

НАУЧНО-ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ ИТМО
ФАКУЛЬТЕТ СИСТЕМ УПРАВЛЕНИЯ И РОБОТОТЕХНИКИ

Отчет по лабораторной работе №1
**Поиск минимума с помощью методов
статического оптимизации**
Вариант 21

Выполнил студент
Проверил преподаватель

Мовчан Игорь Евгеньевич, R3480
Парамонов Алексей Владимирович

Санкт-Петербург
2025

Содержание

1	Поиск глобального минимума	2
1.1	Поиск минимума без ограничений	2
1.2	С ограничением в виде равенства	3
1.3	С ограничением в виде неравенства	4
2	Градиентный поиск минимума	5
2.1	Метод Ньютона-Рафсона	5
2.2	Метод наискорейшего спуска	7
3	Общие выводы	7

1 Поиск глобального минимума

Рассмотрим функцию двух переменных:

$$J(x, u) = 8x^2 + 3u^2 + 6xu + x + 4u - 13$$

Целью работы ставится поиск глобального минимума $J(x, u)$ различными методами. Сперва взглянем на теоретические подходы на основе необходимого и достаточного условий и ограничениях.

1.1 Поиск минимума без ограничений

Для начала примем, что все ограничения отсутствуют. Найдем градиент функции:

$$\nabla J(x, u) = \begin{bmatrix} \frac{\partial J}{\partial x} \\ \frac{\partial J}{\partial u} \end{bmatrix} = \begin{bmatrix} 16x + 6u + 1 \\ 6x + 6u + 4 \end{bmatrix} = 0$$

По необходимому условию градиент должен быть равен нулю, так что в итоге получаем систему на x и u :

$$\begin{cases} 16x + 6u + 1 = 0 \\ 6x + 6u + 4 = 0 \end{cases} \Rightarrow \begin{cases} 10x = 3 \\ 3x + 3u + 2 = 0 \end{cases}$$

Откуда имеем

$$\begin{cases} x = 0.3 \\ 0.9 + 3u = -2 \end{cases} \Rightarrow \begin{cases} x^* = \frac{3}{10} \\ u^* = -\frac{29}{30} \end{cases}$$

Для определения типа полученной точки найдем матрицу Гессе:

$$H(x) = \begin{bmatrix} 16 & 6 \\ 6 & 6 \end{bmatrix} \succcurlyeq 0$$

Она положительно определена, а значит, найденная (x^*, u^*) - точка глобального минимума **выпуклой** функции $J(x, u)$.

1.2 С ограничением в виде равенства

Пусть теперь задано ограничение $c(x, u) = 0$:

$$c(x, u) = x^2 - 8u - 4 = 0 \quad \Rightarrow \quad u = (x^2 - 4)/8$$

Минимум тогда можно найти с помощью функции Лагранжа

$$\begin{aligned} \mathcal{L}(x, u, \lambda) &= J(x, u) + \lambda c(x, u) = \\ &= 8x^2 + 3u^2 + 6xu + x + 4u - 13 + \lambda(x^2 - 8u - 4) \end{aligned}$$

Здесь λ — множитель Лагранжа.

Поиск минимума с ограничениями в виде равенства тогда сводится к решению системы:

$$\begin{cases} \frac{\partial \mathcal{L}}{\partial \lambda} = c(x, u) = 0 \\ \frac{\partial \mathcal{L}}{\partial x} = 0 \\ \frac{\partial \mathcal{L}}{\partial u} = 0 \end{cases} \quad \Rightarrow \quad \begin{cases} x^2 - 8u - 4 = 0 \\ 16x + 6u + 1 + 2\lambda x = 0 \\ 6x + 6u + 4 - 8\lambda = 0 \end{cases}$$

Система имеет единственное вещественное решение

$$\begin{cases} x^* \approx 0.121 \\ u^* \approx -0.498 \\ \lambda^* \approx 0.217 \end{cases}$$

Для проверки того, что найденная точка действительно является точкой глобального минимума подставим полученное из ограничения $c(x, u) = 0$ выражение на u в функцию $J(x, u)$:

$$\begin{aligned} J(x) &= 8x^2 + \frac{3}{64}x^4 - \frac{3}{8}x^2 + \frac{3}{4} + \frac{3}{4}x^3 - 3x + \frac{1}{2}x^2 - 2 - 13 = \\ &= \frac{3}{64}x^4 + \frac{3}{4}x^3 + \frac{65}{8}x^2 - 2x - \frac{57}{4} \end{aligned}$$

Теперь это функция одной переменной с положительным коэффициентом при старшей чётной степени (а это значит, что функция ограничена снизу). Найдём её вторую производную:

$$J'(x) = \frac{3}{16}x^3 + \frac{9}{4}x^2 + \frac{65}{4}x - 2$$

$$J''(x) = \frac{9}{16}x^2 + \frac{9}{2}x + \frac{65}{4}$$

Проверим дискриминант:

$$D = \frac{81}{4} - \frac{9 \cdot 65}{16} = -\frac{261}{16} < 0$$

При этом $J''(x)$ - парабола с ветвями вверх, а значит, вторая производная положительна для всех x . Выходит, найденная (x^*, u^*) действительно является точкой минимума.

1.3 С ограничением в виде неравенства

Пусть задано ограничение $c(x, u) \leq 0$:

$$c(x, u) = x^2 - 8u - 4 \leq 0$$

Для решение задачи введем в рассмотрение функцию Лагранжа

$$\begin{aligned} \mathcal{L}(x, u, \lambda) &= J(x, u) + \lambda c(x, u) = \\ &= 8x^2 + 3u^2 + 6xu + x + 4u - 13 + \lambda(x^2 - 8u - 4) \end{aligned}$$

Найдём экстремальные точки через необходимые условия

$$\left\{ \begin{array}{l} \frac{\partial \mathcal{L}}{\partial \lambda} = c(x, u) \leq 0 \\ \frac{\partial \mathcal{L}}{\partial x} = 0 \\ \frac{\partial \mathcal{L}}{\partial u} = 0 \\ \lambda \geq 0 \\ \lambda c(x, u) = 0 \end{array} \right. \Rightarrow \left\{ \begin{array}{l} x^2 - 8u - 4 \leq 0 \\ 16x + 6u + 1 + 2\lambda x = 0 \\ 6x + 6u + 4 - 8\lambda = 0 \\ \lambda \geq 0 \\ \lambda c(x, u) = 0 \end{array} \right.$$

Сперва рассмотрим случай $\lambda = 0$. Задача тогда сводится к безусловной минимизации $J(x, u)$, которая была уже проведена и для которой получены следующие значения:

$$\begin{cases} x^* = \frac{3}{10} \\ u^* = -\frac{29}{30} \end{cases} \Rightarrow c(x^*, u^*) = \frac{9}{100} + \frac{116}{15} - 4 > 0$$

Найденная точка нарушает ограничение, значит, $\lambda > 0$. Тогда по условию дополнительной нежесткости получаем систему с ограничением в виде равенства, для которой уже была получена точка

$$\begin{cases} x^* \approx 0.121 \\ u^* \approx -0.498 \\ \lambda^* \approx 0.217 \end{cases}$$

Она удовлетворяет неравенству, а значит, является стационарной. Здесь важно, что неравенство $c(x, u) \leq 0$ порождает выпуклое множество, сама $J(x, u)$ тоже является выпуклой, поэтому найденная (x^*, u^*) является точкой минимума.

2 Градиентный поиск минимума

Пусть мы исследуем всё тот же критерий качества

$$J(x, u) = 8x^2 + 3u^2 + 6xu + x + 4u - 13$$

Необходимо динамическими методами найти точки глобального минимума. Далее исследуем два различных подхода.

2.1 Метод Ньютона-Рафсона

Поставленная задача оптимизации с помощью данного метода решается последовательным нахождением точек

$$\bar{x}^{(n+1)} = \bar{x}^{(n)} - H^{-1}(\bar{x}^{(n)}) \nabla J(\bar{x}^{(n)})$$

Здесь были введены вектор переменных и градиент в точке

$$\bar{x}^{(n)} = \begin{bmatrix} x(n) \\ u(n) \end{bmatrix}, \quad \nabla J(\bar{x}^{(n)}) = \begin{bmatrix} \frac{\partial J(\bar{x}^{(n)})}{\partial x} \\ \frac{\partial J(\bar{x}^{(n)})}{\partial u} \end{bmatrix} = \begin{bmatrix} 16x + 6u + 1 = 0 \\ 6x + 6u + 4 = 0 \end{bmatrix}$$

А также матрица Гессе вторых частных производных:

$$H(\bar{x}^{(n)}) = \begin{bmatrix} \frac{\partial}{\partial x} \left(\frac{\partial J(\bar{x}^{(n)})}{\partial x} \right) & \frac{\partial}{\partial x} \left(\frac{\partial J(\bar{x}^{(n)})}{\partial u} \right) \\ \frac{\partial}{\partial u} \left(\frac{\partial J(\bar{x}^{(n)})}{\partial x} \right) & \frac{\partial}{\partial u} \left(\frac{\partial J(\bar{x}^{(n)})}{\partial u} \right) \end{bmatrix} = \begin{bmatrix} 16 & 6 \\ 6 & 6 \end{bmatrix}$$

Суть подхода заключается в движении в направлении наискорейшего убывания - по антиградиенту - с учётом кривизны функции с помощью матрицы Гессе.

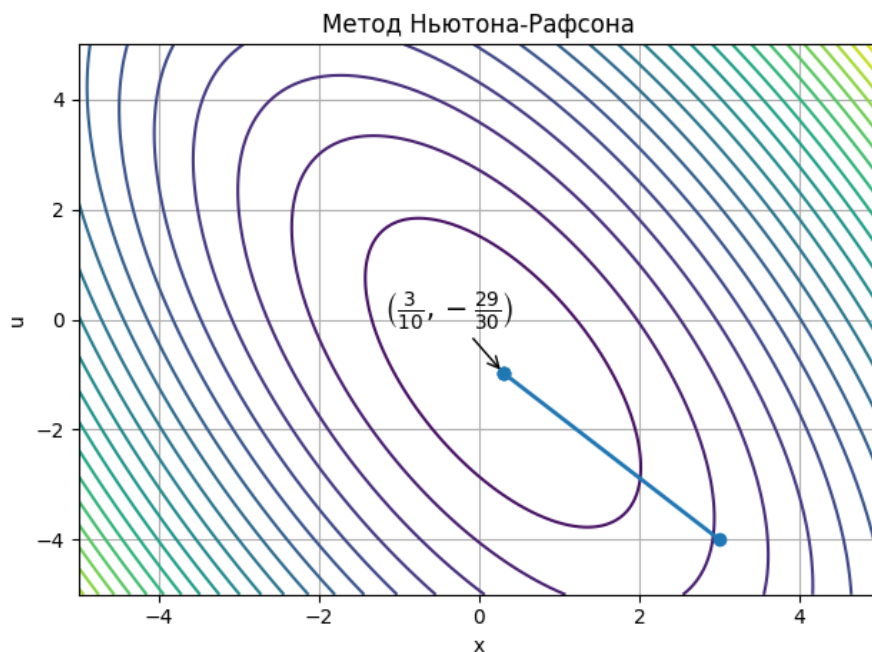


Рис. 1: Применение метода Ньютона-Рафсона для поиска минимума

Применение метода при точке старта $\bar{x}^{(0)} = [3 \ -4]^T$ продемонстрировано на рисунке 1. Критерием останова было выбрано увеличение значения $J(\bar{x}^{(n)})$. Видим, что траектория приходит в уже найденную в первом пункте точку минимума за один шаг. Успех!

2.2 Метод наискорейшего спуска

В методе Ньютона-Рафсона на каждом шаге необходимо считать матрицу вторых производных, а после её обращать, это *очень* тяжело. В связи с этим возникает идея упрощения множителя перед антиградиентом до константы γ , получаем:

$$\bar{x}^{(n+1)} = \bar{x}^{(n)} - \gamma \nabla J(\bar{x}^{(n)})$$

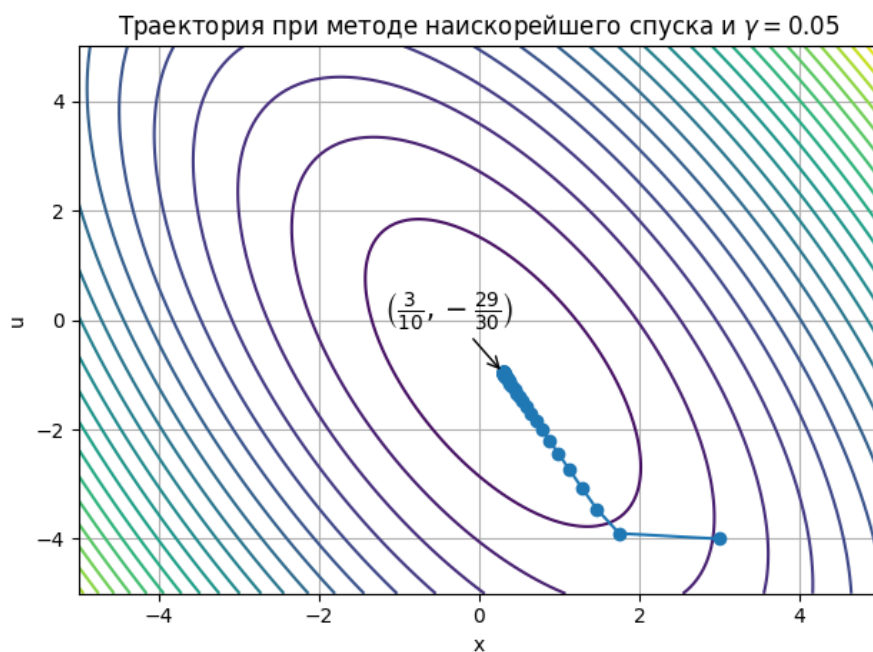
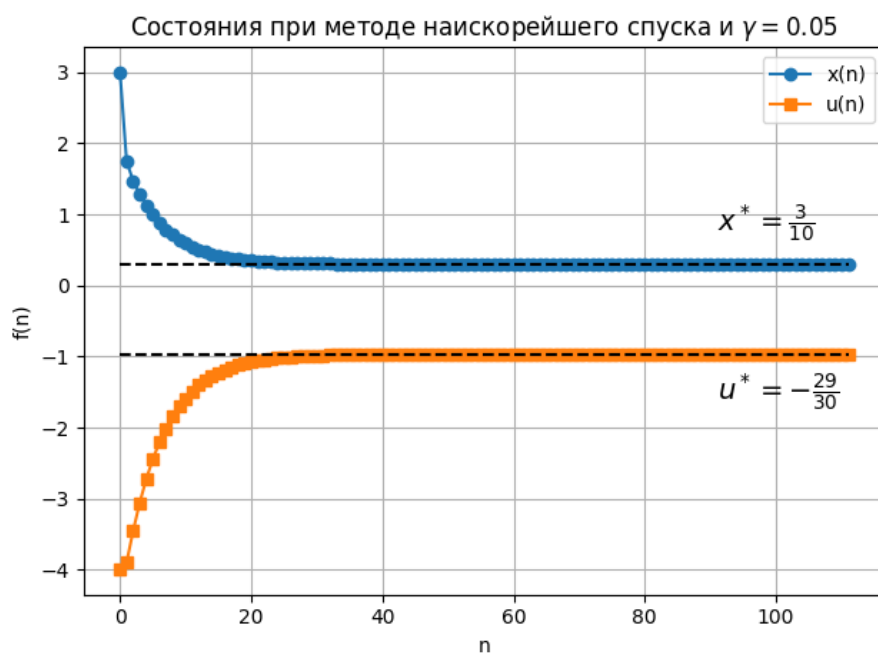
Это и назовём методом наискорейшего спуска или же градиентным спуском. Суть его также заключается в движении в направлении наискорейшего убывания функции, но более просто - без учёта локальной кривизны. Поэтому на практике, если взять слишком большое значение γ , то состояние $\bar{x}^{(n)}$ начнёт «раскачиваться».

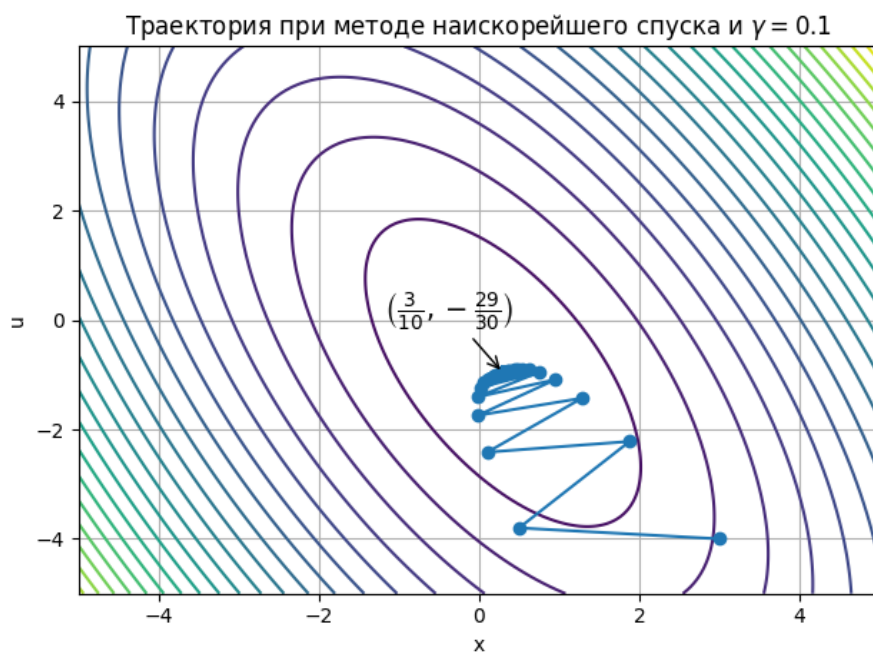
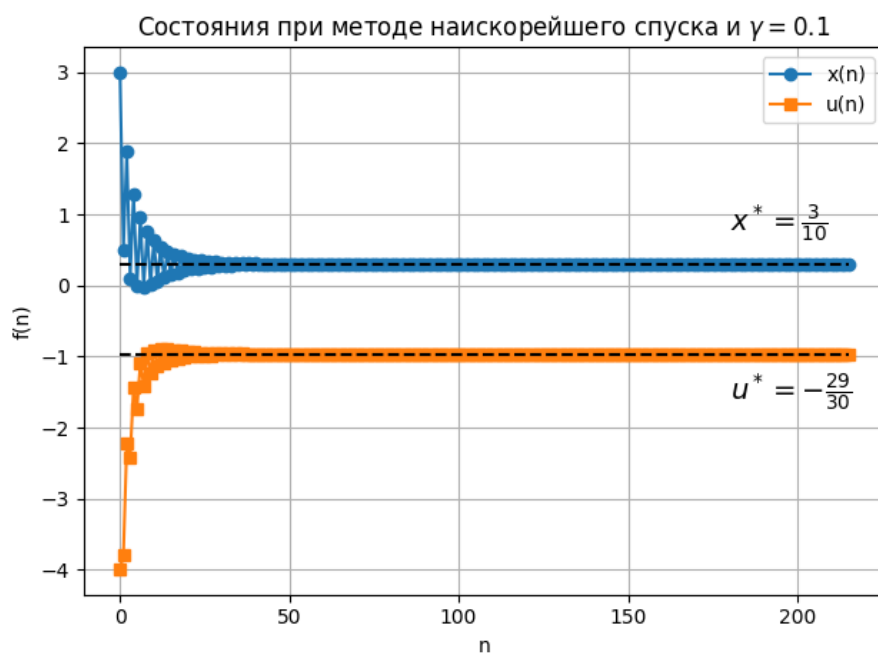
Итак, применение метода при параметрах $\gamma = 0.05$ и $\gamma = 0.125$, начальном условии $\bar{x}^{(0)} = [3 \ -4]^T$ и критерии останова в виде увеличения значения $J(\bar{x}^{(n)})$ приведено на рисунках 2-5.

Можем видеть, что параметр $\gamma = 0.05$ соответствует апериодической сходимости, а его повышение до $\gamma = 0.1$ уже приводит к колебаниям и требует большего числа шагов для попадания в окрестность точки глобального минимума. В целом обе вариации отработали хорошо и дошли до вычисленных стационарных точек.

3 Общие выводы

В ходе лабораторной работы были исследованы аналитические и численные методы поиска глобального минимума квадратичной функции двух переменных. Показано, что для выпуклой функции без ограничений минимум находится точно через систему уравнений,

Рис. 2: Применение градиентного спуска: траектория, $\gamma = 0.05$ Рис. 3: Применение градиентного спуска: состояния, $\gamma = 0.05$

Рис. 4: Применение градиентного спуска: траектория, $\gamma = 0.1$ Рис. 5: Применение градиентного спуска: состояния, $\gamma = 0.1$

задаваемую градиентом, и подтверждается положительной определённой матрицей Гессе. При наличии ограничений — как равенства, так и неравенства — использован метод множителей Лагранжа с учётом условий Куна–Таккера, и найдены соответствующие стационарные точки, удовлетворяющие критериям оптимальности. Численные методы — Ньютона–Рафсона и градиентного спуска — подтвердили аналитические результаты: первый обеспечил сходимость за один шаг благодаря точному учёту кривизны, второй — продемонстрировал чувствительность к выбору шага, также сошёлся, причем с меньшей вычислительной нагрузкой на каждый шаг.