Министерство высшего образования и науки
Российской Федерации
Федеральное государственное бюджетное образовательное
учреждение высшего образования
«Владимирский государственный университет имени Александра Григорьевича и Николая Григорьевича Столетовых»

Абрахин С.И.

Методические указания для выполнения лабораторных работ по дисциплине «Методы оптимизации»

для студентов по направлению подготовки
01.03.02 «Прикладная математика и информатика»

Владимир, 2018 г.

Составитель: Абрахин Сергей Иванович

Методические указания для выполнения лабораторных работ по дисциплине «Методы оптимизации» / Владимирский государственный университет имени Александра Григорьевича и Николая Григорьевича Столетовых; Абрахин Сергей Иванович, – Владимир: Изд-во ВлГУ, 2018 – 39 с.

Методические рекомендации составлены на основе требований ФГОС ВО направления 01.03.02 «Прикладная математика и информатика». В методических указаниях содержатся материалы для лабораторных работ по численным методам решения задач минимизации функций. Описание каждой лабораторной работы включает краткую постановку задачи, описание используемого метода и варианты индивидуальных заданий. Пособие предназначено для студентов, обучающихся по направлению подготовки 01.03.02 «Прикладная математика и информатика».

Для организации эффективной работы студентов использованы методические пособия ведущих вузов России.

Пособие выполнено в рамках государственного задания ВлГУ на 2016/19 года на выполнение государственных работ в сфере научно-технической деятельности.

Рецензент – Давыдов Н.Н.-профессор, доктор технических наук, директор ИПМФИ

СОДЕРЖАНИЕ

Лабораторная работа № 1. Минимизация функций одной переменной методами дихотомии и золотого сечения

Постановка задачи

Используя методы дихотомии и золотого сечения, найти на отрезке $[a, b]$ точку u^*, в которой достигается минимальное значение унимодальной функции $J(u)$ с точностью ε. Вычислить минимальное значение $J(u^*)$. Оценить погрешность вычислений.

Краткая теоретическая часть

Методы дихотомии и золотого сечения применяются к функциям, которые являются унимодальными на отрезке.

Определение 1. Функция $J(u)$ является унимодальной на отрезке $[a, b]$, если $J(u)$ непрерывна на $[a, b]$ и $\exists \alpha, \beta : a \le \alpha \le \beta \le b$, такие, что:

1) $J(u)$ строго монотонно убывает при $a \le u \le \alpha$ (если $a < \alpha$);
2) $J(u)$ строго монотонно возрастает при $\beta \le u \le b$ (если $\beta < b$);
3) $J(u)=const$, при $\alpha \le u \le \beta$.

Случаи, когда некоторые из отрезков $[a, \alpha]$, $[\alpha, \beta]$, $[\beta, b]$ вырождаются в точку, не исключаются (см. рис. 1).

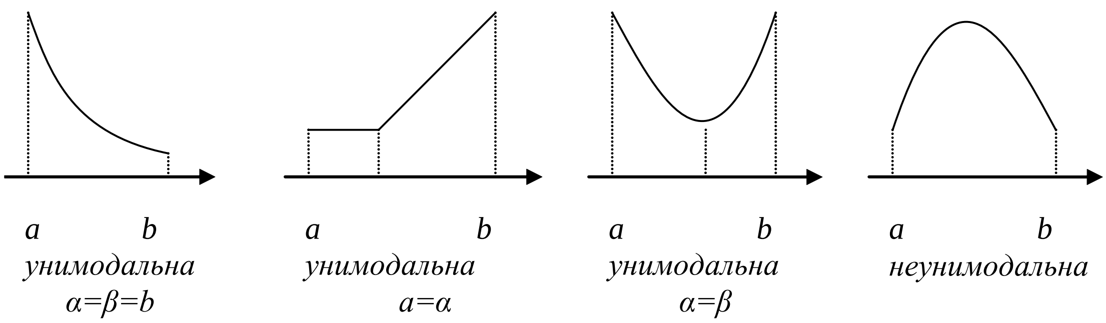

Рис. 1. Примеры унимодальных и не унимодальных функций

Метод дихотомии (деления отрезка пополам)

Алгоритм

Описание алгоритма метода

Выберем параметр метода – число δ: $0 < \delta < b - a$.

Этап 1

Вычислим $u_1 = \frac{a+b-\delta}{2}, u_2 = \frac{a+b+\delta}{2}$ (см. рис. 1) и проверяем условие: если $J(u_1) < J(u_2) \Rightarrow a_1 = a, b_1 = u_2$, если $J(u_1) > J(u_2) \Rightarrow a_1 = u_1, b_1 = b$. Тем самым перешли от отрезка $[a,b]$ к отрезку $[a_1,b_1]$, длина которого $b_1 - a_1 = \frac{b-a-\delta}{2} + \delta < b-a$, т.к. $J(u)$ - унимодальна на $[a,b]$, то ясно, что $[a_1,b_1]$ имеет одну общую точку с множеством U^* точек минимума на отрезке $[a,b]$.

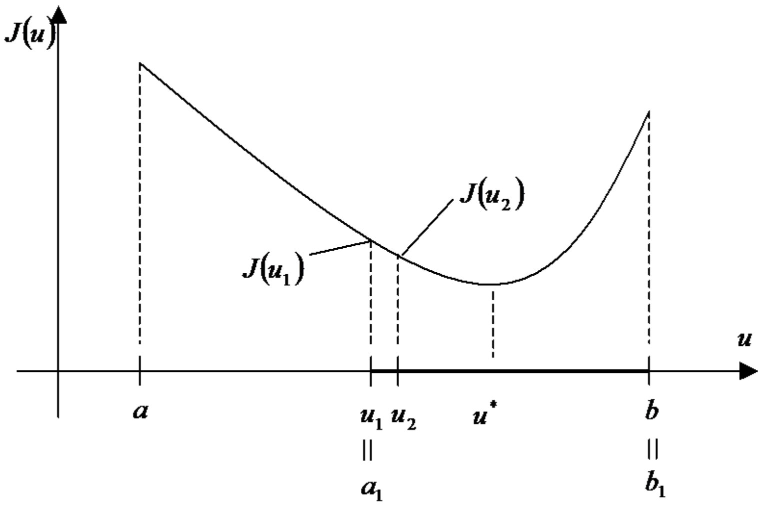

Рис. 2. Определение точек *u₁* и *u₂*

Этап 2

Повторяем действие этапа 1 и приходим к шагу с номером *k*, для которого известен отрезок $[a_{k-1}, b_{k-1}]$, вычисляем $u_{2k-1} = \frac{b_{k-1}+a_{k-1}-\delta}{2}, u_{2k} = \frac{b_{k-1}+a_{k-1}+\delta}{2}$ и выполняем сравнение: если $J(u_{2k-1}) < J(u_{2k}) \Rightarrow a_k = a_{k-1}, b_k = u_{2k}$, если $J(u_{2k-1}) > J(u_{2k}) \Rightarrow a_k = u_{2k-1}, b_k = b_{k-1}$. Тем самым приходим к отрезку $[a_k, b_k]$ и его длина $b_k - a_k = \frac{b-a-\delta}{2^k} + \delta$.

Этап 3

Процесс останавливается при достижении точности: $|b_k - a_k| < \varepsilon$, в качестве u^* берем $\frac{a_k+b_k}{2}$, а $J_* = J(u^*)$. При этом точность вычислений (погрешность) не может превысить длины отрезка $[a_k, b_k]$, т.е. $\frac{b-a-\delta}{2^k}$.

<u>**Замечание 1**</u> Т.к. параметр метода δ на каждой итерации *k* должен быть меньше длины текущего отрезка *[ak, bk]*, а длина отрезка на

заключительной итерации не превосходит значение точности ε, то можно выбрать единый параметр $\delta<\varepsilon$, который будет удовлетворять всем итерациям.

Замечание 2 Из описанного метода следует, что для определения минимума унимодальной функции на отрезке $[a,b]$ требуется 2^k раз вычислений значения функции. В случаях, когда вычисление значения функции в точке сложно желательно использовать методы, требующие меньшего количества вычислений значений функции, и такие методы есть. Рассмотрим их.

Метод золотого сечения

Метод золотого сечения от рассмотренного ранее метода отличается тем, что он позволяет решить задачу минимизации унимодальной на отрезке $[a,b]$ функции с требуемой точностью при меньшем количестве вычислений значений функции.

Определение 2: Золотым сечением отрезка называется деление отрезка на две неравные части так, чтобы отношение длины всего отрезка к длине большей части равнялось отношению длины большей части к длине меньшей части отрезка.

Нетрудно проверить, что золотое сечение отрезка $[a, b]$ производится двумя точками $u_1 = a + \left(3-\sqrt{5}\right)\left(b-a\right)/2$ и $u_2 = a + \left(\sqrt{5}-1\right)\left(b-a\right)/2$, расположенными симметрично относительно середины отрезка, причем $a < u_1 < u_2 < b$.

Точки золотого сечения обладают следующими свойствами, которые используются в методе золотого сечения:

1. Для точек золотого сечения выполняется равенство:

$$u_1=a+b-u_2 \text{ и } u_2=a+b-u_1$$

2. Точка u_2 производит золотое сечение отрезка $[u_1, b]$, так как $u_2-u_1<b-u_2$ и $(b-u_1)/(b-u_2)=(b-u_2)/(u_2-u_1)$. Аналогично точка u_1 производит золотое сечение отрезка $[a, u_2]$.

Алгоритм метода золотого сечения заключается в следующем:

1. Пусть $a_1 = a, b_1 = b$ вычисляем u_1 и u_2 по вышеуказанным формулам и сравниваем $J(u_1)$ и $J(u_2)$:

   Если $J(u_1) \le J(u_2)$, то $a_2 = a_1, b_2 = u_2, \overline{u_2} = u_1$. Если же $J(u_1) > J(u_2)$, то

$a_2 = u_1, b_2 = b_1, \overline{u_2} = u_2$. Получим отрезок $[a_2, b_2]$, на котором функция унимодальна, причем длина отрезка $b_2 - a_2 = (\sqrt{5}-1)\dfrac{b-a}{2}$. Внутри отреза $[a_2, b_2]$ содержится точка $\overline{u_2}$, в которой уже вычислено значение $J(\overline{u_2}) = \min\{J(u_1), J(u_2)\}$, которая, согласно свойству точек золотого сечения, производит золотое сечение отрезка $[a_2, b_2]$.

2. По достижении шага N имеем отрезок $[a_{n-1}, b_{n-1}]$ и точку $\bar{u}_{n-1}$ от предыдущего шага и следовательно $J(\bar{u}_{n-1}) = \min\{J(u_1), ..., J(u_{n-1})\} = \min\limits_{1<i<n-1} J(u_i)$. Тогда $u_n = a_{n-1} + b_{n-1} + \bar{u}_{n-1}$, которая производит золотое сечение отрезка $[a_{n-1}, b_{n-1}]$. Пусть $a_{n-1} < u_n < \bar{u}_{n-1} < b_{n-1}$, тогда если $J(u_n) \le J(u_{n-1})$, то $a_n = a_{n-1}, b_n = \bar{u}_{n-1}, \overline{u_n} = u_n$, если $J(u_n) > J(u_{n-1})$, то $a_n = u_n, b_n = b_{n-1}, \overline{u_n} = u_{n-1}$. Т.о. получили новый отрезок $[a_n, b_n]$, имеющий длину $b_n - a_n = \dfrac{\sqrt{5}-1}{2^n}(b-a)$.

   <u>Замечание.</u> Если $\bar{u}_{n-1} < u_n$, то действуем аналогично с точностью до наоборот.
3. Процесс продолжается до достижения точности $b_n - a_n < \varepsilon$, погрешность будет $\delta \le \dfrac{\sqrt{5}-1}{2^n}(b-a)$.

Сравнение методов дихотомии и золотого сечения

Сравним метод золотого сечения и метод дихотомии.

Метод дихотомии дает погрешность $\dfrac{b-a-\delta}{2^n} \approx \dfrac{b-a}{2^n}$, а метод золотого сечения - $\dfrac{b-a}{2^n}(3-\sqrt{5})$, следовательно, метод золотого сечения дает более высокую точность (в $3-\sqrt{5}$ раза), чем метод дихотомии.

<u>Общее замечание.</u> Методы золотого сечения и дихотомии можно использовать и для не унимодальных функций, в этом случае найдем какой-либо из локальных минимумов данной функции.

Задание к лабораторной работе

1. Убедиться, что заданная в соответствии с вариантом, функция является унимодальной.

2. Разработать программу, реализующую оба описанных метода для функции $J(u)$; найти её минимальное значение на отрезке $[a, b] = [-100,100]$, с точностью $\varepsilon=10^{-4}$. Выводимые значения для каждого метода:
   - Точка минимума u^* и минимальное значение J^*;
   - Количество итераций и количество вызовов функции;
   - Точность вычислений и длина последнего отрезка локализации точки минимума.

Варианты заданий

№ вар.	$J(u)$	α	β	№ вар.	$J(u)$	α	β
1	$u^2 + \alpha e^{\beta u}$	1	-0.85	**16**	$u^4 + \alpha\, arctg\; \beta u$	-0.3	3.5
2		2	-0.65	**17**		-0.1	4.0
3		3	-0.45	**18**		0.2	4.5
4		4	-0.25	**19**		0.4	5.0
5		5	-0.05	**20**		0.8	5.5
6		6	0.15	**21**	$\beta u + e^{\lvert u-\alpha \rvert}$	0.2	-4.0
7		7	0.35	**22**		0.4	-3.4
8		8	0.55	**23**		0.6	-2.8
9		9	0.75	**24**		0.8	-2.2
10		10	0.95	**25**		1.0	-1.6
11	$u^4 + \alpha\, arctg\; \beta u$	-1.5	1.0	**26**		1.2	-1.0
12		-1.3	1.5	**27**		1.4	-0.4
13		-1.1	2.0	**28**		1.6	-0.2
14		-0.9	2.5	**29**		1.8	0.8
15		-0.7	3.0	**30**		2.0	1.4

Содержание отчета

- Титульный лист;
- Цель работы;
- Постановка задачи;
- Краткая теоретическая часть;
- Вариант задания;
- Проверка функции на унимодальность;
- График функции J(u) на заданном интервале;
- Листинг программы;
- Результаты работы программы;
- Вывод.

Контрольные вопросы

1. Дайте определение унимодальной функции.
2. Каким образом в методе дихотомии гарантируется попадание точки минимума u^* в отрезок $[a_k, b_k]$ при любом $k > 0$?
3. Чем отличается метод дихотомии от метода золотого сечения?
4. Почему в методе золотого сечения, начиная со второй итерации, необходимо вычислять только один раз значение функции $J(u)$ вместо двух в методе дихотомии?
5. Какой из рассмотренных в лабораторной работе методов сходится быстрее и почему?

Лабораторная работа №2. Минимизация функций одной переменной методами ломаных касательных

Применение рассмотренных методов дихотомии и золотого сечения без априорных знаний об унимодальности функции, и приводит к какой-либо точке локального минимума (или окрестности этой точки). В связи с этим встает вопрос о методах поиска точки глобального минимума функции одной переменной на заданном отрезке. Методы ломаных и касательных (он же метод Ньютона) позволяют решить эту задачу.

Постановка задачи

Найти при помощи методов ломаных и касательных точку глобального минимума функции $J(u)$ на отрезке $[a, b]$ с заданной точностью ε. Найти минимальное значение функции.

Краткая теоретическая часть

Описание метода ломаных

Метод ломанных применим для функций $J(u)$, удовлетворяющих условию Липшица на отрезке $[a, b]$.

Определение 1. Функция $J(u)$ удовлетворяет условию Липшица (такие функции называют «липшицевыми») на отрезке $[a, b]$, если существует такое число $L>0$, называемое константой Липшица, что для него выполняется условие:

$$\frac{|J(u) - J(v)|}{|u - v|} \le L, \quad \forall u, v \in [a, b].$$

Теперь рассмотрим собственно сам метод.

Из определения липшецевости функции следует, что:

1. Функция $J(u)$ - Липшицева на *отрезке* $[a, b]$=> $J(u)$- непрерывна на отрезке $[a, b]$=>$U^*\neq\emptyset$;
2. Геометрический смысл: условие липшицевости означает, что угловой коэффициент (тангенс угла наклона) $tg(\alpha)=\frac{|J(u)-J(v)|}{|u-v|}$ хорды, соединяющей любые две точки $(u, J(u))$ и $(v, J(v))$ графика функции не превышает значение константы Липшица L, для $\forall u, v \in [a,b]$.

Вычисление константы Липшица

Аналитически константу Липшица можно оценить с помощью утверждения 1.

<u>Утверждение 1.</u> Пусть функция $J(u)\in C^1[a,b]$, т.е. дифференцируема на $[a, b]$ и ее производная $J'(u)$ ограничена на отрезке $[a, b]$. Тогда функция $J(u)$ удовлетворяет условию Липшица на отрезе $[a, b]$ с константой $L = J = \sup\limits_{u\in[a,b]}\left|J'(u)\right|$.

Рассмотрим численный метод вычисления константы Липшица.

Согласно определению:

$$J'(u) = \lim_{\Delta u\to 0}\left(\frac{J(u+\Delta u)-J(u)}{\Delta u}\right) = \left(\frac{J(u+\Delta u)-J(u)}{\Delta u}\right) + O(\Delta u),$$

где $O(\Delta u)$: $\lim\limits_{\Delta u\to 0}\frac{O(\Delta u)}{\Delta u} = 0$, тогда можно считать, что $J'(u) \approx \frac{J(u+\Delta u)-J(u)}{\Delta u}$ при достаточно малом Δu.

Возьмем число $m>0$ и нанесем равномерное разбиение на отрезок $[a, b]$: $u_i = a + \Delta u \cdot i$, где $\Delta u = \frac{b-a}{m}$, $i=0,1,...,m$.

Далее введем обозначение: $L_m = \max\limits_{i=\overline{1,m}}\frac{\left|J(u_i)-J(u_{i-1})\right|}{u_i-u_{i-1}}$.

<u>Утверждение 2.</u> $\lim\limits_{m\to\infty} L_m = L$, следовательно, при достаточно большом числе m в качестве приближенного значения константы Липшица можно взять величину L_m.

Вспомогательная функция g(u,v)

Рассмотрим функцию $g(u,v)=J(v)-L|u-v|$, где v - фиксировано, $u\in[a,b]$. Рассмотрим график этой функции (см. рис. 3). Раскроем модуль:

- Если $u\geq v\Rightarrow g(u,v)=J(v)-L(u-v)$
- Если $u<v\Rightarrow g(u,v)=J(v)-L(u-v)$.

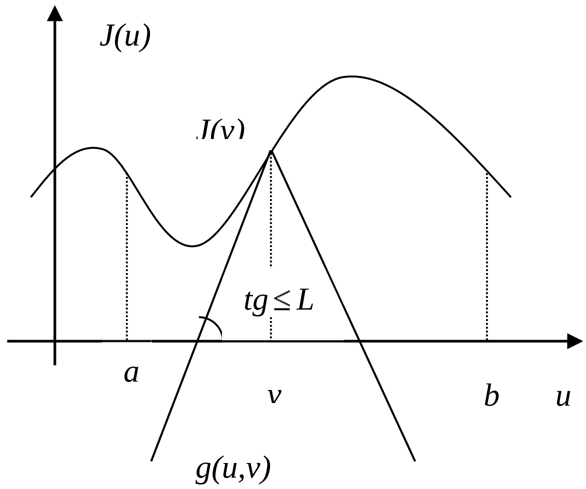

Рис. 3. График функции *g(u,v)*

График функции $J(u)$ будет всегда выше графика функции $g(u,v)$, т.е. $J(u)-g(u,v)\geq 0$ для любого фиксированного $v\in[a,b]$ и $\forall u\in[a,b]$.

Описание алгоритма метода ломаных

Будем предполагать, что константа Липшица L известна.

1. Выбираем произвольно точку $u_0\in[a,b]$ и строим функцию $g(u,u_0)=J(u_0)-L|u-u_0|=p_0$;
2. Следующая точка выбирается из условия $p_0(u_1)=\min\limits_{u\in[a,b]}p_0(u), u_1\in[a,b]$ (очевидно, что, $u_1=a$ или $u_1=b$);
3. Строится функция $p_1(u)=\max\limits_{u\in[a,b]}\{g(u,u_1),p_0(u)\}$;
4. Очередная точка u_2 находится как $p_1(u_2)=\min\limits_{u\in[a,b]}p_1(u), u_2\in[a,b]$;
5. Рассмотрим шаг $n+1$, т.е. $u_0,u_1,...,u_n$ – известны, таким оборазом $p_n(u)=\max\limits_{u\in[a,b]}\{g(u,u_n),p_{n-1}(u)\}$, а u_{n+1} определяем из условия $p_n(u_{n+1})=\min\limits_{u\in[a,b]}p_n(u), u_{n+1}\in[a,b]$ и строим $p_{n+1}(u)=\max\limits_{u\in[a,b]}\{g(u,u_{n+1}),p_n(u)\}$;

6. Процесс останавливается по достижении точности: $|J(u_n)-J(u_{n-1})|<\varepsilon$ (тип I) или $|u_n-u_{n-1}|<\varepsilon$ (тип II).

Графическая интерпретация метода ломаных

Графическая интерпретация метода ломаных приведена на рис. 4.

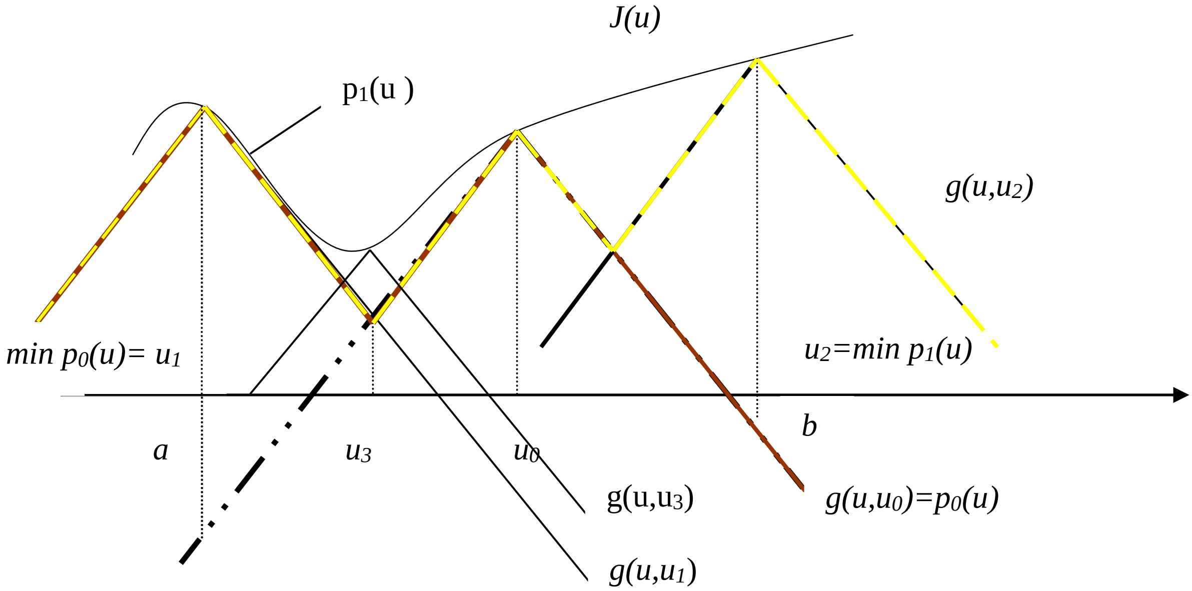

Рис. 4. Графическая интерпретация метода ломаных

Функция $p_n(u)$ – является кусочно-линейной функцией и график ее представляет собой непрерывную ломаную линию, состоящую из отрезков прямых с угловым наклоном $\pm L$, причем $p_n(u)$ удовлетворяет условиям Липшица с той же константой Липшица L , что и функция $J(u)$.

Сходимость и оценка погрешности метода

Метод ломаных сходится, что утверждается в следующей теореме.

Теорема 1. Пусть $J(u)$ – произвольная функция, удовлетворяющая на $[a, b]$ условию Липшица. Тогда последовательность $\{u_n\}$, полученная с помощью метода ломаных такова, что:

1. $\lim\limits_{n\to\infty} J(u_n)=\lim\limits_{n\to\infty} p_n(u_{n+1})=J^*=\inf\limits_{u\in[a,b]} J(u)$, причем справедлива оценка $0\le J(u_{n+1})-J_*\le J(u_{n+1})-p_n(u_{n+1}),\ n=0,1,\ldots$

2. Последовательность $\{u_n\}$ сходится к множеству U^* точек минимума $J(u)$ на отрезке $[a, b]$, т.е. $\lim\limits_{n\to\infty} \rho\left(u_n, U^*\right) = 0$.

Замечание. Заниженная оценка константы Липшица может привести к неправильному нахождению минимального значения и точки минимума функции. Завышенное значение константы Липшица может привести к увеличению количества итераций в методе ломаных.

Метод касательных

Метод касательных, как и метод ломаных, позволяет найти глобальный минимум функции на отрезке. Метод касательных применим для выпуклых на отрезке функций.

Определение 2. Функция $J(u)$ называется выпуклой (выпуклой вниз) на отрезке $[a, b]$, если для любых $u, v \in [a, b], u < v$, и для любого $\alpha \in [0, 1]$ справедливо соотношение:

$$J\left(\alpha u + (1-\alpha)v\right) \le \alpha J(u) + (1-\alpha)J(v).$$

Соотношение имеет следующий геометрический смысл: хорда, соединяющая точки $(u, J(u)), (v, J(v))$ должна быть не ниже значения функции на отрезке $[u, v]$ (Рис. 5 $\alpha = \frac{1}{2}$).

Рис. 5. Геометрический смысл выпуклой функции

Критерии выпуклости

Критерий выпуклости 1. Функция $J(u)$ является выпуклой на $[a, b]$, тогда и только тогда, когда справедливо неравенство:

$$\frac{J(u)-J(v)}{u-v} \le \frac{J(\omega)-J(v)}{\omega-v} \le \frac{J(\omega)-J(u)}{\omega-u},$$

для $\forall u, v, \omega \in [a,b] : a \le v \le u \le \omega \le b$.

Критерий выпуклости 2. Для того, чтобы дифференцируемая функция на отрезке *[a, b]* была выпуклой необходимо и достаточно, чтобы ее производная не убывала на *[a, b]*.

Критерий выпуклости 3. Для того, чтобы дважды дифференцируемая функция на отрезке *[a, b]* была выпуклой необходимо и достаточно, чтобы ее вторая производная была неотрицательной *(J''(u)≥0)*.

Алгоритм метода касательных

Введем обозначение: $g(u,v) = J(v) + J'(v)(u-v)$, где $v \in [a,b]$ - фиксировано, $\forall u \in [a,b]$ по свойству касательной: $J(u) \ge g(u,v)$ при $\forall u \in [a,b]$.

Выберем произвольную точку $u_0 \in [a,b]$. В данной точке построим касательную к графику функции:

$$g(u,u_0) = J(u_0) + J'(u_0)(u-u_0) = p_0(u).$$

После ее построения найдем точку u_1, удовлетворяющую соотношению:

$$p_0(u_1) = \min_{u \in [a,b]} p_0(u).$$

Нулевая итерация завершена.

На следующей итерации построим новую касательную, соответствующую точке u_1: $g(u,u_1) = J(u_1) + J'(u_1)(u-u_1)$, найдем функцию $p_1(u) = \max\{p_0(u), g(u,u_1)\}$ и затем - следующую точку u_2, удовлетворяющую соотношению: $p_1(u_2) = \min\limits_{u \in [a,b]} p_1(u)$.

Рассмотрим k-ю итерацию. К её началу известна точка u_k и функция $p_{k-1}(u)$. Строим касательную $g(u,u_k) = J(u_k) + J'(u_k)(u-u_k)$, затем - функцию $p_k(u) = \max\{p_{k-1}(u), g(u,u_k)\} = \max\limits_{i=\overline{0,k}}\{g(u,u_i)\}$ и находим следующую точку u_{k+1}, удовлетворяющую соотношению: $p_k(u_{k+1}) = \min\limits_{u \in [a,b]} p_k(u)$.

В итоге мы получаем ломанную линию $p_k(u)$, которая по мере увеличения числа k приближается к функции *J(u)* снизу (Рис. 6).

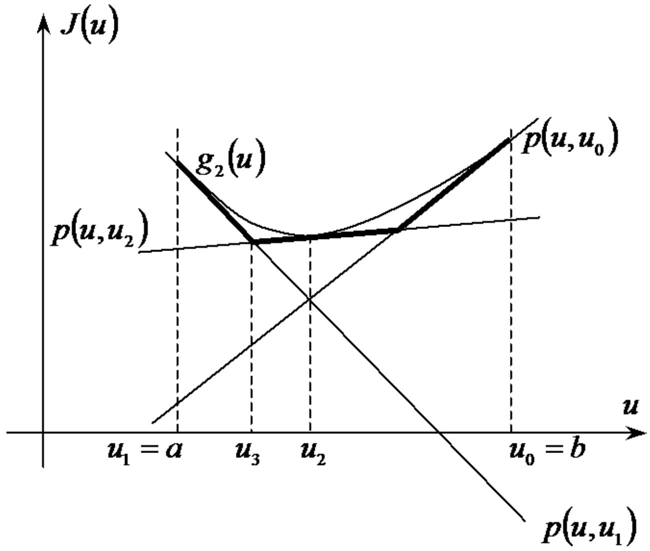

Рис. 6. Графическая интерпретация метода касательных

Итерации выполняются до тех пор, пока не будет достигнута заданная точность ε , аналогично методу ломаных.

Сходимость и оценка погрешности метода

Метод касательных сходится, что утверждается в следующей теореме.

Теорема 1. Пусть $J(u)$ на $[a,b]$ выпукла и дифференцируема, а последовательность $\{u_n\}$ получена описанным выше методом касательных, причем $u_n \notin U_*,\ \ n = 0,1,\ldots$. Тогда:

1. $\lim\limits_{n\to\infty} J(u_n) = \lim\limits_{n\to\infty} p_n(u_{n+1}) = J_*$ и справедлива оценка: $0 \le J(u_{n+1}) - J_* \le J(u_{n+1}) - p_n(u_{n+1}),\ n = 1, 2, \ldots$
2. $\lim\limits_{n\to\infty} \rho(u_n, U_*) = 0$.

Задание к лабораторной работе

Убедиться, что заданная в соответствии с вариантом функция удовлетворяет условиям Липшица (оценить константу Липшица аналитически) и является выпуклой (по критерию выпуклости) на заданном отрезке *[a, b]*. Разработать программу, реализующую метод ломанных и метод касательный. для заданной функции *J(u)* и отрезка *[a, b]]*, с точностью $\varepsilon=10^{-4}$. Выводимые значения для каждого метода:

- Точка минимума *u** и минимальное значение *J*;*

- Количество итераций и точность вычислений;
- Графики функций $J(u)$ и итоговой $p_n(u)$.

Указание. При проверке на выпуклость, если функция является выпуклой только на части отрезка $[a, b]$, то использовать в методе касательных только ту часть отрезка где функция является выпуклой. Если функция не является выпуклой на отрезке $[a, b]$, то необходимо взять ее с обратным знаком ($-J(u)$).

Варианты заданий

№ вар.	$J(u)$	$a;b$	α	β	γ	δ
1	$\operatorname{tg}\alpha u-\beta u$	-0.9; 0.5	1.5773	2.3041		
2		-0.5; 0.7	2.2982	3.2258		
3		-0.4; 0.2	3.7855	5.5300		
4		-0.15; 0.09	9.1484	13.3641		
5		-0.1; 0.2	5.9937	8.7558		
6		-0.15; 0.1	7.8864	11.5207		
7	$\ln(\alpha u)-\beta u+\gamma$	0.4; 3.0	7.6220	8.5900	0.5	
8		1.0; 3.0	6.0976	6.8720	1.0	
9		2.0; 7.0	4.5732	5.1540	1.5	
10		3.5; 7.5	3.9634	4.4868	2.0	
11		4.0; 8.0	3.0488	3.4360	2.5	
12		6.0; 9.9	1.5244	1.7180	3.0	
13	$\alpha\sin\beta u-\gamma u$	-0.3; 4.0	9.3300	6.9770	7.25	
14		-1.8; 6.2	7.6670	5.9830	6.00	
15		-2.1; 6.9	6.6700	5.3870	5.25	
16		-3.0; 7.5	5.6700	4.7940	4.50	
17		-3.5; 8.1	4.3300	4.0080	3.50	
18		-5.0; 9.7	2.6700	3.0440	2.25	
19	$\alpha e^{-\beta u}-u$	-2.1; 2.7	0.9737	0.5067		
20		-2.2; 2.3	0.9286	0.5185		
21		-2.8; 2.2	0.5458	0.5391		
22		-2.4; 2.4	0.7593	0.5683		
23		-1.9; 2.1	0.5909	0.6286		
24		-2.6; 2.0	0.4474	0.6909		
25		-2.2; 2.4	0.1667	0.8571		
26		-1.8; 2.3	0.7308	0.5778		
27		-2.5; 1.9	0.8330	0.5455		
28	$\alpha u^3+\beta u^2+\gamma u+\delta$	-17; 14	0.1697	-0.5693	-1.600	3.73
29		-11; 11	1.0390	-3.1450	-1.940	8.00
30		-3; 8	4.6839	-14.0400	-2.448	23.50

Содержание отчета

- Титульный лист;
- Цель работы;
- Постановка задачи;
- Краткая теоретическая часть;
- Вариант задания;
- Проверка функции на липшицевость и выпуклость;
- Листинг программы;
- Результаты работы программы:
- Вывод.

Контрольные вопросы

1. Дайте определение липшицевой функции.
2. Геометрическая интерпретация липшицевости функции.
3. Будет ли ломанная линия, получаемая при помощи метода ломанных, пересекать график функции $J(u)$ и почему.
4. Из чего следует соотношение для численной оценки константы Липшица?
5. Описание алгоритма метода ломаных.
6. Графическая интерпретация метода ломаных.
7. Условие сходимости метода ломаных.
8. К чему может привести заниженная и завышенная оценка константы Липшица и почему?
9. Дайте определение выпуклой функции.
10. Геометрическая интерпретация выпуклой функции.
11. Критерии выпуклости.
12. Описание алгоритма метода касательных.
13. Графическая интерпретация метода касательных.
14. Условие сходимости метода касательных.
15. Всякая ли липшицева функция является выпуклой, и наоборот, всякая ли выпуклая функция является липшицевой и почему?
16. Сравните метод ломаных и метод касательных по скорости сходимости.

Лабораторная работа №3. Аналитические методы решения задач оптимизации функций многих переменных

Постановка задачи

Решить задачи безусловной и условной оптимизации функций многих переменных классическим методом и методом множителей Лагранжа.

Краткая теоретическая часть

Классический метод основан на том, что в точке экстремума градиент функции равен нулю.

Определение 1. Пусть Функция *J(u)* определенная в ε-окрестности точки *u* $O_\varepsilon(u)=\left\{\upsilon:\upsilon\in E^n,\|u-\upsilon\|<\varepsilon\right\}$. Тогда *J(u)* – дифференцируема в точке *u*, если $\exists$ вектор $J'(u)\in E^n$, называемый градиентом этой функции в точке *u*, такой, что для $\forall$ h: $\|h\|<\varepsilon$ справедливо:

$$J(u+h)-J(u)=(J'(u),h)+\alpha(u,h),$$

где $\lim\limits_{\|h\|\to 0}\dfrac{\alpha(u,h)}{\|h\|}=0$.

Градиент, это вектор состоящий из частных производных:

$$J'(u)=(J'_{u^1}(u),\ldots,J'_{u^n}(u))^T.$$

$$J'_{u^i}(u)=\frac{\partial J(u)}{\partial u^i}=\lim_{a\to 0}\frac{J(u+ae_i)-J(u)}{a},$$

где e_i - базисный вектор.

Определение 2. Пусть функция *J(u)* определена в некоторой ε-окрестности точки *u*. Тогда *J(u)* дважды дифференцируема в точке *u*, если кроме градиента $J'(u)$ $\exists$ матрица $J''(u)$ такая что: $\forall\, h\in E^n:\ \|h\|<\varepsilon$ справедливо:

$$J(u+h)-J(u)=(J'(u),h)+\frac{1}{2}(J(u)h,h)+\beta(u,h),$$

где $\beta:\lim\limits_{\|h\|\to 0}\dfrac{\beta(u,h)}{\|h\|^2}=0$.

Матрица вторых является квадратной матрицей порядка *n*:

$$J''(u)=\begin{vmatrix} \dfrac{\partial^2 J(u)}{(\partial u^1)^2}, & \dots & , & \dfrac{\partial^2 J(u)}{\partial u^1 \partial u^n} \\ \dfrac{\partial^2 J(u)}{\partial u^2 \partial u^1}, & \dots & , & \dfrac{\partial^2 J(u)}{\partial u^2 \partial u^n} \\ \dots & \dots & \dots & \dots \\ \dfrac{\partial^2 J(u)}{\partial u^n \partial u^1}, & \dots & , & \dfrac{\partial^2 J(u)}{(\partial u^n)^2} \end{vmatrix}.$$

Рассмотрим классический метод для задачи на безусловный экстремум:

$$J(u) \to extr,\ U \equiv E^n.$$

Классический метод основывается на следующих утверждениях…

Утверждение 1. Если *J(u)* дифференцируема в точке $u^* \in U^*$, то $J'(u^*)=0$.

Утверждение 2. Пусть точка u^* – стационарная точка, и функция *J(u)* дважды дифференцируема в окрестности точки u^*. Тогда:

1. Если квадратичная форма $(J''(u^*)h,h)$, положительно определена для $\forall h \in E^n$, то u^* является точкой минимума.
2. Если квадратичная форма $(J''(u^*)h,h)$, отрицательно определена для $\forall h \in E^n$, то u^* является точкой максимума.
3. Если квадратичная форма $(J''(u^*)h,h)$, знакопеременная для $\forall h \in E^n$, то u^* - не определена (не является экстремумом).

Утверждение 3 (критерий Сильвестра). Квадратичная форма: $(J''(u^*)h,h))$, $\forall h \in E^n:\ h \neq 0$ является положительно определенной, если положительны все главные миноры матрицы *J''(u*)*, отрицательно определенной, если знаки главных миноров матрицы *J''(u*)* чередуется и является знакопеременной если не выполняется ни одно из предыдущих условий.

Таким обозом, для того чтобы решить задачу на безусловный экстремум классическим методом необходимо:

1. Найти все частные производные, приравнять их к нулю и, решив полученную систему уравнений, найти стационарные точки.

2. Найти матрицу вторых производных для каждой стационарной точки.
3. Исследовать матрицу вторых производных в каждой стационарной точке на положительно/отрицательно определенность по критерию Сильвестра для определения типа экстремума.

Классический метод для задачи на условный экстремум

Рассмотрим классический метод для задачи на условный экстремум. Классический метод позволяет решать задачи на условный экстремум только при наличии ограничений в виде равенств:

$$J(u) \to extr,\ u \in U,$$

где $U = \left\{ u : U \subset E^n,\ g_i(u) = 0,\ i = \overline{1, m} \right\}$, $m<n$.

Множество, на котором ищется решение, описывается системой ограничений в виде равенств:

$$\begin{cases} g_1(u) = 0 \\ g_2(u) = 0 \\ \quad \dots \\ g_m(u) = 0 \end{cases} \qquad (*)$$

Если из системы ограничений $(*)$ можно перейти к эквивалентной системе вида:

$$\begin{cases} u_1 = f_1(u_{m+1}, \dots, u_n) \\ u_2 = f_2(u_{m+1}, \dots, u_n) \\ \dots \\ u_m = f_m(u_{m+1}, \dots, u_n) \end{cases} \qquad (**)$$

выразив из системы ограничений $(*)$ первые m переменных через последующие, то исходную задачу на условный экстремум можно свести к задаче на безусловный экстремум для вспомогательной функции:

$$\varphi(u_{m+1}, \dots, u_n) =$$
$$= J\left(f_1(u_{m+1}, \dots, u_n), f_2(u_{m+1}, \dots, u_n), \dots, f_m(u_{m+1}, \dots, u_n), u_{m+1}, \dots, u_n \right)$$

,

от переменных $\left(u_{m+1}, \ldots, u_n\right) \in E^{n-m}$, а как решать задачу на безусловный рассматривалось ранее.

Таким обозом, для того чтобы решить задачу на условный экстремум классическим методом необходимо:

1. Выразить первые переменные через последующие (прейти от системы ограничений в виде (*) к виду (**)).
2. Подставить получившиеся выражения исходную функцию и перейти к вспомогательной задаче безусловной оптимизации.
3. Решить вспомогательную задачу ранее описанным методом.
4. Сделать обратную подстановку, получить стационарные точки.
5. Исследовать матрицу вторых производных в каждой стационарной точке на положительно/отрицательно определенность по критерию Сильвестра для определения типа экстремума.

Замечание. Классический метод не позволяет решать задачи оптимизации при наличии ограничений в виде неравенств в связи с этим рассмотрим более общий подход.

Метод множителей Лагранжа для задач с ограничениями типа равенств

Метод множителей Лагранжа можно применять для задач условной оптимизации, как при наличии ограничений в виде равенств, так и при наличии ограничений в виде неравенств, а также при наличии обеих видов ограничений.

Рассмотрим более простую задачу при наличии ограничений только в виде равенств:

$$J(u) \to extr,\ u \in U \subset E^n$$

$$U = \left\{u : \left(u_1, u_2, ..., u_n\right)^T \in E^n,\ \ g_i(u) = 0,\ i = \overline{1,m}\right\}.$$

Для решения данной задачи составляется функция Лагранжа:

$$L(u,\lambda) = \lambda_0 J(u) + \lambda_1 g_1(u) + \ldots + \lambda_m g_m(u) = \lambda_0 J(u) + \sum_{i=1}^{m} \lambda_i g_i(u) \qquad (1)$$

где переменные $\lambda_0, \lambda_1, \ldots, \lambda_m$ – называют множителями Лагранжа.

Функция Лагранжа $L(u,\lambda)$ зависит от переменных $(u,\lambda) = (u_1, \ldots, u_n, \lambda_0, \lambda_1, \ldots, \lambda_m)^T \in E^{n+m+1}$.

Утверждение 1. Пусть u^* – точка экстремума, тогда существует такой набор множителей Лагранжа $\lambda^* = (\lambda_0^*, \lambda_1^*, \ldots, \lambda_m^*)^T \neq 0$, такой что:

$$\left.\frac{\partial L(u, \lambda^*)}{\partial u_i}\right|_{u=u^*} = 0, \quad i = \overline{1, n} \qquad (2)$$

Здесь соотношение (2) представляет собой систему из n уравнений с $n+m+1$ неизвестными (n от переменной u, и $m+1$ от множителей Лагранжа λ). Решить такую систему нельзя, необходимо еще $m+1$ уравнение.

Если u^* является точкой экстремума, то для нее будут справедливы ограничения:

$$g_i(u^*) = 0, \quad i = \overline{1, m} \qquad (3)$$

Объединение соотношений (2)(3) задают систему уравнений, в которой $n+m$ уравнений с $n+m+1$ неизвестными. Решить такую систему нельзя, необходимо еще одно уравнение.

Если точка (υ, λ) удовлетворяет набору уравнений (2)(3), то в этом случаи точка $(\upsilon, \alpha\lambda)$, где α произвольное число $\alpha \neq 0$, так же будет удовлетворять набору уравнений (2)(3). В связи с этим набор множителей Лагранжа λ будем подчинять условию нормировки:

$$\|\lambda\|^2 = \sum_0^m \lambda_i^2 = 1, \; \lambda_0 > 0 \qquad (4)$$

Объединяя соотношения (2)(3)(4) получаем набор из $n+s+1$ уравнений с $n+s+1$ неизвестными, решая их, найдем точки (u^*, λ^*), которые подозрительны на экстремум, которые надо исследовать.

Утверждение 2. Пусть точка (u^*, λ^*) является решением уравнение (2)(3)(4), функция Лагранжа дважды дифференцируема в точке u^*, $\lambda_0^* > 0$, тогда если:

- квадратичная форма $\left(L_{uu}(u^*, \lambda^*)h, h\right)$ $\forall h \in E^n : h \neq 0$ положительно определена, то в точке u^* достигается минимум функции $J(u)$ при ограничениях (3);

- квадратичная форма $\left(L_{uu}(u^*,\lambda^*)h,h\right)$ $\forall h \in E^n : h \neq 0$ отрицательно определена, то в точке u^* достигается максимум функции $J(u)$ при ограничениях (3);
- квадратичная форма $\left(L_{uu}(u^*,\lambda^*)h,h\right)$ $\forall h \in E^n : h \neq 0$ знакопеременна, то в точке u^* экстремума нет.

Замечание. Если уравнения (2)(3)(4) могут иметь решение (u^*,λ^*), такое что $\lambda_0^* \neq 0$, то такая задача оптимизации называется невырожденной. В невырожденной задаче условие нормировки (4) можно заменить более простым условием $\lambda_0=1$, следовательно, для невырожденной задачи условие (4) примет вид:

$$\lambda_0=1, \qquad (4')$$

т.е. если задача является не вырожденной, то для поиска точек подозрительных на экстремум необходимо решать уравнения (2)(3)(4'), если задача не является невырожденной,- то решаются уравнения (2)(3)(4).

Проверка задачи на невырожденность осуществляется методом от противного, для этого предполагается, что уравнения (2)(3)(4) могут иметь решения при $\lambda_0=0$ и ищутся несоответствия.

Таким обозом, для того чтобы решить задачу на условный экстремум методом множителей Лагранжа при наличии ограничений в виде равенств необходимо:

1. Составить функцию Лагранжа (1).
2. Найти частные производные функции Лагранжа по компонентам вектора u и приравнять их к нулю, тем самым получив набор уравнений (2).
3. Добавить к уравнениям (2) ограничения (3).
4. Добавить к уравнениям (2)(3) условие нормировки (4).
5. Проверить задачу на невырожденность методом от противного, для этого предположить сто уравнения (2)(3)(4) могут иметь решение при $\lambda_0=0$ и искать несоответствия.
6. Если задача является невырожденной, то решить уравнения (2)(3)(4'), если не является невырожденной, то решить уравнения (2)(3)(4) и найти точки подозрительные на экстремум.
7. Исследовать матрицу вторых производных функции Лагранжа по компонентам вектора u в каждой стационарной точке на

положительно/отрицательно определенность по критерию Сильвестра для определения типа экстремума.

8. При записи ответа использовать только компоненты вектора u.

Метод множителей Лагранжа для задач с ограничениями типа неравенств

Рассмотрим метод множителей Лагранжа для решения задач оптимизации при наличии ограничений, как в виде равенств, так и в виде неравенств:

$$J(u) \to extr,\ u \in U$$

$$U = \{u \in E^n;\ g_i(u) = 0;\ i = \overline{1,m};\ h_j(u) \le 0;\ j = \overline{1,s}\}$$

Для решения данной задачи ограничения типа неравенств $h_j(u) \le 0$, $j = \overline{1,s}$ сводятся к ограничениям типа равенств, путем ввода новых переменных $\upsilon_j,\ j = \overline{1,s}$ следующим образом: $h_j(u) + (\upsilon_j)^2 = 0$ и, решение задач с неравенствами сводится к решению задач с равенствами, с переменными: $(u,\upsilon) = (u_1, \ldots, u_n, \upsilon_1, \ldots, \upsilon_s)^T$.

Таким образом, функция Лагранжа примет вид:

$$L(u,\upsilon,\lambda,\mu) = \lambda_0 J(u) + \sum_{i=1}^{m} \lambda_i g_i(u) + \sum_{j=1}^{s} \mu_j [h_j(u) + (\upsilon_j)^2]. \qquad (5)$$

Функция Лагранжа $L(u,\upsilon,\lambda,\mu)$ зависит от переменных $(u,\upsilon,\lambda,\mu) = (u_1, \ldots, u_n, \upsilon_1, \ldots, \upsilon_s \lambda_0, \lambda_1, \ldots, \lambda_m, \mu_1, \ldots, \mu_s)^T \in E^{n+s+m+1+s}$.

Таким обозом, для того чтобы решить задачу на условный экстремум методом множителей Лагранжа при наличии ограничений в виде неравенств необходимо:

1. Избавиться от ограничений в виде неравенств, заменив их ограничениями в в виде равенств путем добавления новых переменных υ.
2. Составить функцию Лагранжа (5).
3. Найти частные производные функции Лагранжа по компонентам вектора u и новым переменным υ, приравнять их к нулю, тем самым получив набор уравнений (2).
4. Добавить к уравнениям (2) ограничения (3).
5. Добавить к уравнениям (2)(3) условие нормировки (4).

6. Проверить задачу на невырожденность методом от противного, для этого предположить сто уравнения (2)(3)(4) могут иметь решение при $\lambda_0=0$ и искать несоответствия.
7. Если задача является невырожденной, то решить уравнения (2)(3)(4'), если не является невырожденной, то решить уравнения (2)(3)(4) и найти точки подозрительные на экстремум.
8. Исследовать матрицу вторых производных функции Лагранжа по компонентам вектора u и новым переменным υ в каждой стационарной точке на положительно/отрицательно определенность по критерию Сильвестра для определения типа экстремума.
9. При записи ответа использовать только компоненты вектора u.

Задание к лабораторной работе

1. Найти экстремумы функции $J(u)= au_1^2 +b\, u_2^2+ cu_1u_2, u=(u_1, u_2)^T\epsilon E^2$ классическим методом.
2. Найти экстремумы функции $J(u)$ из задания 1 классическим методом и методом множителей Лагранжа при наличии ограничения $(d,u)+e=0$.
3. Найти экстремумы функции $J(u)$ из задания 1 методом множителей Лагранжа при наличии ограничения $u\geq 0$.

Варианты заданий

№ вар.	*a*	*b*	*c*		*d*	*e*
			c_1	c_2		
1	-4	-5	7	2	6	-6
2	2	7	-5	-7	3	-3
3	-5	-7	8	4	-4	1
4	6	-7	-3	-9	3	2
5	-7	1	9	5	-6	-3
6	3	3	7	-1	3	2
7	3	-2	-8	-5	5	-5
8	-9	2	3	0	-8	1
9	4	1	2	4	-4	-9
10	1	6	9	-1	0	-2
11	-4	6	-6	-7	2	-3
12	-4	-7	7	5	9	8
13	-7	7	-4	8	-8	-6
14	-1	-6	-2	-5	8	-9
15	-9	-8	9	6	3	-1

№ вар.	*a*	*b*	*c*		*d*	*e*
			c_1	c_2		
16	1	9	-2	-9	-3	-5
17	4	-6	-2	-9	6	0
18	7	-3	8	8	-5	7
19	2	3	-2	2	-5	2
20	4	1	-1	2	8	0
21	-9	-8	-3	-8	-4	2
22	-3	5	6	-2	4	9
23	-5	-6	9	-9	-1	-3
24	2	-6	9	2	-4	-9
25	3	8	1	-1	-4	-5
26	-4	9	0	-6	1	9
27	-7	-8	-2	7	-8	-7
28	1	-2	5	-2	3	-5
29	8	5	3	1	4	-5
30	-9	-9	-4	-7	-5	9

Содержание отчета

- Титульный лист;
- Цель работы;
- Постановка задачи;
- Краткая теоретическая часть;
- Вариант задания;
- Аналитическое решение задач, с ответом.
- Вывод.

Контрольные вопросы

1. Что такое градиент функции многих переменных, какой его геометрический смысл?
2. Что позволяет определить матрица вторых производный в стационарной точке. ?
3. Что позволяет определить критерий Сильвестра?
4. Какие задачи на условный экстремум можно решить классическим метом.
5. Какие задачи на условный экстремум можно решить методом множителей Лагранжа.
6. Как проверить задачу оптимизации на невырожденность?
7. Как избавиться от ограничений в виде неравенств в методе множителей Лагранжа.

Лабораторная работа №4. Минимизация функции многих переменных градиентными методами

Постановка задачи

Найти при помощи градиентных методов наискорейшего спуска и условного градиента точку минимума дифференцируемой функции многих переменных $J(u)$, $u \in U$, $U \subseteq E^n$, с заданной точностью ε.

Краткая теоретическая часть

Общая идея градиентных методов

В общем виде градиентные методы заключается в построении минимизирующей последовательности $\{u_k\}$ по правилу:

$$u_{k+1} = u_k - \alpha_k J'(u_k), \alpha_k > 0, k = 0,1,2,\ldots$$

где $J^{/}(u_k)$ – градиент функции $J(u)$ в точке u_k, а число α_k должно быть положительным и называется шагом. Именно способом выбора α_k один градиентный метод отличается от другого.

Метод наискорейшего спуска

Метод наискорейшего спуска является одной из разновидностей семейства градиентных методов. Метод наискорейшего спуска позволяет решать только задачи безусловной оптимизации $J(u) \to min$, $u \in U$, $U \equiv E^n$.

В случае градиентного метода наискорейшего спуска шаг α_k определяется из следующего соотношения:

$$f_k(\alpha_k) = \min_{\alpha \ge 0} f_k(\alpha) = f_{k*}, \alpha_k > 0,$$

где $f_k(\alpha) = J\left(u_k - \alpha J^{/}(u_k)\right), \alpha \ge 0$

Это означает, что для нахождения значения α_k на каждом шаге метода наискорейшего спуска приходится решать вспомогательную задачу минимизации функции одной переменной. Такая задача не всегда решается просто. Тем не менее, для некоторых классов функций $J(u)$ значение шага α_k удается найти аналитически.

В настоящей лабораторной работе рассматривается один из этих классов, а именно:

$$J(u) = \frac{1}{2}(Au,u) - (b,u).$$

Здесь A симметричная, положительно определенная матрица порядка n, b - вектор из пространства E^n.

Градиент функции имеет вид $J'(u)=Au-b$. По общей формуле градиентных методов получаем итерационную формулу:

$$u_{k+1} = u_k - \alpha_k (Au_k - b), k = 0,1,\ldots,$$

причем шаг α_k вычисляется по формуле:

$$\alpha_k = \frac{|Au_k - b|^2}{(A(Au_k - b), Au_k - b)} > 0 .$$

В качестве начального приближения u_0 можно выбрать любую точку из множества U на котором ищется решение, в нашем случаи всего пространства E^n.

В качестве критериев окончания итераций можно использовать следующие условия:

- $|J(u_{k+1}) - J(u_k)| < \varepsilon$, для задач типа I;
- $|u_{k+1} - u_k| < \varepsilon$, для задач типа II;
- $|J'(u)| < \varepsilon$, для задач типа I и II.

Метод условного градиента

Метод наискорейшего спуска не пригоден для задач условной оптимизации, когда на переменную u накладываются ограничения вида $u \in U \neq E^n$. Одной из модификаций градиентных методов, позволяющих решать подобные задачи, является метод условного градиента.

Он может быть использован для задач следующего вида:

$$J(u) \to min, \ u \in U, \ U \neq E^n.$$

где U - выпуклое замкнутое ограниченное множество из E^n, функция $J(u) \in C^1(U)$.

Описание метода

Пусть известно k-е приближение $u_k \in U, k \geq 0$. Приращение функции $J(u)$ в точке u_k можно представить в виде:

$$J(u) - J(u_k) = (J'(u_k), u - u_k) + o(|u - u_k|).$$

Возьмем главную линейную часть этого приращения:

$$J_k(u) = (J'(u_k), u - u_k),$$

и определим вспомогательное приближение u_k из условий

$$\bar{u}_k \in U,\; J_k(\bar{u}_k) = \min_U \left(J'(u_k), \bar{u}_k - u_k\right).$$

Так как множество U замкнуто и ограничено, а линейная функция $J_k(u)$ непрерывна, то точка $\bar{u}_k$ всегда существует. Если функция достигает своей нижней грани на U более чем в одной точке, то в качестве $\bar{u}_k$ возьмем любую из них.

Рассмотрим в качестве примера нахождение вспомогательного приближения $\bar{u}_k$ в случае, когда множество U представляет собой n-мерный гиперпараллелепипед: $U = \{u = (u_1, \ldots, u_n) : \alpha_i \le u_i \le \beta_i, i = \overline{1,n}\}$,

В этом случае функция $J_k(u\,)$примет вид: $J_k(u) = \sum_{i=1}^{n} J_{u_i}(u_k)(u_i - u_{i_k})$,

или $J_k(u) = \sum_{i=1}^{n} J_{u_i}(u_k) u_i$, которая является линейной и очевидно, что она достигает своей минимума на множестве U в точке $\bar{u}_k = (\bar{u}_{1_k}, \ldots, \bar{u}_{n_k})$, где

$$\bar{u}_{i_k} = \begin{cases} \alpha_i, \textit{ если } J_{u_i}(u_k) \ge 0, \\ \beta_i, \textit{ если } J_{u_i}(u_k) < 0. \end{cases}$$

Допустим, что вспомогательное приближение $\bar{u}_k$ найдена. Тогда $k+1$-е приближение будем искать в виде

$$u_{k+1} = u_k + \alpha_k(\bar{u}_k - u_k), \text{ где } \alpha_k : 0 \le \alpha_k \le 1.$$

В силу выпуклости множества U всегда $u_{k+1} \in U$.

Теперь рассмотрим один из способов выбора шагаα_k. Данная величина может быть определена из следующих условий:

$$\alpha_k \in [0,1],\; f_k(\alpha_k) = \min_{\alpha \in [0,1]} f_k(\alpha) = f_{k_*},\; f_k(\alpha) = J(u_k + \alpha(\bar{u}_k - u_k)).$$

В частности, для функции $J(u) = \frac{1}{2}(Au, u) - (b, u)$, где A - симметричная положительно определенная матрица порядка n, $b \in E^n$ шаг α_k, вычисляется следующим образом:

1. Вначале вычисляется промежуточная величина:

$$\alpha_k^* = -\frac{\left(J'(u_k), \bar{u}_k - u_k\right)}{\left(A(\bar{u}_k - u_k), \bar{u}_k - u_k\right)}$$

2. Затем находится искомое значение α_k: $\alpha_k = \min\{\max\{0, \alpha_k^*\}, 1\}$.

В качестве начального приближения u_0 можно выбирать любую точку из множества U. В качестве критериев окончания счета можно использовать те же условия, что и в методе наискорейшего спуска.

Задание к лабораторной работе

Разработать программу нахождения минимума функции $J(u)=\frac{1}{2}(Au,u)-(b,u)$, где $u=\begin{pmatrix}u_1\\u_2\end{pmatrix}$, A - симметричная положительно определенная матрица второго порядка, $b\in E^2$, зависящей от двух переменных, при помощи метода наискорейшего спуска и метода условного градиента, с точностью $\varepsilon=10^{-3}$.

Для метода условного градиента в качестве множества U взять прямоугольник:

$$U=\left\{u=\left(u^1,u^2\right):\alpha_i\le u^i\le\beta_i, i=1,2\right\}.$$

Начальное приближение u_0 выбирается произвольно пользователем.

Выводимые значения для каждого метода:

- Координаты точки минимума и минимальное значение J^*;
- Количество итераций и точность вычислений.

Варианты заданий

№ вар.	A, b, $\alpha_1;\beta_1;\alpha_2;\beta_2$	№ вар.	A, b, $\alpha_1;\beta_1;\alpha_2;\beta_2$
1	$\begin{pmatrix}3&2\\2&5\end{pmatrix},\begin{pmatrix}-3\\5\end{pmatrix}$, 0; 2; 3; 7	16	$\begin{pmatrix}11&7\\7&10\end{pmatrix},\begin{pmatrix}5\\6\end{pmatrix}$**, 2; 4; 3; 4**
2	$\begin{pmatrix}6&-3\\-3&2\end{pmatrix},\begin{pmatrix}4\\-6\end{pmatrix}$, 0; 2; 3	17	$\begin{pmatrix}10&-8\\-8&12\end{pmatrix},\begin{pmatrix}-9\\3\end{pmatrix}$, -1; 1; 0; 2
3	$\begin{pmatrix}7&-1\\-1&4\end{pmatrix},\begin{pmatrix}7\\-2\end{pmatrix}$, 1; 3; -4; -1	18	$\begin{pmatrix}8&3\\3&3\end{pmatrix},\begin{pmatrix}7\\-11\end{pmatrix}$, 0; 2; -4; -1
4	$\begin{pmatrix}2&1\\1&3\end{pmatrix},\begin{pmatrix}14\\3\end{pmatrix}$, 0; 5; 0; 3	19	$\begin{pmatrix}9&5\\5&10\end{pmatrix},\begin{pmatrix}7\\2\end{pmatrix}$**, 1; 3; -3; 1**
5	$\begin{pmatrix}9&2\\2&1\end{pmatrix},\begin{pmatrix}9\\5\end{pmatrix}$, 3; 7; 1; 4	20	$\begin{pmatrix}4&-5\\-5&9\end{pmatrix},\begin{pmatrix}3\\3\end{pmatrix}$, 0; 2; 0; 2

№ вар.	A, b, $\alpha_1; \beta_1; \alpha_2; \beta_2$	№ вар.	A, b, $\alpha_1; \beta_1; \alpha_2; \beta_2$
6	$\begin{pmatrix} 8 & -3 \\ -3 & 4 \end{pmatrix}, \begin{pmatrix} 7 \\ 6 \end{pmatrix}$ -1; 1; 0; 2	21	$\begin{pmatrix} 15 & 2 \\ 2 & 1 \end{pmatrix}, \begin{pmatrix} -5 \\ -5 \end{pmatrix}$, -1; 0; -4; -2
7	$\begin{pmatrix} 5 & 2 \\ 2 & 5 \end{pmatrix}, \begin{pmatrix} -3 \\ -7 \end{pmatrix}$, 0; 1; 0; 2	22	$\begin{pmatrix} 1 & 3 \\ 3 & 17 \end{pmatrix}, \begin{pmatrix} 3 \\ 4 \end{pmatrix}$, 1; 2; 0; 2
8	$\begin{pmatrix} 6 & -3 \\ -3 & 6 \end{pmatrix}, \begin{pmatrix} -5 \\ 4 \end{pmatrix}$, 1; 3; 1; 3	23	$\begin{pmatrix} 2 & -4 \\ -4 & 14 \end{pmatrix}, \begin{pmatrix} -5 \\ 11 \end{pmatrix}$, 6; 9; 4; 8
9	$\begin{pmatrix} 11 & 4 \\ 4 & 2 \end{pmatrix}, \begin{pmatrix} 6 \\ 9 \end{pmatrix}$, -3; 1; 0; 7	24	$\begin{pmatrix} 11 & 6 \\ 6 & 11 \end{pmatrix}, \begin{pmatrix} -9 \\ 5 \end{pmatrix}$, -1; 1; 2; 3
10	$\begin{pmatrix} 7 & -5 \\ -5 & 12 \end{pmatrix}, \begin{pmatrix} 5 \\ 5 \end{pmatrix}$, 0; 1; 0; 1	25	$\begin{pmatrix} 9 & 4 \\ 4 & 7 \end{pmatrix}, \begin{pmatrix} -7 \\ 3 \end{pmatrix}$, 1; 5; 0;1
11	$\begin{pmatrix} 13 & 3 \\ 3 & 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 12 \end{pmatrix}$, -2, 11; 20; 30	26	$\begin{pmatrix} 5 & 2 \\ 2 & 7 \end{pmatrix}, \begin{pmatrix} 4 \\ 3 \end{pmatrix}$, 1; 3; 2; 3
12	$\begin{pmatrix} 5 & 2 \\ 2 & 2 \end{pmatrix}, \begin{pmatrix} -4 \\ -3 \end{pmatrix}$, 1; 3; 0; 2	27	$\begin{pmatrix} 3 & 12 \\ 12 & 3 \end{pmatrix}, \begin{pmatrix} -1 \\ 2 \end{pmatrix}$, 1; 3; -3; 1
13	$\begin{pmatrix} 9 & 5 \\ 5 & 7 \end{pmatrix}, \begin{pmatrix} 3 \\ 9 \end{pmatrix}$, 0; 1; 0; 1	28	$\begin{pmatrix} 6 & 4 \\ 4 & 1 \end{pmatrix}, \begin{pmatrix} 0 \\ -7 \end{pmatrix}$, -2; 0; 2; 4
14	$\begin{pmatrix} 7 & 2 \\ 2 & 7 \end{pmatrix}, \begin{pmatrix} 13 \\ 9 \end{pmatrix}$, 2; 6; 1; 5	29	$\begin{pmatrix} 3 & 2 \\ 2 & 8 \end{pmatrix}, \begin{pmatrix} 2 \\ 12 \end{pmatrix}$, -3; -1; 2; 3
15	$\begin{pmatrix} 3 & 4 \\ 4 & 6 \end{pmatrix}, \begin{pmatrix} -1 \\ 15 \end{pmatrix}$, -20; 0; -10; 20	30	$\begin{pmatrix} 9 & 4 \\ 4 & 4 \end{pmatrix}, \begin{pmatrix} 6 \\ 12 \end{pmatrix}$, -5; -2; 0; 2

Содержание отчета

- Титульный лист установленного образца;
- Цель работы;
- Постановка задачи;
- Краткая теоретическая часть;
- Вариант задания;
- График функции (поверхность) и линии уровня функции на заданной области;

- Листинг программы;
- Результаты работы программы:
- Вывод.

Контрольные вопросы

1. Дайте определение градиента функции.
2. В какую сторону направлен вектор градиента.
3. Общая идея всех градиентных методов.
4. Чем отличается метод условного градиента от метода наискорейшего спуска?
5. Условие выбора шага в методе наискорейшего спуска.
6. Геометрический смысл метода наискорейшего спуска.
7. Для какого класса задач используется метод условного градиента?
8. Дайте определения выпуклого множества.
9. Что такое вспомогательное приближение в методе условного градиента и для чего оно используется?
10. Почему точка u_{k+1}, в методе условного градиента, не выходит за пределы выпуклого множества?
11. Геометрический смысл метода условного градиента.

Лабораторная работа №5. Минимизация функции многих переменных методом покоординатного спуска

Постановка задачи

Требуется решить задачу условной оптимизации методом покоординатного спуска:

$$J(u) \to \min, u \in U,$$

где $J(u) \in C^1(U)$, $U = \{u = (u^1, \ldots, u^n): a_i \le u^i \le b_i, i = \overline{1,n}\}$, $a_i < b_i$ - заданные числа, с заданной точностью ε.

Краткая теоретическая часть

Будем использовать набор векторов: $e_i = (0, \ldots, 1, \ldots, 0)$, с i-й координатой, равной 1, $i = \overline{1,n}$.

Пусть u_0 – начальное приближение, $\alpha_0 > 0$ – параметр метода. Допустим, что уже известны $u_k \in U, \alpha_k > 0,$ для $k \ge 0$.

Пусть $p_k = e_{i_k}$, где

$$i_k = k - n\left[\frac{k}{n}\right] + 1,\ k = 0,1,2,\ldots \qquad (*)$$

Здесь $\left[\frac{k}{n}\right]$ – целая часть от $\frac{k}{n}$.

Соотношение (*) обеспечивает циклический перебор координатных векторов $e_1,\ldots,e_n$.

Пусть $u = u_k + \alpha_k p_k$, тогда если $J(u_k + \alpha_k p_k) < J(u_k)$ и $u \in U$, то $u_{k+1} = u_k + \alpha_k p_k, \alpha_{k+1} = \alpha_k$, иначе $u = u_k - \alpha_k p_k$. Если $J(u_k - \alpha_k p_k) < J(u_k)$ и $u \in U$, то $u_{k+1} = u_k - \alpha_k p_k, \alpha_{k+1} = \alpha_k$.

Назовем k-ю итерацию удачной, если удалось перейти в сторону уменьшения функции в k-ом координатном направлении, в противном случаи итерация называется неудачной.

Если k-я итерация неудачна, то:

$$u_{k+1} = u_k,\ \alpha_k = \begin{cases} \lambda\alpha_k,\ u_k = u_{k-n+1}, \\ \alpha_k, \text{ иначе.} \end{cases}$$

Здесь λ, $0 < \lambda < 1$ - параметр метода, обычно $\lambda = 1/2$. Данное соотношение означают, что если при переборе всех направлений (координатных осей $e_1,\ldots,e_n$) с шагом α_k реализовалась хотя бы одна удачная итерация, то шаг α_k не дробится и сохраняется, по крайней мере, в течение следующего цикла из n итераций.

В качестве критериев окончания итераций можно использовать следующие условия:

- $|J(u_{k+1}) - J(u_k)| < \varepsilon$, для задач типа I;
- $|u_{k+1} - u_k| < \varepsilon$, для задач типа II;
- $\alpha_k < \varepsilon$, для задач типа I и II.

Сходимость метода и оценка погрешности

Теорема 1. Пусть функция $J(u)$ выпукла на E^n и принадлежит классу $C^1\lfloor E^n \rfloor$, а начальное приближение u_0 таково, что множество $M(u_0) = \{u \in E^n, J(u) \le J(u_0)\}$ - ограничено. Тогда последовательность $\{u_k\}$, получаемая методом покоординатного спуска, минимизирует $J(u)$ на E^n и сходится к множеству U^*.

Задание к лабораторной работе

Разработать программу, реализующую метод покоординатного спуска для функции $J(u_1,u_2)=au_1+bu_2+e^{cu_1^2+du_2^2}$ (значения a, b, c, d определяются вариантами заданий) и $U=\{(u_1,u_2)\in E^2: 0\le u_1\le 4;$ $-1\le u_2\le 6\}$, $\lambda=1/2$, $\alpha_0=1$. Программа должна выводить: координаты точки минимума u_*; значение функции в данной точке $J(u_*)$; количество итераций, после которых достигается точность $\varepsilon=10^{-3}$.

Выводимые значения для каждого метода:

- Координаты точки минимума и минимальное значение *J*;*
- Количество итераций и количество обращений к функции.
- Точность вычислений.

Варианты заданий

№ вар.	a	b	c	d	№ вар.	a	b	c	d
1	1	-1.4	0.01	0.11	16	16	0.0	1.99	0.26
2	2	-1.3	0.04	0.12	17	17	0.1	2.56	0.27
3	3	-1.2	0.02	0.13	18	18	0.2	2.89	0.28
4	4	-1.1	0.16	0.14	19	19	0.3	3.24	0.29
5	5	-1.0	0.25	0.15	20	20	0.4	3.81	0.30
6	6	-0.9	0.36	0.16	21	21	0.5	4.00	0.31
7	7	-0.8	0.49	0.17	22	22	0.6	5.02	0.32
8	8	-0.7	0.64	0.18	23	23	0.7	4.84	0.33
9	9	-0.6	0.81	0.19	24	24	0.8	5.29	0.34
10	10	-0.5	0.94	0.20	25	25	0.9	5.76	0.35
11	11	-0.4	1.00	0.21	26	26	1.0	6.25	0.36
12	12	-0.3	1.21	0.22	27	27	1.1	6.76	0.37
13	13	-0.2	1.44	0.23	28	28	1.2	6.98	0.38
14	14	-0.1	1.69	0.24	29	29	1.3	7.29	0.39
15	15	0.0	1.96	0.25	30	30	1.4	8.41	0.40

Содержание отчета

- Титульный лист установленного образца;
- Цель работы;
- Постановка задачи;
- Краткая теоретическая часть;
- Вариант задания;

- Проверка условия сходимости метода покоординатного спуска для заданной функции *J(u);*
- Листинг программы;
- Результаты работы программы.
- Вывод.

Контрольные вопросы

1. Опишите идею метода покоординатного спуска.
2. Геометрический смысл метода покоординатного спуска.
3. Приведите формулу перебора координатных направлений.
4. Что называют удачной итерацией?
5. В каких случаях выполняется дробление шага?
6. Каково условие сходимости метода?
7. Дайте характеристику метода покоординатного спуска в сравнении с градиентными методами.
8. Приведите преимущества и недостатки метода покоординатного спуска.

Лабораторная работа №6. Минимизация функции многих переменных методом штрафных функций

Постановка задачи

Требуется решить задачу условной оптимизации:

$$J(u) \to \min,\ u \in U,\ U \neq E^n$$

Краткая теоретическая часть

В методе штрафных функций исходная задача условной оптимизации сводится к серии задач минимизации вида:

$$\Phi_k(u) \to \min, u \in U_o \subseteq E^n, k = 1, 2, \ldots,$$

где множество U_0 содержит в себе исходное множество U ($U \subset U_0$), $\Phi_k(u)$– некоторая вспомогательная функция, причем, $\Phi_k(u)$ должна быть такой, что с ростом k она мало отличается от $J(u)$ на U и быстро возрастает на разности множеств $U_0 \backslash U$. В качестве множества U_0 обычно выбирается множество простой структуры типа параллелепипеда, шара или всего пространства E^n.

Определение 1. Последовательность функций $\{P_k(u)\}$, k=1,2,3,…, определенных и неотрицательных на множестве U_0, содержащем U, называют штрафами или штрафными функциями множества U на множестве U_0, если справедливо:

$$\lim_{k\to\infty} P_k(u) = \begin{cases} 0, u \in U \\ +\infty, u \in U_0 \setminus U \end{cases}$$

Так, для задачи выпуклого программирования, если множество U задано в виде набора ограничений в виде неравенств и равенств и имеет вид:

$$U = \left\{u \in E^n : u \in U_0,\ g_i(u) \le 0,\ i = \overline{1, m};\ g_i(u) = 0,\ i = \overline{m+1, s}\right\},$$

то в качестве штрафной можно использовать следующую функцию:

$$P_k(u) = A_k \sum_{i=1}^{s} \left(g_i^+(u)\right)^p,$$

где $A_k > 0,\ k = 1, 2, \ldots,\ \lim\limits_{k\to\infty} A_k = +\infty;$

$$g_i^+ = \begin{cases} \max\{g_i(u);\ 0\},\ i = \overline{1, m} \\ |g_i(u)|,\ i = \overline{m+1, s} \end{cases};$$

$p \geq 1$.

Пусть множество U_0: $U \subset U_0$, последовательность штрафных функций $\{P_k(u)\}$, k=1,2,3,…, уже выбраны, $J(u)$ определена на U_0. В качестве функции $\Phi_k(u)$ будем брать:

$$\Phi_k(u) = J(u) + P_k(u),\ u \in U_0.$$

Будем считать, что^

$$\Phi_{k*} = \min_{U_0} \Phi_k(u) > -\infty,\ k = 1, 2, 3, \ldots.$$

Если при каждом $k=1,2,\ldots$ нижняя граница достигается, то условия $\Phi_k(u_k) = \Phi_k^*,\ u_k \in U_0$ определяют последовательность $\{u_k\}$. Если же минимальное значение не достигается, то в качестве элементов последовательности $\{u_k\}$ будем брать такие точки u_k, удовлетворяющие условиям: $u_k \in U_0,\ \Phi_k(u_k) \le \Phi_k^* + \varepsilon_k$, где $\varepsilon_k \geq 0$: $\lim\limits_{k\to\infty} \varepsilon_k = 0$.

Никаких рекомендаций по выбору метода решения серии вспомогательных задач $\Phi_k(u) \to \min,\ u \in U_o \subseteq E^n,\ k = 1, 2, \ldots$ нет. Может быть использован любой ранее рассмотренный метод для функций многих переменных (аналитические, градиентные, покоординатного спуска).

Сходимость метода

Теорема 1. Пусть U_0 – замкнутое множество из E^n функции $J(u), g_1(u), \dots, g_m(u), |g_{m+1}(u)|, \dots, |g_s(u)|$ непрерывны на U_0, $J^* = \min\limits_U J(u) > -\infty$. Пусть последовательность $\{u_k\}$ *построена по* методу штрафных функций, имеет хотя бы одну предельную точку. Тогда все предельные точки последовательности $\{u_k\}$ принадлежат множеству U^* всех точек минимума исходной условной задачи оптимизации. Если, кроме того, существует хотя бы одно значение $\delta > 0$ такое, что множество $U_\delta = \{u \in U_0,\ g_i(u) \le \delta,\ i = \overline{1, m};\ |g_i(u)| \le \delta,\ i = \overline{m+1, s}\}$ ограничено, то $\lim\limits_{k\to\infty} \Phi_k(u_k) = \lim\limits_{k\to\infty} \Phi_{k*} = \lim\limits_{k\to\infty} J(u_k) = J^*,\ \lim\limits_{k\to\infty} \rho(u_k, U^*) = 0$, т.е. последовательность $\{u_k\}$ является минимизирующей и сходится к множеству всех точек минимума.

Задание к лабораторной работе

Разработать программу, реализующую метод штрафных функций для минимизации функции $J(u) = \frac{1}{2}(Au, u) - (b, u)$ (варианты функции определяются вариантами заданий к лабораторной работе № 4) на множестве $U = \{u \in U_0 \equiv E^2, (c, u) + d = 0\}$. В качестве штрафной функции следует взять $P_k(u) = k((c, u) + d)^2$.

Для решения вспомогательной задачи использовать аналитический метод и метод покоординатного спуска.

Выводимые значения для каждого метода:

- Координаты точки минимума и минимальное значение *J*;*
- Количество итераций и точность вычислений (при решении вспомогательной задачи методом покоординатного спуска).

Варианты заданий

№ вар.	c	d	№ вар.	c	d
1	(1, 1)	-5	16	(-4, 2)	3
2	(1, -1)	-3	17	(5, 3)	4
3	(3, -2)	4	18	(-1, 2)	0
4	(1, -1)	0	19	(2, -3)	2
5	(3, 4)	-1	20	(-2, 5)	1

№ вар.	c	d	№ вар.	c	d
6	(1, -1)	2	21	(3, 1)	2
7	(1, 4)	5	22	(-3, 2)	-3
8	(5, 2)	-1	23	(1, -2)	1
9	(2, -1)	2	24	(-2, 1)	7
10	(3, 1)	-4	25	(1, -1)	6
11	(2, -1)	44	26	(2, 6)	1
12	(7, 2)	4	27	(5, -1)	2
13	(3, 1)	1	28	(1, -1)	1
14	(1, 4)	8	29	(2, -1)	-1
15	(2, -1)	20	30	(3, 5)	-2

Содержание отчета

- Титульный лист;
- Цель работы;
- Постановка задачи;
- Краткая теоретическая часть;
- Вариант задания;
- Проверка условия сходимости метода штрафных функций для заданной функции $J(u)$;
- Аналитическое решение вспомогательной задачи;
- Листинг программы (для решения вспомогательной задачи методом покоординатного спуска);
- Результаты работы программы.
- Вывод.

Контрольные вопросы

1. Опишите идею метода штрафных функций.
2. Дайте определение штрафной функции.
3. Приведите примеры штрафных функций?
4. Каким образом решается вспомогательная задача в методе штрафных функций, какой способ применен в данной работе?
5. Каково условие сходимости метода?
6. Дайте характеристику метода штрафных функций в сравнении с ранее рассмотренными методами.
7. Приведите преимущества и недостатки метода штрафных функций.

Список использованных источников

1. Методические указания к лабораторным работам по курсу лекций "Методы оптимизации" – Владим. гос. ун-т; Сост: К.В. Демидов, А.В. Духанов. Владимир, 2003 г. 30с.
2. Акулич И.Л. Математическое программирование в примерах и задачах – 3-е издание – М: Лань, 2011 г. 352 с. ISBN: 978-5-8114-0916-7
3. Васильев Ф. П. Методы оптимизации. М.: Факториал Пресс, 2002 г. 824 с. ISBN 5-88688-056-9
4. Васильев Ф.П. Численные методы решения экстремальных задач – 2-е изд., переработ. и доп., М.: Наука, 1988 г. 552 с. ISBN 5-02-013796-0