Санкт-Петербургский политехнический университет Петра Великого
Институт прикладной математики и механики
Высшая школа прикладной математики и вычислительной физики

Лабораторная работа №3

Тема: «Решение задач одномерной минимизации»

по дисциплине

«Методы оптимизаций»

Выполнили студенты
группы 3630102/80401

Мамаева Анастасия Сергеевна
Веденичев Дмитрий Александрович
Тырыкин Ярослав Алексеевич

Руководитель
Доцент, к.ф.-м.н.

Родионова Елена Александровна

Санкт-Петербург
2021

СОДЕРЖАНИЕ

1 Постановка задачи

Пусть имеются две трансцендентные функции:

$$f(x) \ = \ x^2 - 2x + e^{-x} \quad [1; \ 1.5] \quad \varepsilon = 0.05$$

$$f(x) \ = \ x^3 - 3sin(x) \quad [0.5; \ 1] \quad \varepsilon = 0.05$$

Необходимо:

1. Найти минимум данных функций, используя методы дихотомии и Фибоначчи.

2. Провести сравнительный анализ данных методов.

3. Исследовать зависимость числа обращений к функции от задаваемой точности.

4. Вывести аналитическую оценку для числа обращений к функции у обоих методов.

2 Применимость методов

Рассмотрим простейшую математическую модель оптимизации, в которой целевая функция зависит от одной переменной, а допустимым множеством является отрезок вещественной оси:

$$f(x) \longrightarrow min$$

$$x \in [a; \ b]$$

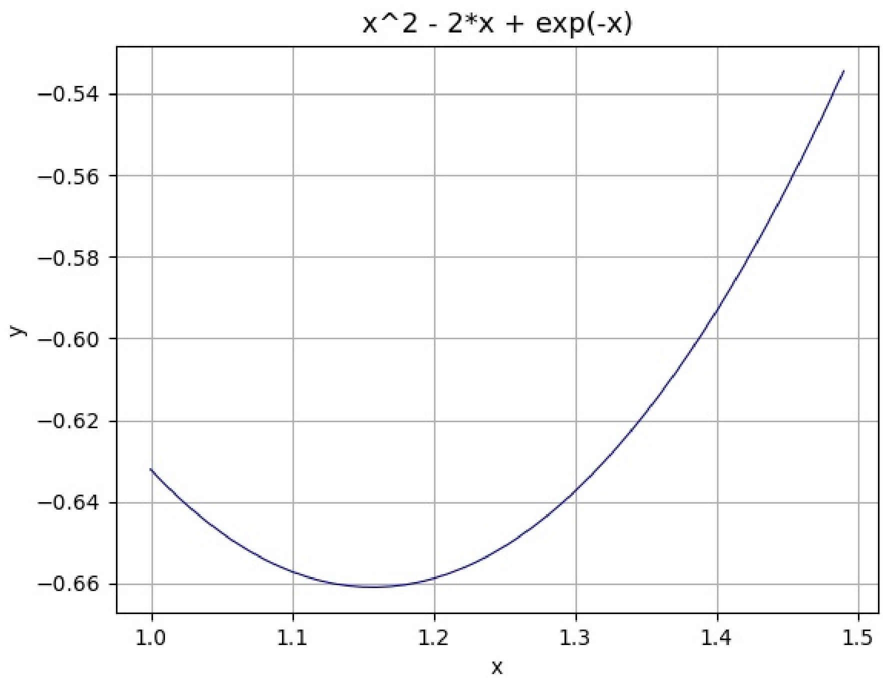
Для решения таких задач на практике, как правило, применяются приближённые методы. Они позволяют найти решение этой задачи с необходимой точностью в результате определения конечного числа значений функции $f(x)$ и её производных в некоторых точках отрезка $[a; \ b]$. Методы использующие только значения функции и не требующие вычисления её производных называются прямыми методами.

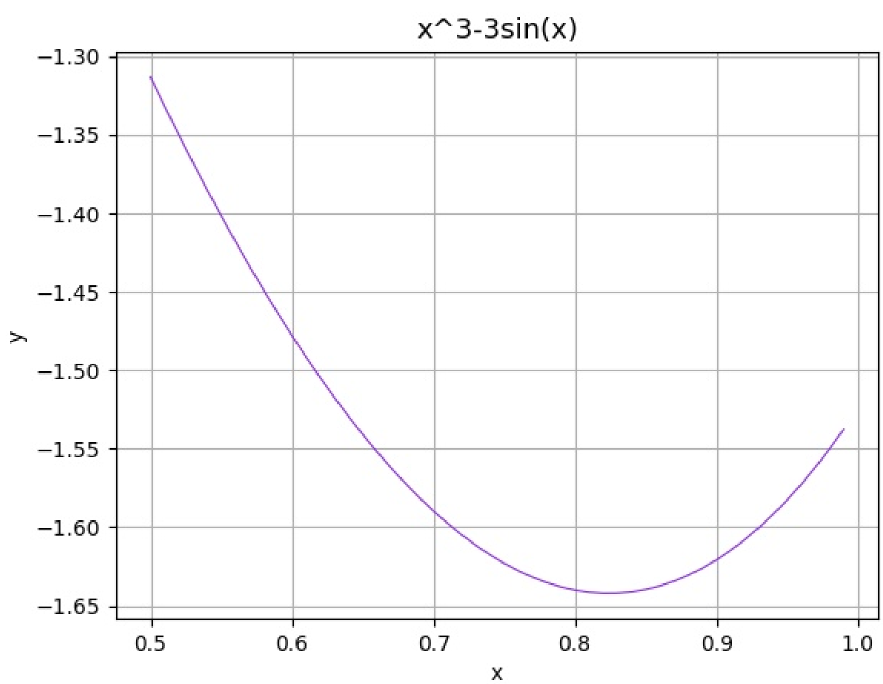
Большим достоинством прямых методов является то, что от целевой функции не требуется дифференцируемости и, более того, она может быть не задана в аналитическом виде. Единственное, на чём основаны алгоритмы прямых методов минимизации, это возможность определения значений $f(x)$ в заданныых точках.

Рассмотрим метод дихотомии и Фибоначчи. Самым слабым требованием на функцию $f(x)$, позволяющим использовать эти методы, является, является её унимодальность.

Опредление 1. *Функция $f(x)$ называется унимодальной, если для $x \in [a;\ b]$ существует единственная точка глобального минимума, слева от которой $f(x)$ монотонно убывает, а справа - монотонно возрастает.*

Поэтому проверим выполнение данного условия для наших функций. Для этого построим их графики.





Видим, что функции являются унимодальными, значит для них можем применить прямые методы поиска минимального значения.

3 Описание алгоритмов

3.1 Метод дихотомии

Пусть $a < x_1 < x_2 < b$. Сравнив значения $f(x)$ в точках x_1 и x_2, можно сократить отрезок поиска точки x^*, перейдя к отрезку $[a;\ x_2]$, если $f(x_1) \leq f(x_2)$, или отрезку $[x_1;\ b]$, если $f(x_1) > f(x_2)$. Описанную процедуру можно повторить необходимое число раз, последовательно уменьшая отрезок, содержащий точку минимума. Когда длина последнего из найденных отрезков станет достаточно малой, следует положить $x^* = \bar{x}$, где $\bar{x}$ – одна из точек этого отрезка, например его середина.

В методе дихотомии точки x_1 и x_2 располагаются близко к середине очередного отрезка $[a;\ b]$, т.е

$$x_1 = \frac{b + a - \delta}{2}$$

$$x_2 = \frac{b + a + \delta}{2}$$

где $\delta > 0$ – малое число.

В конце вычислений в качестве приближённого значения x^* берут середину последнего из найденных отрезков $[a;\ b]$, убедившись предварительно, что достигнуто неравенство $b - a \leq \varepsilon$. Опишем алгоритм метода

1. Вычисляем $x_1 = \frac{b+a-\delta}{2}$ и $x_2 = \frac{b+a+\delta}{2}$. Вычисляем $f(x_1)$ и $f(x_2)$.

2. Сравниваем $f(x_1)$ и $f(x_2)$. Если $f(x_1) \leq f(x_2)$, то перейти к отрезку $[a;\ x_2]$, положив $b = x_2$, иначе – к отрезку $[x_1;\ b]$, положив $a = x_1$.

3. Найти достигнутую точность $\varepsilon_n = b - a$. Если $\varepsilon_n > \varepsilon$, то перейти к следующей итерации, вернувшись к шагу 1. Если $\varepsilon_n \leq \varepsilon$, то завершить поиск x^*, перейдя к шагу 4.

4. Положить $x^* = \bar{x} = \frac{a+b}{2}, \quad f^* = f(\bar{x})$

3.2 Выбор числа δ

Число δ выбирается на интервале $(0; \varepsilon)$ с учётом следующих соображений:

1. Чем меньше δ, тем больше относительное уменьшение длины отрезка на каждой итерации, т.е. при уменьшении δ достигается более высокая скорость сходимости метода дихотомии.

2. При чрезмерно малом δ сравнение значений $f(x)$ в точках x_1 и x_2, отличающихся на величину δ, становится затруднительным. Поэтому выбор δ должен быть согласован с точностью определения $f(x)$ и с количеством верных десятичных знаков при задании аргумента x.

3. В нашей работе будем использовать δ равную 0.1% от текущей длины отрезка $[a_k;\ b_k]$

3.3 Метод Фибоначчи

Суть заключается в том, что здесь генерируется последовательность чисел Фибоначчи:

$$F_0 = F_1 = 1$$

$$F_{k+1} = F_k + F_{k-1}$$

1. Выбираем допустимую конечную длину интервала неопределенности l и константу различимости ε

2. Выбираем общее число вычислений функции n так, что $F_n > \frac{b-a}{l}$

3. Положим $\lambda_1 = a_1 + \frac{F_{n-2}}{F_n}(b-a)$ и $\mu_1 = a_1 + \frac{F_{n-1}}{F_n}(b-a)$

4. Вычислим значения функции в точках λ_1 и μ_1

5. Положим $k = 1$

6. Если значение функции в точке λ_k больше, чем в точке μ_k перейдем к пункту 7, иначе к пункту 8

7. Положим $a_{k+1} = \lambda_k$, $b_{k+1} = b_k$, $\lambda_{k+1} = \mu_k$, $\mu_{k+1} = a_{k+1} + \frac{F_{n-k-1}}{F_{n-k}}(b_{k+1} - a_{k+1})$
   Если $k = n-2$, то перейдем к пункту 10, иначе вычисляем функцию в точке μ_{k+1} и переходим к пункту 9

8. Положим $a_{k+1} = a_k$, $b_{k+1} = \mu_k$, $\mu_{k+1} = \lambda_k$, $\lambda_{k+1} = a_{k+1} + \frac{F_{n-k-2}}{F_{n-k}}(b_{k+1} - a_{k+1})$
   Если $k = n-2$, то перейдем к пункту 10, иначе вычисляем функцию в точке μ_{k+1} и переходим к пункту 9

9. Заменяем k на $k+1$ и переходим к пункту 6

10. Положим $\lambda_n = \lambda_{n-1}$, $\mu_n = \lambda_n + \varepsilon$. Если функция в точке λ_n равна функции в точке μ_n, то положим $a_n = \lambda_n$, $b_n = b_{n-1}$, иначе $a_n = a_{n-1}$, $b_n = \mu_n$. Повторяем пока заданный интервал неопределенности не удовлетворяет заданной точности.

После того, как мы этот метод реализуем многократно, мы получим, что метод сходится в одну точку и она будет иметь следующие координаты $x^* = \frac{a_{n-1}+b_{n-1}}{2}$

4 Результаты решения задачи

Для функции $f(x) = x^2 - 2x + e^{-x}$ на интервале $[1;\ 1.5]$ при заданной точности $\varepsilon = 0.05$ метод дихотомии дал результат:

$$x^* = 1.172$$

$$f(x^*) = -0.661$$

метод Фибоначчи для этой же функции выдал значения, равные:

$$x^* = 1.179$$

$$f(x^*) = -0.660$$

Если рассматривать функцию $f(x) = x^3 - 3sin(x) \quad [0.5;\ 1] \quad \varepsilon = 0.05$, то методом дихотомии получим результат:

$$x^* = 0.828$$

$$f(x^*) = -1.642$$

методом Фибоначчи:

$$x^* = 0.845$$

$$f(x^*) = -1.641$$

5 Аналитическая оценка числа обращений к функции

5.1 Аналитическая оценка числа обращений к функции метода дихотомии

Обозначим длину исходного отрезка $[a;\ b]$ через Δ_0. Длина отрезка, полученного после первой итерации, будет

$$\Delta_1 = \frac{\Delta_0}{2} + \frac{\delta}{2}$$

после второй итерации

$$\Delta_1 = \frac{\Delta_1}{2} + \frac{\delta}{2} = \frac{b-a}{4} + \delta(\frac{1}{4} + \frac{1}{2})$$

после третьей

$$\Delta_3 = \frac{\Delta_2}{2} + \frac{\delta}{2} = \frac{b-a}{8} + \delta(\frac{1}{8} + \frac{1}{4} + \frac{1}{2})$$

и т.д. Таким образом, в результате n итераций длина отрезка поиска точки x^* станет

$$\Delta_n = \frac{b-a}{2^n} + (\frac{1}{2^n} + \frac{1}{2^{n-1}} + \cdots + \frac{1}{2}) = \frac{b-a}{2^n} + (1 - \frac{1}{2^n})\delta$$

При этом будет достигнута точность определения точки минимума $\varepsilon_n = \Delta_n$

$$\varepsilon_n = \frac{b-a}{2^n} + (1 - \frac{1}{2^n})\delta \leq \varepsilon$$

$$\frac{b-a}{2^n} + \delta - \frac{\delta}{2^n} \leq \varepsilon$$

$$\frac{b-a}{2^n} - \frac{\delta}{2^n} \leq \varepsilon - \delta$$

$$\frac{b-a-\delta}{2^n} \leq \varepsilon - \delta$$

$$\frac{b-a-\delta}{\varepsilon - \delta} \leq 2^n$$

$$n \geq log_2 \frac{b-a-\delta}{\varepsilon - \delta}$$

Получили формулу для подсчёта числа итераций. Подставим в полученную формулу наши данные:

$$f(x) \ = \ x^2 - 2x + e^{-x} \quad [1;\ 1.5] \quad \varepsilon = 0.05$$

$$n \geq log_2 \frac{1.5 - 1 - 0.0005}{0.05 - 0.0005} = 3.32$$

Значит $n = 4$. Теперь учтём, что на каждой итерации обращаемся к функции ровно два раза, значит $m = 2n$. Таким образом число вызовов функции равняется 8.

Повторим те же действия для второй функции:

$$f(x) \ = \ x^3 - 3sin(x) \quad [0.5;\ 1] \quad \varepsilon = 0.05$$

$$n \geq log_2 \frac{1 - 0.5 - 0.0005}{0.05 - 0.0005} = 3.32$$

Аналогично предыдущему получаем $n = 4,\ m = 8$.

Результаты, полученные с помощью программного кода, полностью совпали с нашими ожиданиями.

5.2 Аналитическая оценка числа обращений к функции метода Фибоначчи

Число итераций, необходимое для достижения заданной точности ε , можно найти из условия $\varepsilon_n \leq \varepsilon$ с учётом соотношения

$$\varepsilon_k = b_{k-1} - a_{k-1} = \frac{F_{n-k}}{F_{n-k+1}} \frac{F_{n-(k-1)}}{F_{n-(k-2)}}(b_{k-1} - a_{k-1}) = \frac{F_{n-k}}{F_n}(b-a)$$

Пусть $k = n-1$, тогда:

$$\varepsilon \leq \frac{b-a}{F_n}$$

Можем показать, что

$$\frac{1}{F_n} \longrightarrow (\frac{\sqrt{5}-1}{2})^{n-1} = (0.61803)^{n-1}$$

Тогда имеем:

$$\varepsilon \leq (b-a)(0.61803)^{n-1}$$

$$0.61803 \cdot \varepsilon \leq (b-a)(0.61803)^{n}$$

$$\frac{0.61803 \cdot \varepsilon}{b-a} \leq (0.61803)^{n}$$

$$log_2 \frac{0.61803 \cdot \varepsilon}{b-a} \leq log_2 (0.61803)^{n}$$

$$log_2 \frac{0.61803 \cdot \varepsilon}{b-a} \leq n \cdot log_2 (0.61803)$$

$$n \geq \frac{log_2 \frac{0.61803 \cdot \varepsilon}{b-a}}{log_2(0.61803)}$$

Подставим в полученную формулу наши значения:

$$f(x) \ = \ x^2 - 2x + e^{-x} \quad [1;\ 1.5] \quad \varepsilon = 0.05$$

$$n \geq \frac{log_2(\frac{0.61803 \cdot 0.05}{1.5-1})}{log_2(0.61803)} = \frac{log_2(0.061803)}{log_2(0.61803)} = 5.78$$

Если число итераций n, то число вызова функции будет $m = n+1$, так как на первой итерации необходимо вычислить значение функции в 2-ух точках. Значит $n = 6$, а $m = 7$.
Аналогичные действия проделаем для второй функции:

$$f(x) \ = \ x^3 - 3sin(x) \quad [0.5;\ 1] \quad \varepsilon = 0.05$$

$$n \geq \frac{log_2(\frac{0.61803 \cdot 0.05}{1-0.5})}{log_2(0.61803)} = \frac{log_2(0.061803)}{log_2(0.61803)} = 5.78$$

$n = 6$ и $m = 7$

Фактическое и теоретическое количество вызовов функции на нашем примере совпало.

6 Сравнительный анализ числа обращений от заданной точности

Проведём ряд экспериментов, на основе которых составим таблицы фактического обращения к функции от заданного числа ε.

ε	результат $f(x^*)$	число обращений	коэффициент сжатия
	метод дихотомии		
0.1	-0.66	6	0.125751
0.01	-0.661	12	0.01581
0.001	-0.6609	18	0.001988
	метод Фибоначчи		
0.1	-0.66	5	0.125
0.01	-0.661	10	0.0112
0.001	-0.6609	15	0.00101

Таблица 1: $f(x) = x^2 - 2x + e^{-x}$

ε	результат $f(x^*)$	число обращений	коэффициент сжатия
	метод дихотомии		
0.1	-1.64	6	0.125751
0.01	-1.642	12	0.01581
0.001	-1.6421	18	0.001988
	метод Фибоначчи		
0.1	-1.64	5	0.125
0.01	-1.642	10	0.0112
0.001	-1.6421	15	0.00101

Таблица 2: $f(x) = x^3 - 3sin(x)$

7 Оценка достоверности полученного результата

Лемма 1. *Пусть* $f(x)$ *– унимодальная на* $[a;\ b]$. *Пусть* $x_1,\ x_2 \in [a;\ b]$ и $x_1 < x_2$. *Тогда:*

- *если* $f(x_1) \geq f(x_2)$, *то* $x^* \notin [a;\ x_1]$
- *если* $f(x_1) \leq f(x_2)$, *то* $x^* \notin [x_2;\ b]$

Воспользуемся пакетом MATLAB2020b и найдем минимум функции на заданном отрезке

$$f(x) \ = \ x^2 - 2x + e^{-x} \ \ [1; \ 1.5]$$

```
xRes1 =

    1.1572


yRes1 =

   -0.6609


exitflag1 =

     1


output1 =

  struct with fields:

    iterations: 5
     funcCount: 6
     algorithm: 'golden section search, parabolic interpolation'
       message: 'Optimization terminated:↵ the current x satisfies the termination criteria using OPTIONS.TolX of 5.000000e-02
```

Теперь проверим, что найденное значение действительно принадлежит всем интервалам неопределённости, которые находит наша программа метода дихотомии:

$$1.1572 \in [1, 1.5]$$
$$1.1572 \in [1, 1.2505]$$
$$1.1572 \in [1.1249995, 1.2505]$$
$$1.1572 \in [1.1249995, 1.1878752505]$$
$$1.1572 \in [1.1563744994995, 1.1878752505]$$

Аналогично проверим для метода Фибоначчи:

$$1.1572 \in [1, 1.5]$$
$$1.1572 \in [1, 1.3095238095238095]$$
$$1.1572 \in [1.119047619047619, 1.3095238095238095]$$
$$1.1572 \in [1.119047619047619, 1.2380952380952381]$$
$$1.1572 \in [1.119047619047619, 1.1904761904761905]$$
$$1.1572 \in [1.1428571428571428, 1.1904761904761905]$$
$$1.1572 \in [1.1666666666666667, 1.1904761904761905]$$

Для второй функции проведем аналогичные выкладки:

$$f(x) \; = \; x^3 - 3sin(x) \quad [0.5; \; 1]$$

```
xRes2 =

    0.8257


yRes2 =

   -1.6421


exitflag2 =

     1


output2 =

  struct with fields:

    iterations: 4
     funcCount: 5
     algorithm: 'golden section search, parabolic interpolation'
       message: 'Optimization terminated:↵ the current x satisfies the termination criteria using OPTIONS.TolX of 5.000000e-02
```

Метод дихотомии:

$$0.8257 \in [0.5, 1]$$
$$0.8257 \in [0.7495, 1]$$
$$0.8257 \in [0.7495, 0.8750005000000001]$$
$$0.8257 \in [0.8121247495, 0.8750005000000001]$$
$$0.8257 \in [0.8121247495, 0.8436255005005001]$$

Метод Фибоначчи:

$$0.8257 \in [0.5, 1]$$
$$0.8257 \in [0.6904761904761905, 1]$$
$$0.8257 \in [0.6904761904761905, 0.8809523809523809]$$
$$0.8257 \in [0.7619047619047619, 0.8809523809523809]$$
$$0.8257 \in [0.8095238095238095, 0.8809523809523809]$$
$$0.8257 \in [0.8095238095238095, 0.8571428571428571]$$
$$0.8257 \in [0.8333333333333333, 0.8571428571428571]$$

В обоих случаях x^* лежат в полученных интервалах неопределённости, значит можем заключить, что методы работают корректно.

8 Дополнительные исследования

Попытаемся ответить на вопрос: сколько чисел Фибоначчи выгоднее брать 5 или 15 и почему? Так как N вычислений $f(x)$ позволяют выполнить $N-1$ итерации метода Фибоначчи, то достигнутая в результате этих вычислений точность определения x^* составляет

$$\varepsilon(N) = (\frac{\sqrt{5}-1}{2})^{N-1} \cdot (b-a)$$

$$\varepsilon(5) = (0.61803)^{5-1} \cdot (1.5-1) = 0.07294901687$$

$$\varepsilon(15) = (0.61803)^{15-1} \cdot (1.5-1) = 0.00059312064$$

С ростом числа N точность ε будет улучшаться, значит выгоднее брать 15 чисел.

9 Программная реализация

В процессе реализации алгоритмов использовался язык программирования Python3.6. Для проверки полученных решений пользовались пакетом MATLAB2020b.

Исходный код находится в системе контроля версий GitHub
https://github.com/Brightest-Sunshine/Optimization-methods-2021

Список литературы

[1] Лесин В. В., Лисовец Ю. П. Основы методов оптимизации: Учебное пособие. 3-е изд., испр, – СПб.: Издательство «Лань», 2011. – 352с.