НАО «Восточно-Казахстанский технический университет имени Д. Серикбаева»

Дисциплина «Методы оптимизации»

ЛЕКЦИЯ 3

Симплексный метод решения задачи линейного программирования

Тема 3 по рабочей программе. Объём — 2 академических часа. Образовательная программа 6B06103 Математическое и компьютерное моделирование.

1 Цель и связь с предыдущей лекцией

В лекции 2 задача линейного программирования была приведена к канонической форме. Теперь требуется найти оптимальный допустимый план. Симплекс-метод решает эту задачу последовательной заменой базисных переменных. На каждом шаге сохраняются ограничения, а значение целевой функции не уменьшается. Если шаг положителен и выбранная оценка положительна, значение функции строго возрастает.

После изучения материала студент должен уметь строить базисное решение, вычислять оценки небазисных переменных, выбирать входящую и выходящую переменные, обосновывать длину шага и проверять оптимальность результата. Необходимы решение систем линейных уравнений, понятия ранга и обратимой матрицы, а также преобразования из лекции 2.

Табличная организация вычислений рассматривается в теме 4, а построение начального допустимого базиса — в теме 5. Геометрическую идею и вычислительную схему можно сопоставить с разделом 2.3 практикума [1].

2 Каноническая задача и базис

Сохраним обозначения предыдущей лекции:

$$f(x) = c^T x \to \max, \qquad Ax = b, \qquad x \geq 0.$$

Здесь $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, $c, x \in \mathbb{R}^n$. Будем считать, что $\operatorname{rank} A = m \leq n$. Линейно зависимые уравнения сначала проверяют на совместность; избыточные ограничения удаляют, а противоречивые свидетельствуют об отсутствии допустимых решений.

Выберем m линейно независимых столбцов матрицы A. Обозначим множество их индексов через I, остальные индексы — через J. Матрица выбранных столбцов $B = A_I$ называется базисной. Она квадратная и обратимая. Переменные x_I называются базисными, а x_J — небазисными. Порядок индексов в I должен совпадать с порядком столбцов в B.

Разобьём систему и выразим базисные переменные:

$$Bx_I + A_J x_J = b, \qquad x_I = B^{-1}b - B^{-1}A_J x_J.$$

Базисное решение получается при $x_J = 0$:

$$\overline{b} = B^{-1}b, \qquad x_I = \overline{b}, \qquad x_J = 0.$$

Оно является **базисным допустимым решением**, если $\overline{b} \geq 0$. Обратимость B сама по себе не обеспечивает допустимости: отрицательная компонента $B^{-1}b$ нарушает ограничение $x \geq 0$.

Обозначение	Смысл	Размер
I, J	Индексы базисных и небазисных переменных	m и $n-m$ индексов
$B = A_I$	Базисная матрица	$m \times m$
$\overline{b} = B^{-1}b$	Значения базисных переменных	m
$\overline{A}_J = B^{-1}A_J$	Коэффициенты выражения базисных переменных	$m \times (n-m)$
c_I, c_J	Коэффициенты целевой функции	m и $n-m$

Если все компоненты $\overline{b}$ положительны, базисное допустимое решение невырожденное. Если хотя бы одна равна нулю, решение вырожденное. При вырождении одна геометрическая точка может соответствовать нескольким базисам.

3 Почему метод работает с вершинами

Допустимое множество $D = \{x: Ax = b,\ x \geq 0\}$ выпукло. Точка $x \in D$ является крайней, если её нельзя представить как середину двух различных допустимых точек. Для канонической задачи крайние точки и базисные допустимые решения совпадают.

Покажем одну сторону этого утверждения. Пусть x — базисное допустимое решение с небазисными координатами $x_J = 0$ и $x = (u+v)/2$, где $u, v \in D$. Из неотрицательности следует $u_J = v_J = 0$. Поэтому $Bu_I = b$ и $Bv_I = b$. Обратимость B даёт $u_I = v_I = B^{-1}b$, следовательно, $u = v = x$.

Для обратной стороны рассмотрим столбцы A, соответствующие положительным координатам крайней точки. Если они линейно зависимы, существует ненулевой вектор d, для которого $Ad = 0$, а компоненты вне положительных координат равны нулю. При достаточно малом $t > 0$ обе точки $x + td$ и $x - td$ остаются неотрицательными и допустимыми. Это противоречит крайности. Значит, столбцы независимы; их можно дополнить до базиса ранга m, получив базисное допустимое решение.

Если допустимое множество канонической задачи непусто и максимум конечен, существует оптимальное базисное допустимое решение. Симплекс-метод организует поиск среди таких решений. При невырожденных шагах это движение между соседними вершинами; при вырождении базис может смениться без изменения самой точки.

4 Оценки небазисных переменных

Подставим выражение $x_I = \overline{b} - \overline{A}_J x_J$ в целевую функцию:

$$f(x) = c_I^T x_I + c_J^T x_J,$$

$$f(x) = c_I^T \overline{b} + (c_J^T - c_I^T \overline{A}_J) x_J.$$

Обозначим текущее значение через $f_0 = c_I^T \overline{b}$. Для каждого небазисного столбца a_j введём оценку, или приведённую стоимость:

$$\Delta_j = c_j - c_I^T B^{-1} a_j, \qquad j \in J.$$

Тогда

$$f(x) = f_0 + \sum_{j \in J} \Delta_j \, x_j.$$

Число Δ_j показывает изменение целевой функции на единицу увеличения x_j, когда базисные переменные одновременно изменяются так, чтобы сохранить равенства $Ax = b$. Поэтому выбирать переменную только по исходному коэффициенту c_j после первого шага нельзя.

Критерий оптимальности для принятого знака оценок. Если текущее базисное решение допустимо и $\Delta_j \le 0$ для всех $j \in J$, оно оптимально в задаче максимизации.

Доказательство непосредственно следует из $x_j \ge 0$: для любого допустимого плана $f(x) = f_0 + \sum \Delta_j x_j \le f_0$. Текущий план достигает f_0. Это глобальная, а не только локальная проверка.

В некоторых учебниках используется противоположный знак оценки $\tilde{\Delta}_j = c_I^T B^{-1} a_j - c_j$. Тогда знаки критериев меняются. Во всех формулах и программах данного комплекта используется $\Delta_j = c_j - c_I^T B^{-1} a_j$.

5 Допустимое направление и длина шага

Пусть найдена небазисная переменная x_s с $\Delta_s > 0$. Будем увеличивать её от нуля, сохраняя остальные небазисные переменные равными нулю. Найдём вектор h из системы $Bh = a_s$, то есть $h = B^{-1} a_s$.

$$x_s(\theta) = \theta, \qquad x_I(\theta) = \overline{b} - h\theta, \qquad x_j(\theta) = 0 \quad (j \in J,\ j \ne s).$$

При любом θ равенства $Ax = b$ сохраняются. Допустимость ограничивается условиями $\theta \ge 0$ и $\overline{b}_i - h_i \theta \ge 0$. Из них получаем три случая:

1. При $h_i > 0$ базисная переменная убывает и требуется $\theta \le \overline{b}_i / h_i$.
2. При $h_i = 0$ данная компонента не меняется.
3. При $h_i < 0$ данная компонента возрастает, поэтому верхней границы шага эта строка не создаёт.

Если есть хотя бы одно $h_i > 0$, максимальный допустимый шаг равен

$$\theta^* = \min_{i : h_i > 0} \frac{\overline{b}_i}{h_i}.$$

Это правило минимального отношения. Отношения с нулевым или отрицательным знаменателем в минимум не включаются. Числитель может быть нулём: тогда допустим нулевой шаг.

Пусть минимум достигается в строке r. На границе шага переменная x_{I_r} обращается в нуль. Она выходит из базиса, а x_s входит. Новый базис получается заменой столбца B_r на a_s. Поскольку $h_r > 0$, такая замена сохраняет обратимость базисной матрицы.

Приращение цели составляет

$$f_{\text{new}} - f_0 = \Delta_s \theta^* \geq 0.$$

Если $\theta^* > 0$, цель строго возрастает. При $\theta^* = 0$ значение цели и точка могут сохраниться, хотя набор базисных столбцов изменится.

6 Признак неограниченности

Если $\Delta_s > 0$, но $h_i \leq 0$ для всех строк, увеличить x_s можно на любое $\theta \geq 0$. Все базисные компоненты остаются неотрицательными, а $f(x(\theta)) = f_0 + \Delta_s \theta \to +\infty$.

Получен не только вывод об отсутствии конечного максимума, но и его проверяемое свидетельство — допустимый луч:

$$d_s = 1, \qquad d_I = -h, \qquad d_j = 0 \ (j \in J,\ j \neq s),$$

$$Ad = 0, \qquad d \geq 0, \qquad c^T d = \Delta_s > 0.$$

Например, для $\max(x_1 + x_2)$ при $x_1 - x_2 \leq 1, x \geq 0$ допустимы точки $(1 + t, t), t \geq 0$. Целевая функция равна $1 + 2t$ и не ограничена сверху. Допустимость задачи и наличие конечного оптимума — разные свойства.

7 Алгоритм и правило выбора переменных

Входные данные прямого симплекс-метода — каноническая задача и допустимый базис. Если известен только произвольный базис, сначала нужно проверить $B^{-1}b \geq 0$. Если проверка не пройдена, требуется отдельная процедура поиска допустимого старта; это ещё не доказывает несовместность задачи.

Алгоритм состоит из следующих шагов:

1. Сформировать B и решить $B\overline{b} = b$.
2. Решить $B^T \pi = c_I$ и вычислить $\Delta_j = c_j - \pi^T a_j$.
3. Если все небазисные оценки неположительны, выдать оптимальный план.
4. Выбрать индекс $s \in J$ с $\Delta_s > 0$ и решить $Bh = a_s$.
5. Если $h \leq 0$, выдать признак неограниченности и направление d.
6. Найти θ^* по минимальному отношению, выбрать выходящую переменную и заменить столбец базиса.
7. Вернуться к первому шагу.

Формулы с B^{-1} удобны для вывода. В программной реализации достаточно решать системы с матрицами B и B^T; отдельно вычислять всю обратную матрицу не требуется.

Вырождение и конечность. При положительном шаге одно и то же значение цели повториться не может, но нулевые шаги делают возможным зацикливание при неудачном правиле выбора. Для учебной программы используем правило Бланда: входящей становится переменная с наименьшим индексом среди имеющих положительную оценку; при равенстве минимальных отношений выходит переменная с наименьшим собственным индексом. Это правило предотвращает зацикливание в точной арифметике [2, раздел 2.4.2].

Минимальный индекс выходящей переменной относится к её номеру I_r, а не к номеру строки r. На практике вычислительные реализации также учитывают погрешности округления,

масштаб коэффициентов и допуски проверок. В приложенной учебной программе эти вопросы отделены от логики метода использованием точных дробей.

8 Полный пример с двумя заменами базиса

Продолжим задачу из лекции 2:

$$\max f = 3x_1 + 2x_2, \qquad \begin{cases} x_1 + x_2 \le 4, \\ 2x_1 + x_2 \le 5, \\ x_1, x_2 \ge 0. \end{cases}$$

Введём остатки ресурсов $x_3, x_4 \ge 0$:

$$\begin{cases} x_1 + x_2 + x_3 = 4, \\ 2x_1 + x_2 + x_4 = 5. \end{cases} \qquad A = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 2 & 1 & 0 & 1 \end{pmatrix}.$$

8.1 Начальный план

Выберем $I = (3,4)$. Тогда B — единичная матрица, $x_3 = 4$, $x_4 = 5$, а $x_1 = x_2 = 0$. План допустим, $f_0 = 0$. Так как $c_I = (0,0)^T$, оценки равны $\Delta_1 = 3$ и $\Delta_2 = 2$. По правилу Бланда входит x_1.

$$h = B^{-1}a_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \qquad \theta^* = \min\left\{\frac{4}{1}, \frac{5}{2}\right\} = \frac{5}{2}.$$

Из базиса выходит x_4. После шага

$$x^{(1)} = \left(\frac{5}{2}, 0, \frac{3}{2}, 0\right)^T, \qquad f_1 = \frac{15}{2} = 7{,}5.$$

Обратите внимание: минимальное отношение задаёт ресурс, который исчерпывается первым. Первый ресурс ещё имеет остаток $3/2$; второй уже исчерпан.

8.2 Пересчёт выражений после первого шага

Новый порядок базиса $I = (3,1)$, небазисные переменные x_2, x_4. Из второго уравнения выразим x_1, затем подставим его в первое и в целевую функцию:

$$\begin{aligned} x_1 &= \frac{5}{2} - \frac{1}{2}x_2 - \frac{1}{2}x_4, \\ x_3 &= \frac{3}{2} - \frac{1}{2}x_2 + \frac{1}{2}x_4, \\ f &= \frac{15}{2} + \frac{1}{2}x_2 - \frac{3}{2}x_4. \end{aligned}$$

Следовательно, $\Delta_2 = 1/2 > 0$, $\Delta_4 = -3/2 < 0$. Входит x_2. При её увеличении обе базисные переменные убывают с коэффициентом $1/2$:

$$\theta^* = \min\left\{\frac{3/2}{1/2}, \frac{5/2}{1/2}\right\} = \min\{3,5\} = 3.$$

Выходит x_3. Получаем $x_1 = 1$, $x_2 = 3$, $x_3 = x_4 = 0$. Приращение цели $\Delta_2\theta^* = 3/2$ согласуется с $f_2 = 7{,}5 + 1{,}5 = 9$.

8.3 Проверка оптимальности

В финальном базисе $I = (2,1)$ выразим основные переменные через x_3, x_4:

$$\begin{aligned} x_1 &= 1 + x_3 - x_4, \\ x_2 &= 3 - 2x_3 + x_4, \\ f &= 9 - x_3 - x_4. \end{aligned}$$

Обе небазисные оценки равны -1. Поскольку $x_3, x_4 \geq 0$, для любого допустимого плана $f \leq 9$. План (1,3) достигает этой границы и оптимален. Строго отрицательные оценки дополнительно показывают единственность: равенство $f = 9$ требует $x_3 = x_4 = 0$, после чего x_1, x_2 определены однозначно.

Независимая проверка получается сложением исходных неравенств:

$$(x_1 + x_2) + (2x_1 + x_2) \leq 4 + 5, \qquad 3x_1 + 2x_2 \leq 9.$$

Шаг	Базис по порядку	x_1	x_2	f	Действие
0	(x_3, x_4)	0	0	0	Входит x_1, выходит x_4
1	(x_3, x_1)	$5/2$	0	$15/2$	Входит x_2, выходит x_3
2	(x_2, x_1)	1	3	9	Все оценки неположительны

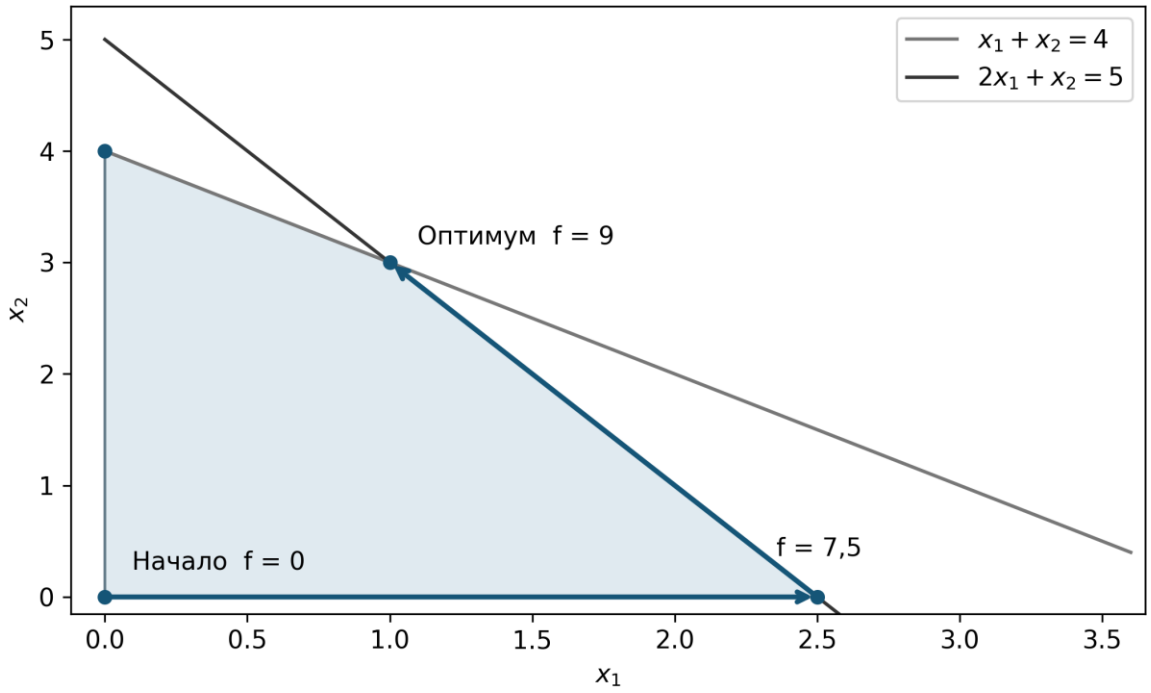


Рисунок 1 — Движение от начального плана к оптимальному

9 Особые ситуации

Нулевой шаг. В задаче $\max(x_1 + x_2)$ при $x_1 \leq 0, x_2 \leq 1, x \geq 0$ начальный план имеет нулевой запас по первому ограничению. При выборе x_1 минимальное отношение равно нулю. Переменная может войти в базис, но значение цели останется прежним. Последующее увеличение x_2 приводит к оптимуму (0,1).

Несколько оптимальных планов. В задаче $\max(x_1 + x_2)$ при $x_1 + x_2 \le 4, x \ge 0$ все точки $(4 - t, t)$, $0 \le t \le 4$, оптимальны. В оптимальном базисе имеется небазисная переменная с нулевой оценкой, которую можно увеличить на положительный допустимый шаг. Одна нулевая оценка без проверки возможности положительного шага ещё не доказывает существования другой оптимальной точки.

Недопустимый начальный базис. Отрицательная компонента $B^{-1}b$ означает, что выбранный базис не подходит для прямого симплекс-метода. Следует найти допустимый старт. Метод искусственного базиса будет рассмотрен в лекции 5.

Результат проверки	Математический смысл	Действие
$\overline{b} \ge 0$, все $\Delta_j \le 0$	Текущий план оптимален	Завершить расчёт
Есть $\Delta_s > 0$ и $h \le 0$	Цель не ограничена сверху	Выдать допустимый луч
Есть $\Delta_s > 0$ и положительные h_i	Возможна замена базиса	Применить минимальное отношение
Есть $\overline{b}_i < 0$	Текущий базис недопустим	Найти допустимый начальный базис

10 Проверка примера на Python

В архиве находится файл simplex_exact.py. Он реализует описанный алгоритм с правилом Бланда и типом Fraction из стандартной библиотеки Python. Программа принимает уже приведённую к каноническому виду задачу и заданный допустимый базис. Фаза поиска начального базиса в этой учебной версии не реализована.

После распаковки архива выполните команду python simplex_exact.py. Для использования функции из другого файла достаточно следующего фрагмента:

```
from simplex_exact import simplex

A = [[1, 1, 1, 0], [2, 1, 0, 1]]
b = [4, 5]
c = [3, 2, 0, 0]
result = simplex(A, b, c, basis=[2, 3])
for row in result['history']:
    print(row['k'], [str(v) for v in row['x']],
          str(row['value']))
```

В Python нумерация начинается с нуля: basis=[2, 3] соответствует математическому базису (x_3, x_4). Для точных нецелых входных коэффициентов используйте строки, например ‘1/3’, либо Fraction. Передача десятичного float сохраняет его двоичное приближение, а не задуманную рациональную дробь.

Контрольный результат: значения цели $0, 15/2, 9$; конечный вектор (1,3,0,0). Для каждой строки истории проверьте $Ax = b, x \ge 0$ и совпадение $c^T x$ с напечатанной целью. При статусе optimal оценки должны быть неположительными; при unbounded дополнительно проверяют направление луча.

11 Задания для самостоятельной работы

1. Для основного примера начните с базиса (x_3, x_1). Найдите B, $B^{-1}b$, оценки и следующий шаг без обращения к готовым преобразованиям раздела 8.
2. Решите задачу $\max(2x_1 + x_2)$ при $x_1 + x_2 \le 4, x_1 \le 2, x \ge 0$, начиная с базиса дополнительных переменных. Подтвердите ответ отдельной верхней оценкой цели.
3. Для $\max(x_1 + x_2)$ при $x_1 + x_2 \le 4$, $x \ge 0$ найдите два различных оптимальных базисных решения и опишите всё множество оптимумов.
4. Для $\max(x_1 + x_2)$ при $x_1 - x_2 \le 1, x \ge 0$ выпишите допустимый луч и покажите рост цели вдоль него.
5. Постройте начальный базис с нулевой базисной переменной для задачи из раздела 9 и выполните нулевой шаг. Объясните, почему это не является арифметической ошибкой.

Ориентиры для проверки. В задании 1: $\overline{b} = (3/2, 5/2)^T$, входит x_2, выходит x_3, шаг 3. В задании 2: $(x_1, x_2) = (2,2)$, максимум 6; оценка $2x_1 + x_2 = (x_1 + x_2) + x_1 \le 4 + 2$. В задании 3: отрезок $(4 - t, t)$, $0 \le t \le 4$. В задании 4: $(1 + t, t)$, $t \ge 0$. В задании 5 первая цель остаётся равной нулю, после положительного шага по x_2 максимум равен 1.

12 Контрольные вопросы

1. Чем базисное решение отличается от базисного допустимого?
2. Почему выбранные базисные столбцы должны быть линейно независимыми?
3. Как связаны крайние точки и базисные допустимые решения?
4. Как вывести формулу приведённой стоимости?
5. Почему неположительные оценки обеспечивают глобальную оптимальность?
6. Почему после смены базиса недостаточно смотреть на исходные коэффициенты цели?
7. Из каких условий получается правило минимального отношения?
8. Почему строки с $h_i \le 0$ не ограничивают увеличение входящей переменной?
9. Как построить свидетельство неограниченности задачи?
10. Почему при вырождении возможен нулевой шаг?
11. Как правило Бланда выбирает входящую и выходящую переменные?
12. Какие проверки подтверждают ответ компьютерной программы?

Литература и материалы для чтения

1. Казанская О. В., Юн С. Г., Альсова О. К. Модели и методы оптимизации. Практикум. — Новосибирск: НГТУ, 2012. — 204 с. Раздел 2.3, с. 86–101. Геометрическое объяснение — с. 86–87; вычислительная схема — с. 88 и далее.
2. Goemans M. X. Linear programming. Учебные материалы MIT. Раздел 2.4.2 о правиле Бланда и конечности симплекс-метода. https://math.mit.edu/~goemans/18310S15/lpnotes310.pdf.