**Метод градиентного спуска.**

 Рассмотрим функцию f, считая для определенности, что она зависит от трех переменных***x,y,z.*** Вычислим ее частные производные *дf/дх,дf/ду, дf/дz* и образуем с их помощью вектор, который называют *градиентом* функции:

grad f(x, *у, z*) = *дf(х, у,z)/дх\*****i****+дf( x, у, z)/ду\*****j****+дf(x, y,z)/дг\*****k****.*

Здесь **i***,* ***j, k*** *-* единичные векторы, параллельные координатным осям. Частные производные характеризуют изменение функции f по каждой независимой переменной в отдельности. Образованный с их помощью вектор градиента дает общее представление о поведении функции в окрестности точки (*х, у,z).* Направление этого вектора является направлением наиболее быстрого возрастания функции в данной точке. Противоположное ему направление, которое часто называют *антиградиентным,* представляет собой направление наиболее быстрого убывания функции. Модуль градиента  
                          \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_                                    
grad*(х, у,z)**дf/дх(х, у,z))2+(дf/ду( x, у, z))2+(дf/дг(x, y,z))2.*

определяет скорость возрастания и убывания функции в направлении градиента и антиградиента. Для всех остальных направлений скорость изменения функции в точке *(х, у, z)* меньше модуля градиента. При переходе от одной точки к другой как направление градиента, так и его модуль, вообще говоря, меняются. Понятие градиента естественным образом переносится на функции любого числа переменных.

Перейдем к описанию метода градиентного спуска. Основная его идея состоит в том, чтобы двигаться к минимуму в направлении наиболее быстрого убывания функции, которое определяется антиградиентом. Эта идея реализуется следующим образом.

Выберем каким-либо способом начальную точку, вычислим в ней градиент рассматриваемой функции и сделаем небольшой шаг в обратном, антиградиентном направлении. В результате мы придем в точку, в которой значение функции будет меньше первоначального. В новой точке повторим процедуру: снова вычислим градиент функции и сделаем шаг в обратном направлении. Продолжая этот процесс, мы будем двигаться в сторону убывания функции. Специальный выбор направления движения на каждом шаге позволяет надеяться на то, что в данном случае приближение к наименьшему значению функции будет более быстрым, чем в методе покоординатного спуска.

Метод градиентного спуска требует вычисления градиента целевой функции на каждом шаге. Если она задана аналитически, то это, как правило, не проблема: для частных производных, определяющих градиент, можно получить явные формулы. В противном случае частные производные в нужных точках приходится вычислять приближенно.

Метод наискорейшего спуска

Суть метода наискорейшего спуска состоит в следующем. Как и прежде, в начальной точке определяется антиградиент минимизируемой функции. Однако теперь в направлении антиградиента делается ни один шаг, а движутся в данном направлении до тех пор, пока целевая функция убывает, достигает в некоторой точке минимума. В этой точке опять определяют антиградиент и ищут новую точку минимума целевой функции и так далее. В данном методе спуск имеет более целеустремлённый характер, производится более крупными шагами и градиент функции вычисляется в меньшем числе точек.

