Квазиньютоновские методы. Идея подходов

Квазиньютоновские методы объединяют в себе преимущества метода наискорейшего спуска и метода Ньютона. Методы этой группы обеспечивают скорость сходимости, свойственную методу Ньютона, и не требуют обращения матрицы Гессе. Направление оптимизации на i – ой *итерации алгоритма квазиньютоновских методов определяет соотношение:*

$$s^i = -A^i \nabla f\left(x^i\right)$$

где $\nabla f\left(x^i\right)$ *– градиент целевой функции в точке* x^i, A^i*– положительно определенная матрица направлений, которую на каждой итерации вычисляют по определенному методом алгоритму.*

Очередное приближение x^{i+1} *рассчитывают следующим образом:*

$$x^{i+1} = x^i + \lambda^i s^i$$

Рассмотрим способы определения матрицы направлений A_i.

Квазиньютоновские методы. Выбор матрицы направлений

Пусть $f(x)$ *– квадратичная целевая функция. Тогда из разложения в ряд Тейлора в окрестности точки* x^i *с исключением всех членов третьего и более высокого порядков имеем:*

$$f(x) \approx f(x^i) + \left(\nabla f(x^i), (x - x^i)\right) + \frac{1}{2}\left((x - x^i), H(x^i)(x - x^i)\right)$$

Если задать $x = x^{i+1}$ *и продифференцировать обе части соотношения ,получим:*

$$\nabla f(x^{i+1}) = \nabla f(x^i) + H(x^i)(x^{i+1} - x^i)$$

$$\nabla f(x^{i+1}) - \nabla f(x^i) = H(x^i)(x^{i+1} - x^i)$$

Тогда, домножив обе части на $H^{-1}(x^i)$*, получим:*

$$x^{i+1} - x^i = H^{-1}(x^i)\left[\nabla f(x^{i+1}) - \nabla f(x^i)\right]$$

Заметим, что матрица $H(x^i) = H$ *постоянная, поскольку* $f(x)$ *– квадратичная функция.*

Перейдем к аппроксимации обратной матрицы Гессе. С помощью информации, полученной на предыдущем i*-ом шаге, определим:*

$$H^{-1}(x^{k+1}) \approx \omega A^{k+1} = \omega\left(A^i + \Delta A^i\right)$$

где ω *– масштабный множитель,* A^i *– матрица, аппроксимирующая* $H^{-1}(x^i)$ *на предыдущем шаге,* ΔA^i *– поправочная матрица.*

Квазиньютоновские методы. Выбор матрицы направлений

Таким образом, имеем:

$$x^{i+1} - x^i = H^{-1}\left(x^i\right)\left[\nabla f\left(x^{i+1}\right) - \nabla f\left(x^i\right)\right]$$

$$H^{-1}\left(x^{k+1}\right) \approx \omega A^{k+1} = \omega\left(A^i + \Delta A^i\right)$$

Заметим, что способ выбора матрицы ΔA^i и определяет конкретный квазиньютоновский метод.
Пусть $g^i = \nabla f\left(x^{i+1}\right)$. Тогда из указанных соотношений получим:

$$\Delta x^i = \omega A^{i+1}\left[\nabla f\left(x^{i+1}\right) - \nabla f\left(x^i\right)\right] = \omega A^{i+1} \Delta g^i$$

$$\Downarrow$$

$$A^{i+1} \Delta g^i = \frac{1}{\omega} \Delta x^i$$

Так как $A^{i+1} = A^i + \Delta A^i$, составим и решим уравнение относительно ΔA^i:

$$\Delta A^i \Delta g^i = \frac{1}{\omega} \Delta x^i - A^i \Delta g^i$$

$$\Downarrow$$

$$\Delta A^i = \frac{1}{\omega} \frac{\Delta x^i}{} \Delta x^i - A^i \Delta g^i$$

Квазиньютоновские методы. Выбор матрицы направлений

Для уравнения:

$$\Delta A^i \Delta g^i = \frac{1}{\omega}\Delta x^i - A^i \Delta g^i$$

Можем показать существование следующего решения:

$$\Delta A^i = \frac{1}{\omega}\frac{y^T}{y^T \Delta g^i}\Delta x^i - A^i \frac{\Delta g^i z^T}{z^T \Delta g^i},$$

где $y, z \in \mathbb{R}^n$ *– произвольные векторы*

В частности, если задать $\omega = 1, y = z = \Delta x^i - A^i \Delta g^i$*, получим:*

$$\Delta A^i = \frac{\left(\Delta x^i - A^i \Delta g^i\right)\left(\Delta x^i - A^i \Delta g^i\right)^T}{\left(\Delta x^i - A^i \Delta g^i\right)^T}$$ *– Метод Бройдена*

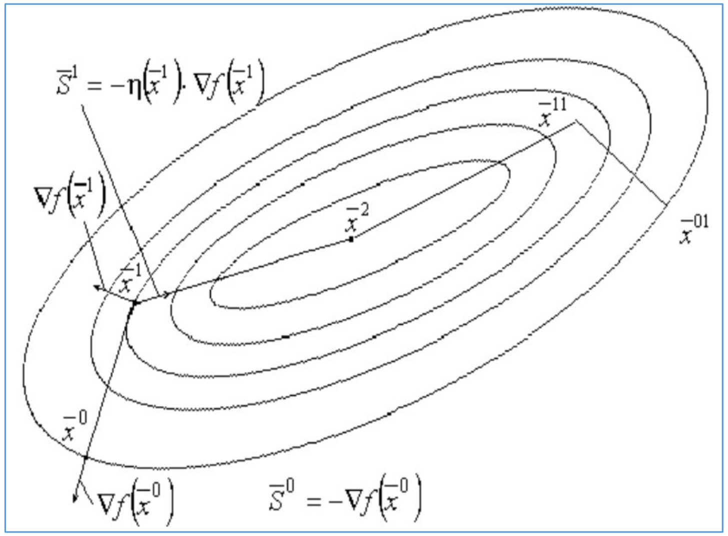
А если задать $y = \Delta x^i,\ z = A^i \Delta g^i$*:*

$$\Delta A^i = \frac{1}{\omega}\frac{\left(\Delta x^i\right)^T}{\left(\Delta x^i\right)^T \Delta g^i}\Delta x^i - A^i \frac{\Delta g^i \left(A^i \Delta g^i\right)^T}{\left(A^i \Delta g^i\right)^T \Delta g^i}$$ *– Метод Дэвидсона-Флетчера-Пауэлла*

Метод Бройдена

Broyden's Method

Идея подхода: *при построении релаксационной последовательности используем значения антиградиента и матрицы направлений, определенной способом Бройдена.*



Подготовительный этап:

Зададим начальное приближение x^0 *и точность расчета* ε: $|\nabla f(x^i)| < \varepsilon$. *Определим значение градиента* $\nabla f(x^0)$ *целевой функции в точк* x^0. *Зададим начальную положительно определенную матрицу* A^0.

Шаг 1:

Вычислим приближение к точке минимума x^{i+1}:

$$x^{i+1} = x^i - \lambda^i A^i\left(x^i\right)\nabla f\left(x^i\right),$$

где $\lambda^i = \min\limits_{\lambda>0} f\left(x^i - \lambda A^i\left(x^i\right)\nabla f\left(x^i\right)\right)$.

Определим очередное значение матрицы направлений:

$$A^{i+1} = A^i + \Delta A^i,$$

где

$$\Delta A^i = \frac{\left(\Delta x^i - A^i \Delta g^i\right)\left(\Delta x^i - A^i \Delta g^i\right)^T}{\left(\Delta x^i - A^i \Delta g^i\right)^T}$$

Шаг 2:

Проверим выполнение условия останова:

Если $|\nabla f(x^{i+1})| \geq \varepsilon$, *то* $i = i + 1$ *и возвращаемся к шагу 1.*

Иначе – завершаем поиск со значением $x^* = x^{i+1}$.

Метод Бройдена

Примечание

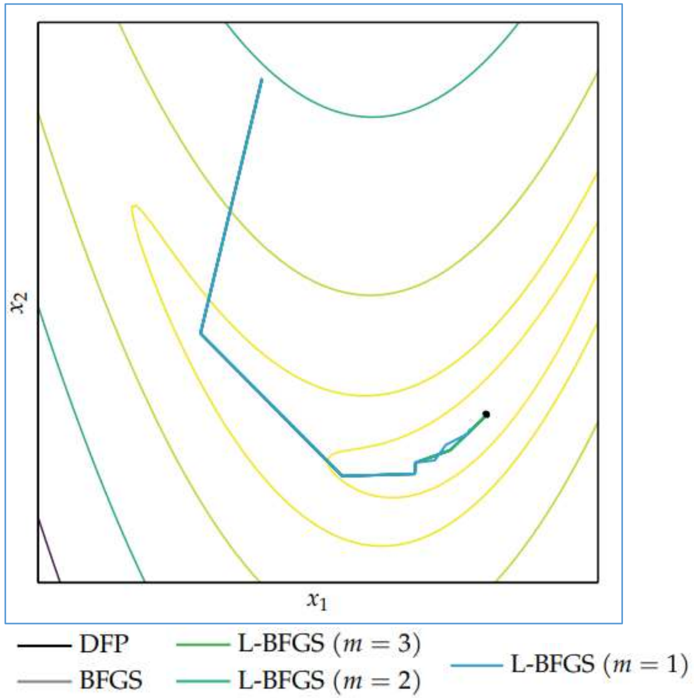
Примечание: *каждую* $i = n + 1$ *итерацию* n*-мерной оптимизации алгоритм Бройдена необходимо обновлять (начинать сначала), чтобы избежать следующих негативных последствий:*

- *Матрица* A^i *может потерять положительную определенность.*
- *Матрица* A^i *может стать неопределенной или сингулярной при совпадении (близком приближении) направлений поиска на текущей и предыдущей итерациях.*

Метод Давидона-Флетчера-Пауэлла

Davidon-Fletcher-Powell's Method (DFP)

Идея подхода: *при построении релаксационной последовательности используем значения антиградиента и матрицы направлений, определенной способом DFP.*



Подготовительный этап:

Зададим начальное приближение x^0 *и точность расчета* ε: $|\nabla f(x^i)| < \varepsilon$. *Определим значение градиента* $\nabla f(x^0)$ целевой функции в точке x^0. *Зададим начальную положительно определенную матрицу* A^0.

Шаг 1:

Вычислим приближение к точке минимума x^{i+1}:

$$x^{i+1} = x^i - \lambda^i A^i\left(x^i\right)\nabla f\left(x^i\right),$$

где $\lambda^i = \min\limits_{\lambda>0} f\left(x^i - \lambda^i A^i\left(x^i\right)\nabla f\left(x^i\right)\right)$.

Определим очередное значение матрицы направлений:

$$A^{i+1} = A^i + \Delta A^i,$$

где

$$\Delta A^i = \frac{1}{\omega}\frac{\left(\Delta x^i\right)^T}{\left(\Delta x^i\right)^T \Delta g^i}\Delta x^i - A^i\frac{\Delta g^i\left(A^i \Delta g^i\right)^T}{\left(A^i \Delta g^i\right)^T \Delta g^i}$$

Шаг 2:

Проверим выполнение условия останова:

Если $\left|\nabla f\left(x^{i+1}\right)\right| \geq \varepsilon$, *то* $i = i + 1$ *и возвращаемся к шагу 1.*

Иначе – завершаем поиск со значением $x^* = x^{i+1}$.

Метод Давидона-Флетчера-Пауэлла

Пример

Задание: *с помощью метода Дэвидона-Флетчера-Пауэлла найти точку минимума и минимальное значение целевой функции* $f(x_1, x_2) = x_1 - x_2 + 2x_1^2 + 2x_1x_2 + x_2^2$, *завершив вычисления при* $\left|\nabla f\left(x^i\right)\right| < \varepsilon = 0.05$

0) Положим начальное приближение $x^0 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$. *Определим значение градиента* $\nabla f(x^0)$ *и зададим начальное значение матрицы направлений* A^0*:*

Градиент:

$$\nabla f(x) = \begin{bmatrix} 1 + 4x_1 + 2x_2 \\ -1 + 2x_1 + 2x_2 \end{bmatrix}$$

$$\nabla f\left(x^0\right) = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \quad \text{и} \quad \left|\nabla f\left(x^0\right)\right| = \sqrt{1^2 + (-1)^2} > \varepsilon$$

Матрица направлений:

$$A^0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Метод Давидона-Флетчера-Пауэлла

Пример

Задание: *с помощью метода Дэвидона-Флетчера-Пауэлла найти точку минимума и минимальное значение целевой функции* $f(x_1, x_2) = x_1 - x_2 + 2x_1^2 + 2x_1x_2 + x_2^2$, *завершив вычисления при* $\left|\nabla f\left(x^i\right)\right| < \varepsilon = 0.05$

1) Найдем приближение к точке минимума x^1*:*

$$x^1 = x^0 - \lambda^0 A^0\left(x^0\right)\nabla f\left(x^0\right)$$

Определим длину шага спуска λ^0*. С помощью выбранного метода одномерной оптимизации найдем точку минимума функции* $f\left(x^0 - \lambda x^0(x^0)\nabla f(x^0)\right)$*:*

$$\lambda^0 = \min_{\lambda>0} f\left(x^0 - \lambda A^0\left(x^0\right)\nabla f\left(x^0\right)\right)$$

$$\lambda^0 = \min_{\lambda>0} f\left(\begin{pmatrix}0\\0\end{pmatrix} - \lambda\begin{pmatrix}1 & 0\\0 & 1\end{pmatrix}\begin{pmatrix}1\\-1\end{pmatrix}\right) = \min_{\lambda>0} f\left(\lambda_1^2 - 2\lambda\right)$$

$$\lambda^0 = 1$$

Тогда приближение к точке минимума:

$$x^1 = \begin{pmatrix}0\\0\end{pmatrix} - \lambda\begin{pmatrix}1 & 0\\0 & 1\end{pmatrix}\begin{pmatrix}1\\-1\end{pmatrix} = \begin{pmatrix}-1\\1\end{pmatrix}$$

Метод Давидона-Флетчера-Пауэлла

Пример

Задание: *с помощью метода Дэвидона-Флетчера-Пауэлла найти точку минимума и минимальное значение целевой функции* $f(x_1,x_2) = x_1 - x_2 + 2x_1^2 + 2x_1x_2 + x_2^2$*, завершив вычисления при* $\left|\nabla f\left(x^i\right)\right| < \varepsilon = 0.05$

$$x^1 = \begin{pmatrix} 0 \\ 0 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

Проверим выполнение условия останова:

$$\nabla f(x) = \begin{bmatrix} 1 + 4x_1 + 2x_2 \\ -1 + 2x_1 + 2x_2 \end{bmatrix}$$

$$\nabla f\left(x^1\right) = \begin{bmatrix} -1 \\ -1 \end{bmatrix} \quad \text{и} \quad \left|\nabla f\left(x^0\right)\right| = \sqrt{(-1)^2 + (-1)^2} > \varepsilon$$

Продолжаем поиск. Определим очередное значение матрицы направлений:

$$A^1 = A^0 + M^0 + N^0$$

$$M^0 = \lambda_1^* \frac{-A^0\nabla f\left(x^0\right)\left(-A^0\nabla f\left(x^0\right)\right)^T}{\left(-A^0\nabla f\left(x^0\right)\right)^T\left(\nabla f\left(x^1\right) - \nabla f\left(x^0\right)\right)}$$

$$N^0 = \lambda_1^* \frac{\left(-A^0\left(\nabla f\left(x^1\right) - \nabla f\left(x^0\right)\right)\right)\left(-A^0\left(\nabla f\left(x^1\right) - \nabla f\left(x^0\right)\right)\right)^T}{\left(\nabla f\left(x^1\right) - \nabla f\left(x^0\right)\right)^T\left(-A^0\left(\nabla f\left(x^1\right) - \nabla f\left(x^0\right)\right)\right)}$$

Метод Давидона-Флетчера-Пауэлла

Пример

Задание: *с помощью метода Дэвидона-Флетчера-Пауэлла найти точку минимума и минимальное значение целевой функции* $f(x_1, x_2) = x_1 - x_2 + 2x_1^2 + 2x_1x_2 + x_2^2$*, завершив вычисления при* $\left|\nabla f\left(x^i\right)\right| < \varepsilon = 0.05$

Вычислим значение матрицы M^0*:*

$$M^0 = \lambda_1^* \frac{-A^0\nabla f\left(x^0\right)\left(-A^0\nabla f\left(x^0\right)\right)^T}{\left(-A^0\nabla f\left(x^0\right)\right)^T\left(\nabla f\left(x^1\right)-\nabla f\left(x^0\right)\right)} =$$

$$= 1\left(\frac{1}{2}\right)\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} 0.5 & -0.5 \\ -0.5 & 0.5 \end{pmatrix}$$

$$g^0 = \nabla f\left(x^1\right) - \nabla f\left(x^0\right) = \begin{pmatrix} -1 \\ -1 \end{pmatrix} - \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} -2 \\ 0 \end{pmatrix}$$

$$-A^0\nabla f\left(x^0\right)\left(-A^0\nabla f\left(x^0\right)\right)^T = \begin{pmatrix} -1 \\ 1 \end{pmatrix}\begin{pmatrix} -1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$$

$$\left(-A^0\nabla f\left(x^0\right)\right)^T g^0 = \begin{pmatrix} -1 & 1 \end{pmatrix}\begin{pmatrix} -2 \\ 0 \end{pmatrix} = 2$$

Вычислим значение матрицы N^0*:*

$$N^0 = \lambda_1^* \frac{\left(-A^0\left(\nabla f\left(x^1\right)-\nabla f\left(x^0\right)\right)\right)\left(-A^0\left(\nabla f\left(x^1\right)-\nabla f\left(x^0\right)\right)\right)^T}{\left(\nabla f\left(x^1\right)-\nabla f\left(x^0\right)\right)^T\left(-A^0\left(\nabla f\left(x^1\right)-\nabla f\left(x^0\right)\right)\right)} =$$

$$= \frac{\begin{pmatrix} -2 \\ 0 \end{pmatrix}\begin{pmatrix} -2 & 0 \end{pmatrix}}{4} = -\frac{1}{4}\begin{pmatrix} 4 & 0 \\ 0 & 0 \end{pmatrix} = -\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

$$-A^0\left(\nabla f\left(x^1\right)-\nabla f\left(x^0\right)\right) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} -2 \\ 0 \end{pmatrix} = \begin{pmatrix} -2 \\ 0 \end{pmatrix}$$

$$\left(-A^0\left(\nabla f\left(x^1\right)-\nabla f\left(x^0\right)\right)\right)^T = \begin{pmatrix} -2 & 0 \end{pmatrix}$$

$$\left(\nabla f\left(x^1\right)-\nabla f\left(x^0\right)\right)^T\left(-A^0\left(\nabla f\left(x^1\right)-\nabla f\left(x^0\right)\right)\right) = \begin{pmatrix} -2 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} -2 \\ 0 \end{pmatrix} = 4$$

Метод Давидона-Флетчера-Пауэлла

Пример

Задание: *с помощью метода Дэвидона-Флетчера-Пауэлла найти точку минимума и минимальное значение целевой функции* $f(x_1, x_2) = x_1 - x_2 + 2x_1^2 + 2x_1x_2 + x_2^2$, *завершив вычисления при* $\left|\nabla f\left(x^i\right)\right| < \varepsilon = 0.05$

Определим очередное значение матрицы направлений:

$$A^1 = A^0 + M^0 + N^0$$

$$A^1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 0.5 & -0.5 \\ -0.5 & 0.5 \end{pmatrix} + \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0.5 & -0.5 \\ -0.5 & 1.5 \end{pmatrix}$$

Метод Давидона-Флетчера-Пауэлла

Пример

Задание: *с помощью метода Дэвидона-Флетчера-Пауэлла найти точку минимума и минимальное значение целевой функции* $f(x_1, x_2) = x_1 - x_2 + 2x_1^2 + 2x_1x_2 + x_2^2$, *завершив вычисления при* $\left|\nabla f\left(x^i\right)\right| < \varepsilon = 0.05$

2) Найдем приближение к точке минимума x^2*:*

$$x^2 = x^1 - \lambda^1 A^1\left(x^1\right)\nabla f\left(x^1\right)$$

Определим длину шага спуска λ^1*. С помощью выбранного метода одномерной оптимизации найдем точку минимума функции* $f\left(x^1 - \lambda x^1(x^1)\nabla f(x^1)\right)$*:*

$$\lambda^1 = \min_{\lambda>0} f\left(x^1 - \lambda A^1\left(x^1\right)\nabla f\left(x^1\right)\right)$$

$$\lambda^1 = \min_{\lambda>0} f\left(\begin{pmatrix} -1 \\ 1 \end{pmatrix} - \lambda\begin{pmatrix} 0.5 & -0.5 \\ -0.5 & 1.5 \end{pmatrix}\begin{pmatrix} -1 \\ -1 \end{pmatrix}\right) = \min_{\lambda>0} f\left(\lambda^2 - \lambda - 1\right)$$

$$\lambda^1 = 0.5$$

Тогда приближение к точке минимума:

$$x^2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix} - 0.5\begin{pmatrix} 0.5 & -0.5 \\ -0.5 & 1.5 \end{pmatrix}\begin{pmatrix} -1 \\ -1 \end{pmatrix} = \begin{pmatrix} -1 \\ 1.5 \end{pmatrix}$$

Метод Давидона-Флетчера-Пауэлла

Пример

Задание: *с помощью метода Дэвидона-Флетчера-Пауэлла найти точку минимума и минимальное значение целевой функции* $f(x_1, x_2) = x_1 - x_2 + 2x_1^2 + 2x_1x_2 + x_2^2$, *завершив вычисления при* $\left|\nabla f\left(x^i\right)\right| < \varepsilon = 0.05$

$$x^2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix} - 0.5 \begin{pmatrix} 0.5 & -0.5 \\ -0.5 & 1.5 \end{pmatrix} \begin{pmatrix} -1 \\ -1 \end{pmatrix} = \begin{pmatrix} -1 \\ 1.5 \end{pmatrix}$$

Проверим выполнение условия останова:

$$\nabla f(x) = \begin{bmatrix} 1 + 4x_1 + 2x_2 \\ -1 + 2x_1 + 2x_2 \end{bmatrix}$$

$$\nabla f\left(x^2\right) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

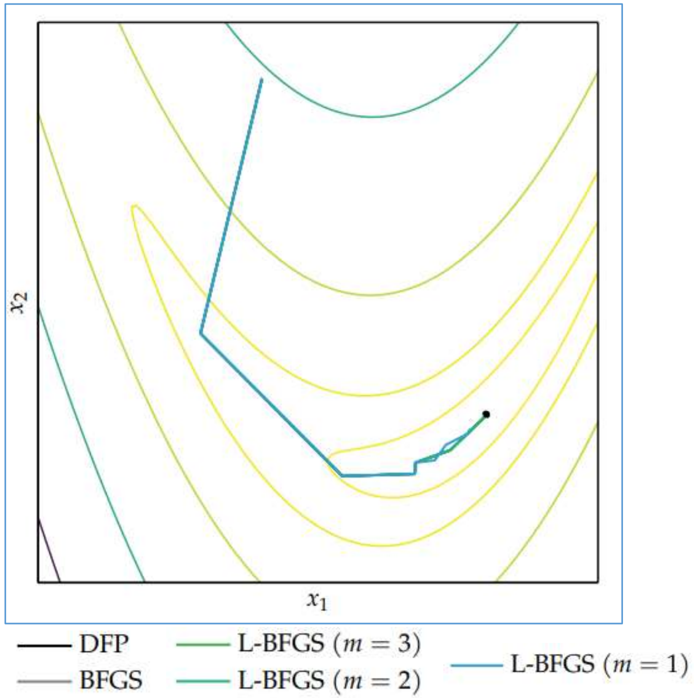
$$\left|\nabla f\left(x^0\right)\right| = 0 < \varepsilon \Rightarrow$$ *завершаем поиск со значениями:*

$$x^* = \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \quad f\left(x^*\right) = -1$$

Метод Бройдена-Флетчера-Гольдфарба-Шанно

Broyden-Fletcher-Goldfarb-Shanno's Method (BFGS)

Идея подхода: *улучшенный метод DFP с особым способом определения матрицы направлений.*



Подготовительный этап:

Зададим начальное приближение x^0 *и точность расчета* ε: $\left|\nabla f(x^i)\right| < \varepsilon$. *Определим значение градиента* $\nabla f(x^0)$ целевой функции в точке x^0. *Зададим начальную положительно определенную матрицу* A^0.

Шаг 1:

Вычислим приближение к точке минимума x^{i+1}:

$$x^{i+1} = x^i - \lambda^i A^i\left(x^i\right)\nabla f\left(x^i\right),$$

где $\lambda^i = \min\limits_{\lambda>0} f\left(x^i - \lambda^i A^i\left(x^i\right)\nabla f\left(x^i\right)\right)$.

Определим очередное значение матрицы направлений:

$$A^{i+1} = A^i + \Delta A^i,$$

где

$$\Delta A^i = \left(1 + \frac{\left(\Delta g^i\right)^T A^{i-1}\Delta g^i}{\left(\Delta x^i\right)^T \Delta g^i}\right)\frac{\Delta x^i\left(\Delta x^i\right)^T}{\left(\Delta x^i\right)^T \Delta g^i} - \frac{\Delta x^i\left(\Delta g^i\right)^T A^{i-1}}{\left(\Delta x^i\right)^T \Delta g^i} - \frac{A^{i-1}\Delta g^i\left(\Delta x^i\right)^T}{\left(\Delta x^i\right)^T \Delta g^i}$$

Шаг 2:

Проверим выполнение условия останова:

Если $\left|\nabla f(x^{i+1})\right| \geq \varepsilon$, *то* $i = i + 1$ *и возвращаемся к шагу 1.*

Иначе – завершаем поиск со значением $x^* = x^{i+1}$.

Метод Бройдена-Флетчера-Гольдфарба-Шанно

Пример

Задание: *с помощью метода Бройдена-Флетчера-Гольдфарба-Шанно найти точку минимума и минимальное значение целевой функции* $f(x_1, x_2) = x_1 - x_2 + 2x_1^2 + 2x_1x_2 + x_2^2$*, завершив вычисления при* $|\nabla f(x^i)| < \varepsilon = 0.05$

0) Положим начальное приближение $x^0 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$*. Определим значение градиента* $\nabla f(x^0)$ *и зададим начальное значение матрицы направлений* A^0*:*

Градиент:

$$\nabla f(x) = \begin{bmatrix} 1 + 4x_1 + 2x_2 \\ -1 + 2x_1 + 2x_2 \end{bmatrix}$$

$$\nabla f(x^0) = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \quad \text{и} \quad |\nabla f(x^0)| = \sqrt{1^2 + (-1)^2} > \varepsilon$$

Матрица направлений:

$$A^0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Метод Бройдена-Флетчера-Гольдфарба-Шанно

Пример

Задание: *с помощью метода Бройдена-Флетчера-Гольдфарба-Шанно найти точку минимума и минимальное значение целевой функции* $f(x_1, x_2) = x_1 - x_2 + 2x_1^2 + 2x_1x_2 + x_2^2$, *завершив вычисления при* $\left|\nabla f\left(x^i\right)\right| < \varepsilon = 0.05$

1) Найдем приближение к точке минимума x^1*:*

$$x^1 = x^0 - \lambda^0 A^0\left(x^0\right)\nabla f\left(x^0\right)$$

Определим длину шага спуска λ^0*. С помощью выбранного метода одномерной оптимизации найдем точку минимума функции* $f\left(x^0 - \lambda x^0(x^0)\nabla f(x^0)\right)$*:*

$$\lambda^0 = \min_{\lambda>0} f\left(x^0 - \lambda A^0\left(x^0\right)\nabla f\left(x^0\right)\right)$$

$$\lambda^0 = \min_{\lambda>0} f\left(\begin{pmatrix}0\\0\end{pmatrix} - \lambda\begin{pmatrix}1 & 0\\0 & 1\end{pmatrix}\begin{pmatrix}1\\-1\end{pmatrix}\right) = \min_{\lambda>0} f\left(\lambda_1^2 - 2\lambda\right)$$

$$\lambda^0 = 1$$

Тогда приближение к точке минимума:

$$x^1 = \begin{pmatrix}0\\0\end{pmatrix} - \lambda\begin{pmatrix}1 & 0\\0 & 1\end{pmatrix}\begin{pmatrix}1\\-1\end{pmatrix} = \begin{pmatrix}-1\\1\end{pmatrix}$$

Метод Бройдена-Флетчера-Гольдфарба-Шанно

Пример

Задание: *с помощью метода Бройдена-Флетчера-Гольдфарба-Шанно найти точку минимума и минимальное значение целевой функции* $f(x_1, x_2) = x_1 - x_2 + 2x_1^2 + 2x_1x_2 + x_2^2$*, завершив вычисления при* $\left|\nabla f\left(x^i\right)\right| < \varepsilon = 0.05$

$$x^1 = \begin{pmatrix} 0 \\ 0 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

Проверим выполнение условия останова:

$$\nabla f(x) = \begin{bmatrix} 1 + 4x_1 + 2x_2 \\ -1 + 2x_1 + 2x_2 \end{bmatrix}$$

$$\nabla f\left(x^1\right) = \begin{bmatrix} -1 \\ -1 \end{bmatrix} \quad \text{и} \quad \left|\nabla f\left(x^0\right)\right| = \sqrt{(-1)^2 + (-1)^2} > \varepsilon$$

Продолжаем поиск. Определим очередное значение матрицы направлений:

$$A^1 = A^0 + \Delta A^0,$$

$$\Delta A^{i-1} = \left(1 + \frac{\left(\Delta g^i\right)^T A^{i-1} \Delta g^i}{\left(\Delta x^i\right)^T \Delta g^i}\right) \frac{\Delta x^i \left(\Delta x^i\right)^T}{\left(\Delta x^i\right)^T \Delta g^i} - \frac{\Delta x^i \left(\Delta g^i\right)^T A^{i-1}}{\left(\Delta x^i\right)^T \Delta g^i} - \frac{A^{i-1} \Delta g^i \left(\Delta x^i\right)^T}{\left(\Delta x^i\right)^T \Delta g^i}$$

Метод Бройдена-Флетчера-Гольдфарба-Шанно

Пример

Задание: *с помощью метода Бройдена-Флетчера-Гольдфарба-Шанно найти точку минимума и минимальное значение целевой функции* $f(x_1, x_2) = x_1 - x_2 + 2x_1^2 + 2x_1x_2 + x_2^2$, *завершив вычисления при* $\left|\nabla f\left(x^i\right)\right| < \varepsilon = 0.05$

$$\Delta A^{i-1} = \left(1 + \frac{\left(\Delta g^i\right)^T A^{i-1} \Delta g^i}{\left(\Delta x^i\right)^T \Delta g^i}\right) \frac{\Delta x^i \left(\Delta x^i\right)^T}{\left(\Delta x^i\right)^T \Delta g^i} - \frac{\Delta x^i \left(\Delta g^i\right)^T A^{i-1}}{\left(\Delta x^i\right)^T \Delta g^i} - \frac{A^{i-1} \Delta g^i \left(\Delta x^i\right)^T}{\left(\Delta x^i\right)^T \Delta g^i}$$

Последовательно вычислим:

$$\Delta g^1 = \nabla f\left(x^1\right) - \nabla f\left(x^0\right) = \begin{pmatrix} -1 \\ -1 \end{pmatrix} - \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} -2 \\ 0 \end{pmatrix}$$

$$\Delta x^1 = x^1 - x^0 = \lambda^0 A^0 \nabla f\left(x^0\right) = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

Тогда

$$\Delta A^0 = \begin{pmatrix} 1.5 & -1.5 \\ -1.5 & 1.5 \end{pmatrix} - \begin{pmatrix} 1 & 0 \\ -1 & 0 \end{pmatrix} - \begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix}$$

$$A^1 = A^0 + \Delta A^0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 1.5 & -1.5 \\ -1.5 & 1.5 \end{pmatrix} - \begin{pmatrix} 1 & 0 \\ -1 & 0 \end{pmatrix} - \begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0.5 & -0.5 \\ -0.5 & 2.5 \end{pmatrix}$$

Метод Бройдена-Флетчера-Гольдфарба-Шанно

Пример

Задание: *с помощью метода Бройдена-Флетчера-Гольдфарба-Шанно найти точку минимума и минимальное значение целевой функции* $f(x_1, x_2) = x_1 - x_2 + 2x_1^2 + 2x_1x_2 + x_2^2$*, завершив вычисления при* $\left|\nabla f\left(x^i\right)\right| < \varepsilon = 0.05$

2) Найдем приближение к точке минимума x^2*:*

$$x^2 = x^1 - \lambda^1 A^1\left(x^1\right)\nabla f\left(x^1\right)$$

Определим длину шага спуска λ^1*. С помощью выбранного метода одномерной оптимизации найдем точку минимума функции* $f\left(x^1 - \lambda x^1(x^1)\nabla f(x^1)\right)$*:*

$$\lambda^1 = \min_{\lambda>0} f\left(x^1 - \lambda A^1\left(x^1\right)\nabla f\left(x^1\right)\right)$$

$$\lambda^1 = \min_{\lambda>0} f\left(\begin{pmatrix}-1\\1\end{pmatrix} - \lambda\begin{pmatrix}0.5 & -0.5\\-0.5 & 2.5\end{pmatrix}\begin{pmatrix}-1\\-1\end{pmatrix}\right) = \min_{\lambda>0} f\left(4\lambda^2 - 2\lambda - 1\right)$$

$$\lambda^1 = 0.25$$

Тогда приближение к точке минимума:

$$x^2 = \begin{pmatrix}-1\\1\end{pmatrix} - 0.25\begin{pmatrix}0.5 & -0.5\\-0.5 & 2.5\end{pmatrix}\begin{pmatrix}-1\\-1\end{pmatrix} = \begin{pmatrix}-1\\1.5\end{pmatrix}$$

Метод Бройдена-Флетчера-Гольдфарба-Шанно

Пример

Задание: *с помощью метода Бройдена-Флетчера-Гольдфарба-Шанно найти точку минимума и минимальное значение целевой функции* $f(x_1, x_2) = x_1 - x_2 + 2x_1^2 + 2x_1x_2 + x_2^2$, *завершив вычисления при* $\left|\nabla f\left(x^i\right)\right| < \varepsilon = 0.05$

$$x^2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix} - 0.25 \begin{pmatrix} 0.5 & -0.5 \\ -0.5 & 2.5 \end{pmatrix} \begin{pmatrix} -1 \\ -1 \end{pmatrix} = \begin{pmatrix} -1 \\ 1.5 \end{pmatrix}$$

Проверим выполнение условия останова:

$$\nabla f(x) = \begin{bmatrix} 1 + 4x_1 + 2x_2 \\ -1 + 2x_1 + 2x_2 \end{bmatrix}$$

$$\nabla f\left(x^2\right) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$\left|\nabla f\left(x^0\right)\right| = 0 < \varepsilon \Rightarrow$ *завершаем поиск со значениями:*

$$x^* = \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \quad f\left(x^*\right) = -1$$