \lambda) \Phi(\vec{v}).$$

Тогда функция $\Phi(\cdot)$ называется *строго выпуклой*.

В одномерном случае свойство выпуклости имеет наглядную геометрическую интерпретацию: график функции $y = f(x)$ на интервале с концами в точках u и v , лежит ниже

(или строго ниже) хорды, соединяющей точки $(u, f(u))$ и $(v, f(v))$ графика. Примеры строго выпуклых функций: $y = e^x$ и $y = x^2$.

Матрица Гессе $G(\vec{u})$ дважды непрерывно дифференцируемой функции $\Phi(\vec{u})$ определяется

равенством

$$G(\vec{u}) = \begin{pmatrix} \frac{\partial^2 \Phi}{\partial u_1^2} & \frac{\partial^2 \Phi}{\partial u_1 \partial u_2} & \frac{\partial^2 \Phi}{\partial u_1 \partial u_3} & \cdots & \frac{\partial^2 \Phi}{\partial u_1 \partial u_n} \\ \frac{\partial^2 \Phi}{\partial u_2 \partial u_1} & \frac{\partial^2 \Phi}{\partial u_2^2} & \frac{\partial^2 \Phi}{\partial u_2 \partial u_3} & \cdots & \frac{\partial^2 \Phi}{\partial u_2 \partial u_n} \\ \frac{\partial^2 \Phi}{\partial u_3 \partial u_1} & \frac{\partial^2 \Phi}{\partial u_3 \partial u_2} & \frac{\partial^2 \Phi}{\partial u_3^2} & \cdots & \frac{\partial^2 \Phi}{\partial u_3 \partial u_n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 \Phi}{\partial u_n \partial u_1} & \frac{\partial^2 \Phi}{\partial u_n \partial u_2} & \frac{\partial^2 \Phi}{\partial u_n \partial u_3} & \cdots & \frac{\partial^2 \Phi}{\partial u_n^2} \end{pmatrix}$$

В случае одной переменной эта матрица состоит в точности из одного элемента — второй производной функции $\Phi(u)$.

Теорема. Если матрица Гессе $G(\vec{u})$ дважды непрерывно дифференцируемой функции $\Phi(\vec{u})$ положительно определена, то функция $\Phi(\vec{u})$ строго выпукла.

Отметим, что матрица $G(\vec{u})$ симметрична и поэтому является положительно определенной в том и только том случае, если все ее главные миноры строго положительны.

Теорема (о минимуме выпуклой функции).

Пусть $\Phi(\vec{u})$ — выпуклая функция на выпуклом множестве U . Тогда любой ее локальный минимум является глобальным на U .

Для строго выпуклой функции $\Phi(\vec{u})$ ее глобальный минимум на выпуклом множестве достигается в единственной точке.

Доказательство. Предположим, что \vec{u}_0 — это точка локального минимума выпуклой функции на выпуклом множестве U , а \vec{u}_* — это какая-нибудь точка ее глобального минимума на том же множестве.

Если $\vec{u}_0 \neq \vec{u}_*$, то есть если \vec{u}_0 не является точкой глобального минимума, то имеем $\Phi(\vec{u}_0) > \Phi(\vec{u}_*)$.

Получаем далее с учетом выпуклости $\Phi(\vec{u})$ для любого числа λ с условием $0 < \lambda < 1$:

$$\begin{aligned}\Phi(\lambda \vec{u}_* + (1 - \lambda) \vec{u}_0) &\leq \lambda \Phi(\vec{u}_*) + (1 - \lambda) \Phi(\vec{u}_0) < \\ &< \lambda \Phi(\vec{u}_0) + (1 - \lambda) \Phi(\vec{u}_0) = \Phi(\vec{u}_0).\end{aligned}$$

В этом строгом неравенстве λ — любое число с условием $0 < \lambda < 1$. Полагая здесь $\lambda = \frac{1}{n}$, получаем, что в любой сколь угодно малой

окрестности точки \vec{u}_0 локального минимума
найдется точка

$$\vec{u}_n = \frac{1}{n} \vec{u}_* + \left(1 - \frac{1}{n}\right) \vec{u}_0,$$

в которой значение функции $\Phi(\vec{u})$ строго
меньше значения $\Phi(\vec{u}_0)$. Это противоречит
условию, что в \vec{u}_0 функция достигает ло-
кального минимума.

Полученное противоречие доказывает, что в условиях теоремы с необходимостью выполняется неравенство $\Phi(\vec{u}_0) \leq \Phi(\vec{u}_*)$. Следовательно и по определению точки глобального минимума, имеем равенство $\Phi(\vec{u}_0) = \Phi(\vec{u}_*)$. Таким образом, первая часть теоремы доказана.

Пусть теперь строго выпуклая функция $\Phi(\vec{u})$ имеет две разные точки глобального мини-

мума \vec{u}_1 и \vec{u}_2 , $\vec{u}_1 \neq \vec{u}_2$. Тогда для $\forall \lambda: 0 < \lambda < 1$ выполняется оценка

$$\Phi(\lambda \vec{u}_1 + (1 - \lambda) \vec{u}_2) < \lambda \Phi(\vec{u}_1) + (1 - \lambda) \Phi(\vec{u}_2).$$

Но по условию

$$\Phi(\vec{u}_1) = \Phi(\vec{u}_2) = \min_{\vec{u} \in U} \Phi(\vec{u}) \equiv \Phi^*.$$

Поэтому имеем

$$\Phi(\lambda \vec{u}_1 + (1 - \lambda) \vec{u}_2) < \lambda \Phi^* + (1 - \lambda) \Phi^* = \Phi^*.$$

Это противоречит условию, что \vec{u}_1 и \vec{u}_2 — это точки глобального минимума.

Таким образом, предположение, что $\vec{u}_1 \neq \vec{u}_2$ неверно, то есть $\vec{u}_1 = \vec{u}_2$ и точка глобального минимума строго выпуклой функции $\Phi(\vec{u})$ на выпуклом множестве U единственна. □

5⁰. Часть теории экстремальных задач, в которой минимизируют линейные функции при дополнительных линейных ограничениях, которых может быть много, называется линейным программированием.

Общий вид линейной функции многих переменных

$$\Phi(\vec{u}) = \sum_{i=1}^n C_i u_i + C_0.$$

В задаче минимизации удобно полагать, что $C_0 = 0$, так как значение C_0 на положение точки минимума никак не влияет.

Линейные ограничения общего вида — это либо равенства, либо неравенства следующего вида:

$$\sum_{i=1}^n a_{ij} u_i = b_j, \quad 1 \leq j \leq J_1;$$

$$\sum_{i=1}^n a_{ij} u_i \leq b_j, \quad J_1 + 1 \leq j \leq J_2.$$

Каждое из этих неравенств определяет в \mathbb{R}^n полупространство, границей которого является гиперплоскость размерности $(n - 1)$.

Совместное выполнение этих условий означает пересечение соответствующих полупространств, то есть принадлежность варьируе-

мой переменной некоторому выпуклому многограннику M .

Дополнительные условия на варьируемые переменные в виде линейных равенств выделяют в \mathbb{R}^n плоскость размерности меньшей n . Пересечение этой плоскости с многогранником M дает выпуклый многогранник G размерности $< n$.

Таким образом, задача состоит в минимизации линейной функции $\Phi(\vec{u}) = \sum_{i=1}^n C_i u_i$, являющейся выпуклой на выпуклом многограннике G .

Если многогранник G выпуклый, и возможно неограниченный, то внутри него линейная функция достигать минимального значения не может: у нее нет стационарных точек.

Если все же $\min_{\vec{u} \in G} \Phi(\vec{u})$ существует, то достигаться он может лишь в какой-то из вершин G . Таким образом, теоретически задача линейного программирования проста: необходимо вычислить значения минимизируемой функции в конечном числе точек — в вершинах многогранника G . Затем следует сравнить вычисленные значения $\Phi(\cdot)$ между собой и выбрать из них минимальное.

Трудности в решении этих задач порождаются тем, что количество переменных может оказаться очень большим. Например, в экономических задачах количество переменных достигает 10^2 – 10^4 .

Пример. Пусть $\Phi(\vec{u}) = C_1u_1 + C_2u_2$. Найти минимум при дополнительных условиях

$$\begin{cases} a_{11}u_1 + a_{12}u_2 = b_1, & u_1 \geq 0, \\ a_{21}u_1 + a_{22}u_2 = b_2, & u_2 \geq 0. \end{cases}$$

Предположим, что система из дополнительных условий

$$\begin{cases} a_{11}u_1 + a_{12}u_2 = b_1 \\ a_{21}u_1 + a_{22}u_2 = b_2 \end{cases}$$

имеет единственное решение $\vec{u}^* = (u_1^*, u_2^*)$.

Если $u_1^* \geq 0$ и $u_2^* \geq 0$, то $\min_{\vec{u} \in G} \Phi(\vec{u}) = \Phi(u_1^*, u_2^*)$.

Если же $u_1^* < 0$ или $u_2^* < 0$, то задача решения не имеет.

Тема : Интерполяция функций

1⁰. Дискретизация и интерполяция непрерывной функции. Погрешность (потеря информации). 2⁰. Кусочно-линейная интерполяция: графическое и аналитическое определения. Теорема о погрешности кусочно-линейной интерполяции. Порядок сходимости. Кусочно-линейный интерполянт как линейная комбинация базисных функций. 3⁰. Достаточное условие интерполяции функции обобщенным полиномом. Переход к системе линейных уравнений с матрицей Грама. Пример дискретной ортогональной системы функций на отрезке с равномерной сеткой. 4⁰. Условие интерполяции функции алгебраическим полиномом. Интерполянт как линейная комбинация базисных полиномов Лагранжа. Интерполяционная формула Лагранжа, полином Лагранжа. Свойства лагранжевой интерполяции. 5⁰. Теорема о погрешности интерполяционной формулы Лагранжа. Следствие для равномерной сетки узлов. 6⁰. Эквивалентный вид полинома Лагранжа. Пример: интерполяция $\sin x$ на отрезке $[0, \pi/2]$ по трем узлам.

1⁰. Пусть на отрезке $[a, b]$ числовой оси задана непрерывная функция $f = f(x)$, $x \in [a, b]$. Значения функции $f(x)$ известны в точках сетки, состоящих из равноотстоящих узлов

$$x_k = a + k\tau, \quad k = 0, 1, 2, \dots, N,$$

где $\tau = \frac{b-a}{N}$ — шаг сетки. Таким образом, имеется числовой вектор

$$\vec{f}_N = \{f(x_0), f(x_1), f(x_2), \dots, f(x_N)\},$$

называемый сеточной проекцией функции $f(x)$.
Оператор $R: f(x) \rightarrow \vec{f}_N$ называется оператором ограничений функции на сетку, или же оператором рестрикции.

Замена функции $f(x)$ ее сеточной проекцией называется **дискретизацией**.

Процесс восстановления непрерывной функции по известной последовательности ее се-

точных проекций $\{\vec{f}_N\}$, $N = 2, 3, \dots$, называется **интерполяцией**.

Имеется множество возможностей интерполировать функцию по известным ее значениям в точках сетки, то есть множество правил вида

$$\{f(x_k) \mid k = 0, 1, 2, \dots, N\} \rightarrow f^*(x) \in C(\mathbb{R}). \quad (IF)$$

Восстановленная по такого рода интерполяционной формуле непрерывная функция $f^*(x)$ в общем случае с исходной функцией $f(x)$ не совпадает.

Таким образом, возникает разница между $f(x)$ и $f^*(x)$, о которой принято говорить как о **погрешности интерполяции**.

В связи с процессом “функция — дискретизация — интерполяция — новая функция”

приходится решать вопрос о выборе таких шага сетки (**шага дискретизации**) и последующего способа интерполяции (*IF*), при которых погрешность оказывается наименьшей. Широко используются варианты согласованных между собой дискретизации и интерполяции, при которых для достаточно емких классов функций (например, полиномов не выше определенной степени) погрешности вообще не возникает.

2^0 . Простейший способ интерполяции — это **кусочно-линейная**. Определить этот способ проще всего графически.

Сопоставив каждому узлу x_k сетки и соответствующему значению $f(x_k)$ точку на плоскости, следует затем соединить соседние точки отрезками прямых.

В результате получим **ломаную** непрерывную линию — это и есть искомая непрерывная функция $f^*(x)$.

Аналитическое задание интерполанта $f^*(x)$ произведем следующим образом:

$$\left\{ \begin{array}{l} f_k(x) = \frac{f_{k+1} \cdot (x - x_k) + f_k \cdot (x_{k+1} - x)}{x_{k+1} - x_k}, \quad x_k \leq x \leq x_{k+1}, \\ \text{где} \quad k = 0, 1, 2, \dots, N - 1 \quad \text{и} \quad f_k = f(x_k). \end{array} \right.$$

Теорема (погрешность кусочно-линейной интерполяции). Пусть функция $f(x)$ удовлетворяет на $[a, b]$ условию Липшица:

$$|f(x^*) - f(x^{**})| \leq L|x^* - x^{**}| \quad \forall x^*, x^{**} \in [a, b].$$

Тогда $|f(x) - f^*(x)| \leq \frac{L}{2}\tau$, где

$$\tau = \max_{0 \leq k \leq N-1} (x_{k+1} - x_k).$$

Доказательство. Из определения кусочно линейного интерполянта получаем равенства:

$$\forall x : x_k \leq x \leq x_{k+1} \quad \Rightarrow \quad x = x_k + \alpha \tau_k,$$

где $\tau_k = x_{k+1} - x_k$. При этом

$$\alpha = \alpha(x) = \frac{x - x_k}{x_{k+1} - x_k},$$

то есть $0 \leq \alpha \leq 1$. Кроме того

$$f^*(x) = \alpha f_{k+1} + (1 - \alpha) f_k; \quad k = 0, 1, \dots, N - 1.$$

Таким образом, для $x_k \leq x \leq x_{k+1}$ имеем оценку

$$\begin{aligned} |f^*(x) - f(x)| &= |\alpha f_{k+1} + (1 - \alpha)f_k - \alpha f(x) - (1 - \alpha)f(x)| \leq \\ &\leq \alpha |f_{k+1} - f(x)| + (1 - \alpha) |f_k - f(x)|. \quad (I) \end{aligned}$$

Учитывая, что $f_{k+1} = f(x_k + \tau_k)$, имеем далее по условию Липшица

$$\begin{aligned} |f_{k+1} - f(x)| &= |f(x_k + \tau_k) - f(x_k + \alpha\tau_k)| \leq \\ &\leq L|\tau_k - \alpha\tau_k| = L(1 - \alpha)\tau_k. \end{aligned}$$

Последнее равенство справедливо в силу соотношений $0 \leq \alpha \leq 1$. Аналогично

$$|f_k - f(x)| = |f(x_k) - f(x_k + \alpha\tau_k)| \leq L|\alpha\tau_k| = L\alpha\tau_k.$$

Подставляя полученные оценки в правую часть неравенства (**I**), получаем

$$\begin{aligned} |f_{k+1}(x) - f(x)| &\leq L\alpha(1 - \alpha)\tau_k + L(1 - \alpha)\alpha\tau_k = \\ &= 2\alpha(1 - \alpha)L\tau_k, \quad k = 0, 1, 2, \dots, N - 1. \end{aligned}$$

Но при $0 \leq \alpha \leq 1$ справедлива оценка

$$0 \leq 2\alpha(1 - \alpha) \leq \max_{0 \leq \alpha \leq 1} 2\alpha(1 - \alpha) = \frac{1}{2}.$$

(В самом деле, если $\varphi(\alpha) \equiv 2\alpha(1 - \alpha)$, то $\varphi'(\alpha) = 2 - 4\alpha$ и $\varphi'(\alpha) = 0$ при $\alpha = \frac{1}{2}$. Следовательно, в точке $\alpha = \frac{1}{2}$ функция $\varphi(\alpha)$ максимальна).

Таким образом, окончательная оценка погрешности кусочно-линейной интерполяции

имеет следующий вид:

$$|f(x) - f^*(x)| \leq \frac{L}{2}\tau.$$

Здесь $\tau = \max_{0 \leq k \leq 1} \tau_k = \max_{0 \leq k \leq 1} (x_{k+1} - x_k)$. □

Следствие. При равномерном разбиении отрезка узлами интерполяции, то есть при

$$\tau_0 = \tau_1 = \dots = \tau_{N-1} = \frac{b-a}{N},$$

для погрешности интерполяции справедлива оценка $|f(x) - f^*(x)| \leq \frac{L(b-a)}{2N}$.

Имея в виду эту оценку, говорят, что кусочно-линейная интерполяция имеет **первый** порядок сходимости.

Для линейной функции $f(x) = Ax + B$ справедливо точное равенство $f(x) \equiv f^*(x)$.

Таким образом, кусочно-линейная интерполяция **точна на полиномах первой степени**.

Для кусочно-линейной интерполяции используется также несколько иная форма записи.

Именно, любому узлу x_k исходной равномерной сетки сопоставляется кусочно-линейная функция $\varphi_k = \varphi_k(x)$.

Ее значение в узле x_k равно 1, а во всех остальных узлах сетки эта функция принимает нулевое значение, то есть $\varphi_k(x_m) = \delta_k^m$, где δ_k^m — это символ Кронекера.

Всего таких базисных функций будет $N + 1$:

$$\varphi_0(x), \quad \varphi_1(x), \quad \varphi_2(x), \quad \dots, \quad \varphi_N(x).$$

Интерполяционная формула при этом принимает вид

$$f(x) \approx f^*(x) = \sum_{k=0}^N f_k \varphi_k(x).$$