Содержание

1 Введение

Определение (Методы оптимизации)**.** Раздел прикладной математики, содержание которого составляет теория и методы решения оптимизационных задач

Определение (Оптимизационная задача)**.** Задача выбора наилучшего варианта (в некотором смысле) из имеющихся

Определение (Задача оптимизации)**.** $\begin{cases} f(x) \to \min(\max) \\ x \in D \end{cases}$

D - множество допустимых решений, $f : D \to \mathbb{R}$

Определение (Задача МП)**.** $\begin{cases} (1) f(x) \to \min(\max)[extr](opt) \\ (2) g_i(x) \# 0, i = 1, \dots, m - \text{ограничения} \\ (3) x \in \mathbb{R}^n \end{cases}$ $\quad x = (x_1, ..., x_n)\ f(x) : \mathbb{R}^n \to \mathbb{R},\ g_i(x) : \mathbb{R}^n \to \mathbb{R}$

Определение (Допустимое решение)**.** $x \in \mathbb{R}^n$, удовл (2), называется допустимым решением задачи.

Определение (Оптимальное решение)**.** Допустимое решение $x^* \in D$ задачи 1 - 3 называется оптимальным решением, если $f(x) \leq f(x^*) \, \forall x \in D$ в случае задачи максимизации и $f(x) \geq f(x^*) \, \forall x \in D$ в случае задачи минимизации

Глобальный оптимум - x^*

Определение (Локальный оптимум)**.** Допустимое решение $\widetilde{x} \in D$ задачи 1 - 3 называется локальным оптимумом, если $f(x) \leq f(\widetilde{x})$ для всех х из некоторой окрестности $\widetilde{x}$ в случае задачи максимизации и $f(x) \geq f(\widetilde{x})$ для всех х из некоторой окрестности $\widetilde{x}$ в случае задачи минимизации

Определение (Разрешимая/неразрешимая)**.** Задача 1 - 3, которая обладает оптимальным решением, называется разрешимой, иначе неразрешимой

2 Линейное программирование

2.1 Постановка задачи (ЛП), теоремы эквивалентности

Определение (Общая задача ЛП)**.** $\begin{cases} f(x) = c_0 + \sum_{j=1}^n c_j x_j \to \max(\min) \\ \sum_{j=1}^n a_{ij} x_j \# b_i, \, i = 1, \ldots, m \\ x_j \geq 0, j \in J \subseteq \{1, \ldots, n\} \end{cases}$, где $x = (x_1, ..., x_n) \in \mathbb{R}^n$ - вектор переменных

Матричная запись:

$$\begin{cases} f(x) = (c, x) \to \max(\min) \\ Ax\#b \\ x_j \geq 0, j \in J \subseteq \{1, \ldots, n\} \end{cases}, \; x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, b = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}, A = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \ldots & a_{mn} \end{pmatrix}$$

Определение (Стандартная (симметрическая) форма)**.** $\begin{cases} f(x) = c_0 + \sum_{j=1}^n c_j x_j \to \max \\ \sum_{j=1}^n a_{ij} x_j \leq b_i, \, i = 1, \ldots, m \\ x_j \geq 0, j = 1, \ldots, n \end{cases}$

Определение (КЗЛП)**.** $\begin{cases} f(x) = c_0 + \sum_{j=1}^n c_j x_j \to \max \\ \sum_{j=1}^n a_{ij} x_j = b_i, \, i = 1, \ldots, m \\ x_j \geq 0, j = 1, \ldots, n \end{cases}$

Определение (Основная задача ЛП)**.** $\begin{cases} f(x) = c_0 + \sum_{j=1}^n c_j x_j \to \max \\ \sum_{j=1}^n a_{ij} x_j \leq b_i, \, i = 1, \ldots, m \end{cases}$

Определение (Эквивалентные ЗЛП (ЗМП))**.** Две задачи ЛП P_1, P_2 называются *эквивалентными*, если любому допустимому решению задачи P_1 соответствует некоторое допустимое решение задачи P_2 и наоборот, причем оптимальному решению одной задачи соответствует оптимальное решение другой задачи.

Теорема 2.1 (Первая теорема эквивалентности)**.** *Для любой ЗЛП существует эквивалентная ей каноническая ЗЛП.*

Теорема 2.2 (Вторая теорема эквивалентности)**.** *Для любой ЗЛП существует эквивалентная ей симметрическая ЗЛП.*

2.2 Каноническая задача ЗЛП. Базисные решения

Определение (Базисное решение)**.** Пусть $\overline{x}$ - решение $Ax = B$. Тогда вектор $\overline{x}$ называется базисным решением СЛАУ, если система вектор-столбцов матрицы А, соответствующая ненулевым компонентам вектора $\overline{x}$, ЛНЗ

Замечание. Если система однородная, то х = $\overline{0}$ - базисное решение

Определение. Неотрицательное базисное решение СЛУ называется базисным решением канонической задачи ЛП

Определение (Вырожденное БР)**.** $\overline{x}$ - БР КЗЛП называется вырожденным, если число ненулевых компонент меньше ранга матрицы А, иначе невырожденное

Лемма 2.1. *Если x и x' - Б.Р. КЗЛП, $x \neq x'$, то*

$$J(x) \neq J(x'), J(x) \subset J(x'), J(x) \supset J(x'),$$

где $J(x) = \{j | x_j \neq 0, j = 1 \ldots n\}$

Теорема 2.3 (О конечности множества базисных решений)**.** *Число базисных решений КЗЛП конечно*

Теорема 2.4 (О существовании оптимальных БР)**.** *Если КЗЛП разрешима, то существует ее оптимальное БР*

2.3 Симплекс-метод

Рассмотрим КЗЛП.

2.3.1 Симплекс-метод для приведенной ЗЛП

Определение (Система с базисом)**.** СЛАУ - СЛАУ с базисом, если в каждом уравнении имеется переменная с коэффициентом $+1$, отсутствующая в других уравнениях. Такие переменные будем называть базисными, остальные не базисными

Определение (ПЗЛП)**.** КЗЛП называется приведенной, если

1. СЛАУ $Ax = B$ является системой с базисом
2. Целевая функция выражена через небазисные переменные

Определение (Прямо допустимая симплексная таблица)**.** СТ называется прямо допустимой, если $a_{i0} \geq 0, i = 1, \ldots, m$ (bшки)

Определение (Двойственно допустимая симплексная таблица)**.** СТ называется двойственно допустимой, если $a_{0j} \geq 0, i = 1, \ldots, n+m$ (сшки)

Теорема 2.5. *Если симплекс-таблица является прямо допустимой и $a_{0j} \geq 0, j = 1 \ldots, n+m$, то соответствующее базисное решение является оптимальным*

Теорема 2.6. *Если в симплекс-таблице существует $a_{0q} < 0, a_{iq} \leq 0, \forall i = 1 \ldots, m$, то задача неразрешима, потому что f неограничена на множестве допустимых решений*

Теорема 2.7. *Если ведущая строка выбирается из условия минимума ключевого отношения, то следующаяя симплексная таблица будет прямо допустимой*

Теорема 2.8 (Об улучшении базисного решения)**.** *Если $\exists a_{0j} < 0, j = 1 \ldots n+m$, то возможен переход к новой прямо допустимой симплекс таблице, причем $f(x) \leq f(x')$, где x - БР старой таблицы, x'- БР новой таблицы, f(x) = a_{00} старой таблицы, f(x') = $a_{00} - \frac{a_{p0}a_{0q}}{apq}, a_{p0} = 0$ - вырожденное решение*

2.4 Каноническая ЗЛП

Метод искусственного базиса

Определение (искусственные)**.** $t_i \geq 0$ - искусственные переменные

Замечание (Свойства ВЗЛП)**.**

1. ВЗЛП почти приведенная (нужно выразить t_i)
2. $h(x,t) \leq 0 \quad \forall (x,t) \in \widetilde{D}$
3. $\widetilde{D} \neq 0$ (например, есть $(0, \ldots, n, b_1, \ldots, b_m)$, n нулей)
4. ВЗЛП всегда разрешима

Теорема 2.9 (О существовании допустимого решения исходной КЗЛП)**.**

$$D \neq 0 \Leftrightarrow h^*(x,t) = 0$$

Теорема 2.10 (О преобразовании КЗЛП в эквивалентную ей приведенную)**.** *Если множество допустимых решений исходной КЗЛП непусто, то ПЗЛП, эквивалентная исходной КЗЛП, может быть получена из последней симплекс таблицы - таблицы ВЗЛП*

2.5 Двойственность в ЛП

Определение. Будем говорить, что знаки линейных ограничений ЗЛП согласованы с целевой функцией, если в задаче на max ограничения неравенства имеют вид "$\leq$ а в задаче на min ограничения на неравенство имеют вид "$\geq$"

Определение (Двойственная задача). Для ЗЛП I двойственной задачей II является ЗЛП вида:

$$f(x) = \sum_{j=1}^{n} c_j x_j \to \max \leftrightarrow g(y) = \sum_{i=1}^{m} b_i y_i \to \min,$$

$$\sum_{j=1}^{n} a_{ij} x_j \leq b_i, i = 1, \dots, l \leftrightarrow y_i \geq 0, i = 1...l,$$

$$\sum_{j=1}^{n} a_{ij} x_j = b_i, i = l+1, \dots m \leftrightarrow y_i \in \mathbb{R}, i = l+1, \dots, m,$$

$$x_j \geq 0, i = 1, \dots p \leftrightarrow \sum_{i=1}^{m} a_{ij} y_i \geq c_j, j = 1, \dots, p$$

$$x_j \in \mathbb{R}, j = p+1, \dots n \leftrightarrow \sum_{i=1}^{m} a_{ij} y_i = c_j, j = p+1, \dots, n$$

Задачу I называют прямой, а II - двойственной. Стрелки соответствуют сопряженным ограничениям

Теорема 2.11 (Основное неравенство двойственности).

$$\forall x \in D_I, \forall y \in D_{II}, f(x) \leq g(y)$$

2.6 Теоремы двойственности

Лемма 2.2 (основная лемма). *Пусть* $\forall x \in D_I \neq \varnothing, f(x) \leq M < +\infty \implies \exists y \in D_{II}\, g(y) \leq M$

Теорема 2.12 (Первая теорема двойственности). *Если одна из пары двойственных задач разрешима, то разрешима и другая, причем оптимальное значение целевых функций совпадает, т.е* $f(x^*) = g(y^*)$*, где* x^*, y^* *- оптимальные решения задач I, II соответственно*

Теорема 2.13. *Вектор* $x^* \in D_I$ *является оптимальным решением задачи I* $\Leftrightarrow \exists y^* \in D_{II}$ *т.ч* $g(y^*) = f(x^*)$

Определение (Условия дополняющей нежесткости). Будем говорить, что $x \in D_I, y \in D_{II}$ удовлетворяют УДН, если при подстановке в любую пару сопряженных неравенств хотя бы одно из них обращается в равенство. Это означает, что следующие характеристические произведения обращаются в 0:

$$(\sum_{j=1}^{n} a_{ij} x_j - b_i) y_i = 0, i = 1, \dots m$$

$$x_i (\sum_{i=1}^{m} a_{ij} y_i - c_j) = 0, j = 1, \dots n$$

Теорема 2.14 (Вторая теорема двойственности). $x^* \in D_I, y^* \in D_{II}$. *оптимальны в задачах I, II тогда и только тогда, когда они удовлетворяют УДН.*

Теорема 2.15 (Второй критерий оптимальности (следствие)). $x^* \in D_I$ *является оптимальным решением I* $\Leftrightarrow$ $\exists y^* \in D_{II}$ *т.ч.* x^* *и* y^* *удовлетворяют УДН*

2.7 Критерий разрешимости ЛП

Определение (Точная верхняя грань функции). M^* называется точной верхней гранью функции f(x) на множестве D, если

1. $\forall x \in D \quad f(x) \leq M^*$
2. $\forall M < M^* \quad \exists x \in D \quad f(x) > M$

Лемма 2.3 (О точной верхней грани функции g(y) на D_{II}). $M^* < +\infty$ *- точная верхняя грань* $f(x)$ *на* D_I*, тогда* $\forall y \in D_{II} \quad g(y) \geq M^*$

Теорема 2.16 (Критерий разрешимости). *Целевая функция задачи ЛП ограничена сверху (снизу) на непустом множестве допустимых решений тогда и только тогда, когда задача максимизации (минимизации) разрешима*

2.8 Классификация пар двойственных задач

Теорема 2.17 (Малая теорема двойственности). *Если $D_I \neq \varnothing, D_{II} \neq \varnothing \implies$ обе задачи точно разрешимы*

Теорема 2.18 (О причинах неразрешимости ЗЛП). *$D_I \neq \varnothing$, целевая функция неограничена сверху на D_I тогда и только тогда, когда II неразрешима, так как $D_{II} = \varnothing$*

Классификация

1. $D_I \neq \varnothing, D_{II} \neq \varnothing$ обе задачи разрешимы, т.к $f(x^*) = g(y^*)$
2. $D_I \neq \varnothing, D_{II} = \varnothing$ обе неразрешимы, т.к $f(x)$ неограничена, $D_{II} = \varnothing$
3. $D_I = \varnothing, D_{II} \neq \varnothing$ обе неразрешимы, т.к $D_I = \varnothing, g \to +\infty$ на D_{II}
4. $D_I = \varnothing, D_{II} = \varnothing$ обе неразрешимы

2.9 Экономическая интерпретация двойственной задачи и теорема двойственности

Экономический смысл двойственной переменной и задачи Линейные ограничения двойственной задачи связывают задачи всех ресурсов, идущих на производство 1 ед. продукции, с прибылью от продажи этой единицы продукции $\implies y_i$ измеряются в ед. стоимости

Т.к y_i соответствует ресурсам, то y_i - некая цена ресурса. Будем называть ее условной ценой (двойственной оценкой на ресурсы).

Для интерпретации двойственной задачи посмотрим на предприятие как **на продавца ресурсов.**

Задача (II) определяет справедливые цены на ресурсы, в которой требуется определить набор оценок всех ресурсов, при котором для каждого вида продукции ресурсов затрачено на производство 1 ед. продукции **не меньше** дохода от ее реализации, при этом суммарная оценка ресурсов будет минимальна

Теорема 2.19 (1). *Суммарная прибыль от продажи произведенной продукции = суммарной оценке всех ресурсов*

y_i^* - ценность i-того ресурса для производителя - доход, который может быть получен от 1 единицы использованного i-того ресурса

Теорема 2.20 (2). • *ресурс 1 и 2 расходуется полностью - их называют дефицитными - они соответствуют $y_i^* \geq 0$*

- *$x_1^* > 0, x_2^* > 0$, т.е продукция произвед. $\implies$ расходы ресурсов равны стоимости продажи этих продуктов если стоимость ресурсов, требуемых для производства 1 ед. прод > прибыль*

2.10 Анализ на чувствительность модели ЛП

Определение (Анализ чувствительности модели ЛП). Анализ чувствительности модели ЛП - исследование влияния изменения входных данных на оптимальное решение

Рассмотрим частную задачу - анализ изменения оптимального решения при изменении запаса только одного ресурса.

y_i^* рассмотрим как потенциальную возможность получить доп. доход.

Рассмотрим три задачи:

$$\begin{cases} (I) f = (c, x) \to \max \\ Ax \leq b \\ x \geq 0 \\ b = (b_1, \ldots, b_m) \end{cases}$$

$$\begin{cases} (I') f = (c, x) \to \max \\ Ax \leq b' \\ x \geq 0 \\ b' = (b_1 + \Delta b_1, \ldots, b_m), D_I \subset I' \end{cases}$$

$$\begin{cases} (\overline{I}) f = (c, x) \to \max \\ Ax \leq \overline{b} \\ x \geq 0 \\ \overline{b} = \alpha b + (1 - \alpha) b', \alpha \in (0, 1) \end{cases}$$

Определение (решения, имеющие одинаковую структуру)**.** Будем говорить, что решения $x \in D_I$ и $x' \in D_{I'}$ имеют одинаковую структуру, если

1. совпадают по ассортименту, т.е. $x_j = 0 \Leftrightarrow x_j{}' = 0 j = 1, \dots, n$
2. имеют одни и те же дефицитные ресурсы, т.е i-тое ограничение I выполняется на равенство тогда и только тогда, когда i-тое ограничение I' задачи выполняется на равенство

Лемма 2.4 (О планах одинаковой структуры)**.** *Пусть x^* - опт решение I и $x' \in D_I'$ - решение той же структуры, тогда*

1. *x' - опт решение задачи I';*
2. *для любого $\alpha \in (0, 1)$ существует оптимальное решение $\overline{I}$ имеющее эту же структуру*

Замечание. Изменять запас ресурса P_1 можно до тех пор, пока в задаче I' будет существовать оптимальный план той же структуры, что и в I

Определение (Малое (допустимое) изменение)**.** Малое (допустимое) изменение ресурса P1 - такое изменение $\Delta b_1 = b_1' - b_1$ для кот в задаче I' существует оптимальное решение той же структуры, что и оптимальное решение исходной задачи I

В силу леммы, если Δb_1 - допустимое изменение ресурса, то и все меньшие изменения также допустимы.
Пусть $F(b_1, \dots, b_m)$ - max доход, который можно получить при запасах ресурсов b_i

Определение (3-я теорема двойственности)**.** При допустимом изменении i-того ресурса приращение целевой функции прямо пропорционально изменению ресурса с коэффициентом пропорциональности, равным y_i^*

$$\Delta_i F = \Delta b_i y_i^*, \Delta_i F = F(b_1, \dots, b_{i-1}, b_i + \Delta b_i, \dots, b_m) - F(b_1, \dots, b_{i-1}, b_i, \dots, b_m)$$

2.11 О конечности симплекс-метода

Определение (вырожденная КЗЛП)**.** КЗЛП является вырожденной, если среди ее БР имеются вырожденные.

1. Если КЗЛП не является вырожденной - в процессе работы симплекс-метода $f(x_1) < \dots < f(x^*)$ (с-метод конечен)
2. $a_{p0} = 0 \implies f(x) = f(x')$ БР сохраняется, но меняется набор базисных переменных

после некоторого числа итераций возможен возврат к уже встречавшимся ранее наборам базисных переменных - с-м может зациклиться

Уточняющие правила

1. Правило Данцига - выбирается столбец
$$a_{0q} = \min_{j : a_{0j} < 0} a_{0j}$$
2. правило наибольшего приращения: выбираем такое q, при котором приращение наибольшее
$$a_{00}' = a_{00} - \frac{a_{0q} a_{p0}}{a_{pq}}$$
3. Правило Бленда

   Строка и столбец выбираются в соответствии с обычными правилами выбора, причем каждый раз из возможных выбирается переменная с наименьшим номером
4. Лексикографическое правило выбора ведущей строки
$$\frac{\overrightarrow{a_p}}{a_{pq}} = \min_{a_{iq} > 0} \frac{\overrightarrow{a_i}}{a_{iq}}$$

2.12 Двойственный симплекс-метод

мне по. чисто по.

3 Целочисленное линейное программирование

3.1 Задачи ЦЛП

Определение (Задача ЦЛП).

$$f(x) = \sum_{j=1}^{n} c_j x_j \to \max \tag{1}$$

$$\sum_{j=1}^{n} a_{ij} x_j \# b_i, i = 1, \dots, m \tag{2}$$

$$x_j \geq 0, j = 1, \dots, n \tag{3}$$

$$x_j \in \mathbb{Z}, j = 1, \dots, n \tag{4}$$

$c_j, b_i, a_{ij} \in \mathbb{Z}$ или $\mathbb{Q}$

Определение (Релаксационная задача). (1-3) - задача ЛП, которая называется соответствующей непрерывной или релаксационной задачей.

D - область допустимых решений (1 - 3), а D_Z - множество всех целочисленных точек области D.

3.2 Метод отсечения

шаг 1 Решается задача ЛП (1-3). Если она не имеет решения, то и задача ЦЛП не имеет решения. **СТОП**

шаг 2 Пусть $x^0 \in D$ - оптимальное решение задачи ЛП. Если оно из D_Z - то оно оптимальное решение задачи ЦЛП. **СТОП**

шаг 3 Строится дополнительно линейное ограничение (отсечение)

$$\sum_{j=1}^{n} \alpha_j x_j \geq \beta$$

Отсечение добавляется к задаче ЛП. После этого осуществляется возврат на шаг 1, на котором решается задача ЛП.

Определение (Правильное отсечение). Доп. ограничение - правильное, если

1. оно отсекает часть области D, содержащее нецелочисленное оптимальное решение x^0 текущей задачи ЛП.
2. В отсекаемой части области не должно быть ни одного допустимого решения задачи ЦЛП (ограничение сохраняет все допустимые целочисленные решения)

1. $\sum_{j=1}^{n} \alpha_j x_j^0 < \beta$
2. $\sum_{j=1}^{n} \alpha_j x_j \geq \beta \quad \forall x \in D_Z$

Отсечение Гомори Имеем оптимальную с-таблицу $a_{ij, i=0,\dots,m, j=0,\dots,n}$
Рассмотрим $a_{l0} \notin \mathbb{Z}$. l выбираем с наибольшей дробной частью по правилу "первая сверху"($l \in \{0, \dots, n\}$)
Дробная часть: $\{\frac{5}{4}\} = \frac{1}{4}$, $\{-\frac{5}{4}\} = \frac{3}{4}$
Дополнительное ограничение:

$$\sum_{j \in Nb} \{a_{lj}\} x_j \geq \{a_{l0}\}$$

Приводится к канон. виду и добавляется в ограничение

Теорема 3.1. *Отсечение Гомори является правильным.*

Первый алгоритм Гомори

1. Все ЗЛП решаются ЛДСМ (кроме, быть может, самой первой)
2. Специальное правило выбора производящей строки - "первая сверху"
3. Отсечение Гомори добавляется снизу к симплекс-таблице, причем таблица имеет размерность $(n+m+2) \times (n-1)$

   Применяем ЛДСМ, выбирая ведущей строку отсечения s1, после выполнения итерации строка s1 становится тривиальной - можно удалить => размер таблицы не растет

Теорема 3.2. *$D_Z \neq \varnothing$ или f ограничена снизу на D, то первый алгоритм Гомори конечен.*

3.3 Метод ветвей и границ (МВ и Г)

МВ и Г используется для решения различных классов оптимизационных задач, в основном для задач дискретной оптимизации (в которых D конечно или счетно).

Алгоритмы ветвей и границ основаны на последовательном разбиении допустимого множества решений на подмножества (ветвление) и вычислении оценок значений целевой функции на них (вычислении границ), позволяющий отбрасывать подмножества не содержащие оптимального решения, что может существенно сократить перебор.

$$\begin{cases} f(x) \to \max \\ x \in D \end{cases}$$

Определение (Стандартная задача ЦЛП)**.**

$$f(x) = \sum_{j=1}^{n} c_j x_j \to \max \tag{5}$$

$$\sum_{j=1}^{n} a_{ij} x_j \leq b_i, i = 1, \ldots, m \tag{6}$$

$$x_j \geq 0, j = 1, \ldots, n \tag{7}$$

$$x_j \in \mathbb{Z}, j = 1, \ldots, n \tag{8}$$

Определение (Верхняя оценка целевой функции)**.** Пусть $\bar{D} \subset D \implies \phi(\bar{D})$ - верхняя оценка целевой функции f(x) на $\bar{D}$, если

$$f(x) \leq \varphi(\bar{D}) \quad \forall x \in \bar{D}$$

$\bar{D} : \varphi(\bar{D}) \leq Record \implies \bar{D}$ отбрасываем как неперспективное
$\lhd$ алгоритм Лэнд и Дойг для ЗЦЛП
$\lhd$ задачу (1)-(4)
Случай D - множество дополнительных решений (1)-(4) является ограниченным:

Алгоритм Лэнд и Дойг

шаг 0 Положим $Record = -\infty$

Список задач-кандидатов на ветвление = $\varnothing$.

Решим релаксационную задачу (1)-(3) [5-7], $\bar{x}$ - ее оптимальное решение, если $\bar{x} \in Z^n$, то оптимальное решение найдено, $x^* = \bar{x}$ **СТОП.**

Иначе обозначим текущую задачу через $\bar{P}$ и объявим ее задачей для ветвления $\varphi(\bar{P}) = f(\bar{x})$, переходим на шаг 1

шаг 1 Ветвление

Определим номер k: $\bar{x_k} \in Z$

Сформируем 2 задачи

$$\begin{cases} P_1 = \bar{P} \& (x_k \leq [\bar{x_k}]) \\ P_2 = \bar{P} \& (x_k \geq [\bar{x_k}] + 1) \end{cases}$$

шаг 2 Решить P_1, аналогично P_2, например, с-методом.

Возможны следующие ситуации:

(a) $P_1(P_2)$ неразрешима - $P_1(P_2)$ исключаем из рассмотрения
если $P_1\&P_2$ обе неразрешимы - переходим на шаг 4

(b) Пусть $x^1(x^2)$ - оптимальное решение $P_1(P_2)$. Если $x^1 \in Z^n$, тогда P_1 включается в список кандидатов для ветвления $\varphi(P_1) = f(x^1)$, с x^2 аналогично. Переход на шаг 4

(c) Если $x_1 \in Z^n$ и $f(x^1) > Record \implies Record$ полагаем равным $f(x^1)$, задача P_1 исключается из рассмотрения (аналогично P_2)
Если $Record$ был изменен в П3, то на *шаг 3*, иначе на *шаг 4*

шаг 3 Исключение неперспективных множеств.

Из списка кандидатов на ветвление исключаются задачи $\bar{P}$ по правилу $\varphi(\bar{P}) \leq Record$

шаг 4 Если список кандидатов на ветвление пуст, то задача пуста. Лучшее найденное решение является оптимальным $f^* = Record$. **СТОП.**

Иначе - выбираем из списка кандидатов на ветвление задачу $\bar{P} : \varphi(\bar{P}) = \max_{p' \text{ из списка}} \varphi(p')$

$\bar{P}$ удаляется из списка кандидатов на ветвление, переход на шаг 3 с задачей $\bar{P}$

4 Выпуклое программирование

4.1 Выпуклое множество и выпуклая функция

Определение (Выпуклое множество)**.** Множество называется выпуклым, если вместе с двумя его точками оно содержит отрезок, их соединяющий, или

$$\forall x^1, x^2 \in D \quad \forall \lambda \in (0,1) \quad x^* = (1-\lambda)x^1 + \lambda x^2 \in D$$

Определение (Выпуклая функция)**.** Функция $f : D \to R$ (D - выпкуло) называется выпуклой, если

$$\forall x^1, x^2 \in D, \forall \lambda \in (0,1) \quad f((1-\lambda)x^1 + \lambda x^2) \leq (1-\lambda)f(x^1) + \lambda f(x^2)$$

Функция строго выпуклая - строгое $<$ Если неравенство $\geq$ - ф-я вогнута $>$ строго

Утверждение. 1. Пересечение выпуклых множеств выпукло.

2. Коническая комбинация выпуклых функций выпуклая.

Теорема 4.1. *Локальный минимум выпуклой функции на выпуклом множестве совпадает с глобальным минимумом*

Доказательство. Пусть $f(x)$ выпукла на D x^* - локальный минимум $f(x)$, то есть существует такая окрестность $O_{x^*} \subseteq D$ такая, что $\forall x \in O_{x^*} \quad f(x) \geq f(x^*)$. Докажем, что x^* - точка глобального минимума функции $f(x)$ на D, т.е $\forall x \in D \quad f(x^*) \leq f(x)$

от противного пусть $\exists x' \in D : f(x') < f(x^*)$. Рассмотрим отрезок x^*x'

$$\forall \lambda \in (0,1) \quad f((1-\lambda)x^* + x') \leq^{\text{выпуклость}} (1-\lambda)f(x^*) + \lambda f(x') < (1-\lambda)f(x^*) + \lambda f(x^*) = f(x^*)$$

но существует такое λ, что

$$x^\lambda = (1-\lambda)x^* + \lambda x' \in O_{x^*} \implies f(x^\lambda) \geq f(x^*) \implies_{\text{противоречие}}$$

►

4.2 Задача выпуклого программирования (ВП)

Определение (Задача ВП)**.** :

$$f(x) \to \min$$
$$\phi_i(x) \le 0, i = 1, \dots, m$$
$$x \in G$$

Здесь ϕ_i, f - выпуклые в G функции, G - выпуклое замкнутое множество $(\mathbb{R}^n, \mathbb{R}^n_+)$

Рассмотрим задачу ВП. Будем предполагать выполненным условие Слейтера **(УС)**

$$\exists \overline{x} \in G, \phi_i(\overline{x}) < 0.$$

$$D = \{x \in G | \phi_i \le 0, i = 1, \dots, m\} - \textbf{множество допустимых решений задачи ВП.}$$

УС гарантирует существование внутренних точек множества D.

Лемма 4.1 (Утверждение)**.** *Множество доп. решений ЗВП является выпуклым*

Доказательство.

$$D_i = \{x \in G | \phi_i(x) \le 0\} \implies D = \cap_{i=1}^m D_i$$

Докажем, что D_i выпукло $i = 1, \dots, m$

$$x^1, x^2 \in D_i \quad \forall \lambda \in (0,1) \quad (1-\lambda)x^1 + \lambda x^2 \in D_i$$

$$\phi_i((1-\lambda)x^1 + \lambda x^2) \le_{\phi_i}^{\text{вып для}} (1-\lambda)\phi_i(x^1) + \lambda \phi_i(x^2) \le 0 \implies$$

D_i - выпукло для всех i, поэтому D также выпукло (по свойствам выпуклых множеств) ▶

Пример (Задача размещения магазина)**.** m точек - пункты размещения магазинов.
$\{P_1, \dots, P_m\}$ - множество точек.
$P_i(x_i, y_i), \quad i = 1, \dots, m$
требуется найти точку P, суммарное расстояние которой до заданных точек минимально

$$\begin{cases} \sum_{i=1}^m \sqrt{(x-x_i)^2 + (y-y_i)^2} \to \min \\ (x, y) \in R^2 \end{cases}$$

Основные подходы к решению задач ВП

1. Модификация численных методов для задач безусловной оптимизации

   Градиентный метод
2. Обобщение метода множителей Лагранжа
3. Метод штрафных функций
4. Методы линеаризации

4.3 Свойства градиента. Идея градиентных методов

Определение. Функция $f(x_1, \dots, x_n)$, определенная в некоторой окрестности O_{x^0} называется дифференцируемой в точке x^0, если $\exists \ \nabla f(x^0)$

$$f(x) = f(x^0) + (\nabla f(x^0), x - x_0) + o(\|x - x_0\|)$$

$$\nabla f(x^0) = (\frac{\partial f}{\partial x_1}(x^0), \dots, \frac{\partial f}{\partial x_n}(x^0)) - \textbf{градиент}$$

Определение. Рассмотрим функцию $f(x)$ и $z \in \mathbb{R}^n$

Производной функции $f(x)$ в точке x_0 по направлению z называется

$$\lim_{\lambda \to 0+0} \frac{f(x^0 + \lambda z) - f(x^0)}{\lambda},$$

если он существует.

Теорема 4.2 (О градиенте и производной по направлению). *Если $f(x)$ дифференцируема в точке x^0, то предел*

$$\lim_{\lambda\to 0+0} \frac{f(x^0+\lambda z)-f(x^0)}{\lambda},$$

существует и равен

$$f'_z(x^0) = (\nabla f(x^0), z)$$

Из определения градиента следует:

1. Если $(\nabla f(x^0), z) > 0$, то при достаточно малом шаге вдоль этого направления z $f(x)$ увеличивается
2. Если $(\nabla f(x^0), z) 0 <$, то $f(x)$ уменьшается

Идея градиентных методов $x^0 \in D$ начинаем с допустимой точки Выбираем направление z, составляющее тупой угол с $\nabla f(x^0)$ и вдоль этого направления делаем достаточно малый шаг h:

1. $x^1 = x^0 + hz \in D$
2. $f(x^1) < f(x^0)$

далее действия совершаются с x^1 и так далее.

4.4 Возможные и прогрессивные направления

Задача ВП:

$$f(x) \to \min \tag{9}$$

$$\phi_i(x) \le 0, i = 1, \dots, m \tag{10}$$

$$x \in \mathbb{R}^n \tag{11}$$

ϕ_i, f- выпуклые и непрерывные дифференцируемые функции в R^n, выполнено УС.

На границе возникают проблемы...

Пусть задана точка $x_0 \in D$. $I_0 = \{i \mid \phi_i(x^0) = 0\}$ - множество индексов *активных* ограничений

Определение. Направление z называется возможным (допустимым) в x^0, если $(\nabla \phi_i(x^0), z) < 0 \quad \forall i \in I_0$

Замечание. Если z - возможное направление в x^0, то сделав достаточно малый шаг в направлении z, мы останемся в области.

Если $I_0 = \varnothing$, но любое направление является возможным.

Определение. Направление z называется прогрессивным в точке x^0, если

$$\begin{cases} (\nabla \phi_i(x^0), z) < 0 \quad \forall i \in I_0 \\ (\nabla f(x^0), z) < 0 \end{cases}$$

Замечание. Если направление z прогрессивно, то сделав достаточно малый шаг h из x^0 вдоль z, получаем: $x^1 = x^0 + hz$,

1. $x^1 \in D$
2. $f(x^1) < f(x^0)$

4.5 Критерий оптимальности

Теорема 4.3 (Критерий оптимальности). *$x^* \in D$ - оптимальное решение задачи ВП $\Leftrightarrow$ в точке x^* нет прогрессивного направления, т.е не существует $z \in R^n$:*

$$\begin{cases} (\nabla \phi_i(x^0), z) < 0 \quad \forall i \in I_0 \\ (\nabla f(x^0), z) < 0 \end{cases}$$

Доказательство.

Лемма 4.2. *пусть $f(x)$ выпукла на D и дифференцируема в точке x^** $\implies (\nabla f(x^*), x - x^*) \le \Delta f(x^*) = f(x) - f(x^*)$

Доказательство. Пусть $z = x - x^*$.

$$(\nabla f(x^*), z) = f'_z(x^*) = \lim_{\lambda \to 0+0} \frac{f(x^* + \lambda z) - f(x^*)}{\lambda} = \lim_{\lambda \to 0+0} \frac{f(x^* + \lambda(x - x^*)) - f(x^*)}{\lambda} =$$

$$= \lim_{\lambda \to 0+0} \frac{f((1-\lambda)x^* + \lambda x) - f(x^*)}{\lambda} \leq \lim_{\lambda \to 0+0} \frac{(1-\lambda)f(x^*) + \lambda f(x) - f(x^*)}{\lambda} = \lim_{\lambda \to 0+0} \frac{\lambda(f(x) - f(x^*))}{\lambda} = f(x) - f(x^*)$$

▶

$\Rightarrow$ От противного. Если в $x^* \exists$ прогрессивное направление, то в силу зам. 2, сделав достаточно малый шаг в направлении z, мы получим точку с меньшим значением целевой функции.

$\Leftarrow$ Пусть в точке x^* нет прогрессивного направления. Докажем, что x^* оптимальное решение.

Предположим, что это не так и существует $x' \in D$ $f(x') < f(x^*)$

В силу непрерывности функции $f(x)$ существует окрестность $O_{x'} : \forall x \in O_{x'}$ $f(x) < f(x^*)$

Рассмотрим $(.)\bar{x}$ - точка Слейтера, т.е. $\forall \phi_i(\bar{x}) < 0 \forall i$ и рассмотрим $[x', \bar{x}]$

Любая точка этого отрезка записывается как $x^\lambda = (1-\lambda)x' + \lambda\bar{x}\lambda \in (0,1)$

Для некоторого λ_0 соответствует точка

$$x^{\lambda_0} = (1-\lambda_0)x' + \lambda_0\bar{x} \in O_{x'},$$

т.е $f(x^{\lambda_0}) < f(x^*)$

Докажем, что x^{λ_0} является внутренней точкой D.

$$\forall i \phi_i(x^{\lambda_0}) = \phi_i((1-\lambda_0)x' + \lambda_0\bar{x}) \leq^{\text{выпуклость}} (1-\lambda_0)\phi_i(x')_{\leq 0} + \lambda_0\phi_i(\bar{x})_{<0} < 0 \implies$$

x^{λ_0} внутренняя точка

Пусть $z = x^{\lambda_0} - x^*$ Докажем, что z - прогрессивное направление в x^*

$$I_0 = \{i | \phi_i(x^*) = 0\}$$
$$\forall i \in I_0 (\nabla\phi_i(x^*), z) = (\nabla\phi_i(x^*), x^{\lambda_0} - x^*) \leq^{\text{лемма}} \phi_i(x^{\lambda_0} - \phi_i(x^*)) < 0 \implies$$

z - возможное направление

$$\leq (\nabla f(x^*), x^{\lambda_0} - x^*) \leq^{\text{лемма}} f(x^{\lambda_0})_{<0} - f(x^*)_{\geq 0} < 0 \implies$$

z - прогрессивное направление, противоречие ▶

4.6 Каноническая задача ВП

Определение. Канонической задачей ВП называется задача ВП с линейной целевой функцией, т.е $f(x) = (c, x) \to \min$

Свойства

1. $\nabla f = c = const$

2. оптимальное значение функции достигается на границе

Теорема 4.4 (эквивалентности). *Для любой задачи ВП, $\exists$ эквивалентная ей каноническая ЗВП*

Идея:

$$f(x) \to min \quad x_{n+1} \to \min$$
$$\phi_i(x) \leq 0, i = 1, \ldots, m \quad f(x) - x_{n+1} \leq 0, \phi_i(x) \leq 0, i = 1, \ldots, m$$
$$x \in G = R^n \quad x \in G = R^{n+1}$$

4.7 Метод возможных направлений(модификация Зойтендейка)

Рассмотрим каноническую задачу ВП (I)

$$f(x) = (c, x) \to \min$$
$$\phi_i(x) \le 0, i = 1, \ldots, m$$
$$x \in R^n$$

ϕ_i−выпуклая, непрерывно дифф. Выполнено условие Слейтера, x^0- начальная допустимая точка $\in D$

Задаем начальный уровень допустимой близости к границе $\delta^0 > 0$

$$I_0 = \{i | \phi_i(x^0) = 0\}$$

$$I_\delta = \{i | -\delta^0 < \phi_i(x^0) \le 0\}$$

Выпишем одну итерацию алгоритма.

шаг 0 (проверка близости x_0 к границе D)

Определяем для текущих x^0, δ^0 множества I_0, I_{δ_0} (безусловно активные и условно активные ограничения)

[На практиках отрезок $-\delta^0 \le \phi(x^0) \le 0$]

шаг 1 Нахождение прогрессивного направления z в x^0

Составляем вспомогательную задачу ЛП:

$$y \to \min$$

$$(\nabla\phi_i(x^0), z) \le y, \forall i \in I_\delta$$

$(c, z) \le y,$ **может быть и не линейное ограничение actually**

тогда просто скалярное произведение градиента и z

$|z_j| \le 1, j = 1, \ldots, n -$ условие нормировки, **зам 1**

$y \le 0,$ **зам 2**

Решаем задачу ЛП (она везде разрешима, **зам 3**)

[Замены $|z_j| \le 1 \implies u_j = x_j - 1$, тем самым получаем неотрицательную переменную]

Возможны следующие случаи.

(a)
 i. $y^0 = y^* < -\delta^0 \implies z^0$ - прогрессивное направление и на следующей итерации полагаем $\delta^1 = \delta^0$
 ii. $-\delta^0 \le y^0 < 0 \implies z^0$ - также прогрессивное, но на след итерации $\delta^1 = \delta^0/2(\delta^0$ характеризует скорость убывания)

 Переходим на шаг 2

(b) Если $y^0 = y^* = 0$, то строим аналогичную задачу ЛП, только индексы i из I_0. Если и в этом случае $y^* = 0$, то прогрессивного направления нет и текущая точка оптимальна, **СТОП**

 Если же $y^* < 0 \implies$ соотв. z - прогрессивное направление и на след. итерацию $\delta^1 = \delta^0/2$, шаг 2

шаг 2 (Поиск величины шага в прогрессивном направлении)

$$x^1 = x^0 + zh$$

h надо найти, величина шага определяется как наименьший положительный корень уравнений $\phi_i(x^0 + zh) = 0, \forall i = 1, \ldots, m$

Далее выполняем шаг величины h в направлении z, $f(x^1) < f(x^0)$, т.к шаг делает в прогрессивном направлении $\implies x^1, \delta^1$ - вход для следующей итерации, переходим к ней.

Замечания

1. $|z_j| \leq 1, j = 1, \ldots, n$. Без этого условия соответствующая задача ЛП неразрешима, z - прогрессивное направление, y - значение целевой функции. $\implies$ tz - прогрессивное направление

$$(\nabla \phi_i(x^0), tz) \leq (ty), i \in I_s$$
$$(\nabla f(x^0), tz) = ty \implies$$

ty - соотв. значение целевой функции.

$ty \to -\infty$ при $t \to +\infty \implies$ функция неограничена на D.

2. $y \leq 0$ - зачем? для удобства получения канонической задачи ЛП. Можно так сделать, так как рассматриваем нулевое решение - которое является допустимым $(0, \overline{0}) \implies y^* \leq 0$

Наличие этого условия не влияет на оптимальное значение целевой функции.

3. ВЗЛП является разрешимой, так как $D \neq \varnothing$ и целевая функция ограничена как непрерывная функция, заданная на замкнутом ограниченном множестве.

Поиск начального допустимого решения Пусть $\delta_0 > 0$ - произвольный начальный допустимый уровень.

Составляем вспомогательную задачу (она каноническая, t линейно):

$$t \to \min$$

$$\phi_i(x) - t \leq 0, i = 1, \ldots, m$$
$$x \in R^n$$
$$t \in R$$

шаг 1 Берем x^0 - произвольный вектор, $t^0 = \max_{i=1,\ldots,m} \phi_i(x^0), (t^0, x^0)$ - допустимое решение.

шаг 2 Решаем задачу вспомог. с начальным решением (t^0, x^0) до тех пор, пока $t^k \leq 0 \implies x^k$ - доп. решение исходной задачи. (такое t^k в силу выполнения условия Слейтера)

Теорема 4.5 (о сходимости метода возможных направлений). *Рассмотрим каноническую задачу выпуклого программирования, в которой $\phi_i(x), i = 1, \ldots, m$ являются непрерывными и дифференцируемы, выполнено УС, D - ограничена.*

Если есть некая последовательность точек $\{x^k\}_{k=1}^{\infty}$, получаемых методом возможных направлений (МВН), тогда

1. $\{f(x^k)\}_{k=1}^{\infty} \to f^*$
2. *Любая предельная точка $\{x^k\}_{k=1}^{\infty}$ является оптимальным решением ЗВП.*

Доказательство. Без доказательства. ▶

4.8 Теорема Куна-Таккера о седловой точке

Рассмотрим задачу выпуклого программирования следующего вида:

$$f \to \min$$

$$\phi_i(x) \leq 0, i = 1, \ldots, m$$
$$x \in G \subseteq \mathbb{R}^n$$

f, ϕ_i - выпуклые в множестве G, G - выпуклое и замкнутое. Выполнено УС.

Для этой задачи определим функцию Лагранжа

$$L(x, y) = f(x) + \sum_{i=1}^{m} y_i \phi_i(x), \text{где } y_i \geq 0, i = 1, \ldots, m, \text{ в отличие от классического метода множителей Лагранжа}$$

Определение. (x^*, y^*), где $x^* \in G, y^* \geq \overline{0}$, называется седловой точкой функции Лагранжа, если

$$L(x^*, y) \leq L(x^*, y^*) \leq L(x, y^*), \forall x \in G, \forall y \geq 0$$

Теорема 4.6 (Теорема Куна-Таккера о седловой точке). *$x^* \in G$ - оптимальное решение задачи выпуклого программирования тогда и только тогда, когда существует $y^* \geq 0$, такое что (x^*, y^*) является седловой точкой функции Лагранжа*

Доказательство. При доказательстве используем следующее утверждение

Теорема 4.7 (Теорема Фана(ударение на первую а)). *Пусть G - выпуклое, замкнутое множество в $R^n, g_i, i = 0, \ldots, m$ - выпуклые функции на G. Предположим, что система*

$$\begin{cases} g_0(x) < 0 \\ g_i(x) \leq 0, i = 1, \ldots, m \end{cases}$$

не имеет решений, тогда существует такой вектор c: $(\|c\| = 1, c_j \geq 0, i = 0, \ldots, n)$ такой, что

$$\sum_{i=0}^{m} c_i g_i(x) \geq 0, \forall x \in G$$

$\implies$ Пусть $x^* \in G$ - опт. решение ЗВП. $\implies$ $\forall x \in G f(x) \geq f(x^*)$

Докажем, что $\exists y^* \geq 0 : (x^*, y^*)$ - седловая точка функции Лагранжа

$$\begin{cases} f(x) - f(x^*) < 0 \\ \phi_i(x) \leq 0, i = 1, \ldots, m \end{cases}$$

несовместная система

Тогда по теореме Фана существует с $(\|c\| = 1, c_i \geq 0)$

$$c_0(f(x) - f(x^*)) + \sum_{i=1}^{m} c_i \phi_i(x) \geq 0, \forall x \in G$$

1. $c_0 > 0$

   c_0 не может быть равным нулю, подставляем точку Слейтера неравенство не выполняется (одно слагаемое 0, другое строго меньше 0 (из УС)).

   Значит, можно делить на c_0

2. Разделим неравенство на c_0, назовем $y_i^* = \frac{c_i}{c_0} \geq 0$

$$(f(x) - f(x^*)) + \sum_{i=1}^{m} y_i^* \phi_i(x) \geq 0, \forall x \in G$$

3. Докажем, что

$$\sum_{i=1}^{m} y_i^* \phi_i(x^*) = 0$$

   (типа условие дополняющей нежесткости)

   Подставим в последнее полученное большое неравенство $x = x^*$

$$\sum_{i=1}^{m} y_i^* \phi_i(x^*) \geq 0$$

   Но при $x = x^*$ $\phi_i(x^*) \leq 0, y_i^* \geq 0 \implies$

$$\sum_{i=1}^{m} y_i^* \phi_i(x^*) \leq 0$$

   Доказано.

4. $L(x^*, y) \leq L(x^*, y^*) \leq L(x, y^*)$

   $L(x^*, y) = f(x^*) + \sum_{i=}^{m} y_i \phi_i(x^*) \leq f(x^*) = f(x^*) + \sum_{i=1}^{m} y_i^* \phi_i(x^*)$ Левое неравенство доказано.

   Давайте правое.

   (как идет так идет................)

   $L(x^*, y^*) = f(x^*) + \sum_{i=1}^{m} y_i^* \phi_i(x^*) \leq f(x) + \sum_{i=1}^{m} y_i^* \phi_i(x) - L(x, y^*)$ доказано.

В другую сторону:
$x^* \in G, y^* \geq 0 : (x^*, y^*)$ - седловая точка
Доказать, что x^* - оптимальное решение ЗВП.
$x^* \in G, y^* \geq 0$ образуют седловую точку функции функции Лагранжа

$$L(x^*, y) \leq^{(*)} L(x^*, y^*) \leq^(**)L(x, y^*) \forall x \in G, \forall y \geq 0$$

Доказать, что x^* - опт решение (I)
Докажем, что $x^* \in G$, т.е $\phi_i(x^*) \leq 0 \forall i = 1 \dots m$
От противного $\exists k \phi_k(x^*) > 0$
Применим (*) для y = (0, ..., t, 0...), t >0, t на k-атом месте

$$f(x^*) + t\phi_k(x^*) \leq f(x^*) + \sum_{i=1}^{m} y_i^* \phi_i(x^*)$$

$f(x^*)$ - сократится

$$t\phi_k(x^*) \leq \sum_{i=1}^{m} y_i^* \phi_i(x^*)$$

$\phi_k(x^*) > 0$ при t стремящемся к бесконечности стремится к бесконечности, а сумма является константой - получаем противоречие
$\implies \phi_i(x^*) \leq 0 \; \forall i = 1, \dots, m \implies x^* \in D$
Докажем, что $\sum_{i=1}^{m} y_i^* \phi_i(x^*) = 0$
Подставим в (*) y =0, $f(x^*) \leq f(x^*) + \sum_{i=1}^{m} y_i^* \phi_i(x^*) \implies \sum_{i=1}^{m} y_i^* \phi_i(x^*) \geq 0$
но

$$\phi_i(x^*) \leq 0, \forall i, y_i^* \geq 0 \implies \sum_{i=1}^{m} y_i^* \phi_i(x^*) \leq 0 \implies \sum_{i=1}^{m} y_i^* \phi_i(x^*) = 0$$

Докажем, что для любого x из D $f(x) \geq f(x^*)$
Рассмотрим (**)

$$f(x^*) + \sum_{i=1}^{m} y_i^* \phi_i(x^*) \leq f(x) + \sum_{i=1}^{m} y_i^* \phi_i(x)$$

$$y_i^* \geq 0, \phi_i(x) \leq 0 \implies \sum_{i=1}^{m} y_i^* \phi_i(x) \leq 0 \implies$$

$$f(x^*) \leq f(x) \implies x^* - \textbf{оптимальное}$$

▶

4.9 Теорема Куна-Таккера в дифференциальной форме

4.9.1 Дифференциальная форма 1

Рассмотрим задачу ВП (I)

$$f(x) \to \min$$

$$\phi_i(x) \leq 0, i = 1, ..., m$$

$$x_j \geq 0, j = 1, ..., n$$

$$G = \{x \in R^n, x_j \geq 0, j = 1, ..., n\}$$

f, ϕ_i - непрерывно дифференцируемые на G
Выполнено УС

$$L(x, y) = f(x) + \sum_{i=1}^{m} y_i \phi_i(x), y_i \geq 0, i = 1, \dots, m$$

Обозначим

$$\nabla_x L(x^*, y^*) = (\frac{\partial L}{\partial x_1}, ..., \frac{\partial L}{\partial x_n})|_{(x^*, y^*)}$$

$$\nabla_y L(x^*, y^*) = (\frac{\partial L}{\partial y_1}, ..., \frac{\partial L}{\partial y_m})|_{(x^*, y^*)} = (\phi_1(x^*), \dots, \phi_m(x^*))$$

Теорема 4.8 (Куна-Таккера в дифференциальной форме 1)**.** *Точка $x^* \geq 0$ является оптимальным решением задачи (I) тогда и только тогда, когда существует $y^* \geq 0$ такой, что выполняются следующие условия:*

1. $\nabla_x L(x^*, y^*) \geq 0$, *т.е.* $\frac{\partial L(x,y^*)}{\partial x_j}|_{x=x^*} \geq 0, \forall j = 1, \ldots, n$

2. $(x^*, \nabla_x L(x^*, y^*)) = 0$ *т.е* $\sum_{i=1}^n x_j^* \frac{\partial L(x,y^*)}{\partial x_j}|_{x^*} = 0$

3. $\nabla_y L(x^*, y^*) \leq 0$, *т.е* $\phi_i(x^*) \leq 0, i = 1, ..., m$

4. $(y^*, \nabla_y L(x^*, y^*)) = 0$, *т.е* $\sum_{i=1}^m y_i^* \phi_i(x^*) = 0$

Доказательство. Докажем слева направо

$x^* \geq 0$ - оптимальное решение. По теореме Куна Таккера существует $y^* \geq 0$

$$L(x^*, y) \leq L(x^*, y^*) \leq L(x, y^*), \forall x \geq 0, \forall y \leq 0$$

Для каждого $j = 1, \ldots, n$ рассмотрим точку $x^j = (x_1^*, \ldots, x_{j-1}^*, x_j^*, \ldots, x_n^*)$. Из (**) следует, x^j - глобальный минимум функции $L(x^j, y^*)$ - функция от одной переменной на множестве $\{x_j \in R : x_j \geq 0\}$

Условия 1), 2) теоремы - необходимые условия локального минимума на неотрицательной полуоси

$x_j^* > 0, \frac{\partial L}{\partial x_j}|_{(x^*,y^*)}$

$x_j^* = 0 \frac{\partial L}{\partial x_j}|_{(x^*,y^*)} \geq 0$

По аналогии из (*) получаются неравенства 3), 4) теоремы как необходимые условия локального максимума

Справа налево: Пусть выполняются условия 1)-4) для некоторого $x^* \geq 0, y^* \geq 0$

Докажем, что (x^*, y^*) - седловая точка функции Лагранжа

При фиксированном $y^* \geq 0$ функция Лагранжа

$$L(x, y^*) = f(x) + \sum_{i=1}^m y_i^* \phi_i(x)$$

- выпуклая функция

По лемме о приращении и градиенте для выпуклой функции

$$\forall x \geq 0, L(x, y^*) - L(x^*, y^*) \geq (\nabla_x L(x^*, y^*), x - x^*) = (\nabla_x L(x^*, y^*), x) - (\nabla_x L(x^*, y^*), x^*)$$

$$(\nabla_x L(x^*, y^*), x^*) = 0$$

по 2 условию из Теоремы

$$(\nabla_x L(x^*, y^*) \geq 0, \text{по } 1, x \geq 0)$$
$$(\nabla_x L(x^*, y^*), x) - (\nabla_x L(x^*, y^*), x^*) \geq 0$$

доказали (**)

$$L(x^*, y) - L(x^*, y^*) = f(x^*) + \sum_{i=1}^m y_i \phi_i(x^*) - f(x^*) - \sum_{i=1}^m y_i^* \phi_i(x^*) \leq 0 \implies$$

вып (*)

[Применили 3 и 4 для для слагаемых-сумм, последняя сумма равна 0, а у предпоследней $y_i \geq 0, \phi_i \leq 0$] ▶

4.9.2 Дифференциальная форма 2

Рассмотрим задачу ВП (II)

$$f(x) \to \min$$
$$\phi_i(x) \leq 0, i = 1, ..., m$$
$$x \in R^n$$
$$G = R^n$$

f, ϕ_i - непрерывно дифференцируемые на G

Выполнено УС

Теорема 4.9 (Куна-Таккера в дифференциальной форме 2)**.** *Точка $x^* \in R^n$ является оптимальной точкой задачи (II) тогда и только тогда, когда существует $y^* \geq 0$ такое, что выполняются условия*

1. $\nabla_x L(x^*, y^*) = 0$, *т.е.* $\frac{\partial L(x,y^*)}{\partial x_j}|_{x=x^*} = 0, \forall j = 1, \ldots, n$
2. $\nabla_y L(x^*, y^*) \leq 0$, *т.е* $\phi_i(x^*) \leq 0, i = 1, ..., m$
3. $(y^*, \nabla_y L(x^*, y^*)) = 0$, *т.е* $\sum_{i=1}^m y_i^* \phi_i(x^*) = 0$, *или* $\forall i : y_i^* \phi_i(x^*) = 0$

Доказательство. самостоятельно ▶

5 Квадратичное программирование

5.1 Квадратичная функция

Определение (Квадратичная форма)**.** Квадратичной формой от n переменных называется функция вида

$$g(x) = (cx, x) = \sum_{i=1}^{n} \sum_{j=1}^{n} c_{ij} x_i x_j$$

, где C - симметричная матрица, $c_{ij} = c_{ji}$

Диагональные элементы - коэффициенты при квадратах переменных, а остальные - половине коэффициентов

Пример.

$$g(x_1, x_2) = 3x_1^2 - 3x_1x_2 + 4x_2^2, C = \begin{pmatrix} 3 & -\frac{3}{2} \\ -\frac{3}{2} & 4 \end{pmatrix}$$

Определение. Квадратичная форма называется неотрицательно определенной, если $\{\forall x \in R^n | x \neq 0\}$ выполняется $g(x) \geq 0$, и положительно определенной если $\{\forall x \in R^n | x \neq 0\}$ выполняется $g(x) > 0$

Способы проверки неотрицательной/положительной определенности

1. Критерий Сильвестра

   Квадратичная форма является положительно определенной, если ее главные миноры положительны, и неотрицательно определенной, если ее главные миноры неотрицательны

2. Через собственные числа

   (a) Квадратичная форма (Cx, x) положительно определена, если корни характеристического многочлена ($|C - \lambda E| = 0$) положительны, и неотрицательно определена, если корни характеристического многочлена неотрицательны, и хотя бы 1 корень равен 0.

Определение (Квадратичная функция)**.** Функция вида $f(x) = \sum_{i=1}^n \sum_{j=1}^n c_{ij} x_i x_j + \sum_{i=1}^n p_i x + p_0 = (Cx, x) + (p, x) + p_0, p \in R^n, p_0 \in R$

Теорема 5.1. *Квадратичная функция f(x) выпукла тогда и только тогда, когда квадратичная форма (Cx, x) неотрицательно определена*

Доказательство. Так как

$$(p, x) + p_0$$

не нарушает выпуклости, будем доказывать только для нелинейной части, т.е для g(x) = (Cx, x).

Лемма 5.1. *Для любых $x^1, x^2 \forall \lambda \in (0, 1)$*

$$(1-\lambda)g(x^1) + \lambda g(x^2) - g((1-\lambda)x^1 + \lambda x^2) = \lambda(1-\lambda)g(x^1 - x^2)$$

$$(1-\lambda)(Cx^1, x^1) + \lambda(Cx^2, x^2) - (C((1-\lambda)x^1 + \lambda x^2), (1-\lambda)x^1 + \lambda x^2) =$$

$$= (1-\lambda)(Cx^1, x^1) + \lambda(Cx^2, x^2) - (1-\lambda)^2(Cx^1, x^1) - (1-\lambda)\lambda(Cx^1, x^2) - \lambda(1-\lambda)(Cx^2, x^1) - \lambda^2(Cx^2, x^2) =$$

$$= (1-\lambda)(\lambda)(Cx^1, x^1) - \lambda(1-\lambda)(Cx^2, x^2) + \lambda(1-\lambda)(Cx^2, x^1) - \lambda(1-\lambda)(Cx^1, x^2) =$$

$$\lambda(1-\lambda)((Cx^1, x^1 - x^2) - C(x^2, x^1 - x^2)) = \lambda(1-\lambda)g(x^1 - x^2)$$

$\Rightarrow$ Если функция g(x) выпукла, то левая часть $\lambda(1-\lambda)g(x^1 - x^2) \geq 0 \implies g(x^1 - x^2) \geq 0$ для любых x^1, x^2

Любой вектор $x \in R^n$ можно представить как разность $x^1, x^2 \implies$ неопределенно определена

$\Leftarrow$ C неотрицательно определена, аналогично $\lambda(1-\lambda)g(x^1 - x^2) \geq 0 \implies$ левая часть выражения в лемме больше, либо равна $0 \implies g(x)$ ▶

5.2 Задача квадратичного программирования

Определение. Задача $f(x) = \sum_{i=1}^{n}\sum_{j=1}^{n} c_{ij}x_ix_j + \sum_{j=1}^{n} p_jx_j + p_0 \to \min$

$$\sum_{j=1}^{n} a_{ij}x_j \leq b_i, i = 1, \dots, m$$

$$x_j \geq 0, j = 1, \dots, n$$

называется задачей квадратичного программирования

В матричном виде

$$f(x) = (Cx, x) + (p, x) + p_0$$

$$Ax \leq b$$

$$x \geq 0$$

Замечание (1). Далее $p_0 = 0$

Замечание (2). $c \geq 0$ задача кв. программирования является задачей выпуклого программирования, верна теорема Куна-Таккера (в дифференциальной форме 1, так как x из $\mathbb{R}^n_+$)

Теорема 5.2. *Вектор $x^* \geq 0$ является оптимальным решением задачи КП $\Leftrightarrow$ $\exists$ неотрицательные векторы $y^*, u^* \in R^m$ и $v^* \in R^n$ что выполняется*

1. $2Cx^* + A^Ty^* - v^* = -p$
2. $Ax^* + u^* = b$
3. $(x^*, v^*) = 0$
4. $(y^*, u^*) = 0$

Доказательство. Рассмотрим задачу КП как задачу ВП

$$f = (Cx, x) + (p, x) \to \min$$

$$Ax - b \leq 0$$

$$x \geq 0$$

$$L(x, y) = (Cx, x) + (p, x) + (y, Ax - b) \forall y \geq 0, \forall x \geq 0$$

- функция Лагранжа

Используем теорему Куна-Таккера в диф. форме 1:

Точка $x^* \geq 0$ является оптимальным решением задачи (I) тогда и только тогда, когда существует $y^* \geq 0$ такой, что выполняются следующие условия:

1. $\nabla_x L(x^*, y^*) \geq 0$, т.е. $\frac{\partial L(x,y^*)}{\partial x_j}|_{x=x^*} \geq 0, \forall j = 1, \dots, n$
2. $(x^*, \nabla_x L(x^*, y^*)) = 0$ т.е $\sum_{i=1}^{n} x_j^* \frac{\partial L(x,y^*)}{\partial x_j}|_{x^*} = 0$
3. $\nabla_y L(x^*, y^*) \leq 0$, т.е $\phi_i(x^*) \leq 0, i = 1, ..., m$
4. $(y^*, \nabla_y L(x^*, y^*)) = 0$, т.е $\sum_{i=1}^{m} y_i^* \phi_i(x^*) = 0$

$\implies$ применим условия 1 и 3 к нашей задаче и приведем их к равенствам

1. $2Cx^* + A^Ty^* - v^* = -p$
2. $Ax^*u^* = b$

Применим условие 2 и 4 к нашей задаче

$\nabla_x L(x, y)|_{(x^*,y^*)} = 2Cx^* + A^Ty^* + p = v^*$

$2)(x^*, v^*) = 0$

$\nabla_y L(x, y)|_{(x^*,y^*)} = Ax^* - b = -u^*$

$4)(y^*, -u^*) = 0 \implies (y^*, u^*) = 0$

Условие теоремы эквивалентно условиям теоремы Куна-Таккера ►

Вывод: для решения задачи квадратичного программирования достаточно найти неотрицательные векторы x^*, y^*, u^*, v^* :

1. $2Cx^* + A^T y^* - v^* = -p$
2. $Ax^* + u^* = b$
3. $(x^*, v^*) = 0$, т.е $x_j^* v_j^* = 0 \; \forall j = 1, \dots, n$
4. $(y^*, u^*) = 0$ т.е $y_i^* u_i^* = 0 \; \forall i = 1, \dots, m$

Замечание. 1. 1, 2 - СЛУ n+m уравнений, 2n+2m переменных

Теорема 5.3. *Если ∃ неотрицательное решение системы 1-2, уд. 3-4, то существует и базисное неотрицательное решение этой же системы, удовлетворяющее тем же свойствам*

Доказательство. ▶

5.3 О методах решения задач квадратичного программирования

Все методы основаны на методах решения системы 1 - 4

1. метод Баранкина-Дорфмана
2. Франка-Вулфа
3. метод Вулфа

в 1-2 с помощью некоторого базиса строится некоторое базисное решение 1-2, которое не обязательно удовлетворяет 3-4, а затем с помощью вспомогательных симплексных преобразований добиваются выполнения условий 3-4,

а в 3 с помощью введения искусственных переменных строится вспомогательная ЗЛП, которая решается симплекс методом, при этом начальное базисное решение удовлетворяет 3-4, и далее решаем эту задачу, следя за выполнением 3-4 на каждой итерации (если C положительно определена, то метод сходится)