§4 ЛИНЕЙНОЕ ПРОГРАММИРОВАНИЕ

Пусть n, $m \in \mathbb{N}$ и $I = \{1, 2, \ldots, n\}$, $J = \{1, 2, \ldots, m\}$. Задачей линейного программирования в общей форме называется следующая задача

$$\begin{cases} \sum\limits_{i=1}^{n} c_i x_i \to \min; \\ \sum\limits_{i=1}^{n} a_{ij} x_i \geq b_j, \quad j \in J_1; \\ \sum\limits_{i=1}^{n} a_{ij} x_i = b_j, \quad j \in J_2; \\ x_i \geq 0, i \in I_1; \end{cases} \tag{4.1}$$

где $c = (c_1, c_2, \ldots, c_n)^\mathsf{T} \in \mathbb{R}^n$, $b = (b_1, b_2, \ldots, b_m)^\mathsf{T} \in \mathbb{R}^m$, $A = (a_{ij})_{i,j=1}^{n,m} \in \mathbb{R}^{n \times m}$, а $I = I_1 \sqcup I_2$ и $J = J_1 \sqcup J_2$ — некоторые разбиения множеств I и J, соответственно. Очевидно, что задача (4.1) — частный случай задачи выпуклого программирования. Если $J_1 = J$ (а значит, $J_2 = \varnothing$) и $I_1 = I$, то задача (4.1) называется стандартной (или симметричной), если же $J_2 = J$ и $I_1 = I$, то задача (4.1) — канонической (или основной).

Пример 4.1 (Расстояние землекопа)**.** Пусть дано два конечных множества $A = \{a_i\}_{i=1}^m$ и $B = \{b_j\}_{j=1}^k \subset \mathbb{R}^n$ точек. Требуется каким-либо разумным образом измерить расстояние (меру схожести) между этими множествами. Рассмотрим так называемое расстояние землекопа (Earth mover's distance). В каждую точку $a_i \in A$ поместим кучу песка объёма $1/|A|$, а в каждой точке $b_j \in B$ выкопаем яму объёма $1/|B|$ (очевидно, что общий объём песка в точках множества A равен общему объёму выкопанных ям в точках множества B). Будем считать, что стоимость перемещения песка объёма v из точки a_i в точку b_j равна $vd(a_i, b_j)$, где $d(a_i, b_j)$ — расстояние между точками a_i и b_j. Расстояние землекопа между множествами A и B равно минимальной стоимости, за которую можно засыпать ямы в точках множества B песком из точек множества A. Расстояние землекопа может быть найдено решением следующей задачи линейного программирования:

$$\begin{cases} \sum\limits_{i=1}^{m} \sum\limits_{j=1}^{k} v_{ij} d(a_i, b_j) \to \min; \\ \sum\limits_{j=1}^{k} v_{ij} = \dfrac{1}{m}, \quad 1 \leq i \leq m; \\ \sum\limits_{i=1}^{m} v_{ij} = \dfrac{1}{k}, \quad 1 \leq j \leq k; \\ v_{ij} \geq 0, \quad 1 \leq i \leq m,\ 1 \leq j \leq k. \end{cases}$$

Пример 4.2 (Линейная регрессия)**.** Пусть дана обучающая выборка $S = \{(x_i, y_i\}_{i=1}^N$. Задача линейной регрессии заключается в том, чтобы найти вектор $a \in \mathbb{R}^n$ и число b, такие что $y \approx a^\mathsf{T} x + b$. Как правило, поиск параметров a и b сводится к решению задачи

$$\sum_{i=1}^{N} (y_i - a^\mathsf{T} x_i - b_i)^2 \to \min_{a,b}.$$

Однако, при таком подходе даже единственный выброс может существенно исказить искомые параметры. Для уменьшения влияния выбросов переходят к следующей задаче

$$\sum_{i=1}^{N} |y_i - a^\mathsf{T} x_i - b_i| \to \min_{a,b},$$

которая, очевидно, равносильна задаче линейного программирования

$$\begin{cases} \sum\limits_{i=1}^{N} \xi_i \to \min; \\ \xi_i \geq y_i - a^\mathsf{T} x_i - b_i \geq -\xi_i, \quad 1 \leq i \leq N; \\ \xi \geq \mathbf{0}. \end{cases}$$

Каждой задаче (4.1) соответствуют так называемая двойственная задача. Мотивировкой введения двойственной задачи являются следующие рассуждения. Пусть $x \in \mathbb{R}^n$ — произвольный допустимый вектор задачи (4.1). Рассмотрим произвольный вектор $y = (y_1, y_2, \ldots, y_m)^\mathsf{T} \in \mathbb{R}^m$, такой что $y_j \geq 0$, если $j \in J_1$. Умножим каждое ограничение задачи (4.1) на соответствующую компоненту вектора y и сложим их. В итоге получаем, что

$$\sum_{j=1}^{m} b_j y_j \leq \sum_{j=1}^{m} \sum_{i=1}^{n} a_{ij} x_i y_j = \sum_{i=1}^{n} \left(\sum_{j=1}^{m} a_{ij} y_j \right) x_i.$$

Потребуем, чтобы $\sum\limits_{j=1}^{m} a_{ij} y_j \leq c_i$ при $i \in I_1$ и $\sum\limits_{j=1}^{m} a_{ij} y_j = c_i$ при $i \in I_2$. Тогда справедливо неравенство $\sum\limits_{i=1}^{n} c_i x_i \geq \sum\limits_{j=1}^{m} b_j y_j$. Другими словами, чем больше величина $\sum\limits_{j=1}^{m} b_j y_j$, тем точнее оценка снизу оптимального значения целевой функции $c^\mathsf{T} x$. Таким образом, рассмотрим следующую задачу линейного программирования

$$\begin{cases} \sum\limits_{j=1}^{m} b_j y_j \to \max; \\ \sum\limits_{j=1}^{m} a_{ij} y_j \leq c_i, \quad i \in I_1; \\ \sum\limits_{j=1}^{m} a_{ij} y_j = c_i, \quad i \in I_2; \\ y_j \geq 0, \quad j \in J_1. \end{cases} \tag{4.2}$$

Задача (4.2) называется двойственной, а задача (4.1) — прямой (или исходной). Сведём воедино правила составления двойственной задачи:

1. Каждому j-му ограничению исходной задачи соответствует переменная y_j двойственной задачи и, наоборот, каждому i-му ограничению двойственной задачи соответствует переменная x_i исходной задачи.

2. Матрица ограничений A заменяется на транспонированную A^{T}.
3. Свободные члены ограничений одной из задач являются коэффициентами при соответствующих переменных в целевой функции другой задачи. При этом поиск минимума заменяется на поиск максимума и наоборот.
4. В каждой из задач ограничения-неравенства следует записывать со знаком "$\geq$" при поиске минимума и со знаком "$\leq$" при поиске максимума.
5. Каждому j-му ограничению-неравенству прямой задачи в двойственной задаче соответствует условие неотрицательности ($y_j \geq 0$), а равенству — переменная y_j без ограничение на знак. Наоборот, неотрицательной переменной $x_i \geq 0$ соответствует в двойственной задаче i-е ограничение-неравенство, а произвольной переменной-равенство.

Нетрудно видеть, что двойственной задачей для задачи (4.2) является задача (4.1). Между решениями задач (4.1) и (4.2) имеется ряд нетривиальных связей, образующих теорию двойственности. Для доказательства некоторых утверждений этой теории нам понадобится лемма Фаркаша. Но сперва установим, что всякий конечно порождённый выпуклый конус в пространстве $\mathbb{R}^n$ является замкнутым множеством.

Опр. 4.1. *Подмножество C векторного пространства V называется выпуклым конусом, если $\alpha x + \beta y \in C$ для любых неотрицательных чисел α, β и любых векторов x, $y \in C$. Другими словами, выпуклый конус — это подмножество векторного пространства, замкнутое относительно сложения и умножения на неотрицательные числа. Конус C называется конечно порождённым, если найдутся такие векторы v_1, v_2, ..., v_m, что*

$$C = \{\alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_m v_m \colon \alpha_i \geq 0, 1 \leq i \leq m\}.$$

Лемма 4.1. *Пусть F и $H \subset \mathbb{R}^n$ — замкнутые подмножества, такие что $f \perp h$ для любых $f \in F$ и $h \in H$. Тогда множество $F + H \stackrel{\text{def}}{=} \{f + h \colon f \in F, h \in H\}$ замкнуто.*

$\triangleright$ Пусть $(f_k)_{k\in\mathbb{N}} \subset F$ и $(h_k)_{k\in\mathbb{N}} \subset H$ — две последовательности, такие что $\lim\limits_{k\to+\infty}(f_k + h_k) = x$. Так как

$$\|f_m + h_m - (f_k + h_k)\|^2 = \|(f_m - f_k) + (h_m - h_k)\|^2 = \|f_m - f_k\|^2 + \|h_m - h_k\|^2 \to 0,$$

то $\|f_m - f_k\|^2 \to 0$ и $\|h_m - h_k\|^2 \to 0$. Следовательно,

$$\lim_{k\to+\infty} f_k = f \in F \quad \text{и} \quad \lim_{k\to+\infty} h_k = h \in H,$$

а значит, $x = f + h \in F + H$. $\triangleleft$

Лемма 4.2. *Пусть a_1, a_2, ..., $a_m \in \mathbb{R}^n$ — столбцы матрицы $A \in \mathbb{R}^{n\times m}$. Тогда конус $C \stackrel{\text{def}}{=} \{Ax \colon x \in \mathbb{R}^m_+\}$ замкнут.*

$\triangleright$ Докажем замкнутость конуса $C_s \stackrel{\text{def}}{=} \{\lambda_1 a_1 + \lambda_2 a_2 + \ldots + \lambda_s a_s \colon \lambda_i \geq 0\}$ индукцией по s. Очевидно, что $C_1 = \{\lambda_1 a_1 \colon \lambda_1 \geq 0\}$ — замкнутое множество. Предположим, что мы доказали, что конус C_s замкнут для любых векторов a_1, a_2, …, $a_s \in \mathbb{R}^n$. Покажем, что тогда замкнут конус C_{s+1}. Рассмотрим произвольную последовательность $(c_k)_{k\in\mathbb{N}} \subset C_{s+1}$,

$$c_k = \lambda_1^k a_1 + \lambda_2^k a_2 + \ldots + \lambda_{s+1}^k a_{s+1},$$

такую что $\lim\limits_{k\to+\infty} c_k = c$, и докажем, что $c \in C_{s+1}$. Если все последовательности чисел $(\lambda_i^k)_{k\in\mathbb{N}}$ ограничены, то без нарушения общности будем считать, что они сходятся: $\lim\limits_{k\to+\infty} \lambda_i^k = \lambda_i$. Следовательно, имеет место равенство

$$c = \lambda_1 a_1 + \lambda_2 a_2 + \ldots + \lambda_{s+1} a_{s+1} \in C_{s+1}.$$

Рассмотрим случай, когда хотя бы одна из последовательностей чисел $(\lambda_i^k)_{k\in\mathbb{N}}$ неограничена, например, с номером $s+1$. Без нарушения общности будем считать, что $\lambda_{s+1}^k \uparrow +\infty$ и $\lambda_{s+1}^k \geq \lambda_i^k$, $1 \leq i \leq s$. Но тогда последовательности $(\lambda_i^k/\lambda_{s+1}^k)_{k\in\mathbb{N}}$ ограничены, а значит, без нарушения общности будем считать, что они сходятся: $\lim\limits_{k\to+\infty} \lambda_i^k/\lambda_{s+1}^k = \lambda_i$, $1 \leq i \leq s$. Следовательно, $\lambda_1 a_1 + \lambda_2 a_2 + \ldots + a_{s+1} = \mathbf{0}$, т.е.

$$-a_{s+1} = \lambda_1 a_1 + \lambda_2 a_2 + \ldots + \lambda_s a_s \in C_s \subset C_{s+1}.$$

Пусть $P \colon \mathbb{R}^n \to (a_{s+1})^\perp$ — проекция на ортогональное дополнение к одномерному подпространству, натянутому на вектор a_{s+1}. Тогда

$$PC_{s+1} = \{\lambda_1 Pa_1 + \lambda_2 Pa_2 + \ldots + \lambda_s Pa_s \colon \lambda_i \geq 0\}$$

в силу равенства $P(\lambda_1 a_1 + \lambda_2 a_2 + \ldots + \lambda_{s+1} a_{s+1}) = \lambda_1 Pa_1 + \lambda_2 Pa_2 + \ldots + \lambda_s Pa_s$. Более того, $PC_{s+1} \subset C_{s+1}$, так как $Pa_i = a_i - \dfrac{(a_i, a_{s+1})}{\|a_{s+1}\|^2} a_{s+1}$. Согласно предположению индукции PC_{s+1} — замкнутое множество. Наконец, так как $C_{s+1} = PC_{s+1} + \mathbb{R}a_{s+1}$, то согласно лемме 4.1 множество C_{s+1} замкнуто. $\triangleleft$

Лемма 4.3 (Фаркаш)**.** *Пусть $A \in \mathbb{R}^{n\times m}$ и $b \in \mathbb{R}^n$. Тогда либо $Ax = b$, для некоторого $x \in \mathbb{R}^m_+$, либо найдётся такой вектор $y \in \mathbb{R}^n$, что $y^\top A \leq \mathbf{0}$ и $y^\top b > 0$.*

$\triangleright$ Пусть выполнена первая из альтернатив, т.е. существует вектор $x \in \mathbb{R}^m_+$, такой что $Ax = b$. Предположим, что $y^\top A \leq 0$ для некоторого вектора $y \in \mathbb{R}^n$, тогда $y^\top b = y^\top Ax \leq 0$.

Предположим теперь, что такого $x \in \mathbb{R}^m_+$ не существует. Рассмотрим выпуклый конус $C = \{Ax \colon x \in \mathbb{R}^m_+\}$, который согласно лемме 4.2 является замкнутым множеством. По предположению $b \notin C$, а значит, точка b строго отделима от C, т.е. существуют ненулевой вектор $y \in \mathbb{R}^n$ и число d, такие что $y^\top b > d > y^\top c$, $c \in C$. Так как $\mathbf{0} \in C$, то $y^\top b > d > 0$. С другой стороны $d \geq y^\top Ax = (A^\top y)^\top x$. Так как компоненты вектора x могут быть сколь угодно большими, то $y^\top A \leq \mathbf{0}$. Таким образом, выполнена вторая альтернатива. $\triangleleft$

Следствие 4.1. *Если векторы b, a_1, a_2, $\ldots$, $a_m \in \mathbb{R}^n$, такие что для каждого вектора $x \in \mathbb{R}^n$ из неравенств $x^\mathsf{T} a_i \geq 0$, $1 \leq i \leq m$, следует $x^\mathsf{T} b \geq 0$, то найдутся такие числа $\lambda_i \geq 0$, что $b = \sum\limits_{i=1}^{m} \lambda_i a_i$.*

Следствие 4.2. *Пусть $\mathbb{R}^n = V \oplus V^\perp$. Если векторы b, a_1, a_2, $\ldots$, $a_s \in \mathbb{R}^n$, такие что для каждого вектора $x \in V^\perp$ из неравенств $x^\mathsf{T} a_i \geq 0$, $1 \leq i \leq m$, следует неравенство $x^\mathsf{T} b \geq 0$, то найдутся такие числа $\lambda_i \geq 0$ и вектор $v \in V$, что $b = v + \sum\limits_{i=1}^{m} \lambda_i a_i$.*

$\triangleright$ Пусть b', a'_1, $a'_2, \ldots, a'_m \in V^\perp$ — проекция векторов b, a_1, $a_2, \ldots, a_m$ на подпространство $V^\perp$. Тогда для любого вектора $x \in \mathbb{R}^n$ из неравенств $x^\mathsf{T} a'_i \geq 0$ следует $x^\mathsf{T} b' \geq 0$, а значит, найдутся такие числа $\lambda_i \geq 0$, что $b' = \sum\limits_{i=1}^{m} \lambda_i a'_i$. Возвращаясь к исходным векторам, получаем, что $b = v + \sum\limits_{i=1}^{m} \lambda_i a_i$, где $v \in V$. $\triangleleft$

Пример 4.3 (Арбитраж)**.** Предположим, что имеется две биржи, на которых продаются и покупаются товары n различных видов. Торговец хочет заработать, покупая товары на одной бирже и продавая их на второй. В зависимости от конъюнктуры возможны m ситуаций. Через c_{ij} обозначим разницу цен за единицу товара i при наступлении ситуации j (от цены на второй бирже отнимается цена на первой). Произвольный вектор $y = (y_1, y_2, \ldots, y_n)^\mathsf{T} \in \mathbb{R}^n$ назовём стратегией торговца, где через y_i обозначено количество товара i, которое покупает и продаёт торговец. Доход торговца для стратегии y в ситуации j, очевидно, равен $\sum\limits_{i=1}^{n} c_{ij} y_i$.

Теорема 4.1 (Де Финетти)**.** *Верно ровно одно из следующих утверждений:*

1. *существует такое распределение $p = (p_1, p_2, \ldots, p_m)$, $p_j \geq 0$ и $\sum\limits_{j=1}^{m} p_j = 1$, на множестве ситуаций, что математическое ожидание дохода торговца для любого товара равно нулю, т.е. $\sum\limits_{j=1}^{m} c_{ij} p_j = 0$, $1 \leq i \leq n$;*
2. *существует такая стратегия y, что доход торговца положительный вне зависимости от ситуации, т.е. $\sum\limits_{i=1}^{n} c_{ij} y_i > 0$, $1 \leq j \leq m$.*

$\triangleright$ Рассмотрим вектор $b = (0, 0, \ldots, 0, -1)^\mathsf{T} \in \mathbb{R}^{n+1}$ и матрицу размера $(n+1) \times m$

$$A = \begin{pmatrix} c_{11} & \ldots & c_{1m} \\ \vdots & \ddots & \vdots \\ c_{n1} & \ldots & c_{nm} \\ -1 & \ldots & -1 \end{pmatrix}$$

Если какой-либо вектор $p \in \mathbb{R}^m_+$ удовлетворяет системе $Ap = b$, то справедливы равенства $\sum\limits_{j=1}^{m} c_{ij} p_j = 0$, $1 \leq i \leq n$, и $\sum\limits_{j=1}^{m} -p_j = -1$, т.е. p — искомое распределение. Согласно лемме Фаркаша либо такой вектор p существует, либо для некоторого $\widetilde{y} \in \mathbb{R}^{n+1}$ верно $\widetilde{y}^\mathsf{T} A \geq \mathbf{0}$,

$\widetilde{y}^{\mathsf{T}} b < 0$. Согласно определению матрицы A и вектора b имеем

$$\sum_{i=1}^{n} c_{ij}\widetilde{y}_i \geq \widetilde{y}_{n+1}, \quad 1 \leq j \leq m, \quad \text{и} \quad -\widetilde{y}_{n+1} < 0.$$

Другими словами, стратегия $y = (\widetilde{y}_1, \widetilde{y}_2, \ldots, \widetilde{y}_n)^{\mathsf{T}}$ искомая. $\lhd$

Упражнения

13. Решите следующую задачу линейного программирования

$$\begin{cases} c^{\mathsf{T}} x \to \max; \\ x_1 + x_2 + \ldots + x_n = 1; \\ 0 \leq x_i \leq 1, \quad 1 \leq i \leq n. \end{cases}$$

14. (Гордон) Докажите, что для произвольной матрицы $A \in \mathbb{R}^{n\times m}$ выполнена ровно одна из альтернатив:
    (a) существует такой вектор $x \in \mathbb{R}^m$, что $Ax < \mathbf{0}$;
    (b) существует такой ненулевой вектор $y \in \mathbb{R}^n$, что $A^{\mathsf{T}} y = \mathbf{0}$ и $y \geq \mathbf{0}$.
15. Пусть $P = (p_{ij})_{i,j=1}^n$ — стохастическая (слева) матрица, т.е. матрица с неотрицательными элементами $p_{ij} \geq 0$ и для всех j от 1 до n выполнено равенство $\sum\limits_{i=1}^{n} p_{ij} = 1$ (другими словами, в каждом столбце сумма стоящих в нём элементов равна 1). Используя лемму Фаркаша, докажите, что существует такой вектор $y = (y_1, y_2, \ldots, y_n)^{\mathsf{T}} \in \mathbb{R}^n_+$, что $Py = y$ и $\sum\limits_{i=1}^{n} y_i = 1$.
16. Докажите, что для того чтобы точка $x^* \in X$ была точкой минимума выпуклой дифференцируемой функции f на множестве $X = \{x\colon a_j^{\mathsf{T}} x \leq b_j\}$ необходимо и достаточно существование таких чисел $y_j \geq 0$, что

$$-\nabla f(x^*) = \sum_{j\in I(x^*)} y_j a_j.$$