Лабораторная работа 2.

Методы минимизации функций одной переменной, использующие информацию о производных целевой функции.

Постановка задачи: Требуется найти безусловный минимум функции одной переменной $f(x)$, т.е. такую точку $x^* \in U$, что $f(x^*) = \min_{x \in U} f(x)$. Значение точки минимума вычислить приближенно с заданной точностью ε.

В лабораторной работе №1 были изучены прямые методы решения задачи минимизации функции одной переменной. В данной работе будут рассмотрены методы минимизации, которые используют информацию о производной целевой функции.

Пусть на предварительно выбранном интервале неопределенности $U_0 = [a;b]$ целевая функция $f(x)$ является выпуклой дифференцируемой функцией. Тогда необходимым и достаточным условием глобального минимума является равенство нулю первой производной функции:

$$f'(x) = 0, \quad x \in U_0 = [a;b] \qquad (1)$$

Метод средней точки.

Стратегия поиска: Если определение производной целевой функции не представляет затруднений, то в процедуре исключения отрезков методом дихотомии вычисление двух значений $f(x)$ вблизи середины очередного отрезка можно заменить вычислением одного значения $f'(x)$ в его средней точке $\bar{x} = \frac{a+b}{2}$. Сравнивая $f'(\bar{x})$ с нулем, делим отрезок поиска точки x^* ровно вдвое, причем на каждой итерации вычисляется только одно значение $f'(x)$.

Алгоритм:

1. Выбрать начальный интервал неопределенности $U_0 = [a;b]$ и точность ε.
2. Положить $\bar{x} = \frac{a+b}{2}$. Вычислить $f'(\bar{x})$.
3. Проверка на окончание поиска: если $|f'(\bar{x})| \le \varepsilon$, то положить $x^* = \bar{x}$, $f* = f(\bar{x})$ и завершить поиск, иначе – перейти к шагу 4.
4. Сравнить $f'(\bar{x})$ с нулем. Если $f'(\bar{x}) > 0$, то продолжить поиск на отрезке $[a;\bar{x}]$, положив $b = \bar{x}$, иначе – перейти к отрезку $[\bar{x};b]$, положив $a = \bar{x}$. Перейти к шагу 2.

Метод хорд.

Стратегия поиска: Пусть на концах отрезка $[a;b]$ производная $f'(x)$ имеет разные знаки, т.е. $f'(a) \cdot f'(b) < 0$. Тогда на интервале $(a;b)$ найдется точка, в которой $f'(x)$ обращается в нуль. В этом случае поиск точки минимума $f(x)$ на отрезке $[a;b]$ эквивалентен решению уравнения $f'(x) = 0$ на интервале $(a;b)$.

Сущность метода хорд приближенного решения уравнения $f'(x) = 0$ на отрезке $[a;b]$ при $f'(a) \cdot f'(b) < 0$ состоит в исключении отрезков путем определения точки $\tilde{x}$ - точки пересечения с осью Ox хорды графика функции $f'(x)$ на $[a;b]$.

Координата точки $\tilde{x}$ равна:

$$\tilde{x} = a - \frac{f'(a)}{f'(a) - f'(b)}(a - b)\,. \qquad (2)$$

Отрезок дальнейшего поиска точки x^* (отрезок $[a;\tilde{x}]$ или $[\tilde{x};b]$) выбирается в зависимости от знака $f'(\tilde{x})$ так же, как в методе средней точки. На каждой итерации, кроме первой, необходимо вычислять только одно новое значение $f'(x)$.

Алгоритм:

1. Выбрать начальный интервал неопределенности $U_0 = [a;b]$ и точность ε.
2. Вычислить $f'(a)$ и $f'(b)$. Проверить условие $f'(a)\cdot f'(b) < 0$. Если оно выполняется перейти к шагу 3, иначе завершить алгоритм, выдав соответствующий предупреждение.
3. Найти $\tilde{x}$ по формуле (2). Вычислить $f'(\tilde{x})$.
4. Проверка на окончание поиска: если $|f'(\tilde{x})| \le \varepsilon$, то положить $x^* = \bar{x}$, $f* = f(\bar{x})$ и завершить поиск, иначе – перейти к шагу 5.
5. Переход к новому отрезку. Если $f'(\tilde{x}) > 0$, то положить $b = \tilde{x}$, $f'(b) = f'(\tilde{x})$, иначе – положить $a = \tilde{x}$, $f'(a) = f'(\tilde{x})$. Перейти к шагу 3.

Метод Ньютона.

Стратегия поиска: Пусть функция является дважды непрерывно дифференцируемой на $[a;b]$. Тогда для приближенного решения уравнения (1) можно применить метод Ньютона, или метод касательных.

Пусть $x_0 \in [a;b]$ - нулевое, или начальное приближение к искомой точке x^*. Линеаризуем функцию $F(x) = f'(x)$ в окрестности начальной точки, приближенно заменив дугу графика этой функции касательной в точке $(x_0\,, f'(x_0))$:

$$F(x) \approx F(x_0) + F'(x_0)\cdot(x - x_0)\,. \qquad (3)$$

Выберем в качестве следующего приближения к x^* точку x_1 пересечения касательной с осью абсцисс. Приравнивая к нулю правую часть в (3), получим первый элемент $x_1 = x_0 - \dfrac{F(x_0)}{F'(x_0)}$ итерационной последовательности $\{x_k\},\quad k = 1,2\ldots$.

В очередной точке x_k строится линейная аппроксимация функции $F(x)$ и точка, в которой эта аппроксимирующая функция обращается в нуль, используется в качестве следующего приближения x_{k+1}.

Уравнение касательной к графику $F(x)$ в точке $x = x_k$ имеет вид $y = F(x_k) + F'(x_k)\cdot(x - x_k)$, поэтому точка $x = x_{k+1}$, найденная из условия $y = 0$, определяется формулой $x_{k+1} = x_k - \dfrac{F(x_k)}{F'(x_k)}$. Возвращаясь к обозначению $F(x) = f'(x)$, получим, что для решения уравнения $f'(x) = 0$ необходимо построить последовательность

$$x_{k+1} = x_k - \frac{f'(x_k)}{f''(x_k)},\quad k = 0,1,\ldots \qquad (4)$$

где x_0 - точка, выбранная в качестве начального приближения. Вычисления по формуле (4) производятся до тех пор, пока не выполнится неравенство $|f'(x_k)| \le \varepsilon$, после чего полагают $x^* \approx x_k$, $f^* \approx f(x_k)$.

Возможные модификации метода Ньютона.

1. Метод Ньютона-Рафсона:

При переходе к новой итерации новая точка x_{k+1} рассчитывается по формуле:

$$x_{k+1} = x_k - \tau_k \frac{f'(x_k)}{f''(x_k)}, \quad 0 < \tau_k \le 1.$$

В простейшем варианте метода $\tau_k = \tau = const$ (значение $\tau = 1$ соответствует исходному методу Ньютона). Оптимальный набор параметров τ_k может быть найден из решения задачи минимизации:

$$\varphi(\tau) = f\left(x_k - \tau \frac{f'(x_k)}{f''(x_k)}\right) \to \min.$$

На практике для параметров τ_k обычно используется приближенное решение последней задачи:

$$\tau_k = \frac{(f'(x_k))^2}{(f'(x_k))^2 + (f'(\tilde{x}))^2}, \text{ где } \tilde{x} = x_k - \frac{f'(x_k)}{f''(x_k)}.$$

2. Метод Марквардта:

При переходе к новой итерации новая точка x_{k+1} рассчитывается по формуле:

$$x_{k+1} = x_k - \frac{f'(x_k)}{f''(x_k) + \mu_k}, \quad \mu_k > 0$$

Значение параметра μ_0 выбирается как минимум на порядок больше значения $f''(x_0)$.

При переходе к новой итерации новое значение μ_{k+1} полагают равным $\mu_{k+1} = \mu_k / 2$, если $f(x_{k+1}) < f(x_k)$, либо $\mu_{k+1} = 2\mu_k$ в противном случае.

Метод перебора (минимизации многомодальных функций).

Стратегия поиска: Применение этого метода строго обосновано лишь для унимодальной на $[a;b]$ функции $f(x)$. Однако, если вместо унимодальности потребовать, чтобы функция $f(x)$ удовлетворяла на $[a;b]$ условию Липшица

$$|f(x_1) - f(x_2)| \le L|x_1 - x_2|, \text{ для всех } x_1, x_2 \in [a;b], \tag{5}$$

то можно гарантировать определение минимального значения f^* методом перебора с любой заданной точностью. Сформулируем утверждение более строго.

Пусть функция $f(x)$ удовлетворяет на отрезке $[a;b]$ условию Липшица (5) с константой L и приближенные значения $x^* \approx x_m$, $f^* \approx f(x_m)$ найдены методом перебора с разбиением отрезка $[a;b]$ на n частей. Тогда для погрешности δ_n определения минимального значения f^* справедлива оценка

$$\delta_n = f(x_m) - f^* \le L \cdot \frac{b-a}{2n}. \tag{6}$$

Замечание. Если функция $f(x)$ многомодальна, то погрешность определения ее точки минимума может быть значительной, несмотря на то, что сам минимум f^* найден достаточно точно.

Однако во многих случаях практический интерес представляют те значения аргумента x (возможно, далекие от x^*), при которых целевая функция принимает значения, достаточно близкие к минимальному. Это позволяет использовать метод перебора для многомодальных функций.

Метод ломаных (минимизации многомодальных функций).

Ознакомиться со стратегией и алгоритмом метода ломаных по главе 3 пособия на электронном носителе или по учебному пособию на бумажном носителе.

Задания

1. Написать в среде MATLAB функции, реализующие метод средней точки, метод хорд и метод Ньютона.
2. Выбрать для выполнения лабораторной работы тестовую функцию, номер которой соответствует номеру Вашего компьютера. Например, для компьютера №3 это будет функция 3), для компьютера №13 – функция 4): 13-9=4; для компьютера №23 это будет функция 5): 23-9$\times$2=5.

1) $f(x) = x^3 - 3\sin x \to \min, \quad x \in [0;1].$
2) $f(x) = x^4 + x^2 + x + 1 \to \min, \quad x \in [-1;0].$
3) $f(x) = e^x + \dfrac{1}{x} \to \min, \quad x \in [0{,}5;1{,}5].$
4) $f(x) = x^2 - 2x + e^{-x} \to \min, \quad x \in [-1;1{,}5].$
5) $f(x) = x\sin x + 2\cos x \to \min, \quad x \in [-6;-4].$
6) $f(x) = x + \dfrac{1}{x^2} \to \min, \quad x \in [1;2].$
7) $f(x) = 10x\ln x - \dfrac{x^2}{2} \to \min, \quad x \in [0{,}1;1].$
8) $f(x) = e^x - \dfrac{1}{3}x^3 + 2x \to \min, \quad x \in [-2{,}5;-1].$
9) $f(x) = x^2 - 2x - 2\cos x \to \min, \quad x \in [-0{,}5;1].$

3. Для выбранной функции (построить ее график!) и для каждого указанного в пункте 1 метода изучить зависимость скорости работы (числа вычислений производной функции N) от заданного значения точности ε. Провести сравнение методов друг с другом. Сравнить полученные результаты с результатами выполнения Лабораторной работы №1. Объяснить полученные результаты.
4. Используя встроенную в MATLAB функцию fminsearch, вычислить координату минимума выбранной функции.
5. Определить, сколько вычислений производной функции потребуется каждому методу, чтобы модуль разности координаты минимума и результата, полученного в пункте 4, не превосходил 10^{-4}.
6. С помощью метода Ньютона решить задачу минимизации функции

$$f(x) = x\,arctg\,x - \frac{1}{2}\ln(1+x^2)$$

для нескольких вариантов выбора начального приближения. Убедиться в том, что при выборе начального приближения не достаточно близко от точки минимума метод Ньютона может расходиться. Найти диапазон начальных приближений, при которых метод сходиться к точке минимума функции.

Решить ту же задачу с теми же начальными приближениями с помощью модификаций метода Ньютона (метода Марквардта и метода Ньютона-Рафсона). Объяснить полученные результаты.

7. Составить программу нахождения глобального минимума многомодальных функций методом перебора и методом ломаных. При построении программы метода ломаных можно пользоваться кодом, предложенном в приложении. Проверить ее работоспособность на примере следующих функций (построить их графики!):

а) $f(x) = \dfrac{\cos x}{x^2} \to \min, \quad x \in [1;12];$

$f(x) = \dfrac{1}{10}x + 2\sin 4x \to \min, \quad x \in [0;4];$

б) $f(x) = \dfrac{\cos(10x)}{e^x} \to \min, \quad x \in [1;5]\,;$

$f(x) = 0.3\cos 2x + 2\sin(4x) \to \min, \quad x \in [0;4]\,.$

Сделать выводы о сравнительных достоинствах и недостатках метода перебора и метода ломаных.

8. Сдать лабораторную работу преподавателю, ответив предварительно на все следующие контрольные вопросы.

Контрольные вопросы к лабораторной работе 2.

1. Пусть $f(x)$ - выпуклая дифференцируемая функция и $|f'(\bar{x})| \le \varepsilon$. Можно ли указать погрешности определения точки минимума x^* и минимального значения f^* по формулам $x^* = \bar{x}$, $f^* = f(\bar{x})$? Ответ пояснить рисунком.
2. Является ли условие $f'(\bar{x}) = 0$ достаточным для того, чтобы число $\bar{x}$ было точкой минимума унимодальной, но не выпуклой функции $f(x)$? Ответ сопроводить примером.
3. Указать класс функций, для которых точное определение точки минимума гарантировано в результате всего одной итерации метода Ньютона.
4. Сформулировать достаточные условия сходимости метода Ньютона.
5. Сформулировать достаточные условия монотонной сходимости метода Ньютона. Всегда ли в этом случае скорость сходимости будет квадратичной?
6. Для каких выпуклых дважды дифференцируемых функций метод золотого сечения приводит к цели за меньшее количество итераций, чем метод Ньютона?
7. Минимизировать функцию $f(x) = (x-1)^8 \to \min, x \in [0;2]$ с помощью методов Ньютона и золотого сечения. Сравнить эти методы.

8. Сформулировать оценку погрешности определения минимума f^* многомодальной функции методом перебора.
9. Увеличение используемого значения константы Липшица L при реализации метода ломаных приводит к замедлению сходимости метода. Объяснить этот факт с помощью геометрической иллюстрации.
10. Показать с помощью рисунка, что если в методе ломаных используется ошибочно заниженное значение константы Липшица L, то задача минимизации может быть решена неверно.

Приложение

В приложении представлен код алгоритма метода ломаных в виде m-файла.

```
function [Xmin,Fmin,count] = lom(func,U0,eps,L)
% Функция реализует алгоритм метода ломаных.
% Входные параметры:
% func - целевая функция (должны быть введены как inline
функции,
% например func = inline('x.^3-3*sin(x)'))
% U0  - начальный интервал неопределённости
% eps - точность вычислений
% L   - одна из констант Липшица целевой функции
% Возвращаемае значения
% Xmin  - координата минимума
% Fmin  - значение функции в точке Xmin
% count - количество вычислений функции.

count = 0;  % количество вычислений функции
N = 0;      % длина массива ломанной

x = [U0(1):0.01:U0(2)];
y = func(x);
% шаг 1
fa = func(U0(1)); count = count + 1;
fb = func(U0(2)); count = count + 1;

% добавляем первый минимум ломанной (x0,y0)
x0 = (fa - fb + L*(U0(1) + U0(2)))/(2*L);
y0 = (fa + fb + L*(U0(1) - U0(2)))/(2);
N = N + 1;
lom_x(N) = x0;
lom_y(N) = y0;
```

```
while 1;
    % шаг 2
    % поиск минимума в ломаной
    min_idx = 1;
    for i = 2:N
        if lom_y(i) < lom_y(min_idx)
            min_idx = i;
        end
    end
    % шаг 3
    % вычисляем delta
    fx    = func(lom_x(min_idx)); count = count + 1;
    px    = lom_y(min_idx);
    delta = (fx - px)/(2*L);
    % проверяем условие выхода из цикла
    if 2*L*delta <= eps
        Xmin = lom_x(min_idx);
        Fmin = fx;
        break;
    end
    % шаг 4
    % вводим в расмотрение ещё 2 минимума, вместо текущего
    min_x1 = lom_x(min_idx) - delta;
    min_x2 = lom_x(min_idx) + delta;
    min_y  = (fx + px)/2;
    % для экономии массива, заменяем значение в текущем
минимуме
    lom_x(min_idx) = min_x1;
    lom_y(min_idx) = min_y;
    % добавляем новый минимум в конец массива
    N = N + 1;
    lom_x(N) = min_x2;
    lom_y(N) = min_y;
    plot(x,y,'b',lom_x,lom_y,'r+');
    pause;
end
```