Лекция № 3

Прямые методы поиска безусловного экстремума ФМП

Решение задачи о поиске безусловного экстремума $f(X)$ с помощью необходимых и достаточных условий, приводит к необходимости решать систему n нелинейных уравнений с n неизвестными, с последующей проверкой знакоопределенности $H(X^*)$. Как правило, для достаточно сложных функций такая процедура решения задачи достаточно трудоемка и подразумевает численное решение нескольких задач. Поэтому возникает необходимость использовать так называемые прямые или численные методы безусловной оптимизации, которые позволяют найти стационарные точки функции, не используя аппарат необходимых и достаточных условий экстремума.

Замечание. Во всех прямых методах рассматривается задача о поиске минимума функции $f(X)$, задача поиска максимума функции $f(X)$ может быть решена заменой целевой функции на $-f(X)$.

Все прямые методы ищут точки как предел последовательности точек $\{X^k\}$ при $k \to \infty$.

Определение 1. Последовательность $\{X^k\}$ называется минимизирующей, если $\lim\limits_{k\to\infty} f(X^k) = f(X^*)$, т.е. последовательность сходится к нижней грани функции $f(X)$.

Не всякая минимизирующая последовательность дает возможность найти искомую точку минимума.

Определение 2. Последовательность $\{X^k\}$ называется сходящейся к точке минимума, если $\lim\limits_{k\to\infty} X^k = X^*$.

Все прямые методы имеют один и тот же алгоритм: $X^{k+1} = X^k + t_k d^k$, где

- X^k - текущая точка последовательности, причем X^0 – задается из физического содержания задачи или произвольно;
- X^{k+1} - последующая точка последовательности;
- d^k - приемлемое направление перехода из точки в точку – направление спуска. Приемлемым при решении задачи поиска минимума функции будет только то направление, для которого $f(X^{k+1}) < f(X^k)$, что обеспечивается выполнением условия $(\nabla f(X^k), d^k) < 0$;
- t_k - шаг (число >0).

Процесс преобразования точки X^k в точку X^{k+1} называют *<u>итерацией</u>*, а сам процесс построения последовательности точек по определенному правилу, оканчивающийся согласно выполнению критерия окончания – *<u>итерационным</u>*.

Все прямые методы отличаются друг от друга способом задания d^k и выбором t_k.

В зависимости от наивысшего порядка частных производных функции $f(X)$, используемых для формирования направления d^k, численные методы решения задачи безусловной минимизации принято делить на три группы:

1. Методы первого порядка, использующие информацию о 1-х производных функции $f(X)$.
2. Методы второго порядка, использующие для своей реализации информацию о 1-х и 2-х производных функции $f(X)$.
3. Методы нулевого порядка, использующие информацию только о значении функции $f(X)$.

Методы первого порядка

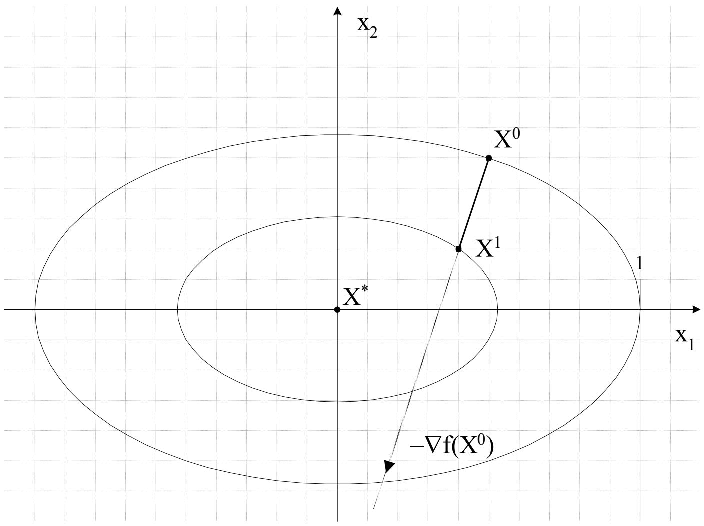
(1) Метод градиентного спуска

Алгоритм метода: $\boxed{X^{k+1} = X^k - t_k \nabla f(X^k)}$ (1.1)

здесь:

- $d^k = -\nabla f(X^k)$ - направление антиградиента функции;
- t_k - шаг выбирается из условия убывания функции в точках последовательности:
  $f(X^{k+1}) < f(X^k)$. (1.2)

Геометрическая интерпретация метода:



Основной критерий окончания метода:
Построение последовательности заканчивается в точке, для которой:

$\left\|\nabla f(X^k)\right\| < \varepsilon$, где ε - заданное малое положительное число. (1.3)

Дополнительные критерии окончания метода:

- при выполнении предельного числа итераций: $k = M$ (1.4)
- при однократном или двукратном одновременном выполнении двух условий:
  $\left\|X^{k+1} - X^k\right\| < \tilde{\varepsilon}$, $\left|f(X^{k+1}) - f(X^k)\right| < \tilde{\varepsilon}$, где $\tilde{\varepsilon}$ - малое положительное число. (1.5)

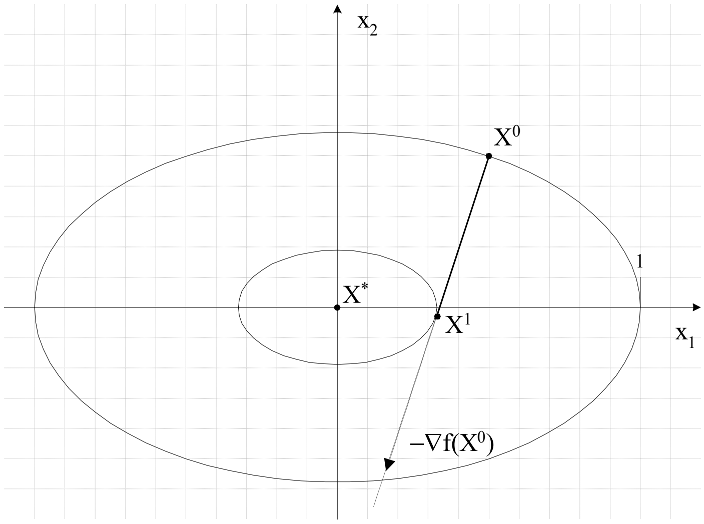
(2) Метод градиентного наискорейшего спуска

Алгоритм метода: $\boxed{X^{k+1} = X^k - t_k \nabla f(X^k)}$ (2.1)

здесь

- $d^k = -\nabla f(X^k)$ - направление антиградиента функции
- t_k - шаг вычисляется из условия наибольшего убывания функции в точках последовательности, построенной по закону (2.1): $t_k = \arg\min[f(X^{k+1})]$ (2.2)

Геометрическая интерпретация метода



Как видно из чертежа, точка X^1 принимает на направлении спуска $d^0 = -\nabla f(X^0)$ предельное положение, которое характеризуется тем, что линия уровня, проходящая через точку X^1 касается направления спуска, а, следовательно, в точках минимизирующей последовательности, построенной по методу градиентного наискорейшего спуска выполняется условие:

$$\nabla f(X^k) \perp \nabla f(X^{k+1}) \quad \Rightarrow \quad (\nabla f(X^k), \nabla f(X^{k+1})) = 0 \tag{2.3}$$

Критерии окончания метода такие же, как и в методе градиентного спуска.

Вычисление шага t_k может быть выполнено тремя способами:

Способ А заключается в вычислении функции $f(X^{k+1})$, эта функция оказывается функцией одной переменной t_k, т.е. $f(X^{k+1}) = \varphi(t_k)$, и последующем использовании необходимых и достаточных условий безусловного экстремума этой функции: $\frac{d\varphi(t_k)}{dt_k} = 0;\quad \frac{d^2\varphi(t_k)}{d^2 t_k} > 0$. Этот способ может быть использован в случаях, когда функция $f(X)$ достаточно проста.

Способ В заключается в использовании условия (2.3) для вычисления шага. Способ также может быть использован в случаях, когда функция $f(X)$ достаточно проста.

Способ С предполагает численное решение задачи $f(X^{k+1}) = \varphi(t_k) \to \min$ методом дихотомии (см. [2], глава II, §5, п. 5.1.4.) на отрезке $[a, b]$ с заданной точностью ε_D.

(3) Метод покоординатного спуска

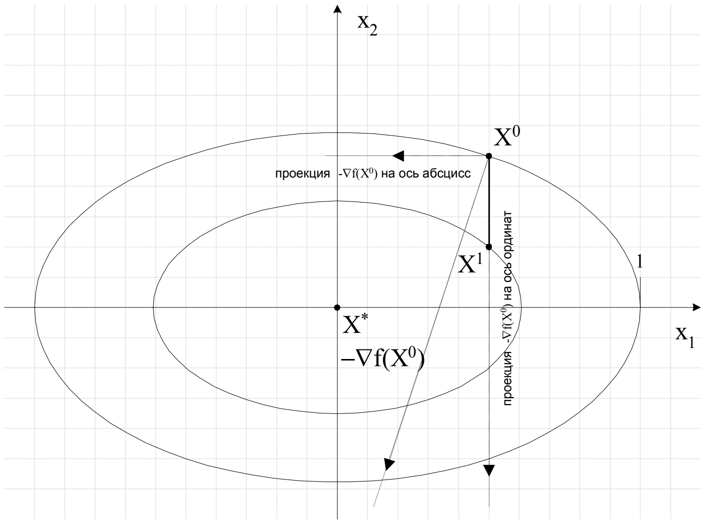
Алгоритм метода: $\boxed{X^{k+1} = X^k - t_k \left[\nabla f(X^k)\right]_{\text{пр на } x_i}}$ (3.1)

здесь:

- $d^k = -\left[\nabla f(X^k)\right]_{\text{пр на } x_i}$ - проекция на ось x_i антиградиента функции
- t_k - шаг выбирается из условия убывания функции в точках последовательности:

$$f(X^{k+1}) < f(X^k) \tag{3.2}$$

Геометрическая интерпретация метода



Критерии окончания метода такие же, как и в методе градиентного спуска.

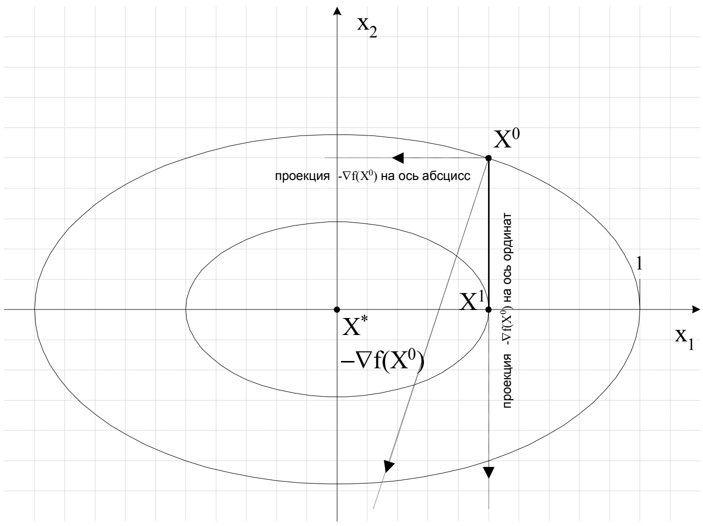
(4) Метод Гаусса-Зейделя (наискорейшего покоординатного спуска)

Алгоритм метода: $\boxed{X^{k+1} = X^k - t_k \left[\nabla f(X^k)\right]_{\text{пр на } x_i}}$ (4.1)

здесь:

- $d^k = -\left[\nabla f(X^k)\right]_{\text{пр на } x_i}$ - проекция на ось x_i антиградиента функции
- t_k - шаг вычисляется из условия наибольшего убывания функции в точках последовательности, построенной по закону (4.1): $t_k = \arg\min[f(X^{k+1})]$ (4.2)

Геометрическая интерпретация метода



Критерии окончания метода такие же, как и в методе градиентного спуска.
Вычисление шага t_k может быть выполнено способами А и С, описанными в методе градиентного наискорейшего спуска.

Сходимость метода градиентного спуска (градиентного наискорейшего спуска, покоординатного спуска и метода Гаусса-Зейделя) регламентируется теоремой.

Теорема.

Если функция $f(X)$ ограничена снизу, а ее градиент удовлетворяет условию Липшица:

$$\|\nabla f(X)-\nabla f(Y)\| \le L\|X-Y\|, \quad \forall X, Y \in R^n, \quad L>0,$$

то метод градиентного спуска гарантирует $\|\nabla f(X^k)\| \to 0$ при $k \to \infty$.

Поскольку критерий окончания гарантирует выполнение в точке $X^* = \lim\limits_{k\to\infty} X^k$ только необходимых условий минимума, но не достаточных, можно утверждать, что метод обеспечивает сходимость к стационарной точке функции, в которой необходимо проверить достаточные условия минимума.

Сходимость именно к точке минимума гарантируется только для выпуклых функций.

(5) Метод сопряженных градиентов (Флетчера-Ривса)

Алгоритм метода: $\boxed{X^{k+1} = X^k + t_k d^k}$ (5.1)

здесь:

- $d^0 = -\nabla f(X^0)$ (5.2)

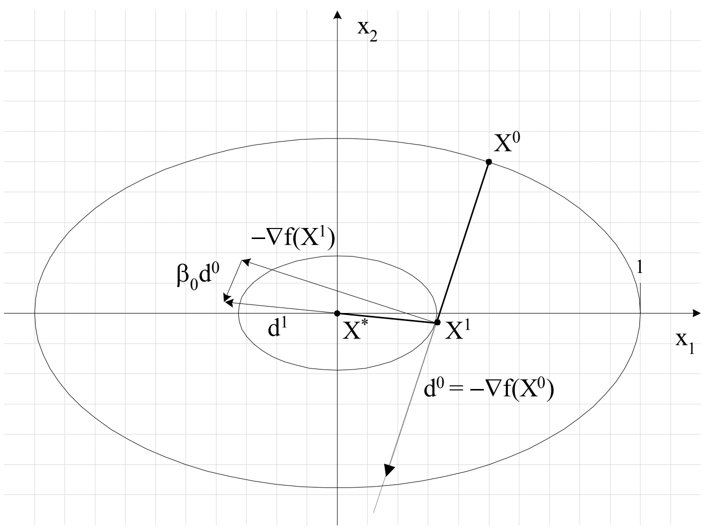
  $d^k = -\nabla f(X^k) + \beta_{k-1} d^{k-1}$ (5.3)

  $$\beta_{k-1} = \frac{\|\nabla f(X^k)\|^2}{\|\nabla f(X^{k-1})\|^2} \qquad (5.4)$$

- t_k - шаг вычисляется из условия наибольшего убывания функции в точках последовательности, построенной по закону (5.1): $t_k = \arg\min[f(X^{k+1})]$ (5.5)

Из формул (5.2) и (5.5) следует, что первая итерация метода сопряженных градиентов совпадает с первой итерацией метода наискорейшего спуска.

Геометрическая интерпретация метода



Критерии окончания метода такие же, как и в методе градиентного спуска.

Вычисление шага t_k может быть выполнено способами A и C, описанными в методе градиентного наискорейшего спуска.

Для квадратичных функций метод сопряженных градиентов называется <u>методом Флетчера-Ривса</u>.

Доказано, что для функций, имеющих минимум, метод Флетчера-Ривса <u>сходится</u> за конечное число шагов, не превышающее число переменных функции $f(X)$.

Для не квадратичных функций, используется модификация метода сопряженных градиентов, называемая <u>методом Полака-Рибьера</u> (см.[2], гл.II, §6. п.6.5).