НАПРАВЛЕННЫЙ ПОИСК ГЛОБАЛЬНОГО ЭКСТРЕМУМА МНОГОМЕРНЫХ ФУНКЦИЙ

Чипига А.Ф., Колков Д.А.
Северо-Кавказский государственный технический университет
Ставрополь, Россия
zik@ncstu.ru

Для осуществления направленного поиска глобального экстремума многомерной функции применимы градиентные методы [1, 4]. Они используют только первые производные целевой функции и являются методами линейной аппроксимации на каждом шаге, т.е. целевая функция на каждом шаге заменяется касательной гиперплоскостью к ее графику в текущей точке.

На k-м этапе градиентных методов переход из точки X^k в точку X^{k+1} описывается соотношением $X^{k+1} = X^k + l^k \cdot S^k$, где l^k - величина шага, S^k - вектор в направлении $X^{k+1} \to X^k$.

Способ определения l^k и S^k на каждой итерации связан с особенностями применяемого метода. Обычно выбор l^k осуществляется путем решения задачи минимизации целевой функции ($f(X)$) в направлении S^k. Поэтому при реализации градиентных методов необходимо использовать эффективные алгоритмы одномерной минимизации.

В основе простейшего градиентного метода [2] лежит следующая итерационная модификация формулы:

$X^{k+1} = X^k - a \cdot \nabla f(X^k)$, где a - заданный положительный коэффициент; $\nabla f(X^k)$ - градиент целевой функции первого порядка.

Недостатки:

- необходимость выбора подходящего значения a ;
- медленная сходимость к точке минимума ввиду малости $\nabla f(X^k)$ в окрестности этой точки.

Идея метода наискорейшего спуска проста [1, 2]: градиент целевой функции $f(X)$ в любой точке есть вектор в направлении наибольшего возрастания значения функции. Следовательно, антиградиент будет направлен в сторону наибольшего убывания функции и является направлением наискорейшего спуска. Антиградиент (и градиент) ортогонален поверхности уровня $f(X)$ в точке X. Если в $X^{k+1} = X^k + l^k \cdot S^k$ ввести направление $\overline{S_k} = -\frac{\nabla f(X^k)}{\left\|\nabla f(X^k)\right\|}$, то это будет направление наискорейшего спуска в точке X^k.

Тогда формула перехода из X^k в X^{k+1}:

$$X^{k+1} = X^k - l^k \frac{\nabla f(X^k)}{\left\|\nabla f(X^k)\right\|} = X^k - m \cdot \nabla f(X^k).$$

Антиградиент дает только направление спуска, но не величину шага. В общем случае один шаг не дает точку минимума, поэтому процедура спуска должна применяться несколько раз. В точке минимума все компоненты градиента равны нулю

Направление поиска, соответствующее наискорейшему спуску, связано с линейной аппроксимацией целевой функции. Методы, использующие вторые производные, возникли из квадратичной аппроксимации целевой функции. При разложении функции в ряд Тейлора отбрасываются члены третьего и более высоких порядков

$$f(x) \approx f(X^k) + \nabla^T f(X^k)(X - X^k) + \frac{1}{2}(X - X^k)^T G(X^k)(X - X^k),$$

где $G(X^k)$ - матрица Гессе.

Минимум правой части (если он существует) достигается там же, где и минимум квадратичной формы. Тогда формула для определения направления поиска $\overline{S}$:

$$f(X^k + \overline{S}) = f(X^k) + \nabla^T f(X^k)\overline{S} - \frac{1}{2}\overline{S}^T G(X^k)\overline{S}.$$

Минимум достигается при $G(X^k)\overline{S} = -\nabla f(X^k), (\overline{S} = -G^{-1}(X^k)\nabla f(X^k))$.

Алгоритм оптимизации, в котором направление поиска определяется из этого соотношения, называется методом Ньютона [1, 4], а направление $\overline{S}$ - ньютоновским направлением.

В задачах поиска минимума произвольной квадратичной функции с положительной матрицей вторых производных метод Ньютона дает решение за одну итерацию независимо от выбора начальной точки.

Идея алгоритма Пауэлла [3] состоит в том, что на каждом этапе поиска определяется минимум квадратичной функции, которой аппроксимируется целевая функция, вдоль каждого из сопряженных ко всем предыдущим. Затем выбирается новая система направлений с использованием результатов поиска и утверждения.

Процедура Пауэлла сходится к точке, в которой градиент равен нулю, если $f(X)$ - строго выпуклая функция. Эта точка является локальным экстремумом. Метод очень чувствителен к способу построения сопряженных направления и поэтому зависит от точности используемого одномерного поиска.

Алгоритм метода сопряженных направлений Пауэлла заключается в следующем.

Шаг 1. Задать начальную точку x^o и систему N линейно независимых направлений s^i ; целесообразно принять $s^i = e^i$, $i=1,2,...,N$.

Шаг 2. Минимизировать $f(X)$ при последовательном движении по $N+1$ направлениям; при этом полученная ранее точка минимума берется в качестве исходной, а направление s^N используется как при первом, так и при последнем поиске.

Шаг 3. Определить новое сопряженное направление с помощью обобщенного свойства параллельного подпространства.

Шаг 4. Заменить s^1 на s^2 и так далее. Заменить s^N сопряженным направлением. Перейти к шагу 2.

Для применения метода на практике его необходимо дополнить процедурами проверки сходимости и линейной независимости системы направлений. Если целевая функция квадратична и обладает минимумом, то точка минимума находится в результате реализации N циклов, включающих шаги 2-4. Здесь N - число переменных. Если же функция $f(X)$ не является квадратичной, то требуется более чем N циклов.

Литература

1. http://m-16rulez.narod.ru/Ucheba/MOP/Bo4.doc.

2. Иркитова М.А. Методы безусловной оптимизации, 2001.

3. http://m-16rulez.narod.ru/Ucheba/MOP/Bo3.doc.

4. Горелик В.А. Математическое программирование.