Chapter 1

Численная многомерная безусловная оптимизация

Рассматриваем функции без ограничений.

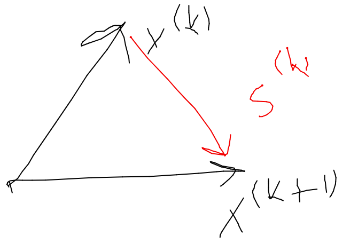
Она основывается на построении последовательностей точек $\{x^{(k)}\}$ в $\mathbb{R}^n$ которые будут сходится к x^*.

Нужно задать правило перехода к точке $x^{(k+1)}$ и построить последовательность, которая будет обладать следующими свойствами:

$$1)\ \lim_{k\to\infty} x^{(k)} = x^*$$

$$2)\ f(x^{(0)}) > f(x^{(1)}) > ... > f(x^{(k)}) > f(x^{(k+1)})$$

Если



то можно считать

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

Последовательность, которая будет обладать такими свойствами, будет называться релаксиционной или итерационной.

Вектор $S^{(k)}$ будем называть направлением спуска функции $f(x)$, если существует $\alpha^0 > 0$ такое что для любого $\alpha \in (0, \alpha^{(0)})$ верно

$$f(x^{(k)} + \alpha S^{(k)}) < f(x^{(k)})$$

Основная задача методов спуска: подобрать правильно параметры $S^{(k)}$ и $\alpha^{(k)}$ обеспечив при этом убывание целевой функции в точках $x^{(k)}$.

Лемма 1: (Направление спуска безусловной оптимизации) Пусть $x^{(k)}$ не является критической точкой. Тогда в этой точке может существовать бесконечно много направлений спуска. Для того, чтобы вектор S был направлением спуска достаточно, чтобы

$$(\nabla f(x^{(k)}), S) < 0$$

Доказательство: Определим току вида

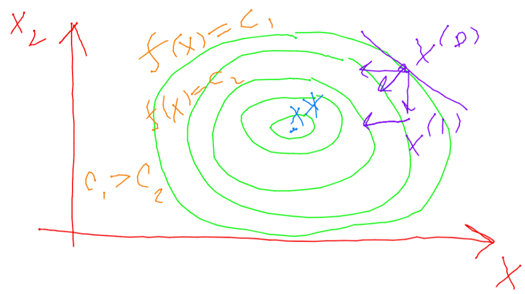
$$x = x^{(k)} + \alpha^{(k)} S^{(k)}$$

таким образом задали альфа-окрестность точки $x^{(k)}$.

Используя формулу Тейлора

$$f(x) - f(x^{(k)}) = (\nabla f(x^{(k)}), (x - x^{(k)})) + o(||x - x^{(k)}||^2) =$$
$$= \alpha(\nabla f(x^{(k)}), S^{(k)}) + o(\alpha^2)$$

$o(\alpha^2) \rightarrow 0$ и скалярное произведение с коэффицентом меньше нуля. Тогда и правая часть меньше нуля. Т.е. S - направление спуска. ч.т.д.



Замечу, что чем меньше скалярное произведение, тем ближе к минимуму будет шаг.

Нетрудно доказать, что

$$-||\nabla f(x^{(k)})|| \cdot ||S^{(k)}|| \leq (\nabla f(x^{(k)}, S^{(k)})) \leq ||\nabla f(x^{(k)})|| \cdot ||S^{(k)}||$$

надо подобрать некую матрицу $\mathcal{F}_k$ которая будет удовлетворять неравенству

$$\forall y \in E^n : r||y||^2 \leq (\mathcal{F}_k y, y) \leq R||y||^2 \qquad (1)$$

где $0 < r < R$, такую что будет приближать равенство

$$S^{(k)} = -\nabla f(x^{(k)})$$

т.е. будем строить по формуле

$$S^{(k)} = -\mathcal{F}_k \cdot f(x^{(k)}) \qquad (2)$$

$\mathcal{F}_k$ - матрица спуска.

Лемма 2: Пусть $x^{(k)}$ - не критическая точка, и система матриц $\mathcal{F}_k$ удовлетворяет неравенству (1) тогда вектор $S^{(k)}$ полученный по (2) является направлением спуска.

Доказательство: аналогично по формуле Тейлора Ч.т.д.

Отсюда, из предыдущих лемм, будем искать $x^{(k+1)}$ как

$$x^{(k+1)} = x^{(k)} - \alpha^{(k)} \mathcal{F}_k \nabla f(x^{(k)}) \tag{3}$$

При выборе метода - нужно подобрать $\alpha^{(k)}$ и $\mathcal{F}_k$.

Важно при использовании (3) указать способ выбора матрицы $\mathcal{F}_k$, который дает новый метод, а также нужно указать алгоритм выбора шага $\alpha^{(k)}$, который предусматривает новый алгоритм метода.

1.1 Способы выбора матрицы спуска

Если $\mathcal{F}_k = E$ то по (1) будем иметь $r = R = 1$ - методы такого вида будем назвывать градиентными.

Второй вариант: Пусть функция дважды непрерывно-дифференцируема, тогда матрица Гессе будет удовлетворять (1).

$$m||y||^2 \leq (H(x^{(k)})y, y) \leq M||y||^2$$

$0 < m < M < 1$

В таком случае если взять $\mathcal{F}_k = (H(x^{(k)}))^{(-1)}$, то тогда $r = \frac{1}{M}$ и $R = \frac{1}{m}$.

Методы с таким выбором $\mathcal{F}_k$ назовем методами Ньютона или обобщенными методами Ньютона.

Третий вариант: В том случае когда в начале используем градиентный метод, а далее - Ньютоновские, то такие методы называем методами Дэвида-Фитчера-Гауэлла.