Санкт-Петербургский Политехнический университет
Петра Великого Институт прикладной математики и
механики
Кафедра «Прикладная математика»

Курсовая
работа по
дисциплине
"Методы оптимизации"

Тема: общий сравнительный анализ эффективности применения метода одномерной минимизации для выбора шага в методе наискорейшего спуска

Выполнили студенты гр.3630102/70301
Камянский Д.В.
Лебедев К.С.
Иванкин А.С.
Ли Жуйци

Руководитель

доцент, к.ф.-м.н. Родионова Е.А.

Санкт-Петербург 2020

1. Постановка задачи курсовой работы

Целью курсовой работы является оценка и сравнение методов Золотого сечения и Фибоначчи по их вычислительной эффективности для выбора шага в методе наискорейшего спуска.

Достижение указанной цели осуществлялось путем решения следующих основных задач:

1. Выбор функции, для которой выполняются условия применимости методов, и точки начального приближения.
2. Подсчет параметров: число итераций основного алгоритма; общее число обращений к функции, общее количество арифметических операций.
3. Анализ полученных результатов.

2. Описание методов одномерной минимизации

а) Постановка задачи одномерной минимизации

$$f: R \to R$$

$$\exists!\, x^* \in |a, b| \text{ доставляющая минимум } f(x)$$

$$\forall x, y \in |a, b| \quad \begin{cases} x^* < x \le y \Rightarrow f(x^*) \le f(x) \le f(y) \\ x^* > x \ge y \Rightarrow f(x^*) \le f(x) \le f(y) \end{cases}$$

$$\text{Найти } x^* = \operatorname{argmin}\{f(x), x \in |a, b|\} \text{ с точностью } \epsilon$$

Методы одномерной минимизации основаны на построении последовательности вложенных отрезков $[a_n, b_n]$, сходящихся к искомой точке x^*.

b) Метод золотого сечения

В данном методе строится последовательность вложенных отрезков, осуществляющих золотое сечение.

На каждом шаге алгоритма рассматривается три отрезка $[a_n, \alpha_n]$, $[\alpha_n, \beta_n]$, $[\beta_n, b_n]$ где $\alpha_n = b_n - \frac{b_n - a_n}{\varphi}$, $\beta_n = a_n + \frac{b_n - a_n}{\varphi}$ $\varphi = \frac{1+\sqrt{5}}{2}$

Исходя из ограничений, наложенных на целевую функцию, определяется следующее приближение $[a_{n+1}, b_{n+1}]$ и переопределяются α_{n+1}, β_{n+1}

$a_{n+1} = \alpha_n, \alpha_{n+1} = \beta_n, \beta_{n+1} = b_{n+1} - (\alpha_{n+1} - a_{n+1}), b_{n+1} = b_n$ if $f(\alpha_n) > f(\beta_n)$

$a_{n+1} = a_n, \alpha_{n+1} = a_{n+1} + (b_{n+1} - \beta_{n+1}), \beta_{n+1} = \alpha_n, b_{n+1} = \beta_n$ if $f(\alpha_n) \leq f(\beta_n)$

Можно выделить два важных свойства метода золотого сечения:

1: $|b_n - a_n| = \left(\frac{-1+\sqrt{5}}{2}\right)^n (b_1 - a_1)$

2: На каждом шаге алгоритма достаточно пересчитать значение $f(x)$ лишь в одной точке, α_n *или* β_n

c) Метод Фибоначчи

Метод Фибоначчи является оптимальным последовательным методом, т.е. методом, обеспечивающим максимальное гарантированное сокращение отрезка локализации при заданном числе N вычислений функции. Этот метод основан на использовании чисел Фибоначчи F_n, задаваемых рекуррентной формулой

$F_n = F_{n-1} + F_{n-2}$ $(n \geq 2)$

Метод Фибоначчи состоит из N-1 шагов. Очередной (k+1)-й шаг выполняют здесь аналогично (k+1)-й итерации метода деления отрезка пополам. В отличие от него точки α^{k}, β^{k} находят по формулам

$$\alpha^{k} = \alpha^{k} + \frac{F_{N-k-1}}{F_{N-k+1}}\Delta^{k} \qquad \beta^{k} = \alpha^{k} + \frac{F_{N-k}}{F_{N-k+1}}\Delta^{k}$$

Новый отрезок локализации определяют по тому же правилу:

Если $f(\alpha^{k}) \leq f(\beta^{k})$. то $[a^{k+1}, b^{k+1}] = [\alpha^{k}, \beta^{k}]$

Если $f(\alpha^{k}) > f(\beta^{k})$. то $[a^{k+1}, b^{k+1}] = [a^{k}, b^{k}]$ [i]

3. Метод наискорейшего спуска

$$f(x_{k} - \alpha_{k}\nabla f(x_k)) = \min_{\alpha \geq 0} f(x_k - \alpha\nabla f(x_k))$$

<u>Начальный этап</u>

Выберем $\boldsymbol{eps} > 0$ параметр. характеризующий условие окончание вычислений, $x_0 -$ начальное приближение. Положим k= 0 и перейдем к основному этапу

<u>Основной этап</u>

Шаг 1: вычисляем V/(т).

Шаг 2: определяем α_{k}, исходя из условия $\min\limits_{\alpha \geq 0} f(x_k - \alpha\nabla f(x_k))$.

Шаг 3: полагаем $x_{k+1} = x_k - \alpha_{k}\nabla f(x_k)$, заменяем К на К +1 и переходим к шагу 1.

<u>Условие окончания вычислений</u> $\|\nabla f\|^2 < \boldsymbol{eps}$

<u>Теорема 1</u> Для ограниченной снизу функции $\boldsymbol{f}(x) \in C^1$,градиент $\nabla f(x)$ которой удовлетворяет условию Липшица с постоянной L:

$\|\nabla f(x) - \nabla f(y)\| \leq L\|x - y\|$ при любых $x, y \in R^{n}$. значение параметра

α_k градиентном методе выбираются по формуле $f(x) - f(x_k) \leq \varepsilon\alpha\left|\left|\nabla f(x_k)\right|\right|^2$; последовательность $(x\}$. определяемая процессом ($x_{k+1} = x_k - \alpha_k \nabla f(x_k)$ x_0 – дано $k = 0,1,2, \ldots$), такова, что неравенство ($\left|\left|\nabla f(x_k)\right|\right|^2 <$ eps) при $k \to \infty$ выполнено при любом малом положительном eps для любого начального приближения x_0

<u>Теорема 2</u> Если Функция f(x) удовлетворяет Теореме 1 при условии определения параметра α_k по формуле ($f\left(x_k - \alpha_k \nabla f(x_k)\right) = \min_{\alpha \geq 0} f(x_k - \alpha \nabla f(x_k))$ в оптимальном градиентном методе, то $\left|\left|\nabla f(x_k)\right|\right| \to 0$ при $k \to \infty$ при любом начальном x_0

<u>Теорема 3</u> Пусть функция $f(x) \in C^2$, причем квадратичная форма $y^T H(x) y$. связанная с матрицей Гессе вторых производных $H(x)$, такова, что при любых т $f(x) \in R^n$

$$m\left|\left|y\right|\right|^2 \leq y^T H(x) y \leq M||y||^2 \; 0 < m < M$$

Пусть $\{x_k\}$ последовательность, определяемая процессом $x_{k+1} = x_k - \alpha_k \nabla f(x_k)$,Тогда для любой начальной точки то последовательность $\{x_k\}$ при$k \to \infty$ сходится к единственной точке минимума x_0, со скоростью, определяемое следующим образом:

$f_k - f_* \leq q^k (f_0 - f_*)$, $\left|\left|x_k - x_*\right|\right| \leq C q^{\frac{k}{2}}$, $0 < q < 1$ [ii]

В данной работе будем рассматривать функцию:

$$f(x_1, x_2) = 2x_1^2 + 3x_2^2 + \frac{\sin(2x_1 + 7x_2)}{49} + 3x_1 + 2x_2$$

с начальным приближением $x_0 = \{-100; 100\}$

Вывод оценок для теоремы 3:

$$\mathrm{y}^{\mathrm{T}}H(x)y = 4y_1^2 + 6y_2^2 - \frac{(2y_1 + 7y_2)^2}{49} * \sin(2x_1 + 7x_2)$$

$$\mathrm{y}^{\mathrm{T}}H(x)y \geq 4y_1^2 + 6y_2^2 - \frac{(2y_1 + 7y_2)^2}{49} \geq 4y_1^2 + 6y_2^2 - \frac{(2|y_1| + 7|y_2|)^2}{49}$$

$=4y_1^2 + 6y_2^2 - \frac{4}{49}y_1^2 - y_2^2 - \frac{4}{7}|y_1 - y_2| \geq 3y_1^2 + 5y_2^2 - \frac{4}{7}|y_1 y_2| \geq 3y_1^2 +$
$5y_2^2 - \frac{2}{7}y_1^2 - \frac{2}{7}y_2^2 \geq 2y_1^2 + 4y_2^2 \geq 2 * \left||y|\right|^2, \mathbf{m = 2}$

$\mathrm{y}^{\mathrm{T}}H(x)y \leq 4y_1^2 + 6y_2^2 \leq 6\left||y|\right|^2, \mathbf{M = 6}$

$$2\left||y|\right|^2 \leq \mathrm{y}^{\mathrm{T}}H(x)y \leq 6\left||y|\right|^2$$

4. Результаты и выводы

Метод Золотого сечения

ε	i_n	i_f	i_m
1e-1	4	48	140
1e-3	6	72	210
1e-5	7	84	245
1e-7	8	96	280
1e-10	10	120	350

Метод Фибоначчи

ε	i_n	i_f	i_m
1e-1	4	60	236
1e-3	6	75	295

1e-5	6	90	354
1e-7	7	105	413
1e-10	10	120	472

i_n – число итераций основного алгоритма

i_f –общее число обращений к функции

i_m – общее количество арифметических операций

Вывод: из полученных результатов следует, что для данной задачи МЗС показывает себя эффективнее МФ: i_f в МЗС не превышает i_f в МФ, а для невысоких € значительно ниже; i_m в МЗС для всех рассмотренных € значительно ниже, чем в МФ. Но для высоких точностей i, ниже в МФ по сравнению с МЗС, что делает его предпочтительнее при высокой сложности вычислений в основном алгоритме (напр., вычисление градиента).

5. Список литературы

1) Курс лекций по предмету Методы оптимизации, Родионова Е.А. 2020.
2) Петухов Л.В., Серёгин Г.А., Родионова Е.А. Методы оптимизации. Задачи выпуклого программирования: Учеб.пособие - СПб. : Изд-во Политехнического ун-та, 2014.
3) Лыткина Л.И., Сафонов К.В., Хоролич Г.Б. Методы оптимизации и вариационное исчисление: Учеб. пособие. Сиб. гос. аэрокосмич. ун-т. – Красноярск, 2012. – 116 с.
4) Амосов А.А Вычислительные методы для инженеров: Учеб.пособие / Амосов А.А., Дубинский Ю.А., Копченова Н.В.М.: Высш. шк.,1994. - 544 с.:ил.

[i] Амосов А.А Вычислительные методы для инженеров §9.3
[ii] Все теоремы и алгоритмы взяты из лекций по предмету Методы оптимизации Родионовой Е.А.