Тема 2

Двойственность в задаче линейного программирования.

Выше мы определили метрику Канторовича на симплексе вероятностных мер. Сдвинем этот симплекс так, чтобы его центр совпал с началом координат, и рассмотрим наименьшее линейное подпространство, которое содержит сдвинутый симплекс. Мы покажем, что на этом подпространстве можно задать норму, для которой порожденное ей расстояние совпадает с метрикой Канторовича. В частности, прямолинейные отрезки, соединяющие вероятностные меры, являются кратчайшими кривыми относительно транспортной метрики. Чтобы это проделать, нам будет удобно рассматривать двойственный вариант задачи Канторовича, который является частным случаем конструкции двойственности в задачах линейного программирования. Перейдем к деталям.

Обозначим через $M(m,n)$ семейство всех вещественных матриц размера $m \times n$ (т.е. с m строками и n столбцами), причем если $m = n$, то для краткости положим $M(n) = M(m,n)$. Каждый вектор $v \in \mathbb{R}^n$ будем воспринимать как вектор-столбец, тогда v^t — соответствующая вектор-строка. Таким образом, $v \in M(n,1)$, а $v^t \in M(1,n)$. Для $A \in M(m,n)$ и $B \in M(n,k)$ через AB обозначаем стандартное произведение матриц. Таким образом, $AB \in M(n,k)$, и для векторов $v, w \in \mathbb{R}^n$ выражение $v^t w$ есть ни что иное, как декартово произведение этих векторов.

Через $\le$ будем обозначать стандартный (покомпонентный) частичный порядок на $\mathbb{R}^n$. Определим исходную и двойственную задачи линейного программирования (обозначения взяты нами из [6], хотя формулировка и доказательство теоремы двойственности мы излагаем по [17]).

Имеется несколько видов пар двойственных задач. Начнем с симметричного.

2.1 Симметричная форма двойственности

Дано: $A \in M(m,n)$, $c \in \mathbb{R}^n$, $b \in \mathbb{R}^m$. Требуется:

Исходная задача (1)		**Двойственная задача (2)**	
максимизировать	$c^t x,\ x \in \mathbb{R}^n,$	минимизировать	$b^t y,\ y \in \mathbb{R}^m,$
при условиях	$Ax \le b,$	при условиях	$A^t y \ge c,$
	$x \ge 0.$		$y \ge 0.$

Каждый $x \in \mathbb{R}^n$ (каждый $y \in \mathbb{R}^m$), удовлетворяющий условиям исходной (двойственной) задачи называется *допустимым решением*; функции $c^t x$ и $b^t y$ называются *целевыми*; допустимое решение *оптимально*, если оно максимизирует (минимизирует) целевую функцию в исходной (двойственной) задаче; значение целевой функции на оптимальном решении назовем *оптимальным значением*; целевую функцию назовем *неограниченной*, если в случае исходной задачи она может принимать на допустимых x сколь угодно большие значения, а в случае двойственной задачи — она может принимать на допустимых y сколь угодно маленькие значения.

Из рассмотренных только что задач линейного программирования можно получить много других, в некотором смысле эквивалентных, изменяя тем или иным способом данные матрицы A, данные векторы b и c, переменные векторы x и y, а также иногда меняя местами min и max. Мы опишем эти преобразования следующим образом. Сначала мы будем приводить идею модификации, выписывая задачу, с которой мы начинаем, а затем будем менять форму этой задачи без введения переобозначений. После этого мы будем переписывать полученную в результате задачу в новых обозначениях, а также указывать, как новые и старые обозначения связаны. Отметим, что все приводимые трансформации приводят к задачам, одновременно с начальной разрешимым или нет, причем оптимальные решения полученной задачи легко выписываются через оптимальные решения начальной (эти формулы мы также напишем). Наконец, всем рассматриваемым типам задач мы присвоим свои номера и в дальнейшем будем на эти номера ссылаться.

Переход между задачами (1) и (2)

Начнем с задачи (1) и заменим знаки у A, b и c на противоположные, а max поменяем на min. Тогда мы получим задачу (2).

$$(1)\ \begin{cases} c^t x \to \max \\ Ax \le b \\ x \ge 0 \end{cases} \quad \rightsquigarrow \quad \begin{cases} -c^t x \to \min \\ -Ax \le -b \\ x \ge 0 \end{cases} \qquad (2)\ \begin{cases} \bar{c}^t \bar{x} \to \min \\ \bar{A}^t \bar{x} \ge \bar{b} \\ \bar{x} \ge 0 \end{cases} \qquad \begin{matrix} \bar{c} = -c \\ \bar{A}^t = -A \\ \bar{b} = -b \\ \bar{x} = x \end{matrix}$$

Отметим, что у этих двух задач множества и допустимых, и оптимальных решений совпадают, в частности, они одновременно или имеют, или не имеют решений (в этом смысле эти задачи эквивалентны). Тем не менее, оптимальные значения у этих задач, вообще говоря, различны.

Аналогично только что проделанному преобразованию, можно перевести задачу (2) в задачу (1).

Задачи (1.1), (2.1), а также переход $(*) \rightsquigarrow (*.1)$

Снова начнем с задачи (1) и покажем, как можно получить эквивалентную задачу, заменив неравенство в условии $Ax \le b$ на равенство. Полученная задача обозначается нами как (1.1). Аналогично определим задачу (2.1) и покажем, как можно перейти к ней из задачи (2). Чтобы осуществить эти операции, нам понадобится ввести дополнительные переменные векторы: в случае задачи (1) вектор $u \in \mathbb{R}^n$, а в случае задачи (2) — вектор $w \in \mathbb{R}^m$. Также через E будем обозначать единичную матрицу, размер которой легко вычисляется по тому, на какой вектор она умножается.

$$(1) \begin{cases} c^t x \to \max \\ Ax \le b \\ x \ge 0 \end{cases} \rightsquigarrow \begin{cases} c^t x + 0^t u \to \max \\ Ax + Eu = b \\ x \ge 0 \\ u \ge 0 \end{cases} \quad (1.1) \begin{cases} \bar{c}^t \bar{x} \to \max \\ \bar{A}\bar{x} = \bar{b} \\ \bar{x} \ge 0 \end{cases} \quad \begin{array}{l} \bar{c}^t = (c^t, 0) \\ \bar{A} = (A, E) \\ \bar{b} = b \\ \bar{x} = \begin{pmatrix} x \\ u \end{pmatrix} \end{array}$$

$$(2) \begin{cases} b^t y \to \min \\ A^t y \ge c \\ y \ge 0 \end{cases} \rightsquigarrow \begin{cases} b^t y + 0^t w \to \min \\ A^t y - Ew = c \\ y \ge 0 \\ w \ge 0 \end{cases} \quad (2.1) \begin{cases} \bar{b}^t \bar{y} \to \min \\ \bar{A}^t \bar{y} = \bar{c} \\ \bar{y} \ge 0 \end{cases} \quad \begin{array}{l} \bar{b}^t = (b^t, 0) \\ \bar{A}^t = (A^t, -E) \\ \bar{c} = c \\ \bar{y} = \begin{pmatrix} y \\ w \end{pmatrix} \end{array}$$

Что касается связи допустимых и оптимальных решений, то для каждого допустимого (оптимального) решения x задачи (1) задача (1.1) имеет допустимое (оптимальное) решение $\begin{pmatrix} x \\ b - Ax \end{pmatrix}$. Обратно, если $\begin{pmatrix} x \\ u \end{pmatrix}$ — допустимое (оптимальное) решение (1.1), то x — допустимое (оптимальное) решение (1).

Задачи $(1')$, $(2')$, а также переход $(*') \rightsquigarrow (*)$

Под задачей $(1')$ будем понимать задачу (1), в которой опущено условие $x \ge 0$. Аналогично определяется задача $(2')$. Мы введем два переменных вектора $u, v \in \mathbb{R}^n$ для задачи $(1')$, и соответственно два переменных вектора $w, z \in \mathbb{R}^m$ для задачи $(2')$.

$$(1') \begin{cases} c^t x \to \max \\ Ax \le b \end{cases} \rightsquigarrow \begin{cases} c^t (u - v) \to \max \\ A(u - v) \le b \\ u \ge 0 \\ v \ge 0 \end{cases} \quad (1) \begin{cases} \bar{c}^t \bar{x} \to \max \\ \bar{A}\bar{x} \le \bar{b} \\ \bar{x} \ge 0 \end{cases} \quad \begin{array}{l} \bar{c}^t = (c^t, -c^t) \\ \bar{A} = (A, -A) \\ \bar{b} = b \\ \bar{x} = \begin{pmatrix} u \\ v \end{pmatrix} \end{array}$$

$$(2') \begin{cases} b^t y \to \min \\ A^t y \ge c \end{cases} \rightsquigarrow \begin{cases} b^t (w - z) \to \min \\ A^t (w - z) \ge c \\ w \ge 0 \\ z \ge 0 \end{cases} \quad (2) \begin{cases} \bar{b}^t \bar{y} \to \min \\ \bar{A}^t \bar{y} \ge \bar{c} \\ \bar{y} \ge 0 \end{cases} \quad \begin{array}{l} \bar{b}^t = (b^t, 0) \\ \bar{A}^t = (A^t, -E) \\ \bar{c} = c \\ \bar{y} = \begin{pmatrix} w \\ z \end{pmatrix} \end{array}$$

Выясним, как связаны допустимые (оптимальные) решения задач $(1')$ и (1). Пусть x — допустимое (оптимальное) решение задачи $(1')$. Положим $x^+ = \max(x, 0)$, $x^- =$

$\max(-x, 0)$, тогда $x^+ \geq 0$, $x^- \geq 0$ и $x = x^+ - x^-$. Таким образом, положив $u = x^+$ и $v = x^-$, мы получим допустимое (оптимальное) решение задачи (1). Обратно, если $\begin{pmatrix} u \\ v \end{pmatrix}$ — допустимое (оптимальное) решение задачи (1), то $x = u - v$ — допустимое (оптимальное) решение задачи $(1')$.

Отметим, что эти допустимые (оптимальные) решения задачи (1) можно легко модифицировать в классе допустимых (оптимальных) решений, добавляя к u и v любые одинаковые неотрицательные векторы, при этом соответствующий $x = u - v$ остается неизменным.

Задачи $(1.1')$, $(2.1')$, а также переходы $(*.1) \rightsquigarrow (*)$ и $(*.1') \rightsquigarrow (*')$

Под задачей $(1.1')$ будем понимать задачу (1.1), в которой опущено условие $x \geq 0$. Аналогично определяется задача $(2.1')$. Обозначим через $\boxed{}$ или присутствующее, или отсутствующее условие $x \geq 0$. Таким образом, заменяя $x \geq 0$ на прямоугольник $\boxed{}$, мы одновременно кодируем обе задачи (1.1) и $(1.1')$. Ниже мы покажем, как перейти от этих задач соответственно к (1) и $(1')$. Переход от (2.1) и $(2.1')$ к (2) и $(2')$ осуществляется аналогично.

$$\begin{matrix}(1.1)\\(1.1')\end{matrix} \quad \begin{cases} c^t x \to \max \\ Ax = b \\ \boxed{} \end{cases} \quad \rightsquigarrow \quad \begin{cases} c^t x \to \max \\ Ax \leq b \\ -Ax \leq -b \\ \boxed{} \end{cases} \qquad \begin{matrix}(1)\\(1')\end{matrix} \quad \begin{cases} \bar{c}^t \bar{x} \to \max \\ \bar{A}\bar{x} \leq \bar{b} \\ \boxed{} \end{cases} \qquad \begin{matrix} \bar{c} = c \\ \bar{A} = \begin{pmatrix} A \\ -A \end{pmatrix} \\ \bar{b} = \begin{pmatrix} b \\ -b \end{pmatrix} \\ \bar{x} = x \end{matrix}$$

На сей раз обе задачи имеют одинаковые множества допустимых (оптимальных) решений.

2.2 Асимметричная форма двойственности

Следующая диаграмма является гипотетической и приводит к паре задач, которые мы будем называть исходной и двойственной в асимметричной форме. Начнем с двойственной задачи (2.1), преобразуем ее в задачу (2), затем рассмотрим исходную задачу для (2), т.е. задачу (1), которая, как окажется, будет записана в форме $(1')$. Конечная и начальная задачи в этой цепочке и будут парой двойственных задач в асимметричной форме. Приведем соответствующую диаграмму.

$$
(2.1)\ \begin{cases} b^t y \to \min \\ A^t y = c \\ y \geq 0 \end{cases} \quad \rightsquigarrow \quad (2)\ \begin{cases} b^t y \to \min \\ \begin{pmatrix} A^t \\ -A^t \end{pmatrix} y \geq \begin{pmatrix} c \\ -c \end{pmatrix} \\ y \geq 0 \end{cases} \quad \rightsquigarrow
$$

$$
\rightsquigarrow \quad (1)\ \begin{cases} c^t(u-v) \to \max \\ A(u-v) \leq b \\ u \geq 0 \\ v \geq 0 \end{cases} \quad \rightsquigarrow \quad (1')\ \begin{cases} c^t x \to \max \\ Ax \leq b \end{cases}
$$

Таким образом, пара двойственных задач линейного программирования в асимметричной форме имеет следующий вид.

Исходная задача (I)		**Двойственная задача (II)**	
максимизировать	$c^t x,\ x \in \mathbb{R}^n,$	минимизировать	$b^t y,\ y \in \mathbb{R}^m,$
при условиях	$Ax \leq b.$	при условиях	$A^t y = c,$
			$y \geq 0.$

2.3 Теорема двойственности для асимметричной формы

Мы докажем теорему двойственности для задач в асимметричной форме.

Теорема 2.1 (см. например [17])**.** *Если одна из двух задач* (I), (II) *имеет оптимальное решение, то вторая также имеет оптимальное решение, причем значения целевых функций на этих оптимальных решениях совпадают. Если целевая функция одной из этих задач неограничена, то оставшаяся задача не имеет допустимых решений.*

Доказательство. Начнем со следующей леммы.

Лемма 2.2. *Пусть x и y — допустимые решения задач* (I) *и* (II) *соответственно. Тогда $c^t x \leq b^t y$.*

Доказательство. Действительно,

$$c^t x = y^t A x \leq y^t b,$$

где последнее неравенство вытекает из того, что $Ax \leq b$ и $y \geq 0$. □

Таким образом, если у обеих задач (I), (II) имеются допустимые решения, то множество значений целевой функции $c^t x$ по всем допустимым x лежит слева от множества всех значений целевой функции $b^t y$ по всем допустимым y. Отсюда мгновенно получается доказательство второй части теоремы.

Предположим теперь, что задача (II) имеет оптимальное решение y_0. Положим

$$C = \Big\{ \big(r := b^t(sy_0 - y),\ w := sc - A^t y\big) : y \geq 0,\ s \geq 0 \Big\} \subset \mathbb{R}^{m+1}.$$

Так как, при фиксированных A, b и c, отображение $(y,s) \mapsto \big(b^t(sy_0 - y),\ sc - A^t y\big)$ линейно, то C, будучи образом неотрицательного ортанта в $\mathbb{R}^{m+1}$, является замкнутым и выпуклым.

Лемма 2.3. *Точка $(1,0) \in \mathbb{R} \times \mathbb{R}^m$, где 0 — начало координат в $\mathbb{R}^m$, не содержится во множестве C.*

Доказательство. Предположим противное, т.е. для некоторых $s \geq 0$ и $y \geq 0$ выполняется $b^t(sy_0 - y) = 1$ и $sc - A^t y = 0$.

Пусть сначала $s \neq 0$, тогда y/s — допустимое решение задачи (II). Так как y_0 — оптимальное решение задачи (II), то $b^t y_0 - b^t(y/s) \leq 0$, однако левая часть этого неравенства равна $1/s > 0$, противоречие.

Пусть теперь $s = 0$, тогда $A^t y = 0$ и $b^t y = -1$. Отсюда вытекает, что для любого $\alpha \geq 0$ выполняется одновременно $A^t(y_0 + \alpha y) = c$ и $y_0 + \alpha y \geq 0$, поэтому $y_0 + \alpha y$ — допустимое решение задачи (II). Но $b^t(y_0 + \alpha y) = b^t y_0 - \alpha$, так что при $\alpha > 0$ значение целевой функции задачи (II) на допустимом решении $y_0 + \alpha y$ меньше, чем $b^t y_0$, а это противоречит оптимальности y_0. □

Лемма 2.4. *Пусть $X, Y \subset \mathbb{R}^d$ — непересекающиеся непустые выпуклые замкнутые множества, причем X — ограниченно (и, значит, компактно). Тогда существует такая гиперплоскость π, что X и Y лежат в разных открытых полупространствах относительно π (в этом случае говорят, что π **строго разделяет** X **и** Y).*

Доказательство. Функция $f : \mathbb{R}^d \to \mathbb{R}$, $f(x) = \inf\big\{|xy| : y \in Y\big\}$, непрерывна, поэтому, в силу компактности X, существует x_0 такая, что $f(x_0) = \inf_{x\in X} f(x)$. Выберем произвольное $r > f(x_0)$, и пусть $B_r(x_0)$ — замкнутый шар с центром в x_0 и радиусом r. Положим $Z = Y \cap B_r(x_0)$, тогда Z — непустой компакт, и $\inf_{y\in Y} |x_0 y| = \inf_{y \in Z} |x_0 y|$. Так как Z — компакт, то существует $y_0 \in Z$ такое, что $d := |x_0 y_0| = \inf_{y\in Z} |x_0 y|$. Так как $X \cap Y = X \cap Z = \emptyset$, то $|x_0 y_0| > 0$. Пусть z — середина отрезка $[x_0, y_0]$, и π — гиперплоскость, перпендикулярная $[x_0, y_0]$ и проходящая через z. Покажем, что π — искомая гиперплоскость.

Отметим, что внутри шара $B_d(x_0)$ нет точек из Y, так как для всех $x \in X$ и $y \in Y$ имеем $|xy| \geq d$. Обозначим через π_y гиперплоскость, проходящую через y_0 перпендикулярно $[x_0, y_0]$, и пусть Π_y — то полупространство относительно π_y, которое содержит шар $B_d(x_0)$. Покажем, что внутри Π_y не могут содержаться точки из Y. Действительно, если это не так и некоторая $y' \in Y$ содержится в Π_y, то, в силу выпуклости Y, отрезок $[y_0, y']$ содержится в Y. Однако, каждый такой отрезок пересекает внутренность шара $B_d(x_0)$, противоречие. Итак, Y содержится в противоположном Π_y полупространстве Π'_y. Аналогично определяется полупространство Π'_x, содержащее X. Но $\Pi'_y \cap \Pi'_x = \emptyset$, а $\Pi'_y \cup \Pi'_x$ не пересекает π, поэтому $\Pi'_y \supset Y$ и $\Pi'_x \supset X$ лежат в разных открытых полупространствах относительно π. □

Итак, в силу леммы 2.4, существуют такой вектор $(\alpha, \beta) \in \mathbb{R} \times \mathbb{R}^m$, что некоторая гиперплоскость, перпендикулярная этому вектору, строго разделяет C и точку $(1, 0)$. Выбрав (α, β) так, чтобы он был направлен в полупространство, содержащее C, получим, что

$$\alpha = (\alpha, \beta)^t(1, 0) < d = \inf\{\alpha r + \beta^t w : (r, w) \in C\}.$$

Заметим, что $d \geq 0$, так как вместе с каждой точкой (r, w) множество C содержит и точку вида $\gamma(r, w)$, $\gamma \geq 0$, поэтому если для некоторой точки $(r, w) \in C$ выполняется $\alpha r + \beta^t w < 0$, то $d = -\infty$, противоречие. С другой стороны, $(0, 0) \in C$, поэтому $d \leq 0$ и, значит, $d = 0$, а $\alpha < 0$. Умножая, если необходимо, α и β на подходящее положительное число, добьемся того, чтобы α было равно -1. Тем самым, мы доказали существование $\beta \in \mathbb{R}^m$ такого, что $-r + \beta^t w \geq 0$ при всех $(r, w) \in C$, или, вспоминая определения r и w, получаем

$$b^t(y - sy_0) + \beta^t(sc - A^t y) = (b^t - \beta^t A^t)y + s(\beta^t c - b^t y_0) \geq 0 \tag{2.1}$$

при всех $y \geq 0$ и $s \geq 0$.

Положим $s = 0$, тогда неравенство (2.1) превратится в $(b^t - \beta^t A^t)y \geq 0$ при всех $y \geq 0$, что равносильно $b^t - \beta^t A^t \geq 0$ или $A\beta \leq b$. Но это в точности означает, что β — допустимое решение задачи (I).

Положим теперь $s = 1$ и $y = 0$, тогда неравенство (2.1) запишется в виде $\beta^t c \geq b^t y_0$, откуда, в силу леммы 2.2, заключаем, что β — оптимальное решение задачи (I) и при этом $c^t\beta = b^t y_0$, т.е. значения целевых функций на оптимальных решениях совпадают.

Докажем теперь, что если задача (I) имеет оптимальное решение, то и задача (II) также имеет оптимальное решение, причем значения целевых функций на этих решениях совпадают. Для этого воспользуемся обсуждениями перед формулировкой теоремы. Рассмотрим следующую диаграмму.

$$\begin{matrix}(\mathrm{I})\\(1')\end{matrix}\ \begin{cases} c^t x \to \max \\ Ax \leq b \end{cases} \rightsquigarrow (1) \begin{cases} c^t(u - v) \to \max \\ A(u - v) \leq b \\ u \geq 0 \\ v \geq 0 \end{cases} \rightsquigarrow (2) \begin{cases} -c^t(u - v) \to \min \\ -A(u - v) \geq -b \\ u \geq 0 \\ v \geq 0 \end{cases} \rightsquigarrow$$

$$\rightsquigarrow (2.1) \begin{cases} -c^t(u - v) + 0^t w \to \min \\ -A(u - v) - Ew = -b \\ u \geq 0 \\ v \geq 0 \\ w \geq 0 \end{cases} \rightsquigarrow (1') \begin{cases} -b^t y \to \max \\ -A^t y \leq -c \\ A^t y \leq c \\ -Ey \leq 0 \end{cases} \rightsquigarrow \begin{matrix}(\mathrm{II})\\(2.1)\end{matrix}\ \begin{cases} b^t y \to \min \\ A^t y = c \\ y \geq 0 \end{cases}$$

На этой диаграмме все переходы, кроме $(2.1) \to (1')$, описаны перед теоремой, а переход $(2.1) \to (1')$ был обоснован в первой части доказательства теоремы. Таким образом, если x_0 — оптимальное решение задачи (I), то $u = x^+$, $v = x^-$ — оптимальное решение следующих за ней задач (1) и (2), а $\big(u, v, w := b - A(u - v)\big)$ — оптимальное решение задачи (2.1) в силу соображений, описанных перед теоремой. Далее, в силу доказанного выше, существует оптимальное решение y_0 задачи $(1')$ (переход $(2.1) \to (1')$

от двойственной задачи к исходной). Снова воспользовавшись обсуждениями до теоремы, получим, что y_0 является оптимальным решением задачи (II). Тем самым, мы доказали, что если задача (I) имеет оптимальное решение, то и задача (II) также имеет оптимальное решение. Осталось применить первую часть доказательства, в соответствии с которой оптимальные значения целевых функций совпадают. $\square$