Математическое программирование.
Лекция 1

Краткое введение

Математическое программирование (следует отличать от программирования для ЭВМ) это раздел математики, в основном, разрабатывающий теорию и численные методы решения многомерных экстремальных задач. Данный курс посвящен скорее практическому применению данных методов, поскольку методы математического программирования достаточно давно и широко реализованы в различных программных пакетах компьютерной алгебры, а так же в виде библиотек к языкам программирования. Данный курс несколько расширен в сторону некоторых численных методов и современных методов обработки информации (нейронные сети), которые традиционно не включаются в курсы математического программирования, которые обычно посвящены исключительно решению экстремальных задач.

Применение численных методов встроенных в ПО, позволяет специалистам, не являющимся математиками, получать решение своих практических задач относительно легко, однако необходимость изучения применимости различных методов МП для решения различных классов задач все равно остается. В настоящее время даже перечисление систем компьютерной алгебры и языков программирования, использующих библиотеки для вычислений, представляет собой непростую задачу. Можно отметить такие системы как Matlab, Octave, SciLab, R. Широко развито применение всевозможных библиотек в таком очень популярном в настоящее время языке программирования как Python. Можно найти библиотеки к таким классическим языкам программирования как Fortran, C/C++, Pascal (например, библиотека GSL).

Численные методы, в отличие от аналитических решений различных задач, по своей природе являются приближенными, обеспечивающими поиск решения задачи с заданной точностью. Кроме того, следует понимать, что компьютер это устройство реализующее любые вычисления с действительными числами с конечной точностью.

Решение уравнения вида $f(x)=0, \quad x \in R$

Для того, чтобы выяснить, есть ли вообще корни у этого уравнения, используются различные *алгоритмы отделения корней*. В основном, идея основана на нахождении таких интервалов разбиения внутри интервала интереса, чтобы

$$f(a)f(b) < 0, \quad a < b \quad ,$$

где значения *a* и *b* принадлежать соседним интервалам разбиения, и поскольку произведения меньше нуля, функция меняет знак где-то между этими значениями. При использовании системы компьютерной алгебры, часто достаточно просто построить график внутри интервала интереса и посмотреть на него. Пусть для примера в интервале [-5,5] нам нужны корни уравнения

$$0.03x^3 - 0.3x^2 + 0.1 + 5\sin(x) = 0 \quad .$$

Сценарий в Octave (почти также в Matlab и SciLab) представлен на Рис. 1.

Метод дихотомии

Метод годится для отыскания единственного корня уравнения $f(x)=0$ на интервале [*a*,*b*]. В предыдущем примере таким отрезком может быть, например, отрезок [2,4]. Алгоритм можно описать в виде следующей последовательности шагов:
0. вычисляем значения f(*a*) и f(*b*), можно убедиться, что $f(a)f(b)<0$;
1. делим отрезок [*a*,*b*] пополам, найдя его середину $x=(a+b)/2$;

```
x=-5:0.05:5;

y=0.03*x.^3-0.5*x.^2+0.1+5*sin(x);

plot(x,y,'r',x,zeros(size(x)),'b');
```

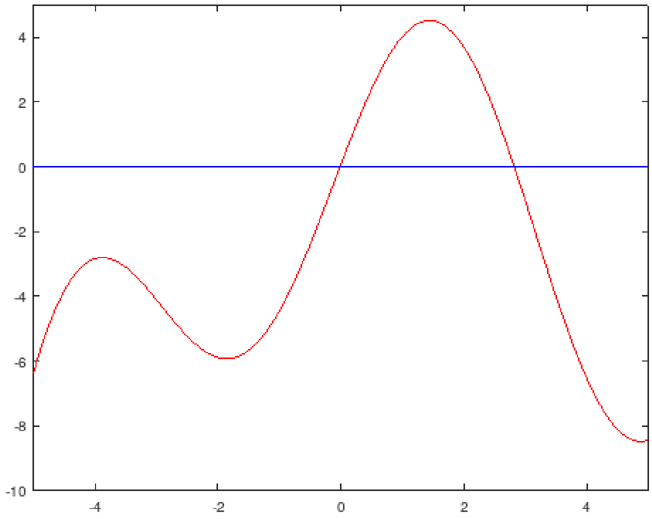

Рис.1 Сценарий и график функции в Octave.

2. вычисляем $f(x)$;
3. если $f(x)f(a) > 0$, то новое значение $a = x$ и переходим к шагу 5;
4. если $f(x)f(a) < 0$, то новое значение $b = x$;
5. проверяем достигнутую точность, если $|a - b| < \varepsilon$, то заканчиваем вычисления, если нет переходим к шагу 1.

В качестве решения, полученного с заданной точностью можно взять среднее значение последнего интервала или вычислить решение поточнее с помощью линейной интерполяции.

Пример реализации в Octave:

```
function y = f(x);
  y = 0.03*x.^3 - 0.3*x.^2 + 0.1 + 5*sin(x);
endfunction;

a = 2;
b = 4;

fa = f(a);
fb = f(b);
tol = 0.01;
if (fa*fb > 0)
  disp("Неверно выбранный интервал!");
  stop
endif;

while ((b - a) > tol)
 x = (a+b)/2;
 if (fa*f(x) > 0)
   a = x;
 else
   b = x;
 endif
endwhile
disp("Root is found");
disp((a+b)/2);
```

Можно нарисовать график и положить на него решение...

Метод простых итераций

Метод используется в тех случаях, когда уравнение $f(x)=0$ можно представить в виде $x=\varphi(x)$. Метод заключается в том, чтобы выбрать первое приближение $x_1 \in [a,b]$ и вычислить следующее как $x_2=\varphi(x_1)$ и так далее, что дает общую формулу процесса в виде

$$x_{n+1}=\varphi(x_n) \quad .$$

Условие останова можно сформулировать как неравенство $|x_{n+1}-\varphi(x_n)|<\varepsilon$. Метод не всегда сходится к решению, поэтому важно понимать когда можно воспользоваться этим методом, а когда нет. Процесс итераций сходиться при условии $|\varphi'(x)|<1$, следовательно имеет смысл посчитать производные на концах интервала и начать с того, где это условие выполняется. Геометрическую интерпретацию метода для случаев, когда процесс сходится, а когда нет, можно увидеть на Рис.2

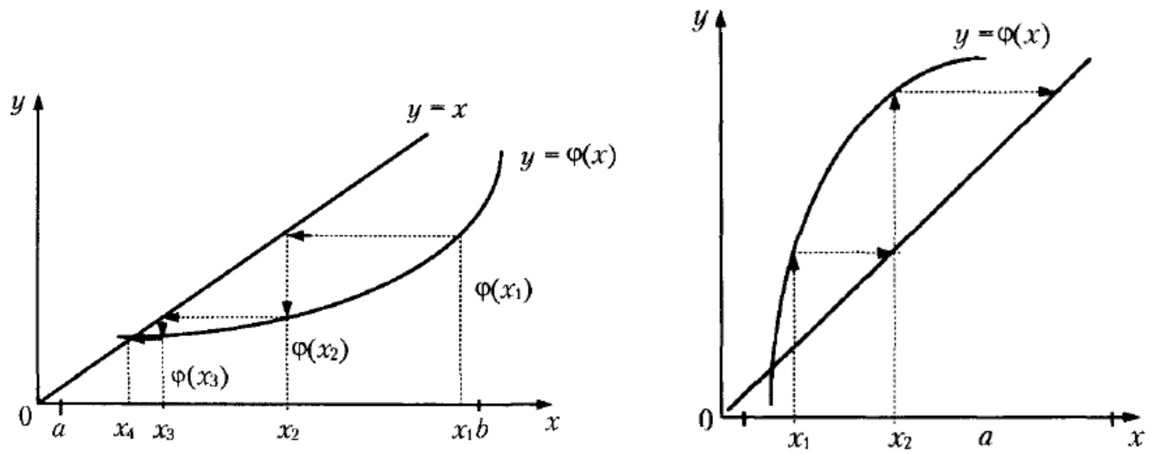

Рис. 2. Случаи когда метод простой итерации сходится (слева) и расходится (справа).

Метод касательных (метод Ньютона)

Продолжим рассматривать ситуацию, когда уравнение $f(x)=0$ имеет единственный корень на отрезке [*a*,*b*], причем первая и вторая производные функции $f(x)$ определены, непрерывны и сохраняют свой знак на отрезке [*a*,*b*]. Метод проиллюстрирован на Рис.3.

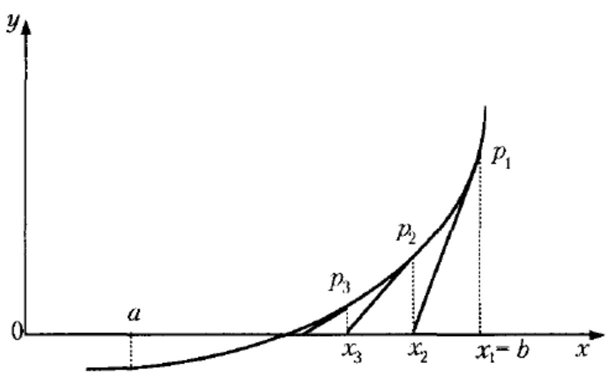

Рис. 3. Иллюстрация метода касательных

В качестве начального приближения выбирается точка $x_1=b$ и через точку p_1 с координатами $(x_1, f(x_1))$ проводится касательная к кривой $y=f(x)$, затем находится ее пересечение с осью x, абсцисса этой точки x_2 становится вторым приближением и т. д. Уравнение касательной

$$y = f(x_1) + f'(x_1)(x - x_1) ,$$

точка x_2 определяется из условия $y = 0$, что приводит к выражению

$$x_2 = x_1 - f(x_1)/f'(x_1) .$$

Рекуррентная же формула процесс имеет вид

$$x_{n+1} = x_n - f(x_n)/f'(x_n) .$$

Если же в качестве начального приближения в ситуации соответствующей Рис. 3 взять точку a, то пересечение касательной с осью x окажется вне интервала $[a,b]$. Поэтому в качестве начального приближения выбирается тот конец интервала, для которого $f(x)f''(x) > 0$, это условие сходимости метода является достаточным, но не необходимым (при невыполнении метод, однако, может все таки сойтись).

Метод хорд

Пусть уравнение $f(x) = 0$ имеет единственный корень на отрезке $[a,b]$, причем первая и вторая производные функции $f(x)$ определены, непрерывны и сохраняют свой знак на отрезке $[a,b]$. На Рис. 4 представлена геометрическая иллюстрация метода.

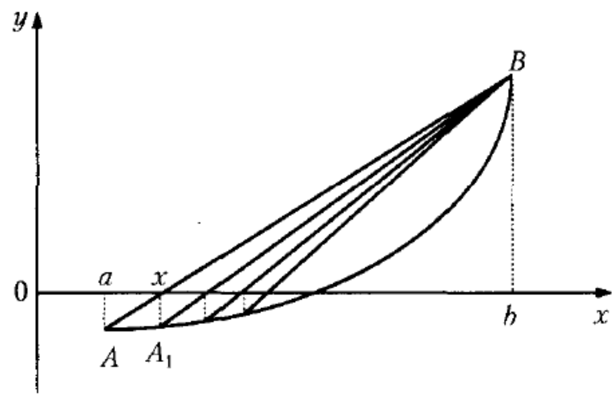

Рис. 4. Геометрическая интерпретация метода хорд.

В этом методе проводится прямая через точки, соответствующие концам интервала на оси x и значениям функции для них. Новый конец интервала определяется в точке пересечения очередной прямой с осью абсцисс и т. д. Для нашей иллюстрации точка a является подвижной, а точка b остается той же. В качестве подвижной выбирается точка для которой выполняется условие $f(x)f''(x) < 0$. Уравнение прямой, проходящей через две точки

$$y = f(a) + \frac{f(b) - f(a)}{b - a}(x - a) ,$$

а точка пересечения с осью абсцисс определяется из условия $y = 0$, т. е.

$$x = a - f(a) \frac{b - a}{f(b) - f(a)} ,$$

после чего определяется новое $a = x$ и процесс повторяется до достижения условия останова.

Здесь представлены не все возможные методы, далее при изучении элементов выпуклого анализа и нелинейной оптимизации будут рассмотрены еще два: метод квадратичной интерполяции и метод золотого сечения.