Семинар 6. Методы оптимизации. МФТИ. Осень 2016. Тренин С.А.

Градиентные методы.

Пусть $f: \mathbb{R}^n \to \mathbb{R}$ - непрерывно дифференцируемая функция. Направление Δx в точке x называется *направлением убывания (роста)*, если $f(x + \varepsilon\Delta x) - f(x) < 0$ ($f(x + \varepsilon\Delta x) - f(x) > 0$) при всех достаточно малых ε.

1. Докажите, что если $\nabla f(x)^T \Delta x < 0$, то Δx является направлением убывания, а если $\nabla f(x)^T \Delta x > 0$, то Δx является направлением роста. Приведите примеры, когда $\nabla f(x)^T \Delta x = 0$, а Δx является направлением убывания, направлением роста, или ни тем, ни другим.

Методом спуска называется любой итеративный метод решения задачи безусловной минимизации $f(x) \to min$, при котором x_{k+1} выбирается по правилу $x_{k+1} = x_k + t_k \Delta x_k$, где Δx_k есть направление убывания в точке x_k, а $t_k > 0$. Величина t_k (*размер шага*) может выбираться по различным правилам:

a) *Правило одномерной минимизации*: $t_k = \operatorname{argmin}_{s \geq 0} f(x_k + s\Delta x_k)$.

b) Величина есть некоторая постоянная, либо заранее заданная последовательность. Например: $t_k = \frac{1}{\sqrt{k+1}}$.

c) *Правило обратного одномерного поиска* (Армихо). Заданы два числа $0 < \alpha < 0.5$, $0 < \beta < 1$. Размер шага вычисляется по следующей процедуре:

$$
\begin{aligned}
& s = 1; \\
& while\ (f(x + s\Delta x) > f(x) + \alpha s \nabla f(x)^T \Delta x)\ \{ \\
& \quad s = \beta s; \\
& \} \\
& t_k = s
\end{aligned}
$$

2. Покажите, что процедура выбора шага по правилу Армихо завершается.
3. Покажите, что, если условие $f(x + s\Delta x) \leq f(x) + \alpha s \nabla f(x)^T \Delta x$ истинно для всех $s \in (0, s_0]$ 0 до s_0, то процедура Армихо завершается при значении s принадлежащем интервалу $(\beta s_0, s_0]$ или при s = 1.

Градиентным методом называется любой метод спуска, в котором Δx_k является антиградиентом, т.е. $x_{k+1} = x_k - t_k \nabla f(x_k)$. Если при этом t_k выбирается по правилу одномерной минимизации, то метод называется *методом скорейшего спуска*.

4. Опишите метод скорейшего спуска для квадратичной функции $f(x) = \frac{1}{2}x^T Ax - b^T x$. А именно, напишите формулу, выражающую t_k через x_k.
5. Пусть функция $f: \mathbb{R}^2 \to \mathbb{R}$ определяется формулой
$$f(x, y) = 2x^2 + xy + 3y^2.$$
Сделайте 1 шаг градиентного метода решения задачи $f(x) \to min$ из начальной точки $(1, -1)$ с выбором шага по правилу Армихо с параметрами $\alpha = 0.25$, $\beta = 0.5$.

Метод Ньютона

Методом Ньютона решения уравнения $f(x) = 0$ на прямой называют итеративный метод, в котором точка x_{k+1} выбирается, как точка пересечения касательной к графику в точке x_k и оси абсцисс.

6. Покажите, что указанная точка определяется условием
$$x_{k+1} = x_k - (\nabla f(x_k))^{-1} f(x_k).$$
7. Приведите пример, когда метод Ньютона на прямой зацикливается.
8. Приведите пример, когда метод Ньютона расходится.

Метод Ньютона можно обобщить на многомерные функции и применить к задачам оптимизации для поиска стационарной точки, в которой $\nabla f(x) = 0$. Таким образом, итерация метода Ньютона задаётся формулой

$$x_{k+1} = x_k - (\nabla^2 f(x_k))^{-1} \nabla f(x_k).$$

9. Покажите, что такая же формула получится при определении как точки минимума квадратичного приближения к функции в точке.
10. Пусть функция $f: \mathbb{R}^2 \to \mathbb{R}$ определяется формулой $f(x, y) = x^2 + e^{y^2}$. Сделайте 2 шага метода Ньютона решения задачи $f(x) \to min$ из начальной точки $(1, 1)$.
11. Всегда ли направление метода Ньютона является направлением спуска?