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

Метод Марквардта

МЕТОД МАРКВАРДТА

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

СТРАТЕГИЯ ПОИСКА

Стратегия метода Марквардта [D. W. Marquardt] 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 $\{x^k\}$ вычисляются по правилу

$$x^{k+1} = x^k - [H(x^k) + \mu^k E]^{-1} \nabla f(x^k),\ k = 0, 1, \ldots, \qquad (7.7)$$

где точка x^0 задается пользователем; E — единичная матрица; μ^k — последовательность положительных чисел таких, что матрица $[H(x^k) + \mu^k E]^{-1}$ положительно определена. Как правило, число μ^0 назначается как минимум на порядок больше, чем самый большой элемент матрицы $H(x^0)$, а в ряде стандартных программ полагается $\mu^0 = 10^4$ [36]. Если $f(x^k - (H(x^k) + \mu^k E)^{-1} \nabla f(x^k)) < f(x^k)$, то $\mu^{k+1} = \dfrac{\mu^k}{2}$. В противном случае $\mu^{k+1} = 2\mu^k$. Легко видеть, что алгоритм Марквардта в зависимости от величины μ^k на каждом шаге по своим свойствам либо приближается к алгоритму Ньютона, либо к алгоритму градиентного спуска.

Построение последовательности $\{x^k\}$ заканчивается, когда либо $\|\nabla f(x^k)\| < \varepsilon_1$, либо число итераций $k \ge M$, где ε_1 — малое положительное число, а M — предельное число итераций.

Вопрос о том, может ли точка x^k рассматривается как найденное приближение искомой точки минимума, решается путем проведения дополнительного исследования, которое описано ниже.

АЛГОРИТМ

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

Шаг 2. Положить $k = 0$, $\mu^k = \mu^0$.

Шаг 3. Вычислить $\nabla f(x^k)$.

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

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

б) если нет, перейти к шагу 5.

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

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

б) если нет, перейти к шагу 6.

Шаг 6. Вычислить $H(x^k)$.

Шаг 7. Вычислить $H(x^k) + \mu^k E$.

Шаг 8. Вычислить $[H(x^k) + \mu^k E]^{-1}$.

Шаг 9. Вычислить $d_k = -[H(x^k) + \mu^k E]^{-1} \nabla f(x^k)$.

Шаг 10. Вычислить $x^{k+1} = x^k - [H(x^k) + \mu^k E]^{-1} \nabla f(x^k)$.

Шаг 11. Проверить выполнение условия $f(x^{k+1}) < f(x^k)$:

а) если неравенство выполняется, то перейти к шагу 12;

б) если нет, перейти к шагу 13.

Шаг 12. Положить $k = k + 1$, $\mu^{k+1} = \dfrac{\mu^k}{2}$ и перейти к шагу 3.

Шаг 13. Положить $\mu^k = 2\mu^k$ и перейти к шагу 7.

ПРОЦЕДУРА РЕШЕНИЯ ЗАДАЧИ

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

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.

Замечания 7.3.

1. Метод Марквардта за счет выбора μ^k обеспечивает построение последовательности $\{x^k\}$ такой, что $f(x^{k+1}) < f(x^k)$, $k = 0, 1, \ldots$ [29].

2. В окрестности точки минимума x^* метод Марквардта обладает скоростью сходимости, близкой к квадратичной [29].

Пример 7.3. Найти локальный минимум функции $f(x) = 2x_1^2 + x_1 x_2 + x_2^2$.

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

1. Зададим $x^0 = (0{,}5;\ 1)^T$; $\varepsilon_1 = 0{,}1$; $M = 10$. 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 $k = 0$, $\mu^0 = 20$.

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 $H(x^0) + \mu^0 E$:

$$H(x^0) + \mu^0 E = \begin{pmatrix} 24 & 1 \\ 1 & 22 \end{pmatrix}.$$

8^0. Вычислим $[H(x^0) + \mu^0 E]^{-1}$:

$$[H(x^0) + \mu^0 E]^{-1} = \begin{bmatrix} 0{,}0417 & -0{,}0019 \\ -0{,}0019 & 0{,}0455 \end{bmatrix}.$$

9^0. Вычислим $d^0 = -[H(x^0) + \mu^0 E]^{-1} \nabla f(x^0)$: $d^0 = (-0{,}119;\ -0{,}108)^T$.

10^0. Вычислим $x^1 = x^0 - [H(x^0) + \mu^0 E]^{-1} \nabla f(x^0)$: $x^1 = (0{,}381;\ 0{,}892)^T$.

11^0. Проверим выполнение условия $f(x^1) < f(x^0)$: $f(x^1) = 1{,}438 < 2 = f(x^0)$.

12^0. Полагаем $k = 1$, $\mu^1 = \dfrac{\mu^0}{2} = 10$ и переходим к шагу 3.

3^1. Вычислим $\nabla f(x^1)$: $\nabla f(x^1) = (2{,}41;\ 2{,}16)^T$.

4^1. Проверим выполнение условия $\|\nabla f(x^1)\| < \varepsilon_1$: $\|\nabla f(x^1)\| = 3{,}18 > 0{,}1$. Переходим к шагу 5.

5^1. Проверим выполнение условия $k \ge M$: $k = 1 < 10$. Переходим к шагу 6.

6^1. Вычислим $H(x^1)$:

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

7^1. Вычислим $H(x^1) + \mu^1 E$:

$$H(x^1) + \mu^1 E = \begin{pmatrix} 14 & 1 \\ 1 & 12 \end{pmatrix}.$$

8^1. Вычислим $[H(x^1) + \mu^1 E]^{-1}$:

$$[H(x^1) + \mu^1 E]^{-1} = \begin{bmatrix} 0{,}072 & -0{,}0059 \\ -0{,}0059 & 0{,}084 \end{bmatrix}.$$

9^1. Вычислим $d^1 = -[H(x^1) + \mu^1 E]^{-1} \nabla f(x^1)$: $d^1 = (-0{,}160;\ -0{,}168)^T$.

10^1. Вычислим $x^2 = x^1 - [H(x^1) + \mu^1 E]^{-1} \nabla f(x^1)$:

$$x^2 = (0{,}381;\ 0{,}892)^T - (0{,}160;\ 0{,}168)^T = (0{,}221;\ 0{,}724)^T.$$

11^1. Проверим выполнение условия $f(x^2) < f(x^1)$:

$$f(x^2) = 0{,}791 < 1{,}438 = f(x^1).$$

12^1. Полагаем $k = 2$, $\mu^2 = \dfrac{\mu^1}{2} = 5$ и переходим к шагу 3.

3^2. Вычислим $\nabla f(x^2)$: $\nabla f(x^2) = (1{,}60;\ 1{,}67)^T$.

4^2. Проверим выполнение условия $\|\nabla f(x^2)\| < \varepsilon_1$: $\|\nabla f(x^2)\| = 2{,}31 > 0{,}1$. Переходим к шагу 5.

5^2. Проверим выполнение условия $k \ge M$: $k = 2 < 10$. Переходим к шагу 6.

6^2. Вычислим $H(x^2)$:

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

7^2. Вычислим $H(x^2) + \mu^2 E$:

$$H(x^2) + \mu^2 E = \begin{pmatrix} 9 & 1 \\ 1 & 7 \end{pmatrix}.$$

8^2. Вычислим $[H(x^2) + \mu^2 E]^{-1}$:

$$[H(x^2) + \mu^2 E]^{-1} = \begin{pmatrix} 0{,}113 & -0{,}016 \\ -0{,}016 & 0{,}145 \end{pmatrix}.$$

9^2. Вычислим $d^2 = -[H(x^2) + \mu^2 E]^{-1} \nabla f(x^2)$: $d^2 = (-0{,}155;\ -0{,}217)^T$.

10^2. Вычислим $x^3 = x^2 - [H(x^2) + \mu^2 E]^{-1} \nabla f(x^2)$:

$$x^3 = (0{,}221;\ 0{,}724)^T - (0{,}155;\ 0{,}217)^T = (0{,}07;\ 0{,}51)^T.$$

11^2. Проверим выполнение условия $f(x^3) < f(x^2)$:

$$f(x^3) = 0{,}3 < 0{,}791 = f(x^2).$$

12^2. Полагаем $k = 3$, $\mu^3 = \dfrac{\mu^2}{2} = 2{,}5$ и переходим к шагу 3.

3^3. Вычислим $\nabla f(x^3)$: $\nabla f(x^3) = (0{,}79;\ 1{,}09)^T$.

4^3. Проверим выполнение условия $\|\nabla f(x^3)\| < \varepsilon_1$: $\|\nabla f(x^3)\| = 1{,}34 > 0{,}1$. Переходим к шагу 5.

5^3. Проверим выполнение условия $k \ge M$: $k = 3 < 10$. Переходим к шагу 6.

6^3. Вычислим $H(x^3)$:

$$H(x^3) = \begin{pmatrix} 4 & 1 \\ 1 & 2 \end{pmatrix}.$$

7^3. Вычислим $H(x^3) + \mu^3 E$:

$$H(x^3) + \mu^3 E = \begin{pmatrix} 6{,}5 & 1 \\ 1 & 4{,}5 \end{pmatrix}.$$

8^3. Вычислим $[H(x^3) + \mu^3 E]^{-1}$:

$$[H(x^3) + \mu^3 E]^{-1} = \begin{pmatrix} 0{,}159 & -0{,}035 \\ -0{,}035 & 0{,}23 \end{pmatrix}.$$

9^3. Вычислим $d^3 = -[H(x^3) + \mu^3 E]^{-1} \nabla f(x^3)$: $d^3 = (-0{,}078;\ -0{,}22)^T$

10^3. Вычислим $x^4 = x^3 - [H(x^3) + \mu^3 E]^{-1} \nabla f(x^3)$:

$$x^4 = (0{,}07;\ 0{,}51)^T - (0{,}078;\ 0{,}22)^T = (-0{,}008;\ 0{,}29)^T.$$

11^3. Проверим выполнение условия $f(x^4) < f(x^3)$:

$$f(x^4) = 0{,}082 < 0{,}3 = f(x^3).$$

12^3. Полагаем $k = 4$, $\mu^4 = \dfrac{\mu^3}{2} = 1{,}25$ и переходим к шагу 3.

3^4. Вычислим $\nabla f(x^4)$: $\nabla f(x^4) = (0{,}26;\ 0{,}57)^T$.

4^4. Проверим выполнение условия $\|\nabla f(x^4)\| < \varepsilon_1$: $\|\nabla f(x^4)\| = 0{,}62 > 0{,}1$. Переходим к шагу 5.

5^4. Проверим выполнение условия $k \ge M$: $k = 4 < 10$. Переходим к шагу 6.

6^4. Вычислим $H(x^4)$:

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

7^4. Вычислим $H(x^4) + \mu^4 E$:

$$H(x^4) + \mu^4 E = \begin{pmatrix} 5{,}25 & 1 \\ 1 & 3{,}25 \end{pmatrix}.$$

8^4. Вычислим $[H(x^4) + \mu^4 E]^{-1}$:

$$[H(x^4) + \mu^4 E]^{-1} = \begin{pmatrix} 0{,}203 & -0{,}0623 \\ -0{,}0623 & 0{,}327 \end{pmatrix}.$$

9^4. Вычислим $d^4 = -[H(x^4) + \mu^4 E]^{-1} \nabla f(x^4)$: $d^4 = (-0{,}017;\ -0{,}17)^T$.

10^4. Вычислим $x^5 = x^4 - [H(x^4) + \mu^4 E]^{-1} \nabla f(x^4)$:

$$x^5 = (-0{,}008;\ 0{,}29)^T - (0{,}017;\ 0{,}17)^T = (-0{,}025;\ 0{,}12)^T.$$

11^4. Проверим выполнение условия $f(x^5) < f(x^4)$:

$$f(x^5) = 0{,}012 < 0{,}082 = f(x^4).$$

12^4. Полагаем $k = 5$, $\mu^5 = \dfrac{\mu^4}{2} = 0{,}625$ и переходим к шагу 3.

3^5. Вычислим $\nabla f(x^5)$: $\nabla f(x^5) = (0{,}02;\ 0{,}22)^T$.

4^5. Проверим выполнение условия $\|\nabla f(x^5)\| < \varepsilon_1$: $\|\nabla f(x^5)\| = 0{,}22 > 0{,}1$. Переходим к шагу 5.

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

6^5. Вычислим $H(x^5)$:

$$H(x^5) = \begin{pmatrix} 4 & 1 \\ 1 & 2 \end{pmatrix}.$$

7^5. Вычислим $H(x^5) + \mu^5 E$:

$$H(x^5) + \mu^5 E = \begin{pmatrix} 4{,}625 & 1 \\ 1 & 2{,}625 \end{pmatrix}.$$

8^5. Вычислим $[H(x^5) + \mu^5 E]^{-1}$:

$$[H(x^5) + \mu^5 E]^{-1} = \begin{pmatrix} 0{,}236 & -0{,}09 \\ -0{,}09 & 0{,}416 \end{pmatrix}.$$

9^5. Вычислим $d^5 = -[H(x^5) + \mu^5 E]^{-1} \nabla f(x^5)$: $d^5 = (0{,}015;\ -0{,}090)^T$.

10^5. Вычислим $x^6 = x^5 - [H(x^5) + \mu^5 E]^{-1} \nabla f(x^5)$:

$$x^6 = (-0{,}025;\ 0{,}12)^T - (0{,}015;\ -0{,}09)^T = (-0{,}01;\ 0{,}03)^T.$$

11^5. Проверим выполнение условия $f(x^6) < f(x^5)$:

$$f(x^6) = 0{,}0006 < 0{,}012 = f(x^5).$$

12^5. Полагаем $k = 6$, $\mu^6 = \dfrac{\mu^5}{2} = 0{,}311$ и переходим к шагу 3.

3^6. Вычислим $\nabla f(x^6)$: $\nabla f(x^6) = (-0{,}01;\ 0{,}05)^T$.

4^6. Проверим выполнение условия $\|\nabla f(x^6)\| < \varepsilon_1$: $\|\nabla f(x^6)\| = 0{,}051 < 0{,}1$. Расчет окончен.

II. Анализ точки x^6.

Точка $x^6 = (-0{,}01;\ 0{,}03)^T$ (см. рис. 7.2) является найденным приближением точки минимума x^*, так как функция $f(x) = 2x_1^2 + x_1 x_2 + x_2^2$ является строго выпуклой (см. пример 7.1). На рисунке 7.2 полученная траектория спуска изображена штриховой линией.

Вопросы

ПРЕДЫДУЩИЙ ЭЛЕМЕНТ КУРСА

◄ ВТ. Сравнительный анализ эффективности численных методов первого порядка для поиска безусловного экстремума

Перейти на...

СЛЕДУЮЩИЙ ЭЛЕМЕНТ КУРСА

ВТ. Задание №13. Метод Ньютона ►