6. ЛЕКЦИЯ. Методы первого порядка. Метод Флетчера-Ривса

Идея метода Флетчера—Ривса состоит в том, что на каждом шаге в качестве направления спуска используется линейная комбинация вектора градиента с прежним направлением спуска. Последовательность приближений

к точке минимума целевой функции *f(x)* определяется по формуле $\overline{x^{(k+1)}} = \overline{x^{(k)}} + h_k \overline{p^{(k)}}$, *k = 0,1,..., где k — номер итерации (k = 0,1,...),*

$\overline{x^{(0)}}$—начальное приближение,

h_k — величина шага,

$\overline{p^{(k)}}$ — направление спуска.

Направление спуска $\overline{p^{(k)}}$ определяется в зависимости от номера текущей итерации *k:*

• Если значение $k = 0$ направление спуска $\overline{p^{(0)}}$ совпадает с направлением вектора антиградиента целевой функции — $\nabla f(\overline{x^{(0)}})$в точке $\overline{x^{(0)}}$.

•Если значение $k = 1,2,\ldots$ направление спуска определяется по формуле

$$\overline{p^{(k)}} = \nabla f(\overline{x^{(k)}}) + \beta_{k-1}\overline{p^{(k-1)}},$$

где $$\beta_{k-1} = \frac{\left\|\nabla f\left(\overline{x^{(k)}}\right)\right\|^2}{\left\|\nabla f\left(\overline{x^{(k-1)}}\right)\right\|^2} = \frac{\sum_{i=1}^{n}\left(\frac{\partial f\left(\overline{x^{(k)}}\right)}{dx_i}\right)^2}{\sum_{i=1}^{n}\left(\frac{\partial f\left(\overline{x^{(k-1)}}\right)}{dx_i}\right)^2}$$

и называется сопряженным.

В качестве условия окончания поиска используется близость к нулю нормы градиента $\left\|\nabla f\left(\overline{x^{(k+1)}}\right)\right\| \le \varepsilon$.

Для минимизации квадратичной целевой функции n независимых переменных методом Флетчера –Ривса требуется не более n шагов. Для уменьшения влияния накапливающихся погрешностей вычислений через каждые n итераций проводят обновление метода, полагая $\overline{x^{(0)}} = \overline{x^{(n)}}$.

Иллюстрация последовательных приближений к точке минимума представлена на рис. 2.22.

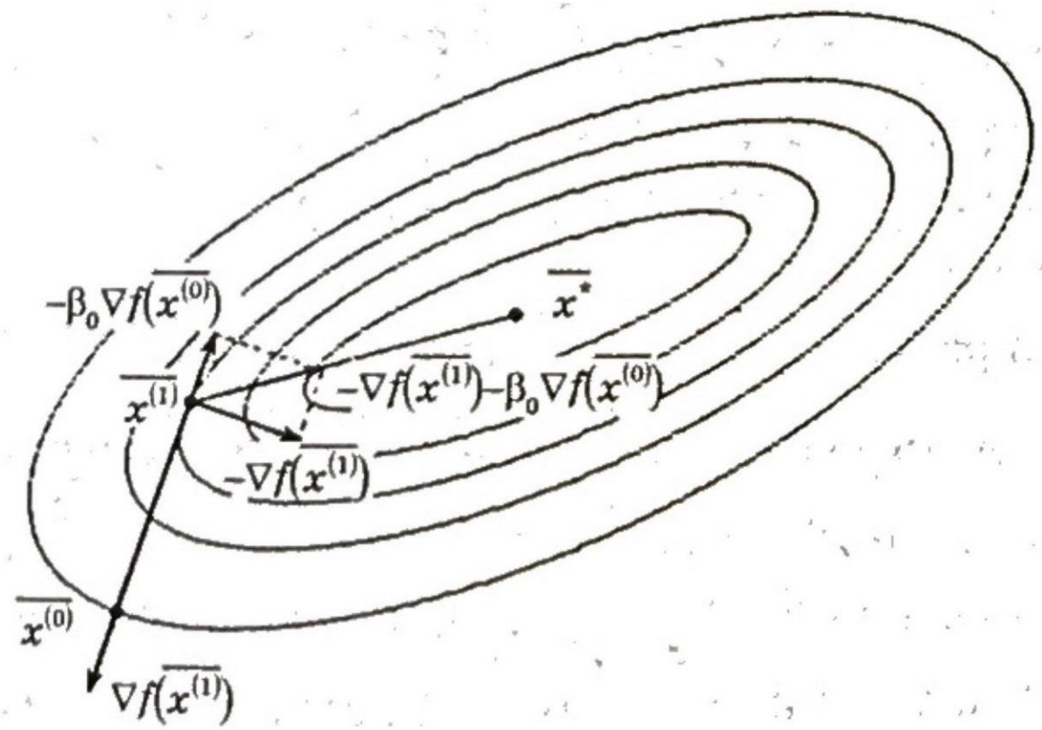

Рис. 2.22. Траектория поиска минимума методом Флетчера—Ривса

Величина шага h_k, определяется из решения вспомогательной одномерной задачи минимизации $\varphi(h_k)= f(\overline{x^{(k)}}-h_k\nabla f(\overline{x^{(k)}})) \rightarrow \min$ которая может быть решена аналитически или численно. При квадратичной интерполяции целевой функции величину шага можно определить по формуле, аналогичной формуле метода наискорейшего градиентного спуска.

$$h_k=\frac{\left(\nabla f\overline{(x^{(k)})},\overline{p^{(k)}}\right)}{\left(H\overline{\left(x^{(k)}\right)p^{(k)}},\overline{p^{(k)}}\right)},$$

где $H\overline{(x^{(k)})}$ — матрица Гессе, вычисленная в точке $\overline{x^{(k)}}$.

Алгоритм метода Флетчера—Ривса

1. Задать размерность задачи оптимизации n, координаты начальной точки $\overline{x^{(0)}} = (x_1^{(0)}, x_2^{(0)},\ldots, x_n^{(0)})$, точность поиска ε.

2. Положить счетчик числа итераций $k = 0$.

3. Определить направление вектора градиента

$\nabla f\overline{(x^{(k)})}$=$\left(\frac{\partial f\overline{(x^{(k)}}}{dx_1}, \frac{\partial f\overline{(x^{(k)}}}{dx_2}, \dots, \frac{\partial f\overline{(x^{(k)}}}{dx_n}\right)$ целевой функции $\overline{f(x)}$ в точке $\overline{x^{(k)}}$ = ($x_1^{(k)}$, $x_2^{(k)}$,…, $x_n^{(k)}$).

4. Проверить условие окончания поиска

$\left\|\nabla f\overline{(x^{(k)})}\right\| = \sqrt{\sum_{i=1}^{n}\left(\frac{\partial f\overline{(x^{(k)}}}{dx_i}\right)^2} \leq \varepsilon$. Если условие выполнено то расчет окончен

$\overline{x^*} = \overline{x^{(k)}}$, иначе перечти к пункту 5.

5. Если k = 0, то вычислить $\overline{p^{(0)}}$ = -$\nabla\overline{f(x^{(0)})}$ и перейти к шагу 8, иначе – перейти к шагу 6.

6. Вычислить β_{k-1}=$\frac{\left\|\nabla \mathrm{f}\overline{(x^{(k)})}\right\|^2}{\left\|\nabla \mathrm{f}\overline{(x^{(k-1)})}\right\|^2} = \frac{\sum_{i=1}^{n}\left(\frac{\partial \mathrm{f}\overline{(x^{(k)}}}{dx_i}\right)^2}{\sum_{i=1}^{n}\left(\frac{\partial \mathrm{f}\overline{(x^{(k-1)}}}{dx_i}\right)^2}$

7. Вычислить $\overline{p^{(k)}} = -\nabla f\overline{(x^{(k)})} + \beta_{k-1}\overline{p^{(k-1)}}$.

8. Вычислить шаг h_k формуле,

$$h_k=\frac{\left(\nabla f\overline{(x^{(k)})},\overline{p^{(k)}}\right)}{\left(H\overline{\left(x^{(k)}\right)}\overline{p^{(k)}},\overline{p^{(k)}}\right)},$$

используя результаты пункта 3 и разностные формулы.

Это вторые частные производные по формуле

$$\frac{\partial f^2\left(\overline{x^{(k)}}\right)}{\partial x_i^2} \approx \frac{u_1-2u_2+u_2}{(\Delta x_i^{(k)})^2}, \tag{2.5}$$

И смешанные производные по формуле

$$\frac{\partial f\left(\overline{x^{(k)}}\right)}{\partial x_i \partial x_j} \approx \frac{u_1-u_2-u_3+u_4}{\Delta x_i^{(k)}\Delta x_j^{(k)}}, \tag{2.6}$$

9. Определить координаты точки $\overline{x^{(k+1)}} = \overline{x^{(k)}} + h_k\overline{p^{(k)}}$,

Положить $k = k + 1$ и перейти к пункту 3.

Пример. Найти минимум целевой функции

$$f(\overline{x}) = x_1^2 - x_1x_2 + 3x_2^2 - x_1$$

методом Флетчера – Ривса с точностью ε = 0.1.

Решение. За начальную точку примем

$$\overline{x^{(0)}}=(x_1^{(0)}, x_2^{(0)})^T = (0,0)^T$$

Найдем градиент функции в произвольной точке $\overline{x^{(k)}} = (x_1^{(k)}, x_2^{(k)})^T$

$$\nabla f\left(\overline{x^{(k)}}\right) = \left(\frac{\partial f\left(\overline{x^{(k)}}\right)}{dx_1};\ \frac{\partial f\left(\overline{x^{(k)}}\right)}{dx_2}\right)^T = (2x_1^{(k)} - x_x^{(k)} - 1;\ -x_1^{(k)} + 6x_2^{(k)})^T.$$

Итерация k = 0. Определим координаты точки $\overline{x^{(1)}} = \overline{x^{(0)}} + h_0\overline{p^{(0)}}$.

Для этого вычислим градиент функции $\nabla f\left(\overline{x^{(0)}}\right)$ в начальной точке

$\overline{x^{(0)}}$: $\nabla f\left(\overline{x^{(0)}}\right) = (-1;\ 0)^T$. *Положим* $\overline{p^{(0)}} = -\nabla f\left(\overline{x^{(0)}}\right) = (1;\ 0)^T$

Вычислим координаты точки $\overline{x^{(1)}}$

$$\overline{x^{(1)}} = \overline{x^{(0)}} - h_0\overline{p^{(0)}} = (0;\ 0)^T - h_0(-1;\ 0)^T = (h_0,\ 0)^T.$$

Найдем значение h_0 из условия $\varphi(h0)= f(\overline{x^{(0)}}-h0\nabla f(\overline{x^{(0)}})) \to \min\limits_{h_0>0}$

Подставляя координаты точки $\overline{x^{(1)}}$ в целевую функцию $f(\overline{(x}\overline{)}$, получаем

$\varphi(h_0) = f(\overline{x^{(0)}}) = f(h_0,\ 0) = h_0^2 - h_0$ - функция одной переменной.

Запишем необходимое условие безусловного экстремума функции $\varphi(h_0)$:

$\frac{\partial\varphi(h_0)}{\partial h_0} = 2h_0 - 1 = 0.$ В результате решения определим $h_0 = \frac{1}{2}$. Проверим выполнение достаточных условий экстремума: $\frac{\partial^2\varphi(h_0)}{\partial {h_0}^2} = 2 > 0$

Найденное значение шага h_0 обеспечивает минимум функции $\varphi(h_0)$. По данной величине шага находим координаты точки $\overline{x^{(1)}}$:

$$\overline{x^{(1)}} = \overline{x^{(0)}} - h_0\overline{p^{(0)}} = (0;\,0)^T - \tfrac{1}{2}(-1;\,0)^T = (\tfrac{1}{2},\,0)^T.$$

Итерация k = 1.

Определим координаты точки по формуле:

$$\overline{x^{(2)}} = \overline{x^{(1)}} + h_1\overline{p^{(1)}}$$

$$Где\ \overline{p^{(1)}} = -\nabla f\overline{(x^{(1)})} + \beta_0\overline{p^{(0)}}.$$

Определим составляющие вектора $\overline{\mathrm{p}^{(1)}}$:

Вычислим градиент целевой функции $\nabla f\overline{(x^{(1)})}$ в точке

$$\overline{\mathrm{x}^{(1)}}:\ \nabla f\overline{(x^{(1)})} = (0;\ -\tfrac{1}{2})^T.$$

Найдем координаты вектора антиградиента

$$-\nabla f\overline{(\mathrm{x}^{(1)})} = (0;\,\tfrac{1}{2})^{\mathrm{T}}\ ;\ \beta_0 = \frac{\left\|\nabla \mathrm{f}\overline{(\mathrm{x}^{(1)})}\right\|^2}{\left\|\nabla \mathrm{f}\overline{(\mathrm{x}^{(0)})}\right\|^2} = \frac{0+\left(\frac{1}{2}\right)}{(-1)^2+0} = \frac{1}{4}.$$

$$\overline{x^{(2)}} = \overline{x^{(1)}} + h_1\overline{p^{(1)}} = \left(\frac{1}{2};0\right)^T + h_1\left[\left(0;\,\frac{1}{2}\right)^T + \frac{1}{4}(1;0)^T\right] =$$

$$\left(\frac{1}{2} + \frac{h_1}{4};\ \frac{h_1}{2}\right)^T.$$

Найдем значение h_1 из условия $\varphi(h_1) = f(\overline{\mathrm{x}^{(1)}} - h_1\nabla f\overline{(\mathrm{x}^{(1)})}) \to \min\limits_{h_1>0}$

Подставляя координаты точки $\overline{x^{(2)}}$ в целевую функцию $f(\overline{x})$ получаем

$\varphi(h_1) = f\overline{(x^{(2)})} = f(h_0;\,0) = \frac{11}{16}h^2 - \frac{1}{4}h_1 = \frac{1}{4}$ – функция одной переменной.

Запишем необходимое условие безусловного экстремума функции $\varphi(h_1)$:

$\frac{\partial\varphi(h_1)}{\partial h_1} = \frac{11}{8}h_1 - \frac{1}{4} = 0$. В результате решения уравнения определим $h_1 = \frac{2}{11}$.

Проверим выполнение достаточных условий экстремума.

$$\frac{\partial^2\varphi(h_0)}{\partial h_0^2} = \frac{11}{8} > 0.$$

Найденное значение величины шага h_1 обеспечивает минимум функции $\varphi(h_1)$. По данной величине шага находим координаты точки $\overline{x^{(2)}}$.

$$\overline{x^{(2)}} = \overline{x^{(1)}} + h_1\overline{p^{(1)}} = \left(\frac{1}{2};0\right)^T + \frac{2}{11}\left[\left(0;-\frac{1}{2}\right)^T + \frac{1}{4}(1;0)^T\right] =$$

$$\left(\frac{6}{11};\frac{1}{11}\right)^T.$$

Точка $\overline{x^{(2)}}$ есть найденное приближение точки минимума $\overline{x^*} = \left(\frac{6}{11};\frac{1}{11}\right)^T$,

$f\overline{(x^{(1)})} = -0.27$.

Методы второго порядка

В методах второго порядка при поиске минимума функции многих переменных используют информацию о частных производных целевой функции первого и второго порядка. К этой группе относят метод Ньютона и его модификации.

Метод Ньютона

В основе метода Ньютона лежит квадратичная’ аппроксимация целевой функции. Последовательность итераций строится таким образом, чтобы во вновь получаемой точке градиент аппроксимирующей функции обращался в нуль.

Последовательность приближений строится в соответствии с формулой

$$\overline{x^{(k+1)}} = \overline{x^{(k)}} + \overline{p^{(k)}},$$

где k — номер итерации (k = 0, 1, …),

$\overline{x^{(0)}}$ — начальное приближение,

$\overline{p^{(k)}} = -H^{-1}(\overline{x^{(k)}})\nabla f(\overline{x^{(k)}})$ — вектор направления спуска.

Здесь $H(\overline{x^{(k)}})$ — матрица Гессе.

Направление спуска $\overline{p^{(k)}}$ ведет к убыванию целевой функции только при положительной определенности матрицы Гессе $H(\overline{x^{(k)}}) > 0$. В тех итерациях, в которых матрица Гессе отрицательно определена $H(\overline{x^{(k)}}) < 0$, последовательность приближений к точке минимума строится по методу наискорейшего градиентного спуска. С этой целью проводится замена вектора направления спуска на антиградиентное $\overline{p^{(k)}} = -h_k \nabla f(\overline{x^{(k)}})$.

Алгоритм метода Ньютона.

1. Задать размерность задачи оптимизации n, координаты начальной точки $\overline{x^{(0)}} = (x_1^{(0)}, x_2^{(0)}, \ldots, x_n^{(0)})$, точность поиска Ɛ.

2. Положить счетчик итераций $k = 0$.

3. Определить направление вектора градиента целевой функции

$f(\overline{x})\nabla f(\overline{x^{(k)}}) = \left(\frac{\partial f(\overline{x^{(k)}})}{\partial x_1}, \frac{\partial f(\overline{x^{(k)}})}{\partial x_2}, \ldots, \frac{\partial f(\overline{x^{(k)}})}{\partial x_n}\right)$ в точке

$\overline{x^{(k)}} = (x_1^{(k)}, x_2^{(k)}, \ldots, x_n^{(k)})$.

Для вычисления координат вектора градиента использовать разностную формулу (2.3).

Первые частные производные вычисляются по формуле (2.3).

$$\frac{\partial f(\overline{x^{(k)}})}{\partial x_i} \approx \qquad (2.3)$$

$$\approx \frac{f(x_1{}^{(k)}, x_2{}^{(k)}, \ldots, x_i{}^{(k)} + \Delta x_i{}^{(k)}, \ldots, x_n{}^{(k)}) - f(x_1{}^{(k)}, x_2{}^{(k)}, \ldots, x_i{}^{(k)} \ldots, x_n{}^{(k)})}{\Delta x_i{}^{(k)}}$$

4. Проверить условие окончания поиска

$$\left\|\nabla f(\overline{x^{(k)}})\right\| = \sqrt{\sum_{i=1}^{n} \left(\frac{\partial f(\overline{x^{(k)}})}{\partial x_i}\right)^2} \leqslant \varepsilon .$$

Если условие выполнено, то расчет окончен $\overline{x^*} = \overline{x^{(k)}}$, иначе перейти к пункту 5.

5. Сформировать матрицу Гессе $H(\overline{x^{(k)}})$, используя разностные формулы вычисления вторых (2.5) и смешанных производных (2.6).

6. Проверить положительную определенность матрицы Гессе $H(\overline{x^{(k)}})$. Если матрица положительно определена $H(\overline{x^{(k)}}) > 0$, то перейти к пункту 7, иначе — к пункту 8.

7. Определить координаты точки $\overline{x^{(k+1)}} = \overline{x^{(k)}} - H^{-1}(\overline{x^{(k)}})\nabla f(\overline{x^{(k)}})$ и перейти к пункту 10.

8. Вычислить шаг h_k по формуле

(2.4), используя результаты вычислений пункта 3 и разностные формулы (2.5), (2.6).

9. Определить координаты точки $\overline{x^{(k+1)}} = \overline{x^{(k)}} - h_k \nabla f(\overline{x^{(k)}})$ по методу наискорейшего градиентного спуска.

10. Положить $k = k + 1$ и перейти к пункту 3.

Пример. Найти минимум целевой функции

$$f(\overline{x}) = x_1^2 - x_1 x_2 + 3x_2^2 - x_1$$

методом Ньютона с точностью ε = 0,1.

Решение. В качестве начального приближения возьмем точку $\overline{x^{(0)}} = \left(x_1^{(0)}, x_2^{(0)}\right)^T = (0,0)^T$.

Найдем градиент функции и матрицу Гессе в произвольной точке $\overline{x^{(k)}} = (x_1^{(k)}, x_2^{(k)})^T$:

$$\nabla f(\overline{x^{(k)}}) = \left(\frac{\partial f(\overline{x^{(k)}})}{\partial x_1}; \frac{\partial f(\overline{x^{(k)}})}{\partial x_2}\right)^T =$$

$$(2x_1^{(k)} - x_2^{(k)} - 1; \ \ -x_1^{(k)} + 6x_2^{(k)})^T.$$

$$H(\overline{x^{(k)}}) = \begin{bmatrix} \dfrac{\partial^2 f(\overline{x^{(k)}})}{\partial x_1^2} & \dfrac{\partial^2 f(\overline{x^{(k)}})}{\partial x_1 \partial x_2} \\ \dfrac{\partial^2 f(\overline{x^{(k)}})}{\partial x_2 \partial x_1} & \dfrac{\partial^2 f(\overline{x^{(k)}})}{\partial x_2^2} \end{bmatrix} = \begin{bmatrix} 2 & -1 \\ -1 & 6 \end{bmatrix}.$$

Итерация $\boldsymbol{k = 0}$. Определим в формуле $\overline{x^{(1)}} = \overline{x^{(0)}} + \overline{p^{(0)}}$ направление спуска $\overline{p^{(0)}}$. Для этого проведем анализ матрицы Гессе в точке $\overline{x^{(0)}}$: $H(\overline{x^{(0)}}) = \begin{bmatrix} 2 & -1 \\ -1 & 6 \end{bmatrix}$. Вычислим угловые миноры матрицы Гессе:

$$\Delta_1 = 2 > 0,\ \ \Delta_2 = \begin{vmatrix} 2 & -1 \\ -1 & 6 \end{vmatrix} = 11 > 0.$$

Так как знаки угловых миноров строго положительны, то согласно критерию Сильвестра матрица Гессе положительна определена. Следовательно, направление спуска определяем по формуле Ньютона

$$\overline{p^{(0)}} = -H^{-1}(\overline{x^{(0)}})\nabla f(\overline{x^{(0)}})\ .$$

Вычислим составляющие вектора $\overline{p^{(0)}}$: $\nabla f(\overline{x^{(0)}}) = (-1; 0)^T$,

$H^{-1}(\overline{x^{(0)}}) = \frac{1}{11}\begin{bmatrix} 6 & 1 \\ 1 & 2 \end{bmatrix}$.

Вычислим координаты точки $\overline{x^{(1)}}$

$$\overline{x^{(1)}} = \overline{x^{(0)}} - H^{-1}(\overline{x^{(0)}})\nabla f(\overline{x^{(0)}}) = \begin{pmatrix} 0 \\ 0 \end{pmatrix} - \frac{1}{11}\begin{bmatrix} 6 & 1 \\ 1 & 2 \end{bmatrix}\begin{pmatrix} -1 \\ 0 \end{pmatrix} = = \begin{pmatrix} \frac{6}{11} & \frac{1}{11} \end{pmatrix}^T.$$

Проверим условие окончания поиска. Для этого вычислим вектор градиента $\nabla f(\overline{x^{(1)}}) = (10^{-6}, -10^{-6})^T$ в точке $\overline{x^{(1)}} = \left(x_1^{(1)}, x_2^{(1)}\right)^T = \left(\frac{6}{11} \quad \frac{1}{11}\right)^T$. Так как норма вектора градиента $\left\|\nabla f(\overline{x^{(1)}})\right\| = 1{,}41 \cdot 10^{-6} < \varepsilon$, то расчет окончен.

Точка $\overline{x^{(1)}}$ является искомой точкой минимума $\overline{x^*} = \overline{x^{(1)}}$,$f(\overline{x^*}) = -0{,}273$.

Методы второго порядка.
Метод Ньютона–Рафсона

Метод Ньютона—Рафсона обеспечивает более устойчивую сходимость последовательности приближений, чем рассмотренный выше метод Ньютона. Последовательность приближений строится по модифицированной формуле

$$\overline{x^{(k+1)}} = \overline{x^{(k)}} + h_k \overline{p^{(k)}}\ ,$$

Где k — номер итерации (k = 0, 1,…),

$\overline{x^{(0)}}$ — начальное приближение,

h_k — величина шага,

$\overline{p^{(k)}} = -H^{-1}(\overline{x^{(k)}})\nabla f\,\overline{(x^{(k)})}$ — вектор направления спуска.

Здесь $H^{-1}\overline{(x^{(k)})}$ — обратная матрица для матрицы Гессе.

Направление спуска $\overline{p^{(k)}}$ ведет к убыванию целевой функции только при положительной определенности матрицы Гессе $H\overline{(x^{(k)})} > 0$. В тех итерациях, в которых матрица Гессе отрицательно определена, последовательность приближений к точке минимума строится по методу наискорейшего градиентного спуска. С этой целью проводится замена вектора направления спуска на анти-градиентное $\overline{p^{(k)}} = -h_k \nabla f\overline{(x^{(k)})}$.

Выбор величины шага h_k осуществляется из решения одномерной задачи оптимизации

$$\varphi(h_k) = f(\overline{x^{(k)}} + h_k \overline{p^{(k)}}) \to \min.\ h_k>0$$

Величина шага h_k может быть найдена в явном виде, аналогично формуле метода наискорейшего градиентного спуска

$$h_k = \frac{\nabla f(\overline{x^{(k)}}),\ \overline{p^{(k)}}}{H\left(\overline{x^{(k)}}\right)\overline{p^{(k)}},\ \ \overline{p^{(k)}}}$$

где $H\left(\overline{x^{(k)}}\right)$— матрица Гессе, вычисленная в точке $\overline{x^{(k)}}$.

Алгоритм метода Ньютона—Рафсона

1. Задать размерность задачи оптимизации n, координаты начальной точки $\overline{x^{(0)}} = (x_1^{(0)}, x_2^{(0)}, \dots, x_n^{(0)})$, точность поиска ε.
2. Положить счетчик числа итераций k = 0.
3. Определить направление вектора градиента целевой функции

$$f(\overline{x})\ \nabla f\left(\overline{x^{(k)}}\right) = \left(\frac{\partial f(\overline{x^{(k)}})}{\partial x_1}, \frac{\partial f(\overline{x^{(k)}})}{\partial x_2}, \dots, \frac{\partial f(\overline{x^{(k)}})}{\partial x_n}\right)$$

в точке $\overline{x^{(k)}} = (x_1^{(k)}, x_2^{(k)}, \dots, x_n^{(k)})$. Для вычисления координат вектора градиента использовать разностную формулу (2.3).

4. Проверить условие окончания поиска

$$\left\|\nabla \mathrm{f}\left(\overline{x^{(k)}}\right)\right\| = \sqrt{\sum_{i=1}^{n}\left(\frac{\partial \mathrm{f}(\overline{x^{(k)}})}{\partial x_1}\right)^2} \le \varepsilon.$$

Если условие выполнено, то расчет окончен $\overline{x^*} = \overline{x^{(k)}}$, иначе перейти к пункту 5.

5. Сформировать матрицу Гессе $\mathrm{H}\left(\overline{x^{(k)}}\right)$, используя разностные формулы вычисления вторых и смешанных производных.

6. Проверить положительную определенность матрицы Гессе $\mathrm{H}\left(\overline{x^{(k)}}\right)$. Если матрица положительно определена $\mathrm{H}\left(\overline{x^{(k)}}\right) > 0$, то перейти к пункту 7, иначе — положить $\mathrm{H}^{-1}(\overline{\mathrm{x}^{(k)}}) = \mathrm{E}$ и выполнить пункт 7.

7. Вычислить шаг h_k по формуле (2.4), используя результаты вычислений пункта 3 и разностные формулы (2.5), (2.6).

8. Определить координаты точки

$\overline{x^{(k+1)}} = \overline{x^{(k)}} - h_k H^{-1}(\overline{x^{(k)}})\nabla f(\overline{x^{(k)}})$, положить k = k + 1 и перейти к пункту 3.

Пример. Найти минимум целевой функции

$$f(\overline{x}) = x_1^2 - x_1x_2 + 3x_2^2 - x_1$$

Методом Ньютона—Рафсона с точностью ε = 0,1.

Решение. В качестве начального приближения возьмем точку $\overline{x^{(0)}} = (x_1^{(0)}, x_2^{(0)})^T = (0,0)^T$.

Найдем градиент функции и матрицу Гессе в произвольной точке

$\overline{x^{(k)}} = (x_1^{(k)}, x_2^{(k)})^T$:

$$\nabla \mathrm{f}\left(\overline{x^{(k)}}\right) = \left(\frac{\partial \mathrm{f}\left(\overline{x^{(k)}}\right)}{\partial x_1}, \frac{\partial \mathrm{f}\left(\overline{x^{(k)}}\right)}{\partial x_2}\right)^T = (2x_1^{(k)} - x_2^{(k)} - 1; \quad -x_1^{(k)} + 6x_2^{(k)})^T.$$

$$H\left(\overline{x^{(k)}}\right) = \begin{bmatrix} \dfrac{\partial^2 \mathrm{f}\left(\overline{x^{(k)}}\right)}{\partial x_1^2} & \dfrac{\partial^2 \mathrm{f}\left(\overline{x^{(k)}}\right)}{\partial x_1\,\partial x_2} \\ \dfrac{\partial^2 \mathrm{f}\left(\overline{x^{(k)}}\right)}{\partial x_2 \delta x_1} & \dfrac{\partial^2 \mathrm{f}\left(\overline{x^{(k)}}\right)}{\partial x_2^2} \end{bmatrix} = \begin{bmatrix} 2 & -1 \\ -1 & 6 \end{bmatrix}.$$

Итерация k = 0. Определим в формуле $\overline{x^{(1)}} = \overline{x^{(0)}} + h_0\overline{p^{(0)}}$ направление спуска $\overline{p^{(0)}}$. Для этого проведем анализ матрицы Гессе в точке $\overline{x^{(0)}}$:

$H\left(\overline{x^{(0)}}\right) = \begin{bmatrix} 2 & -1 \\ -1 & 6 \end{bmatrix}$. Вычислим угловые миноры матрицы Гессе:

$$\Delta_1 = 2 > 0, \Delta_2 = \begin{bmatrix} 2 & -1 \\ -1 & 6 \end{bmatrix} = 11 > 0.$$

Так как знаки угловых миноров строго положительны, то согласно критерию Сильвестра матрица Гессе положительно определена. Следовательно, направление спуска определяем по формуле Ньютона $\overline{p^{(0)}} = -H^{-1}\left(\overline{x^{(0)}}\right)\nabla f\left(\overline{x^{(0)}}\right)$.

Вычислим составляющие вектора $\overline{p^{(0)}}$: $\nabla f\left(\overline{x^{(0)}}\right) = (-1; 0)^T$,

$H^{-1}\left(\overline{x^{(0)}}\right) = \frac{1}{11}\begin{bmatrix} 6 & 1 \\ 1 & 2 \end{bmatrix}$.

Вычислим координаты точки $\overline{x^{(1)}}$

$$\overline{x^{(1)}} = \overline{x^{(0)}} - h_0 H^{-1}\left(\overline{x^{(0)}}\right)\nabla f\left(\overline{x^{(0)}}\right) = \begin{pmatrix} 0 \\ 0 \end{pmatrix} * \frac{h_0}{11}\begin{bmatrix} 6 & 1 \\ 1 & 2 \end{bmatrix} * \begin{pmatrix} -1 \\ 0 \end{pmatrix}$$

$$= \left(\frac{6}{11}h_0 ; \frac{1}{11}h_0\right)^T.$$

Определим h_0 из решения одномерной задачи

$$\varphi(h_0) = f\left(\overline{x^{(0)}} - h_0 H^{-1}\left(\overline{x^{(0)}}\right)\nabla f\left(\overline{x^{(0)}}\right)\right) \to min.$$

Получаем

$$\varphi(h_0) = \left(\frac{6}{11}\right)^2 h_0^2 - \frac{6}{11} * \frac{1}{11}h_0^2 + 3\left(\frac{1}{11}\right)^2 h_0^2 - \frac{6}{11}h_0 = \frac{33}{121}h_0^2 - \frac{6}{11}h_0.$$

Запишем необходимое условие экстремума

$$\frac{d\varphi(h_0)}{dh_0} = \frac{66}{121}h_0 - \frac{6}{11} = 0.$$

Откуда получаем $h_0 = 1$.

Поскольку

$$\frac{d\varphi^2(h_0)}{dh_0^2} = \frac{66}{121} > 0,$$

то найденное значение h_0 обеспечивает минимум функции $\varphi(h_0)$. По данной величине шага находим координаты точки $\overline{x^{(1)}}$:

$$\overline{x^{(1)}} = \left(\frac{6}{11}h_0 \quad \frac{1}{11}h_0\right)^T = \left(\frac{6}{11} \quad \frac{1}{11}\right)^T.$$

Проверим условие окончания поиска. Для этого вычислим вектор градиента $\nabla f\left(\overline{x^{(1)}}\right) = (10^{-6}, 10^{-6})^T$ в точке $\overline{x^{(1)}} = (x_1^{(1)}, x_2^{(1)})^T = \left(\frac{6}{11} \quad \frac{1}{11}\right)^T$. Так как норма вектора градиента $\left\|\nabla \mathrm{f}\left(\overline{x^{(1)}}\right)\right\| = 1{,}41 * 10^{-6} < \varepsilon$, то расчет окончен. Точка $\overline{x^{(1)}}$ является искомой точкой минимума $\overline{x^*} = \overline{x^{(1)}}, f\left(\overline{x^*}\right) = -0{,}273.$