Градиентные методы

Выполнил студент группы 467
Павлов Валерий

Метод градиентного спуска

Суть метода градиентного спуска заключается в том, что в каждой i-й точке алгоритма вычисляется градиент [*a = z1 - 2*z2 + z3*], определяются направление движения и шаг. Так как за один шаг невозможно достичь точки минимума целевой функции, то строится последовательность точек, переходя от одной точки к другой, достигают точки минимума. В точке минимума все элементы вектора градиента принимают значение нуля. Для определения координат очередной точки используют направление, противоположное градиенту (антиградиент), а размер шага можно определить по различным правилам.

Два основных класса правил определения размера шага

С фиксированным коэффициентом изменения размера шага и с оптимальным подбором размера шага. Каждый класс правил содержит несколько конкретных методов поиска минимума. Для случая с фиксированным коэффициентом изменения размера шага координаты точки на k-м шаге определяются по формуле: *x^k=x^(k-1)-Sk*

Знак минус определяет направление движения против градиента (при поиске минимума целевой функции). Размер шага Sk на k-й итерации определяется по формуле: *Sk=dk*grad f(x^(k-1))* где dk— коэффициент изменения шага, как правило, меньше единицы.

Алгоритм метода градиентного спуска с использованием фиксированного коэффициента изменения шага

1. Задать координаты стартовой точки x_i^0
2. Задать значения p, d_k, ε.
3. Вычислить значение целевой функции, значения первых производных в текущей точке по каждой координате и определить антиградиент
4. Определить, достигнут ли минимум целевой функции, т. е. выполняется ли неравенство $\left|-\operatorname{grad} f\left(x^k\right)\right| < \varepsilon$, Если «да», то перейти к шагу 8. Если «нет», то перейти к шагу 5

Алгоритм метода градиентного спуска с использованием фиксированного коэффициента изменения шага

5. Вычислить размер шага по формуле *Sk=dk*grad f(x^(k-1))*
6. *Определить, надо ли уменьшать коэффициент изменения шага. Если неравенство* $\left|f\left(x^{k-1}\right)-f\left(x^{k}\right)\right| \geq d_{k}\left|-\operatorname{grad} f\left(x^{k-1}\right)\right|$ *не выполняется, то коэффициент изменения шага уменьшают в 2 раза, dk = 0,5 d(k-1) и переходят к шагу 5. Если неравенство выполняется, то переходят к шагу 7*
7. Определить координаты текущей точки по формуле *Sk=dk*grad f(x^(k-1))* и перейти к шагу 3.
8. Вывод координат точки минимума x и значения целевой функции в точке минимума *f*(x). Останов алгоритма.

Пример

Найти минимум целевой функции $F(x) = x_1^2 + 4x_2^2$ при $\varepsilon \le 0,01$ и $d_0 = 1$.

Сначала решим задачу при постоянном значении коэффициента (табл. 6.1) изменения шага $d = 1$ (пункт 6 алгоритма не выполняется).

Таблица 6.1

Итерация	x_1	x_2	$\frac{\partial f(x^k)}{\partial x_1}$	$\frac{\partial f(x^k)}{\partial x_2}$	$\mathrm{grad}\, f(x^k)$	Размер шага		d
						x_1	x_2	
0	2	2	4	16	16,5	–0,24	–0,97	1
1	1,76	1,03	3,52	8,24	8,96	–0,39	–0,92	1
2	1,37	0,11	2,74	0,88	2,88	–0,95	–0,31	1
3	0,42	–0,2	0,84	–1,6	1,8	–0,46	–0,89	1
4	–0,04	–1,09	–0,08	–8,72	8,72	0,009	1	1
5	–0,03	–0,09	–0,06	–0,72	0,72	0,08	1	1
6	0,05	–0,91	0,1	–7,28	7,28	–0,01	–1	1
7	0,04	–1,91	0,08	–15,28	15,28	–0,01	1	1
8	0,03	–0,91	0,06	–7,28	7,28	–0,01	1	1
9	0,02	0,09	0,04	0,72	0,72	–0,06	–1	1
10	–0,04	–0,91	–0,08	–7,28	7,28	0,01	1	1

Решение

Начиная с четвертой итерации, вокруг точки минимума целевой функции начинаются колебания, слабо сходящиеся к точке минимума.

Использование п. 6 алгоритма приводит к точке минимума целевой функции по асимптотическому закону, т. е. градиент функции плавно убывает (табл. 6.2).

Таблица 6.2

Итерация	x_1	x_2	$\frac{\partial f(x^k)}{\partial x_1}$	$\frac{\partial f(x^k)}{\partial x_2}$	$\text{grad} f(x^k)$	Значение целевой функции		d	Размер шага	
						текущее	предшествующее		x_1	x_2
0	2	2	4	16	16,5	20	0	1	–0,24	–0,97
1	1,76	1,03	3,52	8,24	8,96	7,34	20	1	–0,39	–0,92
2	1,37	0,11	2,74	0,88	2,88	1,93	7,34	1	–0,95	–0,31
3	0,42	–0,2	0,84	–1,6	1,81	0,34	1,93			
	Нарушено условие (3.40), для итерации 3							0,5	–0,48	–0,16
3	0,89	–0,05	1,78	0,4	1,82	0,8	1,93	0,5	–0,49	–0,11
4	0,4	–0,16	0,8	–1,28	1,51	0,26	0,8			
	Нарушено условие (3.40), для итерации 4							0,25	–0,25	–0,06
4	0,64	–0,11	1,28	–0,88	1,55	0,45	0,8			
	Нарушено условие (3.40), для итерации 4							0,125	–0,12	–0,03
4	0,77	–0,08	1,54	–0,64	1,67	0,62	0,8	0,125	–0,12	0,05
5	0,65	–0,03	1,3	–0,24	1,32	0,42	0,62	0,125	–0,12	0,02
6	0,53	–0,01	1,06	–0,08	1,06	0,28	0,42	0,125	–0,13	0,01
7	0,4	0	0,16	0	0,16	0,03	0,28	0,125	–0,31	0
8	0,09	0	0,18	0	0,18	0,01	0,03			
	Нарушено условие (3.40), для итерации 8							0,0625	–0,16	0
8	0,24	0	0,48	0	0,24	0,06	0,03	0,0625	–0,06	0
9	0,18	0	0,36	0	0,18	0,03	0,06	0,0625	–0,06	0
10	0,12	0	0,24	0	0,12	0,01	0,03	0,0625	–0,13	0
11	–0,01	0	–0,02	0	0,01	0,01	0,01			

Главный минус метода градиентного спуска

Процесс поиска минимума целевой функции может закончиться в стационарной точке (точка перегиба, седловая точка). Для выхода из стационарной точки необходимо предпринять дополнительные меры — определить тип точки (стационарная или нет) и в случае положительного ответа применить любой не градиентный метод для выхода из стационарной точки, и далее продолжить поиск методом градиентного спуска.