

Документ подписан простой электронной подписью

Информация о владельце:

ФИО: Локтионова Оксана Геннадьевна

Должность: проректор по учебной работе

Дата подписания: 09.07.2024 09:32:54

Уникальный программный ключ:

0b817ca911e6668abb13a5d426d39e5f6c11eabbf73e943df4a4851fda56d089

МИНОБРНАУКИ РОССИИ

Федеральное государственное бюджетное  
образовательное учреждение высшего образования  
«Юго-Западный государственный университет»  
(ЮЗГУ)

Кафедра вычислительной техники



## ЧИСЛЕННОЕ РЕШЕНИЕ ЗАДАЧ БЕЗУСЛОВНОЙ МИНИМИЗАЦИИ МЕТОДАМИ ПЕРВОГО ПОРЯДКА

Методические указания к лабораторным и практическим занятиям для  
студентов направлений подготовки 09.03.01 и 09.04.01

Информатика и вычислительная техника



## 1. Цель работы

Изучение численных методов многомерной безусловной минимизации методами первого порядка.

## 2. Постановка задачи

Рассмотрим задачу безусловной минимизации дифференцируемой функции многих переменных  $\mathbf{f}(\mathbf{x})$ ,  $\mathbf{f} \in \mathbb{R}$ ,  $\mathbf{x} \in \mathbb{R}^n$ .

Пусть  $\mathbf{x}^{(k)}$  – приближение к точке минимума  $\mathbf{x}_*$  и  $\mathbf{g}^{(k)} = \mathbf{g}(\mathbf{x}^{(k)})$ . Как мы знаем, что в малой окрестности  $\mathbf{x}^{(k)}$  на направление наискорейшего убывания функции  $\mathbf{f}(\mathbf{x})$  указывает антиградиент  $-\mathbf{g}^{(k)}$ .

В градиентных методах направление спуска из точки  $\mathbf{x}^{(k)}$  выбирается  $(p)^{(k)} = -\mathbf{g}^{(k)}$ .

Таким образом в градиентных методах

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} - \alpha_k \mathbf{g}^{(k)}, \quad k = 0, 1, 2, \dots \quad (1)$$

В зависимости от того, как выбирается шаг  $\alpha_k$ , существуют различные модификации метода. Эти методы относятся к методам первого порядка.

## 3. Метод наискорейшего спуска

Пусть

$$\varphi_k(\alpha) = \mathbf{f}(\mathbf{x}^{(k)} - \alpha \mathbf{g}^{(k)}) \quad (2)$$

есть функция скалярной переменной  $\alpha \geq 0$ .

В методе *наискорейшего градиентного спуска* шаг  $\alpha_k$  выбирается из условия

$$\alpha_k = \min_{\alpha > 0} \varphi_k(\alpha).$$

Метод был предложен в 1845 г. французским математиком О.Коши.

**Теорема 1.** Пусть  $\mathbf{f}(\mathbf{x})$  дифференцируема. Тогда в итерационном процессе

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} - \alpha_k \mathbf{g}^{(k)}, \quad \alpha_k = \min_{\alpha > 0} \varphi_k(\alpha), \quad \varphi_k(\alpha) = \mathbf{f}(\mathbf{x} - \alpha \mathbf{g}^{(k)}), \quad k = 0, 1, 2, \dots$$

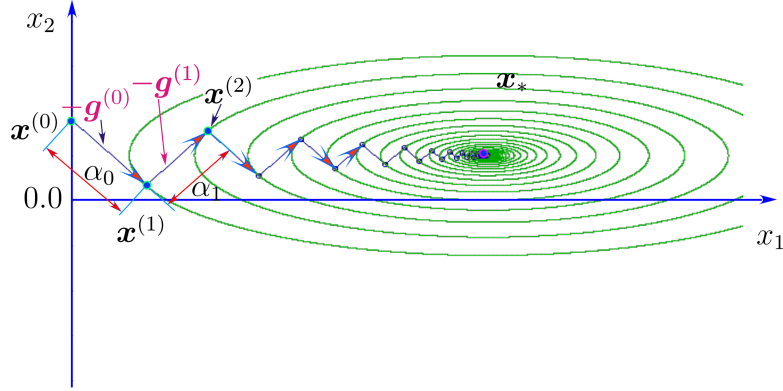


Рис. 1.

для любого  $k \geq 1$  справедливо

$$\left( \nabla \mathbf{f}(\mathbf{x}^{(k+1)}), \mathbf{p}^{(k)} \right) = 0, \quad \mathbf{x} = (x_1, x_2, \dots, x_n)^T, \quad \mathbf{p} = (p_1, p_2, \dots, p_n)^T, \quad (3)$$

где « $T$ » – знак транспонирования.

*Доказательство.* Запишем необходимое условие минимума функции  $\varphi(\alpha)$  (см. (2))

$$\varphi'_k(\alpha) = 0,$$

используя правило дифференцирования сложной функции

$$\varphi'_k(\alpha) = \frac{d\varphi_k(\alpha)}{d\alpha} = \sum_{i=1}^n \frac{\partial \mathbf{f}(\mathbf{x}_i^{k+1})}{\partial x_i} \cdot \frac{\partial x^{k+1}}{\partial \alpha} = 0.$$

Принимая во внимание, что

$$x_i^{(k+1)} = x_i^{(k)} + \alpha p_i^{(k)},$$

получаем условие (3). □

На рис. 1 изображена геометрическая иллюстрация метода наискорейшего градиентного спуска.

Из начальной точки  $\mathbf{x}^{(0)}$  в направлении  $\mathbf{p}^{(0)} = -\mathbf{g}^{(0)}$  спуск продолжается до тех пор, пока либо не будет достигнута стационарная точка, в которой  $\nabla \mathbf{f}(\mathbf{x}^{(k+1)}) = 0$ , или прямая не коснется в точке  $\mathbf{x}^{(k+1)}$  некоторой линии уровня. Равенство (3) и есть условие касания.

#### 4. Минимизации квадратичной функции методом наискорейшего градиентного спуска

Решим задачу минимизации квадратичной функции

$$\mathbf{f}(\mathbf{x}) = \frac{1}{2} (A\mathbf{x}, \mathbf{x}) - (\mathbf{b}, \mathbf{x}), \quad A \in \mathbb{R}^{n \times n}, \quad \mathbf{b}, \mathbf{x} \in \mathbb{R}^n. \quad (4)$$

Здесь  $A$  – симметричная положительно-определенная матрица.

Как мы уже знаем, градиент квадратичной функции

$$\nabla \mathbf{f}(\mathbf{x}) = \mathbf{g}(\mathbf{x}) = A\mathbf{x} - \mathbf{b}.$$

Тогда формула (1) принимает вид

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} - \alpha_k \left( A\mathbf{x}^{(k)} - \mathbf{b} \right), \quad k = 0, 1, 2, \dots$$

Найдем функцию  $\varphi_k(\alpha)$ :

$$\varphi_k(\alpha) = \frac{1}{2} \left( A(\mathbf{x}^{(k)} - \alpha \mathbf{g}^{(k)}), \mathbf{x}^{(k)} - \alpha \mathbf{g}^{(k)} \right) - (\mathbf{b}, \mathbf{x}^{(k)} - \alpha \mathbf{g}^{(k)}).$$

Отсюда отсюда получим квадратичную функцию переменной  $\alpha$

$$\varphi_k(\alpha) = \frac{1}{2} \left( A\mathbf{g}^{(k)}, \mathbf{g}^{(k)} \right) \cdot \alpha^2 - (\mathbf{g}^{(k)}, \mathbf{g}^{(k)}) \cdot \alpha + \frac{1}{2} \left( A\mathbf{x}^{(k)}, \mathbf{x}^{(k)} \right) - (\mathbf{b}, \mathbf{x}^{(k)})$$

Запишем необходимое условие минимума

$$\varphi'_k(\alpha) = 0,$$

где

$$\varphi'_k(\alpha) = \left( A\mathbf{g}^{(k)}, \mathbf{g}^{(k)} \right) \cdot \alpha - (\mathbf{g}^{(k)}, \mathbf{g}^{(k)}).$$

Таким образом,

$$\alpha_k = \frac{(\mathbf{g}^{(k)}, \mathbf{g}^{(k)})}{(A\mathbf{g}^{(k)}, \mathbf{g}^{(k)})}.$$

**Алгоритм**

$\varepsilon \leftarrow 10^{-12}$ ;

$\mathbf{x} \leftarrow \mathbf{x}_0$ ;

REPEAT

$\mathbf{x}_0 \leftarrow \mathbf{x}$ ;

$\mathbf{g}_0 \leftarrow A\mathbf{x}_0 - \mathbf{b}$ ;

$\alpha_0 = \frac{(\mathbf{g}_0, \mathbf{g}_0)}{(A\mathbf{g}_0, \mathbf{g}_0)}$ ;

$\mathbf{x} \leftarrow \mathbf{x}_0 - \alpha_0(A\mathbf{x}_0 - \mathbf{b})$ ;

UNTIL  $\|\mathbf{g}_0\| < \varepsilon$ ;

## 5. Метод наискорейшего градиентного спуска

Итак, подведем итог. Пусть дана функция многих переменных  $\mathbf{f}(\mathbf{x})$ ,  $\mathbf{f} \in \mathbb{R}$ ,  $\mathbf{x} \in \mathbb{R}^n$ , которая имеет непрерывные частные производные во всех его точках.

Требуется найти локальный минимум функции  $\mathbf{f}(\mathbf{x})$ , т.е. такую точку  $\mathbf{x}_* \in \mathbb{R}^n$ , в которой она принимает минимальное значение

$$\mathbf{x}_* = \min_{\mathbf{x} \in \mathbb{R}^n} \mathbf{f}(\mathbf{x}).$$

Решение задачи, как мы выяснили выше, состоит в построении последовательности приближений  $\mathbf{x}^{(0)}$ ,  $\mathbf{x}^{(1)}$ ,  $\mathbf{x}^{(2)}$ , ... к точке  $\mathbf{x}_*$ , таких что  $\mathbf{f}(\mathbf{x}^{k+1}) < \mathbf{f}(\mathbf{x}^k)$ ,  $k = 0, 1, 2, \dots$  (т.е. выполняется условие убывания).

Такая последовательность вычисляется

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} - \alpha_k \mathbf{g}^{(k)}, \quad \mathbf{g}^{(k)} = \nabla \mathbf{f}(\mathbf{x}^k), \quad k = 0, 1, 2, \dots \quad (5)$$

Шаг  $\alpha_k$  рассчитывается для каждого шага  $k$  из условия

$$\alpha_k = \min_{\alpha > 0} \varphi_k(\alpha), \quad \varphi_k(\alpha) = \mathbf{f}(\mathbf{x}^{(k)} - \alpha \mathbf{g}^{(k)}).$$

Построение последовательности приближений  $\{\mathbf{x}^{(k)}\}$ ,  $k = 0, 1, 2, \dots$  заканчивается в точке  $\mathbf{x}^{(k)}$ , в которой  $\|\nabla \mathbf{f}(\mathbf{x}^k)\| < \varepsilon_1$  или принудительно, когда  $k \geq m$ , где  $m$  – предельное число итераций.

Можно заканчивать при одновременном выполнении двух условий  $\|\mathbf{f}(\mathbf{x}^{k+1}) - \mathbf{f}(\mathbf{x}^k)\| < \varepsilon_2$  и  $\|\mathbf{x}^{k+1} - \mathbf{x}^k\| < \varepsilon_2$ .

С другой стороны, выбор критерия останова – задача не совсем простая.

**Алгоритм** в общем случае включает

- Шаг 1. Задать  $\mathbf{x}^{(0)}$ ,  $\varepsilon_1 > 0$ ,  $\varepsilon_2 > 0$  и найти найти градиент функции в произвольной точке  $\nabla \mathbf{f}(\mathbf{x} = \left( \frac{\partial \mathbf{f}(\mathbf{x})}{\partial x_1}, \dots, \frac{\partial \mathbf{f}(\mathbf{x})}{\partial x_n} \right)$ . Это означает, что надо выписать частные производные первого порядка  $\mathbf{f}(\mathbf{x})$  по всем компонентам  $x_1, x_2, \dots, x_n$  вектора  $\mathbf{x}$  или составить алгоритм численного нахождения градиента функции  $\mathbf{f}(\mathbf{x})$  методом «конечно-разностной» аппроксимации производных.
- Шаг 2. Положить  $k = 0$ .
- Шаг 3. Вычислить  $\nabla \mathbf{f}(\mathbf{x}^{(k)})$ .

- Шаг 4. Проверить условие останова  $\|\nabla \mathbf{f}(\mathbf{x}^k)\| < \varepsilon_1$ :
  - (а) если условие выполняется, то  $\mathbf{x}_* = \mathbf{x}^{(k)}$ ;
  - (б) если нет, то перейти к шагу 5.
- Шаг 5. Проверить условие  $k > m$ :
  - (а) если условие выполняется, то мы говорим, что принудительно завершаем работу алгоритма (вопрос, насколько близко  $\mathbf{x}^{(k)}$  к  $\mathbf{x}_*$ , требует дополнительного исследования);
  - (б) если нет, то перейти к шагу 6.
- Шаг 6. Вычислить шаг  $\alpha_k$  как

$$\alpha_k = \min_{\alpha > 0} \varphi_k(\alpha), \quad \varphi_k(\alpha) = \mathbf{f}(\mathbf{x}^{(k)}) - \alpha \mathbf{g}^{(k)}.$$

- Шаг 7. Вычислить

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} - \alpha_k \nabla \mathbf{f}(\mathbf{x}^k).$$

- Шаг 8. Проверить выполнение условий  $\|\mathbf{f}(\mathbf{x}^{k+1}) - \mathbf{f}(\mathbf{x}^k)\| < \varepsilon_2$  и  $\|\mathbf{x}^{k+1} - \mathbf{x}^k\| < \varepsilon_2$ :
  - (а) если условия выполняются то  $\mathbf{x}_* = \mathbf{x}^{(k+1)}$  и закончить поиск.
  - (б) если хотя бы одно из условий не выполняется, то положить  $k = k + 1$  и перейти к шагу 3.

## 6. Метод сопряженных градиентов

Рассмотрим задачу минимизации квадратичной функции

$$\mathbf{f}(\mathbf{x}) = \frac{1}{2} (\mathbf{A}\mathbf{x}, \mathbf{x}) - (\mathbf{b}, \mathbf{x}), \quad \mathbf{A} \in \mathbb{R}^{n \times n}, \quad \mathbf{b}, \mathbf{x} \in \mathbb{R}^n.$$

Ненулевые векторы  $\mathbf{p}^{(0)}, \mathbf{p}^{(1)}, \dots, \mathbf{p}^{(n-1)}$  называются взаимно-сопряженными относительно  $\mathbf{A}$ , если

$$\left( \mathbf{p}^{(k)}, \mathbf{p}^{(i)} \right) = 0,$$

для всех  $k \neq i$ .

В методе сопряженных направлений

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \alpha_k \mathbf{p}^{(k)}, \quad k = 0, 1, 2, \dots, n - 1, \quad (6)$$

где векторы  $\mathbf{p}^{(0)}, \mathbf{p}^{(1)}, \dots, \mathbf{p}^{(n-1)}$  взаимно сопряжены, а  $\alpha_k$ :

$$\alpha_k = -\frac{(\mathbf{g}^{(k)}, \mathbf{p}^{(k)})}{(A\mathbf{p}^{(k)}, \mathbf{p}^{(k)})}$$

находится как решение

$$\alpha_k = \min_{\alpha > 0} \varphi_k(\alpha), \quad \varphi_k(\alpha) = \mathbf{f}(\mathbf{x} + \alpha \mathbf{p}^{(k)}).$$

### 6.1. Метод сопряженных градиентов (метод Флетчера -Ривса)

В этом методе векторы  $\mathbf{p}^{(0)}, \mathbf{p}^{(1)}, \dots, \mathbf{p}^{(n-1)}$  находятся

$$\mathbf{p}^{(0)} = -\mathbf{g}^{(0)}, \quad \mathbf{p}^{(k)} = -\mathbf{g}^{(k)} + \beta_{k-1} \mathbf{g}^{(k-1)}, \quad k \geq 1,$$

где

$$\beta_{k-1} = \frac{(A\mathbf{p}^{(k-1)}, \mathbf{p}^{(k)})}{(A\mathbf{p}^{(k-1)}, \mathbf{p}^{(k-1)})}.$$

Таким образом,

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \alpha_k \mathbf{p}^{(k)}, \quad k = 0, 1, 2, \dots, \quad (7)$$

$$\mathbf{p}^{(0)} = -\mathbf{g}^{(0)}, \quad \mathbf{p}^{(k)} = -\mathbf{g}^{(k)} + \beta_{k-1} \mathbf{g}^{(k-1)}, \quad k \geq 1, \quad (8)$$

$$\alpha_k = \min_{\alpha > 0} \varphi_k(\alpha), \quad \varphi_k(\alpha) = \mathbf{f}(\mathbf{x} + \alpha \mathbf{p}^{(k)}). \quad (9)$$

Здесь в случае неквадратичных функций

$$\beta_{k-1} = \frac{(\mathbf{g}^{(k)}, \mathbf{g}^{(k)} - \mathbf{g}^{(k-1)})}{|\mathbf{p}^{(k-1)}|^2}$$

или

$$\beta_{k-1} = \frac{|\mathbf{g}^{(k)}|^2}{|\mathbf{p}^{(k-1)}|^2}.$$

ЗАМЕЧАНИЕ:

1. Метод сопряженных градиентов сходится к решению задачи не более чем за  $n$  шагов, где  $n$  – размерность вектора  $\mathbf{x}$ .



2. Если функция неквадратичная, то метод уже не сходится за конечное число итераций. Часто, для уменьшения влияния вычислительной погрешности, при  $k = n, 2n, 3n, \dots$ , коэффициент  $\beta_{k-1}$  обнуляется. Такая процедура называется «обновлением алгоритма».

### Алгоритм минимизации квадратичной функции

Задать  $\mathbf{x}_0, \varepsilon$ ;

1.  $\mathbf{x} \leftarrow \mathbf{x}_0; \mathbf{g}_0 \leftarrow A\mathbf{x}_0 - \mathbf{b}$ ;

$\mathbf{p} \leftarrow -\mathbf{g}_0; k \leftarrow 0$ ;

REPEAT

2.  $\mathbf{x}_0 \leftarrow \mathbf{x}$ ;

3.  $\mathbf{g}_0 \leftarrow A\mathbf{x}_0 - \mathbf{b}$ ;

4.  $\mathbf{p}_0 \leftarrow \mathbf{p}$ ;

5. Выбрать направление  $\mathbf{p}$  поиска:

If  $k = 0$  Then

$\mathbf{p} \leftarrow \mathbf{p}_0$

Else Begin

$\beta \leftarrow \frac{(A\mathbf{p}_0, \mathbf{g}_0)}{(A\mathbf{p}_0, \mathbf{p}_0)}$ ;

$\mathbf{p} \leftarrow -\mathbf{g}_0 + \beta\mathbf{p}_0$

End;

6. Рассчитать шаг  $\alpha$ :

$\alpha \leftarrow -\frac{(\mathbf{g}_0, \mathbf{p}_0)}{(A\mathbf{p}_0, \mathbf{p}_0)}$ ;

7. Рассчитать  $\mathbf{x}$ :

$\mathbf{x} \leftarrow \mathbf{x}_0 + \alpha\mathbf{p}$ ;

8.  $k \leftarrow k + 1$ ;

UNTIL  $\|\mathbf{g}_0\| < \varepsilon$ ;

## Задания к лабораторным и практическим занятиям

1. Решить численно задачу минимизации функции

$$f(\mathbf{x}) = \frac{1}{2} \mathbf{x}^T A \mathbf{x} - \mathbf{b}^T \mathbf{x}$$

методами наискорейшего градиентного спуска и методом сопряженных градиентов. Решение задачи проиллюстрировать графически.

Здесь

$$A = \begin{bmatrix} N & -N \\ -N & N + 1 \end{bmatrix}$$
$$\mathbf{b} = \begin{bmatrix} 3N - 1 \\ 5N + 3 \end{bmatrix}.$$

Параметр  $N$  равен номеру студента в списке группы.

2. Решить аналитически задачу минимизации функции

$$f(\mathbf{x}) = \frac{1}{2} \mathbf{x}^T A \mathbf{x} - \mathbf{b}^T \mathbf{x}$$

методом сопряженных градиентов.

Здесь

$$A = \begin{bmatrix} N & -N \\ -N & N + 1 \end{bmatrix}$$
$$\mathbf{b} = \begin{bmatrix} 3N - 1 \\ 5N + 3 \end{bmatrix}.$$

Параметр  $N$  равен номеру студента в списке группы.

3. Выполнить 3 итерации методом сопряженных градиентов для

$$f(\mathbf{x}) = \frac{1}{2} \mathbf{x}^T A \mathbf{x} - \mathbf{b}^T \mathbf{x}.$$

Здесь

$$A = \begin{bmatrix} N & -N \\ -N & N + 1 \end{bmatrix}$$
$$\mathbf{b} = \begin{bmatrix} 3N - 1 \\ 5N + 3 \end{bmatrix}.$$

Параметр  $N$  равен номеру студента в списке группы.

4. Решить задачу минимизации функций

$$f(\mathbf{x}) = -x_1 x_2 e^{-(x_1+x_2)}, \mathbf{x}^{(0)} = (0, 1)^T;$$

$$f(\mathbf{x}) = (x_1^2 + x_2^2 - 1)^2 + (x_1 + x_2 - 1)^2, \mathbf{x}^{(0)} = (0, 3)^T.$$

методами наискорейшего градиентного спуска и Флетчера-Ривса.

### Библиографический список

1. *Амосов, А.А.* Вычислительные методы для инженеров: Учебное пособие / А.А. Амосов, Ю.А. Дубинский, Н.В. Копченова – М.: Высш. шк., 1994. – 544 с

2. *Аттетков, А. В.* Введение в методы оптимизации [Текст]: учебное пособие / А. В. Аттетков, В. С. Зарубин, А. Н. Канатников. - М. : Финансы и статистика, 2008. - 272 с.

3. *Гончаров, В. А.* Методы оптимизации [Текст] : учебное пособие / В. А. Гончаров. - М. : Юрайт, 2010. - 191 с.

4. *Пантелеев, А. В.* Методы оптимизации в примерах и задачах [Текст] : учебное пособие / А. В. Пантелеев, Т. А. Летова. - 2-е изд., испр. - М. : Высшая школа, 2005. - 544 с.