Теоретические сведения. Численные методы второго порядка ("высокая" траектория)

Численные методы второго порядка

МЕТОД НЬЮТОНА

ПОСТАНОВКА ЗАДАЧИ

Пусть дана функция $f(x)$, ограниченная снизу на множестве R^n и имеющая непрерывные частные производные во всех его точках.

Требуется найти локальный минимум функции $f(x)$ на множестве допустимых решений $X = R^n$, т. е. найти такую точку $x^* \in R^n$, что

$$f(x^*) = \min_{x \in R^n} f(x), \quad f(x) \in C^2.$$

СТРАТЕГИЯ ПОИСКА

Стратегия метода Ньютона [I. Newton] состоит в построении последовательности точек $\{x^k\}$, $k = 0, 1, \ldots$, таких, что $f(x^{k+1}) < f(x^k)$, $k = 0, 1, \ldots$ Точки последовательности вычисляются по правилу

$$x^{k+1} = x^k + d^k, \; k = 0, 1, \ldots, \tag{7.1}$$

где x^0 — задается пользователем, а направление спуска d^k определяется для каждого значения k по формуле

$$d^k = -H^{-1}(x^k)\nabla f(x^k). \tag{7.2}$$

Выбор d^k по формуле (7.2) гарантирует выполнение требования $f(x^{k+1}) < f(x^k)$ при условии, что $H(x^k) > 0$. Формула (7.2) получена из следующих соображений.

1. Функция $f(x)$ аппроксимируется в каждой точке последовательности $\{x^k\}$ квадратичной функцией

$$F_k = f(x^k) + (\nabla f(x^k),\ d^k) + \frac{1}{2}(d^k,\ H(x^k)d^k).$$

2. Направление d^k определяется из необходимого условия экстремума первого порядка:

$$\frac{dF_k}{dd^k} = 0.$$

Таким образом, при выполнении требования $H(x^k) > 0$ последовательность является последовательностью точек минимумов квадратичных функций F_k, $k = 0, 1, \ldots$ (рис. 7.1). Чтобы обеспечить выполнение требования $f(x^{k+1}) < f(x^k)$, $k = 0, 1, \ldots$, даже в тех случаях, когда для каких-либо значений матрица Гессе $H(x^k)$ не окажется положительно определенной, рекомендуется для соответствующих значений k вычислить точку x^{k+1} по методу градиентного спуска $x^{k+1} = x^k - t_k\nabla f(x^k)$ с выбором величины шага t_k из условия $f(x^k - t_k\nabla f(x^k)) < f(x^k)$.

Построение последовательности $\{x^k\}$ заканчивается в точке x^k, для которой $\|\nabla f(x^k)\| < \varepsilon_1$, где ε_1 — заданное малое положительное число, или при $k \ge M$ (M — предельное число итераций), или при двукратном одновременном выполнении двух неравенств $\|x^{k+1} - x^k\| < \varepsilon_2$, $|f(x^{k+1}) - f(x^k)| < \varepsilon_2$, где ε_2 — малое положительное число. Вопрос о том, может ли точка x^k рассматриваться как найденное приближение искомой точки минимума, решается путем проведения дополнительного исследования, которое описано ниже.

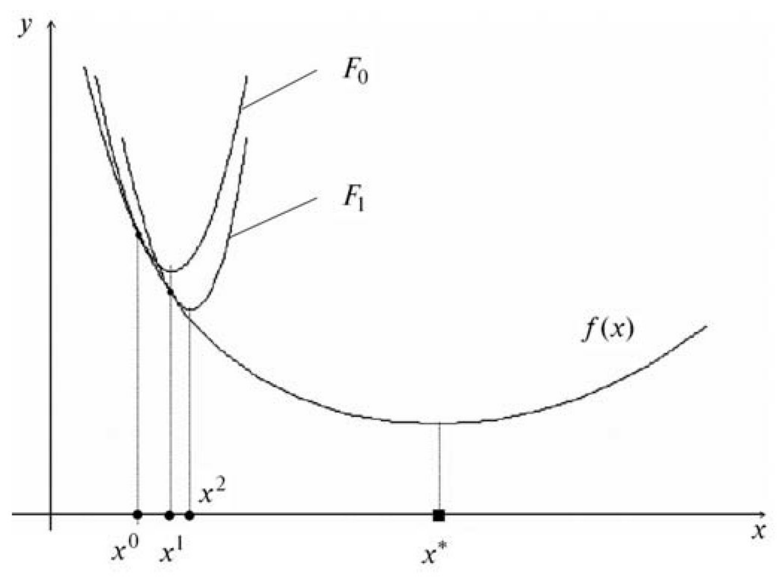

Рис. 7.1

АЛГОРИТМ

Шаг 1. Задать x^0, $\varepsilon_1 > 0$, $\varepsilon_2 > 0$, M — предельное число итераций. Найти градиент $\nabla f(x)$ и матрицу Гессе $H(x)$.

Шаг 2. Положить $k = 0$.

Шаг 3. Вычислить $\nabla f(x^k)$.

Шаг 4. Проверить выполнение критерия окончания $\|\nabla f(x^k)\| \le \varepsilon_1$:

а) если неравенство выполнено, то расчет окончен и $x^* = x^k$;

б) в противном случае перейти к шагу 5.

Шаг 5. Проверить выполнение неравенства $k \ge M$:

а) если неравенство выполнено, расчет окончен и $x^* = x^k$;

б) если нет, перейти к шагу 6.

Шаг 6. Вычислить матрицу $H(x^k)$.

Шаг 7. Вычислить матрицу $H^{-1}(x^k)$.

Шаг 8. Проверить выполнение условия $H^{-1}(x^k) > 0$:

а) если $H^{-1}(x^k) > 0$, то перейти к шагу 9;

б) если нет, то перейти к шагу 10, положив $d^k = -\nabla f(x^k)$.

Шаг 9. Определить $d_k = -H^{-1}(x^k)\nabla f(x^k)$.

Шаг 10. Найти точку $x^{k+1} = x^k + t_k d^k$, положив $t_k = 1$, если $d_k = -H^{-1}(x^k)\nabla f(x^k)$, или выбрав t_k из условия $f(x^{k+1}) < f(x^k)$, если $d^k = -\nabla f(x^k)$.

Шаг 11. Проверить выполнение условий $\|x^{k+1} - x^k\| < \varepsilon_2$, $|f(x^{k+1}) - f(x^k)| < \varepsilon_2$:

а) если оба условия выполнены при текущем значении k и $k = k - 1$, то расчет окончен, $x^* = x^{k+1}$;

б) в противном случае положить $k = k + 1$ и перейти к шагу 3.

СХОДИМОСТЬ

Утверждение 7.1. Пусть $f(x)$ дважды непрерывно дифференцируемая сильно выпуклая функция с константой $l > 0$ на R^n и удовлетворяет условию

$$\|H(x) - H(y)\| \le L\|x - y\| \quad \forall\ x, y \in R^n,$$

где $L > 0$, а начальная точка такова, что

$$\|\nabla f(x^0)\| \le 8\frac{l^2}{L},$$

т. е.

$$\|\nabla f(x^0)\| = \frac{8l^2 q}{L}, \tag{7.3}$$

где $q \in (0, 1)$. Тогда последовательность $\{x^k\}$ сходится к точке минимума с квадратичной скоростью $\|x^k - x^*\| \le \frac{4lq^{2^k}}{L}$ [39].

Замечания 7.1.

1. Сходимость метода Ньютона доказана лишь для сильно выпуклых функций и для достаточно хорошего начального приближения, определяемого условием (7.3), практическое использование которого крайне затруднено, так как постоянные l и L, как правило, неизвестны или требуют трудоемкого исследования для их определения. Поэтому при практическом использовании метода Ньютона следует:

а) анализировать матрицу $H(x^k)$ на выполнение условия $H(x^k) > 0\ \forall\ k = 0, 1, \ldots$ и заменять формулу $x^{k+1} = x^k - H^{-1}(x^k)\nabla f(x^k)$ на формулу $x^{k+1} = x^k - t_k\nabla f(x^k)$ в случае его невыполнения;

б) производить анализ точки x^k с целью выяснения, является ли она найденным приближением искомой точки x^*.

2. При решении задачи поиска безусловного максимума формула (7.2) не изменяется, так как в этом случае $H(x^k) < 0$.

ПРОЦЕДУРА РЕШЕНИЯ ЗАДАЧИ

1. Используя алгоритм Ньютона, построить точку x^k, в которой выполняется по крайней мере один критерий окончания расчета.

2. Так как $f(x) \in C^2$, то осуществить проверку выполнения достаточных условий минимума $H(x^k) > 0$. Если условие выполнено, то точка x^k может рассматриваться как найденное приближение точки минимума x^*. Проверку выполнения достаточных условий минимума можно заменить проверкой функции $f(x)$ на выпуклость.

Пример 7.1. Найти локальный минимум функции $f(x) = 2x_1^2 + x_1x_2 + x_2^2$.

☐ I. Определение точки x^k, в которой выполняется по крайней мере один критерий окончания расчетов.

1. Зададим x^0, ε_1, ε_2, M: $x^0 = (0{,}5; 1)^T$; $\varepsilon_1 = 0{,}1$, $\varepsilon_2 = 0{,}15$, $M = 10$. Найдем градиент функции $\nabla f(x) = (4x_1 + x_2; x_1 + 2x_2)^T$ и матрицу Гессе

$$H(x) = \begin{pmatrix} 4 & 1 \\ 1 & 2 \end{pmatrix}.$$

2. Положим $k = 0$.

3^0. Вычислим $\nabla f(x^0)$: $\nabla f(x^0) = (3; 2{,}5)^T$.

4^0. Проверим выполнение условия $\|\nabla f(x^0)\| \le \varepsilon_1$: $\|\nabla f(x^0)\| = 3{,}9 > 0{,}1$. Переходим к шагу 5.

5^0. Проверим выполнение условия $k \ge M$: $k = 0 < 10$. Переходим к шагу 6.

6^0. Вычислим $H(x^0)$:

$$H(x^0) = \begin{pmatrix} 4 & 1 \\ 1 & 2 \end{pmatrix}.$$

7^0. Вычислим $H^{-1}(x^0)$:

$$H^{-1}(x^0) = \begin{pmatrix} \frac{2}{7} & -\frac{1}{7} \\ -\frac{1}{7} & \frac{4}{7} \end{pmatrix}.$$

8^0. Проверим выполнение условия $H^{-1}(x^0) > 0$. Так как

$$\Delta_1 = \frac{2}{7} > 0, \quad \Delta_2 = \frac{1}{7} > 0,$$

то согласно критерию Сильвестра $H^{-1}(x^0) > 0$ (см. § 2).

9^0. Определим

$$d_0 = -H^{-1}(x^0)\nabla f(x^0) = -\begin{pmatrix} \frac{2}{7} & -\frac{1}{7} \\ -\frac{1}{7} & \frac{4}{7} \end{pmatrix}\begin{pmatrix} 3 \\ 2{,}5 \end{pmatrix} = \begin{pmatrix} -\frac{1}{2} \\ -1 \end{pmatrix}.$$

10^0. Вычислим $x^1 = \left(\frac{1}{2}, 1\right)^T + \left(-\frac{1}{2}, -1\right)^T = (0, 0)^T$.

11^0. Проверим выполнение условий $\|x^1 - x^0\| < \varepsilon_2$, $|f(x^1) - f(x^0)| < \varepsilon_2$:

$$\|x^1 - x^0\| = 1{,}12 > 0{,}15; \; |f(x^1) - f(x^0)| = 2 > 0{,}15.$$

Полагаем $k = 1$, переходим к шагу 3.

3^1. Вычислим $\nabla f(x^1)$: $\nabla f(x^1) = (0, 0)^T$.

4^1. Проверим выполнение условия $\|\nabla f(x^1)\| < \varepsilon_1$: $\|\nabla f(x^1)\| = 0 < 0{,}1$. Расчет окончен. Заметим, что в точке x^1 выполняется необходимое условие первого порядка, поэтому она является стационарной точкой.

II. Анализ точки x^1.

Функция $f(x) = 2x_1^2 + x_1x_2 + x_2^2$ является строго выпуклой, так как ее матрица вторых производных

$$H = \begin{pmatrix} 4 & 1 \\ 1 & 2 \end{pmatrix} > 0$$

в силу того, что $\Delta_1 = 4 > 0$, $\Delta_2 = 7 > 0$. Найденная точка $x^1 = (0, 0)^T$ есть точка локального и одновременно глобального минимума функции $f(x)$.

На рисунке 7.2 траектория спуска изображена сплошной линией. ■

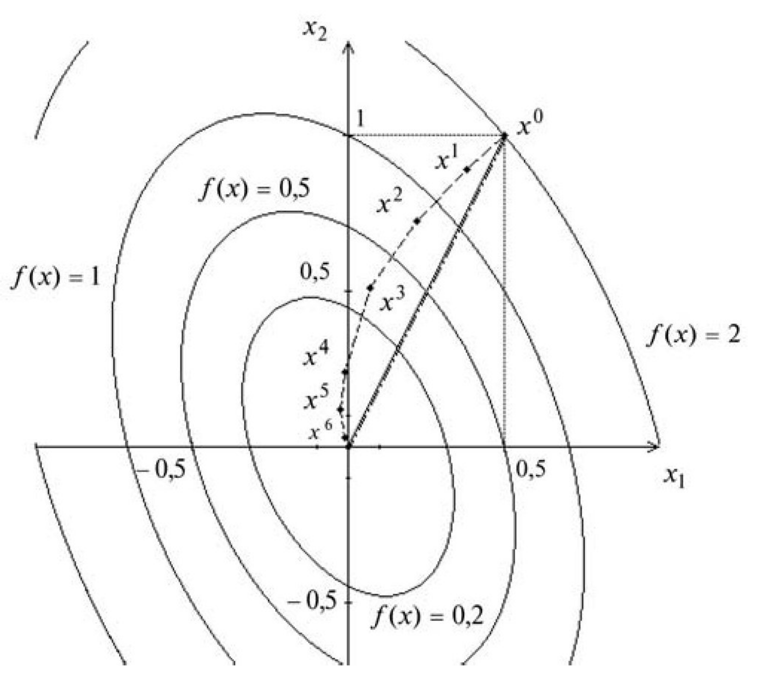

Рис. 7.2