

***Федеральное государственное бюджетное образовательное учреждение
высшего образования**

**«НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ МОСКОВСКИЙ
ГОСУДАРСТВЕННЫЙ СТРОИТЕЛЬНЫЙ УНИВЕРСИТЕТ»**

Кафедра Информатики и прикладной математики

Методические указания к компьютерному практикуму

Методы оптимизации

Москва 2024

Для студентов заочного отделения номер варианта задания определяется по последним цифрам зачетной книжки студента, G=1:

Последние цифры	№ S	Последние цифры	№ S	Последние цифры	№ S	Последние цифры	№ S
01	1	26	26	51	21	76	16
02	2	27	27	52	22	77	17
03	3	28	28	53	23	78	18
04	4	29	29	54	24	79	19
05	5	30	30	55	25	80	20
06	6	31	1	56	26	81	21
07	7	32	2	57	27	82	22
08	8	33	3	58	28	83	23
09	9	34	4	59	29	84	24
10	10	35	5	60	30	85	25
11	11	36	6	61	1	86	26
12	12	37	7	62	2	87	27
13	13	38	8	63	3	88	28
14	14	39	9	64	4	89	29
15	15	40	10	65	5	90	30
16	16	41	11	66	6	91	1
17	17	42	12	67	7	92	2
18	18	43	13	68	8	93	3
19	19	44	14	69	9	94	4
20	20	45	15	70	10	95	5
21	21	46	16	71	11	96	6
22	22	47	17	72	12	97	7
23	23	48	18	73	13	98	8
24	24	49	19	74	14	99	9
25	25	50	20	75	15	00	10

ВАРИАЦИОННОЕ ИСЧИСЛЕНИЕ. ПРЯМЫЕ МЕТОДЫ ВАРИАЦИОННОГО ИСЧИСЛЕНИЯ. ОПТИМАЛЬНЫЙ РАСЧЁТ СТРОИТЕЛЬНЫХ КОНСТРУКЦИЙ.

РЕШЕНИЕ ЗАДАЧ ВАРИАЦИОННОГО ИСЧИСЛЕНИЯ ДЛЯ ФУНКЦИОНАЛОВ ПРОСТЕЙШЕГО ВИДА

Пример. Найти экстремаль функционала

$$I = \int_0^{\pi/4} (y'^2 - 4y^2 - 2xe^x y) dx,$$

удовлетворяющую граничным условиям $y(0) = 0$, $y(\pi/4) = 0$.

Частные производные от подынтегральной функции равны:

$$F_y = -8y - 2xe^x, \quad F_{y'} = 2y', \quad F_{y'y'} = 2, \quad F_{yy'} = F_{xy'} = 0.$$

Поэтому уравнение Эйлера запишется в виде

$$2y'' + 8y + 2xe^x = 0 \quad \text{или} \quad y'' + 4y = -xe^x.$$

Это линейное неоднородное дифференциальное уравнение второго порядка с постоянными коэффициентами. Общее решение однородного уравнения будет

$$y_0 = C_1 \cos 2x + C_2 \sin 2x.$$

Частное решение y_* неоднородного уравнения находим методом подбора, принимая его в виде

$$y_* = (Ax + B)e^x.$$

Подставляя y_* в уравнение, находим $A = -1/5$, $B = 2/25$.

Поэтому общее решение уравнения (семейство экстремалей функционала) запишется в виде

$$y_0 = C_1 \cos 2x + C_2 \sin 2x + \frac{1}{5} \left(-x + \frac{2}{5} \right) e^x.$$

Произвольные постоянные C_1 и C_2 находим из граничных условий:

$$C_1 = -\frac{2}{25}, \quad C_2 = \frac{1}{5} \left(\frac{\pi}{4} - \frac{2}{5} \right) e^{\pi/4}.$$

Поэтому искомая экстремаль будет

$$y = -\frac{2}{25} \cos 2x + \frac{5\pi - 8}{100} e^{\pi/4} \sin 2x + \frac{2 - 5x}{25} e^x.$$

РЕШЕНИЕ ЗАДАЧ ВАРИАЦИОННОГО ИСЧИСЛЕНИЯ ДЛЯ ФУНКЦИОНАЛОВ РАЗЛИЧНОГО ТИПА С РАЗЛИЧНЫМИ ГРАНИЧНЫМИ УСЛОВИЯМИ

Пример. Найти экстремаль функционала

$$I = \int_{-1}^1 (2xy - y'^2 + z'^3 / 3) dx,$$

удовлетворяющую граничным условиям

$$y(-1) = 2, \quad y(1) = 0, \quad z(-1) = -1, \quad z(1) = 1.$$

Вычисляем частные производные функции $F = 2xy - y'^2 + \frac{z'^3}{3}$:

$$F_y = 2x, \quad F_{y'} = -2y', \quad F_{yy'} = -2, \quad F_{yy'} = F_{xy'} = 0,$$

$$F_z = 0, \quad F_{z'} = z'^2, \quad F_{z'z'} = 2z', \quad F_{zz'} = F_{xz'} = 0.$$

Поэтому уравнения Эйлера примут вид

$$y'' = -x,$$

$$z'z'' = 0.$$

Эти уравнения независимы.

Интегрируя первое уравнение системы, получим

$$y = -\frac{x^3}{6} + C_1x + C_2.$$

Второе уравнение равносильно совокупности двух

$$\begin{cases} z' = 0, \\ z'' = 0. \end{cases}$$

Общее решение уравнения $z' = 0$: $z = x + C$, очевидно, принадлежит общему решению последнего уравнения $z = C_3x + C_4$.

Определяя произвольные постоянные из граничных условий, находим

$$C_1 = -\frac{5}{6}, \quad C_2 = 1, \quad C_3 = 1, \quad C_4 = 0.$$

Поэтому экстремали данного функционала будут

$$y(x) = -\frac{x}{6}(x^2 + 5) + 1,$$

$$z(x) = x.$$

Пример. Найти экстремаль функционала

$$I = \int_0^{\pi/2} (2yz - 2y^2 + y'^2 - z'^2) dx,$$

удовлетворяющую граничным условиям

$$y(0) = 0, \quad y(\pi/2) = 0, \quad z(0) = 0, \quad z(\pi/2) = 1.$$

Система уравнений Эйлера для данного функционала приводится к виду

$$\begin{cases} y'' + 2y = z, \\ z'' = -y. \end{cases}$$

Решаем её методом исключения. Дифференцируя дважды первое из уравнений и подставляя z'' из второго уравнения, получим однородное линейное уравнение 4-го порядка относительно y :

$$y^{IV} + 2y'' + y = 0,$$

общее решение которого будет

$$y = C_1 \cos x + C_2 \sin x + C_3 x \cos x + C_4 x \sin x.$$

Из граничных условий относительно функции $y(x)$ находим

$$C_1 = 0, \quad C_4 = -\frac{2}{\pi} C_2. \quad \text{Поэтому}$$

$$y = C_2 \sin x + C_3 x \cos x - \frac{2}{\pi} C_2 x \sin x.$$

Далее из первого уравнения системы следует

$$z = y'' + 2y = (C_2 - 2C_3) \sin x + C_3 x \cos x - \frac{4}{\pi} C_2 \cos x - \frac{2}{\pi} C_2 x \sin x.$$

Из граничных условий относительно $z(x)$ находим $C_2 = 0$,

$$C_3 = -\frac{1}{2}. \quad \text{Следовательно, } z = \sin x - \frac{1}{2}x \cos x.$$

Итак, экстремали данного функционала

$$y = -\frac{1}{2}x \cos x, \quad z = \sin x - \frac{1}{2}x \cos x.$$

РЕШЕНИЕ ЗАДАЧ ВАРИАЦИОННОГО ИСЧИСЛЕНИЯ НА УСЛОВНЫЙ ЭКСТРЕМУМ.

Пример. Найти экстремаль функционала

$$I = \int_0^{\pi/4} (y'^2 - 4y^2) dx, \quad \text{удовлетворяющую} \quad \text{граничным} \quad \text{условиям}$$

$$y(0) = 0, \quad y(\pi/4) = 0 \quad \text{и интегральной связи} \quad K = 4 \int_0^{\pi/4} y dx = 3.$$

Составляем расширенный функционал $I^* = \int_0^{\pi/4} (y'^2 - 4y^2 + 4\lambda y) dx.$

Для него уравнение Эйлера приводится к виду $y'' + 4y = 2\lambda.$ Его общее решение

$$y = C_1 \cos 2x + C_2 \sin 2x + \frac{\lambda}{2}.$$

Подставляя это решение в граничные условия, получим

$$C_1 = C_2 = -\frac{\lambda}{2}.$$

$$y = \frac{\lambda}{2}(1 - \cos 2x - \sin 2x).$$

Постоянную λ определяем теперь из условия

$$K = 4 \int_0^{\pi/4} y dx = 4 \int_0^{\pi/4} \frac{\lambda}{2}(1 - \cos 2x - \sin 2x) dx = 3.$$

Интегрируя, находим неопределённый множитель $\lambda = \frac{6}{\pi - 4}.$

Поэтому искомая экстремаль будет

$$y = \frac{3}{\pi - 4} (1 - \cos 2x - \sin 2x).$$

Пример. Найти экстремаль функционала $I = \int_0^{\pi/2} y \cos x dx$, удовлетворяющую граничным условиям $y(0) = 0$, $y(\pi/2) = \pi$ и интегральной связи $K = \int_0^{\pi/2} y'^2 dx = \frac{5\pi}{2}$.

Составляем функцию Лагранжа. Так как $F = y \cos x$, $G = y'^2$, то $L = y \cos x + \lambda y'^2$.

Составляем уравнение Эйлера для расширенного функционала

$$I^* = \int_0^{\pi/2} L dx = \int_0^{\pi/2} (y \cos x + \lambda y'^2) dx.$$

Так как $L_y = \cos x$, $L_{y'} = 2\lambda y'$, $L_{y'y'} = 2\lambda$, то уравнение Эйлера приводится к виду $y'' = \frac{1}{2\lambda} \cos x$. Интегрируя его методом понижения порядка, получим общее решение в виде

$$y = -\frac{1}{2\lambda} \cos x + C_1 x + C_2.$$

Из уравнения связи имеем

$$\int_0^{\pi/2} y'^2 dx = \int_0^{\pi/2} \left(C_1 + \frac{1}{2\lambda} \sin x \right)^2 dx = \frac{\pi}{2} C_1^2 + \frac{1}{\lambda} C_1 + \frac{\pi}{16\lambda^2} = \frac{5\pi}{2}.$$

Из граничных условий следует

$$-\frac{1}{2\lambda} + C_2 = 0, \quad \frac{\pi}{2} C_1 + C_2 = \pi.$$

В результате получаем систему трёх уравнений относительно трёх неизвестных C_1 , C_2 , λ :

$$\begin{cases} -\frac{1}{2\lambda} + C_2 = 0, \\ \frac{\pi}{2} C_1 + C_2 = \pi, \\ \frac{\pi}{2} C_1^2 + \frac{1}{\lambda} C_1 = \frac{5\pi}{2} - \frac{\pi}{16\lambda^2}. \end{cases}$$

Из первых двух уравнений находим $C_1 = 2 - \frac{1}{\pi\lambda}$, $C_2 = \frac{1}{2\lambda}$.

Подставляя C_1 и C_2 в третье уравнение системы, получим два значения $\lambda = \pm \frac{\sqrt{\pi^2 - 8}}{2\sqrt{2}\pi}$.

При $\lambda = \frac{\sqrt{\pi^2 - 8}}{2\sqrt{2}\pi}$ $C_1 = 2 \left(1 - \frac{\sqrt{2}}{\sqrt{\pi^2 - 8}} \right)$, $C_2 = \frac{\pi\sqrt{2}}{\sqrt{\pi^2 - 8}}$, и уравнение экстремали будет

$$y = \frac{1}{\sqrt{\pi^2 - 8}} [-\pi\sqrt{2} \cos x + 2(\sqrt{\pi^2 - 8} - \sqrt{2})x + \pi\sqrt{2}].$$

При $\lambda = -\frac{\sqrt{\pi^2 - 8}}{2\sqrt{2}\pi}$ $C_1 = 2 \left(1 + \frac{\sqrt{2}}{\sqrt{\pi^2 - 8}} \right)$, $C_2 = -\frac{\pi\sqrt{2}}{\sqrt{\pi^2 - 8}}$, и уравнение экстремали запишется в виде

$$y = \frac{1}{\sqrt{\pi^2 - 8}} [\pi\sqrt{2} \cos x + 2(\sqrt{\pi^2 - 8} + \sqrt{2})x - \pi\sqrt{2}].$$

ОПТИМАЛЬНОЕ ПРОЕКТИРОВАНИЕ БАЛОК И СТЕРЖНЕЙ ПЕРЕМЕННОГО СЕЧЕНИЯ.

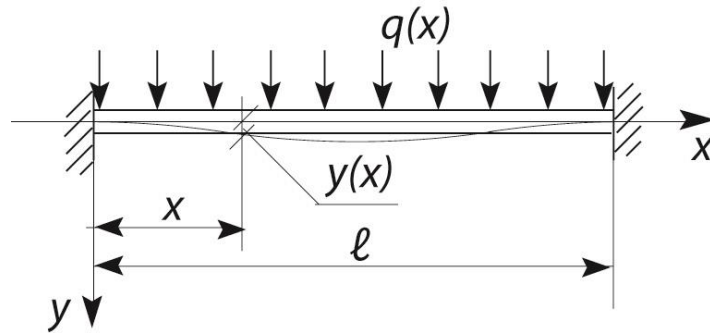
Изгиб защемлённого стержня

Граничные условия для защемлённого по обоим концам стержня могут быть записаны в виде

$$y(0) = y(l) = y'(0) = y'(l) = 0.$$

Решение методом Ритца

Ограничиваясь для простоты одночленной аппроксимацией решения, примем выражение для прогиба в виде



$$y(x) = a_1 \left(1 - \cos \frac{2\pi x}{l} \right). \quad (4.22)$$

При этом удовлетворяются геометрические граничные условия.

Выражение для функционала полной энергии стержня при изгибе

$$\begin{aligned} \mathfrak{D} &= \int_0^l \left[\frac{EI}{2} a_1^2 \left(\frac{2\pi}{l} \right)^4 \cos^2 \frac{2\pi x}{l} - q(x) a_1 \left(1 - \cos \frac{2\pi x}{l} \right) \right] dx = \\ &= \frac{8\pi^4 EI}{l^4} a_1^2 \int_0^l \frac{1}{2} (1 + \cos \frac{4\pi x}{l}) dx - a_1 \int_0^l q(x) \left(1 - \cos \frac{2\pi x}{l} \right) dx. \end{aligned}$$

В случае равномерно распределённой нагрузки $q(x) = q_0 = \text{const}$ и после

интегрирования получим $\mathfrak{D} = \frac{4\pi^4 EI}{l^3} a_1^2 - q_0 a_1 l$.

Из условия $\frac{d\mathfrak{D}}{da_1} = 0$ находим $a_1 = \frac{q_0 l^4}{8\pi^4 EI}$.

Поэтому решение задачи запишется в виде

$$y = \frac{q_0 l^4}{8\pi^4 EI} \left(1 - \cos \frac{2\pi x}{l} \right).$$

В частности, максимальный прогиб в середине пролета будет равен

$$y_{\max} = y\left(\frac{l}{2}\right) = \frac{q_0 l^4}{4\pi^4 EI} \approx \frac{q_0 l^4}{396 EI}.$$

Как видно, полученное решение даже в одночленном приближении оказывается достаточно близким к точному решению для защемлённого стержня:

$$y_{\max} = \frac{q_0 l^4}{384 EI}.$$

Решение методом Галеркина

$$L(y) = -\frac{16\pi^4 EI}{l^4} a_1 \cos \frac{2\pi x}{l} - q(x).$$

Ортогонализируя её к выбранной координатной функции приходим к уравнению

$$\int_0^l \left(\frac{16\pi^4 EI}{l^4} a_1 \cos \frac{2\pi x}{l} + q(x) \right) \left(1 - \cos \frac{2\pi x}{l} \right) dx = 0.$$

Выполняя интегрирование при $q(x) = q_0 = \text{const}$, находим значение

$$a_1 = \frac{q_0 l^4}{8\pi^4 EI}, \text{ совпадающее с полученным методом Ритца.}$$

Поэтому окончательное решение задачи методом Галеркина а запишется, как и в методе Ритца, в виде

$$y = \frac{q_0 l^4}{8\pi^4 EI} \left(1 - \cos \frac{2\pi x}{l} \right).$$

РЕШЕНИЕ ЗАДАЧ ВАРИАЦИОННОГО ИСЧИСЛЕНИЯ С ПОМОЩЬЮ МЕТОДА РИТЦА, МЕТОДА ГАЛЕРКИНА И ДР.

Решение задач вариационного исчисления с помощью метода Ритца, метода Галеркина

Задание. Найти решение краевой задачи.

$$y'' + k_1 y' + k_0 y = f(x),$$

Граничные условия

$$y(x_a) = G$$

$$y(x_b) = S$$

$$x_a = 0, \quad x_b = 1$$

$$k_1 = 0, \quad k_0 = \left(2 + \frac{S}{20}\right)^2, \quad f(x) = Sx$$

Демидович Б.П., Марон И.А., Шувалова Э.З. Численные методы анализа. 3-е изд. 1967 стр.257 и далее

Ручной счет (с элементами автоматизации на Python)

Выберем для примера номер группы $G=6$, номер студента $S=2$

$$y'' + \left(2 + \frac{S}{20}\right)^2 y = Sx, \quad \text{Граничные условия } y(0) = G, \quad y(1) = S$$

$$y'' + 2.1^2 y = 2x, \quad \text{Граничные условия } y(0) = 6, \quad y(1) = 2$$

Решение ищем в виде полного многочлена (полинома) 3-ей степени

$$y(x) = \alpha_0 + \alpha_1 x + \alpha_2 x^2 + \alpha_3 x^3. \quad \text{Число слагаемых } n=4.$$

Сначала удовлетворим наше решение двум граничным условиям

$$\begin{cases} 6 = \alpha_0 + \alpha_1 \cdot 0 + \alpha_2 \cdot 0^2 + \alpha_3 \cdot 0^3 \\ 2 = \alpha_0 + \alpha_1 \cdot 1 + \alpha_2 \cdot 1^2 + \alpha_3 \cdot 1^3 \end{cases} \Rightarrow \begin{cases} \alpha_0 = 6 \\ \alpha_0 + \alpha_1 + \alpha_2 + \alpha_3 = 2 \end{cases}$$

Выражаем два коэффициента, через остальные два. Наверное правильное было бы здесь сказать, что мы приводим систему к ступенчатому виду (здесь уже изначально такой вид существует, правда не привычный нижнетреугольный, а верхнетреугольный) и, например оставляем слева неизвестные α_0, α_1 при не равном

нулю определителе $\begin{vmatrix} 1 & 0 \\ 1 & 1 \end{vmatrix}$, а направо переносим неизвестные α_2, α_3 . И таким образом

мы смогли выразить зависимые переменные α_0, α_1 , через независимые α_2, α_3 .

Итак, как мы выяснили, удобнее всего здесь выразить α_0, α_1 через α_2, α_3

$$\alpha_0 = 6$$

$$\alpha_1 = 2 - 6 - \alpha_2 - \alpha_3, \quad \text{далее подставляем их в } y(x) = \alpha_0 + \alpha_1 x + \alpha_2 x^2 + \alpha_3 x^3$$

$$y(x) = 6 + (-4 - \alpha_2 - \alpha_3)x + \alpha_2 x^2 + \alpha_3 x^3, \quad \text{приводим к коэффициентам } \alpha_2, \alpha_3$$

$$y(x) = 6 - 4x + \alpha_2(-x + x^2) + \alpha_3(-x + x^3)$$

Изменим имена коэффициентов α на c и перенумеруем

$$y(x) = 6 - 4x + c_1(-x + x^2) + c_2(-x + x^3)$$

Введем обозначения $\phi_0(x) = 6 - 4x$; $\phi_1(x) = -x + x^2$; $\phi_2(x) = -x + x^3$

с учетом обозначений наше решение будет $y(x) = \phi_0(x) + c_1\phi_1(x) + c_2\phi_2(x)$.

Интересно, что всегда при подобных преобразованиях выше

$\phi_0(x)$ будет удовлетворять исходным (неоднородным) ГУ, то есть , а все остальные $\phi_i(x)$ будут удовлетворять однородным ГУ (в исходных мы обнулили правую часть)
 $\phi_i(0) = 0, \phi_i(1) = 0, i = 1, 2, \dots$

Конечно, если мы не ошиблись выше в поисках $\phi_0(x), \phi_1(x), \phi_2(x)$, наше решение $y(x) = \phi_0(x) + c_1\phi_1(x) + c_2\phi_2(x)$ будет удовлетворять исходным (неоднородным) ГУ
 $y(0) = 6, y(1) = 2$

Функции $\phi_1(x), \phi_2(x)$ при коэффициентах c_1, c_2 в разложении $y(x) = \phi_0(x) + c_1\phi_1(x) + c_2\phi_2(x)$, мы будем называть базисными, как это было в аналогичных решениях СЛАУ (в ФСР – в Фундаментальной системе решений)

Найдем сначала точное решение $y'' + 2.1^2 y = 2x, y(0) = 6, y(1) = 2$,

Это линейное дифференциальное уравнение 2-го порядка с постоянными коэффициентами и со специальной правой частью

Решаем его или вручную или с помощью функции dsolve

```
# Точное решение краевой задачи Boundary value problem (Exact Solution)
import sympy as sy
import numpy as np
y = sy.Function('y');
x = sy.symbols('x')
deq = sy.Derivative(y(x), x, x) + sy.simplify(2.1)**2*y(x) - 2*x
dsol=sy.dsolve(deq, y(x)).rhs
C=sy.solve([dsol.subs(x,0)-6,dsol.subs(x,1)-2])
print(C)
y=dsol.subs(C)
print('y(x)= ',y)
print('y(0.5)= ', y.subs(x,0.5)) # найдем точное значение посредине двух границ y(0.5)
print('y(0.5)= ', np.float64(y.subs(x,0.5))) # найдем точное значение посредине двух границ y(0.5)

{C1: (682 - 2646*cos(21/10))/(441*sin(21/10)), C2: 6}
y(x)= 200*x/441 + (682 - 2646*cos(21/10))*sin(21*x/10)/(441*sin(21/10)) + 6*cos(21*x/10)
y(0.5)= 3.21218365696487 + 0.00196694608978235*(682 - 2646*cos(21/10))/sin(21/10)
y(0.5)= 7.810081693594386
```

Метод Рунца. Подставляем $y(x) = \phi_0(x) + c_1\phi_1(x) + c_2\phi_2(x)$ в функционал

$$I(c_1, c_2) = \int_0^1 F(x, y(x), y'(x), c_1, c_2) dx \rightarrow \text{extr}, \text{ где}$$

$$F(x, y(x), y'(x), c_1, c_2) = \frac{y'^2}{2} - 2.1^2 \frac{y^2}{2} + 2xy. \text{ Снова получаем невязку } R(x, c_1, c_2)$$

$I(c_1, c_2) \rightarrow \text{extr}$, и решаем задачу на экстремум функции двух переменных $I(c_1, c_2)$ (здесь это будет минимум)

$$\begin{cases} \frac{\partial I(c_1, c_2)}{\partial c_1} = 0 \\ \frac{\partial I(c_1, c_2)}{\partial c_2} = 0 \end{cases}, \text{ решаем систему и находим } c_1, c_2$$

```
# метод Рунца
import sympy as sy
sy.var('x, c1, c2')
y=6-4*x+c1*(-x+x**2)+c2*(-x+x**3)
F=sy.diff(x)**2/2-2.1**2*y**2/2+y*2*x # для y'+2.1^2*y-2x
print(F)
J=sy.integrate(F, (x,0,1))
print(J)
eqs=[J.diff(c1), J.diff(c2)]
print(eqs)
C=sy.solve(eqs)
print(C)
y=y.subs(C)
print(y)
print(y.subs(x,0.5))
```

```
x*(2*c1*(x**2 - x) + 2*c2*(x**3 - x) - 8*x + 12) + (c1*(2*x - 1) + c2*(3*x**2 - 1) - 4)**2/2 - 2.205*(c1*(x**2 - x) + c2*(x**3 - x) - 4*x + 6)**2
0.09316666666666666*c1**2 + 0.2795*c1*c2 + 2.773333333333333*c1 + 0.232*c2**2 + 3.996333333333333*c2 - 26.886666666666667
{0.18633333333333333*c1 + 0.2795*c2 + 2.773333333333333, 0.2795*c1 + 0.464*c2 + 3.996333333333333}
{c1: -20.3697544497861, c2: 3.65735567970235}
3.65735567970235*x**3 - 20.3697544497861*x**2 + 12.7123987700838*x + 6
7.72093023255815
```

ЛИНЕЙНОЕ ПРОГРАММИРОВАНИЕ

СИМПЛЕКС-АЛГОРИТМ РЕШЕНИЯ ЗАДАЧ ЛИНЕЙНОГО ПРОГРАММИРОВАНИЯ.

ЗАДАЧА ЛИНЕЙНОГО ПРОГРАММИРОВАНИЯ

Задание: Решить задачу линейного программирования. Постановка задачи (вариант G=____, S=____)

Найти максимум и точку максимума функции Z

$$Z = Sx_1 + Gx_2$$

при ограничениях

$$\begin{cases} -\frac{x_1}{G} - \frac{x_2}{G} + \frac{1}{4} \leq 0 \\ -\frac{x_1}{G} - \frac{x_2}{G} + \frac{1}{4} \leq 0 \\ 2Sx_1 - Gx_1 - GS \leq 0 \\ x_1 \geq 0 \\ x_2 \geq 0 \end{cases}$$

1. Решить задачу на ЭВМ с помощью стандартной подпрограммы linprog.
2. Решить задачу вручную геометрическим методом.
3. Решить задачу вручную симплекс-методом

Симплекс-метод – универсальный метод решения задач линейного программирования

Суть метода - целенаправленный перебор

Симплекс-метод был разработан и впервые применен для решения задач в 1947 г. американским математиком Дж. Даницигом

Процесс применения симплексного метода предполагает реализацию трех основных элементов

- 1) способ определения какого-либо первоначального допустимого базисного решения задачи
- 2) правило перехода к лучшему (точнее, не худшему) решению
- 3) критерий проверки оптимальности найденного решения

Пример 1. Решить симплекс-методом задачу:

$$Z = 2x_1 + x_2 - x_3 + x_4 - x_5 \rightarrow \max$$

при ограничениях:

$$\begin{cases} x_1 + x_2 + x_3 = 5, \\ 2x_1 + x_2 + x_4 = 9, \\ x_1 + 2x_2 + x_5 = 7, \\ x_j \geq 0, \quad j = 1, 2, \dots, 5 \end{cases}$$

и дать геометрическую интерпретацию процесса решения.

Решение. Прежде всего необходимо перейти к минимизации целевой функции. Для этого введем в рассмотрение функцию

$$Z_1 = -Z.$$

Тогда $Z_1 = -2x_1 - x_2 + x_3 - x_4 + x_5 \rightarrow \min.$

В системе ограничений уже выделены базисные переменные x_3 , x_4 и x_5 . Теперь все исходные данные поместим в таблицу:

БП	Переменные					Свободный член
	x_1	x_2	x_3	x_4	x_5	
x_3	1	1	<u>1</u>	0	0	5
x_4	2	1	0	1	0	9
x_5	1	2	0	0	1	7
Z_1	-2	-1	1	-1	1	0

В клетку на пересечении Z_1 -строки и столбца “Свободный член” запишем свободный член целевой функции с противоположным знаком. В данном случае он равен 0.

Приравняв свободные переменные нулю $x_1=x_2=0$, получим базисное решение системы ограничений:

$$X^1 = (0, 0, 5, 9, 7)^T,$$

где $x_3=5$, $x_4=9$, $x_5=7$.

Так как $x_j \geq 0$, $j = 1, 2, \dots, 5$, то базисное решение X^1 является опорным решением (планом) задачи. Итак, первоначальное опорное решение найдено.

Для проверки его на оптимальность выразим функцию цели Z_1 через свободные переменные x_1 и x_2 . Для исключения переменной x_3 из функции цели Z возьмем за разрешающий элемент $a_{13}=1$ в третьем столбце предыдущей таблицы. Сделав один шаг преобразований Жордана–Гаусса, придем к таблице вида:

БП	Переменные					Свободный член
	x_1	x_2	x_3	x_4	x_5	
x_3	1	1	1	0	0	5
x_4	2	1	0	1	0	9
x_5	1	2	0	0	1	7
Z_1	-3	-2	0	-1	1	-5

Сделав аналогично еще два шага исключения неизвестных x_4 и x_5 из функции цели Z_1 , придем к таблице:

БП	Переменные					Свободный член
	x_1	x_2	x_3	x_4	x_5	
x_1	1	1	1	0	0	5
x_4	2	1	0	1	0	9

x_5	1	$\boxed{2}$	0	0	0	7
Z_1	-2	-3	0	0	0	-3

Из таблицы видно, что

$$Z_1(X^{(1)})=3, \quad r=(-2; -3).$$

Так как условие $r \geq 0$ не выполняется, то полученный опорный план $X^{(1)}$ не является оптимальным.

Необходимо перейти к новому опорному плану, для которого значение функции уменьшится. Для перехода к следующему опорному плану нужно одну из свободных переменных (x_1 либо x_2) перевести в базис, а одну из базисных (x_3 , либо x_4 , либо x_5) перевести в свободные. Выбор свободной переменной, вводимой в базис, определяется максимальной по модулю отрицательной компонентой вектора $r = (-2; -3)$.

Компоненте $r_2 = -3$ соответствует в таблице переменная x_2 , которая вводится в базис, и второй столбец становится разрешающим.

Для выбора разрешающей строки вычислим

$$\min \left\{ \frac{5}{1}, \frac{9}{1}, \frac{7}{2} \right\} = \frac{7}{2}.$$

Итак, третья строка стала разрешающей, а разрешающим элементом стал элемент $a_{32}=2$. Сделав один шаг в симплекс-таблице, получим новое опорное решение:

$$X^{(2)} = (0, 7/2, 3/2, 11/2, 0)^T.$$

Это решение также не является оптимальным, так как в вектор $r = (-1/2; 3/2)$ имеет отрицательную компоненту:

БП	Переменные					Свободный член
	x_1	x_2	x_3	x_4	x_5	
x_3	$\boxed{1/2}$	0	1	0	-1/2	3/2
x_4	3/2	0	0	1	-1/2	11/2
x_5	1/2	1	0	0	1/2	7/2
Z_1	-1/2	0	0	0	3/2	15/2

Значение же функции цели Z_1 уменьшилось от значения $Z_1 = 3$ до значения $Z_1 = -15/2$. (Напомним, что значение функции цели из таблицы берется с противоположным знаком.)

Очевидно, что на следующем шаге необходимо в базис ввести переменную x_1 , соответствующую отрицательной компоненте $r_1 = -1/2$. И первый столбец в последней таблице становится разрешающим. Для выбора разрешающей строки найдем:

$$\min \left\{ \frac{3/2}{1/2}, \frac{11/2}{3/2}, \frac{7/2}{1/2} \right\} = 3.$$

Минимальное из отношений соответствует первой строке. Сделав один шаг симплексных преобразований с разрешающим элементом $a_{11}=1/2$, получим таблицу:

БП	Переменные					Свободный член
	x_1	x_2	x_3	x_4	x_5	
x_3	1	0	2	0	-1	3
x_4	0	0	-3	1	-2	1
x_5	0	1	-1	0	1	2
Z_1	0	0	1	0	1	9

Из таблицы следует, что полученный опорный план:

$$X^{(3)} = (3, 2, 0, 1, 0)^T$$

является единственным оптимальным планом задачи, так как $r = (1; 1) > 0$. При этом значение функции уменьшилось:

$$Z_{1min} = Z_1(X^{(3)}) = -9.$$

Итак, решением исходной задачи является:

$$X^* = (3, 2, 0, 1, 0)^T, \text{ а } Z_{max} = -Z_{1min} = 9.$$

Дадим геометрическую интерпретацию процесса нахождения решения. Для этого прежде всего перейдем от канонической формы модели:

$$Z = 2x_1 + x_2 - x_3 + x_4 - x_5 \rightarrow \max$$

при ограничениях:

$$\begin{cases} x_1 + x_2 + x_3 = 5, \\ 2x_1 + x_2 + x_4 = 9, \\ x_1 + 2x_2 + x_5 = 7, \end{cases}$$

$$x_j \geq 0, \quad j = 1, 2, \dots, 5$$

к стандартной форме модели данной задачи. Это нетрудно сделать, так как система ограничений задачи уже приведена к единичному базису.

Выразим из системы ограничений базисные переменные через свободные:

$$x_3 = 5 - x_1 - x_2 \geq 0,$$

$$x_4 = 9 - 2x_1 - x_2 \geq 0,$$

$$x_5 = 7 - x_1 - 2x_2 \geq 0,$$

и, подставив вместо x_3, x_4, x_5 их выражения в функцию цели Z , получим:

$$Z = -3 + 2x_1 + 3x_2 \rightarrow \max.$$

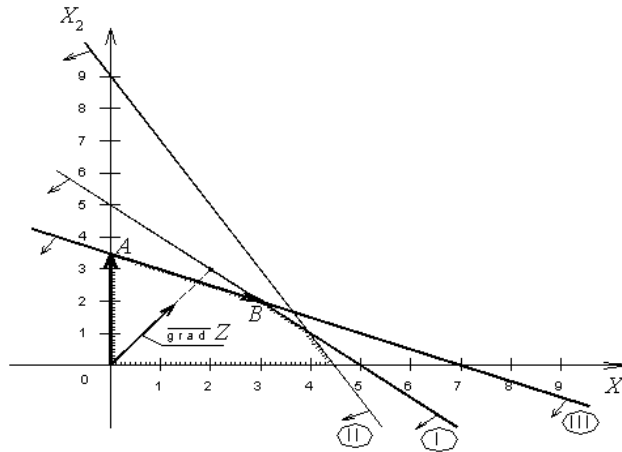
Стандартная форма модели примет вид

$$Z = -3 + 2x_1 + 3x_2 \rightarrow \max$$

при ограничениях:

$$\begin{cases} x_1 + x_2 \leq 5, \\ 2x_1 + x_2 \leq 9, \\ x_1 + 2x_2 \leq 7, \\ x_1, x_2 \geq 0. \end{cases}$$

Решение последней задачи можно найти, используя геометрическую интерпретацию решения ЗЛП, приведенную на рисунке.



Геометрическая интерпретация симплекс-метода

Исходный опорный план

$$X^{(1)} = (0, 0, 5, 9, 7)^T$$

соответствует точке $O(0,0)$ ОДР.

После одного шага симплекс-метода (одной итерации) был получен новый опорный план

$$X^{(2)} = (0, \frac{7}{2}, \frac{3}{2}, \frac{11}{2}, 0)^T,$$

которому соответствует точка $A(0, \frac{7}{2})$.

На рисунке переход от одного опорного плана (вершины) к другому опорному плану происходит по ребру OA .

После второй итерации получен план

$$X^* = (3, 2, 0, 1, 0)^T,$$

который является оптимальным. Он соответствует на чертеже точке B , т.е. осуществлен переход от точки A к точке B по ребру AB .

Пример 2. Решить симплекс-методом ЗЛП:

$$Z = 2x_1 + x_2 + x_4 \rightarrow \max,$$

$$\begin{cases} x_1 + x_2 + x_3 + x_4 = 2, \\ x_1 - x_2 - x_3 = 0, \\ x_1 + x_2 + 2x_3 = 1, \\ x_j \geq 0, \quad (j = \overline{1,4}). \end{cases}$$

Решение. Особенностью данной задачи является то, что система ограничений не приведена к базисному виду, нет разграничения переменных на базисные и свободные (есть только одна базисная переменная x_4). Поэтому прежде всего с помощью преобразований Жордана–Гаусса приведем систему к базисному виду. Для этого коэффициенты системы ограничений и целевой функции $Z_1 = -Z$ занесем в таблицу:

БП	Переменные				Свободный член
	x_1	x_2	x_3	x_4	
x_4	1	1	1	1	2
	<u>1</u>	-1	-1	0	0
	1	1	2	0	1
Z_1	-2	-1	0	-1	0

За разрешающий столбец выберем любой столбец, содержащий хотя бы один положительный элемент, например первый.

Выбор разрешающей строки осуществим по тому же правилу, что при переходе от одного опорного решения к другому. Это правило гарантирует, что свободные члены новой таблицы останутся неотрицательными.

Имеем $\min \left\{ \frac{2}{1}, \frac{0}{-1}, \frac{1}{1} \right\} = \frac{0}{1}$,

т.е. за разрешающую строку примем вторую, а за разрешающий элемент $a_{21}=1$. Выполнив шаг преобразований Жордана–Гаусса, придем к таблице:

БП	Переменные				Свободный член
	x_1	x_2	x_3	x_4	
x_4	0	2	2	1	2
x_1	1	-1	-1	0	0
	0	<u>2</u>	3	0	1
Z_1	0	-3	-2	-1	0

Система, записанная в данной таблице, разрешена относительно x_1 и x_4 , но так как она содержит три уравнения, то необходимо ввести в базис еще одну переменную (x_2 или x_3) и при этом иметь в качестве разрешающей третью строку. Посмотрим, какой столбец с этой точки зрения следует брать за разрешающий. Если возьмем второй

столбец, то за разрешающую строку следует выбрать третью (благоприятный случай):

$$\min \left\{ \frac{2}{2}, \frac{1}{2} \right\} = \frac{1}{2}.$$

Если же возьмем третий столбец, то за разрешающую строку также следует выбрать третью:

$$\min \left\{ \frac{2}{2}, \frac{1}{3} \right\} = \frac{1}{3},$$

следовательно, можно выбрать любой из них. Выберем, например, второй столбец и, выполнив шаг преобразований Жордана–Гаусса с разрешающим элементом $a_{32} = 2$, придем к таблице:

Б П	Переменные				Свобо дный член
	x_1	x_2	x_3	x_4	
x_4	0	0	-1	<u>1</u>	1
x_1	1	0	$\frac{1}{2}$	0	1/2
x_2	0	1	$\frac{3}{2}$	0	1/2
Z_1	0	0	$\frac{5}{2}$	-1	3/2

Теперь система приведена к базисному виду, базисными переменными являются x_1, x_2, x_4 , свободной – x_3 . Все свободные члены положительны, поэтому можем записать первоначальный опорный план:

$$X_{on}^1 = (1/2, 1/2, 0, 1)^T.$$

Чтобы ответить, является ли он оптимальным, необходимо выполнить еще один шаг преобразований с разрешающим элементом $a_{14} = 1$, чтобы выразить целевую функцию Z_1 только через свободную переменную x_3 . Получим новую таблицу:

Б П	Переменные				Своб од- ный член
	x_1	x_2	x_3	x_4	
x_4	0	0	-1	1	1
x_1	1	0	$\frac{1}{2}$	0	1/2
x_2	0	1	$\frac{3}{2}$	0	1/2
Z_1	0	0	$\frac{3}{2}$	0	5/2

Опорное решение

$$X_{on}^1 = (1/2, 1/2, 0, 1)^T,$$

содержащееся в таблице, является оптимальным, так как вектор $r = (3/2)$ содержит одну положительную компоненту. Минимальное значение целевой функции:

$$Z_{\min} = -5/2 \text{ и, очевидно, } Z_{\max} = 5/2.$$

Итак, $X^* = (1/2, 1/2, 0, 1)^T$, $Z_{\max} = 5/2$.

```
: # coding: cp1251
# Импорт пакета
from scipy import optimize
import numpy as np
# Целевая функция
c = np.array([3,12])
# левая часть ограничений
A_ub = np.array([[ -1/12, -1/3], [-3,24],[ 6,-12]])
# правая часть ограничений
B_ub = np.array([ -0.25,36,36])
# Стандартная функция linprog ищет минимум
# Для поиска максимума целевую функцию следует умножить на -1
# то есть сменить знак!
# В полученном ответе знак «вернуть» обратно
res = optimize.linprog(-c,A_ub,B_ub, method='highs') # -с т.к. max (а не min)
print(res)
```

```
message: Optimization terminated successfully. (HiGHS Status 7: Optimal)
success: True
status: 0
  fun: -72.0
   x: [ 1.200e+01  3.000e+00]
  nit: 0
lower: residual: [ 1.200e+01  3.000e+00]
      marginals: [ 0.000e+00  0.000e+00]
upper: residual: [          inf          inf]
      marginals: [ 0.000e+00  0.000e+00]
eqlin: residual: []
      marginals: []
ineqlin: residual: [ 1.750e+00  0.000e+00  0.000e+00]
        marginals: [-0.000e+00 -1.000e+00 -1.000e+00]
mip_node_count: 0
mip_dual_bound: 0.0
mip_gap: 0.0
```

МЕТОДЫ РЕШЕНИЯ ТРАНСПОРТНЫХ ЗАДАЧ

Пример

Составить план перевозок грузов с наименьшей стоимостью от четырёх поставщиков соответственно в количествах 100, 250, 200, 300 единиц к пяти потребителям соответственно в количествах 200, 200, 100, 100, 250 единиц; стоимости перевозок единицы груза приведены в табл. 3.

Таблица 1

Поставщики	Потребители					Запасы
	B_1	B_2	B_3	B_4	B_5	
A_1	10	7	4	1	4	100
A_2	2	7	10	6	11	250
A_3	8	5	3	2	2	200
A_4	11	8	12	16	13	300
Потребности	200	200	100	100	250	850

В табл. 2 отображено построение первоначального опорного плана методом северо-западного угла.

Таблица 2

Поставщики	Потребители					Запасы
	B_1	B_2	B_3	B_4	B_5	
A_1	10 100	7 —	4 —	1 —	4 —	100
A_2	2 100	7 150	10 —	6 —	11 —	250
A_3	8 —	5 50	3 100	2 50	2 —	200
A_4	11 —	8 —	12 —	16 50	13 250	300
Потребности	200	200	100	100	250	850

Без учета стоимости перевозки единицы груза, начинается удовлетворение потребностей первого потребителя B_1 за счёт запаса поставщика A_1 . Для этого сравнивается $a_1=100$ с $b_1=200$, $a_1 < b_1$, меньший из объёмов, т.е. = 100 ед. записываем в левый нижний угол клетки A_1B_1 . Запасы первого поставщика полностью израсходованы, поэтому остальные клетки первой таблицы прочёркиваем. Потребности B_1 остались неудовлетворенными на $200 - 100 = 100$ ед. Сравнивая

этот остаток с запасами поставщика A_2 (так как $100 < 250$), записывается в клетку A_2B_1 , чем полностью удовлетворяются потребности потребителя B_1 , а оставшиеся клетки в первом столбце прочёркиваются.

У поставщика A_2 осталось 150 ед. груза. Потребности потребителя B_2 удовлетворяются за счёт оставшегося у поставщика A_2 груза. Для этого сравнивается этот остаток с потребностями потребителя B_2 : $150 < 200$, записывается 150 ед. в клетку A_2B_2 и, так как запасы A_2 полностью израсходованы, прочёркиваются остальные клетки второй строки. Потребности B_2 остались неудовлетворёнными на 50 ед. они удовлетворяются за счёт поставщика A_3 и затем удовлетворяются потребности B_3 за счёт остатка, имеющегося у поставщика A_3 , и т.д. процесс продолжается до тех пор, пока не будут удовлетворены все потребители за счёт запасов поставщиков. На этом построение первоначального опорного плана заканчивается.

Таким образом, в табл.4 в правых верхних углах клеток стоят числа, определяющие стоимость перевозки единицы грузов, а в левых нижних углах – числа, определяющие план перевозок, так как их сумма по столбцам – потребности соответствующего потребителя.

Для проверки, является ли план, построенный в табл.4, опорным, нужно просмотреть, можно ли получить цикл. Так, начиная движение от занятой клетки A_1B_1 , невозможно вернуться не только в неё, но и в любую другую занятую клетку, двигаясь только по занятым клеткам.

Общая стоимость составленного опорного плана находится как сумма произведений объёмов перевозок, стоящих в левом верхнем углу занятых клеток, на соответствующие стоимости в этих же клетках:

$$Z = 100 \cdot 10 + 100 \cdot 2 + 150 \cdot 7 + 50 \cdot 5 + 100 \cdot 3 + 50 \cdot 2 + 50 \cdot 16 + 250 \cdot 13 = 6950 (\text{ед.стоимости})$$

Если при составлении опорного плана учитываются стоимости перевозок единицы груза, то такой план будет значительно ближе к оптимальному.

МЕТОДЫ РЕШЕНИЯ НЕЛИНЕЙНЫХ ЗАДАЧ МАТЕМАТИЧЕСКОГО ПРОГРАММИРОВАНИЯ.

МЕТОД СОПРЯЖЁННЫХ НАПРАВЛЕНИЙ С ИСПОЛЬЗОВАНИЕМ МАТРИЦЫ ГЕССЕ (МЕТОД ДЭВИДОНА — ФЛЕТЧЕРА — ПАУЭЛЛА)

H_k - положительно определённая симметрическая матрица. Метод ДФП основан на использовании идей метода Ньютона и методов, использующих сопряжённые направления. В методе ДФП матрицы H_k рекуррентно определяются так, чтобы последовательные приближения

$$\bar{x}_{k+1} = \bar{x}_k - \lambda_k H_k \nabla f(\bar{x}_k)$$

минимизировали квадратичную функцию за конечное число шагов.

В отличие от метода Ньютона в методе ДФП используется только первые производные и не требуется на каждом шаге обращать матрицу H_k . Кроме того, рекуррентные соотношения для матрицы H_k строятся таким образом, чтобы последовательность матриц H_k сходилась к $H^{-1}(\bar{x}^*)$.

Начальный этап. Задать начальную точку $\bar{x}_1$ и симметрическую положительно определённую матрицу $\bar{H}_1$. Положить $\bar{y}_1 = \bar{x}_1$, $k = j = 1$. Задать $\varepsilon > 0$ - параметр окончания счёта.

Основной этап.

Шаг 1. Вычислить $\bar{S}_j = -H_j \nabla f(\bar{y}_j)$.

Шаг 2. Вычислить $\lambda_j = \arg \min f(\bar{y}_j + \lambda \bar{S}_j)$, положить $\bar{y}_{j+1} = \bar{y}_j + \lambda_j \bar{S}_j$.

Шаг 3. Если $j = n$, то перейти к шагу 5.

Шаг 4. Вычислить $\bar{P}_j = \lambda_j \bar{S}_j$, $\bar{q}_j = \nabla f(\bar{y}_{j+1}) - \nabla f(\bar{y}_j)$, $\Delta H_j = \frac{\bar{P}_j' \bar{P}_j}{\bar{P}_j' \bar{q}_j} - \frac{H_j \bar{q}_j' \bar{q}_j H_j}{\bar{q}_j H_j \bar{q}_j'}$;

$H_{j+1} = H_j + \Delta H_j$, положить $j = j + 1$ и перейти к шагу 1.

Шаг 5. Положить $\bar{x}_{k+1} = \bar{y}_{n+1}$. Если $\|\nabla f(\bar{x}_{k+1})\| < \varepsilon$, то положить $\bar{x}^* = \bar{x}_{k+1}$ и остановиться, иначе положить $j = 1$, $\bar{y}_1 = \bar{x}_{k+1}$, $k = k + 1$ и перейти к шагу 1.

В условиях алгоритма ДФП справедлива теорема.

Теорема 1. Направления $\bar{S}_j$ ($j = \overline{1, n}$), генерируемые алгоритмом ДФП, являются направлениями спуска, а матрицы H_j ($j = \overline{1, n}$)- симметрическими и положительно определёнными.

Теперь рассмотрим задачу безусловной оптимизации квадратичной функции:

$$f(\bar{x}) = (\bar{c}, \bar{x}) + \frac{1}{2}(\bar{x}, H\bar{x}) \rightarrow \min, \quad \bar{x} \in E_n, \quad (12.18)$$

где H - симметрическая положительно определённая матрица.

Теорема 2. Пусть задача (12.18) решается методом ДФП. Тогда, если $\nabla f(\bar{y}_j) \neq \bar{0}$ для всех j , то

- 1) направления $\bar{S}_1, \dots, \bar{S}_n$ являются H - сопряжёнными;
- 2) $\bar{y}_{n+1} = \bar{x}^*$ - является решением задачи (2.18);
- 3) $H_{n+1} = H^{-1}$.

Из **теоремы 25.** следует, что задача (12.18) с помощью алгоритма ДФП решается за одну итерацию ($k = 1$).

Пример. Найти минимум функции

$$f(\bar{x}) = 10x_1^2 + 10x_1x_2 + 3x_2^2.$$

Начальный этап. Пусть $\bar{x}_1 = (-0,6; 1)$, $H = 1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, положим $k = j = 1$, $\bar{y}_1 = \bar{x}_1$,

вычислим $\nabla f(\bar{x}) = (20x_1 + 10x_2, 10x_1 + 6x_2)$.

Основной этап.

Итерация 1

($k = 1$).

Шаг 1. Вычислим $\bar{S}_1 = -H_1 \nabla f(\bar{x}_1) = -\nabla f(\bar{x}_1) = (2; 0)$.

Шаг 2. Вычислим $\lambda_1 = \arg \min f(\bar{y}_1 + \lambda \bar{S}_1) = \arg \min f(-0.6 + 2\lambda; 1) = 0.05$, тогда

$$\bar{y}_2 = \bar{y}_1 + \lambda_1 \bar{S}_1 = (-0.5; 1.0).$$

Шаг 3. Так как $j = 1 \neq 2$, то переходим к шагу 4.

Шаг 4. Вычислим $\bar{P}_1 = \lambda_1 \bar{S}_1 = (0.1; 0)$, $\nabla f(\bar{y}_2) = (0; 1)$, $\bar{q}_1 = (0; 1) - (-2; 0) = (2; 1)$.

$$\frac{\bar{P}_1 \bar{P}_1}{\bar{P}_1 \bar{q}_1'} = \frac{\begin{pmatrix} 0.01 & 0 \\ 0 & 0 \end{pmatrix}}{0.2} = \begin{pmatrix} 0.05 & 0 \\ 0 & 0 \end{pmatrix}$$

$$\frac{H_1 \bar{q}_1' \bar{q}_1 H_1}{\bar{q}_1 H_1 \bar{q}_1'} = \frac{\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \end{pmatrix} (2; 1) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}}{(2; 1) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \end{pmatrix}} = \begin{pmatrix} 4/5 & 2/5 \\ 2/5 & 1/5 \end{pmatrix}$$

$$H_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 0.05 & 0 \\ 0 & 0 \end{pmatrix} - \begin{pmatrix} 4/5 & 2/5 \\ 2/5 & 1/5 \end{pmatrix} = \begin{pmatrix} 0.25 & -0.4 \\ -0.4 & 0.8 \end{pmatrix}.$$

Положим $j = j + 1 = 2$ и перейдём к шагу 1.

Шаг 1. Вычислим $\bar{S}_2 = -H_2 \nabla f(\bar{y}_2) = -\begin{pmatrix} 0.25 & -0.4 \\ -0.4 & 0.8 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = (0.4; -0.8)$.

Шаг 2. Вычислим $\lambda_2 = \arg \min f(\bar{y}_2 + \lambda \bar{S}_2) = \arg \min f(2/5 \lambda - 1/2; 1 - 4/5 \lambda) = 5/4$, тогда

$$\bar{y}_3 = \bar{y}_2 + \lambda_2 \bar{S}_2 = (0; 0).$$

Шаг 3. Так как $j = 2 = 2$, то переходим к шагу 5.

Шаг 5. Положим $\bar{x}_2 = \bar{y}_3$. Так как $\nabla f(\bar{x}_2) = (0; 0)$, то $\bar{x}_2 = \bar{x}^*$ - решение задачи.

Покажем, что матрица H_3 , которая была бы получена на шаге 4, равна H^{-1} .

Действительно, так как $H = \begin{pmatrix} 20 & 10 \\ 10 & 6 \end{pmatrix}$, то $H^{-1} = \begin{pmatrix} 0.3 & -0.5 \\ -0.5 & 1 \end{pmatrix}$.

С другой стороны, по формулам шага 4 имеем $\bar{P}_2 = 5/4 (2/5; -4/5) = (1/2; -1)$;

$$\bar{q}_2 = (0; 0) - (0; 1) = (0; -1).$$

$$\frac{\bar{P}_2' \bar{P}_2}{\bar{P}_2 \bar{Q}_2'} = \frac{\begin{pmatrix} 1/2 & 0 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} 1/2 & -1 \\ 0 & 0 \end{pmatrix}}{(1/2; -1)(0; -1)} = \frac{\begin{pmatrix} 1/4 & -1/2 \\ -1/2 & 1 \end{pmatrix}}{1} = \begin{pmatrix} 1/4 & -1/2 \\ -1/2 & 1 \end{pmatrix},$$

$$\frac{H_2 \bar{Q}_2' \bar{Q}_2 H_2}{\bar{Q}_2 H_2 \bar{Q}_2'} = \frac{\begin{pmatrix} 1/4 & -2/5 \\ -2/5 & 4/5 \end{pmatrix} \begin{pmatrix} 0 \\ -1 \end{pmatrix} (0; -1) \begin{pmatrix} 1/4 & -2/5 \\ -2/5 & 4/5 \end{pmatrix}}{(0; -1) \begin{pmatrix} 1/4 & -2/5 \\ -2/5 & 4/5 \end{pmatrix} \begin{pmatrix} 0 \\ -1 \end{pmatrix}} = \begin{pmatrix} 1/5 & -2/5 \\ -2/5 & 4/5 \end{pmatrix},$$

$$H_3 = H_2 + H_2 = \begin{pmatrix} 1/4 & -2/5 \\ -2/5 & 4/5 \end{pmatrix} + \begin{pmatrix} 1/4 & -1/2 \\ -1/2 & 1 \end{pmatrix} - \begin{pmatrix} 1/5 & -2/5 \\ -2/5 & 4/5 \end{pmatrix} =$$

$$= \begin{pmatrix} 3/10 & -1/2 \\ -1/2 & 1 \end{pmatrix} = H^{-1}.$$

МЕТОД МНОЖИТЕЛЕЙ ЛАГРАНЖА ДЛЯ ПОИСКА УСЛОВНОГО ЭКСТРЕМУМА ПРИ НАЛИЧИИ УРАВНЕНИЙ СВЯЗИ

Практическая работа 3. Решение задач на безусловный экстремум

$f(x_1, x_2) = Gx_1^2 + x_1x_2 + Sx_1^2 \rightarrow \min$. Найти (x_1, x_2) .

G-номер группы

S-номер студента

Пример: Продукция трех видов производится в объеме x_1 , x_2 , x_3 и реализуется по цене $10 - x_1$, $20 - x_2$, $60 - 2x_3 - x_2$ соответственно. Определить объемы производства, обеспечивающие наибольший доход.

$$\bar{x} = (x_1, x_2, x_3)$$

$$F(\bar{x}) = (10 - x_1)x_1 + (20 - x_2)x_2 + (60 - 2x_3 - x_2)x_3 = 10x_1 + 20x_2 + 60x_3 - x_1^2 - x_2^2 - 2x_3^2 - x_2x_3 \rightarrow \max$$

Определим стационарные точки функции

$$\overline{\nabla f} = 0:$$

$$\frac{\partial f}{\partial x_1} = 10 - 2x_1 = 0$$

$$\frac{\partial f}{\partial x_2} = 20 - 2x_2 - x_3 = 0$$

$$\frac{\partial f}{\partial x_3} = 60 - 4x_3 - x_2 = 0$$

Решением этой системы линейных алгебраических уравнений является вектор

$$\overline{x^o} = (5, \frac{20}{7}, \frac{100}{7}) = (5, 2.857, 14.2857)$$

Проверим достаточное условие оптимальности. Вычислим матрицу Гессе в полученной стационарной точке.

$$H = \left(\frac{\partial^2 f}{\partial x_i \partial x_j} \right) = \left(\begin{array}{ccc} -2 & 0 & 0 \\ 0 & -2 & -1 \\ 0 & -1 & -4 \end{array} \right) \bigg|_{\overline{x} = \overline{x^o}}$$

Угловые миноры матрицы имеют чередующиеся знаки

$$M_1 = |a_{11}| = |-2| < 0 \quad M_2 = \begin{vmatrix} -2 & 0 \\ 0 & -2 \end{vmatrix} = 4 > 0 \quad M_3 = \begin{vmatrix} -2 & 0 & 0 \\ 0 & -2 & -1 \\ 0 & -1 & -4 \end{vmatrix} = -14 < 0,$$

значит, матрица Гессе отрицательно определена и точка $\overline{x^o} = (5, 2.857, 14.2857)$ – точка локального максимума. При полученных объемах производства будет достигнут наибольший доход $F(\overline{x^o}) = 482.145$.

Проверим отрицательную определенность матрицы вторым способом. Найдем собственные числа матрицы Гессе

$$\|H - \lambda E\| = \left\| \begin{pmatrix} -2-\lambda & 0 & 0 \\ 0 & -2-\lambda & -1 \\ 0 & -1 & -4-\lambda \end{pmatrix} \right\| = (-2-\lambda)[(-2-\lambda)(-4-\lambda)-1] = 0$$

$$-\lambda E = \begin{pmatrix} -\lambda & 0 & 0 \\ 0 & -\lambda & 0 \\ 0 & 0 & -\lambda \end{pmatrix}$$

$$P_3(\lambda) = [(2+\lambda)(4+\lambda)-1](2+\lambda) = 0$$

$$\lambda_1 = -2$$

$$(2+\lambda)(4+\lambda)-1 = 0$$

$$\lambda^2 + 6\lambda + 7 = 0$$

$$\lambda_{2,3} = \frac{-6 \pm \sqrt{36 - 4 \cdot 7}}{2}$$

$$\lambda_2 = -1.59$$

$$\lambda_3 = -4.42$$

Так как все собственные числа матрицы $\lambda_i < 0$, то матрица отрицательно определена и $\overline{x^o}$ – локальный максимум.

```

import numpy as np
import matplotlib.pyplot as plt
from scipy.optimize import minimize
from mpl_toolkits.mplot3d import Axes3D

# Функция Розенброка
def rosenbrock(x):
    return (1 - x[0])**2 + 100 * (x[1] - x[0]**2)**2

# Начальное приближение
x0 = [-1, 2]

# Минимизация с использованием разных методов
# res = minimize(rosenbrock, x0, method='CG')
# res = minimize(rosenbrock, x0, method='Newton-CG', jac=lambda x: 2*x + 10*np.pi*np.sin(2*np.pi*x))
res = minimize(rosenbrock, x0, method='Newton-CG', jac=lambda x:
    [-2*(1 - x[0])**1-2*x[0]**1*100 * 2*(x[1] - x[0]**2)**1, 200 * (x[1] - x[0]**2)**1])
# res = minimize(rosenbrock, x0, method='Nelder-Mead')
# res = minimize(rosenbrock, x0, method='Powell')
# res = minimize(rosenbrock, x0, method='BFGS')

print(res)

# Генерация данных для графика
x = np.linspace(-2, 2, 400)
y = np.linspace(-1, 3, 400)
X, Y = np.meshgrid(x, y)
Z = rosenbrock([X, Y])

# Построение трехмерного графика
fig = plt.figure(figsize=(12, 8))
ax = fig.add_subplot(111, projection='3d')
surf = ax.plot_surface(X, Y, Z, cmap='viridis', edgecolor='none', alpha=0.8)

# Маркировка результата оптимизации
ax.set_xlabel('X координата')
ax.set_ylabel('Y координата')
ax.set_zlabel('Z значение')
ax.set_title('Трёхмерный график функции Розенброка')

# Отображение графика
plt.show()

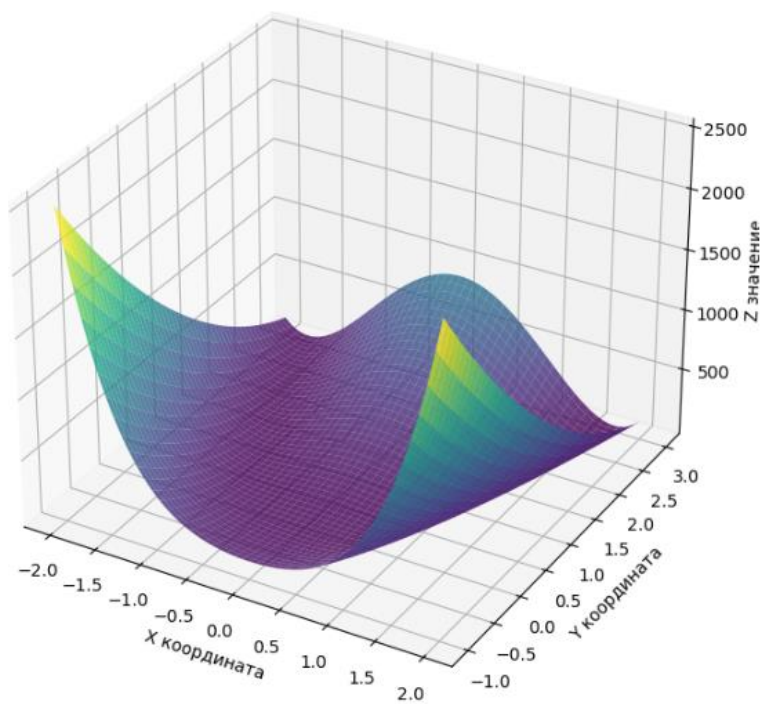
```

```

message: Optimization terminated successfully.
success: True
status: 0
  fun: 4.335397881914305e-11
   x: [ 1.000e+00  1.000e+00]
  nit: 140
  jac: [ 2.176e-03 -1.092e-03]
 nfev: 153
 njev: 437
  nhev: 0

```

Трёхмерный график функции Розенброка



АЛГОРИТМИЧЕСКИЕ ОТОБРАЖЕНИЯ В ЧИСЛЕННЫХ МЕТОДАХ ОПТИМИЗАЦИИ

Применение необходимых и достаточных условий безусловного экстремума, эффективно для решения ограниченного числа примеров, в которых вытекающие из условий соотношения имеют аналитическое решение. Для решения большинства практических задач они не могут быть рекомендованы по следующим причинам.

1. Целевая функция $f(x)$ может не иметь непрерывных производных до второго порядка включительно.
2. Использование необходимого условия первого порядка связано с решением системы n в общем случае нелинейных алгебраических уравнений, что представляет собой самостоятельную задачу, трудоемкость решения которой сравнима с трудоемкостью численного решения, поставленной задачи поиска экстремума.
3. Возможны случаи, когда о целевой функции известно лишь то, что ее значение может быть вычислено с нужной точностью, а сама функция задана неявно.

Подавляющее большинство численных методов оптимизации относится к классу *итерационных*, т. е. порождающих последовательность точек в соответствии с предписанным набором правил, включающим критерий окончания. При заданной начальной точке x^0 методы генерируют последовательность $x^0, x^1, x^2, \dots$. Преобразование точки x^k в x^{k+1} представляет собой *итерацию*.

Для определенности рассмотрим задачу поиска безусловного локального минимума:

Общее правило построения последовательности $\{x^k\}$ имеет вид

$$x^{k+1} = x^k + t_k d^k, \quad k = 0, 1, \dots,$$

где точка x^0 — *начальная точка* поиска; d^k — приемлемое направление перехода из точки x^k в точку x^{k+1} , обеспечивающее выполнение условия выше и называемое *направлением спуска*; t_k — *величина шага*.

Начальная точка поиска x задается, исходя из физического содержания решаемой задачи и наличия априорной информации о положении точек экстремума.

Приемлемое направление спуска d^k должно удовлетворять условию

$$(\nabla f(x^k), d^k) < 0, \quad k = 0, 1, \dots,$$

обеспечивающему убывание функции $f(x)$. Примером приемлемого направления является направление вектора антиградиента $d^k = -\nabla f(x^k)$.

Величина шага $t_k > 0$ выбирается либо из условия либо из условия минимума функции вдоль направления спуска:

$$f(x^k + t_k d^k) \rightarrow \min_{t_k}$$

Выбор шага t_k из этого условия делает спуск в направлении d^k *наискорейшим*.

Определение 4.1. Последовательность $\{x^k\}$ называется *минимизирующей*,

если $\lim_{k \rightarrow \infty} f(x^k) = f^*$ т. е. последовательность сходится к нижней грани $f^* = \inf_{x \in R^n} f(x)$

Определение 4.2. Последовательность $\{x^k\}$ называется *сходящейся к точке минимума x^** , если $\|x^k - x^*\| \rightarrow 0$ при $k \rightarrow \infty$.

Сходимость последовательности $\{x^k\}$ при выборе приемлемого направления d^k и величины шага t_k зависит от характера функции $f(x)$ и от выбора начальной точки x^k .

МЕТОДЫ ПОКООРИНАТНОГО СПУСКА И ХУКА-ДЖИВСА.

Метод покоординатного спуска

Пусть дана функция $f(x)$, ограниченная снизу на множестве R^n и имеющая непрерывные частные производные во всех его точках.

Требуется найти локальный минимум функции $f(x)$ на множестве допустимых решений $X = R^n$, т. е. найти такую точку $x^* \in R^n$ что $f(x^*) \in \min_{x \in R^n} (f(x))$.

Стратегия поиска

Стратегия решения задачи состоит в построении последовательности точек $\{x^k\}$, $k = 0, 1, \dots$, таких, что $f(x^{k+1}) < f(x^k)$, $k = 0, 1, \dots$. Точки последовательности $\{x^k\}$ вычисляются по циклам в соответствии с правилом

$$x^{jk+1} = x^{jk} - t_k \left(\frac{\partial f(x)}{\partial k_{k+1}} \right) \Big|_{x=x^{jk}} \square e_{k+1}$$

где j — номер цикла вычислений; $j = 0, 1, 2, \dots$; k — номер итерации внутри цикла, $k = 0, 1, \dots, n-1$; e_{k+1} , $k = 0, 1, \dots, n-1$ — единичный вектор, $(k+1)$ -я проекция которого равна 1; точка x^{00} задается пользователем, величина шага t_k выбирается из условия

$$f \left(x^{jk} - t_k \left(\frac{\partial f(x)}{\partial k_{k+1}} \right) \square e_{k+1} \right) - f(x^{jk}) < 0 \text{ или}$$

$$f(x^{jk+1}) - f(x^{jk}) < -\varepsilon \left\| \nabla f(x^k) \right\|^2.$$

Если выбранное условие при текущем t_k не выполняется, шаг уменьшается

вдвое и вычисляется заново.

Алгоритм

Шаг 1. Задать x^0 , $\varepsilon > 0$, $\varepsilon_1 > 0$, $\varepsilon_2 > 0$, M — предельное число итераций, предельное число M циклов счета, кратное n , где n — размерность вектора x . Найти градиент $\nabla f(x)$

Шаг 2. Задать номер цикла $j = 0$.

Шаг 3. Проверить условие $j \geq M$:

а) если $j \geq M$, то $x^* = x^{jk}$, расчет окончен;

б) если нет, то перейти к шагу 4.

Шаг 4. Задать $k = 0$.

Шаг 5. Проверить условие $k \leq n-1$:

а) если $k \leq n-1$, то перейти к шагу 6;

б) если $k = n$, то положить $j = j + 1$ и $x^{j+1,k} = x^{jn}$ и перейти к шагу 3.

Шаг 6. Вычислить $\nabla f(x)$

Шаг 7. Проверить выполнение критерия окончания $\|\nabla f(x^k)\| < \varepsilon_1$:

а) если критерий выполнен, $x^* = x^{jk}$, расчет окончен;

б) если нет, то перейти к шагу 8.

Шаг 8. Задать t_k .

Шаг 9. Вычислить точку x^{jk+1} : $x^{jk+1} = x^{jk} - t_k \left(\frac{\partial f(x)}{\partial k_{k+1}} \right) \Big|_{x=x^{jk}} \square e_{k+1}$

Шаг 10. Проверить выполнение условия

$$f\left(x^{jk} - t_k \left(\frac{\partial f(x)}{\partial k_{k+1}} \right) \square e_{k+1}\right) - f(x^{jk}) < 0 \text{ или } f(x^{jk+1}) - f(x^{jk}) < -\varepsilon \|\nabla f(x^k)\|^2 :$$

а) если условие выполнено, то перейти к шагу 11;

б) если нет, то положить $t_k = \frac{t_k}{2}$ и перейти к шагу 9.

Шаг 11. Проверить выполнение условий $\|x^{k+1} - x^k\| < \varepsilon_2$, $|f(x^{jk+1}) - f(x^{jk})| < \varepsilon_2$: а) если в двух последовательных циклах с номерами j и $j-1$ оба условия

выполняются, то расчет в точке x^{jk+1} окончен и $x^* = x^{jk+1}$;

б) если хотя бы одно из условий не выполнено, положить $k = k + 1$ и перейти к шагу 5.

Метод конфигураций (хука — дживса)

Требуется найти безусловный минимум функции $f(x)$ многих переменных, т. е. найти такую точку $x^* \in R^n$ что $f(x^*) \in \min_{x \in R^n} (f(x))$.

Стратегия поиска

Метод конфигураций (метод Хука — Дживса [R. Hooke, T. A. Jeeves]) представляет собой комбинацию *исследующего поиска* с циклическим изменением переменных и ускоряющего *поиска по образцу*. Исследующий поиск ориентирован на выявление локального поведения целевой функции и определение направления ее убывания вдоль «оврагов». Полученная информация используется при поиске по образцу при движении вдоль «оврагов».

Исследующий поиск начинается в некоторой начальной точке x^0 , называемой *старым базисом*. В качестве множества направлений поиска выбирается множество координатных направлений. Задается величина шага, которая может быть различной для разных координатных направлений и переменной в процессе поиска. Фиксируется первое координатное направление и делается шаг в сторону увеличения соответствующей переменной. Если значение функции в пробной точке меньше значения функции в исходной точке, шаг считается удачным. В противном случае необходимо вернуться в предыдущую точку и сделать шаг в противоположном направлении с последующей проверкой поведения функции. После перебора всех координат исследующий поиск завершается. Полученная точка называется *новым базисом*. Если исследующий поиск с данной величиной шага неудачен, то она уменьшается и процедура продолжается. Поиск заканчивается, когда текущая величина шага станет меньше некоторой величины.

Поиск по образцу заключается в движении по направлению от старого базиса к новому (от точки x^0 через точку x^1 , из точки x^1 через точку x^2 , из x^2 через x^3 на рис. 5.12). Величина ускоряющего шага задается ускоряющим множителем λ . Успех поиска по образцу определяется с помощью исследующего поиска из полученной точки. Если при этом значение в наилучшей точке меньше, чем в точке предыдущего базиса, то поиск по образцу удачен.

Обозначим через $d_1, \dots, d_n$ — координатные направления

$$d_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, d_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \dots, d_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$$

При поиске по направлению d_i меняется только переменная x_i , а остальные переменные остаются зафиксированными.

Алгоритм

Шаг 1. Задать начальную точку x^0 , число $\varepsilon > 0$ для остановки алгоритма, начальные величины шагов по координатным направлениям $\Delta_1, \dots, \Delta_n > \varepsilon$, ускоряющий множитель $\lambda > 0$, коэффициент уменьшения шага $\alpha > 1$. Положить $y^1 = x^0$, $i = 1$, $k = 0$.

Шаг 2. Осуществить исследующий поиск по выбранному координатному направлению:

а) если $f(y^i + \Delta_i d_i) < f(y^i)$,

шаг считается удачным. В этом случае следует положить $y^{i+1} = y^i + \Delta_i d_i$ и перейти к шагу 3;

б) если в п. «а» шаг неудачен, то делается шаг в противоположном направлении. Если $f(y^i - \Delta_i d_i) < f(y^i)$, шаг считается удачным. В этом случае следует положить $y^{i+1} = y^i - \Delta_i d_i$ и перейти к шагу 3;

в) если в п. «а» и «б» шаги неудачны, положить $y^{i+1} = y^i$.

Шаг 3. Проверить условия:

а) если $i < n$, то положить $i = i + 1$ и перейти к шагу 2 (продолжить исследующий поиск по оставшимся направлениям);

б) если $i = n$, проверить успешность исследующего поиска:

- если $f(y^{n+1}) < f(x^k)$, перейти к шагу 4;
- если $f(y^{n+1}) \geq f(x^k)$, перейти к шагу 5.

Шаг 4. Провести поиск по образцу. Положить

$$x^{k+1} = y^{n+1}, y^1 = x^{k+1} + \lambda(x^{k+1} - x^k), i = 1, k = k + 1$$

и перейти к шагу 2.

Шаг 5. Проверить условие окончания:

а) если все $\Delta_i > \varepsilon$, то поиск закончить: $x^* \cong x^k$;

б) для тех i , для которых $\Delta_i > \varepsilon$, уменьшить величину шага:

$$\Delta_i = \frac{\Delta_i}{\alpha}$$

$y^1 = x^k, x^{k+1} = x^k, i = 1, k = k + 1$ и перейти к шагу 2.

ОБЩАЯ СХЕМА ГРАДИЕНТНОГО СПУСКА

Пусть дана функция $f(x)$, ограниченная снизу на множестве R^n и имеющая непрерывные частные производные во всех его точках.

Требуется найти локальный минимум функции $f(x)$ на множестве допустимых решений $X = R^n$, т. е. найти такую точку $x^* \in R^n$ что $f(x^*) \in \min_{x \in R^n} (f(x))$

Стратегия поиска

Стратегия решения задачи состоит в построении последовательности точек $\{x^k\}$, $k = 0, 1, \dots$, таких, что $f(x^{k+1}) < f(x^k)$, $k = 0, 1, \dots$. Точки последовательности $\{x^k\}$ вычисляются по правилу

$$x^{k+1} = x^k - t_k \nabla f(x^k), \quad k = 0, 1, \dots,$$

где точка x^0 задается пользователем; $\nabla f(x^k)$ — градиент функции $f(x)$, вычисленный в точке x^k ; величина шага t_k задается пользователем и остается постоянной до тех пор, пока функция убывает в точках последовательности, что контролируется путем проверки выполнения условия $f(x^{k+1}) - f(x^k) < 0$ или $f(x^{k+1}) - f(x^k) < -\varepsilon \|\nabla f(x_k)\|^2$, $0 < \varepsilon < 1$.

Построение последовательности $\{x^k\}$ заканчивается в точке x^k , для которой $\|\nabla f(x^k)\| < \varepsilon_1$, где ε_1 — заданное малое положительное число, или $k \geq M$, где M — предельное число итераций, или при двукратном одновременном выполнении двух неравенств $\|x^{k+1} - x^k\| < \varepsilon_2$, $|f(x^{k+1}) - f(x^k)| < \varepsilon_2$, где ε_2 — малое положительное число. Вопрос о том, может ли точка x^k рассматриваться как найденное приближение искомой точки минимума, решается путем проведения дополнительного исследования, которое описано ниже.

Алгоритм

Шаг 1. Задать x^0 , $0 < \varepsilon < 1$, $\varepsilon_1 > 0$, $\varepsilon_2 > 0$, M — предельное число итераций.

Найти градиент функции в произвольной точке $\nabla f(x) = \left(\frac{\partial f(x)}{\partial x_1}, \dots, \frac{\partial f(x)}{\partial x_n} \right)^T$

Шаг 2. Положить $k = 0$.

Шаг 3. Вычислить $\nabla f(x^k)$.

Шаг 4. Проверить выполнение критерия окончания $\|\nabla f(x^k)\| < \varepsilon_1$:

а) если критерий выполнен, расчет закончен, $x^* = x^k$;

б) если критерий не выполнен, то перейти к шагу 5.

Шаг 5. Проверить выполнение неравенства $k \geq M$:

а) если неравенство выполнено, то расчет окончен: $x^* = x^k$;

б) если нет, то перейти к шагу 6.

Шаг 6. Задать величину шага t_k .

Шаг 7. Вычислить $x^{k+1} = x^k - t_k \nabla f(x^k)$.

Шаг 8. Проверить выполнение условия $f(x^{k+1}) - f(x^k) < 0$ или

$$f(x^{k+1}) - f(x^k) < -\varepsilon \|\nabla f(x_k)\|^2 :$$

а) если условие выполнено, то перейти к шагу 9;

б) если условие не выполнено, положить $t_k = \frac{t_k}{2}$ и перейти к шагу 7.

Шаг 9. Проверить выполнение условий $\|x^{k+1} - x^k\| < \varepsilon_2$, $|f(x^{k+1}) - f(x^k)| < \varepsilon_2$:

а) если оба условия выполнены при текущем значении k и $k = k - 1$, то расчет окончен, $x^* = x^{k+1}$;

б) если хотя бы одно из условий не выполнено, положить $k = k + 1$ и перейти к шагу 3.

МЕТОД НАИСКОРЕЙШЕГО СПУСКА

Пусть дана функция $f(x)$, ограниченная снизу на множестве R^n и имеющая непрерывные частные производные во всех его точках.

Требуется найти локальный минимум функции $f(x)$ на множестве допустимых решений $X = R^n$, т. е. найти такую точку $x^* \in R^n$ что $f(x^*) \in \min_{x \in R^n} (f(x))$

Стратегия поиска

Стратегия решения задачи состоит в построении последовательности точек $\{x^k\}$, $k = 0, 1, \dots$, таких, что $f(x^{k+1}) < f(x^k)$, $k = 0, 1, \dots$. Точки последовательности $\{x^k\}$ вычисляются по правилу

$$x^{k+1} = x^k - t_k \nabla f(x^k), \quad k = 0, 1, \dots,$$

где точка x^0 задается пользователем; величина шага t_k определяется для каждого значения k из условия

$$\varphi(t_k) = f(x^k - t_k \nabla f(x^k)) \rightarrow \min_{t_k}$$

Если возможно, то применяют точный метод нахождения минимума функции одного переменного или это может быть численный метод.

Построение последовательности $\{x^k\}$ заканчивается в точке x^k , для которой $\|\nabla f(x^k)\| < \varepsilon_1$, где ε_1 — заданное малое положительное число, или $k \geq M$, где M — предельное число итераций, или при двукратном одновременном выполнении двух неравенств $\|x^{k+1} - x^k\| < \varepsilon_2$, $|f(x^{k+1}) - f(x^k)| < \varepsilon_2$, где ε_2 — малое положительное число.

Алгоритм

Шаг 1. Задать x^0 , $\varepsilon_1 > 0$, $\varepsilon_2 > 0$, M — предельное число итераций.

Найти градиент функции в произвольной точке $\nabla f(x) = \left(\frac{\partial f(x)}{\partial x_1}, \dots, \frac{\partial f(x)}{\partial x_n} \right)^T$

Шаг 2. Положить $k = 0$.

Шаг 3. Вычислить $\nabla f(x^k)$.

Шаг 4. Проверить выполнение критерия окончания $\|\nabla f(x^k)\| < \varepsilon_1$:

а) если критерий выполнен, расчет закончен, $x^* = x^k$;

б) если критерий не выполнен, то перейти к шагу 5.

Шаг 5. Проверить выполнение неравенства $k \geq M$:

а) если неравенство выполнено, то расчет окончен: $x^* = x^k$;

б) если нет, то перейти к шагу 6.

Шаг 6. Вычислить величину шага t_k^* из условия $\varphi(t_k) = f(x^k - t_k \nabla f(x^k)) \rightarrow \min_{t_k}$.

Шаг 7. Вычислить $x^{k+1} = x^k - t_k^* \nabla f(x^k)$.

Шаг 8. Проверить выполнение условий $\|x^{k+1} - x^k\| < \varepsilon_2$, $|f(x^{k+1}) - f(x^k)| < \varepsilon_2$:

а) если оба условия выполнены при текущем значении k и $k = k - 1$, то расчет окончен, $x^* = x^{k+1}$;

б) если хотя бы одно из условий не выполнено, положить $k = k + 1$ и перейти к шагу 3.

МЕТОД СОПРЯЖЕННЫХ НАПРАВЛЕНИЙ (Метод Пауэлла).

Требуется найти безусловный минимум функции $f(x)$ многих переменных, т. е. найти такую точку $x^* \in R^n$ что $f(x^*) \in \min_{x \in R^n} (f(x))$.

Определение. Пусть H — симметрическая матрица размеров $(n \times n)$. Векторы $d_1, d_2, \dots, d_n$ называются H -сопряженными, или просто *сопряженными*, если $d^T H d_j = 0$ при всех $i \neq j$.

Стратегия поиска

В методе сопряженных направлений (методе Пауэлла [M.J.D. Powell]) используется факт, что минимум квадратичной функции может быть найден не более чем за n шагов при условии, что поиск ведется вдоль сопряженных относительно матрицы Гессе направлений. Так как достаточно большой класс целевых функций может быть представлен в окрестности точки минимума своей квадратичной аппроксимацией, описанная идея применяется и для неквадратичных функций. Задается начальная точка и направления $d_1, d_2, \dots, d_n$, совпадающие с координатными. Находится минимум $f(x)$ при последовательном движении по $(n + 1)$ направлениям с помощью одного из методов одномерной минимизации. При этом полученная ранее точка минимума берется в качестве исходной для поиска по следующему направлению, а направление d_n используется как при первом ($d_0 = d_n$), так и последнем поиске. Находится новое направление поиска, сопряженное с d_n . Оно проходит через точки, полученные при первом и последнем поиске. Заменяется d_1 на d_2 , d_2 на d_3 и т. д. Направление d_n заменяется сопряженным направлением, после чего повторяется поиск по $(n + 1)$ направлениям, уже не содержащим старого направления d_1 .

Для квадратичных функций последовательность n^2 одномерных поисков приводит к точке минимума (если все операции выполнены точно).

Алгоритм

Шаг 1. Задать начальную точку x^0 , число $\varepsilon > 0$ для окончания алгоритма, начальные направления поиска

$$d_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, d_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \dots, d_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$$

Положим $d_0 = d_n, i = 0, y^0 = x^0, k = 0$.

Шаг 2. Найти $y^{i+1} = y^i + t_i d_i$, где параметр t_i (неограниченный по знаку) находится в результате поиска минимума функции $f(y^i + t_i d_i)$ по t_i одним из методов одномерной минимизации.

Шаг 3. Проверить выполнение условий:

а) если $i < n - 1$, положить $i = i + 1$ и перейти к шагу 2;

б) если $i = n - 1$, проверить успешность поиска по n первым направлениям. Если $y^n = y^0$, то поиск завершить: $x^* \cong y^n$, иначе положить $i = i + 1$ и перейти к шагу 2;

в) если $i = n$, проверить успешность поиска по n последним направлениям. Если $y^{n+1} = y^1$, поиск завершить: $x^* \cong y^{n+1}$, иначе перейти к шагу 4 для построения сопряженного направления.

Шаг 4. Положить $x^{k+1} = y^{n+1}$ и проверить условие окончания: а) если $\|x^{k+1} - x^k\| < \varepsilon$, то поиск завершить: $x^* \cong x^{k+1}$;

б) иначе положить: (новое направление); $\bar{d}_i = \bar{d}_{i+1}$, $i = 1, \dots, n - 1$ (исключается старое направление).

Если при этом $rg(\bar{d}_0, \dots, \bar{d}_n) = n$, то новая система направлений линейно

независима. В этом случае положить: $\bar{d}_i = \bar{d}_{i+1}$, $i = 0, 1, \dots, n$; $k = k + 1$, $i = 0$,

$y^0 = x^{k+1}$ и перейти к шагу 2.

Если $rg(\bar{d}_0, \dots, \bar{d}_n) < n$, то новая система направлений линейно зависима.

Тогда следует продолжать поиск в старых направлениях. Для этого положить:

$d_i = \bar{d}_i$, $i = 0, 1, \dots, n$; $y^0 = x^{k+1}$, $k = k + 1$, $i = 0$ и перейти к шагу 2.

МЕТОД НЬЮТОНА

Пусть дана функция $f(x)$, ограниченная снизу на множестве R^n и имеющая непрерывные частные производные во всех его точках.

Требуется найти локальный минимум функции $f(x)$ на множестве допустимых решений $X = R^n$, т. е. найти такую точку $x^* \in R^n$ что $f(x^*) \in \min_{x \in R^n} (f(x))$, $f(x) \in C^2$.

Стратегия поиска

Стратегия метода Ньютона [I. Newton] состоит в построении последовательности точек $\{x^k\}$, $k = 0, 1, \dots$, таких, что $f(x^{k+1}) < f(x^k)$, $k = 0, 1, \dots$. Точки последовательности вычисляются по правилу

$$x^{k+1} = x^k + d^k, k = 0, 1, \dots,$$

где x^0 — задается пользователем, а направление спуска d^k определяется для каждого значения k по формуле

$$d^k = -H^{-1}(x^k) \nabla f(x^k).$$

Выбор d^k гарантирует выполнение требования $f(x^{k+1}) < f(x^k)$ при условии, что $H(x^k) > 0$. Формула (7.2) получена из следующих соображений.

1. Функция $f(x)$ аппроксимируется в каждой точке последовательности $\{x^k\}$ квадратичной функцией

$$F_k = f(x^k) + (\nabla f(x^k), dk) + \frac{1}{2} (dk, H(x^k) dk)$$

2. Направление d^k определяется из необходимого условия экстремума первого порядка:

$$\frac{dF_k}{dd^k} = 0$$

Таким образом, при выполнении требования $H(x^k) > 0$ последовательность является последовательностью точек минимумов квадратичных функций $F_k, k = 0, 1, \dots$ Чтобы обеспечить выполнение требования $f(x^{k+1}) < f(x^k), k = 0, 1, \dots$, даже в тех случаях, когда для каких-либо значений матрица Гессе $H(x^k)$ не окажется положительно определенной, рекомендуется для соответствующих значений k вычислить точку x^{k+1} по методу градиентного спуска $x^{k+1} = x^k - t_k \nabla f(x^k)$ с выбором величины шага t_k из условия $f(x^k - t_k \nabla f(x^k)) < f(x^k)$. Построение последовательности $\{x^k\}$ заканчивается в точке x^k для которой $\|\nabla f(x^k)\| < \varepsilon_1$, где ε_1 — заданное малое положительное число, или при $k \geq M$ (M - предельное число итераций), или при двукратном одновременном выполнении двух неравенств $\|x^{k+1} - x^k\| < \varepsilon_2, |f(x^{k+1}) - f(x^k)| < \varepsilon_2$, где ε_2 — малое положительное число. Вопрос о том, может ли точка x^k рассматриваться как найденное приближение искомой точки минимума, решается путем проведения дополнительного исследования, которое описано ниже.

Алгоритм

Шаг 1. Задать $x^0, \varepsilon_1 > 0, \varepsilon_2 > 0, M$ — предельное число итераций. Найти градиент $\nabla f(x)$ и матрицу Гессе $H(x)$.

Шаг 2. Положить $k = 0$.

Шаг 3. Вычислить $\nabla f(x^k)$

Шаг 4. Проверить выполнение критерия окончания $\|\nabla f(x)\| < \varepsilon_1$: а) если

неравенство выполнено, то расчет окончен и $x^* = x^k$;

б) в противном случае перейти к шагу 5.

Шаг 5. Проверить выполнение неравенства $k \geq M$:

а) если неравенство выполнено, расчет окончен и $x^* = x^k$; б) если нет, перейти к шагу 6.

Шаг 6. Вычислить матрицу $H(x^k)$.

Шаг 7. Вычислить матрицу $H^{-1}(x^k)$.

Шаг 8. Проверить выполнение условия $H^{-1}(x^k) > 0$:

а) если $H^{-1}(x^k) > 0$, то перейти к шагу 9;

б) если нет, то перейти к шагу 10, положив $d^k = -\nabla f(x^k)$.

Шаг 9. Определить $d_k = -H^{-1}(x^k)\nabla f(x^k)$.

Шаг 10. Найти точку $x^{k+1} = x^k + t_k d^k$, положив $t_k = 1$, если $d_k = -H^{-1}(x^k)\nabla f(x^k)$, или выбрав t_k из условия $f(x^{k+1}) < f(x^k)$, если $d^k = -\nabla f(x^k)$.

Шаг 11. Проверить выполнение условий $\|x^{k+1} - x^k\| < \varepsilon_2$, $|f(x^{k+1}) - f(x^k)| < \varepsilon_2$:

а) если оба условия выполнены при текущем значении k и $k = k - 1$, то расчет окончен, $x^* = x^{k+1}$;

б) в противном случае положить $k = k + 1$ и перейти к шагу 3.

МЕТОД ПРОЕКЦИИ ГРАДИЕНТА

Пример. $f(\bar{x}) = x_1 + x_2 \rightarrow \min_P$

$$P = \left\{ \bar{x} \in E_2 : \begin{array}{l} g_1(\bar{x}) = x_1^2 - x_2 \leq 0 \\ g_2(\bar{x}) = x_2 - 1 \leq 0 \end{array} \right\}$$

Решим эту задачу методом возможных направлений, сформулировав её в каноническом виде:

$$-f(\bar{x}) = -x_1 - x_2 \rightarrow \max$$

$$g_1(\bar{x}) = x_1^2 - x_2 \leq 0$$

$$g_2(\bar{x}) = x_2 - 1 \leq 0$$

Начальный этап. Выберем в качестве начальной точку

$$\bar{x}_0 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

Основной этап.

Итерация 1.

Вычисляем вектор-градиент в точке $\bar{x}_0$: $\nabla f(\bar{x}) = \begin{pmatrix} -1 \\ -1 \end{pmatrix}$, $\nabla f(\bar{x}_0) = \begin{pmatrix} -1 \\ -1 \end{pmatrix}$.

Вычисляем $g_1(\bar{x}_0) = 0$ (точка на границе), $g_2(\bar{x}_0) = 0$ (точка на границе).

Тогда $\nabla g_1(\bar{x}_0) = (-2, -1)^T$; $\nabla g_2(\bar{x}_0) = (0, 1)^T$.

Сформируем множество $I(\bar{x}_0)$:

$$I(\bar{x}_0) = \{1, 2\}.$$

$I(\bar{x}_0) \neq \emptyset$, следовательно, для нахождения возможного направления подъёма $\bar{S}_0$ функции $f(\bar{x})$ в точке $\bar{x}_0$, лежащей на границах $x_1^2 - x_2 = 0$ и $x_2 - 1 = 0$ области P , сформулируем ЗЛП:

$$\sigma \rightarrow \max$$

$$-2S_1 - S_2 + \sigma \leq 0$$

$$S_2 + \sigma \leq 0$$

$$S_1 + S_2 + \sigma \leq 0$$

$$-1 \leq S_{1,2} \leq 1$$

или в каноническом виде:

$$\sigma \rightarrow \max$$

$$-2\tilde{S}_1 + 2\tilde{\tilde{S}}_1 - \tilde{S}_2 + \tilde{\tilde{S}}_2 + \sigma + y_1 = 0$$

$$\tilde{S}_2 - \tilde{\tilde{S}}_2 + \sigma + y_2 = 0$$

$$\tilde{S}_1 - \tilde{\tilde{S}}_1 + \tilde{S}_2 - \tilde{\tilde{S}}_2 + \sigma + y_3 = 0$$

$$\tilde{S}_1 - \tilde{\tilde{S}}_1 + y_4 = 1$$

$$-\tilde{S}_1 + \tilde{\tilde{S}}_1 + y_5 = 1$$

$$\tilde{S}_2 - \tilde{\tilde{S}}_2 + y_6 = 1$$

$$-\tilde{S}_2 + \tilde{\tilde{S}}_2 + y_7 = 1$$

$$\tilde{S}_1, \tilde{\tilde{S}}_1, \tilde{S}_2, \tilde{\tilde{S}}_2, \sigma, y_1, y_2, y_3, y_4, y_5, y_6, y_7 \geq 0$$

Решаем эту ЗЛП симплексным методом:

				0	0	0	0	1	0	0	0	0	0	0	0
$\bar{z}$	$\bar{N}$	$\bar{C}_{\bar{N}}$	$\bar{X}_{\bar{N}}$	$\tilde{S}_1$	$\tilde{\tilde{S}}_1$	$\tilde{S}_2$	$\tilde{\tilde{S}}_2$	$\bar{\sigma}$	$\bar{y}_1$	$\bar{y}_2$	$\bar{y}_3$	$\bar{y}_4$	$\bar{y}_5$	$\bar{y}_6$	$\bar{y}_7$
0	6	0	0	-2	2	-1	1	1	1	0	0	0	0	0	0
	7	0	0	0	0	1	-1	1	0	1	0	0	0	0	0
	8	0	1	-1	1	-1	1	0	0	1	0	0	0	0	0
	9	0	1	1	-1	0	0	0	0	0	0	1	0	0	0
	10	0	1	-1	1	0	0	0	0	0	0	0	1	0	0
	11	0	1	0	0	1	-1	0	0	0	0	0	0	1	0
	12	0	1	0	-1	1	0	0	0	0	0	0	0	0	1
	Δ_j		0	0	0	0	0	-1							
	5	1	0	-2	2	-1	1	1	1	0	0	0	0	0	0
	7	0	0	2	-2	2	-2	0	-1	1	0	0	0	0	0

1	8	0	0	3	-3	2	-2	0	-1	0	1	0	0	0	0
	9	0	1	1	-1	0	0	0	0	0	0	1	0	0	0
	10	0	1	-1	1	0	0	0	0	0	0	0	1	0	0
	11	0	1	0	0	1	-1	0	0	0	0	0	0	1	0
	12	0	1	0	0	-1	1	0	0	0	0	0	0	0	1
	Δ_j		0	-2	2	-1	1	-	1						
2	5	1	0	0	0	1/3	-1/3	1	1/3	0	2/3	0	0	0	0
	7	0	0	0	0	2/3	-2/3	0	-1/3	1	-2/3	0	0	0	0
	1	0	0	1	-1	2/3	-2/3	0	-1/3	0	1/3	1	0	0	0
	9	0	1	0	0	2/3	2/3	0	1/3	0	-1/3	1	0	0	0
	10	0	1	0	0	2/3	-2/3	0	-1/3	0	-1/3	0	1	0	0
	11	0	1	0	0	1	-1	0	0	0	0	0	0	1	0
	12	0	1	0	0	-1	1	0	0	0	0	0	0	0	1
	Δ_j		0	-	0	1/3	-1/3	-	1/3	-	2/3	-	-	-	-
3	5	1	1/3	0	0	0	0	1	1/3	0	2/3	0	0	0	1/3
	7	0	2/3	0	0	0	0	0	-1/3	1	-2/3	0	0	0	2/3
	1	0	2/3	1	-1	0	0	0	-1/3	0	1/3	0	0	0	2/3
	9	0	1/3	0	0	0	0	0	1/3	0	-1/3	1	0	0	- 2/3
	10	0	5/3	0	0	0	0	0	-1/3	0	1/3	0	1	0	2/3
	11	0	2	0	0	0	0	0	0	0	0	0	0	1	1
	4	0	1	0	0	-1	1	0	0	0	0	0	0	0	1
	Δ_j		5/3	-	0	0	-	-	1/3	-	-	-	-	-	2/3

Итак, $\sigma_0(\bar{x}_0) = \frac{1}{3}$, $\bar{S}_0 = \begin{pmatrix} \tilde{S}_1 & -\tilde{S}_1 \\ \tilde{S}_2 & -\tilde{S}_2 \end{pmatrix} = \begin{pmatrix} 2/3 \\ -1 \end{pmatrix}$, $\sigma_0(\bar{x}_0) > 0$, следовательно,

$$\bar{x}_1 = \bar{x}_0 + \beta_0 \bar{S}_0 = \begin{pmatrix} -1 + 2/3 \beta_0 \\ 1 - \beta_0 \end{pmatrix}.$$

Вычисляем значение β_0 :

$$g_1(\bar{x}_1) = \left(2/3 \beta_0 - 1\right)^2 - (1 - \beta_0) = 0$$

$$4/9 \beta_0^2 - 4/3 \beta_0 + 1 - 1 + \beta_0 = 0$$

$$4/9 \beta_0^2 - 1/3 \beta_0 = 0; \beta_0 = 0; \beta_0 = \frac{3}{4}.$$

Решим теперь одномерную задачу

$$f(\bar{x}_1) = 1 - \frac{2}{3} \beta_0 - 1 + \beta_0 \rightarrow \max$$

$$0 \leq \beta_0 \leq \frac{3}{4} \text{ или}$$

$$\frac{1}{3} \beta_0 \rightarrow \max$$

$$0 \leq \beta_0 \leq \frac{3}{4}.$$

Получаем $\beta_0 = \frac{3}{4}$. Отсюда $\bar{x}_1 = \begin{pmatrix} -\frac{1}{2}; \frac{1}{4} \end{pmatrix}$.

Итерация 2.

Вычисляем $g_1(\bar{x}_1)$, $g_2(\bar{x}_1)$:

$$g_1(\bar{x}_1) = 0, \quad g_2(\bar{x}_1) = -\frac{3}{4} < 0.$$

Тогда $\nabla g_1(\bar{x}_1) = (-1; -1)^T$.

Сформируем множество $I(\bar{x}_1)$

$$I(\bar{x}_1) = \{1\}.$$

$I(\bar{x}_1) \neq \emptyset$, следовательно, для нахождения возможного направления подъёма $\bar{S}_1$ функции $f(\bar{x})$ в точке $\bar{x}_1$, лежащей на границе $x_1^2 - x_2 = 0$ области P , сформулируем вновь ЗЛП:

$$\sigma \rightarrow \max$$

$$-S_1 - S_2 + \sigma \leq 0$$

$$S_1 + S_2 + \sigma \leq 0$$

$$-1 \leq S_{1,2} \leq 1$$

или в каноническом виде:

$$\sigma \rightarrow \max$$

$$-\tilde{S}_1 + \tilde{\tilde{S}}_1 - \tilde{S}_2 + \tilde{\tilde{S}}_2 + \sigma + y_1 = 0$$

$$\tilde{S}_1 - \tilde{\tilde{S}}_1 + \tilde{S}_2 - \tilde{\tilde{S}}_2 + \sigma + y_2 = 0$$

$$\tilde{S}_1 - \tilde{\tilde{S}}_1 + y_3 = 1$$

$$-\tilde{S}_1 + \tilde{\tilde{S}}_1 + y_4 = 1$$

$$\tilde{S}_2 - \tilde{\tilde{S}}_2 + y_5 = 1$$

$$-\tilde{S}_2 + \tilde{\tilde{S}}_2 + y_6 = 1$$

$$\tilde{S}_1, \tilde{\tilde{S}}_1, \tilde{S}_2, \tilde{\tilde{S}}_2, \sigma, y_1, y_2, y_3, y_4, y_5, y_6 \geq 0.$$

Решаем задачу симплексным методом:

$\bar{z}$	$\bar{N}$	$\bar{C}_{\bar{N}}$	$\bar{X}_{\bar{N}}$	$\tilde{S}_1$	$\tilde{\tilde{S}}_1$	$\tilde{S}_2$	$\tilde{\tilde{S}}_2$	$\bar{\sigma}$	$\bar{y}_1$	$\bar{y}_2$	$\bar{y}_3$	$\bar{y}_4$	$\bar{y}_5$	$\bar{y}_6$
0	6	0	0	-1	1	-1	1	1	1	0	0	0	0	0
	7	0	0	1	-1	1	-1	1	0	1	0	0	0	0
	8	0	1	1	-1	0	0	0	0	0	1	0	0	0

	9	0	1	-1	1	0	0	0	0	0	0	1	0	0
	10	0	1	0	0	1	-1	0	0	0	0	0	1	0
	11	0	1	0	0	-1	1	0	0	0	0	0	0	1
	Δ_j		0	0	0	0	0	-1	-	-	-	-	-	-
1	5	1	0	-1	1	-1	1	1	1	0	0	0	0	0
	7	0	0	2	-2	2	-2	0	-1	1	0	0	0	0
	8	0	1	1	-1	0	0	0	0	0	1	0	0	0
	9	0	1	-1	1	0	0	0	0	0	0	1	0	0
	10	0	1	0	0	1	-1	0	0	0	0	0	1	0
	11	0	1	0	0	0	-1	1	0	0	0	0	0	1
	Δ_j		0	-1	1	-1	1	-	1	-	-	-	-	-
2	5	1	0	0	0	0	0	1	1/2	1/2	0	0	0	0
	1	0	0	1	-1	1	-1	0	-1/2	1/2	0	0	0	0
	3	0	1	0	0	-1	1	0	1/2	-1/2	1	0	0	0
	9	0	1	0	0	1	-1	0	-1/2	1/2	0	1	0	0
	10	0	1	0	0	1	-1	0	0	0	0	0	1	0
	11	0	1	0	0	-1	1	0	0	0	0	0	0	1
	Δ_j		0	-	0	0	0	-	1/2	1/2	-	-	-	-

Итак, $\sigma_1(\bar{x}_1) = 0$, следовательно, $\bar{x}^* = \bar{x}_1 = \left(-\frac{1}{2}; \frac{1}{4}\right)$.

Ответ: $\bar{x}^* = \left(-\frac{1}{2}; \frac{1}{4}\right)$, $f(\bar{x}^*) = -\frac{1}{4}$.

МЕТОДЫ ШТРАФНЫХ ФУНКЦИЙ

$$f(x_1, x_2) = Gx_1^2 + x_1x_2 + Sx_2^2 \rightarrow \min.$$

При ограничении $g(x_1, x_2) = Gx_1 + Sx_2 - 1 = 0$ Найти (x_1, x_2) .

G-номер группы

S-номер студента

Пример1.

Минимизировать $f(\mathbf{x}) = (x_1 - 4)^2 + (x_2 - 4)^2$

при ограничении $h(\mathbf{x}) = x_1 + x_2 - 5 = 0$.

Используя квадратичный штраф, введем штрафную функцию

$P(\mathbf{x}, R) = (x_1 - 4)^2 + (x_2 - 4)^2 + R(x_1 + x_2 - 5)^2$. Определим минимум полученной функции классическим методом. Запишем необходимые условия минимума функции $P(\mathbf{x}, R)$:

$$\left. \begin{aligned} \frac{\partial P}{\partial x_1} &= 2(x_1 - 4) + 2R(x_1 + x_2 - 5) = 0 \\ \frac{\partial P}{\partial x_2} &= 2(x_2 - 4) + 2R(x_1 + x_2 - 5) = 0 \end{aligned} \right\}$$

Находим координаты оптимальной точки в зависимости от R :

$x_1 = x_2 = \left(10 + \frac{8}{R}\right) / \left(4 + \frac{2}{R}\right)$. Переходя в этом выражении к пределу, получаем:

$\lim_{R \rightarrow \infty} \left[\left(10 + \frac{8}{R}\right) / \left(4 + \frac{2}{R}\right) \right] = \frac{10}{4} = 2.5$. Таким образом, метод сходится к точке $\mathbf{x}^* = (2.5; 2.5)$

, в которой $h(\mathbf{x}^*) = 2.5 + 2.5 - 5 = 0$; $f(\mathbf{x}^*) = 4.5$.

Численная реализация метода штрафных функций состоит в последовательном применении метода безусловной минимизации к функции $P(\mathbf{x}, R)$, при значениях R изменяющихся от 0 к ∞ .

Посчитаем оптимальные точки функции $P(\mathbf{x}, R)$ при различных значениях R .

a) $R = 0$

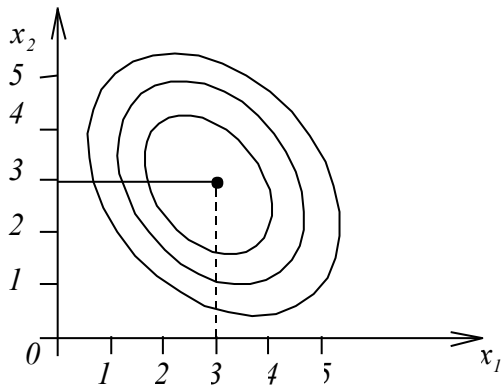
Имеем задачу минимизации функции $P(\mathbf{x}, 0) = (x_1 - 4)^2 + (x_2 - 4)^2 = f(\mathbf{x})$.

Минимум достигается в точке $\mathbf{x}^V = (4, 4)$, $f(\mathbf{x}^V) = 0$. Заметим, что в вычисленной точке значение функции меньше оптимального из-за нарушения условия $h(\mathbf{x}^V) = 0$. Действительно, $h(\mathbf{x}^V) = 4 + 4 - 5 = 3$.

b) $R = 1$.

Функция $P(\mathbf{x}, 1) = (x_1 - 4)^2 + (x_2 - 4)^2 + (x_1 + x_2 - 5)$ имеет минимум в точке $\mathbf{x} = (3, 3)$. Значение функции $f(3, 3) = 3$.

Изобразим линии уровня функции $P(\mathbf{x}, 1)$.



Точка оптимума приближается к реальной, а отклонение функции $h(x)$ от нуля заметно уменьшается: $h(x) = (3 + 3 - 5) = 1$. По существу, коэффициент R играет роль весового коэффициента, определяющего относительную значимость ограничения и целевой функции.

c) $R = 10$.

$P(x, 10) = (x_1 - 4)^2 + (x_2 - 4)^2 + 10(x_1 + x_2 - 5)$. Точка оптимума $x = (2.57, 2.57)$, $f(x) = 4.28$.

Таким образом, при увеличении R точка минимума смещается все больше в сторону действительно оптимальной точки.

d) $R = 100$

$P(x, 100) =$

$x = (2.51, 2.51)$, $f(x) = 4.47$

Таким образом, при $R = 100$ в случае точного решения задачи мы получаем оптимальную точку с двумя верными знаками.

Возникает вопрос: зачем тогда начинать поиск оптимальной точки со значения $R = 0$? Не лучше ли начать прямо с $R = 100$? Такое намерение наталкивается, однако, на ощутимые трудности, связанные с тем, что форма линий уровня функции $P(x, R)$ зависит от R . С ростом R линии уровня функции $P(x, R)$ становятся все более вытянутыми вдоль прямой $h(x) = 0$, и поэтому сложность решения задачи безусловной минимизации резко возрастает. Поэтому предпочтительнее становится стратегия многошагового последовательного приближения к решению задачи. Проблема выбора рациональной последовательности значений штрафных коэффициентов R является основной проблемой, влияющей на эффективность вычислительного процесса.

Пример 2.

Заменим в задаче 1 ограничение-равенство на ограничение-неравенство

$$f(x) = (x_1 - 4)^2 + (x_2 - 4)^2$$

$$g(x) = 5 - x_1 - x_2 \geq 0$$

.

При этом решение x^* задачи остается прежним. Запишем штрафную функцию, используя штраф типа квадрата срезки. Имеем $P(x, R) = f(x) + R\langle g(x) \rangle^2$ или в данном случае $P(x, R) = (x_1 - 4)^2 + (x_2 - 4)^2 + 2R\langle 5 - x_1 - x_2 \rangle^2$.

Необходимые условия существования минимума имеют вид:

$$\frac{\partial P}{\partial x_1} = 2(x_1 - 4) - 2R\langle 5 - x_1 - x_2 \rangle = 0$$

$$\frac{\partial P}{\partial x_2} = 2(x_2 - 4) - 2R\langle 5 - x_1 - x_2 \rangle = 0$$

Откуда, в силу симметричности выражений следует, что $x_1 = x_2$.

С учетом этого получаем $(x_1 - 4) - R\langle 5 - 2x_1 \rangle = 0$. Проанализируем результат, полагая аргумент оператора срезки равным нулю, большим нуля и меньшим нуля.

а) $5 - 2x_1 = 0 \rightarrow \left(x_1 = \frac{2}{5} \wedge \langle 5 - 2x_1 \rangle = 0 \right) \rightarrow \left(x_1 = \frac{2}{5} \wedge x_1 - 4 = 0 \right)$ - противоречие.

б) $5 - 2x_1 > 0 \rightarrow \langle 5 - 2x_1 \rangle = 0 \rightarrow (x_1 - 4 = 0) \rightarrow (x_1 = 4) \wedge (5 - 2x_1 > 0)$ - противоречие.

в) $5 - 2x_1 \leq 0 \rightarrow \langle 5 - 2x_1 \rangle = 5 - 2x_1 \rightarrow (x_1 - 4) - R(5 - 2x_1) = 0$, откуда

$$x_1 = (5R + 4)/(2R + 1).$$

Переходим к пределу $\lim_{R \rightarrow \infty} x_1 = \frac{5}{2} = 2.5$

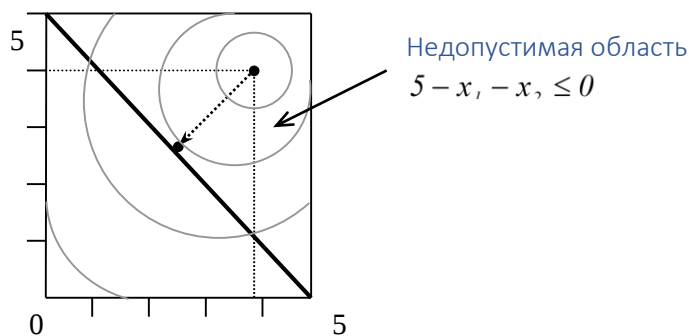
Сделанное предположение (в) не нарушается. При итерационном приближении к точке минимума при возрастающих значениях R получим следующую последовательность точек x :

R	0	1	10	100
$x_1 = x_2$	4	3	2.59	2.5

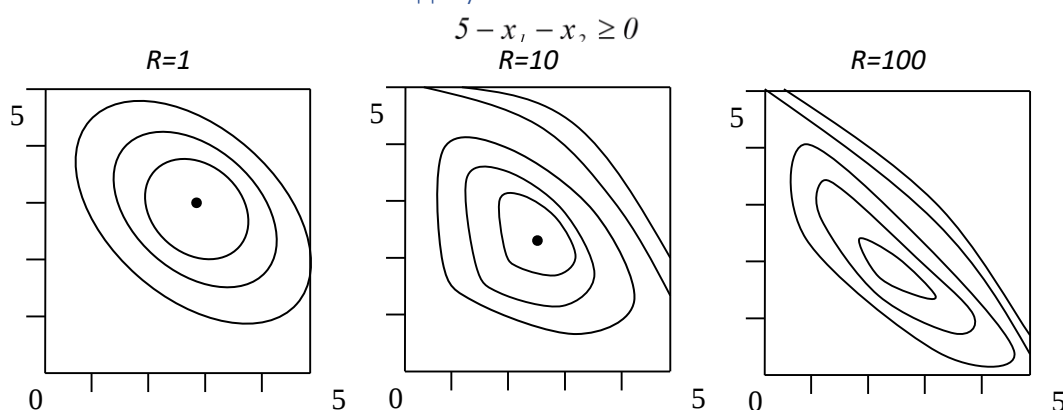
Из решения видно, что при любом конечном R полученная точка не входит в допустимую область, так как не выполняется неравенство $g(x) = 5 - x_1 - x_2 \geq 0$.

По этой причине метод, использующий штраф в виде квадрата срезки называется методом внешней точки.

Траектория движения к решению задачи представляет собой отрезок прямой, соединяющей точку безусловного минимума функции $f(x)$ $x^* = (4; 4)$ с точкой ее условного минимума $x^* = (2.5; 2.5)$.



Область допустимых значений



Используем теперь применительно к этой же задаче логарифмический штраф $P(x, R) = f(x) - R \cdot \ln(g(x))$, т.е. $P(x, R) = (x_1 - 4)^2 - R \cdot \ln(5 - x_1 - x_2)$.

Условия минимума функции $P(x, R)$

$$\frac{\partial P}{\partial x_1} = 2(x_1 - 4) - R \left[\frac{1}{5 - x_1 - x_2} \right] = 0$$

$$\frac{\partial P}{\partial x_2} = 2(x_2 - 4) - R \left[\frac{1}{5 - x_1 - x_2} \right] = 0$$

В силу симметричности выражений $x_1 = x_2$.

С учетом этого имеем $2(x_1 - 4) + [R/(5 - 2x_1)] = 0$, откуда $2x_1^2 - 13x_1 + 20 - R/2 = 0$.

Один из корней дает координату искомой стационарной точки $x_1 = \frac{13}{4} - \frac{1}{4}\sqrt{9 + 4R}$.

Устремляя R к нулю получаем в пределе $x_1 = x_2 = \lim_{R \rightarrow +0} x_1 = \frac{13}{4} - \frac{3}{4} = 2.5$.

Координаты оптимальных точек при различных значениях R :

$x_1 = x_2$	-1.8	1.5	2.34	2.48	2.5
R	100	10	1	0.1	0

Убеждаемся, что стационарные точки штрафной функции сходятся к x^* при $R \rightarrow 0$. При этом каждая стационарная точка является допустимой. Поэтому методы, использующие штрафы, в виде логарифмической функции называются методами внутренней точки. Приведенные примеры демонстрируют принципиальную возможность использования методов штрафных функций для

решения задач условной оптимизации. Однако, применяя его, необходимо иметь в виду два обстоятельства:

во-первых, сходимость метода зависит от степени вытянутости линий уровня штрафной функции;

во-вторых, метод может быть эффективным лишь в том случае, если все подзадачи безусловной минимизации могут быть решены с достаточной степенью точности.

Алгоритм метода штрафных функций

Обозначим

ε_1 - параметр окончания одномерного поиска;

ε_2 - параметр окончания безусловной оптимизации;

ε_3 - параметр окончания работы алгоритма;

$x_0^{(0)}$ - начальное приближение для x ;

$R^{(0)}$ - начальный вектор штрафных параметров;

```
: from scipy.sparse.linalg import spsolve
import numpy as np, sympy as sy, scipy as sp

from numpy import r_, c_, ix_

# f=lambda x: 2*x**2
# g=lambda x: 2*x-2
# F=lambda x,r: f(x)+r/2*g(x)**2;
# x=lambda r: r/(1+r)

f=lambda x: max(0,x**2-4*x)
g=lambda x: x-1
F=lambda x,r: f(x)+r/2*g(x)**2;
x=lambda r: (4+r)/(2+r)

rk=r_[1, 2, 10, 100, 1000]

xk=np.zeros(rk.shape)
Fk=np.zeros(rk.shape)
lk=np.zeros(rk.shape)

k=0
for r in rk:
    xk[k]=x(r)
    Fk[k]=F(xk[k],r)
    lk[k]=r*g(xk[k])
    k=k+1

print('\nx=',xk,'\nF=',Fk,'\nl=',lk)
```

x=	[1.66666667 1.5	1.16666667 1.01960784 1.00199601]
F=	[0.22222222 0.25	0.13888889 0.01922338 0.00199202]
l=	[0.66666667 1.	1.66666667 1.96078431 1.99600798]

МЕТОДЫ ОПТИМИЗАЦИИ В МАШИННОМ ОБУЧЕНИИ

ГРАДИЕНТНЫЕ МЕТОДЫ И МЕТОД НЬЮТОНА.

ОПТИМИЗАЦИЯ В ПРОСТРАНСТВЕ БОЛЬШОЙ РАЗМЕРНОСТИ: ОБЩИЙ МЕТОД СОПРЯЖЁННЫХ ГРАДИЕНТОВ И НЕТОЧНЫЙ (БЕЗГЕССИАННЫЙ) МЕТОД НЬЮТОНА

Обучение нейросети включает в себя минимизацию функции ошибки (loss function) по отношению к ее параметрам (weights), которых может быть очень много. Поэтому существуют множество методов оптимизации для решения данной проблемы.

Градиентный спуск (Gradient Descent)

Инициализация: случайно выбираем параметры x_0

1. Вычисляем градиент: $\nabla f(x_i)$
2. Изменяем параметры в сторону отрицательного градиента: $x_{i+1} = x_i - \alpha \nabla f(x_i)$, где α — некоторый параметр скорости обучения (learning rate)
3. Повторяем предыдущие шаги пока градиент не станет достаточно близок к нулю

Метод Ньютона (Newton's Method)

$$x_{i+1} = x_i - H(f(x_i))^{-1} \nabla f(x_i)$$

$$H(f(x)) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1 \partial x_1} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_n} \end{pmatrix}$$

Hessian-Free оптимизация

$$h(\theta) = \frac{1}{|S|} \sum_{(x,y) \in S} L(t, f(x, \theta))$$

согласно методу Ньютона, определим квадратичную функцию, полученную путем разложения в ряд Тейлора второго порядка:

$$h(\theta + \delta) = M(\delta) = h(\theta) + \nabla h(\theta)^T \delta + \frac{1}{2} \delta^T H \delta \quad (4.1)$$

Далее взяв производную по δ от формулы выше и приравнявая её к нулю получаем:

$$\nabla M(\delta) = \nabla h(\theta) + H \delta = 0 \quad (4.2)$$

Что бы найти δ будем использовать метод сопряженных градиентов (conjugate gradient method).

Метод сопряженных градиентов (CG) — это итерационный метод решения систем линейных уравнений типа $Ax=b$:

алгоритм CG:

Входные данные: $b, A, x_0, i=0$ — шаг алгоритма CG

Инициализация:

1. $r_0 = b - Ax_0$ — вектор ошибки (residual)
2. $d_0 = r_0$ — вектор направления поиска (search direction)

Повторяем пока не выполнится условие остановки:

1. $\alpha_i = \frac{r_i^T r_i}{d_i^T A d_i}$
2. $x_{i+1} = x_i + \alpha_i d_i$
3. $r_{i+1} = r_i - \alpha_i A d_i$
4. $\beta_i = \frac{r_{i+1}^T r_{i+1}}{r_i^T r_i}$
5. $d_{i+1} = r_{i+1} + \beta_i d_i$
6. $i = i + 1$

Умножение гессиана на вектор

Отсюда видно, чтобы посчитать произведение гессиана на вектор надо вычислить:

$$Hv = R(\nabla h(\theta))$$

2. Дампинг (damping).

$$\hat{M}(\delta) = M(\delta) + \frac{\lambda}{2} \delta^T \delta = f(\theta) + \nabla h(\theta)^T \delta + \frac{1}{2} \delta^T \hat{H} \delta$$

где , $\hat{H} = H + \lambda I$, λ — параметр дампинга. Вычисление Hv производится так:

$$\hat{H}v = (H + \lambda I)v = Hv + \lambda Iv \quad (4.6)$$

3. Эвристика Левенберга-Марквардта (Levenberg-Marquardt heuristic).

Для дампинга Тихонова характерно динамическое подстраивание параметра λ .

$$\rho = \frac{h(\theta_k + \delta_k) - h(\delta_k)}{M_k(\delta_k) - M_k(0)} = \frac{h(\theta_k + \delta_k) - h(\delta_k)}{\nabla h(\theta_k)^T \delta_k + \frac{1}{2} \delta_k^T H \delta_k}$$

где k — номер шага НФ алгоритма, δ_k — результат работы CG минимизации. правило обновления λ :

4. Начальное условие для алгоритма сопряженных градиентов (preconditioning).

Инициализация:

1. $r_0 = b - Ax_0$ — вектор ошибки (residual)
2. y_0 — решение уравнения $Py_0 = r_0$
3. $d_0 = y_0$ — вектор направления поиска (search direction)

Повторяем пока не выполнится условие остановки:

1. $\alpha_i = \frac{r_i^T y_i}{d_i^T A d_i}$
2. $x_{i+1} = x_i + \alpha_i d_i$
3. $r_{i+1} = r_i - \alpha_i A d_i$
4. y_i — решение уравнения $Py_i = r_i$

$$5. \quad \beta_i = \frac{r_{i+1}^T y_{i+1}}{r_i^T y_i}$$

$$6. \quad d_{i+1} = y_{i+1} + \beta_i d_i$$

$$7. \quad i = i + 1$$

$$P = \text{diag}(\bar{F}) = \frac{1}{|S|} \sum_{(x,y) \in S} (\nabla L(t, f(x, \theta)))^2 \quad (4.9)$$

5. Инициализация CG — алгоритма.

Полный алгоритм Hessian-Free оптимизации:

Входные данные: θ , λ параметр дампинга, i — шаг итерации алгоритма

Инициализация:

$$1. \quad \delta_0 = 0$$

Основной цикл HF оптимизации:

1. Вычисляем матрицу P .
2. Находим δ_k решая задачу оптимизации с помощью CG или PCG

$$\delta_k = CG(-\nabla h(\theta_k), H, \varepsilon \delta_k, P)$$
3. Обновляем параметр дампинга λ с помощью эвристики Левенберга-Марквардта
4. $\theta_{k+1} = \theta_k + \alpha \delta_k$ — параметр скорости обучения (learning rate)
5. $k = k + 1$

Таким образом метод Hessian-Free оптимизации позволяет решать задачи нахождения минимума функции большой размерности. При этом не требуется нахождения матрицы Гессе напрямую.