Московский авиационный институт
(национальный исследовательский университет)

Факультет информационных технологий и прикладной математики

Кафедра вычислительной математики и программирования

Лабораторная работа №1 по курсу «Методы оптимизации»

Студент:	Л. Я. Вельтман
Преподаватель:	Т. И. Короткова
Группа:	М8О-307Б
Дата:	
Оценка:	
Подпись:	

Москва, 2020

1 Метод конфигураций

- **Класс задач:** безусловная оптимизация
- **Формулировка задачи:** Найти минимум функции 2-ух переменных $f(x1, x2)$ методом кофигурации.
- **Что вычисляется в процессе решения:** В процессе итерации находятся точки базиса, в которых происходит исследующий поиск и поиск по образцу.
- **Алгоритм:** *Описание алгоритма:*

  1. Задается начальная точка $x^0 = (x_1^0, x_2^0, ..., x_n^0)$ и начальные значения приращений $dx_1^0, dx_2^0, ..., dx_n^0$, а также минимальная длина шага ξ для останова и необязательный параметр останова - максимальное количество иттераций N_{max}. Для корректной работы алгоритма $dx_i^0 > \xi$ для $\forall i \in \{1, ..., n\}$. Точка x^0 называется точкой старого базиса.
  2. Проводится исследующий поиск, в результате которого каждого координата новой точки x^{k+1} вычисляется по алгоритму:
     (a) Для $\forall i \in \{1, ..., n\}$:

$$x_i^{k+1} = \begin{cases} x_i^k + dx_i^k, if \quad f(x_1^k, ..., x_i^k + dx_i^k, ..., x_n^k) < f(x_1^k, ..., x_i^k, ..., x_n^j) \\ x_i^k - dx_i^k, if \quad f(x_1^k, ..., x_i^k - dx_i^k, ..., x_n^k) < min(f(x_1^k, ..., x_i^k, ..., x_n^k), f(x_1^k, ..., x_i^k + dx_i^k, ..., x_n^k)) \\ x_i^k, otherwise \end{cases}$$

     В результате исследующего поиска получается точка x^{k+1}.
     Если при этом $x^{k+1} \neq x^k$, то x^{k+1} - точка нового базиса.
     Если $x^{k+1} = x^k$, то исследующий поиск неудачен. В этом случае необходимо уменьшить значения приращений $dx_1^j, dx_2^j, ..., dx_n^j$ и повторить исследующий поиск.

  3. Из точки нового базиса может быть:
     (a) продолжен исследующий поиск со старыми или новыми значения приращений (шаг 2) алгоритма).
     (b) проведен поиск по образцу по алгоритму: $x^{(obr)} = x^k + t_k(x^k - x^{k-1})$, , где t_k - параметр движения(зависит от реализаций, обычно $t_k = 2$).

     В точке $x^{(obr)}$ значение функции не вычисляется, из этой точки проводится исследующий поиск, в результате которого получается точка $x^{(ip)}$.

     Если $x^{(ip)} \neq x^{(obr)}$, то точка $x^{k+1} = x^{(ip)}$ становится точкой нового базиса, а x^k - точкой старого базиса. Если $x^{(ip)} = x^{(obr)}$, то поиск по образцу считается неудачным, точки $x^{(ip)}, x^{(obr)}$ - аннулируются, при этом точка x^k остается точкой нового базиса, а x^{k-1} - точкой старого базиса.

4. Процедура 3) повторяется до выполнения критерия окончания счета.

   Основной критерий окончания метода: $dx^k \leqslant \xi$, $dy^k \leqslant \xi$

   Изменяемый параметр метода: величины приращений dx^k, dy^k.

- **1-ая итерация:**
  $f(x_1, x_2) = 5x_1^2 + 3x_1x_2 + 6x_2^2 + 2x_1 + x_2 + 9$
  $\xi = 0.01$
  $N_{max} = 8$
  $j = 0$
  $x_1^0 = 11$, $x_2^0 = 7$
  $dx_1^0 = 1$, $dx_2^0 = 1$

  Делаем исследующий поиск:
  $dx_1^0 = 1 > 0.01 = \xi$
  $dx_2^0 = 1 > 0.01 = \xi$
  $j < N_{max}$, значит критерий останова не выполнен.

  $f(x_1^j, x_2^j) = 5 \cdot 121 + 3 \cdot 77 + 6 \cdot 49 + 22 + 7 + 9 = 1168$

  1. Для x_1:
     $f(x_1^j + 1, x_2^j) = 5 \cdot 144 + 3 \cdot 84 + 6 \cdot 49 + 24 + 7 + 9 = 1306 > f(x_1^j, x_2^j)$
     $f(x_1^j - 1, x_2^j) = 5 \cdot 100 + 3 \cdot 70 + 6 \cdot 49 + 20 + 7 + 9 = 1040 < f(x_1^j, x_2^j)$
     Значит $x_1^{j+1} = x_1^j - 1 = 11 - 1 = 10$
  2. Для x_2:
     $f(x_1^j, x_2^j + 1) = 5 \cdot 121 + 3 \cdot 88 + 6 \cdot 64 + 22 + 8 + 9 = 1292 > f(x_1^j, x_2^j)$
     $f(x_1^j, x_2^j - 1) = 5 \cdot 121 + 3 \cdot 66 + 6 \cdot 36 + 22 + 6 + 9 = 1056 < f(x_1^j, x_2^j)$
     Значит, $x_2^{j+1} = x_2^j - 1 = 7 - 1 = 6$

  $dx_1^{j+1} = dx_1^j = 1, dx_2^{j+1} = dx_2^j = 1$
  $j = j + 1 = 1$
  $x^1 = (10, \quad 6)$
  Конец первой итерации.

- **Результат компьютерных вычислений:**

Протокол расчета

Выполнил: Вельтман, группа 80-307, 10.03.2020

Квадратичная функция: $f(x_1,x_2)=5x_1^2+ 3x_1x_2+ 6x_2^2+ 2x_1+ 1x_2+ 9$

Метод конфигураций

Точность метода: 0.01, N_{max} = 8, Количество итераций: 9

$N_{ит}$	x_1	x_2	$f(x_1,x_2)$	dx_1	dx_2	коэф-т k
0	11	7	1168	1	1	
1	10	6	931	1	1	0
2	9	5	722	1	1	0
3	8	4	541	1	1	0
4	7	3	388	1	1	0
5	6	2	263	1	1	0
6	5	1	166	1	1	0
7	4	0	97	1	1	0
8	3	-1	56	1	1	0
9	2	-1	32			

Критерий окончания не выполнен

$||x - x^*||$ = 2.39202

$|f(x) - f(x^*)|$ = 23.20721

2 Метод Марквардта

- **Класс задач:** безусловная оптимизация
- **Формулировка задачи:** Найти минимум функции 2-ух переменных $f(x_1, x_2)$ методом Марквардта.
- **Что вычисляется в процессе решения:** В процессе иттерации для поиска точки x^{k+1} с меньшим значением функции $f(x)$ вычислятся шаг-приращение dx^k для точки x^k при помощи градиента и матрицы Гёссе в этой точке.
- **Алгоритм:** Перед началом описания алгоритма следует ввести следующие обозначения:
Матрица Гёссе:

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

Градиент:

$$\nabla f(x) = \begin{pmatrix} \frac{\partial f(x)}{\partial x_1} \\ \frac{\partial f(x)}{\partial x_2} \\ \vdots \\ \frac{\partial f(x)}{\partial x_n} \end{pmatrix}$$

Функции $f(x)$, где $x = (x_1, x_2, ..., x_n)$.

Описание алгоритма:

1. *Начальная инициализация:* Задать точку для начала движения $x^0 = (x_1^0, x_2^0, ..., x_n^0)$, точность приближения ξ, необязательный парамметр - максимальное количество иттераций N_{max}, а также выбрать значение параметра λ_0, которое, впрочем, можно задать как $\lambda_0 = 50 \cdot max(h_{ij}^0)_{0 \le i,j \le n}$, где h_{ij}^k - элемент матрицы $H_f(x^k)$ на i-ой строке и j-ом столбце. Начиннаем первую иттерацию при $k = 0$.
2. *Проверим критерий останова:* Если выполнено основное условие $\|\nabla f(x^k)\| < \xi$ или дополнительное $k = N_{max}$, где k назовем номером иттерации, выходим из алгоритма с результатом x^k.

3. *Вычисление новой точки:* Вычислим приращение $dx^k = -[H_f(x^k)+\lambda_k E]^{-1}\nabla f(x^k)$ и новую точку $x^{k+1} = x^k + dx^k$.
   - Если $f(x^{k+1}) < f(x^k)$, значит мы успешно нашли новую точку и можем уменьшить значение параметра λ_{k+1}, например $\lambda_{k+1} = \dfrac{\lambda_k}{2}$. Положим $k = k+1$ и перейдем к шагу (2).
   - Если $f(x^{k+1}) \geqslant f(x^k)$, значит поиск неудачен и следует увеличить значение параметра λ_k, например $\lambda_k = \lambda_k \cdot 2$. Параметр k не меняется и повторим шаг (3) для поиска x^{k+1} с новым параметром λ_k.

- **1-ая итерация:**
  $f(x_1, x_2) = 5x_1^2 + 3x_1x_2 + 6x_2^2 + 2x_1 + x_2 + 9$
  $\xi = 0.01$
  $N_{max} = 5$ $k = 0$
  $x_1^0 = 11$, $x_2^0 = 7$

$$\nabla f(x) = \begin{pmatrix} 10x_1 + 3x_2 + 2 \\ 12x_2 + 3x_1 + 1 \end{pmatrix}$$

$$H_f(x) = \begin{pmatrix} 10 & 3 \\ 3 & 12 \end{pmatrix}$$

Матрица Гёссе положительно определена, поэтому для ускорения сходмости возьмем небольшой параметр $\lambda_0 = 1$, хотя в общем случае рекоменуется брать значение на порядок больше, чем элементы в матрице Гёссе.

Проверим критерий останова:

$$\|\nabla f(x^k)\| = \left\| \begin{pmatrix} 133 \\ 118 \end{pmatrix} \right\| = \sqrt{133^2 + 119^2} = 177.8004 > 0.01 = \xi$$

$k = 0 < 5 = N_{max}$, значит критерий останова не выполнен.

Вычислим следующую точку:

$$dx^k = -[H_f(x^k)+\lambda_k E]^{-1}\nabla f(x^k) = \left[\begin{pmatrix} 10 & 3 \\ 3 & 12 \end{pmatrix} + \begin{pmatrix} \lambda_k & 0 \\ 0 & \lambda_k \end{pmatrix} \right]^{-1} \begin{pmatrix} 133 \\ 118 \end{pmatrix} = (-10.261, -6.709)$$

$x^{k+1} = x^k + dx^k = (11, 7) + (-10.261, -6.709) = (0.739, 0.291)$
$f(x^{k+1}) = 12.2598 < 1168 = f(x^k)$, значит новая точка найдена успешно.
$\lambda_{k+1} = \dfrac{\lambda_k}{2} = 0.5$
$k = k + 1$
Конец первой итерации.

- **Результат компьютерных вычислений:**

Расчет окончен

Сохранить протокол | Выбрать другой метод | Выход

Протокол расчета

Выполнил: Вельман, группа 80-307, 10.03.2020

Квадратичная функция: $f(x_1,x_2)=5x_1^2+ 3x_1x_2+ 6x_2^2+ 2x_1+ 1x_2+ 9$

Метод Марквардта

Точность метода: 0.01, N_{max} = 5, Количество итераций: 5

$N_{ит}$	шаг μ	x_1	x_2	$f(x_1,x_2)$	f'_{x_1}	f'_{x_2}	$\|\|\nabla f(x_1,x_2)\|\|$
0	0	11	7	1168	133	118	177.80045
1	0	0.90493	0.35494	16.97874	12.11409	7.97407	14.503
2	0	-0.0745	-0.02656	8.86236	1.17531	0.4578	1.26133
3	0	-0.17635	-0.03809	8.79356	0.12222	0.01384	0.123
4	0	-0.18767	-0.03657	8.7928	0.01358	-0.00183	0.01371
5	0	-0.189	-0.03613	8.79279	0.0016	-0.00053	0.00168

Критерий окончания выполнен

$\|\|x - x^*\|\| = 0.00021$

$\|f(x) - f(x^*)\| = 0$

3 Метод Нелдера-Мида

- **Класс задач:** безусловная оптимизация
- **Формулировка задачи:** Найти минимум функции 2-ух переменных $f(x_1, x_2)$ методом Нелдера-Мида.
- **Что вычисляется в процессе решения:** В процессе итераций вычисляется центр тяжести набора точек, с целью заменить «худшую» точку из набора на более близкую к минимуму функции.
- **Алгоритм:** Минимизируем функцию $f(x)$, где $x = (x_1, x_2, ..., x_n)$
  Описание алгоритма:

  1. *Начальная инициализация:* Задать начальную систему из $n+1$ точек (многогранник): $\{x^{0(1)}, x^{0(2)}, ..., x^{0(n+1)}\}$, точность ξ, максимальное количество иттераций N_{max}. Начинаем первую итерацию при $k = 0$.
  2. Вычисляется значение функции во всех точках многогранника и выбирается:
     лучшая точка $x^{(l)} : f(x^{(l)}) = \min\limits_i [f(x^{k(i)})]$
     худшая точка $x^{(x)} : f(x^{(x)}) = \max\limits_i [f(x^{k(i)})]$
     Далее заданная система точек перестраивается, для этого:
  3. *Найдем центр тяжести:* Найдем центр тяжести $x^{(c)} = \dfrac{(\sum\limits_{i=1}^{n+1} x^{k(i)} - x^x)}{n}$
     (для функции 2-х переменных точка $x^{(c)}$ - середина отрезка, соединяющего точки за исключением худшей)
  4. *Отражение точки:* Получим точку отражения $x^{(otr)} = (1 + \alpha)x^{(c)} - \alpha x^x$.

     здесь $\alpha > 0$ - параметр отражения (рекомендуемое значение $\alpha = 1$)
  5. Формируется новая система точек (многогранник). Для этого в точке $x^{(otr)}$ вычисляется значение функции, полученное значение сравнивается с $f(x^{(l)})$:
     - Если $f(x^{(otr)}) < f(x^{(l)})$, значит выполняется операция <u>растяжение</u>:

       $x^{(rst)} = x^{(c)} + \gamma(x^{(otr)} - x^{(c)})$

       здесь $\gamma > 0$ ($\gamma \neq 0$) - параметр растяжения (рекомендованное значение $\gamma \in [2, 3]$) При этом, если $f(x^{(rst)}) < f(x^{(otr)})$, то в новой системе точек точка $x^{(x)}$ будет заменена на $x^{(rst)}$, если же $f(x^{(rst)}) \geqslant f(x^{(otr)})$, то в новой системе точек точка $x^{(x)}$ будет заменена на $x^{(otr)}$.

- Если $f(x^{(l)}) \leqslant f(x^{(otr)}) < f(x^{(x)})$ выполняется операция сжатие:

  $x^{(szh)} = x^{(c)} + \beta(x^{(x)} - x^{(c)})$

  здесь $\beta > 0$ $(\beta \neq 0)$ - параметр сжатия (рекомендованное значение $\beta \in [0.4, 0.6]$).

  При этом, если $f(x^{(szh)}) < f(x^{(otr)})$, то в новой системе точек точка $x^{(x)}$ будет заменена на $x^{(szh)}$, если $f(x^{(szh)}) \geqslant f(x^{(otr)})$, то в новой системе точек точка $x^{(x)}$ будет заменена на $x^{(otr)}$.
- Если $f(x^{(otr)}) \geqslant f(x^{(x)})$ выполняется операция редукции:

  при этом формируется новый многогранник, содержащий точку $x^{(l)}$ с уменьшенными вдвое сторонами:

  $x^{k+1(i)} = x^{(l)} + 0.5(x^{k(i)} - x^{(l)}), i = 1, ..., n+1$

  Таким образом, в результате выполнения этого пункта алгоритма формируется новая система точек(многогранник), причем в случае возникновения операций растяжения и сжатия перестраивается только одна точка - $x^{(x)}$, в случае возникновения операции редукции - все точки, за исключением $x^{(l)}$.

6. Процедура 2) - 5) повторяется до выполнения критерия окончания счета

- Критерий останова: Обозначим $\overline{f} = \dfrac{\sum_{j=1}^{n+1} f(x^{k(j)})}{n+1}$.
  Если $\sqrt{\dfrac{\sum_{i=1}^{n+1} \left| f(x^{k(i)}) - \overline{f} \right|^2}{n+1}} < \xi$ или $k = N_{max}$, то выходим из алгоритма с значением x_l, иначе $k = k+1$.

- **1-ая итерация:**
  $f(x_1, x_2) = 5x_1^2 + 3x_1x_2 + 6x_2^2 + 2x_1 + x_2 + 9$
  $\xi = 0.01$
  $N_{max} = 5$ $k = 0$
  $x_1^0 = 11$, $x_2^0 = 7$
  $\alpha = 1$, $\gamma = 2.8$, $\beta = 0.5$
  $x^{0(1)} = (-3, 3)$, $x^{0(2)} = (-3, -2)$, $x^{0(3)} = (11, 7)$

Выберем 3 точки:
$x_l = (-3, 3)$, $x_h = (11, 7)$, $x_g = (-3, -2)$

Найдем центр тяжести:
$x_c = \dfrac{x_l + x_g}{2} = (\dfrac{-3-3}{2}, \dfrac{3-2}{2}) = (-3, 0.5)$

Отражение точки:
$x_r = (1+\alpha)x_c - \alpha x_h = (-6, 1) - (11, 7) = (-17, -6)$
$f(x_r) = 1936$
$f(x_r) > 1168 = f(x_h)$, значит делаем сжатие.

Сжатие точки:
$x_s = \beta x_h + (1-\beta)x_c = (5.5, 3.5) + (-1.5, 0.25) = (4, 3.75)$
$f(x_s) = 230.125$
$f(x_s) < 1168 = f(x_h)$, значит заменяем точку $x^{k(3)} = x_h$ в наборе на точку $x_s = (4, 3.75)$.

Проверим критерий окончания:
$f(x^{k(1)}) = 78$, $f(x^{k(2)}) = 88$, $f(x^{k(3)}) = 230.125$
$\overline{f} = \dfrac{78 + 88 + 230.125}{3} = 132.041667$

$$\sqrt{\frac{\sum\limits_{i=1}^{3} \left| f(x^{k(i)}) - \overline{f} \right|^2}{n+1}} = 69.4754403 > 0.01 = \xi$$
$k = 0 < 5 = N_{max}$
$k = k + 1$

Конец первой итерации.

- **Результат компьютерных вычислений:**

Протокол расчета

Выполнил: Вельтман, группа 80-307, 10.03.2020

Квадратичная функция: $f(x_1,x_2)=5x_1^2+3x_1x_2+6x_2^2+2x_1+1x_2+9$

Метод Нелдера-Мида

Точность метода: 0.01, $N_{max} = 5$, Количество итераций: 6

$N_{ит}$	α	операция	коэффициент	x_1	x_2	$f(x_1,x_2)$
0	1	редукция		-3 -3 11	3 -2 7	78 88 1168
1	1	редукция		-3 -3 4	0.5 3 5	45.5 78 312
2	1	редукция		-3 -3 0.5	0.5 1.75 2.75	45.5 52.375 63.5
3	1	растяжение	2.8	-1.25 -3 -3	1.625 0.5 1.125	25.6875 45.5 46.59375
4	1	редукция		0.325 -1.25 -3	0.8875 1.625 0.5	16.65688 25.6875 45.5
5	1	растяжение	2.8	-1.3375 0.325 -0.4625	0.69375 0.8875 1.25625	16.06734 16.65688 18.12672
6				-0.55 -1.3375 0.325	0.325 0.69375 0.8875	9.835 16.06734 16.65688

Критерий окончания не выполнен

$\|x - x^*\| = 0.51042$

$|f(x) - f(x^*)| = 1.04221$

4 Метод сопряженных градиентов

(Для квадратичных функций метод сопряженных градиентов называется методом Флетчера-Ривса).

- **Класс задач:** безусловная оптимизация
- **Формулировка задачи:** Найти минимум функции 2-ух переменных $f(x_1, x_2)$ методом сопряженных градиентов.
- **Что вычисляется в процессе решения:** В процессе иттерации для поиска точки x^{k+1} с меньшим значением функции $f(x)$ вычислятся длина шага t_k и его направление dx^k для точки x^k при помощи градиента, направления dx^{k-1} с предыдущей иттерации и метода дихотомии.
- **Алгоритм:** Перед началом описания алгоритма следует ввести следующие обозначения:

  Градиент:

$$\nabla f(x) = \begin{pmatrix} \frac{\partial f(x)}{\partial x_1} \\ \frac{\partial f(x)}{\partial x_2} \\ \vdots \\ \frac{\partial f(x)}{\partial x_n} \end{pmatrix}$$

  Функции $f(x)$, где $x = (x_1, x_2, ..., x_n)$.

  Описание алгоритма:

  1. *Начальная инициализация:*
     Задать точку для начала движения $x^0 = (x_1^0, x_2^0, ..., x_n^0)$, точность приближения ξ, необязательный параммeтр - максимальное количество иттераций N_{max}, а также выбрать значение параметра точности поиска ϵ и отрезка $[a, b]$ для расчета длин шагов. Начинаем первую иттерацию при $k = 0$ и $dx^0 = -\nabla f(x^0)$.
  2. *Длина шага:*
     Найдем длину шага $t_k = \underset{t \in [a,b]}{\operatorname{argmin}} |f(x^k + t \cdot dx^k)|$, поиск значения которой производится методом дихотомии на заданном отрезке $[a, b]$ с точностью ϵ.
  3. *Вычисление следующей точки:*
     Найдем новую точку $x^{k+1} = x^k + t_k \cdot dx^k$ и инкрементируем счетчик $k = k + 1$.

4. *Основной критерий окончания метода:*
   Если выполнено основное условие $\|\nabla f(x^k)\| < \xi$ или дополнительное $k = N_{max}$, где k - номер иттерации, выходим из алгоритма с результатом x^k.

5. *Направление поиска:*
   Найдем новое направление движения
   $dx^k = -\nabla f(x^k) + \beta_{k-1} dx^{k-1}$, где коеффицент $\beta_{k-1} = \dfrac{\|\nabla f(x^k)\|^2}{\|\nabla f(x^{k-1})\|^2}$ и перейдем к шагу (2).

- **1-ая итерация:**
  $f(x_1, x_2) = 5x_1^2 + 3x_1x_2 + 6x_2^2 + 2x_1 + x_2 + 9$
  $\xi = 0.01$
  $\epsilon = 0.001$
  $[a, b] = [0.05, 0.5]$
  $N_{max} = 5$
  $k = 0$
  $x_1^0 = 11$, $x_2^0 = 7$

  Градиент функции расчитывается по формуле:

$$\nabla f(x) = \begin{pmatrix} 10x_1 + 3x_2 + 2 \\ 12x_2 + 3x_1 + 1 \end{pmatrix}$$

  Найдем направление движения:
  Определим начальное направление движения $dx^0 = -\nabla f(x^k) = (-133, -118)$.

  Найдем длину шага:
  Методом дихотомии с точностью ϵ на отрезке $[a, b]$ определили оптимальное значение как $t_k = t_0 = 0.0722$.

  Вычислим следующую точку:
  Найдем новое значение точки:
  $x^{k+1} = x^k + t_k \cdot dx^k = (11, 7) + 0.0722 \cdot (-133, -118) = (1.4021, -1.5155)$
  И увеличим значение $k = k + 1$

  Проверим критерий останова:

$$\|\nabla f(x^k)\| = \left\| \begin{pmatrix} 11.4743 \\ -12.9793 \end{pmatrix} \right\| = \sqrt{(11.4743)^2 + (-12.9793)^2} = 17.3240 > 0.01 = \xi$$

$k = 1 < 5 = N_{max}$, значит критерий останова не выполнен.

Конец первой итерации.

- **Результат компьютерных вычислений:**

Протокол расчета

Выполнил: Вельтман, группа 80-307, 10.03.2020

Квадратичная функция: $f(x_1,x_2)=5x_1^2+ 3x_1x_2+ 6x_2^2+ 2x_1+ 1x_2+ 9$

Метод сопряженных градиентов

Точность метода: 0.01, N_{max} = 2, Количество итераций: 3

$N_{ит}$	шаг t	x_1	x_2	$f(x_1,x_2)$	f'_{x_1}	f'_{x_2}	$\|\|\nabla f(x_1,x_2)\|\|$
0	0.07216	11	7	1168	133	118	177.80045
1	0.12484	1.40207	-1.51546	27.52301	11.47428	-12.97932	17.32403
2	0.0709	-0.18807	-0.03492	8.79281	0.01456	0.01674	0.02218
3	0	-0.1891	-0.03611	8.79279	0.00066	-0.00059	0.00089

Критерий окончания не выполнен

$||x - x^*|| = 0.00011$

$|f(x) - f(x^*)| = 0$

5 Метод Ньютона-Рафсона

- **Класс задач:** безусловная оптимизация
- **Формулировка задачи:** Найти минимум функции 2-ух переменных $f(x_1, x_2)$ методом Ньютона-Рафсона.
- **Что вычисляется в процессе решения:** В процессе иттерации для поиска точки x^{k+1} с меньшим значением функции $f(x)$ вычислятся длина t_k и направление dx^k шага для точки x^k при помощи градиента и матрицы Гёссе в этой точке.
- **Алгоритм:** Перед началом описания алгоритма следует ввести следующие обозначения:
  Матрица Гёссе:
  $$H_f(x) = \begin{pmatrix} \frac{\partial^2 f(x)}{\partial x_1^2} & \frac{\partial^2 f(x)}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f(x)}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f(x)}{\partial x_2 \partial x_1} & \frac{\partial^2 f(x)}{\partial x_2^2} & \cdots & \frac{\partial^2 f(x)}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f(x)}{\partial x_n \partial x_1} & \frac{\partial^2 f(x)}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f(x)}{\partial x_n^2} \end{pmatrix}$$
  Градиент:
  $$\nabla f(x) = \begin{pmatrix} \frac{\partial f(x)}{\partial x_1} \\ \frac{\partial f(x)}{\partial x_2} \\ \vdots \\ \frac{\partial f(x)}{\partial x_n} \end{pmatrix}$$
  Функции $f(x)$, где $x = (x_1, x_2, ..., x_n)$.

  Описание алгоритма:

  1. *Начальная инициализация:*
     Задать точку для начала движения $x^0 = (x_1^0, x_2^0, ..., x_n^0)$, точность приближения ξ, необязательный параметр - максимальное количество иттераций N_{max}, а также выбрать значение параметра точности поиска ϵ и отрезка $[a, b]$ для расчета длин шагов. Начинаем первую иттерацию при $k = 0$.
  2. *Критерий останова:*
     Если выполнено основное условие $\|\nabla f(x^k)\| < \xi$ или дополнительное $k = N_{max}$, где k назовем номером иттерации, выходим из алгоритма с результатом x^k.

3. *Направление поиска:*
   Найдем новое направление спуска
   $dx^k = -H_f^{-1}(x^k)\nabla f(x^k)$.
4. *Длина шага:*
   Найдем длину шага $t_k = \underset{t\in[a,b]}{\operatorname{argmin}}|f(x^k + t \cdot dx^k)|$, поиск значения которой производится методом дихотомии на заданном отрезке $[a, b]$ с точностью ϵ.
5. *Вычисление следующей точки:*
   Найдем новую точку $x^{k+1} = x^k + t_k \cdot dx^k$, инкрементируем счетчик $k = k+1$ и перейдем к шагу (2).

- **1-ая итерация:**
  $f(x_1, x_2) = 5x_1^2 + 3x_1x_2 + 6x_2^2 + 2x_1 + x_2 + 9$
  $\xi = 0.01$
  $\epsilon = 0.001$
  $[a, b] = [0.05, 0.5]$
  $N_{max} = 5$
  $k = 0$
  $x_1^0 = 11$, $x_2^0 = 7$

  Градиент:

  $$\nabla f(x) = \begin{pmatrix} 10x_1 + 3x_2 + 2 \\ 12x_2 + 3x_1 + 1 \end{pmatrix}$$

  Матрица Гёссе:

  $$H_f(x) = \begin{pmatrix} 10 & 3 \\ 3 & 12 \end{pmatrix}$$

  Проверим критерий останова:

  $$\|\nabla f(x^k)\| = \left\| \begin{pmatrix} 133 \\ 118 \end{pmatrix} \right\| = \sqrt{133^2 + 119^2} = 177.8004 > 0.01 = \xi$$

  $k = 0 < 5 = N_{max}$, значит критерий останова не выполнен.

  Найдем направление движения:

  $$dx^k = -H_f^{-1}(x^k)\nabla f(x^k) = \begin{pmatrix} 10 & 3 \\ 3 & 12 \end{pmatrix}^{-1} \begin{pmatrix} 133 \\ 118 \end{pmatrix} = (-11.1913, -7.0408)$$

Найдем длину шага:
Методом дихотомии определили оптимальное значение как $t_k = t_0 = 0.5$.

Вычислим следующую точку:
$x^{k+1} = x^k + t_k \cdot dx^k = (11, 7) + 0.5 \cdot (-11.1913, -7.0408) = (5.4054, 3.4820)$
$f(x^{k+1}) = 298.5946 < 1168 = f(x^k)$, значит новая точка найдена успешно.
$k = k + 1$

Конец первой итерации.

- **Результат компьютерных вычислений:**

Протокол расчета

Выполнил: Вельман, группа 80-307, 10.03.2020

Квадратичная функция: $f(x_1,x_2)=5x_1^2+ 3x_1x_2+ 6x_2^2+ 2x_1+ 1x_2+ 9$

Метод Ньютона-Рафсона

Точность метода: 0.01, N_{max} = 5, Количество итераций: 6

$N_{ит}$	шаг t	x_1	x_2	$f(x_1,x_2)$	f'_{x_1}	f'_{x_2}	$\|\|\nabla f(x_1,x_2)\|\|$
0	0.5	11	7	1168	133	118	177.80045
1	0.5	5.40541	3.48198	298.59459	66.5	59	88.90022
2	0.5	2.60811	1.72297	81.24324	33.25	29.5	44.45011
3	0.5	1.20946	0.84347	26.90541	16.625	14.75	22.22506
4	0.5	0.51014	0.40372	13.32095	8.3125	7.375	11.11253
5	0.5	0.16047	0.18384	9.92483	4.15625	3.6875	5.55626
6	0	-0.01436	0.0739	9.0758	2.07813	1.84375	2.77813

Критерий окончания не выполнен

$||x - x^*|| = 0.20652$

$|f(x) - f(x^*)| = 0.28301$

6 Выводы

Подводя итог, я узнала новые алгоритмы безусловной оптимизации, выполнив лабораторную работу по курсу «Методы оптимизации». Эти методы позволяют находить минимумы функций в процессе иттерационных приближений. Также я выявила для себя их преимущества и недостатки.

Все изученные алгоритмы показались мне довольно интересными. Выделился алгоритм Нелдера-Мида, так как он не требовал вычисления градиента и матрицы Гёссе. Вследствие этого, появляется возможность работать с негладкими функциями и к тому же исчезают ограничения на положительную определенность матрицы.

Это было интересным опытом, надеюсь, что мне предоставится возможность в будущем применить свои знания, полученные при выполнении данной лабораторной.

Список литературы

[1] *Методы дихотомии*
URL: http://www.machinelearning.ru/wiki/index.php?title=Методы_дихотомии
(дата обращения: 06.03.2020).

[2] *Метод Нелдера-Мида - Википедия*
URL: https://ru.wikipedia.org/wiki/Метод_Нелдера_-_Мида
(дата обращения: 06.03.2020).

[3] *Лекция 10: Многометрическая оптимизация - НОУ ИНТУИТ*
URL: https://www.intuit.ru/studies/courses/1020/188/lecture/4931
(дата обращения: 07.03.2020).

[4] *Алгоритм Левенберга-Марквардта - Википедия*
URL: https://ru.wikipedia.org/wiki/Алгоритм_Левенберга_-_Марквардта
(дата обращения: 09.03.2020).

[5] *Метод Ньютона - Википедия*
URL: https://ru.wikipedia.org/wiki/Метод_Ньютона
(дата обращения: 09.03.2020).

[6] *Метод сопряженных градиентов - Википедия*
URL: https://ru.wikipedia.org/wiki/Метод_сопряжённых_градиентов
(дата обращения: 09.03.2020).