Министерство образования и науки Украины

Донбасская государственная машиностроительная академия

МЕТОДЫ ОПТИМИЗАЦИИ И ИССЛЕДОВАНИЯ ОПЕРАЦИЙ

(МОДУЛИ 3-4)

МЕТОДИЧЕСКИЕ УКАЗАНИЯ

к самостоятельной работе

(для студентов направления «Системный анализ» заочной формы обучения)

Краматорск 2014

Министерство образования и науки Украины

Донбасская государственная машиностроительная академия

МЕТОДЫ ОПТИМИЗАЦИИ И ИССЛЕДОВАНИЯ ОПЕРАЦИЙ

(МОДУЛИ 3-4)

МЕТОДИЧЕСКИЕ УКАЗАНИЯ

к самостоятельной работе

(для студентов направления «Системный анализ» заочной формы обучения)

Краматорск 2014

УДК 621.05

Методические указания к самостоятельной работе по курсу «Методы оптимизации и исследования операций» (для направления 6.040303 «Системный анализ») заочной формы обучения / Сост.: Гитис В.Б. – Краматорск: ДГМА, 2014. – 20 с.

В методических указаниях приводятся задания, руководство к выполнению и краткие теоретические сведения по курсу «Методы оптимизации и исследования операций».

Составитель В.Б. Гитис, к.т.н., доцент каф. ИСПР

Отв. за выпуск В.Б.Гитис, к.т.н., доцент каф. ИСПР

СОДЕРЖАНИЕ

ОБЩИЕ СВЕДЕНИЯ

Контрольная работа предполагает выполнение следующих заданий:

1. Ответ на два теоретических вопроса согласно индивидуальному заданию – по 40 баллов каждый;
2. Решение задачи нахождения экстремума функции одной переменной – 20 баллов.

Контрольная работа считается сданной в случае набора не менее 40 баллов.

Экзамен представляет собой ответы на 2 теоретических вопроса и решение типовой задачи определения экстремума многомерной функции. Экзамен считается сданным в случае набора не менее 55 баллов – как по самому экзамену, так и в среднем между контрольной и экзаменационной работой.

КРАТКИЕ ТЕОРЕТИЧЕСКИЕ СВЕДЕНИЯ

1. Метод сканирования (метод перебора)

Рассмотрим задачу нахождения минимума функции y=f(x) на интервале [a, b].

Наиболее простым способом сужения интервала неопределенности является деление его на некоторое число равных частей с последующим вычислением значений целевой функции в точках разбиения. Пусть n – число элементарных отрезков, h = (b – a)/n – шаг разбиения. Вычисляем значения целевой функции $y_k = f_k(x)$ в узлах $x_k = a + kh$ (k = 0, 1, …, n). Сравнивая полученные значения $f(x_k)$, найдем среди них наименьшее $y_{min} = f_k(x)$.

$$y_{min} = \begin{cases} y_k, y_k < y_{min} \\ y_{min}, y_k \geq y_{min} \end{cases}.$$

Число $y_{min} = y_k$ можно приближенно принять за наименьшее значение целевой функции f(x) на отрезке [a, b]. Близость y_{min} к минимуму зависит от числа точек, т. е. с увеличением числа точек разбиения погрешность в определении минимума стремится к нулю.

В данном методе основная трудность состоит в выборе n и оценке погрешности.

2. Метод половинного деления (метод дихотомии)

Идея метода *дихотомии* заключается в том, чтобы на каждом шаге вдвое уменьшать интервал поиска.

Выбирается пара точек x_1 и x_2 на расстоянии $\delta > 0$ от середины текущего интервала неопределенности (рисунок 1), δ – *параметр метода*. δ принимается равным $\frac{\varepsilon}{2}$, где ε – заданная погрешность.

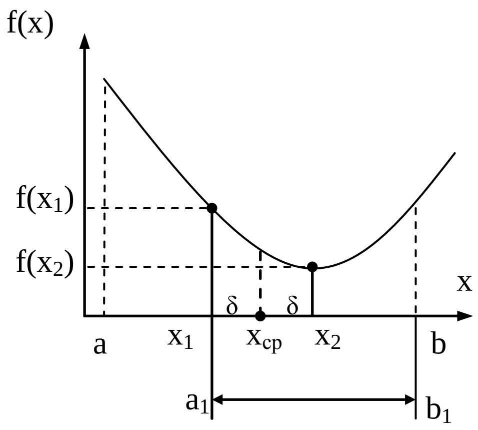

Рисунок 1 – Сужение интервала неопределённости
с помощью метода дихотомии

Точки x_1 и x_2 делят интервал [a; b] на три отрезка: [a; x_1], [x_1; x_2] и [x_2; b]. Правило исключения интервалов, основанное на свойстве унимодальности функции, позволяет исключить какой- либо из интервалов, [a; x_2] или [x_1; b], в котором определенно не может быть точки минимума. А поскольку точки x_1 и x_2 находятся от середины отрезка на расстоянии δ, составляющем половину заданной точности вычислений, то какой-либо из указанных отброшенных интервалов практически равен его половине. Таким образом, на каждом шаге данный метод позволяет исключить половину рассматриваемого интервала.

Алгоритм метода дихотомии

1 Вычисляется срединная точка интервала $x_{ср} = (a + b)/2$.

2 Определяются координаты двух точек: справа и слева вблизи точки $x_{ср}$:

$$x_1 = x_{ср} – \delta;\ x_2 = x_{ср} + \delta;$$

3 Вычисляются значения функции в этих точках $f(x_1)$ и $f(x_2)$;

4 Формируется новый интервал неопределенности:

– если $f(x_1) \le f(x_2)$, то $a_{k+1} = a_k$, $b_{k+1} = x_2$;

– если $f(x_1) > f(x_2)$, то $a_{k+1} = x_1$, $b_{k+1} = b_k$.

5 Если $(b – a) < \varepsilon$, то поиск прекращается, если $(b – a) > \varepsilon$, то следует перейти на этап 1 и повторить процедуру вычислений на следующей итерации.

После k вычислений функции экстремум окажется в интервале

$$l = b_k - a_k = \frac{b - a - \delta}{2^{\frac{k}{2}}} + \delta.$$

В качестве оптимальной точки принимается

$$x^* = \frac{a_k + b_k}{2}.$$

Преимуществом метода дихотомии является то, что на каждой итерации алгоритма исключается практически половина интервала поиска. Недостатком является необходимость вычислений на каждом шаге значений функции в двух точках.

3. Метод золотого сечения

Метод золотого сечения является одним из наиболее эффективных методов, в которых при ограниченном количестве вычислений f(x) достигается наилучшая точность. Он состоит в построении последовательности отрезков $[a_0, b_0]$, $[a_1, b_1]$, …, стягивающихся к точке минимума функции f(x). На каждом шаге, за исключением первого, вычисление значения функции f(x) проводится лишь один раз в определенной точке. Эта точка, называемая золотым сечением, выбирается специальным образом.

Золотым сечением отрезка называют деление отрезка на две неравные части так, чтобы отношение длины всего отрезка к длине большей части было равно отношению длины большей части к меньшей (рисунок 2, *а*). Т. е. выполняется соотношение

$$\frac{b - a}{b - x} = \frac{b - x}{x - a} \approx 1{,}618.$$

Рисунок 2 – Золотые сечения отрезков

Всего на отрезке можно отложить 2 золотых сечения (рисунок 2, *б*).

Преобразовав выражение золотого сечения, получим:

$$x_1 = a + 0{,}382(b - a) = 0{,}618a + 0{,}382b;$$
$$x_2 = a + 0{,}618(b - a) = 0{,}382a + 0{,}618b.$$

Точки x_1 и x_2 располагаются симметрично относительно середины отрезка. Точка x_1 дает золотое сечение отрезка $[a, x_2]$, а точка x_2 дает золотое сечение отрезка $[x_1, b]$.

Алгоритм метода золотого сечения

1 Проводятся два первых вычисления точек:

$$x_1 = 0{,}618a + 0{,}382b;$$
$$x_2 = 0{,}382a + 0{,}618b.$$

2 Вычисляются значения функции в этих точках $f(x_1)$ и $f(x_2)$.

3 Формируется новый интервал неопределенности:

- если $f(x_1) \le f(x_2)$, то $a_{k+1} = a_k$, $b_{k+1} = x_2$;
- если $f(x_1) > f(x_2)$, то $a_{k+1} = x_1$, $b_{k+1} = b_k$.

4 Так как оставшаяся точка дает золотое сечение нового отрезка, то на каждой итерации выбор новой точки $x^н$ и вычисление значения функции производится 1 раз:

- если $f(x_1) \le f(x_2)$, то $x_1^н = 0{,}618a_k + 0{,}382b_k$; $x_2^н = x_1$;
- если $f(x_1) > f(x_2)$, то $x_1^н = x_2$; $x_2^н = 0{,}382a_k + 0{,}618b_k$.

5 Вычисление значения функции $f(x_1^н)$ или $f(x_2^н)$.

6 Если $(b - a) < \varepsilon$, то поиск прекращается, иначе осуществляется переход на этап 3.

После k вычислений функции экстремум окажется в интервале

$$l = b_k - a_k = 0{,}618^k(b - a).$$

В качестве оптимальной точки принимается

$$x^* = \frac{a_k + b_k}{2}.$$

Метод золотого сечения является более эффективным по сравнению с методом дихотомии (и других алгоритмов, использующих процедуру деления интервала на отрезки), поскольку он требует наименьшего числа вычислений значения функции для достижения одной и той же точности.

В то же время уже через небольшое число итераций быстро возрастает погрешность вычисления золотых сечений интервала неопределённости, связанная с ошибками округления. Поэтому через 5 – 10 итераций необходимо заново вычислить 2 золотых сечения, т. е. обнулить метод.

4. Метод оценки с использованием квадратичной аппроксимации

Метод основан на том, что функция, которая принимает минимальное значение во внутренней точке интервала, должна быть, по крайней мере, квадратичной. При реализации метода предполагается, что на ограниченном интервале можно аппроксимировать функцию квадратичным полиномом, а потом использовать построенную аппроксимационную схему для оценки координаты точки истинного минимума функции.

Если задана последовательность точек x1, x2, x3 и известны соответствующие этим точкам значения функции f(x1), f(x2), f(x3), то можно определить постоянные величины a0, a1, a2 таким образом, что значения квадратичной функции q(x) = a0 + a1(x – x1) + a2(x – x1)(x – x2) совпадут со значениями f(x) в трех указанных точках. Перейдем к вычислению q(x) в каждой из трех заданных точек.

1) При x = x1 f(x1) = q(x1) = a0, т. е. a0 = f(x1);
2) При x = x2 f(x2) = q(x2) = f(x1) + a1(x2 – x1). $a1 = \frac{f(x2) - f(x1)}{x2 - x1}$;
3) При x = x3 f(x3) = q(x3) = $f(x1) + \frac{(f(x2) - f(x1))(x3 - x1)}{x2 - x1} + a2(x3 - x1)(x3 - x2)$.

$$a2 = \frac{1}{x3 - x2}\left(\frac{f(x3) - f(x1)}{x3 - x1} - \frac{f(x2) - f(x1)}{x2 - x1}\right)$$

Для стационарной точки функции q(x):

$$\frac{dq(x)}{dx} = a1 + 2a2x - a2x1 - a2x2 = 0\,.$$

Откуда $\bar{x} = \frac{x2 + x1}{2} - \frac{a1}{2a2}$.

Величина $\bar{x}$ является оценкой координаты точки истинного оптимума X*.

5. Метод средней точки (метод Больцано)

Структура поиска основана на исключения интервалов на основании исследования знака производной функции независимо от значений, которые эта производная принимает.

На каждой итерации рассматривается лишь одна пробная точка. Если в пробной точке z выполняется неравенство $f'(z) < 0$, то с учетом предположения об унимодальности, точка минимума не может находиться левее точки z и интервал $x \le z$ подлежит исключению. Если $f'(z) > 0$, то точка минимума не может находиться правее z и исключается интервал $x \ge z$.

Алгоритм метода

Пусть есть ограниченный интервал $a \le x \le b$ и задан параметр сходимости ε. Причем $f'(a) < 0$ и $f'(b) > 0$.

Шаг 1. Вычислить $z = (a + b)/2$ и $f'(z)$.

Шаг 2. Если $|f'(z)| < \varepsilon$, закончить поиск. В противном случае, если $f'(z) < 0$, принять $a = z$, если $f'(z) > 0$, принять $b = z$. Переход к шагу 1.

6. Метод секущих (метод хорд)

Пусть в процессе поиска стационарной точки функции f(x) в интервале [a, b] выявлено, что в точках a и b знаки производной разные. В этом случае алгоритм метода секущих позволяет аппроксимировать функцию $f'(x)$ «секущей прямой» (прямой линией, которая соединяет две точки) и найти точку, в которой секущая графика $f'(x)$ пересекает ось абсцисс. Приближение к стационарной точке X* определяется по формуле:

$$z = b - \frac{f'(b)(b-a)}{f'(b) - f'(a)}$$

Если $|f'(z)| < \varepsilon$, поиск заканчивается. В противном случае необходимо выбрать одну из точек a или b таким образом, чтобы знаки производной в этой точке и точке z были разными, а потом повторить основный шаг алгоритма.

В отличие от метода средней точки метод секущих основан на исследовании не только знака, но и значений производной в пробных точках и потому в ряде случаев позволяет исключить больше половины интервала поиска.

7. Метод Ньютона-Рафсона

Метод Ньютона-Рафсона позволяет улучшить относительно грубую аппроксимацию, чтобы получить корень уравнения $f'(x) = 0$.

Спуск производится по формуле

$$x_{k+1} = x_k - \gamma_k \frac{f'(x_k)}{f''(x_k)},$$

где $k = 0, 1, 2, \ldots$ – номер итерации;

γ_k – шаг на k-й итерации (задаваемая константа).

Итерации продолжаются до достижения необходимой точности ε, когда $|f'(x_k)| < \varepsilon$.

Для квадратичной функции экстремум достигается за 1 итерацию.

8. Понятие выпуклости и вогнутости

Функция f(x), определённая на выпуклом множестве Dx, называется *выпуклой* (рисунок 3), если для любых двух точек x_1 и x_2 из множества Dx выполняется неравенство

$$f(\alpha x_1 + (1 - \alpha)x_2) \le \alpha f(x_1) + (1 - \alpha)f(x_2) \text{ при } 0 \le \alpha \le 1. \quad (1)$$

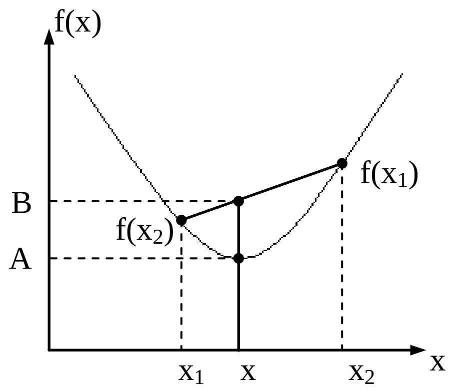

Рисунок 3 – Выпуклая функция

Если неравенство (1) имеет обратный знак, то функция называется *вогнутой* (рисунок 47).

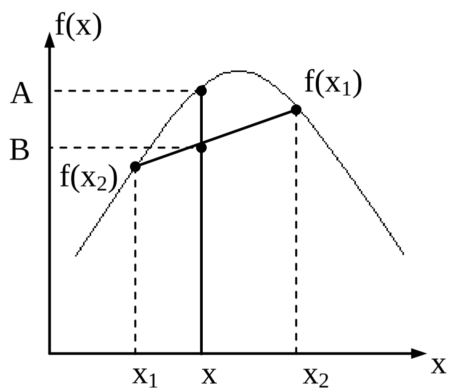

Рисунок 4 – Вогнутая функция

На рисунках 3 и 4 значения A и B соответствуют левой и правой частям неравенства (1). Т. к. для функции, представленной на рисунке 3, неравенство (1) выполняется (A < B), то функция считается выпуклой. Для функции же, приведенной на рисунке 4, неравенство (1) не выполняется (A > B), и функция считается вогнутой.

Если неравенство (1) выполняется как строгое (например, для функций, представленных на рисунках 3 и 4), то функция называется *строго выпуклой* (*строго вогнутой*). В противном случае – *нестрого выпуклой* (*нестрого вогнутой*) (рис. 5).

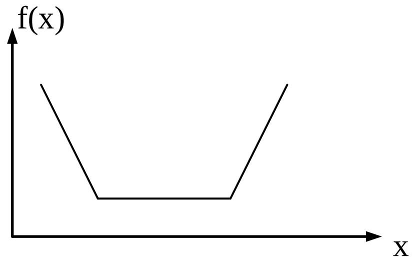

Рисунок 5 – Нестрого выпуклая функция

Выпуклость (вогнутость) функции многих переменных в окрестности точки X* определяется с помощью определителей, построенных на вторых производных функции f(x):

$$f_{11}(X^*);\qquad \begin{vmatrix} f_{11}(X^*) & f_{12}(X^*) \\ f_{21}(X^*) & f_{22}(X^*) \end{vmatrix};\qquad \begin{vmatrix} f_{11}(X^*) & f_{12}(X^*) & f_{13}(X^*) \\ f_{21}(X^*) & f_{22}(X^*) & f_{23}(X^*) \\ f_{31}(X^*) & f_{32}(X^*) & f_{33}(X^*) \end{vmatrix};$$

$$H(X^*) = \begin{vmatrix} f_{11}(X^*) & f_{12}(X^*) & \cdots & f_{1n}(X^*) \\ f_{21}(X^*) & f_{22}(X^*) & \cdots & f_{2n}(X^*) \\ \vdots & \vdots & \vdots & \vdots \\ f_{n1}(X^*) & f_{n2}(X^*) & \cdots & f_{nn}(X^*) \end{vmatrix}, \qquad (2)$$

где $f_{ij}(X^*) = \dfrac{\partial^2 f(X^*)}{\partial x_i \partial x_j}$.

Матрица n-го порядка H(x) (2) называется *матрица Гессе* (или *гессиан*).

Если знаки определителей чередуются, то матрица Гессе называется *отрицательно определённой*. В этом случае функция является строго вогнутой, и в точке X* достигается максимум. Если все определители положительные, то матрица Гессе называется *положительно определённой*, функция f(x) – строго выпуклая, а точка X* является точкой минимума. Если закономерности нет, то функция f(x) в окрестности точки X* называется *седловой* (рисунок 6).

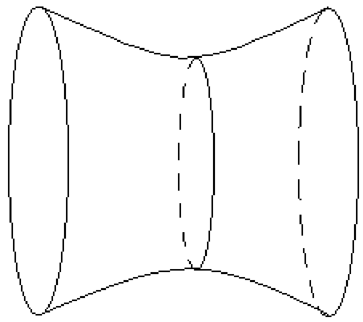

Рисунок 6 – Седловая функция

ПЕРЕЧЕНЬ ВОПРОСОВ ДЛЯ ВЫПОЛНЕНИЯ КОНТРОЛЬНОЙ РАБОТЫ

ТЕОРЕТИЧЕСКИЕ ВОПРОСЫ

1. Постановка задачи оптимизации
2. Понятие экстремума. Виды экстремумов
3. Системы ограничений в математических моделях
4. Классификация методов оптимизации
5. Необходимое и достаточное условие экстремума. Понятие выпуклости и вогнутости
6. Постановка задачи одномерной оптимизации
7. Метод сканирования. Стратегия последовательного поиска
8. Метод дихотомии
9. Метод золотого сечения
10. Методы поиска с использованием полиномиальной аппроксимации
11. Метод Ньютона-Рафсона
12. Метод средней точки
13. Метод секущих

ЗАДАЧИ

1. Построить гессиан для функции $f(x_1, x_2, x_3)$ в точке X;
2. Используя определение выпуклой функции одной переменной, проверить функцию f(x) на выпуклость в интервале [a; b];
3. Определить характер экстремума функции $f(x_1, x_2, x_3)$ в стационарной точке X*;
4. Найти экстремум функции f(x) в интервале [a; b] методом сканирования с точностью ε;
5. Найти экстремум функции f(x) в интервале [a; b] методом дихотомии с точностью ε и параметром метода δ;
6. Найти экстремум функции f(x) в интервале [a; b] методом золотого сечения с точностью ε;

7. Пусть для функции f(x) задана последовательность точек x_1, x_2, x_3 и соответствующие им значения функции $f(x_1)$, $f(x_2)$, $f(x_3)$. Найти экстремум функции f(x) методом оценки с использованием квадратичной аппроксимации;
8. Найти экстремум функции f(x) методом Ньютона-Рафсона. Начальная точка X_0, минимальный шаг γ_{min};
9. Найти экстремум функции f(x) в интервале [a; b] методом Больцано. Параметр сходимости ε;
10. Найти экстремум функции f(x) в интервале [a; b] методом секущих. Параметр сходимости ε.

ПЕРЕЧЕНЬ ВОПРОСОВ ДЛЯ ЭКЗАМЕНА

ТЕОРЕТИЧЕСКИЕ ВОПРОСЫ

1. Общая стратегия многомерной оптимизации
2. Метод покоординатного спуска
3. Метод вращающихся координат
4. Метод параллельных касательных
5. Принцип градиентного спуска
6. Градиентный метод с переменным шагом
7. Метод наискорейшего спуска
8. Партан-метод
9. Метод тяжелого шарика
10. Методы сопряженных градиентов
11. Метод Левенберга-Марквардта
12. Метод Ньютона
13. Квазиньютоновские методы
14. Метод множителей Лагранжа
15. Методы штрафных функций
16. Метод случайного поиска с пересчетом
17. Метод случайного поиска по наилучшей пробе
18. Методы глобальной оптимизации

ЗАДАЧИ

1. Найти экстремум функции $f(x_1, x_2)$ методом покоординатного спуска. Начальная точка X_0, шаг γ;
2. Найти экстремум функции $f(x_1, x_2)$ градиентным методом с переменным шагом. Начальная точка X_0, начальный шаг γ_0, минимальный шаг γ_{min};
3. Найти минимум функции $f(x_1, x_2)$ при $h(x_1, x_2) = 0$ методом множителей Лагранжа.

СПИСОК РЕКОМЕНДУЕМОЙ ЛИТЕРАТУРЫ

1 Бахвалов Н.С. Численные методы / Н.С. Бахвалов. – М.: Наука, 1975. – 275 с.

2 Бажин И.И. Информационные системы менеджмента / И.И. Бажин. – М.: ТУ-ВШЭ, 2000.– 688 с.

3 Березин И.С. Методы вычислений / И.С.Березин, Н.П.Жидков. – М.: Наука, 1966. – 472 с.

4 Васильев Ф.П. Численные методы решения экстремальных задач / Ф.П.Васильев. – М.: Наука. Главная редакция физико-математической литературы, 1980. – 518 с.

5 Волков Е.А. Численные методы / Е.А.Волков. – М.: Наука, 1982. – 194 с.

6 Дьяченко В.Ф. Основные понятия вычислительной математики / В.Ф.Дьяченко. – М.: Наука, 1977. – 362 с.

7 Евтушенко Ю.Г. Методы решения экстремальных задач и их применение в системах оптимизации /Ю.Г. Евтушенко. – М.: Наука, 1982. – 274 с.

8 Калиткин Н.Н. Численные методы / Н.Н.Калиткин. – М.: Наука, 1978. – 384 с.

9 Конспект лекций по курсу «Математические модели в менеджменте и маркетинге» (для студентов специальности 7.050102 очной и заочной формы обучения)/ Сост. Гитис В.Б. – Краматорск: ДГМА, 2004. – 64 с.

10 Крылов В.И. Вычислительные методы / В.И.Крылов, В.В.Бобков, П.И.Монастырный. – М.: Наука, 1977. – 477 с.

11 Пшеничный Б.Н. Численные методы в экстремальных задачах / Б.Н.Пшеничный, Ю.М.Данилин. – М.: Наука, 1975. – 420 с.

12 Самарский А.А. Введение в численные методы / А.А.Самарский. – М.: Наука, 1982. – 233 с.

13 Турчак Л.И. Основы численных методов: учеб. пособие / Л.И.Турчак. – М.: Наука. Гл. ред. физ.-мат. лит., 1987. – 320 с.

14 Численные методы и вопросы организации вычислений / Под. ред. В.П.Ильина, В.Н. Кублановской. – Л.: Наука, 1984. – 184 с.