Метод циклического покоординатного спуска

Методы покоординатного спуска простейшего типа заключаются в изменении каждый раз одной переменной, тогда как другие остаются постоянными. Опишем этот метод для задачи $min\{f(X) \mid X \in E^n\}$.

Пусть X^0 - некоторое начальное приближение, а α_0 - некоторое положительное число, являющееся параметром алгоритма. Допустим, что уже известны точка $X^k \in E^n$ и число $\alpha_k > 0$ при некотором $k > 0$. Обозначим $e^j = (0, \dots, 0, 1, 0, \dots, 0)$ - единичный координатный вектор, у которого j-я координата равна 1, остальные равны нулю, $j = 1, \dots, n$. Положим

$$S^k = e^{j_k},\ j_k = k - n\left[\frac{k}{n}\right] + 1$$

где $[\frac{k}{n}]$ - целая часть числа $\frac{k}{n}$. Условие $S^k = e^{j_k},\ j_k = k - n\left[\frac{k}{n}\right] + 1$ обеспечивает циклический перебор координатных векторов $e^1, \dots, e^n$, т.е.

$$S^0 = e^1, \dots, S^{n-1} = e^n, S^n = e^1, \dots, S^{2n-1} = e^n, S^{2n} = e^1 \dots$$

Вычислим значение функции $f(X)$ в точке $X = X^k + \alpha_k S^k$ и проверим неравенство

$$f(X^k + \alpha_k S^k) < f(X^k)$$

Если оно выполняется, то полагаем

$$X^{k+1} = X^k + \alpha_k S^k, \alpha_{k+1} = \alpha_k$$

В случае, если условие $f(X^k + \alpha_k S^k) < f(X^k)$ не выполняется, вычисляем значение функции $f(X)$ в точке $X = X^k - \alpha_k S^k$ и проверяем неравенство :

$$f(X^k - \alpha_k S^k) < f(X^k)$$

При выполнении условия $f(X^k - \alpha_k S^k) < f(X^k)$ полагаем

$$X^{k+1} = X^k - \alpha_k S^k,\ \alpha_{k+1} = \alpha_k$$

Будем считать $(k+1)$-й этап удачным, если выполнилось хотя бы одно из условий $f(X^k + \alpha_k S^k) < f(X^k)$ или $f(X^k - \alpha_k S^k) < f(X^k)$. Если же ни одно из этих условий не выполнено, считаем $(k+1)$-й этап неудачным и полагаем

$$X^{k+1} = X^k, \alpha_{k+1} = \begin{cases} \lambda\alpha_k, & \text{при } j_k = n, x^k = x^{k-n+1} \\ \alpha_k, & \text{при } j_k \neq n \text{ или } x^k \neq x^{k-n+1}, \text{ или } 0 \leq k \leq n-1 \end{cases}$$

где $\lambda \in (0;1)$ - фиксированное число, являющееся параметром алгоритма. Данное условие означает, что если за один цикл из n этапов при переборе направлений всех координатных векторов $e^1, \ldots, e^n$ с шагом α_k реализовался хотя бы один удачный этап, то длина шага α_k не дробится и сохраняется на протяжении, по крайней мере, следующего цикла из n этапов. Если же среди последних n этапов ни одного удачного не оказалось, то шаг α_k дробится.

Скорость сходимости данного метода невысока. Несмотря на это, метод покоординатного спуска широко применяется на практике благодаря простоте реализации. Заметим, что такой метод работает плохо, если в выражение минимизируемой функции входят произведения $x_i x_j$, т.е. если имеет место взаимодействие между $x_i, i = 1, \ldots, n$.

Указанного недостатка можно избежать с помощью следующей модификации метода покоординатного спуска, известной под названием метода Зейделя.

2.1.4. Метод Гаусса-Зейделя

Этот метод заключается в последовательной минимизации $f(X)$ по направлению каждого из координатных векторов e^j, $j = 1, \ldots, n$ всегда начиная из самой последней точки построенной последовательности. После завершения минимизации по направлению последнего координатного вектора e^n цикл, называемый внешней итерацией, повторяется до тех пор, пока не выполнится одно из возможных условий

окончания поиска:

$\left|f(X^k) - f(X^{k-n})\right| < \varepsilon$ или $\left\|X^k - X^{k-n}\right\| < \varepsilon$,

где $\varepsilon > 0$ - заданный параметр точности.

Для приближенного решения вспомогательной задачи одномерной минимизации

$$\min\{f(X^k - \alpha S^k) \mid \alpha \in R\}$$

на каждом внутреннем шаге внешней итерации целесообразно использовать изученный ранее метод поразрядного поиска. Опишем алгоритм метода Зейделя.

Шаг 0. Задать параметр точности $\varepsilon > 0$, выбрать точку начального приближения $X^0 \in E^n$, вычислить значение функции $f(X^0)$, положить $k = 0$.

Шаг 1. Положить $j = k - n\left[\frac{k}{n}\right] + 1, S^k = e^j$, где $\left[\frac{k}{n}\right]$ - целая часть числа k/n.

Решить задачу одномерного поиска, т е. определить оптимальную величину шага $\alpha_k = \arg\min\{f(X^k + \alpha S^k) | \alpha \in R\}$ Найти новую точку последовательности $X^{k+1} = X^k + \alpha_k S^k$ и вычислить значение функции $f(X^{k+1})$.

Шаг 2. Если $j < n$, то положить $k = k + 1$ и перейти к шагу 1, иначе - перейти к шагу 3.

Шаг 3. Проверить условие достижения заданной точности $\left|f(X^k) - f(X^{k-n})\right| < \varepsilon$ или $\left\|X^k - X^{k-n}\right\| < \varepsilon$. Если оно выполняется, то перейти к шагу 4, иначе - к шагу 1, положив $k = k + 1$.

Шаг 4. Завершить вычисления, положив $X^* \approx X^{k+1}, f^* \approx f(X^{k+1})$.

Эффективность метода Зейделя существенно зависит от свойств целевой функции $f(X)$. Так, если линиями уровня целевой функции двух переменных являются концентрические окружности, то очевидно, что два шага исчерпывающего спуска, приведут в точку минимума из любой точки, т.е. минимум такой функции удается с помощью описанного алгоритма найти точно за конечное число шагов.

Типовой пример

Решить задачу минимизации функции двух переменных $f(X) = 5x_1^2 + 5x_2^2 + 8x_1x_2$ из начальной точки $X^0 = (5; 5)^T$ методом покоординатного спуска Зейделя с точностью 0,01.

0. Найдем значение функции в начальной точке:

$$f(X^0) = 5 * 5^2 + 5 * 5^2 + 8 * 5 * 5 = 450$$

$k = 0$.

ПЕРВЫЙ ШАГ ВНЕШНЕГО ЦИКЛА.

Первый шаг внутреннего цикла.

1. В нашем случае размерность вектора X равна 2, т.е. $n = 2$. =1.

$$j = k - n\left[\frac{k}{n}\right] + 1 = 1; \quad S^0 = \begin{pmatrix}1\\0\end{pmatrix};$$

$$X^1 = X^0 + \alpha S^0 = \begin{pmatrix}5\\5\end{pmatrix} + \alpha\begin{pmatrix}1\\0\end{pmatrix} = \begin{pmatrix}5+\alpha\\5\end{pmatrix}$$

Сформулируем вспомогательную задачу определения величины шага α_0:

$$f(X^1) = f\begin{pmatrix}5+\alpha\\5\end{pmatrix} = 5 * (5+\alpha)^2 + 5 * 5^2 + 8 * (5+\alpha) * 5 \to min$$

Или: $f(X^1) = 5\alpha^2 + 90\alpha + 450 \to min$

Данную задачу можно решить классическим методом:

$$f'(X^1) = (5\alpha^2 + 90\alpha + 450)' = 10\alpha + 90 = 0$$

$$\alpha = -9.$$

Следовательно, точка X^1 имеет координаты: $X^1 = \begin{pmatrix}5-9\\5\end{pmatrix} = \begin{pmatrix}-4\\5\end{pmatrix}$

Значение функции в новой точке: $f(X^1) = 45$. Видно, что значение функции уменьшилось: $f(X^1) < f(X^0)$

Конец первого шага внутреннего цикла.

2. Так как $j < n$ $(1 < 2)$, то спуск произведен еще не по всем координатам (внешний цикл не закончен). Тогда $k = 0 + 1 = 1$. Перейдем к шагу 1.

Второй шаг внутреннего цикла

3. $j = k - n\left[\frac{k}{n}\right] + 1 = 2; \quad S^1 = \begin{pmatrix}0\\1\end{pmatrix};$

Произведем спуск по второй координате из найденной точки X^1.

$$X^2 = X^1 + \alpha S^1 = \begin{pmatrix}-4\\5\end{pmatrix} + \alpha\begin{pmatrix}0\\1\end{pmatrix} = \begin{pmatrix}-4\\5+\alpha\end{pmatrix}$$

Вспомогательная задача:

$$f(X^2) = f\begin{pmatrix} -4 \\ 5+\alpha \end{pmatrix} = 5*(-4)^2 + 5*(5+\alpha)^2 + 8*(-4)*(5+\alpha) \to min$$

Или: $f(X^2) = 5\alpha^2 + 18\alpha + 45 \to min$

Данную задачу можно решить классическим методом:

$$10\alpha + 18 = 0$$

$$\alpha = -1.8$$

Следовательно, точка X^2 имеет координаты: $X^2 = \begin{pmatrix} -4 \\ 3{,}2 \end{pmatrix}$

Значение функции в новой точке: $f(X^2) = 28{,}8$. Видно, что значение функции уменьшилось: $f(X^2) < f(X^1)$

Конец второго шага внутреннего цикла.

4. $j = n$. Следовательно, внешний цикл завершен. Проверяем условие останова:

$$\left| f(X^k - f(X^{k-n}) \right| = \left| f(X^2 - f(X^0) \right| = |28{,}8 - 450| = 421{,}2 > \varepsilon$$

$$\left| f(X^k) - f(X^{k-n}) \right| = |f(X^2) - f(X^0)| = |28{,}8 - 450| = 421{,}2 > \varepsilon$$

Критерий не выполнен, переходим к первому шагу нового внешнего цикла.

КОНЕЦ ПЕРВОГО ШАГА ВНЕШНЕГО ЦИКЛА.

…….

Заметим, что линии уровня данной целевой функции — соосные эллипсы с центром в начале координат, большая ось которых наклонена под углом 135° к оси x_1. Результаты расчетов по приведенному в пункте 3 алгоритму приведены в таблице и графически проиллюстрированы на рис.5. При нахождении очередной точки X^k минимизирующей последовательности происходит смещение по прямой, параллельной одной из координатных осей, до точки с наименьшим на этой прямой значением функции$f(X)$. Очевидно, эта точка будет точкой касания рассматриваемой прямой и соответствующей линии уровня.

k	X1	X2	F(X)
0	5.000	5.000	450.000
1	-4.000	5.000	45.000
2	-4.000	3.200	28.800
3	-2.560	3.200	18.432
4	-2.560	2.050	11.796
5	-1.640	2.050	7.564
6	-1.640	1.310	4.841
7	-1.050	1.310	3.089
8	-1.050	0.840	1.984
9	-0.670	0.840	1.270
10	-0.670	0.540	0.808
11	-0.430	0.540	0.525
12	-0.430	0.340	0.333
13	-0.270	0.340	0.208
14	-0.270	0.220	0.131
15	-0.180	0.220	0.087
16	-0.180	0.140	0.058
17	-0.110	0.140	0.035
18	-0.110	0.090	0.022
19	-0.070	0.090	0.015
20	-0.070	0.060	0.009
21	-0.050	0.060	0.006