***Достоинства и недостатки градиентных методов и метода Ньютона***

|  |  |  |
| --- | --- | --- |
| **Метод** | **Достоинства** | **Недостатки** |
| Градиентный метод | 1. Глобальная сходимость, т.е. слабые требования на начальные приближения точки х0 и к *f*(*x*);  2. Относительная простота вычислений | 1. Медленная сходимость (геометрическая скорость сходимости, порядок сходимости *d*=1);  2. Необходимость вычисления длины шага. |
| Метод Ньютона | 1. Быстрая сходимость. | 1. Локальная сходимость (начальная точка должна быть близка к *х\**); 2. Большой объем вычислений (на любом шаге требуется вычислять матрицу вторых производных и обращать её). 3. Жесткие требования на саму функцию (непрерывная вторая производная). |

*Порядок применения методов*

1. На 1-м этапе – методы 1-ого порядка, т.к. они обеспечивают глобальную сходимость.
2. На 2-м этапе (когда приращения невелики) – выгодно применять методы 2-ого порядка.

Перечисленные методы (градиентные и Ньютона) являются классическими.

Можно предложить методы более высокого порядка, тогда естественно ожидать, что порядок сходимости будет равен *p*.

Как уже отмечалось, к недостаткам метода Ньютона относится:

* локальная сходимость;
* большой объем вычислений;
* жесткие требования на гладкость функции.

В силу названных причин применение классического метода Ньютона не всегда приводит к успеху.

Многочисленные модификации направлены на то, чтобы, *сохраняя* основное достоинство метода Ньютона – *быструю сходимость*, *уменьшить трудоемкость и ослабить* требования на выбор *начального приближения*.

**Метод Ньютона с регулировкой шага**

Рассмотрим метод

,

это метод Ньютона с регулировкой шага.

При *αk* ≡ 1 мы получили классический метод Ньютона.

Выбор *αk* обычно – из условия min функции, вдоль заданного направления, или методом дробления шага, обеспечивающего выполнение условия *ϕ*(*xk*+1) < *ϕ­*(*xk*).

Можно показать, что подобные методы регулировки шага *сходятся* при *любой начальной* точке *x*0∈*Rn*, причем скорость сходимости будет либо *сверхлинейна*, либо *квадратичная* в зависимости от требований, которым удовлетворяет функция *ϕ*.

Таким образом, с помощью регулировки длины шага преодолевается недостаток метода, связанный с необходимостью отыскания хорошего начального приближения.

Однако, трудоемкость вычислений при этом не исчезает.

Более перспективным в этом плане оказывается другой подход, при котором строится аппроксимация матрицы (*ϕ*″(*xk*))-1 на основе информации о значениях градиентов *ϕ*′(*xk*), *ϕ*′(*xk*+1),…

**Квазиньютоновы методы**

Пусть функция *ϕ* дважды дифференцируема. Рассмотрим метод

 (1)

*αk* – шаг, *Hk* – матрица.

1. Если *Hk* = единичная, имеем градиентный метод.
2. Если , то это метод Ньютона (с точностью до шага).
3. Если , то имеем метод, который объединяет достоинства обоих методов.

Заметим, что

.

Полагая невырожденной матрицу *ϕ*″(*xk*+1), отсюда с точностью до членов более высокого порядка малости по сравнению с  имеем:

.

Рассмотрим квадратичную функцию . Для нее

,

и приближенное равенство обращается в точное:

.

Поэтому естественно потребовать, чтобы для матрицы *Hk*+1, приближающей (*ϕ*″(*xk*+1))-1, выполнялось условие:

 (\*)

Это условие называется *квазиньютоновским*. Оно лежит в основе целого ряда методов аппроксимации (*ϕ*″)-1. Соответствующие методы минимизации, для которых на любом шаге выполняется квазиньютоновское условие, также называются квазиньютоновскими.

Пусть приближения к (*ϕ*″)-1 пересчитываются шаг от шага по формуле .

Различные квазиньютоновские методы различаются способом вычисления "добавки" Δ*Hk* таким образом, чтобы удовлетворялось соотношение (\*).

**1. Метод Дэвидона-Флетчера-Пауэлла**

Обозначим:



Метод заключается в построении релаксационной последовательности по следующему правилу:

 (1.1)

Длина шага *αk* в квазиньютоновых методах выбирается так же, как в методе наискорейшего спуска:



Как правило, начальное значение *H*0 = *I*. Вообще, если *H*0 ­– симметричная матрица, то *Hk* ­– симметричная матрица для любого *k*.

**2. Метод Бройдена-Флетчера-Шанно**

Имеем.

Если поставить задачу уточнять обратную матрицу, т.е.  тогда:

 (1.2)

(этот метод более устойчив к ошибкам округления)

Можно доказать, что для квадратичной функции , где *A* – симметричная, положительно определенная матрица, оба метода (1.1) и (1.2) при любом начальном приближении *x*0∈*Rn* генерируют одну и ту же последовательность точек , причем

,

т.е. *квазиньютоновские методы* позволяют найти min *квадратичной* функции *за n-шагов*.

Для неквадратичной функции, это не так. Однако можно показать, что при соответствующих предположениях , причем скорость сходимости сверхлинейна.

Так, например, пусть *ϕ* – дважды непрерывно дифференцируемая функция, сильно выпукла на *Rn*.

Тогда при любом начальном приближении *x*0∈*Rn* последовательность точек {*xk*}, определяемая формулами (1.1.) и (1.2), сходится к *x*\*.

Если, при этом, для всех *x*:, справедливы неравенства

,

то *xk* сходится к *x*\* сверхлинейно

.

*Замечания о квазиньютоновских методах*:

1. Это двухшаговые методы.
2. Для квадратичных функций сходятся за *n*-шагов.
3. Обладают следующими преимуществами:

* небольшая вычислительная сложность;
* более глобальная сходимость, чем в методе Ньютона;
* сверхлинейная скорость сходимости.