

Методы оптимизации

Лекция 2. Численные методы решения задач одномерной безусловной оптимизации.

*Селина Елена
Георгиевна
Ауд. 302*

Литература

1. Банди Б. Методы оптимизации. Вводный курс. - М.: Радио и связь, 1988. - 128 с.
2. Банди Б. Основы линейного программирования. - М.: Радио и связь, 1989. - 176 с.
3. Лесин, В.В. Основы методов оптимизации.— СПб: Лань, 2016. — 344 с.
4. Сухарев, А.Г. Курс методов оптимизации — Москва: Физматлит, 2011.— 384 с

Аналитические методы исследования функции на экстремум можно использовать в тех случаях, когда функция $f(x)$ и ее производные имеют достаточно простой вид. Однако зачастую в практических задачах решение уравнения

$$f'(x) = 0$$

и даже просто вычисление производной $f'(x)$ представляет большие трудности. Кроме того, в практических задачах часто неизвестно, является ли $f(x)$ дифференцируемой функцией. Поэтому существенное значение приобретают численные методы минимизации, не требующие вычисления производной и основанные на исследовании поведения функции в некоторых специально подбираемых точках в соответствии с определенным алгоритмом. Такие методы называются прямыми методами минимизации.

Прямые методы минимизации

Определение. Функция f на действительном отрезке $[a, b]$ называется унимодальной, если она имеет минимум $x^* \in [a, b]$ и если для любых $\alpha, \beta \in [a, b]$ ($\alpha < \beta$) выполняются соотношения:

$f(\alpha) > f(\beta)$ при $\beta \leq x^*$,

$f(\alpha) < f(\beta)$ при $\alpha \geq x^*$.

Численные методы минимизации, как правило, применяются к унимодальным на рассматриваемом отрезке функциям.

Определение. Методы, использующие только значения функции и не требующие вычисления ее производных, называются прямыми методами минимизации.

Метод деления отрезка пополам

Задаются a , b и погрешность ε .

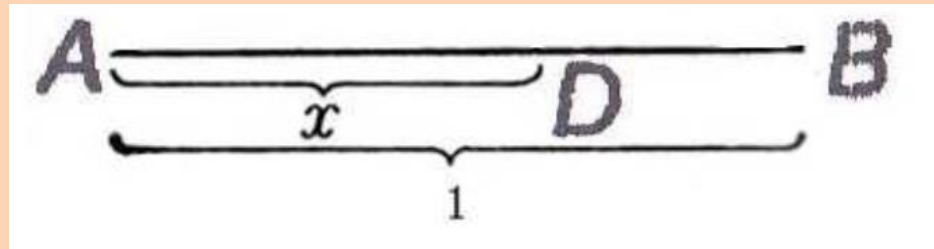
- Берем две точки вблизи середины интервала $[a, b]$:
 $x_1 = (a + b - \varepsilon) / 2$, $x_2 = (a + b + \varepsilon) / 2$.
- Вычисляем $y_1 = f(x_1)$, $y_2 = f(x_2)$.
- Если $y_1 > y_2$, тогда присваивается $a = x_1$,
иначе присваивается $b = x_2$
- Если $b - a > 2\varepsilon$, тогда повторяем с п.1, иначе переходим к пункту 5.
- Вычисляем $x_m = (a + b) / 2$, $y_m = f(x_m)$.
- Конец.

Этот метод прост в реализации, позволяет находить минимум разрывной функции, однако требует большого числа вычислений функции для обеспечения заданной точности.

Метод золотого сечения

Метод золотого сечения основан на делении отрезка локализации «золотым сечением», т.е. таком делении, когда отношение большей части отрезка ко всему отрезку равно отношению меньшей части к большей.

$$\frac{AB}{AD} = \frac{AD}{DB} = \frac{1 + \sqrt{5}}{2}$$



Алгоритм метода схож с методом половинного деления, за исключением только способа расчета точек сравнения x_1 , x_2 , которые теперь вычисляются в пропорции золотого сечения. В этом методе выбор нового интервала неопределенности происходит по результатам сравнения функции в двух точках. Но в отличие от метода половинного деления, на каждой итерации, кроме первой, вычислении функции производится только в одной точке.

На первом шаге (итерации) точки вычисляются по формулам:
 $x_1 = a + 0,382(b-a)$, $x_2 = a + 0,618(b-a)$

Затем вычисляются значения функции в этих точках.

Возможны два случая:

Если $f(x_1) < f(x_2)$, то оставляем отрезок $[a, x_2]$. На второй итерации x_2 полагаем равным x_1 , а x_1 вычисляем по формуле $x_1 = a + 0,382(x_2 - a)$. Значение функции вычисляется только в точке x_1 , так как значение функции в x_2 уже было вычислено на предыдущем шаге.

Если $f(x_1) > f(x_2)$, то оставляем отрезок $[a, x_2]$. На второй итерации x_1 полагаем равным x_2 , а x_2 вычисляем по формуле $x_2 = a + 0,618(b - x_1)$. Значение функции вычисляется только в точке x_2 , так как значение функции в x_1 уже было вычислено на предыдущем шаге.

Вычисления продолжают до тех пор, пока длина интервала не станет меньше требуемой точности.

Метод Фибоначчи

Последовательность чисел Фибоначчи $\{F_n\}$, $n=1,2,3,\dots$ подчиняется соотношению

$$F_{n+2} = F_n + F_{n+1}, \text{ где } F_1 = F_2 = 1$$

и имеет вид

1, 1, 2, 3, 4, 8, 13, 21, 34, 55, 89, 144, 233,...

С помощью метода математической индукции можно показать, что n -е число Фибоначчи вычисляется по формуле Бинэ:

$$F_n = \frac{1}{\sqrt{5}} \left[\left(\frac{1+\sqrt{5}}{2} \right)^n - \left(\frac{1-\sqrt{5}}{2} \right)^n \right], \text{ где } n=1,2,3,\dots$$

В методе Фибоначчи выбор нового интервала неопределенности происходит по результатам сравнения функции в двух точках. Как и в методе золотого сечения, на каждой итерации, кроме первой, вычислении функции производится только в одной точке.

В отличие от других методов число итераций определяется сразу. Для достижения требуемой точности ε число итераций n следует задавать исходя из соотношения

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

Условием окончания вычислений является выполнение заданного количества вычислений n . На первой итерации полагаем $k=1$ и вычисляем точки по формулам

$$x_1 = a + \frac{F_{n-k}}{F_{n+1-k}} (b - a) \quad (18)$$

$$x_2 = a + \frac{F_{n-1-k}}{F_{n+1-k}} (b - a) \quad (19)$$

Затем вычисляются значения функции в этих точках.

Возможны два случая:

$f(x_1) < f(x_2)$ - выбираем отрезок $[a, x_2]$. На второй итерации полагаем $a_1 = a$, $b_1 = x_1$, x_2 равным x_1 , а x_1 вычисляем по формуле (18). Значение функции вычисляется только в точке x_1 , так как значение функции в x_2 уже было вычислено на предыдущем шаге. И переходят к следующей итерации.

$f(x_1) \geq f(x_2)$ - выбираем отрезок $[x_1, b]$. На второй итерации полагаем $a_1 = x_1$, $b_1 = b$, x_1 равным x_2 , а x_2 вычисляем по формуле (19). Значение функции вычисляется только в точке x_1 , так как значение функции в x_2 уже было вычислено на предыдущем шаге. И переходят к следующей итерации.

Вычисления продолжают пока не будут исчерпаны все n .

Методы, использующие информацию о производных целевой функции

Пусть теперь $f(x)$ является дифференцируемой или дважды дифференцируемой выпуклой функцией и возможно вычисление производных $f'(x)$ в произвольно выбранных точках. В этом случае эффективность поиска точки минимума можно существенно повысить.

Рассмотрим три метода минимизации, в которых используются значения производных целевой функции:

- метод средней точки;
- метод хорд;
- метод Ньютона.

Из курса математического анализа известно, что для выпуклой дифференцируемой функции равенство $f'(x)=0$ является не только необходимым, но и достаточным условием глобального минимума. Поэтому, если известно, что является внутренней точкой отрезка, то приближенное равенство $f'(x) \approx 0$ или $|f'(x)| \leq \varepsilon$ может служить условием остановки вычислений в рассматриваемых трех методах.

Метод средней точки

Будем искать минимум функции $f(x)$, непрерывно дифференцируемой и строго унимодальной на отрезке $[a, b]$.

В этом случае единственной точкой $x^* \in [a, b]$ минимума будет стационарная точка, в которой $f'(x^*) = 0$. Отметим, что непрерывно дифференцируемая унимодальная на отрезке функция может иметь на нем более одной стационарной точки. На каждом шаге на отрезке определяются две точки a_k, b_k , в которых производные имеют разные знаки, $f'(a_k) f'(b_k) < 0$. Искомый минимум находится между ними. Делим интервал пополам, из двух интервалов оставляем тот, на концах которого производная имеет разные знаки.

Алгоритм метода средней точки

Шаг 1. Определим точность $\varepsilon > 0$, $a_1 = a$, $b_1 = b$, вычисляем $f'(a_1) < 0$, $f'(b_1) > 0$

Вычисляем $x_1 = \frac{a_1 + b_1}{2}$, $k = 1$

Шаг 2 Вычисляем $f'(x_k)$, если $f'(x_k) = 0$, или $f'(x_k) < \varepsilon$, то $x_k = \frac{a_k + b_k}{2}$ — точка минимума, конец.

Шаг 3 Если $f'(x_k) < 0$, то $a_{k+1} = x_k$, $b_{k+1} = b_k$, иначе $a_{k+1} = a_k$, $b_{k+1} = x_k$

Шаг 4 $x_k = \frac{a_k + b_k}{2}$, переход на пункт 2.

Метод средней точки напоминает метод дихотомии, но сходится к искомому значению x^* быстрее.

Метод хорд

Как уже отмечалось, равенство $f'(x)=0$ является необходимым и достаточным условием глобального минимума выпуклой дифференцируемой функции $f(x)$. Если на концах отрезка $[a,b]$ производная $f'(x)$ имеет разные знаки, т.е. $f'(a) \cdot f'(b) < 0$, и она непрерывна, то на интервале (a,b) найдется точка, в которой $f'(x)$ обращается в нуль. В этом случае поиск точки минимума на отрезке эквивалентен решению уравнения

$$f'(x) = 0, \quad x \in (a, b) \quad (20)$$

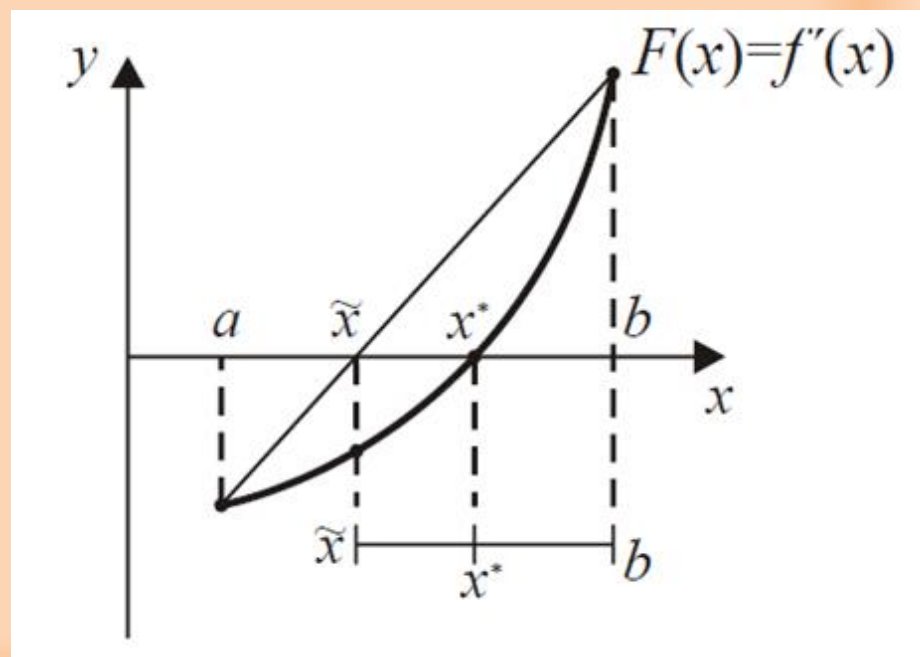
Таким образом, при $f'(a) \cdot f'(b) < 0$, любой приближенный метод решения уравнения (20) можно рассматривать как метод минимизации выпуклой непрерывно дифференцируемой функции $f(x)$ на отрезке $[a,b]$.

Сущность приближенного решения уравнения $F(x)=0$ на отрезке $[a,b]$ при $F(a) \cdot F(b) < 0$ методом хорд состоит в исключении отрезков путем определения \tilde{x} - точки пересечения с осью OX хорды графика функции $F(x)$ на отрезке $[a,b]$, представленного на рисунке.

Полагая $F(x) = f'(x)$, запишем координату точки \tilde{x}

$$\tilde{x} = a - \frac{f'(a)}{f'(a) - f'(b)} (a - b) \quad (21)$$

Отрезок дальнейшего поиска точки \tilde{x} ($[a, \tilde{x}]$ 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 метода хорд

Шаг 1. Находим \tilde{x} по формуле (21). Вычисляем $f'(\tilde{x})$ и переходим к шагу 2.

Шаг 2. Проверка на окончание поиска: если $|f'(\tilde{x})| \leq \varepsilon$ то положить $x^* = \tilde{x}$,

$f^* = f(\tilde{x})$, и завершить поиск, иначе перейти к шагу 3.

Шаг 3. 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 1.

Метод Ньютона

Если выпуклая на отрезке $[a,b]$ функция $f(x)$ дважды непрерывно дифференцируема на этом отрезке, то точку $x \in [a,b]$ минимума этой функции можно найти, решая уравнение $f'(x) = 0$ методом Ньютона (другое название – метод касательных). Пусть $x_0 \in [a,b]$ – нулевое (начальное) приближение к искомой точке x^* . Заменим дугу графика функции $F(x) = f'(x)$ в окрестности начальной точки касательной в точке $(x_0, f'(x_0))$.

$$F(x) \approx F(x_0) + F'(x_0)(x - x_0) \quad (22)$$

Выберем в качестве следующего приближения к x^* точку x_1 пересечения касательной с осью абсцисс.

Приравнивая к нулю правую часть в (22), получим первый элемент

$$x_1 = x_0 - \frac{F(x_0)}{F'(x_0)} \quad (23)$$

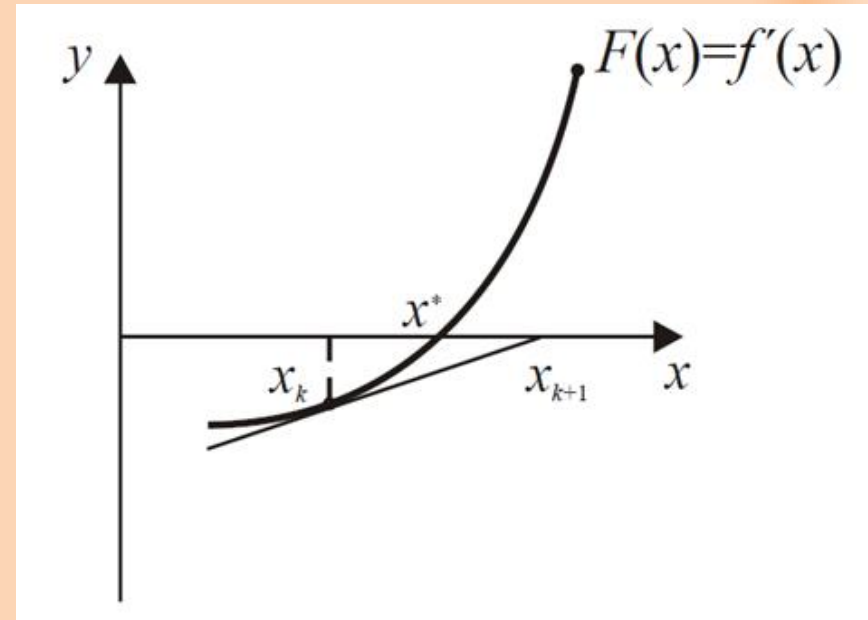
итерационной последовательности $\{x_k\}$, $k=1,2,\dots$

На $(k+1)$ -м шаге по найденной на предыдущем шаге точке x_k можно найти точку

$$x_{k+1} = x_k - \frac{f'(x_k)}{f''(x_k)} \quad (24)$$

Вычисления по формуле (24) производятся до тех пор, пока

не выполнится неравенство $|f'(x_k)| \leq \varepsilon$, после чего полагают $x^* \approx x_k$, $f^* \approx f(x_k)$.



РЕШЕНИЕ ЗАДАЧ ОДНОМЕРНОЙ ОПТИМИЗАЦИИ МЕТОДОМ КВАДРАТИЧНОЙ ИНТЕРПОЛЯЦИИ

В этом методе целевая функция сначала заменяется квадратичным полиномом. После этого положение минимума целевой функции определяется по положению минимума полинома, что значительно проще.

Пусть целевая функция известна в трех точках x_1, x_2, x_3 , то есть даны значения функции $f(x_1), f(x_2), f(x_3)$.

Заменяем целевую функцию в окрестности этих точек квадратичным полиномом ax^2+bx+c , т.е. коэффициенты этого полинома можно определить из системы трех линейных уравнений, которые надо разрешить относительно a, b, c :

$$\begin{cases} ax_1^2 + bx_1 + c = f(x_1) \\ ax_2^2 + bx_2 + c = f(x_2) \\ ax_3^2 + bx_3 + c = f(x_3) \end{cases}$$

При этом непосредственно определять численные значения коэффициентов a , b , c нет необходимости. В задаче нужно только знать положение минимума квадратичного полинома. Известно, что минимум квадратичного

полинома расположен в точке $x^* = -\frac{b}{2a}$

при $a > 0$

Решаем эту систему с помощью определителей (методом Крамера):

$$\Delta = \begin{vmatrix} x_1^2 & x_1 & 1 \\ x_2^2 & x_2 & 1 \\ x_3^2 & x_3 & 1 \end{vmatrix} = (x_2 - x_1)(x_3 - x_1)(x_3 - x_2)$$

$$\Delta_1 = \begin{vmatrix} f(x_1) & x_1 & 1 \\ f(x_2) & x_2 & 1 \\ f(x_3) & x_3 & 1 \end{vmatrix} = f(x_1)(x_3 - x_2) + f(x_2)(x_1 - x_3) + f(x_3)(x_2 - x_1)$$

$$\Delta_2 = (x_2^2 - x_3^2)f(x_1) + (x_3^2 - x_1^2)f(x_2) + (x_1^2 - x_2^2)f(x_3)$$

$$x^* = -\frac{b}{2a}; a = \frac{\Delta_1}{\Delta}; b = \frac{\Delta_2}{\Delta};$$

$$x^* = -\frac{1}{2} \frac{\Delta_2}{\Delta_1}; (1)$$

$$x^* = \frac{1}{2} \frac{(x_2^2 - x_3^2)f(x_1) + (x_3^2 - x_1^2)f(x_2) + (x_1^2 - x_2^2)f(x_3)}{f(x_1)(x_2 - x_3) + f(x_2)(x_3 - x_1) + f(x_3)(x_1 - x_2)}$$

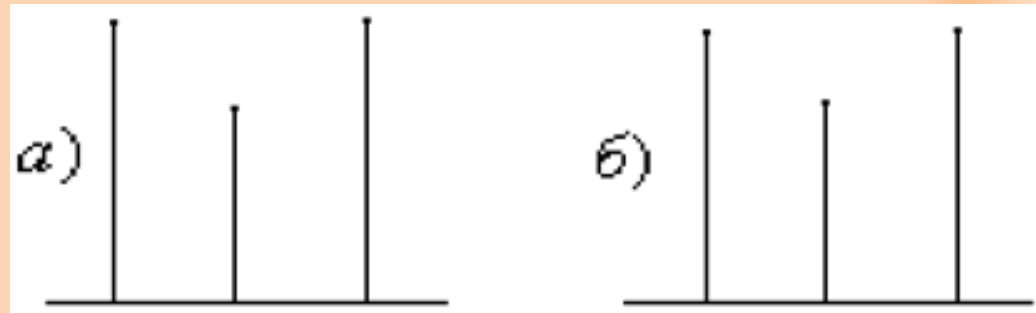
Рассмотрим аспекты практической реализации этого метода.

Пусть задана унимодальная функция $f(x)$, точка x_0 - начальная аппроксимация положения минимума и h - шаг, величина которого имеет тот же порядок, что и расстояние от точки x_0 до истинного минимума. Вычислительные процедуры в этом случае содержат следующие этапы:

1. Вычисляем $f(x_0)$ и $f(x_0 + h)$

2. а). Если $f(x_0) < f(x_0 + h)$, то в качестве третьей аппроксимирующей точки выбирается точка $(x_0 - h)$ и вычисляется $f(x_0 - h)$;

б). Если $f(x_0) > f(x_0 + h)$, то выбирается точка $(x_0 + 2h)$ и вычисляется $f(x_0 + 2h)$.



отбрасываем точку x_3 отбрасываем точку x_4

После проведения некоторого количества вычислений абсциссы x_1, x_2, x_3 точек, а также ординаты этих точек $f(x_1), f(x_2), f(x_3)$ будут достаточно близки друг к другу, поэтому относительная погрешность вычислений будет достаточно большой. Поэтому для вычислений на второй и последующих интерполяциях используется формула в преобразованном виде:

$$\begin{aligned}
 x^* &= \frac{1}{2} \frac{f(x_1)(x_2^2 - x_3^2) + f(x_2)(x_3^2 - x_1^2) + f(x_3)(x_1^2 - x_2^2)}{[f(x_1)(x_2 - x_3) + f(x_2)(x_3 - x_1) + f(x_3)(x_1 - x_2)]} = x^* \\
 &= \frac{1}{2[\dots]} [f(x_1)(x_2 - x_3)(x_2 + x_1 - x_1 + x_3) + f(x_2)(x_3 - x_1)(x_3 + x_2 - x_2 + x_1) + \\
 &+ f(x_3)(x_1 - x_2)(x_1 + x_2)] = \frac{1}{2[\dots]} \{ [f(x_1)(x_2 - x_3)(x_1 + x_2) + \\
 &+ f(x_2)(x_3 - x_1)(x_1 + x_2) + f(x_3)(x_1 - x_2)(x_1 + x_2)] + \\
 &+ [f(x_1)(x_2 - x_3)(x_3 - x_1) + f(x_2)(x_3 - x_1)(x_3 - x_2)] \} = \\
 &= \frac{x_1 + x_2}{2} + \frac{1}{2[\dots]} [(x_3 - x_1)(x_3 - x_2)(f(x_2) - f(x_1))];
 \end{aligned}$$

Численные методы безусловной минимизации функции нескольких переменных

Пусть теперь $f: \mathbb{R}^n \rightarrow \mathbb{R}$, то есть имеем дело с функцией n переменных и нужно по-прежнему найти точку минимума.

В методах прямого поиска минимума целевой функции (или методах нулевого порядка) используют информацию только о значениях этой функции. Из методов этого типа рассмотрим метод покоординатного спуска.

Метод покоординатного спуска

Здесь минимизация осуществляется циклами, по n шагов в каждом. Пусть $x^k = (x_1^k, x_2^k, \dots, x_n^k)^T$.

- Фиксируем значения x_2^k, \dots, x_n^k и проводим минимизацию по переменной x_1 . Получаем $(x_1^{k+1}, x_2^k, \dots, x_n^k)^T$, на этом первый шаг окончен.
- Фиксируем значения $x_1^{k+1}, x_3^k, \dots, x_n^k$ и проводим минимизацию по переменной x_2 . Получаем $(x_1^{k+1}, x_2^{k+1}, \dots, x_n^k)^T$.
- ...
- Фиксируем значения $x_1^{k+1}, x_2^{k+1}, \dots, x_{n-1}^{k+1}$ и проводим минимизацию по переменной x_n . Получаем $(x_1^{k+1}, x_2^{k+1}, \dots, x_n^{k+1})^T$.

Итак, после n шагов, соответствующих каждой из переменных завершаем первый цикл, получая $x^{k+1} = (x_1^{k+1}, x_2^{k+1}, \dots, x_n^{k+1})^T$.

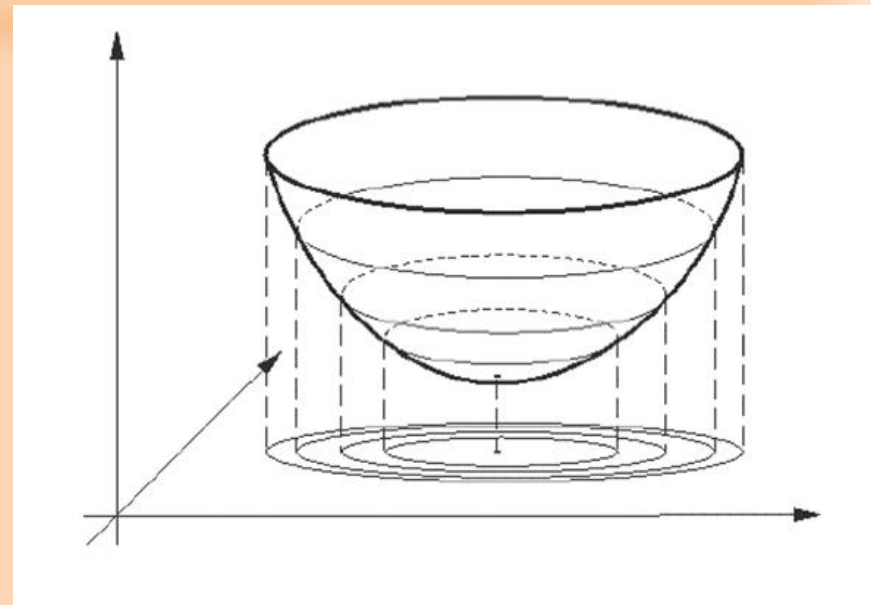
Процесс повторяется до тех пор, пока для некоторого заданного положительного ε не выполнится условие остановки

$$|f(x^{k+1}) - f(x^k)| \leq \varepsilon \text{ или } \|x^{k+1} - x^k\| \leq \varepsilon,$$

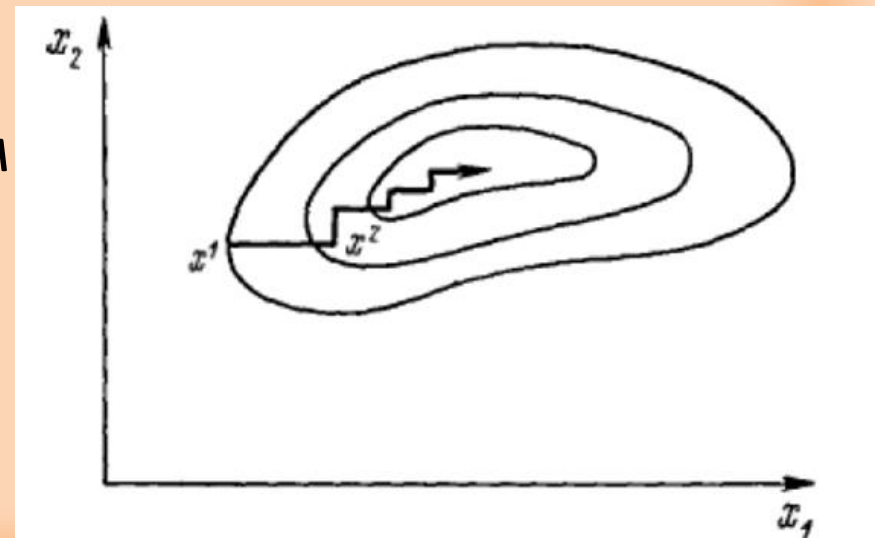
либо их комбинация.

На каждом шаге цикла приходится решать задачу одномерной минимизации, которую мы рассматривали ранее.

Иллюстрируют работу методов многомерной оптимизации обычно с помощью линий уровня – геометрического места точек пространства аргументов, для которых значения исследуемой функции одинаковы.



Для функций двух переменных линии уровня изображаются на плоскости, что позволяет наглядно продемонстрировать работу алгоритма. Так на нижнем рисунке видно, что в методе покоординатного спуска минимизация ведется поочередно по направлениям параллельным осям координат (по ортам).



Метод наискорейшего спуска

Методы 1-го порядка используют информацию о производной функции. Если ограниченная снизу целевая функция $f(x)$ является дифференцируемой на множестве R^n , то алгоритм поиска точки x^* ее минимума можно построить, используя информацию о градиенте этой функции. Такие методы называются градиентными.

Градиентные методы безусловной оптимизации используют только первые производные целевой функции и являются методами линейной аппроксимации на каждом шаге, т. е. целевая функция на каждом шаге заменяется касательной гиперплоскостью к ее графику в текущей точке.

Во всех этих методах предполагается, что $f(x)$, ∇f существуют и непрерывны. Один из таких методов – метод наискорейшего спуска.

Впервые такой метод рассмотрел и применил еще Коши в XVIII в.

Идея его проста: градиент целевой функции $f(x)$ в любой точке есть вектор в направлении наибольшего возрастания значения функции. Следовательно, антиградиент будет направлен в сторону наибольшего убывания функции и является направлением наискорейшего спуска. Антиградиент (и градиент) ортогонален поверхности уровня $f(x)$ в точке x .

Пусть в точке x требуется определить направление наискорейшего спуска, то есть направление наибольшего локального уменьшения $f(x)$. Разложим $f(x)$ в ряд Тейлора в окрестности точки x и отбросим члены второго порядка по Δx и выше :

$$f(x) = f(x^{(k)}) + (\nabla f(x^{(k)}), \Delta x) + \dots$$

Локальное уменьшение $f(x)$ определяется вторым слагаемым, т. е. наибольшее уменьшение $f(x)$ будет тогда, когда $(\nabla f(x^{(k)}), \Delta x)$ будет иметь наибольшую отрицательную величину. Этого можно добиться выбором $S^{(k)} = -\nabla f(x^{(k)})$

Этот случай соответствует наискорейшему локальному спуску

$$x^{(k+1)} = x^{(k)} - \alpha_k \nabla f(x^{(k)})$$

Алгоритм

Шаг 0. Задать $\varepsilon > 0$, $x^{(1)}$ начальную точку. Положить $k = 1$.

Шаг 1. Если $\|\nabla f(x^{(k)})\| < \varepsilon$ то конец и оптимум $x^* = x^{(k)}$,

Иначе вычислить $S^{(k)} = \frac{\nabla f(x^{(k)})}{\|\nabla f(x^{(k)})\|}$.

Шаг 2. Решить задачу одномерной оптимизации

$$\min_{\lambda} f(x^{(k)} + \lambda S^{(k)})$$

Положить $x^{(k+1)} = x^{(k)} + \lambda_k S^{(k)}$

Шаг 3. Положить $k = k + 1$ и перейти на шаг 1.

Пример. Найти максимальное значение функции

$$f = 2x_1 + 4x_2 - x_1^2 - 2x_2^2$$

с точностью $\varepsilon \leq 0,05$.

Решение. Найдем градиент функции:

$$\frac{\partial f}{\partial x_1} = 2 - 2x_1; \quad \frac{\partial f}{\partial x_2} = 4 - 4x_2;$$

$$\nabla f(X) = ((2 - 2x_1); (4 - 4x_2)).$$

Возьмем в качестве первого приближения $X^{(0)} = (0; 0)$, т. е.

$$x_1^{(0)} = x_2^{(0)} = 0.$$

Тогда значение функции $f(X^{(0)}) = 0$, а вектор-строка градиента функции равен $\nabla f(X^{(0)}) = (2; 4)$. Выберем шаг итерации $\lambda = 0,25$ и рассчитаем параметры следующей точки:

$$x_1^{(1)} = x_1^{(0)} + \lambda \cdot \nabla f(x_1^{(0)}) = 0 + 0,25 \cdot 2 = 0,5$$

$$x_2^{(1)} = x_2^{(0)} + \lambda \cdot \nabla f(x_2^{(0)}) = 0 + 0,25 \cdot 4 = 1,0$$

Вычислим значение функции цели в новой точке и определим степень приближения:

$$f(X^{(1)}) = 2 \cdot 0,5 + 4 \cdot 1 - 0,25 - 2 \cdot 1 = 2,75$$

$$|f(X^{(1)}) - f(X^{(0)})| = |2,75 - 0| = 2,75 > \varepsilon$$

Так как заданная точность не достигнута, продолжим итерационный процесс. Градиент функции в новой точке будет определяться вектором-строкой $\nabla f(X^{(1)}) = (1; 0)$. Рассчитаем параметры следующей точки:

$$x_1^{(2)} = x_1^{(1)} + \lambda \cdot \nabla f(x_1^{(1)}) = 0,5 + 0,25 \cdot 1 = 0,75$$

$$x_2^{(2)} = x_2^{(1)} + \lambda \cdot \nabla f(x_2^{(1)}) = 1,0 + 0,25 \cdot 0 = 1,0$$

Значение функции цели в исследуемой точке и степень приближения равны:

$$f(X^{(2)}) = 2 \cdot 0,75 + 4 \cdot 1 - 0,5625 - 2 \cdot 1 = 2,9375$$

$$|f(X^{(2)}) - f(X^{(1)})| = |2,9375 - 2,75| = 0,1875 > \varepsilon$$

Продолжим вычисления. В точке $X^{(2)} = (0,75; 1,0)$ градиент функции будет иметь следующий вектор-строку: $\nabla f(X^{(2)}) = (0,5; 0)$.

Рассчитываем параметры третьей точки итерации:

$$x_1^{(3)} = x_1^{(2)} + \lambda \cdot \nabla f(x_1^{(2)}) = 0,75 + 0,25 \cdot 0,5 = 0,875$$

$$x_2^{(3)} = x_2^{(2)} + \lambda \cdot \nabla f(x_2^{(2)}) = 1,0 + 0,25 \cdot 0 = 1,0$$

Значение функции цели в этой точке и степень приближения равны:

$$f(X^{(3)}) = 2 \cdot 0,875 + 4 \cdot 1 - 0,875^2 - 2 \cdot 1 = 2,9844$$

$$|f(X^{(3)}) - f(X^{(2)})| = |2,9844 - 2,9375| = 0,0469 < \varepsilon$$

Тогда в пределах заданной точности получаем ответ:

$$X^* = (0,875; 1,0)$$

$$f_{max} = 2,9844$$

Спасибо за внимание!