

Алгоритм первый для выбора α . Если задана точка $x^{(0)}$ и каким-то образом рассчитаны $x^{(1)}, x^{(2)}, \dots, x^{(k)}$.

Находим

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

Можно выбрать $\alpha \equiv 1$ или например $\alpha = \text{const} > 0$.

Алгоритм второй для выбора α . Берем $\delta \in (0, 1)$ и задана точка начального приближения $x^{(0)}$ и положим $\alpha = 1$. Затем рассчитываем

$$x = x^{(0)} + \alpha \mathcal{F}_0(-\nabla f(x^{(0)}))$$

и тогда ищем

$$f(x) - f(x^{(0)}) < -\alpha \delta r \cdot \|\nabla f(x^{(0)})\|^2 \quad (*)$$

Если выполняется $(*)$ то полагаем $x^{(1)} = x$, а если нет, то берем $\alpha := \frac{\alpha}{2}$ и снова вычисляем x .

Теорема: Пусть $x^{(0)}$ не критическая точка и для матрицы \mathcal{F}_k верно неравенство (1) и для $H(x^{(0)})$ верно неравенство выведенное ранее из (1). Тогда после некоторого числа дроблений α алгоритм будет корректен.

Доказательство из формулы Тейлора верно

$$\begin{aligned} f(x) - f(x^{(0)}) &= (-\alpha \mathcal{F}_0 \nabla f(x^{(0)}), \nabla f(x^{(0)})) + \\ &+ \frac{1}{2} (H(x^{(0)}), (-\alpha \mathcal{F}_0 \nabla f(x^{(0)}))^2) \leq \\ &\leq -\alpha r \|\nabla f(x^{(0)})\| + \frac{1}{2} \alpha^2 M R^2 \|\nabla f(x^{(0)})\| \leq \\ &\leq -\alpha r \left(1 - \frac{\alpha M R^2}{2r}\right) \|\nabla f(x^{(0)})\| \end{aligned}$$

во второй строке в скалярном произведении используем покомпонентное произведение Адамара.

Должно выполняться

$$\begin{aligned} 1 - \frac{\alpha M R^2}{2r} &> \delta \\ 1 - \delta &> \frac{\alpha M R^2}{2r} \\ \frac{(1 - \delta) 2r}{M R^2} &> \alpha \end{aligned}$$

если последнее неравенство будет выполняться, то алгоритм сходится. Ч.т.д.

Если не указано значение δ берем $\delta = \frac{1}{2}$

Третий алгоритм для нахождения α . Берем $x^{(0)}$ и рассчитываем по некоторому правилу $x^{(1)}, x^{(2)}, \dots, x^{(k)}$ определяем точку вида

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

Надо ввести функцию $g(\alpha) = f(x) = f(x^{(k)} + \alpha^{(k)} \mathcal{F}_k(-\nabla f(x^{(k)})))$ и это будет функция от одного переменного α и ищем минимум функции $g(\alpha)$ на

интервале $(0, 1)$ с помощью уже известных методов одномерной оптимизации. Находим этот минимум $\alpha^{(k)}$.

И в конце проверяем по условиям, является ли $x^{(k)}$ минимумом или нет. Иначе - продолжаем.

Теорема: Пусть $f(x)$ ограничена снизу, дифференцируема и верно неравенство Липшица

$$\forall x', x'' \in \mathbb{R}^n \forall L > 0 : \\ \|\nabla f(x') - \nabla f(x'')\| \leq L \|x' - x''\|$$

Тогда для любой начальной точки $x^{(0)}$ можно $x^{(k+1)}$ представить как

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

и всегда можно найти такое α , что

$$\lim_{k \rightarrow \infty} \|\nabla f(x^{(k)})\| = 0$$

Таким образом строится релаксационная последовательность. Теорема гарантирует сходимость последовательности $\{f(x^{(k)})\}$ либо к некоторому значению функции которая является минимумом, либо к инфимуму функции $f(x)$ (если не будет стационарной точки на отрезке) или она может расходиться.

Доказательства не будет.

Тогда, если выполняется условие Липшица, и существует некоторая непустая область $L(C)$, такая что $f(x) \leq C$ где $\delta \leq L(C)$, это множество не является пустым и является компактным, тогда

$$\forall x^{(0)} \in L(C)$$

наша последовательность будет являться релаксационной.

Лучше всего использовать последний алгоритм. Итерационная процедура спуска с постоянным шагом неудобна, так как при реализации требует знания шага, который выбирается подбором, при этом неудачный выбор шага может привести к нарушению убывания целевой функции, а если взять слишком маленькое α , то алгоритм может сходиться очень долго. Поэтому лучше него будет второй алгоритм - условием завершения алгоритма будет

$$\|x^{(k+1)} - x^{(k)}\| < \varepsilon_1 \\ |f(x^{(k+1)}) - f(x^{(k)})| < \varepsilon_2$$

- продолжаем пока не будут выполнены оба эти условия - здесь $\varepsilon_1 > 0, \varepsilon_2 > 0$.

Если этого оказывается недостаточно для нахождения α , то в качестве завершения будем использовать только условие

$$\|\nabla f(x^{(k)})\| < \varepsilon_3$$

где $\varepsilon_3 > 0$. Последняя рассчитанная точка $x^{(k)}$ будем считать точкой минимума функции.