**Выпуклое программирование**

Рассмотрим задачу математического программирования следующего вида:

(\*) 

**Определение.** Если в задаче (\*) целевая функция *ϕ*(*x*) – выпуклая и допустимое множество *X* – выпукло, то задача (\*) называется задачей выпуклого программирования.

Рассмотрим теперь задачу математического программирования следующего вида:

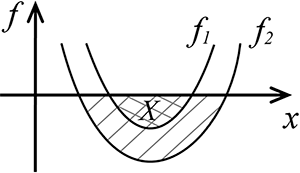
(\*\*)

Покажем, что допустимое множество *X* задачи (\*\*) – выпукло.

Действительно, пусть *x*1, *x*2∈*X*, *α*∈(0,1)

⇒ покажем, что , т.е. ∀*i* *fi*(*x*) ≤ 0.

Имеем, для ∀*i* 

⇒ т.о. ∀*i fi*(*x*) ≤ 0 ⇒ точка *x*∈*X*, т.е. *X* – выпукло.

"Надграфик" выпуклой функции, т.е. множество

 – выпуклое множество.

**Определение.** Задача (\*\*) называется основной задачей выпуклого программирования (ОЗВП).

**Свойства выпуклых функций**

1. **Неравенство Йенсена**

Пусть *f*(*x*) – выпуклая функция на выпуклом множестве *X*. Тогда

 при всех ;

**Доказательство.**Индукция по *m*. Пусть *m* = 1 ⇒ очевидно.

Пусть для *m*= *k* утверждение верно.

Докажем для *m* = *k* + 1.

Пусть .

Если , то .

Если  ⇒ представим *x* в следующем виде:

, где .

Тогда:

, *ч.т.д.*

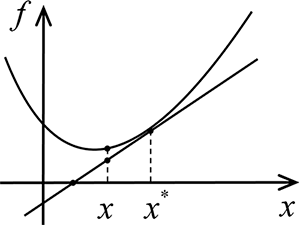
1. Пусть *f*(*x*) – выпуклая на выпуклом множестве ** функция. Тогда любой её локальный минимум на множестве *X* является одновременно и глобальным (доказательство было ранее в лекциях).
2. Пусть *f*(*x*) – выпуклая функция на выпуклом множестве , дифференцируемая в точке *x*\*∈*X*. Тогда .

График *f* лежим не ниже касательной гиперплоскости к графику функции *f* в точке (*x*\*, *f*(*x*\*)).

(*Напоминание: график линейной функции*  *называется касательной гиперплоскостью к графику функции f в точке* (*x*\*, *f*(*x*\*))).

**Доказательство.**По определению выпуклой функции для ∀*x*, *x*\*∈*X*, *λ*∈[0,1] имеем:

.

Преобразуя эту формулу, имеем:



Переходя к пределу при *λ* → 0, имеем искомое соотношение.

1. Пусть *f* – дважды непрерывно дифференцируемая функция на выпуклом множестве Тогда *f* выпукла на *X* ⇔ матрица Гессе *f*″ неотрицательно определена, т.е. ∀*x*\*∈*X*, ∀*h*\*∈*Rn* .

Был ранее без доказательства критерий сильной выпуклой функции *f* с параметром ∅≥0:



**Доказательство.**

*Необходимость.* Пусть *f* – выпукла на *X*.

а) Сначала предположим, что (*x*\* – внутренняя точка множества *X*, т.е. существует *ε*-окрестность точки *x*\*, все точки которой принадлежат *X*). Тогда для ∀*h*\*∈*Rn* имеем  при всех достаточно малых *α* > 0.Поскольку *f* – дважды дифференцируема в *x*\*, то можно записать:

⇒

⇒

.

Переходя к пределу при *α* → 0, имеем требуемое соотношение.

б) Пусть теперь *x*\*∈*X* – произвольная точка ⇒ существует последовательность точек , сходящаяся к *x*\*. По доказанному выше, для ∀*h*∈*Rn* имеем:

,

причем последовательность матриц  сходится к  в силу непрерывности  в *x*\* (*непрерывность всех вторых частных производных*) ⇒ имеем требуемое соотношение.

*Достаточность:* Пусть справедливо . Тогда, рассмотрим произвольные точки  и положим . Используя формулу Тейлора с остаточным членом в форме Лагранжа, имеем для некоторого *α*∈(0,1)

 (по предположению)

Итак, имеем:

 для .

Надо показать, что для  ,

т.е. тем самым показать, что функция f – выпуклая.

Для этого зафиксируем произвольные и рассмотрим точку *х*\*: (*в силу выпуклости X*).

Тогда:



и сложим их: ⇒

⇒,

т.е. *f* – выпуклая, *ч.т.д.*

**Пример.** Пусть  – квадратичная функция, *A* – симметричная матрица. Тогда *f* – выпуклая ⇔ *A* – неотрицательно определена.

Вообще, можно привести *критерии строгой и сильной выпуклости функций* аналогично тем, которые сейчас были доказаны для выпуклой функции (доказательство – аналогичное).

* Итак, для дифференцируемой функции *f*(*x*):
* *строгая выпуклость* эквивалентна неравенству

,

* *сильная выпуклость* эквивалентна неравенству

.

*Графически:*

*выпуклость* – возможно касание касательной плоскости;

*строгая выпуклость* – единственная точка касания с касательной плоскостью;

*сильная выпуклость* – график расположен внутри некоторого параболоида вращения

.

* Для дважды дифференцируемых функций *f*(*x*):
* достаточным условием *строгой выпуклости* *f*(*x*) является *положительная определенность* при ∀*x*∈*X* её матрицы Гессе *f*″(*x*);
* достаточным условием *сильной выпуклости* *f*(*x*) является *положительная определенность* матрицы , где *E* – единичная матрица, а *l* > 0.

Эти критерии в сочетании с критерием Сильвестра дают удобный аппарат для проверки выпуклости функций небольшого числа переменных.

**Функция Лагранжа**

Рассмотрим основную задачу выпуклого программирования:



*ϕ*(*x*), *fi*(*x*) – выпуклые функции.

В основной задаче выпуклого программирования имеем *m* условий, определяющих допустимое множество *X*.

Рассмотрим вектор .

**Определение.** Функция

,

называется *функцией Лагранжа* для основной задачи выпуклого программирования, где .

**Определение.** Пара (*x*\*, *λ*\*) называется *седловой точкой* функции Лагранжа на множестве , если ∀*x*∈*Rn* и ∀*λ*≥ 0

,

т.е.



*Наличие седловой точки означает, что операции минимизации и максимизации можно переставлять местами.*

В задачах *классического* анализа об условном экстремуме (задачи, в которых допустимое множество задается *системой уравнений*) важную роль играет метод множителей Лагранжа: решение исходной задачи ищется среди стационарных точек функции *L*(*x*, *λ*) – точек, удовлетворяющих системе уравнений

.

В задачах выпуклого (и, в частности, линейного) программирования функции Лагранжа также отводится важное место: при весьма общих предположениях задача выпуклого программирования сводится к отысканию седловых точек функции Лагранжа.

**Теорема** *о седловой точке функции Лагранжа* (достаточные условия оптимальности).

Если пара (*x*\*,*λ*\*) является седловой точкой функции Лагранжа  на множестве *x*∈*Rn*, *λ*≥ 0, то *x*\* – оптимальная точка основной ЗВП.

**Доказательство.** По определению седловой точки имеем:

 (\*)

а) Из левого неравенства (\*) убираем *φ*(*x*\*), и получаем:

, для ∀*i*, т.е. *x*\* – допустимая точка.

Действительно, если бы существовал индекс *i*: *fi*(*x*\*) > 0, то слева имели бы неограниченную сумму (т.к. ∀*λi*≥ 0), а справа имеем ограничение ⇒ для ∀*i fi*(*x*\*) ≤ 0.

б) В частности, левое неравенство (\*) верно и для *λ* = 0, тогда имеем, но

 (\*\*)

в) Подставим (\*\*) в правое неравенство (\*):



Поскольку, для ∀*x*∈*X* 

Итак, получили, что для ∀*x*∈*X*  , т.е. *x*\* – оптимальная точка, *ч.т.д.*

Отметим, что при доказательстве теоремы нигде не использовались ни свойства выпуклости функций *ϕ*(*x*), *fi*(*x*), ни свойства выпуклости множества *Rn*, ни какие-либо свойства гладкости.

Т.о., наличие седловой точки (*x*\*, *λ*\*) функции Лагранжа определяет оптимальность точки *x*\* для общей задачи математического программирования. Обратное утверждение, что из оптимальности точки *x*\* следует существование седловой точки (*x*\*, *λ*\*) функции Лагранжа, справедливо лишь для задачи выпуклого программирования при выполнении определенных ограничений относительно допустимого множества *X*.

Сформулируем эти ограничения и саму теорему, известную как *теорема Куна-Таккера*.

**Определение 1.** Рассмотрим допустимое множество

.

Если для всех *i*∈1,…,*m* существует такая точка *xi*∈*X*, что *fi*(*xi*) < 0, то говорят, что допустимое множество *X* удовлетворяет *условию регулярности*.

**Определение 2.** Пусть существует такая точка *x*∈*X*, что для всех *i*∈1,…,*m* выполняется *fi*(*x*) < 0. Тогда говорят, что допустимое множество удовлетворяет *условию регулярности Слейтера*.

Определения (1) и (2) – эквивалентны.

Действительно, из (2) ⇒ (1) – очевидно (*xi* ≡ *x*).

Пусть теперь выполнено (1).

Выберем ,

Неравенство Йенсена для выпуклых функций

тогда для ∀*i*∈1,…,*m* имеем: , *ч.т.д.*

Условие (2) означает, что существует точка внутри допустимого множества.

**Теорема** (Куна-Таккера(необходимые и достаточные условия оптимальности)). Пусть в основной задаче выпуклого программирования допустимое множество *X* обладает свойством регулярности. Тогда необходимым и достаточным условием оптимальности точки *x*\* является существование такого *λ*\*­ ≥ 0, чтобы пара (*x*\*, *λ*\*) была седловой точкой для функции Лагранжа на множестве *x*∈*Rn*, *λ*≥ 0.

**Доказательство.**

*Достаточность* доказана в теореме о седловой точке функции Лагранжа.

*Необходимость*.

Пусть *x*\* – оптимальная точка. Рассмотрим два множества в пространстве *Rm*+1:

– множество ;

– и множество , где ∀*x* .

*Множество P – выпукло.*

Действительно, пусть *z*′,*z*″∈*P* ⇒ рассмотрим *z* = *αz*′ + (1 – *α*)*z*″ ∀*α*∈[0,1] и покажем что z∈*P*.

Положим 



*Множество S – выпукло.*

Действительно, пусть ⇒

Рассмотрим

, ∀*α*∈[0,1]

По определению множества *S*:

⇒ рассмотрим *x*=*αx*′+(1–*α*)*x*″ и покажем, что .

Т.к. *ϕ* – выпуклая функция, то 

Т.к. *fi*(*x*) – выпуклая функция, то  ∀*i*

⇒ ⇒ *S* – выпукло.

Рассмотрим  – множество внутренних точек *P* и покажем, что пересечение *P*0 ∩ *S* = ∅.

* Для ∀*x*∈*X*  (оптимальность), но .
* Для ∀*x*∉*X* , но .

⇒ общих точек в множествах *P*0 и *S* – нет.

Применим к множествам *P* и *S* теорему о разделяющей гиперплоскости.

Существует разделяющая гиперплоскость, т.е. существует

 для  и .

При этом вектор , т.к. компоненты векторов из *P* неограниченны снизу.

Выберем  на границе множества  и .

Тогда получим:

 (\*)

Покажем что *u*0 ≠ 0 (тем самым покажем, что *u*0 > 0, т.к. по условию *u*0 ≥ 0).

Допустим, что *u*0 = 0, тогда (*u*,*f*(*x*)) ≥ 0 ∀*x*∈*Rn*.

При этом, поскольку , то существует индекс *i*: *ui*≠ 0, т.е. *ui*> 0.

С другой стороны, ∀*x*∈*X* *f*(*x*) ≤ 0 ⇒ для *ui*> 0 для ∀*x*∈*X* *fi*(*x*) = 0, что противоречит свойству регулярности.

Итак, *u*0 > 0, и определим .

Для этого вектора соотношение (\*) примет вид:

 (\*\*)

⇒ при *x*=*x*\* (*λ*\*, *f*(*x*\*)) ≥0.

Но т.к. *λ*\* ≥ 0, а .

Далее, для .

Собирая все вместе, получим:

,

или  ∀*λ*≥ 0, ∀*x*∈*Rn*,

т.е. (*x*\*, *λ*\*) – седловая точка функции Лагранжа, *ч.т.д.*

**Замечаниe.**

Теорема Куна-Таккера лежит в *основе теории двойственности* математического программирования. Она также находит *применение в численных методах* решения задач математического программирования. Она позволяет исходную задачу заменить задачей отыскания седловой точки функции Лагранжа, т.е. задачей вида:



"Простые" ограничения этой задачи позволяют применять для её решения методы, во многом схожие с численными методами безусловной оптимизации, достаточно хорошо изученные и апробированные.

**Двойственность.**

В формулировке теоремы Куна-Таккера прямые и двойственные переменные (*x* и *λ*) входят симметричным образом, поэтому можно ожидать, что аналогичная симметрия существует и для задач оптимизации (относительно прямых и двойственных переменных). Действительно, рассмотрим функцию

,

Очевидно, что



⇒ исходная задача  (1)

может быть представлена в виде  (1-a)

Совершенно аналогично, введем функцию  и рассмотрим задачу

 (2)

*Задача (2) называется двойственной, а задача (1) или (1-а) – прямой.*

**Теорема**(двойственность). Справедливы следующие соотношения:

1) Для всех допустимых *x* и *λ* (т.е. *x*∈*X*, *λ*≥0)

 (3)

2) Если прямая задача регулярна, *x*\* – её решение, *λ*\* – множители Лагранжа, то *λ*\* – решение задачи (2) и справедливо

 (4)

3) Если для допустимых *x*\* и *λ*\* имеет место (4), то *x*\* – решение прямой задачи, а *λ*\* – решение двойственной задачи.

**Доказательство.**

1. Если *x*∈*X*, *λ* ≥ 0, то имеем:

, *ч.т.д.*

1. Пусть *x*\* – решение задачи (1), *λ*\* – множители Лагранжа, тогда



для ∀*λ* ≥ 0, т.е. *λ*\* – решение (2), при этом, поскольку , то

, *ч.т.д.*

1. Пусть *x\**∈*X*, *λ\**≥ 0 и выполняется соотношение *ϕ*(*x*\*) = *ψ*(*λ\**), тогда рассмотрим произвольные допустимые *x*, *λ* ⇒ в силу (3) имеем



т.е. *x*\* – решение прямой задачи, *λ*\* – решение двойственной задачи, ч.т.д.

**Замечания.**

1. Можно свести задачу к другой с размерностью *m*, которая может оказаться при

*m* « *n* значительно проще.

1. Неравенство (3) позволяет получить оценку снизу для min в задаче (1) ⇒ можно оценить точность приближенного решения.

Все зависит от того, насколько просто можно вычислить 

**Двойственные задачи линейного программирования**

Рассмотрим задачу линейного программирования в основной форме:



Допустимое множество .

По теореме Куна-Таккера для задачи выпуклого программирования (а ЗЛП есть ЗВП) наличие оптимальной точки *x*\* эквивалентно наличию седловой точки (*x*\*,*λ*\*) функции Лагранжа:

, где



(*предполагаем, что условия регулярности выполняются*).

Если обозначить , то получаем *двойственную задачу*:

.

Построим двойственную задачу к исходной задаче линейного программирования, рассматривая её, *как задачу выпуклого программирования* (напомним, что там допустимое множество задается неравенством вида *fi*(*x*) ≤ 0).

Имеем (*m* + *n*)-ограничений:



Каждому ограничению сопоставим элементы *λi*, - компоненты вектора *λ* > 0.

Тогда функция Лагранжа:



Т.к. , то





Введем функцию *ψ*



Итак,

по условию 

Получаем *двойственную задачу линейного программирования*:

Целевая функция 

Допустимое множество: 

*При этом*:

* размерность исходной ЗЛП (*n*) совпадает с числом ограничений в двойственной: , и наоборот, число ограничений (*m*) в исходной ЗЛП совпадает с размерностью двойственной;
* min меняется на max, знаки неравенств меняются на противоположные.

*Справедливы следующие утверждения:*

1. *Двойственность взаимна*, т.е. задача, двойственная к двойственной – исходная.

Действительно,

рассмотрим задачу, эквивалентную двойственной:



⇒ получили ЗЛП в основной форме. Построим к ней двойственную:



⇒ эта задача эквивалентна исходной:

.

1. Если решение исходной задачи линейного программирования существует, то существует и решение двойственной ЗЛП, причем *экстремумы целевых функций совпадают* (было доказано в теореме о двойственности).
2. Экстремальная точка *λ*\* двойственной задачи является *векторным коэффициентом чувствительности* исходной задачи по вектору *b*.

Рассмотрим видоизмененную задачу с вектором правых частей *b*+Δ*b*:



Для пассивных ограничений  небольшое изменение *bi* не нарушит строгого неравенства.

При этом заметим, что из условия , которое называется условием дополняющей нежесткости, следует, что  для пассивных ограничений. Для активных ограничений  изменение *bi* может привести к большому изменению экстремума.

 *характеризует скорость изменения экстремума, т.е.* .

Действительно, для линейной задачи функция Лагранжа *L* имеет вид

, 



**Пример.**

Провести анализ чувствительности в следующей задаче оптимизации. Для изготовления изделий четырех видов  используют ресурсы трех типов, причем запасы ресурсов ограничены. Исходные данные задачи представлены в таблице.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Тип ресурсов | Расход ресурсов на изготовление одного изделия  при его стоимости | | | | Запасы ресурсов |
| *A*1 | *A*2 | *A*3 | *A*4 |
| *c*=27 | *c*=10 | *c*=15 | *c*=28 |
| I | 3 | 2 | 1 | 2 | 20 |
| II | 3 | 1 | 3 | 4 | 50 |
| III | 2 | 1 | 1 | 2 | 60 |

*Цель*: составить план выпуска изделий , обеспечивающий max стоимость произведенной продукции.

Взяв в качестве управляемых переменных  – количество выпускаемых изделий  получим следующую математическую модель:



Решив задачу симплекс-методом, найдем: *x*\* = (0,0,10,5), т.е. max стоимость произведенной продукции *ϕ*\* = 290 будет получена, если изделия *A*1 и *A*2 не выпускать, а изготовить 10 изделий *A*3 и 5 изделий *A*4.

Двойственная задача имеет вид:



Решив её, получим *y*\* = (12,1,0), *ψ­*(*y*\*) = 290.

Из этого решения видно, что при небольших приращениях Δ*b*­1 запасов ресурса I максимально достижимая стоимость изготовленной продукции *ϕ*\* изменится на величину 12Δ*b*­1

Например, если этот ресурс представляет собой сырье, то увеличение его запасов на 1 кг при оптимальном планировании, вызовет возрастание стоимости изготовленной продукции на 12 руб. То же приращение ресурсов II обеспечит увеличение объема продукции только на 1 руб. И, наконец, изменение в небольших пределах запасов ресурса III вовсе не повлияет на стоимость произведенной продукции.

Это означает, что запасы ресурса III при оптимальном плане расходуются не полностью и являются избыточными.

⇒ Наиболее дефицитным в рассматриваемой задаче является ресурс I, его запасы следует по возможности, увеличивать в первую очередь. Второй ресурс менее дефицитен, а запасы ресурса III превосходят потребности, соответствующие оптимальному плану выпуска изделий.