Лекция № 5.

Безусловная оптимизация многих переменных. Методы сопряжённых градиентов

**Методы сопряжённых направлений**

Высокая эффективность метода Ньютона способствует его популярности и расширению сфер использования. Так с использованием метода Ньютона удаётся минимизировать квадратичную функцию

,

где  – симметрическая положительно определённая квадратичная матрица размера , за один шаг.

Действительно, приведённая в предыдущей лекции схема метода Ньютона, которая имеет следующий вид:

, 

позволяет определить минимум функции  за один шаг

.

Однако наряду с высокой производительностью в методе Ньютона имеется ряд недостатков, среди которых следует выделить необходимость построения матрицы вторых производных минимизируемой функции, обращения этой матрицы или решения с ней системы линейных уравнений. Для обращения такой матрицы или решения с ней системы линейных уравнений, чтобы определить направление спуска, необходимо потребовать её невырожденность.

Поэтому возникает вопрос, а нельзя ли построить методы, которые будут иметь высокую скорость сходимости и требовать лишь первые производные для минимизируемых функций. Такие методы можно построить на основе идеи сопряжённых направлений.

В методах сопряжённых направлений точка минимума такой квадратичной функции может определяться за  шагов. Однако в этих методах не требуется строить и обращать матрицу вторых производных, а используются только первые производные минимизируемых функций.

**Определение.** Векторы  называются сопряжёнными относительно симметрической положительно определённой квадратичной матрицы , если все они отличны от нуля они отличны от нуля и  при .

**Лемма.** Пусть векторы  являются взаимно сопряжёнными. Тогда они линейно независимы.

Доказательство. Пусть это неверно, т.е. для всех  и  справедливо  и  ( и =0,1,…,*n*-1), но существует такое , которое можно выразить через остальные , но тогда , что возможно только при .

**Теорема.** Если векторы  () взаимно сопряжены, то для квадратичной функции

,

с симметрической положительно определённой квадратичной матрицей размера , последовательность одномерных минимизаций вдоль этих направлений, т.е.

, ,

приводит к отысканию минимума этой функции, т.е.



где ,  ̶ линейное подпространство, натянутое на векторы .

Доказательство. Определим 









.

Используя это равенство, получаем









.

Отсюда

.

Таким образом, метод, описанный в условиях теоремы, позволяет найти точку минимума квадратичной функции  не более, чем за  шагов.

**Методы сопряжённых градиентов**

Одну из версий метода сопряжённых градиентов для минимизации квадратичной функции  можно записать в следующем виде:



.

Коэффициент  определяется из условия  сопряжённости векторов  и 

.

Из этого условия коэффициент  определяется следующим образом:

.

Для обоснования того, что метод сопряжённых градиентов является методом сопряжённых направлений докажем следующие две леммы.

Лемма 1. Векторы  и  ортогональны,

Доказательство. Умножим слева уравнение , описывающее алгоритм метода сопряжённых градиентов, на матрицу квадратичной функции  и получим . Учитывая, что  и , получим .

Точка  является точкой минимума квадратичной функции  на прямой , и поэтому . Тогда



Это равенство обеспечивает сходимость метода сопряжённых направлений к минимуму квадратичной функции , поскольку скалярное произведение минимизируемой функции на направление спуска такое же, как и в градиентном методе.

Из равенства нулю  следует, что



Из ранее полученного равенства  следует



Квадратичная функция  при  принимает вид:





Минимум этой функции по  достигается, когда , т.е.

 и

.

Подставляя в выражение  это значение  получаем

,

учитывая, что . Лемма доказана.

Лемма 2. Пусть , точки  и векторы  получены по формулам

, ,



соответственно. Тогда векторы  взаимно сопряжены, а градиенты  взаимно ортогональны.

Доказательство.

Доказательство леммы проведём по методу математической индукции. Предположим, что векторы  взаимно ортогональны, а векторы   взаимно сопряжены.

Покажем, что векторы  взаимно ортогональны.

Как уже было доказано, для всех  было показано, что . Для доказательства этого условия оценим величину для произвольного .



Используемое здесь равенство  было получено и приведено ранее, равенство  справедливо по предположению, равенство  следует из описания метода, а равенство  справедливо по предположению.

Покажем, что векторы  взаимно сопряжены.

Из равенства равенство  имеем при  

.

Для 



Последнее равенство следует по уже ранее доказанному условию о взаимной ортогональности векторов . Лемма доказана.

Из этой леммы следует следующая теорема.

Теорема.

Метод

, ,

, 

позволяет определять минимума квадратичной функции 

,

с симметрической положительно определённой квадратичной матрицей размера , не более чем за шагов.

Методы сопряжённых градиентов являются весьма эффективными методами поиска минимума функций, имеющих непрерывные первые производные. В случае выпуклых функций с помощью методов сопряжённых градиентов определяются глобальные минимумы, а для функций, не являющихся выпуклыми можно определять только локальные минимумы. Для функций, не являющихся квадратичными сходимость к минимуму за шагов не обеспечивается, а после каждых шагов требуется обновление метода. При обновление метода направление поиска  определяется из условия, .

Рассмотрим две версии этого метода.

Итерационная схема метода Флетчера Ривза имеет следующий вид:

,

,

,

,

.

Алгоритм Флетчера-Ривса при некотовых условиях обладает квадратичной сходимостью. Преимуществом этого алгоритма является то, что он не требует обращения матрицы и экономит память ЭВМ, так как ему не нужны матрицы, используемые в Ньютоновских методах, но в то же время почти столь же эффективен как квази-Ньютоновские алгоритмы. Так как направления поиска взаимно сопряжены, то квадратичная функция будет минимизирована не более, чем за *n* шагов. В общем случае используется рестарт, который позволяет достаточно быстро получать хорошие результаты.

Алгоритм Флетчера-Ривса чувствителен к точности одномерного поиска, поэтому при его использовании необходимо устранять любые ошибки округления, которые могут возникнуть. Кроме того, алгоритм может отказать в ситуациях, где Гессиан становится плохо обусловленным. Гарантии сходимости всегда и везде у алгоритма нет, хотя практика показывает, что почти всегда алгоритм приводит к приближенному оптимуму.

**Метод Поллака Райвера.** Итерационная схема этого метода аналогична итерационной схеме метода Флетчера-Ривза. Только направления спуска строятся с помощью следующих соотношений:

,

,

,

,

