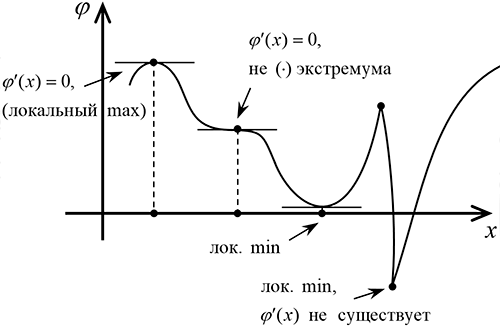
При изучении задач оптимизации в первую очередь возникает вопрос о существовании решения. Напомним в этой связи некоторые результаты из *математического анализа*.

**Теорема 1** (Вейерштрасса).Если *X* – компакт в *Rn* (т.е. замкнутое ограниченное множество), а *ϕ* – непрерывная функция на *X*, то существует *x*\*∈*X*: *x*\* – глобальный минимум *ϕ* на *X*, т.е. *глобальное решение задачи*  *существует*!

**Теорема 2** (необходимые условия локального минимума). Если *ϕ* – дифференцируема в точке *x*\*∈*X* и *x\** – локальный минимум, то *ϕ*′(*x*\*) = 0 (градиент равен нулю).

**Определение.** Точка  в , называется *стационарной* (обратное не верно).

******

**Теорема 3** (достаточное условие локального минимума).

Если ** дважды дифференцируема в точке *x*\*∈*Rn* и выполняется:

1) *ϕ*(*x*\*) = 0;

2) матрица Гессе *ϕ*″(*x*\*) положительно определена,

то *x\** – (строгий) локальный минимум функции **

**Определение.** Матрица *H* называется *положительно определенной*, если .

*Критерий Сильвестра*: *H* положительно определена ⇔ ее главные миноры положительны.

Приведем несколько теорем для выпуклых задач.

**Определение.** Задача минимизации  называется выпуклой, если *X* – выпуклое множество, ** – выпуклая функция на *X*.

Справедливы следующие теоремы:

**Теорема 4.** Если задача минимизации  выпукла, то любое её локальное решение является также глобальным.

**Доказательство.** Пусть *x\** – локальное решение задачи, т.е. при некотором *ε*> 0 выполняется условие:

 при ,

где  – шар радиуса *ε* > 0 с центром в *x\**.

Для любых точек *x*∈*X*: *x* ≠ *x\**, положим .

Тогда . Покажем это.

1. Пусть ,

Если  ⇒ точка 

1. Пусть 



⇒ точка 

и, следовательно,

,

т.е. *x\** – глобальное решение задачи, *ч.т.д.*

Таким образом, *для выпуклых задач* понятия *локального и глобального решений не различаются*.

Второе свойство выпуклых задач можно высказать в виде следующего общего принципа: *необходимые условия оптимальности* в том или ином классе задач минимизации при соответствующих предположениях выпуклости *оказываются и достаточным*.

**Теорема 5.** Пусть функция ** выпукла на *Rn* и дифференцируема в точке *x*\*∈*Rn*. Если *ϕ*′(*x*\*) = 0, то *x*\* – точка минимума функции на *Rn*, т.е. решение задачи минимизации .

**Доказательство.** Для ∀*x*∈*X*, *λ*∈[0,1] имеем

.

Преобразуя эту формулу и, пользуясь дифференцируемостью функции ** в точке *x\**, получаем:

;

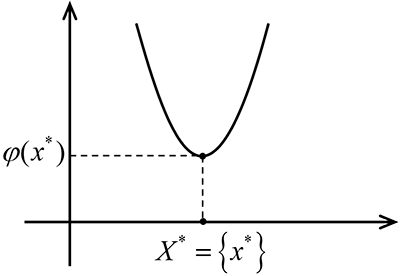
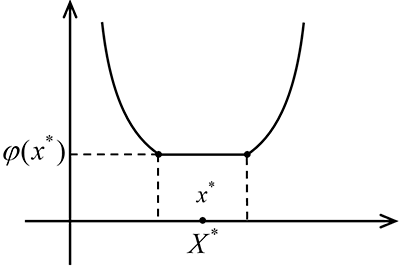
Отсюда предельным переходом при *λ*→ 0 выводим, что , *ч.т.д.* (т.е. для ∀*x*∈*X* ).

Полученные свойства выпуклых задач имеют важное значение не только в теории, но и в численных методах оптимизации. Дело в том, что большинство существующих численных методов позволяет, вообще говоря, находить лишь локальные решения, а точнее – стационарные точки задачи. Теоремы 4 и 5 говорят о том, что для выпуклой задачи *отыскание стационарной точки* автоматически означает *отыскание решения, причем глобального*.

Укажем ещё одно полезное свойство выпуклых задач.

**Теорема 6.** Пусть задача минимизации  выпукла и имеет решение.

Тогда множество её решений  выпукло. Если при этом **(*x*) строго выпукла на *X*, то решение единственно, т.е. *X \**состоит из одной точки.



**Доказательство:**

1. Пусть 

При этом

 (\*)

По определению *X\** неравенство может выполняться только как равенство, поскольку 

, т.е. *X\** – выпукло.

1. Пусть ** – строго выпукла. Если предположить, что в *X\** существуют две различные точки *x*1 и *x*2, то при *λ*∈[0,1] неравенство (\*) должно быть строгим, что невозможно, т.к. *ϕ*\*– min и получается < min.

*Трудности*:

1. В случаях, когда функция ** достаточно проста, теоремы 1-3 помогают решить задачу минимизации даже в явном виде. Однако зачастую задача поиска стационарных точек является нетривиальной. А затем – перебор стационарных точек в поисках точки локального минимума, затем – перебор локальных экстремумов в поисках глобального экстремума.
2. Для задач условной минимизации теоремы 1-3 применимы в случае, когда локальное решение *x\** – внутренняя точка допустимого множества *X*. Если же экстремум достигается в угловых точках границы множества условий, то нарушается дифференцируемость ⇒ неприменимость методов классического анализа.

Т.о., в большинстве случаев задачу min*ϕ*(*x*) приходится *решать численно с применением ЭВМ и специальных методов минимизации*.

**Безусловная минимизация функции**

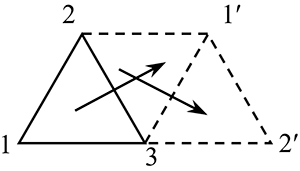
Методы оптимизации функций в *Rn* делятся на:

* локальные методы (поиск локального min, т.е. такой точки *x\**, что существует *δ* > 0, );
* нелокальные (или прямые) методы (поиск глобального min для ограничений снизу функции **(*x*), т.е. если *α\** – нижняя грань, то поиск такой точки *x*\*: *ϕ*(*x*\*) = *α*\*). Для этих методов не требуется аналитического задания функции, надо только уметь вычислять ее значение в любой точке. Обычно – для функций сложной структуры.

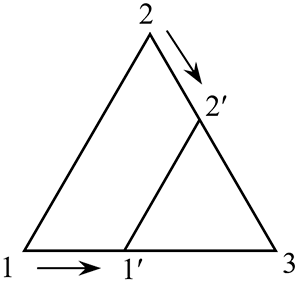
Нелокальные методы сводятся к уменьшению области, внутри которой находится оптимальная точка. Пример нелокального метода – *симплексный метод*.

**Определение.** *Симплекс* – выпуклое тело в *Rn*, состоящее из (*n*+ 1) равноудаленных точек – вершин симплекса, отрезок их соединяющий – ребро симплекса, в *R*2 – треугольник, в *R*3 – тетраэдр.

Неформальное описание симплексного метода: состоит из двух процедур – отражение и сжатие.

– *отражение*: симметричное отражение вершины с наибольшим значением **(*x*) относительно противоположной грани ["перекатывание симплекса"]. Если , то выбирается другая (*i +*1)-я вершина.

Когда зацикливание (все (*n*+ 1)-вершины перебрали), то

– *сжатие*: уменьшение размеров симплекса при сохранении вершины с наименьшим значением **(*x*), затем переход к отражению, и так далее, пока ребро симплекса не станет меньше некоторого числа: .

*Достоинства*: с большой вероятностью метод не распознает локальный минимум ("не остановится").

*Локальные* методы основаны на построении *релаксационной* последовательности {*xi*} такой, что  и .

Поэтому *релаксационные* методы называют также *методами спуска*.

**Классификация релаксационных методов**

С одной стороны,

* *одношаговые* методы:  – каждый шаг (*i*+ 1) зависит только от предыдущей точки *xi* и значения функции *ϕ*(*xi*);
* *двухшаговые* методы:  – зависимость от двух предыдущих точек;
* *и т.д.;*

С другой стороны,

* *методы нулевого порядка:* если используются только значения минимизируемой функции **(*x*);
* *методы первого порядка*: если используются только значение **(*x*) и **′(*x*);
* *методы второго порядка*: если используются значения **(*x*), **′(*x*) и **″(*x*);
* *etc;*

**Градиентные методы (методы первого порядка)**

Итак, будем рассматривать задачу:

 (безусловная минимизация),

предполагая, что функция **(*x*) непрерывно дифференцируема на *Rn*, т.е. *ϕ*(*x*)∈*C*1(*Rn*).

По определению дифференцируемой функции

, (1)

где .

Если , то при достаточно малых  главная часть приращения для ** будет определяться дифференциалом функции . Оценим величину Справедливо неравенство Коши-Буняковского:

,

причем, если *ϕ*′(*x*) ≠ 0, то правое неравенство превращается в равенство, только при *h*=−*αϕ*′(*x*), а левое только при *h*=*αϕ*′(*x*), где *α*= *const*≥ 0.

Отсюда ясно, что при *ϕ*′(*x*) ≠ 0 *направление наибыстрейшего возрастания* функции **(*x*) в точке *x* совпадает с *направлением градиента* **(*x*), а направление наибыстрейшего убывания – с направлением антиградиента –**′(*x*).

Это свойство градиента лежит в основе ряда итерационных методов минимизации функций. Один из таких – *градиентный*. Он предполагает, как, впрочем, и все остальные итерационные методы, наличие априорной точки начального приближения.

Предположим, что начальная точка *x*0 уже выбрана, тогда градиентный метод заключается в построении последовательности {*xk*} по правилу:

 (2)

*αk* – величина шага, *xk* – направление спуска.

Если , то шаг  можно выбрать так, чтобы получить релаксационную последовательность: . Действительно, подставляя (2) в (1), имеем:

,

при всех достаточно малых *αk*> 0.

Если , то *xk* – стационарная точка. В этом случае процесс (2) прекращается и проводятся дополнительные исследования поведения функции в окрестности точки *xk* для выяснения того, достигается ли в точке *xk* минимум функции **(*x*) или не достигается.

Существуют различные *способы выбора величины шага* *­­k* в методе (2). В зависимости от способа выбора *­­k* можно получить различные варианты градиентного метода.

**Метод наискорейшего спуска**

На луче , направленном по антиградиенту, введем функцию одной переменной



и определим *­­k* из условий

.

Другими словами *­­k* выбирается так, чтобы *ϕ*(*xk*+1) в заданном направлении была наименьшей для чего на любом шаге необходимо решать задачу одномерной минимизации функции *ψ* (*α*), например, с помощью .

**Пример.** Рассмотрим задачу



с начальной точкой .

Из общих соображений ясно, что *ϕ*min = 0 при 

*1-й шаг*:



Ищем



Функция *ψ*(*α*) имеет следующий вид:



Решаем уравнение , т.е.

;

.

*2-й шаг*:

;

.

Решаем уравнение  ⇒

; ⇒

.

*3-й шаг*:





Решаем уравнение  ⇒

;

, и.т.д.

*Представим решение задачи графически:*

Из графического представления можно сделать вывод, что имеет место:

|  |  |
| --- | --- |
| C:\Users\user\AppData\Local\Microsoft\Windows\INetCache\Content.Word\15.png | ⇒ а) сходимость к истинной точке минимума  б) взаимная перпендикулярность градиентов |