Лекция № 4.

Безусловная оптимизация многих переменных. Методы скорейшего спуска и сопряжённых направлений.

Задача поиска минимума дифференцируемой функции  сводится к решению уравнения .

Как правило, задача поиска корней уравнения так же сложна, как и задача минимизации функции  и эти задачи приходится решать численно.

Итерационный процесс, используемый в методах оптимизации для поиска точек экстремума функций, имеет вид:

, ,

где  ‒ шаг итерационного процесса,  ‒ направление -го шага в итерационном процессе, при минимизации  ‒ направление убывания или спуска на -м шаге, ‒ начальная точка процесса.

**Условия остановки.**

На практике часто используются следующие условия остановки:

,

,

,

где ,  и  ‒ малые постоянные величины.

**Эффективность метода.**

Эффективность сходящегося метода можно охарактеризовать с помощью понятия скорости сходимости.

Пусть  при . Говорят, что последовательность  линейно сходится к , если существуют такие константы  и , что

 при 

Говорят, что последовательность  сверхлинейно (со сверхлинейной скоростью) сходится к , если

, , при .

Говорят, что последовательность  с квадратичной скоростью сходится к , если существуют такие константы  и , что

 при .

**Градиентный метод**

Итерационный процесс, используемый в методах оптимизации для поиска точек экстремума функций, имеет вид:

, ,

где  ‒ шаг итерационного процесса,  ‒ направление -го шага в итерационном процессе, при минимизации  ‒ направление убывания или спуска на -м шаге, ‒ начальная точка процесса.

В градиентном методе величина  выбирается равной антиградиенту функции  в точке , т.е. . Тогда вычислительная схема градиентного метода принимает вид:

, >0, 

В градиентном методе используются различные способы выбора длины шага .

Если длина шага выбирается из условия минимума функции  вдоль направления антиградиента

: ,

то получаем версию градиентного метода, называемую методом скорейшего спуска.

Рассмотрим более подробно другую версию градиентного метода, в которой выбор величины шага осуществляется путем дробления. В этом случае величина  выбирается из условия:

, ,

причём, если функция  дифференцируема и удовлетворяет неравенству

, ,

которое называется условием Липшица, , то существует такое , что

.

В случае, когда функция  не является выпуклой, градиентный метод может обеспечить лишь сходимость к множеству стационарных точек функция .

**Теорема о сходимости**.

Пусть функция  дифференцируема и ограничена снизу на , а её градиент удовлетворяет условию Липшица

, где .

Тогда для градиентного метода

, >0, ,

, ,

при произвольной начальной точке  имеем

.

Доказательство.

Из условия выбора шага



следует при , что

.

Поскольку функция  ограничена снизу и монотонно убывает, то  при , и поэтому  при . Теорема доказана.

**Теорема о скорости сходимости.**

Пусть функция  дважды непрерывно дифференцируема, причём матрица её вторых производных не вырождена и удовлетворяет условию

,

при любых , а последовательность строится следующим образом

, >0, ,

:  или из условия

, ,

Тогда при любой начальной точке  будет выполняться , где  – точка минимума функция  (единственная). Для скорости сходимости справедливы оценки

,

.

Доказательство.

Используя формулу Тейлора, получаем



или



 поскольку согласно условию теоремы

.

С другой стороны, используя следующее разложение в ряд Тейлора, получаем

,

или с учётом равенства  и приведённого выше условия теоремы имеем

.

Отсюда, согласно этому и ранее полученному неравенствам, выводим



или

, а также

 или .

Тогда из полученного ранее условия



имеем



Отсюда получаем



Подставляя полученную оценку в соотношение для выбора шага, находим

.

Из разложения в ряд Тейлора

,

получаем



Из этого неравенства следует, что условие выбора шага будет заведомо выполняться, если  будет выбрано так, что , т.е. . С учётом этого из неравенства

 путём прибавления и вычитания в левой части неравенства получаем

,

где . Отсюда получаем

.

Из этого неравенства с учётом левой части полученного неравенства



определяем

.

Теорема доказана.

**Метод Ньютона**

Предположим, что функция  выпукла и дважды дифференцируема на , причём матрица  не является вырожденной на . В методе Ньютона последовательность  следующим образом.

Разложим функцию  в ряд Тейлора до второго члена включительно. Такое разложение имеет вид:



Для вычисления минимума функции  используем её приближение квадратичной функцией , и в качестве следующего приближения используется минимум этой функции, т.е. решается задача

.

Необходимое и достаточное условие минимума этой функции имеет вид:

.

Пусть матрица  не вырождена. Решая полученную систему уравнений



и принимая найденную точку минимума квадратичной функции  за , получим следующую схему метода Ньютона:

, .

В такой постановке метод Ньютона сходится к минимуму функции  только из некоторой окрестности её минимума .

Для этой версии метода Ньютона справедлива следующая теорема.

**Теорема.**

Пусть функция  дважды дифференцируема, сильно выпукла с константой  на  и удовлетворяет условию Липшица следующего вида:

, ,

где , а начальная точка  такова, что .

Тогда последовательность  сходится к , т.е. при  с квадратичной скоростью:

.

Существует и другая версия метод Ньютона, которая сходится к минимуму функции  из произвольной начальной точки.

Вычислительная схема этой версии метода Ньютона имеет вид:

, .

Величина шага  может выбираться из условий минимизации шага вдоль выбранного направления :

: 

или с помощью метода дробления шага:

, ,

Можно показать, что эта версия метода Ньютона сходится к минимуму функции  из произвольной начальной точки . Скорость сходимости этой версии будет либо сверхлинейной, либо квадратичной при некоторых дополнительных требований к функции .

Высокую эффективность метода Ньютона можно продемонстрировать на минимизации квадратичной функции

,

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

Приведённая выше схема метода Ньютона имеет следующий вид:

, .

где направление спуска  для квадратичной функции имеют вид:

.

Тогда в соответствии со схемой метода Ньютона имеем

,

т.е. точку минимума этой квадратичной функции удаётся определить с помощью метода Ньютона всего за один шаг.