МИНОБРНАУКИ РОССИИ

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

Кафедра Вычислительная техника

УТВЕРЖДАЮ
Проректор по учебной работе
________ О.Г. Локтионова
«_30_»__06______2017

МИНИМИЗАЦИЯ МЕТОДАМИ ПЕРВОГО ПОРЯДКА

Методические указания к лабораторным занятиям
по дисциплине «Методы оптимизации» для студентов
направления подготовки 09.04.01

Курск 2017

УДК 519.8

Составитель Ж.Т. Жусубалиев

Рецензент
Доктор технических наук, профессор С.А. Филист

Минимизация методами первого порядка: методические указания к практическим занятиям по дисциплине «Методы оптимизации» / Юго-Зап. гос. ун-т; сост.: Ж.Т. Жусубалиев. Курск, 2017. – 16 с.: табл. 2. – Библиограф.: с. 16.

Рассматриваются методы безусловной минимизации функции нескольких переменных методами первого порядка, а также алгоритмы их численной реализации с помощью ЭВМ. Приводятся задачи для упражнений.

Методические указания соответствуют требованиям программы, утвержденной учебно-методическим объединением по направлению подготовки «Информатика и вычислительная техника».

Предназначены для студентов направления подготовки 09.04.01 дневной и заочной форм обучения.

Текст печатается в авторской редакции

Подписано в печать . Формат 60×84 1/16.
Усл. печ. л. . Уч.-изд. л. . Тираж 30 экз. Заказ . Бесплатно.
Юго-Западный государственный университет.
305040, г. Курск, ул. 50 лет Октября, 94..

1. Градиентные методы

Градиентные методы основаны на итерационной процедуре, реализуемой в соответствии с формулой:

$$X_{k+1} = X_k + \alpha_k S_k,$$

где α_k – параметр, характеризующий длину шага; S_k – направление поиска в пространстве управляемых переменных.

Направление поиска в градиентных методах совпадает с антиградиентом целевой функции $S_k = -\nabla f(X_k)$.

Таким образом

$$X_{k+1} = X_k - \alpha_k \nabla f(X_k),\ k = 0, 1, \dots,\ \alpha_k > 0$$

означает, что из точки X_k на k-ой итерации алгоритма происходит спуск с шагом $\alpha_k |\nabla f(X_k)|$.

В зависимости от способа выбора значения α_k выделяют методы градиентного спуска с постоянным шагом и наискорейшего спуска.

1.1. Метод градиентного спуска с постоянным шагом

На первой итерации спуск начинают из выбранной начальной точки X_0. Выбор величины шага α_k осуществляется пользователем и остается постоянным пока функция убывает в рассчитываемых точках последовательности $\{X_k\}$, $k = 0, 1, \dots$, что контролируется проверкой на каждой k-ой итерации выполнения неравенства:

$$f(X_{k+1}) - f(X_k) < 0 \text{ или } f(X_{k+1}) - f(X_k) < \varepsilon \|\nabla f(X_k)\|^2, \qquad (1)$$

где $\varepsilon \in (0, 1)$ – величина, гарантирующая сходимость последовательности $\{\nabla f(X_k)\}$ к нулю.

Если на k-ой итерации не выполнено неравенство (1), то проводится дробление шага $\alpha_k = \alpha_k / 2$.

Построение последовательности $\{X_k\}$ заканчивается в точке X_k, для которой $\|\nabla f(X_k)\| < \varepsilon_1$, где ε_1 - заданное малое положительное число, или $k \ge M$, где M – предельное число итераций, или при двукратном одновременном выполнении двух неравенств

$\|X_{k+1}-X_k\|<\varepsilon_2$, $|f(X_{k+1})-f(X_k)|<\varepsilon_2$, где ε_2 - малое положительное число.

Алгоритм градиентного спуска с постоянным шагом может быть представлен следующим образом.

Шаг 1. Задать X_0, $0<\varepsilon<1$, $\varepsilon_1>0$, $\varepsilon_2>0$, M – предельное число итераций. Найти градиент функции в произвольной точке $\nabla f(X)=\left(\frac{\partial f(X)}{\partial x_1},\ldots,\frac{\partial f(X)}{\partial x_n}\right)^T$.

Шаг 2. Положить $k=0$.

Шаг 3. Вычислить $\nabla f(X_k)$.

Шаг 4. Проверить выполнение критерия окончания $\|\nabla f(X_k)\|<\varepsilon_1$:

а) если неравенство выполнено, то расчет окончен и $X_*=X_k$,

б) в противном случае перейти к шагу 5.

Шаг 5. Проверить выполнение неравенства $k\geq M$:

а) если неравенство выполнено, расчет окончен и $X_*=X_k$,

б) если нет, перейти к шагу 6.

Шаг 6. Задать величину шага α_k.

Шаг 7. Вычислить $X_{k+1}=X_k-\alpha_k\nabla f(X_k)$.

Шаг 8. Проверить выполнение условия $f(X_{k+1})-f(X_k)<0$ (или $f(X_{k+1})-f(X_k)<-\varepsilon\|\nabla f(X_k)\|^2$):

а) если условие выполнено, то перейти к шагу 9;

б) если условие не выполнено, положить $\alpha_k=\frac{\alpha_k}{2}$ и перейти к шагу 7.

Шаг 9. Проверить выполнение условий $\|X_{k+1}-X_k\|<\varepsilon_2$, $|f(x_{k+1})-f(x_k)|<\varepsilon_2$:

а) если оба условия выполнены при текущем значении k и $k=k-1$, то расчет окончен, $X_*=X_{k+1}$,

б) если хотя бы одно из условий не выполнено, положить $k=k+1$ и перейти к шагу 3.

.

Пример. Методом градиентного спуска с постоянным шагом $\alpha_k = 0{,}1$ из начальной точки $X_0 = (-2,\, 1)^T$ найти минимум функции $f(X) = 6x_1^{\,2} - 4x_1x_2 + 3x_2^{\,2} + 4\sqrt{5}(x_1 + 2x_2) + 22$.

Решение.

Шаг 1. Зададим $X_0 = (-2,\, 1)^T$, $\varepsilon_1 = 0{,}01$, $\varepsilon_2 = 10^{-3}$, $\varepsilon_3 = 10^{-4}$, $M = 20$. Градиент функции:

$\nabla f(X) = \left(12x_1 - 4x_2 + 4\sqrt{5},\ -4x_1 + 6x_2 + 4\sqrt{5}\right)^T$.

В таблице 1 приведены результаты вычислений.

Таблица 1

k	α_k	$X_k^{\,T}$	$f(X_k)$	$\lVert\nabla f(X_k)\rVert$
0	0,1	$(-2{,}000;\, 1{,}000)$	57,000	37,148
1	0,1	$(-1{,}047;\, -0{,}594)$	7,150	18,553
2	0,1	$(-0{,}923;\, -2{,}446)$	-15,976	10,309
3	0,1	$(-1{,}688;\, -3{,}136)$	-23,772	5,953
4	0,1	$(-1{,}811;\, -3{,}719)$	-26,494	3,508
5	0,1	$(-2{,}020;\, -4{,}001)$	-27,460	2,087
6	0,1	$(-2{,}091;\, -4{,}197)$	-27,806	1,248
7	0,1	$(-2{,}155;\, -4{,}304)$	-27,930	0,747
8	0,1	$(-2{,}185;\, -4{,}372)$	-27,975	0,448
9	0,1	$(-2{,}206;\, -4{,}412)$	-27,991	0,269
10	0,1	$(-2{,}218;\, -4{,}436)$	-27,997	0,161
11	0,1	$(-2{,}225;\, -4{,}450)$	-27,999	0,097
12	0,1	$(-2{,}230;\, -4{,}459)$	-28,000	0,058
13	0,1	$(-2{,}232;\, -4{,}464)$	-28,000	0,035
14	0,1	$(-2{,}234;\, -4{,}467)$	-28,000	0,021
15	0,1	$(-2{,}235;\, -4{,}469)$	-28,000	0,013
16	0,1	$(-2{,}235;\, -4{,}470)$	-28,000	0,008

На 16-й итерации выполнился критерий окончания $\lVert\nabla f(X_k)\rVert = 0{,}008 < \varepsilon_1$, следовательно $X_* \cong X_{16} = (-2{,}235;\, -4{,}470)^T$.

1.2. Метод наискорейшего градиентного спуска

Эффективность поиска минимума функции, использующего спуск в направлении антиградиента, можно повысить, если значение α_k на каждой итерации выбирать путем минимизации функции:

$$\varphi(\alpha_k) = f(X_k - \alpha_k \nabla f(X_k)) \to \min_{\alpha_k} .$$

Такой метод носит название метода наискорейшего спуска, или метода Коши.

Выбор значения α_k методом наискорейшего спуска гарантирует сходимость последовательности $\{\nabla f(X_k)\}$ к нулю. Поэтому в качестве условия прекращения итераций используется условие $\|\nabla f(X_k)\| < \varepsilon_1$, где ε_1 – заданное число. Построение последовательности $\{X_k\}$, $k = 0, 1, \ldots$ заканчивается в точке X_k, если $k \ge M$, M – предельное число итераций, или при двукратном одновременном выполнении неравенств $|X_{k+1} - X_k| < \varepsilon_2$, $|f(X_{k+1}) - f(X_k)| < \varepsilon_2$, где ε_2 – малое положительное число.

Алгоритм метода наискорейшего спуска включает следующие этапы.

Шаг 1. Задать X_0, $\varepsilon_1 > 0$, $\varepsilon_2 > 0$, M – предельное число итераций. Найти градиент функции в произвольной точке $\nabla f(X) = \left(\frac{\partial f(X)}{\partial x_1}, \ldots, \frac{\partial f(X)}{\partial x_n} \right)^T$.

Шаг 2. Положить $k = 0$.

Шаг 3. Вычислить $\nabla f(X_k)$.

Шаг 4. Проверить выполнение критерия окончания $\|\nabla f(X_k)\| < \varepsilon_1$:

а) если неравенство выполнено, то расчет окончен и $X_* = X_k$,

б) в противном случае перейти к шагу 5.

Шаг 5. Проверить выполнение неравенства $k \ge M$:

а) если неравенство выполнено, расчет окончен и $X_* = X_k$,

б) если нет, перейти к шагу 6.

Шаг 6. Вычислить величину шага α_k из условия

$$\varphi(\alpha_k) = f(X_k - \alpha_k \nabla f(X_k)) \to \min_{\alpha_k}$$

Шаг 7. Вычислить $X_{k+1} = X_k - \alpha_k \nabla f(X_k)$.

Шаг 8. Проверить выполнение условий $\|X_{k+1} - X_k\| < \varepsilon_2$, $|f(X_{k+1}) - f(X_k)| < \varepsilon_2$:

а) если оба условия выполнены при текущем значении k и $k = k - 1$, то расчет окончен, $X_* = X_k$;

б) если хотя бы одно из условий не выполнено, положить $k = k + 1$ и перейти к шагу 3.

Поиск в соответствии с приведенным алгоритмом обеспечивает более высокую надежность по сравнению с градиентным методом с постоянным шагом.

Однако при решении ряда практических задач скорость сходимости метода остается недопустимо низкой. Это объясняется тем, что изменения переменных непосредственно зависят от величины градиента, которая стремится к нулю в окрестности точки минимума, и отсутствует механизм ускорения сходимости к точке минимума на последних итерациях.

Одно из главных преимуществ метода связано с его устойчивостью. Метод обладает важным свойством, которое заключается в том, что при достаточно малой длине шага итерации обеспечивают выполнение неравенства $f(X_{k+1}) < f(X_k)$.

Пример. Методом наискорейшего градиентного спуска из начальной точки $X_0 = (-2, 1)^T$ найти минимум функции $f(X) = 6x_1^2 - 4x_1x_2 + 3x_2^2 + 4\sqrt{5}(x_1 + 2x_2) + 22$ [2].

Решение.

Шаг 2. Зададим $X_0 = (-2, 1)^T$, $\varepsilon_1 = 0{,}01$, $\varepsilon_2 = 10^{-3}$, $\varepsilon_3 = 10^{-4}$, $M = 20$. Градиент функции:

$$\nabla f(X) = \left(12x_1 - 4x_2 + 4\sqrt{5},\ -4x_1 + 6x_2 + 4\sqrt{5}\right)^T.$$

В таблице 2 приведены результаты вычислений.

Таблица 2

k	α_k	X_k^T	$f(X_k)$	$\lVert\nabla f(X_k)\rVert$
0	–	$(-2{,}000;\ 1{,}000)$	57,000	37,148
1	0,09	$(-0{,}283;\ -1{,}872)$	-5,154	15,182
2	0,145	$(-2{,}173;\ -3{,}001)$	-21,860	9,985
3	0,09	$(-1{,}711;\ -3{,}773)$	-26,350	4,081
4	0,145	$(-2{,}219;\ -4{,}077)$	-27,556	2,684
5	0,09	$(-2{,}095;\ -4{,}284)$	-27,881	1,097
6	0,145	$(-2{,}231;\ -4{,}336)$	-27,968	0,721

k	α_k	${X_k}^T$	$f(X_k)$	$\lVert\nabla f(X_k)\rVert$
7	0,09	$(-2{,}198;\ -4{,}422)$	-27,991	0,295
8	0,145	$(-2{,}235;\ -4{,}444)$	-27,998	0,194
9	0,09	$(-2{,}226;\ -4{,}459)$	-27,999	0,079
10	0,145	$(-2{,}236;\ -4{,}464)$	-28,000	0,052
11	0,09	$(-2{,}233;\ -4{,}468)$	-28,000	0,021
12	0,145	$(-2{,}236;\ -4{,}470)$	-28,000	0,014
13	0,09	$(-2{,}235;\ -4{,}471)$	-28,000	0,006

На 13-й итерации выполнился критерий окончания $\lVert\nabla f(X_{13})\rVert = 0{,}006 < \varepsilon_1$, следовательно $X_* \cong X_{13} = (-2{,}235;\ -4{,}471)^T$.

2. Метод Флетчера-Ривса

Среди методов первого порядка поиска минимума целевой функции $f(X)$ особое место занимают методы, основанные на свойствах векторов, сопряженных относительно некоторой квадратной матрицы. Одним из таких методов является метод Флетчера-Ривса [Fletcher R., Reeves C.M.].

Итерации проводятся по правилу:

$$X_{k+1} = X_k + \alpha_k S_k,\ k = 0, 1, \ldots,$$

где α_k – величина шага определяется для каждого значения k из условия $\varphi(\alpha_k) = f(X_k - \alpha_k S_k) \to \min\limits_{\alpha_k}$;

$S_k = -\nabla f(X_k) + \beta_{k-1} S_{k-1}$ – направление поиска, $S_0 = -\nabla f(X_0)$,

$$\beta_{k-1} = \frac{\lVert\nabla f(X_k)\rVert^2}{\lVert\nabla f(X_{k-1})\rVert^2}$$

или

$$\beta_{k-1} = \begin{cases} \dfrac{(\nabla f(X_k), [\nabla f(X_k) - \nabla f(X_{k-1})])}{\lVert\nabla f(X_{k-1})\rVert^2}, & k \notin J; \\ 0, & k \in J, \end{cases}$$

где $J = \{0, n, 2n, \ldots\}$ – множество номеров итераций, называемых моментами обновления алгоритмов. Обновление алгоритма состоит в периодическом через заданное число итераций обнулении β_{k-1}, что позволяет избежать накопления вычислительных погрешностей и уменьшения вероятности построения после каждых n итераций линейно зависимых направлений спуска.

Построение последовательности $\{X_k\}$ заканчивается в точке, для которой $\|\nabla f(X_{jk})\| < \varepsilon_1$, где ε_1 – заданное число, или при $k \geq M$, M – предельное число итераций, или при двукратном одновременном выполнении двух неравенств $|X_{k+1} - X_k| < \varepsilon_2$, $|f(x_{k+1}) - f(x_k)| < \varepsilon_2$, где ε_2 - малое положительное число.

Алгоритм метода включает следующие шаги.

Шаг 1. Задать X_0, $\varepsilon_1 > 0$, $\varepsilon_2 > 0$, M – предельное число итераций. Найти градиент $\nabla f(X)$.

Шаг 2. Положить $k = 0$.

Шаг 3. Вычислить $\nabla f(X_k)$.

Шаг 4. Проверить выполнение критерия окончания $\|\nabla f(X_k)\| < \varepsilon_1$:

а) если неравенство выполнено, то расчет окончен и $X_* = X_k$,

б) в противном случае перейти к шагу 5.

Шаг 5. Проверить выполнение неравенства $k \geq M$:

а) если неравенство выполнено, расчет окончен и $X_* = X_k$,

б) если нет, то при $k = 0$ перейти к шагу 6, а при $k \geq 1$ перейти к шагу 7.

Шаг 6. Определить $S_0 = -\nabla f(X_0)$.

Шаг 7. Определить

$$\beta_{k-1} = \frac{\|\nabla f(X_{k-1})\|^2}{\|\nabla f(X_k)\|^2}, \quad \beta_{k-1} = \begin{cases} \dfrac{(\nabla f(X_k), [\nabla f(X_k) - \nabla f(X_{k-1})])}{\|\nabla f(X_k)\|^2}, & k \notin J, \\ 0, & k \in J, \end{cases}.$$

Шаг 8. Определить $S_k = -\nabla f(X_k) + \beta_{k-1} S_{k-1}$.

Шаг 9. Найти α_k из условия $\varphi(\alpha) = f(X_k + \alpha S_k) \to \min\limits_{\alpha}$.

Шаг 10. Вычислить $X_{k+1} = X_k + \alpha_k S_k$.

Шаг 11. Проверить выполнение условий

$\|X_{k+1} - X_k\| < \varepsilon_2$, $|f(X_{k+1}) - f(X_k)| < \varepsilon_2$:

а) в случае выполнения обоих условий в двух последовательных итерациях с номерами k и $k = k-1$ расчет окончен, и найдена точка $X_* = X_{k+1}$,

б) если не выполняется хотя бы одно из условий, полагаем $k = k+1$ и переходим к шагу 3.

Замечание

Если целевая функция $f(X)$ является квадратичной с неотрицательно определенной матрицей H $(H > 0)$, то метод Флетчера-Ривса обеспечивает отыскание точки минимума не более чем за n шагов.

3. Квазиньютоновские методы

Среди алгоритмов многомерной оптимизации следует выделить группу алгоритмов, которые объединяют достоинства метода наискорейшего спуска и метода Ньютона. Такие алгоритмы принято относить к квазиньютоновским методам. Особенность этих алгоритмов состоит в том, что при их применении нет необходимости вычислять и обращать матрицу Гессе минимизируемой функции и в то же время удается сохранить высокую скорость сходимости алгоритмов, присущую методу Ньютона и его модификациям.

Элементы последовательности $\{X_k\}$, $k = 0, 1, \dots$ при минимизации функции $f(X)$ (т.е. $f(X_{k+1}) < f(X_k)$), непрерывно дифференцируемой в R^n, строят по правилу:

$$X_{k+1} = X_k - \alpha_k A_k \nabla f(X_k),\ k = 0, 1, \dots,$$

где $-\nabla f(X_k)$ – антиградиент функции $f(X)$ в точке X_k; A_k – положительно определенная матрица размера $n \times n$, обновляемая на k-ой итерации.

Матрицы $\{A_k\}$ определяются таким образом, чтобы последовательность $\{A_k\}$ при $k \to \infty$ имела предел, равный $H^{-1}(X_*)$, где – матрица Гессе минимизируемой функции, вычисленная в точке минимума X_*.

На k-ой итерации алгоритма происходит спуск из точки X_k с шагом спуска $|-\alpha_k A_k \nabla f(X_k)|$, величина шага α_k определяется для

каждого значения k из условия $\varphi(\alpha_k) = f(X_k + \alpha_k A_k \nabla f(X_k)) \to \min\limits_{\alpha_k}$.
На первой итерации ($k=1$) спуск начинают из выбранной пользователем точки X_0, при этом в качестве A_1 берут единичную матрицу порядка n. Таким образом первая итерация алгоритма совпадает с итерацией метода наискорейшего спуска.

Используя при конструировании матрицы A_k аппроксимацию матрицы $H^{-1}(X_k)$ с учетом информации о градиенте минимизируемой функции в той же точке, можно существенно упростить процедуру нахождения направления спуска на k-ой итерации. Именно эти соображения лежат в основе построения квазиньютоновских методов. Рассмотрим метод Дэвидона-Флетчера-Пауэлла

Идея метода заключается в построении элементов последовательности $\{X_k\}$, $k = 0, 1, ...$ при минимизации функции $f(X)$ по правилу:

$$X_{k+1} = X_k - \alpha_k A_k \nabla f(X_k),\ k = 0, 1, ...,$$

где матрица A_k вычисляется в соответствии с:

$$A_{k+1} = A_k + \Delta A_k,\ A_0 = E.$$

Здесь ΔA_k – положительно определенная матрица порядка n, называемая поправочной. Способ ее выбора и определяет метод Дэвидона-Флетчера-Пауэлла (ДФП-метод):

$$\Delta A_k = \frac{\Delta X_k (\Delta X_k)^T}{(\Delta X_k)^T \Delta g_k} - \frac{A_k \Delta g_k (\Delta g_k)^T}{(\Delta g_k)^T A_k \Delta g_k},$$

где $\Delta X_k = X_{k+1} - X_k$, $\Delta g_k = \nabla f(X_{k+1}) - \nabla f(X_k)$.

Данный способ пересчета матрицы A_k сохраняет ее свойство положительной определенности на каждой k-ой итерации.

Для неквадратичных функций $f(X)$ алгоритм перестает быть конечным. Чтобы ослабить влияние накапливающихся погрешностей на сходимость ДПФ-метода и уменьшить вероятность появления линейно зависимых направлений спуска, применяют процедуру обновления алгоритма, в соответствие с которой через каждые n итераций в качестве матрицы A_k используют единичную матрицу, т.е.:

$$A_k = \begin{cases} E, & k \in J,\ J = \{0, n, 2n, ...\}, \\ A_{k-1} + \Delta A_{k-1}, & k \notin J. \end{cases}$$

Построение последовательности $\{X_k\}$ заканчивается в точке X_k, для которой $\|\nabla f(X_k)\| < \varepsilon_1$, где ε_1 – заданное малое положительное число, или $k \ge M$, где M – предельное число итераций, или при двукратном одновременном выполнении двух неравенств $\|X_{k+1} - X_k\| < \varepsilon_2$, $|f(X_{k+1}) - f(X_k)| < \varepsilon_2$, где ε_2 – малое положительное число.

Алгоритм ДФП-метода

Шаг 1. Задать X_0, $\varepsilon_1 > 0$, $\varepsilon_2 > 0$, M – предельное число итераций. Найти градиент $\nabla f(X)$.

Шаг 2. Положить $k = 0$, $A_0 = E$.

Шаг 3. Вычислить $\nabla f(X_k)$.

Шаг 4. Проверить выполнение критерия окончания $\|\nabla f(X_k)\| < \varepsilon_1$:

а) если неравенство выполнено, то расчет окончен и $X_* = X_k$,

б) в противном случае перейти к шагу 5.

Шаг 5. Проверить выполнение неравенства $k \ge M$:

а) если неравенство выполнено, расчет окончен и $X_* = X_k$,

б) если нет, то при $k = 0$ перейти к шагу 10, а при $k \ge 1$ перейти к шагу 6.

Шаг 6. Вычислить $\Delta g_{k-1} = \nabla f(X_k) - \nabla f(X_{k-1})$.

Шаг 7. Вычислить $\Delta X_{k-1} = X_k - X_{k-1}$.

Шаг 8. Вычислить $\Delta A_{k-1} = \dfrac{\Delta X_{k-1}(\Delta X_{k-1})^T}{(\Delta X_{k-1})^T \Delta g_{k-1}} - \dfrac{A_{k-1}\Delta g_{k-1}(\Delta g_{k-1})^T}{(\Delta g_{k-1})^T A_{k-1}\Delta g_{k-1}}$.

Шаг 9. Вычислить $A_k = A_{k-1} + \Delta A_{k-1}$.

Шаг 10. Вычислить $S_k = -A_k \nabla f(X_k)$.

Шаг 11. Вычислить α_k из условия: $\varphi(\alpha) = f(X_k + \alpha S_k) \to \min\limits_{\alpha}$.

Шаг 12. Вычислить $X_{k+1} = X_k + \alpha_k S_k$.

Шаг 13. Проверить выполнение условий:
$\|X_{k+1} - X_k\| < \varepsilon_2$, $|f(X_{k+1}) - f(X_k)| < \varepsilon_2$:

а) в случае выполнения обоих условий в двух последовательных итерациях с номерами k и $k = k - 1$ расчет окончен, и найдена точка $X_* = X_{k+1}$, б) если не выполняется хотя бы одно из условий, полагаем $k = k+1$ и переходим к шагу 3.

Замечание

- Алгоритм ДФП-метода в применении к квадратичной функции $f(X) = \dfrac{1}{2}(H(X), X) + (b, X)$ с положительно определенной матрицей Гессе H обеспечивает отыскание минимума $X_* = Hb$ не более чем за n шагов.

Пример. С помощью ДФП-метода найти точку минимума функции $f(X) = 4x_1^2 + 3x_2^2 - 4x_1x_2 + x_1$, $X_0 = (0,\ 0)^T$.

Решение.

Шаг 1. Зададим $X_0 = (0,\ 0)^T$, $\varepsilon_1 = 0{,}1$, $\varepsilon_2 = 0{,}01$, $M = 10$.
$\nabla f(X) = (8x_1 - 4x_2 + 1;\ 6x_2 - 4x_1)^T$.

Шаг 2. $k = 0$, $A_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

Шаг 3. $\nabla f(X_0) = (1;\ 0)^T$.

Шаг 4. $\|\nabla f(X_0)\| = 1 > \varepsilon_1$, переходим к шагу 5.

Шаг 5. $k = 0 < M$, переходим к шагу 10.

Шаг 10. $S_0 = -\nabla f(X_0) = (-1;\ 0)^T$.

Шаг 11. $\varphi(\alpha) = f(-\alpha;\ 0) = 4\alpha^2 - \alpha \to \min\limits_{\alpha}$, $\alpha_0 = 1/8$.

Шаг 12. $X_1 = X_0 + \alpha_0 S_0 = (0;\ 0)^T + \dfrac{1}{8}(-1;\ 0)^T = (-1/8;\ 0)^T$.

Шаг 3. $\nabla f(X_1) = (0;\ 1/2)^T$.

Шаг 4. $\|\nabla f(X_1)\| = 0{,}5 > \varepsilon_1$.

Шаг 5. $k = 1 < M$.

Шаг 6. $\Delta g_0 = \nabla f(X_1) - \nabla f(X_0) = (0;\ 1/2)^T - (1;\ 0)^T = (-1;\ 1/2)^T$.

Шаг 7. $\Delta X_0 = X_1 - X_0 = (-1/8;\ 0)^T - (0;\ 0)^T = (-1/8;\ 0)^T$.

Шаг 8. $\Delta A_0 = \dfrac{\Delta X_0 (\Delta X_0)^T}{(\Delta X_0)^T \Delta g_0} - \dfrac{A_0 \Delta g_0 (\Delta g_0)^T}{(\Delta g_0)^T A_0 \Delta g_0} =$

$$= \frac{(-1/8;\,0)^T(-1/8;\,0)}{(-1/8;\,0)(-1;\,1/2)^T} - \frac{\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}(-1;\,1/2)^T(-1;\,1/2)}{(-1;\,1/2)\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}(-1;\,1/2)^T} =$$

$$\begin{pmatrix}1/8 & 0\\ 0 & 0\end{pmatrix} - \begin{pmatrix}4/5 & -2/5\\ -2/5 & 1/5\end{pmatrix} = \begin{pmatrix}-27/40 & 2/5\\ 2/5 & -1/5\end{pmatrix}.$$

Шаг 9. $A_1 = A_0 + \Delta A_0 = \begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix} + \begin{pmatrix}-27/40 & 2/5\\ 2/5 & -1/5\end{pmatrix} = \begin{pmatrix}13/40 & 2/5\\ 2/5 & 4/5\end{pmatrix}.$

Шаг 10. $S_1 = -A_1\nabla f(X_1) = \begin{pmatrix}-13/40 & -2/5\\ -2/5 & -4/5\end{pmatrix}\begin{pmatrix}0\\ 1/2\end{pmatrix} = \begin{pmatrix}-1/5\\ -2/5\end{pmatrix}.$

Шаг 11. $\varphi(\alpha) = f(-1/8 - 1/5\alpha;\, -2/5\alpha) = 8/25\alpha^2 - 1/5\alpha - 1/16 \to \min\limits_{\alpha}$,

$\alpha_1 = 5/16$.

Шаг 12. $X_2 = X_1 + \alpha_1 S_1 = (-1/8;\,0)^T + \dfrac{5}{16}(-1/5;\,-2/5)^T = (-3/16;\,-1/8)^T$.

Шаг 3. $\nabla f(X_2) = (0;\,0)^T$.

Шаг 4. $\|\nabla f(X_2)\| = 0 < \varepsilon_1$, расчет окончен и $X_* = X_2$.

4. Задачи для упражнений

1. Методами градиентного и наискорейшего градиентного спуска найти минимум функции $f(X) = x_1^3 + x_1x_2 + x_2^2 - 2x_1 + 3x_2 - 4$, $X_0 = (0,\,0)^T$.

2. Методами градиентного и наискорейшего градиентного спуска, методами Флетчера-Ривса и Д-Ф-П найти минимум функции $f(X) = (x_2 + x_1^2)^2 + (1 - x_1)^2$, $X_0 = (0,\,0)^T$.

3. Методами градиентного и наискорейшего градиентного спуска, методами Флетчера-Ривса и Д-Ф-П найти минимум функции $f(X) = \left[(x_2+1)^2 + x_1^2\right]\cdot\left[x_1^2 + (x_2-1)^2\right]$ из начальных точек $X_0 = (0{,}5;\,0)^T$ и $X_0 = (-0{,}1;\,-0{,}5)^T$.

4. Методами градиентного и наискорейшего градиентного спуска, методами Флетчера-Ривса и Д-Ф-П найти минимум функции

$f(X)=\left(x_2{}^2+x_1{}^2-1\right)^2+\left(x_1+x_2-1\right)^2$ из начальных точек $X_0=(0,\,3)^T$ и $X_0=(3,\,0)^T$.

5. Методами градиентного и наискорейшего градиентного спуска, методами Флетчера-Ривса и Д-Ф-П найти минимум функции $f(X)=-x_1{}^2\cdot\exp\left[1-x_1{}^2-20{,}25(x_1-x_2)^2\right]$ из начальной точки $X_0=(0{,}1;\,0{,}5)^T$.

6. Задана функция $f(X)$ и начальная точка X_0. Найдите минимум функции $f(X)$:

1) $f(X)=x_1{}^3-x_1x_2+x_2{}^2-2x_1+x_2-4$, $X_0=(-1;\,1)^T$;
2) $f(X)=x_1{}^2+2x_2{}^2-4x_1-4x_2$, $X_0=(0;\,0)^T$;
3) $f(X)=x_1{}^2+10(x_2-\sin x_1)^2$, $X_0=(1;\,1)^T$;
4) $f(X)=e^{x_1{}^2}+x_2+(x_1-x_2)^2$, $X_0=(0;\,0)^T$;
5) $f(X)=e^{x_2{}^2-x_1}+e^{x_1}$, $X_0=(-1;\,-1)^T$;
6) $f(X)=e^{-x_1}+x_1{}^2+x_2{}^2$, $X_0=(0{,}5;\,0{,}5)^T$;
7) $f(X)=e^{-x_2}+\cos\left(x_1{}^2+x_2\right)$, $X_0=(0;\,0)^T$;
8) $f(X)=e^{x_1}+x_2{}^2-2x_1$, $X_0=(0;\,0)^T$;
9) $f(X)=x_1{}^2-\cos(x_2-1)$, $X_0=(0;\,0)^T$;
10) $f(X)=x_2{}^2+e^{x_1}-3x_1$, $X_0=(0;\,0)^T$;
11) $f(X)=e^{x_1-x_2}+x_1{}^2+x_2{}^2$, $X_0=(1;\,1)^T$;
12) $f(X)=e^{x_1{}^2-x_2}+e^{x_2}$, $X_0=(-1;\,-1)^T$;
13) $f(X)=e^{x_1}+\left(x_1-x_2{}^2\right)^2$, $X_0=(0{,}5;\,0{,}5)^T$;
14) $f(X)=e^{x_1}+x_1{}^2+x_2{}^2$, $X_0=(-1;\,-1)^T$;
15) $f(X)=e^{x_2{}^2-x_1}+x_1{}^2$, $X_0=(0;\,0)^T$;
16) $f(X)=e^{-x_1}-\cos\left(x_1{}^2+x_2{}^2\right)$, $X_0=(0;\,0)^T$;
17) $f(X)=e^{-x_2}+\left(x_2+x_1{}^2\right)^2$, $X_0=(0{,}5;\,0{,}5)^T$;
18) $f(X)=e^{x_1+x_2}+x_1{}^2+x_2{}^2$, $X_0=(-1;\,-1)^T$;

19) $f(X) = x_2{}^2 - \cos(x_1 - 1),\ X_0 = (0;\ 0)^T$;

20) $f(X) = x_1{}^2 + e^{x_2} - 3x_2,\ X_0 = (0;\ 0)^T$;

21) $f(X) = e^{x_2} + x_1{}^2 + x_2{}^2,\ X_0 = (1;\ 1)^T$;

22) $f(X) = e^{-x_1} + x_2{}^2 + 2x_1,\ X_0 = (0;\ 0)^T$.

---

1. Пантелеев, А.В. Методы оптимизации в примерах и задачах [Текст]: учеб. пособие / А.В. Пантелеев, Т.А. Летова; 2-е изд., исправл. – М.: Высш. шк., 2005. – 544 с.: ил.

2. Аттетков, А.В. Методы оптимизации [Текст]: учебник для вузов / А.В. Аттетков, С.В. Галкин, В.С. Зарубин / Под ред. В.С. Зарубина, А.П. Крищенко; 2-е изд. – М.: Изд-во МГТУ им. Н.Э. Баумана, 2003. – 440 с.