Метод Пауэлла

Оглавление

Авторы: Бабичев Денис, Бессонов Трофим.

С помощью метода Пауэлла определяется местонахождение минимума некоторой целевой квадратичной функции f(x), при H(f)>0, где H(f) – матрица Гессе - матрица вторых частных производных.

$$H(f) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1\,\partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1\,\partial x_n} \\ \frac{\partial^2 f}{\partial x_2\,\partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2\,\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n\,\partial x_1} & \frac{\partial^2 f}{\partial x_n\,\partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{bmatrix}$$

x1,x2 – переменные заданной квадратичной функции.

Путем проведения последовательных одномерных поисков (например, метод Золотого сечения), начиная с точки X_0^k (произвольное, задаем сами), вдоль системы <u>сопряженных направлений</u>

Два направления поискаs_i и s_j называются сопряженными, если

$$(\mathbf{s}_j)^T \mathbf{Q} \mathbf{s}_i = 0, \qquad i \neq j,$$

$$(\mathbf{s}_j)^T \mathbf{Q} \mathbf{s}_i \geqslant 0, \qquad i = j,$$

$\hat{s}_0$ - направление поиска экстремума.

Направление, сопряженное к $\hat{s}_0$ (например, $\hat{s}_1$) находится следующим образом: $[\hat{s}_1\ ]\,\hat{}\,T * H * \hat{s}_0 = 0$, или в общем виде:

$\hat{s}_j\ ]\,\hat{}\,T * H * \hat{s}_j\ = 0;\ 0 <= i != j <= n-1$

Идея алгоритма заключается в том, что если на данном этапе поиска определяется минимум квадратичной функции вдоль каждого из сопряженных направлений и если затем в каждом направлении делается некоторый шаг, то полное перемещение от начала до конечного шага сопряжено ко всем поднаправлениям поиска.

Переход из точки $\mathbf{x}_0^{(k)}$ в точку $\mathbf{x}_m^{(k)}$ определяется формулой

$$\mathbf{x}_m^{(k)} = \mathbf{x}_0^{(k)} + \sum_{i=0}^{m-1} \lambda_i^{(k)} \mathbf{s}_i^{(k)}, \quad i = 1, \ldots, m-1.$$

λ_i^k –– длина шага, определяется по формуле(возьмем k= 0):

$$\lambda^{(0)} = -\frac{\nabla^T f(\mathbf{x}^{(0)}) \widehat{\mathbf{s}}^{(0)}}{(\widehat{\mathbf{s}}^{(0)})^T \nabla^2 f(\mathbf{x}^{(0)}) \widehat{\mathbf{s}}^{(0)}}.$$

Поиск в общем случае

В общем случае на к-ом этапе метода Пауэлла используются n линейно независимых направлений поиска; при этом поиск начинается в некоторой точке $x_0^{(k)}$ = $x_{n+1}^{(k-1)}$ и проводится следующим образом

Шаг 1. Начиная из $x_0^{(k)}$, с помощью какого-либо одномерного поиска определяется $\lambda_i^{(k)}$ так, чтобы f($x_0^{(k)} + \lambda_1 s_1^{(k)}$) принимала минимальное значение, и полагается
$x_1^{(k)}$= $x_0^{(k)} + \lambda_1^{(k)} s_1^{(k)}$
Начиная из $x_1^{(k)}$ определяется $\lambda_2^{(k)}$так, чтобы f($x_1^{(k)} + \lambda_2 s_2^{(k)}$) обращалась в минимум, и полагается $x_2^{(k)}$= $x_1^{(k)} + \lambda_2^{(k)} s_2^{(k)}$. Поиск продолжается последовательно в каждом направлении, всегда начиная из самой последней точки последовательности, пока не будут определены все $\lambda_i^{(k)}$ i=1,….n
Величина $\lambda_i^{(k)}$ используется на шаге 4.

Шаг 2
После минимизации f(x) в каждом из n направлений, как на шаге 1, проводится один дополнительный шаг величиной ($x_n^{(k)} - x_0^{(k)}$) соответствующий полному перемещению на к-м этапе. Затем проводится тест (шаг 3), чтобы убедиться, уменьшается ли определитель матрицы направлений поиска путем включения нового направления и отбрасывания старого.

Шаг 3. Обозначим наибольшее уменьшение $f(\mathbf{x})$ в каком-либо направлении поиска на k-м этапе через

$$\Delta^{(k)} = \max_{i=1,\ldots,n} \{f(\mathbf{x}_{i-1}^{(k)}) - f(\mathbf{x}_i^{(k)})\}.$$

Направление поиска, соответствующее этому максимальному изменению $f(\mathbf{x})$, обозначим через $\mathbf{s}_m^{(k)}$. Чтобы сделать обозначения более компактными, положим $f_1 = f(\mathbf{x}_0^{(k)})$, $f_2 = f(\mathbf{x}_n^{(k)})$ и $f_3 = f(2\mathbf{x}_n^{(k)} - \mathbf{x}_0^{(k)})$, где $\mathbf{x}_0^{(k)} = \mathbf{x}_n^{(k-1)}$ и $\mathbf{x}_n^{(k)} = \mathbf{x}_{n-1}^{(k)} + \lambda_n^{(k)}\mathbf{s}_n^{(k)} = \mathbf{x}_0^{(k)} + \sum_{i=1}^{n} \lambda_i^{(k)}\mathbf{s}_i^{(k)}$. Тогда если $f_3 \geqslant f_1$ и (или) $(f_1 - 2f_2 + f_3)(f_1 - f_2 - \Delta^{(k)})^2 \geqslant 0{,}5\Delta^{(k)}(f_1 - f_3)^2$, то следует использовать на $(k+1)$-м этапе те же направления $\mathbf{s}_1^{(k)}, \ldots, \mathbf{s}_n^{(k)}$, что и на k-м этапе, т. е. $\mathbf{s}_i^{(k+1)} = \mathbf{s}_i^{(k)}$ для $i = 1, \ldots, n$, и начать поиск из точки $\mathbf{x}_0^{(k+1)} = \mathbf{x}_n^{(k)}$ [или из $\mathbf{x}_0^{(k+1)} = 2\mathbf{x}_n^{(k)} - \mathbf{x}_0^{(k)} = \mathbf{x}_{n+1}^{(k)}$ в зависимости от того, в какой точке $\mathbf{x}$ функция $f(\mathbf{x})$ принимает наименьшее значение].

Шаг 4. Если тест на шаге 3 не удовлетворен, то ищется минимум f(x) в направлени вектора $s^{(k)}$ проведенного из $x_0^{(k)}$ в $x_n^{(k)}$; точка этого минимума берется в качестве начальной для следующего этапа. На (k+1) этапе будут использоваться следующие направления:

$$[\mathbf{s}_1^{(k+1)}\mathbf{s}_2^{(k+1)} \ldots \mathbf{s}_n^{(k+1)}] = [\mathbf{s}_1^{(k)}\mathbf{s}_2^{(k)} \ldots \mathbf{s}_{m-1}^{(k)}\mathbf{s}_{m+1}^{(k)} \ldots \mathbf{s}_n^{(k)}\mathbf{s}^{(k)}].$$

Шаг 5. Критерий удовлетворительной сходимости для метода Пауэлла, используемый для определения момента окончания поиска в конце любого этапа, состоит в том, что изменение по каждой независимой переменной должно быть меньше, чем заданная точность ε_i, $i = 1, \ldots, n$, или $\| \mathbf{x}_n^{(k)} - \mathbf{x}_0^{(k)} \| \leqslant 0{,}1\ \varepsilon$.

Пример

Рассмотрим вычислительную процедура алгоритма для квадратичной функции $f(x)= 4(x_1-5)^2+(x_2-6)^2$:

1) Задается начальный вектор $X^0= [8;9]^T$
2) Задается направление поиска, как правило это

$p^1 = [1;0]^T$
$p^2 = [0;1]^T$

3) Вычисляется градиент grad f(x) (заданной квадратичной функции) $\nabla f(X) = \begin{pmatrix} 8x_1-40 \\ 2x_2-12 \end{pmatrix}$

4)Критерий остановки - $|\nabla f(X_0)| < \varepsilon$, ε задана

Далее выполняется цикл до тех пор, пока не будет выполнено условие.

Итерация первая:

1) Вычисляем значение градиента в X^0, grad $f(X^0)$, получим вектор значений. $\nabla f(X_0) = \begin{pmatrix} 24 \\ 6 \end{pmatrix}$

2) Проверка условия сходимости

$|\nabla f(X_0)| = \sqrt{24^2+6^2} = 24.739>0.1$

3) Вычислим значение функции в X^0 , $f(X^0)$, Получим вектор значений. $f(X_0) = 45$.

4) Сделаем шаг вдоль направления поиска $p^2 = [0;1]^T$

$X_1 = X_0 + hp^2 = \begin{pmatrix} 8 \\ 9 \end{pmatrix} + h\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 8 \\ h+9 \end{pmatrix}$, где h – длина шага.

$f(X_1) = 4(8-5)^2+((h+9)-6)^2 \rightarrow \min$
$f(X_1) = h^2+6h+45 \rightarrow \min$

5) Найдем такой шаг h, чтобы целевая функция достигала минимума вдоль этого направления. Из необходимого условия существования экстремума функции ($f'(x_1)=0$):

$2h+6 = 0$. Получим шаг: $h = -3$. В коде нужно реализовать любой одномерный поиск (Опять же, метод золотого сечения)

Выполнение этого шага приведет в точку:

$$X_1 = \begin{pmatrix} 8 \\ 9 \end{pmatrix} -3 \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 8 \\ 6 \end{pmatrix}$$

6) Сделаем шаг вдоль другого направления поиска $p^1 = [1;0]^T$

$$X_2 = X_1 + hp^1 = \begin{pmatrix} 8 \\ 6 \end{pmatrix} + h \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} h+8 \\ 6 \end{pmatrix}$$

$f(X_2) = 4((h+8)-5)^2+((6)-6)^2 \rightarrow \min$

$f(X_2) = 4h^2+24h+36 \rightarrow \min$

Найдем такой шаг h, чтобы целевая функция достигала минимума вдоль этого направления. Из необходимого условия существования экстремума функции ($f'(x_2)=0$):

$8h+24 = 0$. Получим шаг: $h = -3$

Выполнение этого шага приведет в точку:

$$X_2 = \begin{pmatrix} 8 \\ 6 \end{pmatrix} -3 \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 5 \\ 6 \end{pmatrix}$$

7) Повторно сделаем шаг вдоль направления поиска $p^2 = [0;1]^T$

$$X_3 = X_2 + hp^2 = \begin{pmatrix} 5 \\ 6 \end{pmatrix} + h\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 5 \\ h+6 \end{pmatrix}$$

$f(X_3) = 4(5-5)^2+((h+6)-6)^2 \to \min$

$f(X_3) = h^2 \to \min$

Найдем такой шаг h, чтобы целевая функция достигала минимума вдоль этого направления. Из необходимого условия существования экстремума функции ($f'(x_3)=0$):

$2h = 0$. Получим шаг: $h = 0$

Выполнение этого шага приведет в точку: $x3 = (5,6)$.

Так как шаг равен нулю, выбираем сопряженное направление.

8) Выбираем сопряженное направление: $p^2 = x^3 - x^1$ $p^2 = [5;6]^T - [8;6]^T = [-3;0]^T$

Итерация вторая:

1) $$\nabla f(X_3) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

   Проверим критерий остановки:

   $|\nabla f(X_3)| < \varepsilon$

   $|\nabla f(X_3)| = \sqrt{0^2+0^2} = 0 < 0.1$

2) Вычислим значение функции в начальной точке $f(X_3) = 0$.

   $X = [5;6]^T$ – минимум функции.

В общем случае, условие может выполниться далеко не на второй итерации. Если бы

$\nabla f(X_3)$

Не был нулевым, то вычисления происходили бы точно так же, только с уже другими сопряженными направлениям.