Коллоквиум по предмету **«Дискретная математика»**
ВШЭ ФКН ПМИ, 2 курс, декабрь 2017г.
Основано на реальных событиях.

Вопросы по линейному программированию

1. Определение задачи линейного программирования

Задача линейного программирования состоит в нахождении максимума или минимума линейной функции при ограничениях, задаваемых линейными равенствами и неравенствами на переменные, которые пробегают действительные числа. Все коэффициенты также действительные числа.

$$\begin{cases} \text{линейная функция} \to \max\,,\ \min\,, \\ \text{лин. ф.} \leqslant 0 \\ \qquad\quad \geqslant 0 \\ \qquad\quad = 0, \\ \qquad\quad \vdots \\ \text{лин. ф.} \leqslant 0 \\ \qquad\quad \geqslant 0 \\ \qquad\quad = 0. \end{cases}$$

Размерностью называется количество пепременных в задаче. **Допустимые решения** – значения переменных, удовлетворяющие всем ограничениям задачи ЛП. **Целевая функция** – функция, которую необходимо оптимизировать.

У целевой функции либо:

— есть допустимые решения, тогда либо:

— целевая функция ограничена: $\begin{cases} x \leqslant 1, \\ x \to \max\,. \end{cases}$

— целевая функция не ограничена: $\begin{cases} x \geqslant 1, \\ x \to \max\,. \end{cases}$

— нет допустимых решений: $\begin{cases} x \leqslant 0, \\ x \geqslant 1 \end{cases}$

2. Три специальных вида задач ЛП: одни неравенства, одни неравенства на неотрицательные переменные, одни равенства на неотрицательные переменные. Приведение любой задачи ЛП к задаче любого из трех специальных видов

Первый вид: *ограничения состоят только из неравенств*. Общую задачу ЛП всегда можно привести к задаче такого вида, поскольку равенство $A = B$ равносильно системе двух неравенств $A \leqslant B$ и $B \leqslant A$. При сведении к этому виду размерность не изменяется, а количество ограничений возрастает не более, чем вдвое: задача ЛП $\begin{cases} x - z \to \max\,, \\ x + y = z + u, \\ y \leqslant x, \\ z \leqslant u. \end{cases}$ равносильна задаче $\begin{cases} x - z \to \max\,, \\ x + y \leqslant z + u, \\ z + u \leqslant x + y, \\ y \leqslant x, \\ z \leqslant u. \end{cases}$

Второй вид: *ограничения состоят только из неравенств, при этом все переменные пробегают неотрицательные числа*. Этот вид является подвидом первого. Чтобы привести задачу ко второму виду, сначала добьемся, чтобы все ограничения были неравенствами (как в первом пункте), а затем заменим в полученной задаче каждую переменную на разность двух новых переменных, после чего наложим требование неотрицательности на все новые переменные. Поскольку каждое действительное число может быть представлено как разность двух неотрицательных чисел, новая задача равносильна исходной. При сведении к этому виду может возрасти как количество ограничений, так и количество переменных. То есть задача ЛП $\begin{cases} x - y \to \max\,, \\ 2x + 1 \leqslant y, \\ x \leqslant y - 1. \end{cases}$ равносильна задаче $\begin{cases} (x_1 - x_2) - (y_1 - y_2) \to \max\,, \\ 2(x_1 - x_2) + 1 \leqslant y_1 - y_2, \\ (x_1 - x_2) \leqslant (y_1 - y_2) - 1, \\ x_1, x_2, y_1, y_2 \geqslant 0. \end{cases}$

Третий вид: *ограничения состоят только из равенств, и при этом все переменные пробегают неотрицательные числа*. Зная, как переходить из первого ко второму, опишем сведение из второго к третьему. Заменимм каждое неравенство $A \leqslant B$ на равенство $A + y = B$, где y новая переменная (своя для каждого неравенства) и добавим требование неотрицательности для каждой добавленной переменной. Поскольку неравенство $A \leqslant B$ эквивалентно существованию неотрицательного числа y, для которого $A + y = B$, полученная ЛП равносильна исходной. Добавленные переменные называют **слабыми**. Задача ЛП $\begin{cases} x - z \to \max, \\ 2z \leqslant x + y, \\ x \leqslant y, \\ x, y, z \geqslant 0. \end{cases}$ равносильна задаче $\begin{cases} x - z \to \max, \\ 2z + u = x + y, \\ x + v = y, \\ x,y,z,u,v \geqslant 0. \end{cases}$

3. Полиэдры, грани полиэдров

Полиэдр в n-мерном пространстве – множество, задаваемое линейными неравенствами в переменных $x_1, \ldots, x_n$.

Пусть полиэдр в n-мерном пространстве задается m ограничениями вида $a_i x \leqslant b_i$ (a_i – n-мерный вектор, b_i – действительное число, $a_i x$ – скалярное произведение вектора a_i и вектора неизвестных $x_1, \ldots, x_n$). Выберем некоторое множество индексов $I \subset \{1, \ldots, m\}$ и рассмотрим множество

$$F = \{x \mid a_i x = b_i \text{ для всех } i \in I, a_i x \leqslant b_i \text{ для всех остальных } i\}.$$

Непустые множества такого вида называются **гранями** полиэдра. В частности, весь полиэдр является своей гранью (для пустого множества индексов I).

4. Достижимость оптимума в любой задаче линейного программирования (доказательство методом исключения переменных)

Теорема. Если значения целевой функции в задаче ЛП ограничены, то оптимальное значение достигается.

Доказательство. Пусть дана задача максимизации некоторой линейной функции $c_1 x_2 + \ldots + c_n x_n$ на полиэдре П, заданном системой линейных ограничений на переменные $x_1, \ldots, x_n$. Докажем, описав алгоритм нахождения оптимума задачи с использованием метода исключения переменных.

Алгоритм. Введем новую переменную x_0 и рассмотрим новый полиэдр П$'$ в $n{+}1$-мерном пространстве, задаваемый исходными ограничениями вместе с новым ограничением $x_0 = c_1 x_1 + \ldots + c_n x_n$. Значения x_0 в точках этого полиэдра – это в точности значения целевой функции $c_1 x_1 + \ldots + c_n x_n$ в точках исходного полиэдра П.

Исключим в системе неравенств, задающей П$'$, все переменные, кроме x_0. В результате мы получим систему линейных неравенств на переменную x_0, которая и задает все возможные значения целевой функции $c_1 x_1 + \ldots + c_n x_n$ на точках исходного полиэдра П. Все полученные неравенства нестрогие и задают либо отрезок, либо луч, либо все множество действительных чисел. Если это множество ограничено сверху, оно является отрезком вида $[e,\ d]$ или лучом вида $(-\infty,\ d]$. В этом случае d и является искомым максимумом.

Пример. Продемонстрируем алгоритм на конкретном примере:

$$\begin{cases} x + y \to \max, \\ 2x + y \leqslant 4, \\ -4 \leqslant 2x + 5y, \\ 3y \leqslant 2x + 4. \end{cases} \xrightarrow{\text{вводим переменную } t} \begin{cases} t = x + y, \\ 2x + y \leqslant 4, \\ -4 \leqslant 2x + 5y, \\ 3y \leqslant 2x + 4. \end{cases} \xrightarrow{\text{исключаем } x} \begin{cases} 2(t - y) + y \leqslant 4, \\ -4 \leqslant 2(t - y) + 5y, \\ 3y \leqslant 2(t - y) + 4. \end{cases} \longrightarrow$$

$$\xrightarrow{\text{разрешаем относительно } y} \begin{cases} 2t - 4 \leqslant y, \\ -\frac{2}{3}t - \frac{4}{3} \leqslant y, \\ y \leqslant \frac{2}{5}t + \frac{4}{5}. \end{cases} \xrightarrow{\text{исключаем } y} \begin{cases} 2t - 4 \leqslant \frac{2}{5}t + \frac{4}{5}, \\ -\frac{2}{3}t - \frac{4}{3} \leqslant \frac{2}{5}t + \frac{4}{5}. \end{cases} \longleftrightarrow \begin{cases} t \leqslant 3, \\ -2 \leqslant t. \end{cases}$$

Таким образом получаем, что наибольшее значение целевой функции равно 3. Отсюда легко найти оптимальные значения, на которых оно достигается: $x = 1$ и $y = 2$. ∎

5. Задача ЛП, двойственная к данной задаче

Двойственной к задаче ЛП, состоящей в максимизации (минимизации) целевой функции cx, называется задача нахождения наименьшего (наибольшего) числа d, для которого неравенство $cx \leqslant d$ ($cx \geqslant d$) является синтаксическим следствием ограничений этой задачи.

Правил построения двойственной задачи достаточно много и их долго описывать, да и вряд ли это потребуется на коллоке, но почитать про это дело можно здесь на страницах 18-21.

Визуализируем симмитричность отношения двойственности:

$$\begin{cases} cx \to \max \\ a_1x \leqslant b_1,\ [y_1] \\ a_2x \geqslant b_2,\ [y_2] \\ a_3x = b_3,\ [y_3] \\ x_1 \leqslant 0,\ [\text{смена знака}] \\ x_2 \geqslant 0,\ [\text{сохранение знака}] \\ x_3 \in \mathbb{R}.\ [\text{не влияет на знак}] \end{cases} \xrightarrow{\text{двойственная}} \begin{cases} by \to \min \\ y_1 \geqslant 0,\ [\text{сохранение знака}] \\ y_2 \leqslant 0,\ [\text{смена знака}] \\ y_3 \in \mathbb{R},\ [\text{не влияет на знак}] \\ d_1y \leqslant c_1,\ [x_1] \\ d_2y \geqslant c_2,\ [x_2] \\ d_3y = c_3.\ [x_3] \end{cases} \xrightarrow{\text{двойственная}} \begin{cases} cx \to \max \\ a_1x \leqslant b_1, \\ a_2x \geqslant b_2, \\ a_3x = b_3, \\ x_1 \leqslant 0, \\ x_2 \geqslant 0, \\ x_3 \in \mathbb{R}. \end{cases}$$

Покажем, почему двойственная задача также является задачей линейного программирования. Пусть исходная задача состоит в максимизации целевой функции cx. Любое неравенство вида $cx \leqslant d$, которое можно вывести из её ограничений, можно получить, складывая их с некоторыми коэффициентами. Нам нужно подобрать эти коэффициенты, которые будут неизвестными в новой задаче так, чтобы после сложения исходных неравенств с этими коэффициентами получилось неравенство $cx \leqslant d$, в левой части которого стоит целевая функция, а свободный член d в правой части был как можно меньше. Первое требование выражается в виде линейных ограничений на новые переменные. К ним ещё нужно добавить требования неотрицательности (или неположительности) тех коэффициентов, на которые умножались неравенства. А правая часть полученного неравенства является линейной комбинацией новых переменных с коэффициентами, являющимися правыми частями исходных ограничений (мы предполагаем, что исходные ограничения записаны в стандартной форме).

6. Принципы двойственности в линейном программировании: первый принцип (критерий совместности) и второй принцип (совпадение оптимума в прямой и двойственной задачах). Примечание: надо знать два доказательства обоих принципов двойственности – доказательство методом исключения переменных и геометрическое доказательство с помощью леммы Фаркаша

Теорема (Первый принцип двойственности, критерий совместности). Если система линейных ограничений несовместна, то из неё можно вывести неравенство $0 \leqslant -1$.

Доказательство (аналитическое). Пусть система, состоящая из неравенств $a_ix \leqslant b_i$ несовместна. Нам нужно вывести из неё некоторое заведомо ложное неравенство. Для этого применим к ней алгоритм проверки совместности методом исключения переменных. Поскольку система несовместна, в некоторый момент мы получим заведомо ложное неравенство. Заметим, что все новые неравенства, возникающие в ходе исключения переменных, являются синтаксическими следствиями исходных неравенств. В самом деле, в методе исключения переменных мы использовали шаги следующего вида: для некоторой переменной x_j мы брали два равенства вида $s \leqslant qx_j$ и $px_j \leqslant r$ (где p, q положительные числа, а r, s – линейные функции от других переменных), разрешали их относительно x_j, получая неравенства $x_j \leqslant \frac{r}{p}$ и $\frac{s}{q} \leqslant x_j$ и затем составляли неравенство $\frac{s}{q} \leqslant \frac{r}{p}$. То же самое равенство можно получить, сложив неравенства $px_j \leqslant r$ и $s \leqslant qx_j$ с положительными коэффициентами $\frac{1}{p}$ и $\frac{1}{q}$ (переменная x_j сократится). Поэтому неравенство $\frac{s}{q} \leqslant \frac{r}{p}$ является синтаксическим следствием равенств $px_j \leqslant r$ и $s \leqslant qx_j$. Значит, все неравенства, появляющиеся в ходе исключения переменных, являются синтаксическими следствиями исходных неравенств. В частности и заведомо ложное неравенство, возникшее в ходе исключения переменных, является синаксическим следствием исходных неравенств, что и требовалось доказать. ■

Доказательство (геометрическое). Пусть дана несовместная система из m линейных ограничений первого специального вида (все неравенства). Сначала с помощью введения слабых переменных перепишем всё в виде равенств, т.е. $a_ix \leqslant b \Longrightarrow a_ix + z_i = b_i,\ z_i \geqslant 0$. После этого заменим каждую переменную x_j на разность новых неотрицательных переменных $u_j - v_j$, получив систему ограничений, состоящую только из равенств в неотрицательных переменных. Поскольку исходная система несовместна, то новая также несовместна. Обозначим через A матрицу исходной системы ограничений, а через A' – матрицу новой системы.

Пусть исходная система ограничений имеет вид $\begin{cases} x + 2y = 3, \\ 4x + 5y \leqslant 6, \\ 7x + 8y \leqslant 9. \end{cases}$ Её матрица равна $A = \begin{pmatrix} 1 & 2 \\ 4 & 5 \\ 7 & 8 \end{pmatrix}$. Тогда новая система ограничений есть $\begin{cases}(u-v)+2(r-s)=3,\\4(u-v)+5(r-s)+a=6,\\7(u-v)+8(r-s)+b=9.\end{cases}$ Её матрица $A'=\begin{pmatrix}1&-1&2&-2&0&0\\4&-4&5&-5&1&0\\7&-7&8&-8&0&1\end{pmatrix}$.

По лемме Фаркаша существует y, для которого $yA'\geqslant 0$ и $yb<0$. Докажем, что $yA=0$. Для этого рассмотрим i-ую координату вектора yA. Она получается умножением y на i-ый столбец матрицы A. Этот столбец входит в матрицу A' дважды – один раз с плюсом, а другой раз с минусом. Поскольку все координаты вектора yA' неотрицательны, произведение y и i-го столбца матрицы A должно быть как неотрицательно, так и неположительно, а значит равно нулю. Следовательно, $yA=0$. Таким образом, складывая исходные ограничения с коэффициентами $y_1,\dots,y_m$, мы получим неравенство, у которого в левой части все переменные сократились, а в правой части стоит отрицательное число (напомним, что $yb<0$).

Осталось доказать, что такое сложение выведет неравенство вида $\leqslant$ (оно будет заведомо ложным). Для этого нужно установить, что $y_i\geqslant 0$. Для этого рассмотрим i-е ограничение $a_ix\leqslant b_i$. После введения слабой переменной оно превратилось в ограничение вида $a_ix+z_i=b_i$. Значит в матрице A' имеется столбец (соответствующей переменной z_i), содержащий нули везде, кроме i-й строки, в которой стоит 1. Поэтому соответствующая этому столбцу координата вектора-строки yA' равна y_i, следовательно $y_i\geqslant 0$. ∎

Теорема (Второй принцип двойственности). Пусть задача ЛП, состоящая в максимизации (минимизации) линейной функции cx при некоторых линейных ограничениях, имеет допустимое решение и значение целевой функции cx ограничено сверху (снизу) при этих ограничениях. Тогда двойственная задача также имеет допустимое решение и оптимальные значения целевых функций в прямой и двойственной задаче совпадают.

Доказательство (аналитическое). Пусть дана задача максимизации линейной функции cx на полиэдре, заданном некоторыми линейными неравенствами $a_ix\leqslant b_i$, $i=1,\dots,m$, причем функция cx ограничена сверху при этих ограничениях. Пусть d обозначает её наибольшее значение. Наша цель – вывести из исходных ограничений неравенство $cx\leqslant d$.

Введем новую переменную x_0 и добавим в систему ограничений новое ограничение $x_0=cx$. Затем исключим из полученной системы все переменные, кроме x_0. В конце мы получим совместную систему ограничений на значение переменной x_0. Некоторые из них ограничивают x_0 сверху конкретным числом (т.к. остальные переменные мы исключили), другие ограничивают x_0 снизу, а третьи вообще не содержат x_0 и являются заведомо истинными неравенствами. При этом должно быть хотя бы одно неравенство первого типа ($x_0\leqslant const$), поскольку иначе x_0 могло бы принимать сколь угодно большие значения. Наибольшее возможное значение x_0, которое мы обозначили ранее через d, равно минимальной из таких верхних оценок x_0. Следовательно, неравенство $x_0\leqslant d$ появилось в некоторый момент в ходе исключения переменных, а значит является синтаксическим следствием нового ограничения $x_0=cx$ и исходных m ограничений. Другими словами, неравенство $x_0\leqslant d$ может быть получено сложением этих ограничений с некоторыми коэффициентами μ_0, $\mu_1\geqslant 0,\dots,\mu_m\geqslant 0$. При этом $\mu_0=1$, так как только ограничение $x_0=cx$ содержит переменную x_0, а переменная x_0 входит в неравенство $x_0\leqslant d$ с единичным коэффициентом.

Итак, после сложения исходных ограничений с коэффициентами $\mu_1,\dots,\mu_m$ и добавления к полученному неравенству равенства $x_0=cx$, мы получаем неравенство $x_0\leqslant d$. Это означает, что до добавления этого равенства мы имели неравенство $cx\leqslant d$, что и требовалось доказать. ∎

Пример. Продемонстрируем доказательство на примере задачи ЛП:

$$\begin{cases}x+y\to\max,\\2x+y\leqslant 4,\\-4\leqslant 2x+5y,\\3y\leqslant 2x+4.\end{cases}\xrightarrow{\text{вводим переменную } t}\begin{cases}t=x+y,\\2x+y\leqslant 4,\\-4\leqslant 2x+5y,\\3y\leqslant 2x+4.\end{cases}\xrightarrow{\text{исключаем } x}\begin{cases}2(t-y)+y\leqslant 4,\\-4\leqslant 2(t-y)+5y,\\3y\leqslant 2(t-y)+4.\end{cases}\longrightarrow$$

$$\xrightarrow{\text{разрешаем относительно } y}\begin{cases}2t-4\leqslant y,\\-\frac{2}{3}t-\frac{4}{3}\leqslant y,\\y\leqslant\frac{2}{5}t+\frac{4}{5}.\end{cases}\xrightarrow{\text{исключаем } y}\begin{cases}2t-4\leqslant\frac{2}{5}t+\frac{4}{5},\\-\frac{2}{3}t-\frac{4}{3}\leqslant\frac{2}{5}t+\frac{4}{5}.\end{cases}\longleftrightarrow\begin{cases}t\leqslant 3,\\-2\leqslant t.\end{cases}$$

Наибольшее значение целевой функции равно 3. Как мы выяснили выше, неравенство $t\leqslant 3$ является следствием исходных неравенств. То есть, если сложить второе, третье и четвертое ограничение с коэффициентами $\frac{5}{8}$, 0, $\frac{1}{8}$, получим неравенство $x+y\leqslant 3$, что и требовалось.

Доказательство (геометрическое). Пусть исходная задача имеет вид $cx \to \max$ при ограничениях $a_i x \leqslant b_i$, $i = 1, \ldots, m$. Пусть x^* есть оптимальное решение этой задачи. Нам нужно установить, что неравенство $cx \leqslant cx^*$ выводимо из исходных m неравенств.

Обозначим через I множество номеров неравенств, насыщаемых в точке x^*: $I = \{i \mid a_i x^* = b_i\}$.

> **Лемма**. Существуют неотрицательные числа λ_i, $i \in I$, для которых вектор $c = \sum_{i \in I} \lambda_i a_i$ (рис. 1).
>
> **Доказательство**. Допустим, что это не так. Тогда по лемме Фаркаша можно отделить вектор c от векторов a_i, $i \in I$, некоторой гиперплоскостью, проходящей через 0. То есть существует вектор u, для которого $cu > 0$, но $a_i u \leqslant 0$ для всех $i \in I$. Докажем, что для достаточно малого положительного t точка $x = x^* + ut$ удовлетворяет всем исходным неравенствам и значение целевой функции в ней больше, чем в x^*.

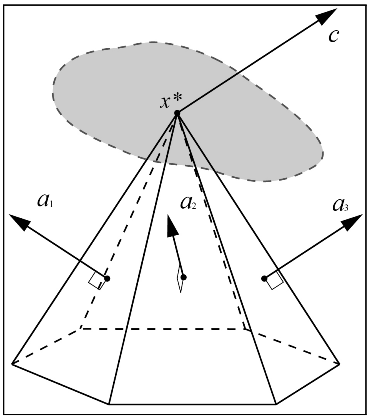

Рис. 1: Точка x^* максимизирует целевую функцию cx на полиэдре. Неравенства $a_i x \leqslant b_i$ насыщены в точке b_i. Векторы a_i перпендикулярны граням полиэдра, задаваемым неравенствами $a_i x \leqslant b_i$, они направлены вне полиэдра. Вектор c выражается через них с неотрицательными коэффициентами.

> Значение целевой функции в x^* равно $cx = c(x^* + ut) = cx^* + tcu$. Поскольку $cu > 0$, это значение строго больше cx^* для любого положительного t.
>
> Теперь нужно показать, что при достаточно малом t точка $x = x^* + ut$ удовлетворяет всем неравенствам. Это делается по-разному для насыщенных и ненасыщенных (в точке x^*) неравенств. Все насыщенные неравенства истинны в точке x при любом положительном t, поскольку для них верно
>
> $$a_i x = a_i(x^* + ut) = a_i x^* + t a_i u = b_i + \underbrace{t}_{>0} \cdot \underbrace{a_i u}_{\leqslant 0 \forall i \in I} \leqslant b_i,$$
>
> а все ненасыщенные неравенства $a_i x \leqslant b_i$ сохраняют свою истинность при достаточно малом t. В самом деле, для них $a_i x^*$ строго меньше b_i. Поэтому при достаточно малом t величина $t a_i u$ меньше разницы между b_i и $a_i x^*$, и поэтому $a_i x = a_i x^* + t a_i u < b_i$. Полученное противоречие доказывает лемму. ■

Теперь докажем, что если сложить все насыщенные неравенства с коэффициентами λ_i из леммы, то как раз и получится искомое неравенство $cx \leqslant cx^*$. В самом деле, при таком сложении в левой части мы получим cx, поскольку $c = \sum_{i \in I} a_i$. А в правой части мы получим $\sum_{i \in I} \lambda_i b_i$, но поскольку в точке x^* все складываемые неравенства насыщены, это то же самое, что $\sum_{i \in I} \lambda_i a_i x^* = cx^*$. ■

7. Теорема о проекции полиэдра

Проекцией $\Pi_n X$ множества $X \subset \mathbb{R}^n$ называется множество таких точек $x' \in \mathbb{R}^{n-1}$, что для некоторого $\lambda \in \mathbb{R}$ выполняется $(x', \lambda) \in X$. Смысл термина "проекция" ясен из (рис. 2).

Лемма. Проекция любого полиэдра снова полиэдр.

Доказательство. Проекция полиэдра состоит в точности из решений системы неравенств, которая получается исключением переменной x_n. ■

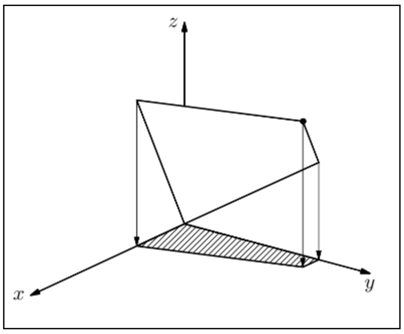

Рис. 2: проекция – тень множества

8. Лемма Фаркаша (знать два доказательства: вывод из первого принципа двойственности и геометрическое доказательство)

Лемма Фаркаша (формулировка из лекции). Пусть в $\mathbb{R}^n$ даны векторы $u_1, \ldots, u_n$ и v. Тогда вектор v нельзя представить в виде $v = \lambda_1 u_1 + \ldots + \lambda_n u_n$, где $\lambda_i \geqslant 0 \iff \exists w : wv > 0$ и $wu_1 \leqslant 0, \ldots, wu_n \leqslant 0$. То есть вектор v невозможно представить в виде неотрицательной линейной комбинации каких-то векторов тогда и только тогда, когда его можно отделить от этих векторов некоторой гиперплоскостью $\{x \mid (v - x_0)x = 0\}$, $v \neq 0$, проходящей через 0.

Лемма Фаркаша (формулировка из конспекта). Система линейных ограничений $Ax = b$, $x \geqslant 0$ несовместна тогда и только тогда, когда найдётся вектор-строка y, удовлетворяющая неравенствам $yA \leqslant 0$, $yb > 0$. То есть вектор b невозможно представить в виде неотрицательной линейной комбинации каких-то векторов (столбцов матрицы A) тогда и только тогда, когда его можно отделить от этих векторов некоторой гиперплоскостью $\{x \mid yx = 0\}$, проходящей через 0 (вектор y задает коэффициенты в уравнении искомой гиперплоскости).

Доказательство (вывод из 1-го принципа двойственности). Для начала вспомним, что такое 1-й принцип двойственности (он же критерий совместности).

Теорема (1-й принцип двойственности, конспект основы). Если система линейных ограничений несовместна, то из неё можно вывести неравенство $0 \leqslant -1$.

Теорема (1-й принцип двойственности, конспект пилота). Система линейных неравенств $Ax \leqslant b$ несовместна тогда и только тогда, когда из неё синтаксически следует неравенство $0 \cdot x \leqslant -1$.

Нетрудно заметить, что лемма Фаркаша является переформулировкой критерия совместности для систем линейных ограничений третьего специального вида (одни равенства на неотрицательные переменные), поэтому докажем аналогично.

Пусть система, состоящая из равенств $a_i x = b_i$ несовместна. Нам нужно вывести из неё некоторое заведомо ложное равенство. Для этого применим к ней алгоритм проверки совместности методом исключения переменных. Поскольку система несовместна, в некоторый момент мы получим заведомо ложное равенство. Заметим, что все новые равенства, возникающие в ходе исключения переменных, являются синтаксическими следствиями исходных равенств. В самом деле, в методе исключения переменных мы использовали шаги следующего вида: для некоторой переменной x_j мы брали два равенства вида $px_j = r$ и $qx_j = s$ (где p, q положительные числа, а r, s – линейные функции от других переменных), разрешали их относительно x_j, получая равенства $x_j = \frac{r}{p}$ и $\frac{s}{q} = x_j$ и затем составляли равенство $\frac{s}{q} = \frac{r}{p}$. То же самое равенство можно получить, сложив равенства $px_j = r$ и $qx_j = s$ с положительными коэффициентами $\frac{1}{p}$ и $\frac{1}{q}$ (переменная x_j сократится). Поэтому равенство $\frac{s}{q} = \frac{r}{p}$ является синтаксическим следствием равенств $px_j = r$ и $qx_j = s$. Значит, все равенства, появляющиеся в ходе исключения переменных, являются синтаксическими следствиями исходных равенств. В частности и заведомо ложное равенство, возникшее в ходе исключения переменных, является синаксическим следствием исходных равенств, что и требовалось доказать. ■

Доказательство (геометрическое, из конспекта). Докажем в обе стороны:

$\Longleftarrow$ Очевидно: если $yA \leqslant 0$, $yb > 0$, то для любого $x \geqslant 0$ выполняется $yAx \leqslant 0$, а так как $yb > 0$, то $Ax \neq b$, поскольку равенство $\underbrace{yAx}_{\leqslant 0} = \underbrace{yb}_{>0}$ не выполняется.

$\Longrightarrow$ Неочевидно: Пусть система линейных ограничений $Ax = b$, $x \geqslant 0$ несовместна. Множество векторов вида Ax, $x \geqslant 0$ состоит из неотрицательных линейных комбинаций столбцов матрицы A и называется выпуклым конусом, порожденным этими столбцами.

Нам важно, что вместе с любым вектором v это множество содержит и все вектора λv для положительных λ, а также, вместе с любыми двумя векторами содержит их сумму. Кроме того важно, что оно замкнуто (в топологическом смысле).

Нам дано, что этот выпуклый конус не содержит точку b. Нам нужно отделить эту точку от выпуклого конуса гиперплоскостью, проходящей через ноль. Для этого возьмём в конусе точку x_0, ближайшую к точке b (*хотя и существование такой точки геометрически очевидно, не менее оно требует доказательства* (не это не точно)).

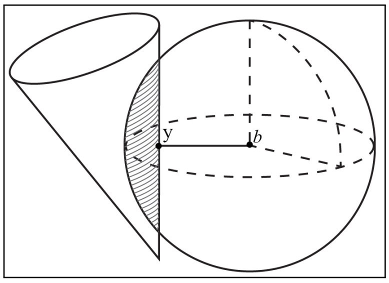

Рис. 3: компакт и ближайшая точка

Доказательство (существования ближайшей точки). Для этого мы пересечём наш конус с шаром в центром в точке b и такого радиуса, чтобы он пересекался с конусом (рис. 3). Это пересечение есть замкнутое и ограниченное множество, то есть **компакт**. А среди расстояний от фиксированной точки до любого непустого компакта есть наименьшее. Ближайшая к b точка конуса в этом шаре (пускай y) и будет ближайшей точкой всего конуса, поскольку все точки конуса вне шара находятся от точки b на расстоянии большем радиуса шара, а найденная точка лежит в шаре, значит её расстояние от b не более радиуса шара. ■

По условию $x_0 \neq b$. Рассмотрим гиперплоскость с уравнением $(b - x_0)x = 0$. Эта гиперплоскость перпендикулярна (ненулевому) вектору $b - x_0$ и содержит точку 0. Эта гиперплоскость и будет искомой отделяющей гиперплоскостью.

Докажем для начала, что $(b - x_0)x_0 = 0$ (то есть эта гиперплоскость содержит точку x_0). Если $x_0 = 0$, то это очевидно. Иначе рассмотрим луч $\{\lambda x_0 \mid \lambda > 0\}$. Этот луч целиком лежит в выпуклом конусе. Поэтому точка x_0, будучи ближайшей к b точкой во всем конусе, является и ближайшей к b точкой на этом луче. Поэтому вектор с началом в x_0 и концом в b (то есть $b - x_0$) перпендикулярен этому лучу, то есть перпендикулярен x_0 (рис. 4). Это и требовалось доказать.

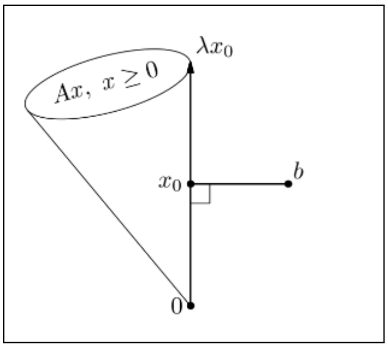

Рис. 4: точка x_0 – ближайшая к b на луче $\lambda x : \lambda > 0$

Итак, мы установили, что $(b - x_0)x_0 = 0$ (то есть гиперплоскость содержит x_0). Наша гиперплоскость делит все пространство на два полупространства: $\{x \mid (b - x_0)x > 0\}$ ("положительное полупространство") и $\{x \mid (b - x_0)x \leqslant 0\}$ ("неположительное полупространство"). Точка b лежит в положительном, поскольку

$$(b - x_0)b = \underbrace{(b - x_0)(b - x_0)}_{\text{скалярный квадрат положителен}} + \underbrace{(b - x_0)x_0}_{=0,\text{ как доказали}}.$$

Осталось установить, что весь конус лежит в неположительном полупространстве. Допустим противное: некоторая точка x_1 из конуса лежит в положительном полупространстве, то есть в $(b - x_0)x_1 > 0$. Рассмотрим луч, состоящий из точек $x_0 + tx_1$ для всех $t \geqslant 0$. Этот луч лежит в конусе и содержит точку x_0. Поскольку точка x_0 есть ближайшая к b точка всего конуса, то она будет и ближайшей к b точкой этого луча. Угол между векторами $b - x_0$ и x_1 острый, поэтому основание перпендикуляра из b

на прямую $(x_0,\, x_0+x_1)$ лежит на этом луче (а не на его продолжении, рис. 5). Получаем противоречие, поскольку основание этого перпендикуляра ближе к точке b, чем начало луча x_0. ■

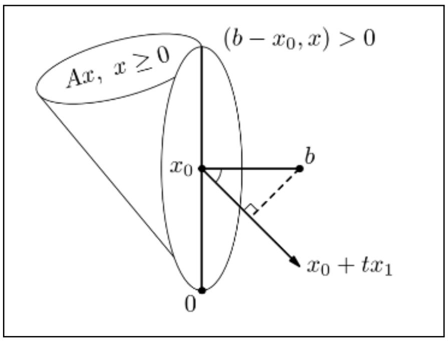

Рис. 5: угол между векторами $(b - x_0)$ и x_1 острый, поэтому основание перпендикуляра из b на прямую $(x_0,\, x_0+tx_1)$ лежит на луче $\{x_0+tx_1 : t > 0\}$

9. Соотношения дополняющей нежесткости

Оптимальные решения прямой и двойственной задач обладают важным свойством: если оптимальное решение не насыщает какое-то неравенство, то соответствующая этому неравенству переменная в оптимальном решении двойственной задачи обращается в 0. В этом случае говорят, что для этого неравенства и этой пары решений выполнено **соотношение дополняющей нежёсткости**.

Теорема. Пусть x^*, y^* – допустимые решения прямой и двойственной задач соответственно. Тогда x^*, y^* – оптимальные тогда и только тогда, когда для всех ограничений прямой и двойственной задач выполнены соотношения дополняющей нежесткости:

$$\begin{cases} y^*(b - Ax^*) = 0, \\ (y^*A - c)x^* = 0. \end{cases}$$

Доказательство. Докажем для задач 2-го специального вида. Пусть имеем

$$\begin{cases} Ax \leqslant b, \\ x \geqslant 0, \\ cx \to \max \end{cases} \xleftrightarrow{\text{двойственная задача}} \begin{cases} yA \geqslant c, \\ y \geqslant 0, \\ yb \to \min \end{cases}$$

Запишем неравенство, доказывающее, что любое возможное значение целевой функции в двойственной задаче ограничивает сверху возможные значения целевой функции в прямой задаче:

$$cx^* \leqslant y^*Ax^* \leqslant y^*b.$$

По принципу двойственности оптимальность пары решений x^*, y^* означает, что $cx^* = y^*b$, поэтому достаточно показать, что соотношения дополняющей нежесткости неравенств прямой задачи означают, что второе неравенство обращается в равенство и наоборот, соотношения дополняющей нежесткости для двойственной задачи означают, что первое неравенство обращается в равенство.

$$cx^* = y^*Ax^* = y^*b.$$

В самом деле, умножим строку y^* на столбец $b - Ax^*$, то есть сложим произведения соответствующих координат этих строки и столбца. Поскольку x^*, y^* – допустимые решения, то все складываемые произведения неотрицательны. Соотношения дополняющей нежесткости для неравенств прямой задачи требуют, чтобы в этой сумме все слагаемые были нулевыми. Поскольку все слагаемые неотрицательны, это равносильно тому, что $y^*(b - Ax^*) = 0$. Аналогично соотношения дополняющей нежесткости для неравенств двойственной задачи означают, что $(y^*A - c)x^* = 0$. ■

Для общего случая мы на лекции не доказывали, так что прочитать можно здесь на странице 32.

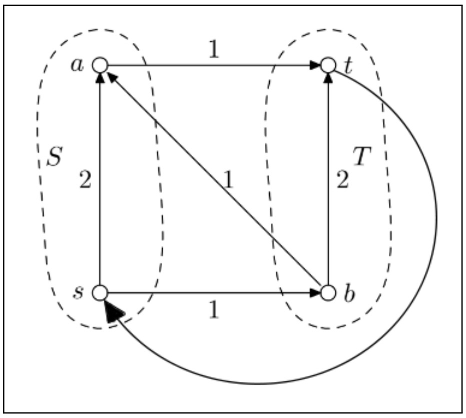

Рис. 6: пример потока в графе

10. Теорема Форда-Фалкерсона о потоках и разрезах (доказательство сведением к задаче ЛП)

Теорема. Максимальный размер потока сети равен минимальной пропускной способности разреза.

Доказательство. Введем дополнительно ребро (t, s) (рис. 6) и потребуем, чтобы дивергенция потока в вершинах s, t была также нулевой, как и на остальных вершинах. Пропускную способность добавленного ребра никак не ограничиваем. Нетрудно понять, что каждому потоку в старой сети соответствует поток в новой сети, в котором количество нефти, пропускаемое через добавленное ребро, равно величине исходного потока. И наоборот, каждому потоку в новой сети соответствует поток в старой сети, размер которого равен количеству нефти, пропускаемому через добавленное ребро.

Таким образом исходная задача эквивалентна задаче максимизации переменной, соответствующей добавленному ребру при тех же ограничениях, что и раньше, плюс два новых равенства нулю дивергенций в s и t. У полученной задачи, очевидно, есть допустимые решения (скажем, всюду нулевой поток).

Распишем прямую задачу для графа:

$$\begin{cases} x_{uv} \geqslant 0, \text{ [количество нефти, которое можно пропустить по ребру]} \\ x_{tx} \geqslant 0, \text{ [также не забываем про добавленное ребро } ts\text{]} \\ x_{uv} \leqslant w_{uw}, \text{ [}w_{uw}\text{– ограничение пропускной способности ребра]} \\ p \in V : \sum\limits_{(p,v)\in E} x_{pv} - \sum\limits_{(u,p)\in E} x_{up} = 0, \text{ [дивергенция вершин]} \\ x_{tx} \to \max \text{ [целевая функция]} \end{cases}$$

А теперь построим двойственную:

$$\begin{cases} z_{uv} \geqslant 0, \text{ [переменная, соответствующая ограничению } x_{uv} \leqslant w_{uv}\text{]} \\ y_v \in \mathbb{R}, \text{ [каждой вершине } v \text{ соответствует переменная } y_v\text{]} \\ z_{uv} + y_u - y_v \geqslant 0, \text{ [ограничение для каждого ребра } (u,v)\text{]}^* \\ y_t - y_s \geqslant 1, \text{ [ограничение на добавленное ребро]} \\ \sum\limits_{u,v\in E} w_{uv} z_{uv} \to \min \text{ [сумма по ребрам, исключая добавленное]} \end{cases}$$

* – переменная x_{uv} входила с плюсом в ограничение $x_{uv} \leqslant w_{uv}$, с минусом в дивергенцию в вершине v и с плюсом в дивергенцию в вершине u.

$x_{ts} \to \max$	$\sum_{(u,v)\in E} w_{u,v} z_{u,v} \to \min,$
$x_{ts} \geqslant 0$	$y_t - y_s \geqslant 1,$
$x_{u,v} \geqslant 0$	$z_{u,v} + y_u - y_v \geqslant 0$
$x_{u,v} \leqslant w_{u,v}$	$z_{u,v} \geqslant 0$
$\mathrm{div}_v = 0$	$y_v \in \mathbb{R}$

Рис. 7: слева – задача о потоке, справа – двойственная к ней; v пробегает все вершины сети, а (u, v) пробегает все ребра.

Значения переменных y_v называют **потенциалами** или **высотами** вершин. Смысл двойственной задачи таков: каждой вершине сети нужно сопоставить некоторый потенциал так, чтобы потенциал t был по крайней мере на единицу больше потенциала s. После этого нужно рассмотреть все ребра (u, v), у которых потенциал конца больше потенциала начала, и для таких ребер (u, v) просуммировать разности потенциалов конца и начала $y_v - y_u$ с коэффициентами w_{uv}. Эту сумму мы хотим сделать как можно меньшей.

Пусть x^*, y^*, z^* – оптимальные решения прямой и двойственной задач, а f – оптимум.

Для доказательства теоремы нам нужно построить разрез пропускной способности f. Для этого выберем произвольное число $r \in [y_s,\ y_t)$ (этот интервал не пуст, так как $y_t \geqslant y_s + 1$) и определим разрез (S, T) по правилу:

$$S = \{v \mid y_v^* \leqslant r\} \text{ [все вершины, потенциал которых не больше } r\text{]}$$
$$T = \{v \mid y_v^* > r\} \text{ [все вершины, потенциал которых строго больше } r\text{]}$$

По построению $s \in S$ и $t \in T$. Мы докажем, что пропускная способность построенного разреза равна f с помощью следующей леммы:

> **Лемма**. Для любого потока и любого разреза размер потока равен общему количеству нефти, перетекающему по ребрам, ведущим из S в T минус общее количество нефти, перетекающее по ребрам, ведущим из T в S. То есть размер потока = поток $(S \to T)$ − поток $(T \to S)$.

А именно, мы установим, что поток ко всем ребрам (u, v) из S в T максимален, то есть $x_{uv}^* = w_{uv}$, а поток ко всем ребрам (u, v) из T в S нулевой, то есть $x_{uv}^* = 0$. Эти два равенства будут следовать из соотношений дополняющей нежесткости.

В самом деле, для ребер (u, v) из S в T высота конца больше высоты начала, поэтому соответствующая переменная z_{uv}^* положительна: $z_{uv}^* \geqslant y_v^* - y_u^* > 0$. Из соотношений дополняющей нежесткости для ограничения $x_{uv} \leqslant w_{uv}$ получаем, что $x_{uv}^* = w_{uv}$.

Для рёбер (u, v) из T в S наоборот, высота конца меньше высоты начала. Поэтому неравенство $z_{uv}^* + y_u^* - y_v^* \geqslant 0$ не насыщено. Из соотношений дополняющей нежесткости для этого неравенства мы получаем, что $x_{uv}^* = 0$. ∎

11. Игры с нулевой суммой. Теорема фонНоймана об играх с нулевой суммой

В отличие от игр с полной информацией, в играх с нулевой суммой игроки делают ход одновременно и поэтому не знают, какой ход выбрать противник. Результат игры полностью определяется сделанными ходами, при этом результаты игроков противоположны: если один игрок выиграл $+1$, то другой проиграл 1, то есть выиграл -1. Игроков будем называть **максимизатором** и **минимизатором**.

В качестве примера смоделируем простейшую матрицу игры:

		min	
max		b_1	b_2
	a_1	0	1
	a_2	-1	2

Если максимизатор делает ход a_1, а минимизатор b_2, то выигрыш максимизатора равен $+1$.

Предположим, что игроки играют много игр подряд. Если минимизатор будет делать ход b_2 все время, то рано или поздно максимизатор сообразит, что ему выгоднее отвечать ходом a_2: в этом случае его выигрыш составит $+2$.

Говорят, что пара (**чистых**) стратегий (a_1, b_2) **не равновесна**. Это означает в общем случае, что один из игроков может улучшить свой результат, отклонившись от выбранной стратегии (в предположении, что второй игрок придерживается той же самой стратегии). **Чистая стратегия** состоит в том, что игрок делает какой-то избранных ход.

Эта игра очень маленькая и легко разобрать все случаи. Оказывается, в ней есть три неравновесных пар чистых стратегии и одна равновесная пара: (a_1, b_1). Действительно, выигрыш в этой стратегии равен 0. Если от неё отклонится максимизатор, а минимизатор будет придерживаться равновесной стратегии, то выигрыш максимизатора станет равным -1, то есть уменьшится. Если отклонится минимизатор, то выигрыш максимизатора станет равным $+1$, т.е. выигрыш минимизатора уменьшится.

Теперь рассмотрим игру "камень-ножницы-бумага":

		min		
		К	Н	Б
max	К	0	1	-1
	Н	-1	0	1
	Б	1	-1	0

В этой игре нет равновесной чистой стратегии (любой игрок может улучшить свой результат, если противник не меняет свою чистую стратегию). В данной игре надежнее всего выбирать ход случайно. Стратегии со случайным выбором хода называются **смешанными**.

Пусть максимизатор выбирает каждый ход с вероятностью $\frac{1}{3}$. Тогда при любом ходе минимизатора математическое ожидание его выигрыша равно 0:

$$\begin{aligned}
\text{К} &: \frac{1}{3} \cdot 0 + \frac{1}{3} \cdot (1) + \frac{1}{3} \cdot (-1) = 0, \\
\text{Н} &: \frac{1}{3} \cdot (-1) + \frac{1}{3} \cdot 0 + \frac{1}{3} \cdot 1 = 0, \\
\text{Б} &: \frac{1}{3} \cdot 1 + \frac{1}{3} \cdot (-1) + \frac{1}{3} \cdot 0 = 0.
\end{aligned}$$

Аналогично и для минимизатора. Таким образом, для этой игры пара смешанных стратегий, в которой обра игрока выбирают ход случайно и равновероятно, является **равновесной**. Но не для любой игры пара таких стратегий равновесна – существует игра, для которой это неверно (пример выше).

Теорема (фон Нойман). Пусть нам дано: $i = 1, \ldots, m$ — количество ходов максимизатора (смешанная стратегия максимизатора определяется выбором вероятностей $p_1, \ldots, p_m$ каждого хода); $j = 1, \ldots, n$ (аналогично смешанная стратегия минимизатора определяется выбором вероятностей $q_1, \ldots, q_n$ каждого хода) – количество ходов минимизатора; G_{ij} – выигрыш максимизатора (проигрыш минимизатора) и математическое ожидание максимизатора при выборе игроками пары стратегий (p, q):

$$W(p,q) = \sum_{i,j} p_i q_j G_{ij} = pGq.$$

Тогда существует такая пара стратегий p^*, q^*, называемая **равновесной**, что для любых p, q выполняется:

$$\underbrace{W(p,q^*) \leqslant}_{\text{гарантия минимизатора}} \quad \underbrace{W(p^*,q^*)}_{\text{цена игры}} \quad \underbrace{\leqslant W(p^*,q)}_{\text{гарантия максимизатора}}$$

Иными словами, в любой антагонистической матричной игре двух игроков есть равновесие в смешанных стратегиях.

Доказательство. Сначала рассмотрим игру с точки зрения максимизатора. Пусть максимизатор хочет гарантировать себе средний выигрыш не меньше некоторого числа u, используя смешанную стратегию $p = (p_1, \ldots, p_m)$:

$$p_1 G_{1j} + p_2 G_{2j} + \ldots + p_m G_{mj} \geqslant u, \; \forall j = 1, \ldots, n$$

В этой системе j-ое неравенство гарантирует выигрыш не меньше u при условии, что минимизатор использует чистую стратегию и всегда делает j-ый ход. Для смешанных стратегий аналогично неравенства записывать не надо, поскольку гарантия для смешанных стратегий следует из гарантий для чистых. В самом деле, просуммируем неравенства этой системы с коэффициентами $q_1, \ldots, q_m$, задающими смешанную стратегию минимизатора. В левой части получится сумма $\sum\limits_{i,j} p_i q_j G_{ij}$, а в правой – $\sum\limits_{j} q_j u$. Поскольку сумма чисел $q_1, \ldots, q_m$ равна единице, мы как раз получим неравенство

$$\sum_{i,j} p_i q_j G_{ij} \geqslant u,$$

означающее, что средний выигрыш максимизатора не меньше u при игре против смешанной стратегии $q_1, \ldots, q_m$. Теперь представим себе, что максимизатор ищет смешанную стратегию, гарантирующую ему максимальный выигрыш. Это означает, что он должен найти оптимальное решение задачи ЛП:

$$\begin{cases}
p_1 G_{1j} + p_2 G_{2j} + \ldots + p_m G_{mj} \geqslant u, \; \forall j = 1, \ldots, n, \\
\sum\limits_{i} p_i = 1, \; p_1, \ldots, p_m \geqslant 0, \\
u \to max
\end{cases}$$

Аналогичным образом можно проанализировать игру с точки зрения минимизатора. Стремясь найти смешанную стратегию, гарантирующую ему наименьший проигрыш, он ищет оптимальное решение задачи ЛП:

$$\begin{cases} q_1G_{i1} + q_2G_{i2} + \ldots + q_nG_{in} \leqslant v, \ \forall i = 1, \ldots, m, \\ \sum\limits_j q_j = 1, \ q_1, \ldots, q_n \geqslant 0, \\ v \to min \end{cases}$$

Докажем, что эти задачи двойственны друг к другу путём построения двойственной задачи к задаче ЛП максимизатора. Перепишем неравенства первой задачи в виде:

$$-p_1G_{1j} - p_2G_{2j} - \ldots - p_mG_{mj} + u \leqslant 0,$$

(чтобы двойственные переменные были положительными). Теперь применим к этой задаче экономный алгоритм (?) построения двойственной задачи. Результат показан на рис. 8. Как видно, это в точности задача, решаемая минимизатором.

Прямая задача	Двойственная задача
$u \to \max$	$v \to \min,$
$p_i \geqslant 0$	$-q_1G_{i1} - q_2G_{i2} - \cdots - p_nG_{in} + v \geqslant 0,$
$-p_1G_{1j} - p_2G_{2j} - \cdots - p_mG_{mj} + u \leqslant 0$	$g_j \geqslant 0$
$p_1 + \cdots + p_n = 1$	$v \in \mathbb{R}$
$u \in \mathbb{R}$	$q_1 + \cdots + q_m = 1$

Рис. 8: прямая задача – задача, решаемая максимизатором; двойственная – минимизатором.

По принципу двойственности существуют их решения с одним и тем же значением целевых функций. Эти решения и дают равновесные стратегии. В самом деле, обозначим через $u^* = v^*$ оптимальные значения u, v в обеих задачах, а через p^*, q^* – их оптимальные решения. Каков средний выигрыш максимизатора, если он придерживается стратегии p^*, а минимизатор – стратегии q^*? Ясно, что он равен u^*, поскольку стратегия p^* гарантирует максимизатору выигрыш не меньше u^*, а стратегия q^* гарантирует минимизатору проигрыш не больше v^*. С другой стороны, эти гарантии означают, что если один из игроков будет отклоняться от оптимальной стратегии, а другой не будет этого делать, то средний выигрыш отклоняющегося игрока не увеличится. ■

12. Алгоритм симплекс метода с данным начальным допустимым решением – общая схема алгоритма и его детализация: как понять, достигнут ли максимум; как находится направление сдвига; как определить, насколько можно сдвинуться в данном направлении; как определить, что целевая функция не ограничена. Оценка числа шагов алгоритма

Общая схема алгоритма.

Имеем $\begin{cases} a_ix \leqslant b_i, \\ cx \to \max \end{cases}$, v_0 – начальное допустимое решение.

0. $v = v_0$
1. $I = \{i \mid a_iv = b_i\}$, то есть I обозначает номера тех неравенств, которые насыщаются в точке v (то есть мы подставляем координаты точки в систему и смотрим, какие из неравенств перешли в равенства).
2. Если c не выражается линейно через a_i, $i \in I$ (проверяем Гауссом), то находим $u \in \mathbb{R}^n : cu > 0$ и $a_iu = 0$ для всех $i \in I$. Переходим к **пункту 5**.
3. Иначе, если существует неотрицательная линейная комбинация $c = \sum\limits_{i \in I} \lambda_i a_i$, то v – точка максимума.

   В этом случае **останавливаемся** и выдаем v и коэффициенты λ_i этой линейной комбинации.
4. Иначе, если c выражается через a_i, но в любой линейной комбинации есть отрицательные числа λ_i, то находим $u \in \mathbb{R}^n : cu > 0$ и $a_iu \leqslant 0$ для всех $i \in I$.
5. Если $\exists i : a_iu > 0$, то находим максимальное t, для которого $\forall i : a_i(v + ut) \leqslant b_i$ (не выйдем из полиэдра). $v = v + ut$. Возвращаемся к **пункту 1**.

6. Если $\forall i : a_i u \leqslant 0$ (или, что то же самое $\forall i : a_i(v + ut) > 0$), то c не ограничена на полиэдре, следовательно **останавливаемся** и говорим, что функция не ограничена.

Как понять, достигнут ли максимум. Максимум будет достигнут тогда, когда целевая функция будет постоянна на опорной грани. Это случается тогда и только тогда, когда вектор c линейно выражается через векторы a_i, где i пробегает множество I номеров всех неравенств, насыщенных в точке v, причем $\lambda_i \geqslant 0$:

$$\forall x : cx = \sum_{i \in I} \lambda_i a_i x \leqslant \sum_{i \in I} \lambda_i b_i = \sum_{i \in I} \lambda_i a_i v = cv \Longleftrightarrow cx \leqslant cv.$$

Как находится направление сдвига. Во-первых, нам нужно, чтобы при движении $v \to v + tu$, где $u, v \in \mathbb{R}^n$ и $t > 0$, целевая функция увеличилась. Разница между значениями целевой функции в точках $v + tu$ и v равна $c(v + tu) - cv = tcu$. Поскольку t положительно, для увеличения целевой функции необходимо и достаточно, чтобы cu было тоже положительным. Это можно проверить, просто вычислив скалярное произведение c и u (то есть надо проверять, что $cu > 0$).

Во-вторых, нам нужно, чтобы при таком движении хотя бы при достаточно малых t мы не вышли из полиэдра. Какие ограничения это накладывает на вектор u? Если i-ое неравенство $a_i x \leqslant b_i$ не насыщено в точке v, то оно не накладывает никаких ограничений на вектор u: при небольшом движении вдоль любого направления строгое неравенство останется истинным. Однако, если i-ое неравенство $a_i x \leqslant bi$ насыщается в точке v, то нам нужно, чтобы вектор u показывал в ту сторону от гиперплоскости $\{x \mid a_i x = b_i\}$, с которой лежит полиэдр. Это условие выполнено, если $a_i u \leqslant 0$. В самом деле: $a_i(v + tu) = a_i v + t a_i u = b_i + t a_i u$, и чтобы последняя величина была меньше или равной b_i необходимо и достаточно, чтобы $a_i u \leqslant 0$. Итак, нам нужно, чтобы $a_i u \leqslant 0$ для всех неравенств $a_i x \leqslant b_i$, насыщенных в точке v. Выяснить для данных i и u, выполнено ли это требование, опять же можно простым подсчётом.

В-третьих, если целевая функция не постоянна на опорной грани, нам нужно, чтобы при небольшом сдвиге в направлении u мы оставались в опорной грани. Как мы уже видели, это означает, что для всех насыщенных в точке v неравенств $a_i x \leqslant b_i$ должно быть выполнено равенство $a_i u = 0$ (а не всего лишь неравенство $a_i u \leqslant 0$, как во втором требовании). Но как узнать, постоянна ли целевая функция на опорной грани? Это случается тогда и только тогда, когда вектор c линейно выражается через векторы a_i, где i пробегает множество I номеров всех неравенств, насыщенных в точке v. В самом деле, если это так, то поскольку для всех $i \in I$ функция $a_i x$ постоянна на текущей опорной грани (она равна b_i), целевая функция cx также постоянна на ней. А иначе, как известно из курса линейной алгебры, система линейных уравнений $cu = 1$, $a_i u = 0$, $i \in I$ неизвестных $u = (u_1, \ldots, u_n)$, имеет решение, и её решение будет искомым направлением. Чтобы выяснить, какой из двух случаев имеет место и найти вектор u (во втором случае), мы методом Гаусса решаем систему линейных уравнений $cu = 1$, $a_i u = 0$, где $i \in I$.

Как определить, насколько можно сдвинуться в данном направлении. Чтобы понять, насколько большое t можно выбрать, двигаясь по лучу $v+tu$, $t > 0$, в каждое ненасыщенное в точке v неравенство $a_i x \leqslant b_i$ подставим $x = v + tu$, а затем разрешим полученное неравенство $a_i v + t a_i u \leqslant b_i$ относительно t. Если $a_i \leqslant 0$, то никакого ограничения на t это неравенство не дает, поскольку вектор u смотрит в ту же сторону от гиперплоскости $\{x \mid a_i x = b_i\}$, с которой лежит полиэдр (ну или просто потому, что сверху t ничем не ограничивается: $t \geqslant \dfrac{b_i - a_i v}{a_i u}$). В этом случае неравенство будет справедливо для всех положительных t. Если же $a_i u > 0$, то вектор u смотрит в другую сторону, а значит двигаясь по лучу $v + tu$, $t > 0$, мы рано или поздно пересечем гиперплоскость $\{x \mid a_i x = b_i\}$. Другими словами, в этом случае неравенство $a_i v + t a_i u \leqslant b_i$ равносильно неравенству $t \leqslant \dfrac{b_i - a_i v}{a_i u}$, в правой части которого стоит положительное число. Если все ненасыщенные неравенства имеют первый тип, то в направлении u можно двигаться неограниченно, и поэтому целевая функция не ограничена сверху на полиэдре. Когда такое направление u и такая точка v найдены, алгоритм останавливается и пара u, v выдается в качестве доказательства неограниченности. Иначе, надо положить t равным наименьшему из чисел $\dfrac{b_i - a_i v}{a_i u}$ для неравенств второго типа и сместиться в точку $v + tu$. В этой точке насытится то самое (ненасыщенное ранее) неравенство, которое дает наименьшее значение $\dfrac{b_i - a_i v}{a_i u}$. В дальнейшем мы будем называть этот переход движением из точки v в направлении u. Это движение может закончится насыщением некоторого нового неравенства (и в этом случае мы говорим, что движемся до упора) или продолжаться неограниченно.

Как определить, что целевая функция не ограничена. Целевая функция будет не ограничена в том случае, когда $\forall i : a_i u \leqslant 0$ (или, что то же самое $\forall i : a_i(v + ut) > 0$), поскольку t будет принимать бесконечно большие значения из $\mathbb{R}$.

Оценка шагов алгоритма. Рассмотрим все грани полиэдра, на которых целевая функция постоянна, будем называть их **особыми**. Если какая-то особая грань в некоторый момент оказалась текущей опорной гранью, то в будущем это не повторится, поскольку значение целевой функции на особых опорных гранях в ходе выполнения алгоритма все время увеличивается. Поэтому такое событие случается не больше раз, чем количество особых граней. Поскольку общее количество граней полиэдра, задаваемого m ограничениями, не превосходит количества способов выбрать подмножество $I \in \{1, \ldots, m\}$, то есть таких событий не больше 2^m. Между любыми двумя такими событиями размерность текущей опорной грани строго уменьшалась. Поэтому количество шагов между любыми двумя такими событиями не больше n (размерности пространства). Поэтому общее количество шагов не больше $n2^m$.

В этом анализе мы не учитывали, что размерность текущей опорной грани неможет увеличиваться. Если же учесть и это обстоятельство, то можно доказать и лучшую оценку 2^m. А именно, отсюда следует, что и любая не особая грань неможет дважды оказаться текущей опорной гранью. В самом деле, на шагах, когда текущая опорная грань не является особой, размерность текущей опорной грани строго уменьшается. А значит размерности всех не особых опорных граней, которые когда-либо были текущими, различны.

13. Алгоритм нахождения отделяющей гиперплоскости в лемме Фаркаша, правило Бленда (доказывать то, что алгоритм не зацикливается, не обязательно)

Конструктивное доказательство леммы Фаркаша (частный случай).
Вход: $c,\ a_1, \ldots, a_k \in \mathbb{R}^n$ такие, что c линейно выражается через $a_1, \ldots, a_k$.
Надо: найти (если есть) $\lambda_1, \ldots, \lambda_k \geqslant 0 : c = \sum\limits_i \lambda_i a_i$, иначе $u \in \mathbb{R}^n : cu > 0,\ a_i u \leqslant 0$ для всех $i = 1, \ldots, k$.

Рассмотрим частный случай, при котором $a_1, \ldots, a_k$ – линейно независимы.

1. С помощью метода Гаусса находим $\lambda_1, \ldots, \lambda_k$ (единственные), такие что $c = \lambda_1 a_1 + \ldots + \lambda_k a_k$;
2. Если $\lambda_1, \ldots, \lambda_k \geqslant 0$, то выдаем $\lambda_1, \ldots, \lambda_k$;
3. Иначе выбираем любое i, такое что $\lambda_i < 0$. Решаем СЛУ $\begin{cases} a_i u = -1, \\ a_j u = 0,\ \forall j \neq i \end{cases}$;

   Тогда $cu = (\sum\limits_i \lambda_i a_i) u = \sum\limits_{l=1}^{k} \lambda_l (a_l u)^{(*)} = \lambda_i \underbrace{a_i u}_{-1} = -\lambda_i > 0$.

   $^{(*)}$ где все $a_l = 0$, кроме $l = i$.

За пояснениями можно обращаться к учебнику (стр. 53-54). ■

Правило Бленда$^{(*)}$ (общий случай).
Пусть B_0 – любой базис (максимальная линейно независимая подсистема $a_1, \ldots, a_l$).
Тогда для $i = 0, 1, \ldots$:

1. $c = \sum\limits_{a_j \in B_i} \lambda_j^{(i)} a_j$;
2. Если $\forall a_j \in B_i : \lambda_j^{(i)} \geqslant 0$, то **стоп**;
3. Иначе находим наименьшее $j^{(*)} : \lambda_j^{(i)} < 0$;
4. Находим $u \in \mathbb{R}^n : \begin{cases} a_l u = 1, \\ a_j u = -1 \end{cases}$ при всех $a_l \in B_i \setminus \{a_j\}$.
   $cu = \sum \lambda_l a_l u = \lambda_j a_j u = -\lambda_j > 0$;
5. Если окажется $a_l u \leqslant 0$ для всех $l = 1, \ldots, k$, то **стоп**.
6. Иначе находим наименьшее $l^{(*)} : a_l u > 0,\ B_{i+1} = B_i \setminus \{a_j\ \cup a_l$.

■

14. Алгоритм нахождения начального допустимого решения в симплекс методе

Выясним, как найти начальное решение системы линейных неравенств (или убедиться, что его нет).

Дано: система неравенств $a_i x \leqslant b_i$
Надо: найти $v \in \mathbb{R}^n : a_i v \leqslant b_i,\ i = 1, \ldots, m$
Введем новую переменную t ассмотрим задачу ЛП:

$$\begin{cases} t \to \max, \\ a_1 x + t \leqslant b_1, \\ \vdots \\ a_m x + t \leqslant b_m \end{cases}, \ v = (\underbrace{0, \ldots, 0}_{x_1, \ldots, x_n},\ \underbrace{\min\{b_1, \ldots, b_m\}}_{t}),$$

где v – допустимое решение. Тогда для начального решения v запустим симплекс-метод. Если найденное значение максимума отрицательно, то исходная система линейных неравенств $a_i x \leqslant b_i$, $i = 1, \ldots, m$ не имеет решения. Иначе, если оно неотрицательно или целевая функция не ограничена, то система имеет решение и из результата работы алгоритма такое решение нетрудно извлечь.

Источники

1. записи (лекция за 05.09.17), конспект (стр. 1);
2. записи (лекция за 19.09.17), конспект (стр. 12-13);
3. конспект (стр. 43);
4. конспект (стр. 11-12);
5. записи (лекция за 15.09.17), конспект (стр. 15-16), импровизация;
6. записи (лекция за 15.09.17, 29.09.17), конспект (стр. 22-29);
7. конспект пилота (стр. 14-15);
8. записи (лекции за 19.09.17, 29.09.17), конспект (стр. 23-30);
9. записи (лекция за 29.09.17), конспект (стр. 31-32);
10. записи (лекция за 03.10.17), конспект (стр. 35-37);
11. записи (лекция за 03.10.17), конспект (стр. 39-42);
12. записи (лекция за 17.10.17), конспект (стр. 43-52);
13. записи (лекция за 17.10.17);
14. записи (лекция за 17.10.17), конспект (стр. 55);

Ченжлоги

v0.0 (03.12.2017) – исходное: добавлены вопросы 1-12 по ЛП
v0.1 (06.12.2017) – исправил коэффициенты матрицы в 9; обновил доказательство второй формы ПД в 6; удалил изображение конуса в 8 и исправил опечатку в лекции; исправил схему в 5; поменян 6 вопрос, убран 9 и переделана нумерация; добавил 12 (не до конца), 13 и 14;
v0.2 (09.12.2017) – исправил опечатку в формулировке первого принципа двойственности в 6 (спасибо Асе); дополнил 12 билет (спасибо Юле); дополнил 9 билет; исправил опечаточки; исправил формулировку второго принципа двойственности в 6; пошаманил с 13;