Лабораторная работа №1
Методы многомерной минимизации
(безусловный экстремум)

Цель работы

Приобретение навыков численного решения задач поиска безусловного экстремума действительной функции от n переменных.

Теоретическая часть

Рассмотрим задачу поиска минимума функции $f : D \to \mathbb{R}$ в виде

$$f(x) \to \min, \qquad x \in D.$$

то есть необходимо найти минимум целевой функции $f(x)$ на допустимом множестве $D \subset \mathbb{R}^n$. По аналогии с одномерной минимизацией, если $D = \mathbb{R}^n$, то задачу оптимизации называют безусловной.

Очевидно, что с ростом количества переменных сильно возрастает объем вычислений, поэтому численные методы для решения задачи безусловной минимизации функции многих переменных значительно сложнее и разнообразнее методов одномерной минимизации. Условно их можно разделить на три класса

1. методы прямого поиска или методы нулевого порядка, которые основаны на использовании информации только о целевой функции;

2. методы первого порядка, в алгоритмах которых используется вычисление целевой функции и ее производных до первого порядка включительно; К таким методам относятся различные градиентные методы;

3. методы второго порядка, в которых используется вычисление целевой функции и ее производных до второго порядка включительно.

Методы первого и второго порядка являются хорошо изученными, теоретически обоснованными, доказана сходимость метода, сделаны оценки точности. Большинство методов прямого поиска носят эвристический характер, мало изучены и не имеют обоснования. Поэтому в данной лабораторной работе для поиска минимума будут использоваться методы первого и второго порядка. Универсального метода решения задачи бесусловной минимизации не существует. Выбор применяемого метода зависит от конкретного вида целевой функции и условий, налагаемых на нее.

Методы первого порядка

К методам первого порядка относятся различные градиентные методы. Предполагается, что целевая функция непрерывна и имеет непрерывные производные первого порядка.Поиск точки минимума обычно прекращается, если выполняется один из **критериев останова**

- $||x^{k+1} - x^k|| \leqslant \varepsilon$;
- $|f(x^{k+1}) - f(x^k)| \leqslant \varepsilon$;
- $\max\limits_{i=\overline{1,n}} \left| \dfrac{\partial f}{\partial x^i} \right| < \varepsilon$;
- $||\nabla f||^2 = \sum\limits_{i=1}^{n} \left(\dfrac{\partial f}{\partial x^i} \right)^2 < \varepsilon$,

где $\varepsilon > 0$ — требуемая точность решения, $||x|| = \sqrt{x_1^2 + x_2^2 + ...x_n^2}$ - евклидова норма вектора x.

Метод градиентного спуска с дробным шагом

1. Зададим следующие значения: $k = 0$; $x^k \in D$; t_k — длина шага; ε; M — допустимое число итераций.
2. Вычислим $\nabla f(x^k)$. Вычислим $x^{k+1} = x^k - t_k \nabla f(x^k)$.
3. Если $f(x^{k+1}) - f(x^k) < 0$, то переходим к пункту 4, в противном случае принимаем $t_k = \dfrac{t_k}{2}$ и переходим к пункту 2.
4. Если выполнен критерий останова или $k+1 > M$, то поиск решения x^* завершен, $x^* = x^{k+1}$. В противном случае увеличим номер шага $k = k + 1$ и переходим к пункту 2.

Метод наискорейшего градиентного спуска

1. Зададим следующие значения: $k = 0$; $x^k \in D$; $\varepsilon > 0$; M — допустимое число итераций.
2. Вычислим $d_k = -\nabla f(x^k)$. Найдем такое λ_k, что $f(x^k + \lambda_k d_k) = \min\limits_{\lambda \geqslant 0} f(x^k + \lambda d_k)$.
3. Вычислим $x^{k+1} = x^k + \lambda_k d_k$.
4. Если выполнен критерий останова или $k+1 > M$, то поиск решения x^* завершен. $X^* = x^{k+1}$ В противном случае увеличим номер шага $k = k + 1$ и переходим к пункту 2.

Метод сопряженных градиентов

1. Зададим следующие значения: $k = 0$; $x^k \in D$; $\varepsilon > 0$; M — допустимое число итераций.

2. Вычислим $d_k = -\nabla f(x^k)$. Найдем такое λ_k, что $f(x^k + \lambda_k d_k) = \min\limits_{\lambda \geqslant 0} f(x^k + \lambda d_k)$.

3. Вычислим $x^{k+1} = x^k + \lambda_k d_k$. Если $||d_k|| < \varepsilon$, то поиск решения x^* завершен. $x^* = x^{k+1}$. В противном случае увеличим номер шага $k = k + 1$ и переходим к пункту 4.

4. Новое направление поиска $d_k = -\nabla f(x^k) + \omega_k d_{k-1}$, где $\omega = \dfrac{||\nabla f(x^k)||^2}{||\nabla f(x^{k-1})||^2}$. Вычислим $x^{k+1} = x^k + \lambda_k d_k$.

5. Если $||d_k|| < \varepsilon$ или $k + 1 > M$, то поиск решения x^* завершен. $x^* = x^{k+1}$. В противном увеличим номер шага $k = k + 1$ и переходим к пункту 2.

Методы второго порядка

Методы второго порядка основаны на использовании второй производной целевой функции, а точнее матрицы Гессе, составленной из этих производных

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

Метод Ньютона

1. Зададим следующие значения: $k = 0$; $x^k \in D$; $\varepsilon > 0$ — требуемая точность решения; M — допустимое число итераций.

2. Если $||\nabla f(x^k)|| < \varepsilon$, то поиск решения x^* завершен, $x^* = x^k$. В противном случае переходим к пункту 3.

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

4. Вычислим $x^{k+1} = x^k + d_k$.

5. Если $||\nabla f(x^{k+1})|| < \varepsilon$ или $k + 1 > M$, то поиск решения x^* завершен. $x^* = x^{k+1}$. В противном случае увеличим номер шага $k = k+1$ и возвращаемся к пункту 3.

Метод Ньютона–Рафсона (Модифицированный метод Ньютона)

1. Зададим следующие значения: $k = 0$; $x^k \in D$; $\varepsilon > 0$ — требуемая точность решения; M — допустимое число итераций.

2. Если $||\nabla f(x^k)|| < \varepsilon$, то поиск решения x^* завершен, $x^* = x^k$. В противном случае переходим к пункту 3.

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

4. Вычислим $x^{k+1} = x^k + d_k$. Найдем такое λ_k, что $f(x^k + \lambda_k d_k) = \min\limits_{\lambda \geqslant 0} f(x^k + \lambda d_k)$.

5. Вычислим $x^{k+1} = x^k + \lambda_k d_k$.

6. Если $||\nabla f(x^{k+1})|| < \varepsilon$ или $k + 1 > M$, то поиск решения x^* завершен. $x^* = x^{k+1}$. В противном случае увеличим номер шага $k = k + 1$ и переходим к пункту 2.

Метод Марквардта

1. Зададим следующие значения: $k = 0$; $x^k \in D$; λ_0 — некоторое большое значение, например, $\lambda_0 = 10^4$;$\varepsilon > 0$ — требуемая точность решения; M — допустимое число итераций.

2. Если $||\nabla f(x^k)|| < \varepsilon$, то поиск решения x^* завершен, $x^* = x^k$. В противном случае переходим к пункту 3.

3. Вычислим $d_k = -[H(x^k) + \lambda_k E]^{-1}\nabla f(x^k)$.

4. Вычислим $x^{k+1} = x^k + d_k$.

5. Проверим выполнение неравенства $f(x^{k+1}) < f(x^k)$. Если выполняется, то следующий шаг 6, если нет — шаг 7.

6. $\lambda_{k+1} = \dfrac{1}{2}\lambda_k$, $k = k + 1$. Перейти к шагу 2.

7. $\lambda_k = 2\lambda_k$. Вернуться к шагу 3.

Анализ решения

После того, как с помощью одного из численных методов оптимизации найдено приближенное решение задачи, необходимо провести анализ того, является ли точка x^* приближением решения поставленной задачи оптимизации.

Если $f(x) \in C^2$, то нужно выполнить проверку достаточных условий минимума: положительноопределенности матрицы Гессе ($H(x^*) > 0$). Если условие верно, то точка x^* является приближением искомого решения задачи оптимизации.

Если $f(x) \in C^1$, то проводится проверка целевой функции на выпуклость в окрестности точки x^* [1]. Функция $f(x) \in C^1$ выпукла в окрестности Q тогда и только тогда, когда $\forall x, y \in Q,\ f(x+y) \geqslant f(x) + (\nabla f(x), y)$. Если функция $f(x)$ выпукла, точка x^* является приближением искомого решения задачи оптимизации.

Для строго выпуклой функции поиск глобального минимума проводится аналогично поиску локального минимума. Если целевая функция имеет несколько локальных минимумов, то глобальный минимум находится после сравнения всех локальных минимумов.

Порядок выполнения работы

1. Ознакомиться с постановкой задачи, определяемой вариантом задания к лабораторной работе.

2. Найти решение задачи безусловной оптимизации для заданной целевой функции, используя теоремы о необходимых и достаточных условиях экстремума.

3. Найти решение задачи безусловной оптимизации для заданной целевой функции в пакете Maple (Optimization).

4. Найти приближенное решение задачи безусловной минимизации численными методами поиска безусловного экстремума согласно варианту, с заданной точностью $\varepsilon = 0.01$.

5. Провести анализ найденного приближенного решения (является ли стационарная точка точкой экстремума).

6. Ответить на вопросы, указанные в задании.

7. По результатам выполненной лабораторной работы составить отчет, содержащий:

   - цель работы;
   - постановку задачи;
   - аналитическое решение поставленной задачи;
   - блок-схемы использованных численных методов поиска безусловного экстремума для решаемой задачи;
   - выводы об эффективности использованного метода;
   - ответы на контрольные вопросы, приведенные в задании.

Контрольные вопросы

1. Необходимые и достаточные условия локального экстремума функции $f : D \to \mathbb{R}, D \subset \mathbb{R}^n$
2. Как определяется порядок численного метода решения задачи оптимизации? Приведите примеры численных методов нулевого, первого и второго порядка.
3. Какие методы одномерной минимизации Вы знаете? В каких численных методах первого и второго порядка они используются?
4. Комбинацией каких методов оптимизации является метод Марквардта?
5. Для каких функций эффективно применение рассмотренных методов первого и второго порядка?

Задания к лабораторной работе №1

Вариант	Целевая функция	Численный метод	
		первого порядка	**второго порядка**
1.	$(x_1 + 6x_2)^2 + (x_1 + 2)^2$	градиентного спуска с дробным шагом	Марквардта
2.	$(x_1 + 4x_2)^2 + (x_1 - 3)^2$	сопряженных градиентов	Марквардта
3.	$(x_1 - 5x_2)^2 + (x_1 + 6)^2$	наискорейшего градиентного спуска	Ньютона
4.	$(x_1 - 6x_2)^2 + (x_1 + 1)^2$	градиентного спуска с дробным шагом	Ньютона-Рафсона
5.	$(x_1 - 2x_2)^2 + (x_1 + 5)^2$	сопряженных градиентов	Ньютона–Рафсона
6.	$(x_1 + 2x_2)^2 + (x_1 - 4)^2$	наискорейшего градиентного спуска	Ньютона–Рафсона
7.	$(x_1 - 2x_2)^2 + (x_2 - 3)^2$	градиентного спуска с дробным шагом	Марквардта
8.	$(x_1 + x_2)^2 + (x_2 - 1)^2$	сопряженных градиентов	Марквардта
9.	$(x_1 + x_2)^2 + (x_2 + 6)^2$	наискорейшего градиентного спуска	Ньютона
10.	$(x_1 - 2x_2)^2 + (x_2 - 9)^2$	градиентного спуска с дробным шагом	Ньютона-Рафсона
11.	$(x_1 + 9x_2)^2 + (x_2 - 1)^2$	сопряженных градиентов	Ньютона–Рафсона
12.	$(x_1 + 3x_2)^2 + (x_2 + 5)^2$	наискорейшего градиентного спуска	Ньютона–Рафсона
13.	$(x_1 + x_2)^2 + (x_2 + 2)^2$	градиентного спуска с дробным шагом	Марквардта
14.	$(x_1 - 2x_2)^2 + (x_2 - 3)^2$	сопряженных градиентов	Марквардта
15.	$(x_1 + 2x_2)^2 + (x_2 - 3)^2$	наискорейшего градиентного спуска	Ньютона
16.	$(x_1 + 5x_2)^2 + (x_2 - 1)^2$	градиентного спуска с дробным шагом	Ньютона