

21

Численные методы квадратичного программирования

В. А. ДАУГАВЕТ

ИЗДАТЕЛЬСТВО
САНКТ-ПЕТЕРБУРГСКОГО
УНИВЕРСИТЕТА





12 - 620

23

Научная библиотека СПбГУ



1001062120

167

САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

В. А. ДАУГАВЕТ

ЧИСЛЕННЫЕ МЕТОДЫ
КВАДРАТИЧНОГО
ПРОГРАММИРОВАНИЯ

Учебное пособие



ИЗДАТЕЛЬСТВО С.-ПЕТЕРБУРГСКОГО УНИВЕРСИТЕТА

2004

УДК 519.8
ББК 22.18
Д21

12-620

23

Рецензенты: д-р физ.-мат. наук, проф. В. М. Рябов (С.-Петербург. гос. ун-т),
канд. физ.-мат. наук, проф. Л. В. Васильев (С.-Петербург. техн.
ун-т)

Печатается по постановлению
Редакционно-издательского совета
С.-Петербургского государственного университета

Даугавет В. А.

Д21 Численные методы квадратичного программирова-
ния: Учеб. пособие. — СПб.: Изд-во С.-Петерб. ун-та, 2004. —
128 с.
ISBN 5-288-03434-6

Пособие, посвященное методам решения задач выпуклого квадратичного программирования, написано на основе спецкурса, читаемого студентам математико-механического факультета СПбГУ. Особое внимание в курсе уделяется обсуждению принципов, положенных в основу методов. В связи с этим в каждой главе рассматриваются группы базисных методов, объединенных одной глобальной идеей. Приводятся описания и строгие обоснования методов Данцига, дополнительного базиса (разработан автором), перебора граней и др.

Пособие предназначено для студентов университетов соответствующих специальностей.

Библиогр. 22 назв. Ил. 2.

ББК 22.18

БИБЛИОТЕКА
Факультета ПМ-ПУ
СПбГУ

ISBN 5-288-03434-6

© В. А. Даугавет, 2004
© Издательство
С.-Петербургского
университета, 2004

ОСНОВНЫЕ ОБОЗНАЧЕНИЯ¹

$\mathbb{N} = \{\dots, -2, -1, 0, 1, 2, \dots\}$ — множество целых чисел;

M, N, P, \dots — произвольные индексные множества (конечные наборы целых чисел);

$|M|$ — число элементов, содержащихся в M ;

\emptyset — пустое множество;

$x = x[N]$ — вектор с компонентами $x[j]$, $j \in N$;

$\mathbb{0} = \mathbb{0}[N]$ — нулевой вектор;

$\mathbb{1} = \mathbb{1}[N]$ — вектор, все компоненты которого равны единице;

$\langle x, y \rangle = x[N] \cdot y[N] = \sum_{j \in N} x[j] \cdot y[j]$ — скалярное произведение векторов x и y ;

$A = A[M, N]$ — матрица с компонентами $A[i, j]$, $i \in M$, $j \in N$;

$A^T = A^T[N, M]$ — транспонированная матрица с компонентами

$$A^T[j, i] = A[i, j], \quad i \in M, \quad j \in N;$$

$\mathbb{0} = \mathbb{0}[M, N]$ — матрица с нулевыми компонентами;

$E = E[N, N]$ — единичная матрица, у которой $E[j, j] = 1$ при всех $j \in N$ и $E[j, k] = 0$ при $k \neq j$;

$A[M_1, N_1]$, где $M_1 \subset M$, $N_1 \subset N$, — подматрица $A[M, N]$;

$A[M, j]$ — j -й столбец матрицы $A[M, N]$;

$A[i, N]$ — i -я строка матрицы $A[M, N]$;

$v = Ax = A[M, N] \cdot x[N]$ — вектор с компонентами

$$v[i] = A[i, N] \cdot x[N] = \sum_{j \in N} A[i, j] \cdot x[j], \quad i \in M;$$

$y = uA = u[M] \cdot A[M, N]$ — вектор с компонентами

$$y[j] = u[M] \cdot A[M, j] = \sum_{i \in M} u[i] \cdot A[i, j], \quad j \in N;$$

$C = AB = A[M, N] \cdot B[N, P]$ — матрица с компонентами

$$C[i, k] = A[i, N] \cdot B[N, k], \quad i \in M, \quad k \in P.$$

Если $N_1 \subset N$, $N_2 = N \setminus N_1$, то имеют место равенства

$$\langle x, y \rangle = x[N] \cdot y[N] = x[N_1] \cdot y[N_1] + x[N_2] \cdot y[N_2],$$

¹См. [2, 3, 14].

$$A[M, N] \cdot y[N] = A[M, N_1] \cdot y[N_1] + A[M, N_2] \cdot y[N_2].$$

Если $N_1 \subset N$, то

$$E[N_1, N] \cdot x[N] = x[N_1], \quad x[N] \cdot E[N, N_1] = x[N_1].$$

Отметим, что векторы uA и $A^T u$ имеют одинаковые компоненты, поэтому

$$\langle u, Ax \rangle = \langle A^T u, x \rangle = \langle uA, x \rangle.$$

Пусть $N = \{1, 2, \dots, n\}$, $N^- = -N = \{-1, -2, \dots, -n\}$, тогда

$$x[N] \cdot y[N^-] = \sum_{j \in N} x[j] \cdot y[-j].$$

ВВЕДЕНИЕ

Задачей квадратичного программирования называется задача минимизации квадратичной функции на выпуклом многогранном множестве Ω :

$$F(x) \rightarrow \inf_{x \in \Omega \subset \mathbb{R}^n}. \quad (1)$$

Квадратичная функция может быть представлена в виде

$$F(x) = \frac{1}{2} \langle Dx, x \rangle + \langle c, x \rangle + \alpha,$$

где D — матрица порядка n , c — вектор из \mathbb{R}^n , α — число.

Заметим, что если матрица D не симметрична, то ее можно заменить на симметричную

$$D_0 = \frac{1}{2}(D + D^T),$$

так как $\langle D_0 x, x \rangle = \frac{1}{2}(\langle Dx, x \rangle + \langle D^T x, x \rangle) = \langle Dx, x \rangle$.

При решении задачи (1) константу α можно отбросить, она на решение не влияет.

Таким образом, нас интересуют способы решения задачи следующего вида:

$$F(x) := \frac{1}{2} \langle Dx, x \rangle + \langle c, x \rangle \rightarrow \inf_{x \in \Omega}, \quad (2)$$

где матрица $D[N, N]$ симметрична, а

$$\Omega = \left\{ x[N] \left| \begin{array}{l} A[M_1, N] \cdot x[N] \geq b[M_1], \\ A[M_2, N] \cdot x[N] = b[M_2], \\ x[N_1] \geq 0, \quad N_1 \subset N \end{array} \right. \right\}.$$

Более того, будем предполагать, что матрица D неотрицательно определена, так что функция $F(x)$ выпукла на Ω . Таким образом, мы будем рассматривать задачу *выпуклого квадратичного программирования* (ВКП). Частным случаем задачи (2), когда $D = \mathbb{O}$, является задача линейного программирования (ЛП).

Задачи ВКП на практике возникают очень часто и в самых разных ситуациях. Приведем некоторые из них.

а) Пусть требуется решить переопределенную систему линейных алгебраических уравнений

$$A[M, N] \cdot x[N] = b[M], \quad |M| > |N|.$$

Такая система, как правило, несовместна, и требуется найти вектор $x[N]$, который в некотором смысле наилучшим образом удовлетворяет этой системе, например такой, который минимизирует функционал

$$F(x) = \frac{1}{2} \|Ax - b\|^2.$$

Под нормой здесь понимается евклидова норма. Нетрудно видеть, что

$$F(x) = \frac{1}{2} \langle Dx, x \rangle + \langle c, x \rangle + \frac{1}{2} \langle b, b \rangle,$$

где $D = A^T A$, $c = -A^T b$. Ясно, что матрица D неотрицательно определена. При этом на переменные могут накладываться дополнительные линейные ограничения вида

$$\alpha[P] \leq B[P, N] \cdot x[N] \leq \beta[P].$$

Таким образом, решение переопределенной системы можно свести к решению задачи ВКП.

б) Пусть требуется найти функцию $w(t)$, минимизирующую интеграл

$$I(w) := \int_0^a ((w'(t))^2 + f(t)w(t)^2 + g(t)w(t)) dt$$

при условиях $w(0) = 0$, $w(a) = 0$.

Разделим промежуток $[0, a]$ на n частей равноотстоящими точками: $0 = t_0 < t_1 < \dots < t_n = a$. Положим

$$z_i = w(t_i), \quad f_i = f(t_i), \quad g_i = g(t_i), \quad h = \frac{a}{n}.$$

Заменим производную $w'(t_i)$ разностным отношением $\frac{z_{i+1} - z_i}{h}$, а интеграл — суммой

$$S(z) := h \sum_{i=0}^{n-1} \left(\left(\frac{z_{i+1} - z_i}{h} \right)^2 + f_i z_i^2 + g_i z_i \right).$$

Вместо исходной задачи будем решать дискретную задачу

$$S(z) \rightarrow \inf_{z \in \mathbf{R}^{(n+1)}}$$

при условиях $z_0 = 0$, $z_n = 0$. Очевидно, что эта задача является задачей ВКП. Ее решение при больших n является достаточно хорошим приближением к решению исходной задачи.

в) В методах нелинейного программирования задачи ВКП возникают при определении направления спуска, например в методах линеаризации [12].

Важно обратить внимание на то, что методы решения задач ВКП могут служить моделями, на основе которых разрабатываются методы нелинейного программирования. Действительно, дважды непрерывно дифференцируемая функция от многих переменных $F(x)$ в окрестности некоторой точки x_0 приближенно представляема в виде отрезка ряда Тейлора

$$F(x_0 + h) \approx F(x_0) + h^T F'(x_0) + \frac{1}{2} h^T F''(x_0) h.$$

Известно, что в окрестности точки минимума матрица Гессе $F''(x)$ неотрицательно определена. Грубо говоря, вблизи точки минимума функцию $F(x)$ можно приблизить выпуклой квадратичной функцией. Обратимся к краткому содержанию книги.

В главе 1 приводятся некоторые результаты из теории линейного и квадратичного программирования, необходимые для дальнейшего изложения методов. Основным моментом в этой главе является вывод условий Куна — Таккера (КТ) для задачи ВКП и выявление некоторых свойств системы КТ.

Глава 2 посвящена базисным методам решения задач ВКП. Известно, что решение задачи (2) сводится к решению системы КТ относительно переменных $x[N]$, $u[M]$, $v[M]$, $y[N]$:

$$\left. \begin{aligned} Dx - A^T u - y &= -c, \\ Ax &- v = b, \end{aligned} \right\} \quad (3)$$

$$\left. \begin{aligned} y[N \setminus N_1] &= \mathbb{O}, \quad v[M_2] = \mathbb{O}, \\ x[N_1] \geq \mathbb{O}, \quad u[M_1] \geq \mathbb{O}, \quad y[N_1] \geq \mathbb{O}, \quad v[M_1] \geq \mathbb{O}, \end{aligned} \right\} \quad (4)$$

$$\langle x, y \rangle = 0, \quad \langle u, v \rangle = 0. \quad (5)$$

Многие известные методы основаны именно на решении системы КТ. Система КТ, вообще говоря, является нелинейной из-за условий дополнителности (5). Способы решения системы КТ могут быть различны, что создает разнообразие методов. Среди многочисленных методов выделяется группа методов, основанных на переборе базисных решений линейной системы (3) до выполнения условий (4) и (5).

Для устранения нелинейных условий (5) Вулф [11, 22] предложил перебирать только такие решения системы (3), у которых одна из пары переменных $x[i]$, $y[i]$ и $u[j]$, $v[j]$ равна нулю. На таких решениях нелинейное условие (5) выполняется автоматически. Для задачи ВКП с ограничениями в первой канонической форме ($Ax = b$, $x \geq 0$) Вулф разработал метод в двух вариантах: простой вариант для частного случая, когда матрица D положительно определена или $c = 0$, и сложный вариант для общей задачи.

В работе [5] была предложена модификация метода Вулфа, в которой перебираются только дополнительные базисы системы (3), т.е. такие, базисное решение на которых удовлетворяет условию (5). Такой базисный подход позволил решать задачу ВКП с общей целевой функцией так же просто, как в указанном частном случае.

Эта же идея Вулфа использована и в методе Данцига для задач ВКП. В работе Данцига [4] при описании и обосновании метода использовалась *невыврожденность* базисных решений системы (3), что ограничивало применимость метода. В работе [7] базисный подход позволил применять метод Данцига к любой задаче ВКП в первой канонической форме.

Заметим, что в методах Вулфа и Данцига предварительно требуется найти базисный план системы $Ax = b$, $x \geq 0$. От этого свободен метод Лемке [1], в котором использована та же идея дополнителности для задач ВКП с ограничениями в симметричной форме: $Ax \geq b$, $x \geq 0$. Базисный подход также позволил применять метод Лемке к любой задаче ВКП в симметричной форме.

Перевод произвольных ограничений в какую-либо стандартную форму может сильно увеличить размерность задачи ВКП, что нежелательно при больших размерах задачи. В связи с этим был разработан метод дополнительного базиса (ДБ) для задач ВКП с любыми линейными ограничениями и неотрицательно определенной матрицей D [6, 8]. Метод использует базисный подход к идее Вулфа и не требует специального предварительного построения базисного плана, как в методах Вулфа и Данцига.

Перечисленные выше методы ВКП сводятся к поиску экстремального базиса среди дополнительных базисов системы КТ. Такие методы будут называться *базисными*. В главе 2 подробно и с полным обоснованием изложены два универсальных базисных метода: метод ДБ и метод Данцига из работы [7]. Методы Вулфа и Лемке представлены как частные случаи метода ДБ.

Обратим внимание на один нетрадиционный подход к описанию методов в данном курсе. В большинстве методов ЛП и ВКП на каждом шаге приходится решать систему линейных уравнений вида $Cx = d$ с невырожденной квадратной матрицей C . Разумеется, если порядок матрицы велик, то эффективность самого метода зависит от способа решения этой системы. Выбор способа решения системы во многом определяется видом матрицы C (она может быть сильно разреженной и иметь четко выраженную структуру). В базисных методах ЛП для решения системы традиционно используется обратная базисная матрица, которая на каждой итерации пересчитывается через предыдущую. Поскольку можно воспользоваться и другими алгоритмами решения системы, то описания методов ВКП будут проводиться без уточнения способа решения этой системы, но будут приведены алгоритмы методов и с использованием обратных базисных матриц.

В главе 3 представлена другая большая группа методов — *методы перебора граней*. К ним относятся методы: Бэста [16], Гилла и Мюррея [20], сопряженных градиентов [13], Флетчера [19] и др. Эти методы описаны для задач ВКП с ограничениями-неравенствами: $A[M, N] \cdot x[N] \geq b[M]$. На каждом шаге этих методов приходится решать линейную систему при некотором $J \subset M$:

$$\begin{aligned} D[N, N] \cdot x[N] - A^T[N, J] \cdot u[J] &= -c[N], \\ A[J, N] \cdot x[N] &= b[J]. \end{aligned}$$

При этом индексное множество J при переходе к следующему шагу может расширяться или сужаться. Как было показано Бэстом [16], перечисленные методы по существу различаются лишь способами решения этой системы. В главе 3 описан и обоснован общий для этих методов алгоритм без указания способа решения данной системы. Кроме того, описаны алгоритмы пересчета обратной матрицы системы при изменении индексного множества J .

В главе 4 рассматриваются итеративные методы, использующие другие подходы к решению системы КТ.

Глава 1 снабжена упражнениями с ответами, главы 2 и 3 — численными примерами.

Сравним задачу ВКП с задачей линейного программирования (ЛП). Формально задача (2) отличается от задачи ЛП лишь наличием квадратичного слагаемого в целевой функции, но это слагаемое меняет всю геометрию задачи. Существенным моментом в ЛП является тот факт, что решение можно искать среди вершин многогранного множества Ω . В задачах ВКП это не так. Например, минимум функции

$$F(x) = \frac{1}{2}(x[1]^2 + x[2]^2)$$

на квадрате с центром в нуле достигается в единственной точке $x_* = 0$, которая является *внутренней* точкой квадрата. Тем не менее, методы ВКП, пригодные для вырожденных матриц D , годятся и для задач ЛП, поэтому некоторые методы ЛП можно рассматривать как частный случай какого-либо метода ВКП. Например, традиционный симплекс-метод для задачи ЛП, в которой множество Ω имеет первую каноническую форму $Ax = b, x \geq 0$, совпадает с методом Данцига для задач ВКП с такими же ограничениями, если $D = 0$. Кроме того, известный метод последовательного улучшения плана для задач ЛП, в которых множество Ω имеет вторую каноническую форму (двойственную к первой) $Ax \geq b$, совпадает с методом перебора граней для ВКП, если $D = 0$.

Методы, представленные в главах 2 и 3, универсальны, они годятся для любой задачи ВКП, в частности для задач с вырожденной матрицей D .

Разумеется, не все известные идейные подходы к методам ВКП затронуты в данном пособии. В частности, не представлены идеи минских математиков Р. Габасова, Ф. М. Кирилловой и их учеников.

Автор благодарит В. Н. Малоземова за поддержку и внимание к этой работе.

Работа выполнена при частичной финансовой поддержке РФФИ (проект № 01-01-00231).

Глава 1

НЕОБХОДИМЫЕ СВЕДЕНИЯ ИЗ ТЕОРИИ ЛИНЕЙНОГО И КВАДРАТИЧНОГО ПРОГРАММИРОВАНИЯ

Пусть заданы индексное множество $N = \{1, 2, \dots, n\}$, симметричная неотрицательно определенная матрица $D[N, N]$, выпуклое многогранное множество Ω из \mathbb{R}^n . Рассмотрим экстремальную задачу

$$F(x) := \frac{1}{2}\langle Dx, x \rangle + \langle c, x \rangle \rightarrow \inf_{x \in \Omega}, \quad (1)$$

где множество Ω определяется следующей системой неравенств и равенств:

$$\begin{aligned} A[M_1, N] \cdot x[N] &\geq b[M_1], \\ A[M_2, N] \cdot x[N] &= b[M_2], \\ x[N_1] &\geq 0. \end{aligned} \quad (2)$$

Здесь $N_1 \subset N$, M_1, M_2 — индексные множества. Вектор $x[N]$ из Ω называется *планом* задачи (1), функция $F(x)$ называется *целевой функцией*. Решением или оптимальным планом задачи (1) называется вектор $x_* \in \Omega$, для которого $F(x_*) \leq F(x)$ для всех $x \in \Omega$.

Очевидно, что $F(x)$ выпукла на Ω , поэтому рассматриваемая задача является задачей выпуклого квадратичного программирования (ВКП), а при $D = 0$ — задачей линейного программирования (ЛП).

§1. Стандартные формы записи задач квадратичного программирования

Во многих случаях методы ВКП и ЛП описываются для задач, представленных в какой-либо стандартной форме. Выделим три стандартных формы записи задачи ВКП.

Первая каноническая форма:

$$\begin{aligned} F(x) &\rightarrow \inf, \\ A[M, N] \cdot x[N] &= b[M], \\ x[N] &\geq 0. \end{aligned}$$

Вторая каноническая форма:

$$F(x) \rightarrow \inf, \\ A[M, N] \cdot x[N] \geq b[M].$$

Симметричная форма:

$$F(x) \rightarrow \inf, \\ A[M, N] \cdot x[N] \geq b[M], \\ x[N] \geq 0.$$

Как известно, общую задачу (1), (2) можно привести к любой из стандартных форм с помощью следующих операций.

При переводе в первую каноническую форму неравенство при $i \in M_1$ можно превратить в равенство путем добавления новой переменной $v[i] \geq 0$, так что

$$A[i, N] \cdot x[N] - v[i] = b[i].$$

При этом увеличиваются порядок матрицы D и размерность вектора c , т.е. для переменной $v[i]$ к матрице D добавляются нулевые строка и столбец, а в вектор c добавляются соответствующие нулевые компоненты.

При переводе во вторую или симметричную формы равенство при $i \in M_2$ можно заменить двумя неравенствами

$$A[i, N] \cdot x[N] \geq b[i], \\ -A[i, N] \cdot x[N] \geq -b[i].$$

Чтобы все переменные сделать неотрицательными, переменные $x[i]$ для $i \in N_2 := N \setminus N_1$ можно представить в виде разности двух неотрицательных переменных $x[i] = x'[i] - x''[i]$ и подставить их в ограничения и целевую функцию. Очевидно, что порядок матрицы D и размерность вектора c увеличатся на $|N_2|$ единиц.

Таким образом, перевод задачи в какую-либо стандартную форму ведет к увеличению либо числа переменных, либо ограничений, либо того и другого. В некоторых случаях это увеличение может оказаться слишком большим, что затруднит численную реализацию выбранного метода. При решении конкретной задачи выбор метода целесообразно связывать с минимальным увеличением размеров задачи.

§2. Признаки разрешимости задачи выпуклого квадратичного программирования

Напомним основной признак разрешимости задачи ВКП.

Теорема 2.1. *Задача (1) разрешима тогда и только тогда, когда множество ее планов Ω непусто и целевая функция $F(x)$ ограничена снизу на множестве планов.*

Доказательство можно найти в работе [3].

Выявим простые достаточные признаки разрешимости задачи (1) с непустым множеством планов.

Теорема 2.2. *Пусть $\Omega \neq \emptyset$. Задача (1) разрешима, если выполнено хотя бы одно из условий:*

- 1) множество Ω ограничено;
- 2) матрица D положительно определена;
- 3) линейная часть целевой функции $\langle c, x \rangle$ ограничена снизу на Ω .

Доказательство. Разрешимость задачи при выполнении первого условия следует из теоремы Вейерштрасса.

Пусть выполнено второе условие. Рассмотрим задачу без ограничений

$$F(x) := \frac{1}{2} \langle Dx, x \rangle + \langle c, x \rangle \rightarrow \min_{x \in \mathbb{R}^n}.$$

Так как матрица D неособенная, то решением этой задачи является вектор $x_0 = -D^{-1}c$, откуда следует неравенство

$$F(x) \geq F(x_0) \quad \forall x \in \mathbb{R}^n \supset \Omega.$$

Осталось воспользоваться теоремой 2.1.

Так как матрица D неотрицательно определена, то из третьего условия также следует ограниченность снизу целевой функции на Ω . ■

Пример. Приведем пример, когда для разрешимой задачи ВКП не выполнено ни одно из условий теоремы 2.2. Рассмотрим задачу

$$F(x) = x[1]^2 - 2x[1] + x[2] \rightarrow \inf_{x \geq 0}.$$

Разрешимость ее очевидна, так как целевая функция имеет представление $F(x) = (x[1] - 1)^2 + x[2] - 1$ и на множестве планов $F(x) \geq -1$. Эта нижняя граница реализуется на плане $x_* = (1, 0)$,

значит, x_* является решением. С другой стороны, в данной задаче матрица $D = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}$ вырождена, множество планов не ограничено, и на луче $x(t) = tx_*$, $t \geq 0$, принадлежащем множеству планов, линейная часть целевой функции не ограничена снизу: $-2x(t)[1] + x(t)[2] = -2t \rightarrow -\infty$ при $t \rightarrow \infty$.

Упражнения

2.1. Будет ли верно утверждение теоремы 2.1, если $F(x)$ — произвольная непрерывная выпуклая функция на Ω ?

2.2. Доказать, что если матрица D неотрицательно определена, то из равенства $\langle Dx, x \rangle = 0$ следует, что $Dx = 0$.

§3. Двойственность в линейном программировании

Напомним некоторые свойства задач линейного программирования, которые далее будут использоваться.²

Каждой задаче ЛП

$$\begin{aligned} f(x) &:= \langle d, x \rangle \rightarrow \min, \\ A[M_1, N] \cdot x[N] &\geq b[M_1], \\ A[M_2, N] \cdot x[N] &= b[M_2], \\ x[N_1] &\geq 0 \end{aligned} \quad (3.1)$$

ставится в соответствие *двойственная* задача ЛП

$$\begin{aligned} g(u) &:= \langle b, u \rangle \rightarrow \max, \\ u[M] \cdot A[M, N_1] &\leq d[N_1], \\ u[M] \cdot A[M, N_2] &= d[N_2], \\ u[M_1] &\geq 0, \end{aligned} \quad (3.2)$$

где $N_2 = N \setminus N_1$, $M = M_1 \cup M_2$. Очевидно, что задача (3.1) является двойственной к (3.2). Нетрудно проверить, что если $x[N]$ — план задачи (3.1), $u[M]$ — план задачи (3.2), то

$$f(x) \geq g(u). \quad (3.3)$$

²По поводу теорем двойственности для задачи ЛП удобнее всего обращаться к книге [3], так как в данном тексте используются такие же обозначения.

Сформулируем первую теорему двойственности.

Теорема 3.1. *Если одна из пары двойственных задач разрешима, то разрешима и другая. При этом значения целевых функций на оптимальных планах у обеих задач совпадают.*

Обозначим множество планов задачи (3.1) через Ω , а множество планов задачи (3.2) — через Λ .

Теорема 3.2. *Пара двойственных задач (3.1), (3.2) разрешима тогда и только тогда, когда множества их планов Ω и Λ непусты.*

Доказательство. Необходимость тривиальна. Докажем достаточность. Согласно (3.3) $f(x) \geq g(u_0) \quad \forall x \in \Omega$, где $u_0 \in \Lambda$. Осталось воспользоваться теоремой 2.1. ■

Из первой теоремы двойственности непосредственно следует первый критерий оптимальности планов пары двойственных задач: *если на планах $x_0 \in \Omega$ и $u_0 \in \Lambda$ значения целевых функций совпадают, т. е. $f(x_0) = g(u_0)$, то эти планы оптимальны.*

Сформулируем вторую теорему двойственности, которая представляет второй критерий оптимальности планов.

Теорема 3.3. *Планы двойственных задач x_0 и u_0 оптимальны тогда и только тогда, когда для них выполнены условия дополнителности («дополняющей нежесткости»):*

$$\begin{aligned} u_0[i] \cdot (b[i] - A[i, N] \cdot x_0[N]) &= 0 \quad \forall i \in M_1, \\ x_0[j] \cdot (u_0[M] \cdot A[M, j] - d[j]) &= 0 \quad \forall j \in N_1. \end{aligned}$$

Упражнения

Воспользовавшись теоремами двойственности, выполнить следующие упражнения.

3.1. Выяснить, при каких $c[N]$ разрешима задача

$$\begin{aligned} \langle c, x \rangle &\rightarrow \inf, \\ \sum_{i \in N} x[i] &= 1. \end{aligned}$$

3.2. Найти решение задачи

$$\begin{aligned} \langle c, x \rangle &\rightarrow \inf, \\ \sum_{i \in N} x[i] &= 1, \quad x[N] \geq 0. \end{aligned}$$

3.3. Доказать, что если задача

$$\langle c, x \rangle \rightarrow \inf, \\ A[M, N] \cdot x[N] = b[M]$$

разрешима, то любой план является ее решением.

§4. Условия оптимальности в форме Куна—Таккера

Рассмотрим общую задачу ВКП

$$F(x) := \frac{1}{2} \langle Dx, x \rangle + \langle c, x \rangle \rightarrow \min_{x \in \Omega}, \quad (4.1)$$

где

$$\Omega = \left\{ x[N] \left| \begin{array}{l} A[M_1, N] \cdot x[N] \geq b[M_1], \\ A[M_2, N] \cdot x[N] = b[M_2], \\ x[N_1] \geq 0 \end{array} \right. \right\}.$$

Нетрудно видеть, что $F'(x) = Dx + c$ и для любых x и h

$$F(x+h) = F(x) + \langle F'(x), h \rangle + \frac{1}{2} \langle Dh, h \rangle. \quad (4.2)$$

Выведем условия оптимальности для задачи ВКП.

Лемма 4.1. Для того чтобы вектор $x_* \in \Omega$ был решением задачи (4.1), необходимо и достаточно выполнение неравенства

$$\langle F'(x_*), (x - x_*) \rangle \geq 0 \quad \forall x \in \Omega. \quad (4.3)$$

Доказательство. Необходимость. Возьмем $x \in \Omega$, отличное от x_* , и положим $h = x - x_*$. Очевидно, что для любого $t \in (0, 1)$ точка $x_* + th$ принадлежит Ω , поэтому $F(x_* + th) - F(x_*) \geq 0$. Из этого неравенства и разложения (4.2) получаем при $t > 0$

$$\langle F'(x_*), h \rangle + t \frac{1}{2} \langle Dh, h \rangle \geq 0.$$

Устремляя t к нулю, видим, что $\langle F'(x_*), h \rangle \geq 0$.

Достаточность. Для произвольного $x \in \Omega$ и $h = x - x_*$ имеем

$$F(x) - F(x_*) = F(x_* + h) - F(x_*) = \langle F'(x_*), h \rangle + \frac{1}{2} \langle Dh, h \rangle.$$

Из неравенства (4.3) и неотрицательной определенности матрицы D следует, что правая часть последнего равенства неотрицательна, поэтому x_* — решение задачи (4.1). ■

Условие (4.3) для $x_* \in \Omega$ равносильно тому, что x_* является решением следующей задачи ЛП:

$$\langle F'(x_*), x \rangle \rightarrow \min, \\ A[M_1, N] \cdot x[N] \geq b[M_1], \\ A[M_2, N] \cdot x[N] = b[M_2], \\ x[N_1] \geq 0. \quad (4.4)$$

Запишем для нее двойственную задачу, полагая $M = M_1 \cup M_2$, $N_2 = N \setminus N_1$:

$$\langle b, u \rangle \rightarrow \max, \\ u[M] \cdot A[M, N_1] \leq F'(x_*)[N_1], \\ u[M] \cdot A[M, N_2] = F'(x_*)[N_2], \\ u[M_1] \geq 0. \quad (4.5)$$

В силу теоремы 3.3 планы пары двойственных задач оптимальны тогда и только тогда, когда для них выполнено условие дополненности. Запись этого условия приводит к признаку оптимальности задачи квадратичного программирования, который носит название условий Куна—Таккера (КТ).

Теорема 4.1 (Куна—Таккера). Для того чтобы вектор $x_*[N]$ был решением задачи (4.1), необходимо и достаточно существование векторов $u_*[M]$, $y_*[N]$, $v_*[M]$, удовлетворяющих следующей системе равенств и неравенств:

$$\begin{array}{l} D[N, N] \cdot x_*[N] - A^T[N, M] \cdot u_*[M] - y_*[N] = -c[N], \\ A[M, N] \cdot x_*[N] - v_*[M] = b[M], \\ y_*[N_2] = 0, \quad v_*[M_2] = 0, \\ x_*[N_1] \geq 0, \quad u_*[M_1] \geq 0, \quad y_*[N_1] \geq 0, \quad v_*[M_1] \geq 0, \\ x_*[N_1] \cdot y_*[N_1] = 0, \quad u_*[M_1] \cdot v_*[M_1] = 0. \end{array} \quad (4.6)$$

При этом $F(x_*) = \frac{1}{2} (\langle c, x_* \rangle + \langle b, u_* \rangle)$.

Доказательство. Необходимость. Пусть $x_*[N]$ — решение задачи (4.1), тогда этот же вектор является решением задачи (4.4). По первой теореме двойственности задача (4.5) тоже имеет решение $u_*[M]$. Положим

$$v_*[M] := A[M, N] \cdot x_*[N] - b[M],$$

$$y_*[N] := D[N, N] \cdot x_*[N] + c[N] - A^T[N, M] \cdot u_*[M].$$

Эти равенства совпадают с первыми двумя равенствами (4.6). Более того, так как $x_* \in \Omega$, а u_* — план двойственной задачи, то

$$x_*[N_1] \geq 0, \quad v_*[M_1] \geq 0, \quad v_*[M_2] = 0,$$

$$u_*[M_1] \geq 0, \quad y_*[N_1] \geq 0, \quad y_*[N_2] = 0.$$

По второй теореме двойственности для решений x_* и u_* выполняются условия дополнителности, которые в данном случае можно записать в виде

$$x_*[N_1] \cdot y_*[N_1] = 0, \quad u_*[M_1] \cdot v_*[M_1] = 0.$$

Итак, выполнение условий (4.6) для x_*, y_*, u_*, v_* доказано.

Определим теперь $F(x_*)$. Умножим первое равенство (4.6) слева на x_*^T , второе — на u_*^T и сложим их. Тогда с учетом третьего и последнего из условий (4.6) получим

$$\langle Dx_*, x_* \rangle = -\langle c, x_* \rangle + \langle b, u_* \rangle,$$

откуда и следует требуемое представление $F(x_*)$.

Достаточность. Из условий (4.6) следует, что x_* является планом задачи (4.4), u_* является планом двойственной задачи (4.5) и для этих планов выполнено условие дополнителности. Значит, x_* является решением задачи (4.4), а поэтому и решением исходной задачи (4.1). ■

Замечание. Если в (4.1) положить $D = 0$, то нетрудно видеть, что условия КТ в этом случае равносильны формулировке второй теоремы двойственности в ЛП.

Итак, решение задач ВКП (4.1) равносильно решению системы КТ (4.6). Если система КТ несовместна, то неразрешима исходная задача, и наоборот.

Многие методы решения задач ВКП основаны на решении системы КТ. Основным недостатком системы КТ является ее нелинейность из-за условия дополнителности — последнего условия в системе (4.6). Ясно, что в тех задачах, где множества N_1 и M_1 пусты, система КТ становится линейной. Это происходит в двух случаях:

1) $\Omega = \mathbb{R}^n$ — в этом случае условия КТ превращаются в равенство $Dx = -c$;

2) $\Omega = \{x | Ax = b\}$, т. е. переменные удовлетворяют только ограничениям-равенствам — тогда условия КТ имеют вид

$$Dx - A^T u = -c,$$

$$Ax = b.$$

В этих двух случаях решение задачи (4.1) сводится к решению линейной системы уравнений.

Так как далее придется часто обращаться к системе (4.6), то представим ее в компактном виде.

Будем считать, что

$$N := \{1, 2, \dots, n\}, \quad M := \{n+1, n+2, \dots, n+m\}.$$

Введем обозначения:

$$P := N \cup M = \{1, 2, \dots, p\}, \quad p = n + m,$$

$$P_1 := N_1 \cup M_1, \quad P_2 := N_2 \cup M_2,$$

$$H[P, P] := \begin{pmatrix} D[N, N] & -A^T[N, M] \\ A[M, N] & 0[M, M] \end{pmatrix}, \quad d[P] := \begin{pmatrix} -c[N] \\ b[M] \end{pmatrix}, \quad (4.7)$$

$$z[P] := (x[N], u[M]), \quad \omega[P] := (y[N], v[M]).$$

В этих обозначениях система КТ запишется в виде

$$H[P, P] \cdot z[P] - E[P, P] \cdot \omega[P] = d[P],$$

$$\omega[P_2] = 0,$$

$$z[P_1] \geq 0, \quad \omega[P_1] \geq 0,$$

$$z[P_1] \cdot \omega[P_1] = 0. \quad (4.8)$$

Здесь $E[P, P]$ — единичная матрица.

Первое равенство в (4.8) представляет линейную систему относительно $z[P]$ и $\omega[P]$. В дальнейшем будем пользоваться номерами

столбцов матрицы этой системы. Для того чтобы различать номера столбцов $H[P, j]$ и $-E[P, j]$ при одном и том же индексе $j \in P$, присвоим столбцам $-E[P, j]$ отрицательные номера $(-j)$. Иными словами, введем индексное множество $P^- = \{-1, -2, \dots, -p\}$, а единичную матрицу $E[P, P]$ запишем с другим индексным множеством для столбцов: $E[P, P^-]$, так что

$$E[i, -k] = \begin{cases} 1, & i = k, \\ 0, & i \neq k. \end{cases}$$

Кроме того, для матрицы рассматриваемой системы и для всех искомым переменных введем единые обозначения. Пусть

$$\begin{aligned} R &= P \cup P^-, \\ T[P, R] &:= (H[P, P], -E[P, P^-]), \\ w[R] &:= (z[P], \omega[P^-]), \end{aligned}$$

т. е. $w[i] = z[i]$, $w[-i] = \omega[i] \quad \forall i \in P$. В этих обозначениях система КТ переписывается в виде

$$T[P, R] \cdot w[R] = d[P], \quad (4.9a)$$

$$w[P_2^-] = \mathbb{O}, \quad (4.9b)$$

$$w[P_1] \geq \mathbb{O}, \quad w[P_1^-] \geq \mathbb{O}, \quad (4.9в)$$

$$w[i] \cdot w[-i] = 0 \quad \forall i \in P_1. \quad (4.9г)$$

В зависимости от удобства будем пользоваться условиями КТ либо в виде (4.6), либо в виде (4.8), либо в виде (4.9).

Для наглядности удобно записывать задачу ВКП в виде таблицы, которую будем называть КТ-таблицей. Способ построения этой таблицы поясним на следующем примере.

Рассмотрим задачу

$$x[1]^2 - x[1] + 2x[2] \rightarrow \min$$

при ограничениях $-x[1] - 3x[2] \geq -2$, $x[1] \geq 0$. Для данной задачи $N = \{1, 2\}$, $M = \{3\}$. Номер 3 получает ограничение и соответствующая ему переменная $w[3]$. В верхней строке КТ-таблицы

1	2	3	-1	-2	-3	$d[P]$
2	0	1	-1	0	0	1
0	0	3	0	-1	0	-2
-1	-3	0	0	0	-1	-2
+	+	*	+	+	0	

указываются номера столбцов матрицы T , ниже записывается сама матрица T , а справа — вектор правых частей системы КТ (см. (4.7)).

Нижняя строка указывает на знаковые ограничения соответствующей переменной. Звездочка означает, что на соответствующую переменную нет знакового ограничения (ее индекс из множества P_2). Ноль означает, что соответствующая переменная должна равняться нулю (ее индекс из множества P_2^-). Знак «плюс» означает, что соответствующая переменная должна быть неотрицательной (ее индекс из $P_1 \cup P_1^-$).

Заметим, что вторая строка в системе КТ несовместна, так что рассматриваемая задача неразрешима.

Упражнения

4.1. Пусть ограничения имеют вид $A[M, N] \cdot x[N] \geq b[M]$. Показать, что условия КТ можно записать в виде

$$\begin{aligned} x_* &\in \Omega, \\ Dx_* + c &= A^T[N, I(x_*)] \cdot u_*[I(x_*)], \\ u_*[I(x_*)] &\geq 0, \end{aligned}$$

где $I(x_*) = \{i \mid A[i, N] \cdot x_*[N] = b[i]\}$.

4.2. Рассмотрим строго выпуклую задачу с ограничениями-равенствами

$$\begin{aligned} \frac{1}{2} \langle Dx, x \rangle + \langle c, x \rangle &\rightarrow \min, \\ A[M, N] \cdot x[N] &= b[M]. \end{aligned}$$

Здесь D положительно определена и $\text{rank } A = |M|$.

Пользуясь условиями КТ, показать, что точка минимума имеет выражение $x^* = D^{-1}A^T(AD^{-1}A^T)^{-1}(b + AD^{-1}c) - D^{-1}c$.

4.3. Определить, для каких значений параметра α разрешима задача ВКП

$$F(x) = 4x[1]^2 + \frac{1}{2}\alpha^2 x[2]^2 + 3x[1] + 2x[2] \rightarrow \min_{x \in \mathbb{R}^2}, \\ x[1] + \alpha x[2] = 1.$$

4.4. Восстановить задачу КП по условиям КТ:

$$\begin{array}{rccccccc} 2x[1] & & & -u[1] & +2u[2] & & = & 1, \\ & x[2] & & +3u[1] & -u[2] & -y[2] & = & -2, \\ & & 2x[3] & +u[1] & +3u[2] & & = & 0, \\ x[1] & -3x[2] & -x[3] & & & & = & 3, \\ -2x[1] & +x[2] & -3x[3] & & & & -v[2] & = -1, \\ \\ x[2] \geq 0, & y[2] \geq 0, & u[2] \geq 0, & v[2] \geq 0, \\ x[2] \cdot y[2] = 0, & u[2] \cdot v[2] = 0. \end{array}$$

§5. Дополнительный и экстремальный базисы системы Куна—Таккера

Основная трудность при решении системы КТ заключена в последнем нелинейном условии дополнительности (4.9г). Вулф [11, 22] предложил искать решение системы КТ (для задачи с ограничениями $Ax = b, x \geq 0$) только среди *дополнительных решений* линейной системы (4.9а), т. е. у которых хотя бы одна из компонент в произведении (4.9г) равна нулю. Используя эту идею, введем понятие дополнительного базиса.

Напомним, что матрица $T[P, R]$ имеет полный ранг, равный p , и столбцов у нее в два раза больше, чем строк.

Базисом системы (4.9а) называются p линейно независимых столбцов матрицы $T[P, R]$. Как обычно, базис фиксируется с помощью индексного множества $Q \subseteq R$, состоящего из номеров базисных столбцов. Это индексное множество Q также будем называть базисом. Так как $R = P \cup P^-$, то базис Q системы (4.9а) имеет следующую структуру:

$$Q = I \cup (-J), \quad I \subseteq P, \quad (-J) \subseteq P^-, \quad \text{т. е. } J \subseteq P.$$

Базис Q , удовлетворяющий условию $I \cap J = \emptyset$, будем называть *дополнительным*.

Иными словами, дополнительный базис — базис, который не содержит одновременно индексов, отличающихся лишь знаком. Примером дополнительного базиса является базис $Q = P^-$ (здесь $I = \emptyset, J = P$).

Базисным решением системы называется вектор $w[R]$, у которого базисная часть $w[Q]$ определяется из системы

$$T[P, Q] \cdot w[Q] = d[P],$$

а небазисная часть равна нулю, т. е. $w[R \setminus Q] = 0$.

Важно отметить тот факт, что *базисное решение на дополнительном базисе автоматически удовлетворяет условию* (4.9г), так как небазисные компоненты равны нулю.

Итак, если ограничиться рассмотрением только *дополнительных базисов* системы (4.9а), то нелинейное условие КТ (4.9г) можно не учитывать, так как оно выполняется автоматически.

Дополнительный базис, базисное решение на котором удовлетворяет еще и условиям (4.9б), (4.9в), будем называть *экстремальным*. Разумеется, базисное решение на экстремальном базисе содержит решение исходной задачи ($x[N] = w[N]$).

Покажем (см. [6]), что у разрешимой системы КТ экстремальный базис существует. А тогда для поиска решения системы КТ можно ограничиться рассмотрением только *дополнительных базисов* системы (4.9а).

Прежде чем доказывать существование экстремального базиса, рассмотрим вопрос о построении дополнительного базиса. Так как всякий базис состоит из p столбцов, то в дополнительном базисе Q представлены все индексы из P , но некоторые из них могут иметь знак «минус». Иначе говоря, для дополнительного базиса $Q = I \cup (-J)$ имеет место равенство $I \cup J = P$. Введем еще одно понятие.

Неполным базисом системы (4.9а) назовем линейно независимые столбцы матрицы $T[P, R]$, число которых меньше $p = |P|$. Неполный базис $\Gamma = I \cup (-J)$ назовем *дополнительным*, если $I \cap J = \emptyset$. В частности, один ненулевой вектор $T[P, j]$ образует неполный дополнительный базис.

Покажем, что к любому неполному дополнительному базису Γ всегда можно добавить не вошедшие в него индексы из P (с плюсом или с минусом) и получить полный дополнительный базис.

Заметим сначала, что несимметричная матрица $H[P, P]$ (см. (4.7)) обладает свойством

$$\begin{cases} \text{для любого } z & \langle Hz, z \rangle \geq 0, \\ \text{если } \langle Hz, z \rangle = 0, & \text{то } (H + H^T)z = \mathbb{O}, \end{cases} \quad (5.1)$$

которое следует из неотрицательной определенности симметричной матрицы

$$H + H^T = \begin{pmatrix} D & -A^T \\ A & \mathbb{O} \end{pmatrix} + \begin{pmatrix} D & A^T \\ -A & \mathbb{O} \end{pmatrix} = \begin{pmatrix} 2D & \mathbb{O} \\ \mathbb{O} & \mathbb{O} \end{pmatrix}.$$

Лемма 5.1. *Неполный дополнительный базис Γ системы (4.9а) можно расширить до полного дополнительного базиса.*

Доказательство. Если $\Gamma = I \cup (-J)$ — неполный дополнительный базис системы (4.9а), то найдется индекс $q \in P$, такой, что $q \notin I$, $q \notin J$. Покажем, что хотя бы один из столбцов $H[P, q]$ или $-E[P, -q]$ линейно независим с базисными столбцами $T[P, \Gamma]$.

Допустим противное, т. е. каждый из этих столбцов является линейной комбинацией базисных:

$$\begin{aligned} H[P, q] &= H[P, I] \cdot \xi_1[I] - E[P, -J] \cdot \eta_1[J], \\ -E[P, -q] &= H[P, I] \cdot \xi_2[I] - E[P, -J] \cdot \eta_2[J]. \end{aligned} \quad (5.2)$$

Доопределим векторы $\xi_1, \xi_2, \eta_1, \eta_2$ на все множество P естественным образом:

$$\xi_1[q] := -1, \quad \eta_2[q] := -1, \quad \xi_2[q] := 0, \quad \eta_1[q] := 0, \quad (5.3)$$

а остальные, еще не определенные компоненты этих векторов, полагаем равными нулю. Тогда оба равенства (5.2) можно переписать в виде

$$H\xi_1 = \eta_1, \quad H\xi_2 = \eta_2. \quad (5.4)$$

В силу дополнительности базиса Γ множества I и J не пересекаются, поэтому с учетом (5.3) имеем

$$\langle \eta_1, \xi_1 \rangle = 0, \quad \langle \eta_1, \xi_2 \rangle = 0, \quad \langle \eta_2, \xi_1 \rangle = 1. \quad (5.5)$$

Из первых равенств (5.4) и (5.5) получаем $\langle H\xi_1, \xi_1 \rangle = 0$, и тогда согласно (5.1) $H\xi_1 = -H^T\xi_1$. Значит,

$$\begin{aligned} 1 &= \langle \eta_2, \xi_1 \rangle = \langle H\xi_2, \xi_1 \rangle = \\ &= \langle \xi_2, H^T\xi_1 \rangle = -\langle \xi_2, H\xi_1 \rangle = -\langle \xi_2, \eta_1 \rangle = 0. \end{aligned}$$

Полученное противоречие показывает, что оба равенства (5.2) одновременно выполняться не могут.

Таким образом, один из векторов $H[P, q]$ или $-E[P, q]$ можно добавить к базису Γ . Новый базис опять дополнительный, и если он не полный, процедуру можно повторить. За конечное число повторений получим полный дополнительный базис. ■

Теорема 5.1. *Экстремальный базис системы КТ существует тогда и только тогда, когда задача (4.1) разрешима.*

Доказательство. Необходимость очевидна.

Достаточность. Пусть задача (4.1) разрешима, тогда разрешима и система КТ. Покажем, что среди ее решений найдется базисное решение с дополнительным базисом.

Пусть $w_0[R]$ — любое решение системы КТ, так что для него выполнены равенства (4.9):

$$T[P, R] \cdot w_0[R] = d[P], \quad (5.6)$$

$$w_0[P_2^-] = \mathbb{O}, \quad (5.7)$$

$$w_0[P_1] \geq \mathbb{O}, \quad w_0[P_1^-] \geq \mathbb{O}, \quad (5.8)$$

$$w_0[i] \cdot w_0[-i] = 0 \quad \forall i \in P_1. \quad (5.9)$$

Если $w_0[P] = \mathbb{O}$, то $w_0[P^-]$ можно воспринимать как базисное решение на дополнительном базисе $Q_* = P^-$, который и является в этом случае экстремальным базисом.

Пусть $w_0[P] \neq \mathbb{O}$. Введем в рассмотрение индексные множества

$$I_0 = \{i \in P \mid w_0[i] \neq 0\}, \quad J_0 = \{i \in P \mid w_0[-i] \neq 0\},$$

$$L_0 = I_0 \cup (-J_0).$$

В силу равенств (5.9) и (5.7) имеем

$$I_0 \cap J_0 = \emptyset. \quad (5.10)$$

Возможны три ситуации:

а) L_0 — полный базис системы (5.6). В силу (5.10) этот базис дополнительный, и так как его базисное решение w_0 удовлетворяет системе КТ, то L_0 — экстремальный базис.

б) L_0 — неполный дополнительный базис системы (5.6). По лемме 5.1 его можно расширить до полного Q_* . Очевидно, что вектор $w_0[R]$ является базисным решением и на базисе Q_* , поэтому базис Q_* экстремальный.

в) L_0 не является базисом системы (5.6), но обладает условием дополнителности (5.10). Построим из L_0 экстремальный базис следующим образом. Так как L_0 не базис, то столбцы матрицы $T[P, L_0]$ линейно зависимы, т. е.

$$T[P, L_0] \cdot s[L_0] = \mathbb{O}, \quad s[L_0] \neq \mathbb{O}.$$

Полагая $s[R \setminus L_0] = \mathbb{O}$, получаем

$$T[P, R] \cdot s[R] = \mathbb{O}. \quad (5.11)$$

С помощью вектора s построим другое решение системы КТ

$$w_1[R] = w_0[R] - \lambda \cdot s[R], \quad (5.12)$$

у которого за счет выбора параметра λ нулевых компонент можно сделать больше, чем у $w_0[R]$. Тогда индексное множество L_1 , построенное аналогично L_0 , но для w_1 , будет обладать свойством

$$L_1 \subset L_0, \quad |L_1| < |L_0|. \quad (5.13)$$

Займемся выбором параметра λ . Заметим сначала, что в силу (5.11) при любом λ для w_1 выполнено равенство (5.6). Кроме того, $s[P_2^-] = \mathbb{O}$, поэтому при любом λ для w_1 выполнено условие (5.7). Значит, λ следует выбрать таким, чтобы выполнялись условия (5.8) и (5.13). Различаются два случая.

1) $s[R_1 \cap L_0] = \mathbb{O}$ — в этом случае условия (5.8) выполняются при любом λ . Найдем нужное λ .

Так как $s[R] \neq \mathbb{O}$, то существует индекс $q \in L_0 \cap R_2$, для которого $s[q] \neq 0$. Положим $\lambda = \frac{w_0[q]}{s[q]}$, тогда из (5.12) имеем $w_1[q] = w_0[q] - \frac{w_0[q]}{s[q]} s[q] = 0$, хотя $w_0[q] \neq 0$, т. е. для w_1 выполнено (5.13).

2) $s[R_1 \cap L_0] \neq \mathbb{O}$ — не умаляя общности можно считать, что у вектора $s[R_1 \cap L_0]$ есть положительные компоненты (иначе у вектора $s[R]$ можно поменять знак). Пусть

$$\lambda = \min \left\{ \frac{w_0[j]}{s[j]} \mid j \in R_1 \cap L_0, s[j] > 0 \right\} = \frac{w_0[q]}{s[q]}.$$

Очевидно, что $\lambda > 0$, так как $w_0[q] > 0$. При таком выборе λ получаем из (5.12)

$$w_1[q] = 0, \quad w_1[R_1 \cap L_0] \geq \mathbb{O}.$$

Значит, для w_1 выполнены условия (5.8) и (5.13).

Итак, получено новое решение системы КТ w_1 , для которого верно (5.13). Если L_1 — полный или неполный базис, то доказательство завершается как в п. а) или б). Если же L_1 еще не базис, то повторяем процедуру в). За конечное число таких повторений базис будет построен. ■

Следствие. Любое решение ВКП можно преобразовать в базисное с дополнительным базисом.

§6. Обобщение симплекс-метода

Многие методы решения экстремальных задач

$$F(x) \rightarrow \inf_{x \in \Omega}$$

строятся по следующей схеме. Выбирается точка x_k из Ω , в ней находится вектор s_k , в направлении которого целевая функция убывает. Если такого вектора нет, то вычисления заканчиваются. В противном случае в направлении s_k вычисляется шаг $t_k \geq 0$, такой, чтобы точка $x_{k+1} = x_k + t_k \cdot s_k$ не выходила из Ω и $F(x_{k+1}) \leq F(x_k)$. Если $F(x)$ — линейная функция, то последнее требование выполнится автоматически.

По этой же схеме работает и симплекс-метод для задачи ЛП. Заметим, что симплекс-метод хорошо иллюстрирует основные приемы, используемые в базисных методах квадратичного программирования.

Рассмотрим задачу линейного программирования, в которой первая каноническая форма нарушена, т. е. не все переменные неотрицательны:

$$F(x) := \langle c, x \rangle \rightarrow \min_{x \in \Omega}, \quad (6.1)$$

$$\Omega = \{x[N] \mid A[M, N] \cdot x[N] = b[M], \quad x[N_1] \geq \mathbb{O}\},$$

где $N_1 \subset N$. Если $N_1 = N$, то задача (6.1) принимает первую каноническую форму.³

Разумеется, путем введения новых переменных эту задачу можно свести к задаче в первой канонической форме и к ней применить стандартный симплекс-метод. Покажем, что можно обойтись и без этого сведения.

Согласно идеи Булавского (см. §7) можно считать, что

$$\text{rank } A = |M|. \quad (6.2)$$

Напомним признак оптимальности плана задачи (6.1).

Положим $N_2 = N \setminus N_1$ и запишем двойственную задачу

$$\begin{aligned} \langle b, u \rangle &\rightarrow \max, \\ A^T[N_1, M] \cdot u[M] &\leq c[N_1], \\ A^T[N_2, M] \cdot u[M] &= c[N_2]. \end{aligned}$$

Введем в рассмотрение *вектор оценок*

$$y[N] := c[N] - A^T[N, M] \cdot u[M].$$

По первой теореме двойственности *план задачи (6.1) $x[N]$ оптимален, если существует вектор $u[M]$, для которого*

$$y[N_1] \geq \mathbb{O}, \quad y[N_2] = \mathbb{O}, \quad (6.3)$$

$$\langle c, x \rangle = \langle b, u \rangle. \quad (6.4)$$

Приведем описание *обобщенного симплекс-метода* с одновременным его обоснованием.

Допустим, что уже проделано k итераций ($k = 0, 1, \dots$) и построена точка $x_k \in \Omega$ — базисный план задачи (6.1), соответствующий столбцовому базису Γ_k матрицы A .⁴ Значит, матрица $A[M, \Gamma_k]$ квадратная невырожденная и

³Задача (6.1) при $|N_1| < |N|$ возникает, в частности, в методах линейной и нелинейной аппроксимации.

⁴Столбцовым базисом матрицы $A[M, N]$ принято называть $|M|$ линейно независимых столбцов этой матрицы или индексное множество, состоящее из номеров этих столбцов. В силу (6.2) столбцовый базис матрицы A существует.

$$A[M, \Gamma_k] \cdot x_k[\Gamma_k] = b[M], \quad x_k[N_1] \geq \mathbb{O}, \quad x_k[N \setminus \Gamma_k] = \mathbb{O}.$$

Описание метода будем вести одновременно для двух численных реализаций: когда известна обратная базисная матрица $B_k[\Gamma_k, M] = (A[M, \Gamma_k])^{-1}$ и когда она не известна. Вариант метода с первой численной реализацией называется *методом обратных матриц*, со второй — *прямым методом*.

Один k -й шаг обобщенного симплекс-метода состоит из четырех операций.

1. *Проверка базиса Γ_k на оптимальность.* Пользуясь базисом Γ_k , найдем вектор двойственных переменных $u_k[M]$. В прямом методе он находится как решение системы линейных уравнений

$$A^T[\Gamma_k, M] \cdot u[M] = c[\Gamma_k], \quad (6.5)$$

а в методе обратных матриц — по формуле

$$u_k[M] = c[\Gamma_k] \cdot B_k[\Gamma_k, M].$$

Заметим, что в силу (6.5)

$$\langle x_k, c \rangle = x_k[\Gamma_k] \cdot A^T[\Gamma_k, M] \cdot u_k[M] = \langle b, u_k \rangle,$$

т. е. для плана x_k и вектора u_k выполнены равенства (6.4). Поэтому для оптимальности плана x_k следует лишь проверить, является ли u_k планом двойственной задачи, т. е. проверить условия (6.3). В силу (6.5) $y_k[\Gamma_k] = \mathbb{O}$, поэтому условия оптимальности принимают вид

$$y_k[N_1 \setminus \Gamma_k] \geq \mathbb{O}, \quad y_k[N_2 \setminus \Gamma_k] = \mathbb{O}. \quad (6.6)$$

Если условия (6.6) выполнены, то $x_k[N]$ является решением задачи (6.1), а $u_k[M]$ — решением двойственной задачи. На этом вычисления заканчиваются.

Если условия (6.6) нарушены, переходим к п. 2.

2. *Выбор направления спуска в точке x_k .* Выделим множество тех небазисных индексов, на которых соответствующее условие оптимальности из (6.6) не выполнено. Положим

$$\begin{aligned} L_k^- &= \{j \in N_1 \mid y_k[j] < 0\}, \\ L_k^0 &= \{j \in N_2 \mid y_k[j] \neq 0\}, \\ L_k &= L_k^- \cup L_k^0. \end{aligned}$$

Ясно, что $L_k \neq \emptyset$. Возьмем любой индекс $l_k \in L_k$. Часто выбирают индекс с максимальной по модулю оценкой, но могут быть и другие соображения. Вычислим $z_k[\Gamma_k]$ — коэффициенты разложения столбца $A[M, l_k]$ по базису Γ_k . Вектор z_k в прямом методе находится как решение системы с базисной матрицей

$$A[M, \Gamma_k] \cdot z[\Gamma_k] = A[M, l_k], \quad (6.7)$$

а в методе обратных матриц — по формуле

$$z_k[\Gamma_k] = B[\Gamma_k, M] \cdot A[M, l_k].$$

Найденный вектор $z_k[\Gamma_k]$ расширим на все индексное множество N , полагая $z_k[l_k] = -1$, а остальные компоненты — равными нулю. Тогда из (6.7) вытекает, что

$$A[M, N] \cdot z_k[N] = \mathbb{O}. \quad (6.8)$$

Покажем, что направление

$$s_k[N] = \text{sgn}(y_k[l_k]) \cdot z_k[N]$$

является направлением строгого убывания функции $F(x)$.

Действительно, $\langle z_k, c \rangle = -c[l_k] + z_k[\Gamma_k] \cdot c[\Gamma_k]$. Воспользуемся равенствами (6.5) и (6.7). Имеем

$$\begin{aligned} \langle z_k, c \rangle &= -c[l_k] + z_k[\Gamma_k] \cdot A^T[\Gamma_k, M] \cdot u_k[M] = \\ &= -c[l_k] + A^T[l_k, M] \cdot u_k[M] = -y_k[l_k]. \end{aligned}$$

Значит, $F(x_k + t s_k) = F(x_k) - t |y_k[l_k]| < F(x_k) \quad \forall t > 0$.

3. *Определение шага в направлении спуска.* Новую точку находим на луче $x(t) = x_k + t \cdot s_k$, $t \geq 0$. Пусть t_k — максимальное значение t , при котором точка луча $x(t)$ принадлежит Ω . В силу (6.8) все точки луча удовлетворяют ограничениям-равенствам из Ω , поэтому принадлежность точки $x(t)$ множеству Ω определяется лишь неравенством $x(t)[N_1] \geq \mathbb{O}$.

Так как $x_k[N_1] \geq \mathbb{O}$, то при выборе шага t_k нас интересуют только те индексы $j \in N_1$, на которых $s_k[j] < 0$. Значит, шаг определяется по формуле

$$t_k = \min \left\{ \frac{x_k[j]}{|s_k[j]|} \mid j \in \Gamma_k \cap N_1, \quad s_k[j] < 0 \right\}. \quad (6.9)$$

Если оказалось, что $s_k[\Gamma_k \cap N_1] \geq \mathbb{O}$, то весь луч $x(t)$, $t \geq 0$, принадлежит множеству Ω и $F(x(t)) \rightarrow -\infty$ при $t \rightarrow \infty$, т. е. исходная задача неразрешима.

В противном случае строим новую точку $x_{k+1} = x_k + t_k \cdot s_k$.

4. *Формирование нового базиса.* Обозначим через P_k множество индексов, на которых реализуется минимум в (6.9). Выберем любой индекс p_k из P_k и положим $\Gamma_{k+1} = \Gamma_k \setminus \{p_k\} \cup \{l_k\}$. Очевидно, что Γ_{k+1} является столбцовым базисом матрицы A , а x_{k+1} — базисным планом. При этом

$$F(x_{k+1}) < F(x_k), \quad \text{если } t_k > 0,$$

$$F(x_{k+1}) = F(x_k), \quad \text{если } t_k = 0.$$

Если реализуется метод обратных матриц, то следует вычислить B_{k+1} через B_k по формулам пересчета (8.4) (см. §8). На этом кончается описание одного шага метода.

§7. Способ построения начального базисного плана

Рассмотрим задачу выпуклого квадратичного программирования

$$F(x) = \frac{1}{2} \langle Dx, x \rangle + \langle c, x \rangle \rightarrow \min_{x \in \Omega}, \quad (7.1)$$

где

$$\Omega = \{x[N] \mid A[M, N] \cdot x[N] = b[M], \quad x[N_1] \geq \mathbb{O}, \quad N_1 \subset N\} \quad (7.2)$$

и матрица $A[M, N]$ произвольна. Некоторые базисные методы ВКП требуют предварительного поиска базисного плана (метод Данцига,

метод Вульфа). Это же относится и к обобщенному симплекс-методу, описанному в предыдущем параграфе.

Покажем, как построить базисный план или убедиться в пустоте множества Ω . Заметим, что если строки матрицы A линейно зависимы, то матрицу A будем расширять, добавляя столбцы-орты. Предлагаемый способ принадлежит В. А. Булавскому [2] и годится для произвольной матрицы A , при этом числа $m = |M|$ и $n = |N|$ могут находиться в любом соотношении.

Напомним, что $N = \{1, 2, \dots, n\}$, $M = \{n+1, n+2, \dots, n+m\}$. Введем новые переменные $x[M]$ и диагональную матрицу $\mathcal{E}[M, M]$ с компонентами

$$\mathcal{E}[i, i] = \begin{cases} \text{sgn } b[i], & \text{если } b[i] \neq 0, \\ 1, & \text{если } b[i] = 0. \end{cases}$$

Рассмотрим вспомогательную задачу ЛП

$$\begin{aligned} \mu(x) &:= \sum_{i \in M} x[i] \rightarrow \min, \\ A[M, N] \cdot x[N] + \mathcal{E}[M, M] \cdot x[M] &= b[M], \\ x[N_1] &\geq 0, \quad x[M] \geq 0. \end{aligned} \quad (7.3)$$

Очевидно, что множество планов этой задачи не пусто и целевая функция ограничена снизу нулем, поэтому задача (7.3) разрешима. Начальный базисный план задачи очевиден:

$$x_0[N] = 0, \quad x_0[i] = |b[i]| \quad \forall i \in M,$$

ему соответствует базис $\Gamma_0 = M$ с базисной матрицей \mathcal{E} . Таким образом, имеется вся нужная информация, чтобы решить задачу (7.3) обобщенным симплекс-методом.

Пусть $x_q[N \cup M]$ — оптимальный базисный план задачи (7.3) с базисом $\Gamma_q = N_q \cup M_q$, где $N_q \subset N$, $M_q \subset M$. Возможны три ситуации:

а) $\mu(x_q) > 0$. Это означает, что $\Omega = \emptyset$, поэтому исходная задача (7.1), (7.2) неразрешима.

б) $\mu(x_q) = 0$ и $M_q = \emptyset$. Это означает, что $x_q[M] = 0$, вектор $x_q[N]$ принадлежит Ω и является требуемым базисным планом (7.2) с базисом $N_q \subseteq N$.

в) $\mu(x_q) = 0$ и $M_q \neq \emptyset$. Это означает, что

$$x_q[M] = 0, \quad x_q[N] \in \Omega,$$

но базисом является $\Gamma_q = N_q \cup M_q$. В этом случае искусственные базисные столбцы $\mathcal{E}[M, j]$, $j \in M_q$, сохраним в базисе, но изменим их знаки. Введем множество

$$\Omega^* = \left\{ x[N \cup M_q] \mid \begin{aligned} &A[M, N] \cdot x[N] - \mathcal{E}[M, M_q] \cdot x[M_q] = b[M], \\ &x[N_1] \geq 0, \quad x[M_q] \geq 0 \end{aligned} \right\}. \quad (7.4)$$

Очевидно, что вектор $x_q[N \cup M_q]$ является базисным планом системы из (7.4) с базисом $\Gamma_q = N_q \cup M_q$.

Для решения исходной задачи (7.1) заменим в ней множество Ω множеством Ω^* , но для этого следует убедиться в следующем.

Теорема 7.1. Для любого вектора из Ω^* искусственная составляющая $x[M_q]$ равна нулю.

Доказательство. Запишем задачу, двойственную задаче (7.3), положив $N_2 = N \setminus N_1$:

$$\begin{aligned} \langle b, u \rangle &\rightarrow \max, \\ u[M] \cdot A[M, N_1] &\leq 0[N_1], \\ u[M] \cdot A[M, N_2] &= 0[N_2], \\ u[M] \cdot \mathcal{E}[M, M] &\leq 0[M]. \end{aligned} \quad (7.5)$$

Воспользуемся оптимальным базисом задачи (7.3) Γ_q и обратной базисной матрицей $B_q[\Gamma_q, M] = (A[M, N_q], \mathcal{E}[M, M_q])^{-1}$. Заметим, что первые n коэффициентов целевой функции задачи (7.3) равны нулю, а все следующие — единице, поэтому оптимальным планом задачи (7.5) является вектор

$$u_q[M] = \sum_{j \in M_q} B_q[j, M].$$

Возьмем произвольный вектор $(x[N \cup M_q]) \in \Omega^*$. Умножим обе части равенства из (7.4) слева на $u_q[M]$. Получим

$$u_q[M] \cdot A[M, N] \cdot x[N] - u_q[M] \cdot \mathcal{E}[M, M_q] \cdot x[M_q] = \langle u_q, b \rangle. \quad (7.6)$$

Из ограничений (7.5) следует, что $u_q[M] \cdot A[M, N] \cdot x[N] \leq 0$. Так как все столбцы матрицы $\mathcal{E}[M, M_q]$ базисные, то

$$u_q[M] \cdot \mathcal{E}[M, M_q] = \mathbb{I}[M_q].$$

Кроме того, по первой теореме двойственности $\langle b, u_q \rangle = \mu(x_q) = 0$. Пользуясь этими равенствами, из (7.6) получаем

$$\sum_{i \in M_q} x[i] \leq 0,$$

но $x[M_q] \geq 0$, значит $x[M_q] = 0$. ■

Согласно доказанной теореме, экстремальную задачу вида

$$F(x) \rightarrow \min_{x \in \Omega \subset \mathbb{R}^n}$$

с непустым множеством планов можно заменить расширенной задачей

$$F(x) \rightarrow \min_{x \in \Omega^* \subset \mathbb{R}^{n+|M_q|}}.$$

Разумеется, матрица D и вектор c в расширенной задаче дополняются нулевыми компонентами.

Если $F(x) = \langle c, x \rangle$, то для дальнейшего решения расширенной задачи можно продолжить симплекс-метод, заменив целевую функцию на исходную $\langle c, x \rangle$ и изменив знак подматрицы $B_q[M_q, M]$.

§8. Пересчет обратной базисной матрицы и базисного решения

Напомним, как пересчитывается обратная базисная матрица, если новый базис отличается от предыдущего лишь одним индексом. Будем рассматривать базисные матрицы, выделенные из прямоугольной матрицы $C[M, N]$. Пусть $\text{rank } C[M, N] = m$ и индексное множество $Q = \{j_1, j_2, \dots, j_m\} \subset N$ является ее базисом. Положим $B[Q, M] = C[M, Q]^{-1}$.

Теорема 8.1. Пусть индексное множество $Q' \subset N$ отличается от Q лишь одним индексом:

$$Q' = \{j_1, j_2, \dots, j_{p-1}, l, j_{p+1}, \dots, j_m\}.$$

Обозначим через $s[Q]$ вектор коэффициентов разложения столбца $C[M, l]$ по столбцам базисной матрицы $C[M, Q]$.

Множество Q' является базисом тогда и только тогда, когда отличен от нуля коэффициент $s[j_p]$. И если $s[j_p] \neq 0$, то соответствующая базису Q' обратная базисная матрица имеет представление

$$B'[Q', M] = \mathcal{M}[Q', Q] \cdot B[Q, M], \quad (8.1)$$

где матрица \mathcal{M} отличается от единичной матрицы лишь p -м столбцом:

$$\mathcal{M} = \begin{pmatrix} 1 & -\frac{s[j_1]}{s[j_p]} & & & \\ & \dots & \dots & & \\ & & \frac{1}{s[j_p]} & & \\ & & \dots & \dots & \\ & & -\frac{s[j_m]}{s[j_p]} & & 1 \end{pmatrix}.$$

Доказательство. Запишем разложение всех столбцов матрицы $C[M, Q']$ по столбцам исходной базисной матрицы

$$C[M, Q'] = C[M, Q] \cdot S[Q, Q'] \quad (8.2)$$

и заметим, что матрица S отличается от единичной лишь p -м столбцом, на месте которого стоит вектор $s[Q]$:

$$S[Q, Q'] = \begin{pmatrix} 1 & s[j_1] & & & \\ & \dots & \dots & & \\ & & s[j_p] & & \\ & & \dots & \dots & \\ & & s[j_m] & & 1 \end{pmatrix}.$$

Очевидно, что определитель матрицы S равен $s[j_p]$, поэтому матрица $C[M, Q']$ неособенная тогда и только тогда, когда

$$s[j_p] \neq 0.$$

Далее, из (8.2) следует, что $B' = S^{-1} \cdot B$. Нетрудно видеть, что $S^{-1} = \mathcal{M}$, что и требовалось доказать. ■

Глава 2

БАЗИСНЫЕ МЕТОДЫ

Равенство (8.1) равносильно следующим пересчетным формулам:

$$\begin{aligned} B'[l, j] &= B[j_p, j]/s[j_p], \\ B'[i, j] &= B[i, j] - B'[l, j] \cdot s[i], \\ & i \in Q' \setminus \{l\}, \quad j \in M. \end{aligned} \quad (8.3)$$

Обратимся к пересчету базисного решения системы $Cw = d$. Так как

$$w'[Q'] = B'[Q', P] \cdot d[P],$$

то из (8.1) следует $w'[Q'] = M[Q', Q] \cdot w[Q]$. Последнее равенство равносильно следующим пересчетным формулам:

$$\begin{aligned} w'[l] &= w[j_p]/s[j_p], \\ w'[i] &= w[i] - w'[l] \cdot s[i], \quad i \in Q' \setminus \{l\}. \end{aligned} \quad (8.4)$$

К базисным методам относятся методы выпуклого квадратичного программирования (ВКП), сводящиеся к поиску экстремального базиса среди дополнительных базисов системы КТ. Пути поиска экстремального базиса среди дополнительных могут быть разными, что порождает многообразие базисных методов. В данной главе приведены описания и обоснования основных базисных методов — метода дополнительного базиса (ДБ) [6, 8] и метода Данцига [4, 7]. Будет показано, что известные методы Вульфа [5, 22] и Лемке (см. [1]) являются частными случаями метода ДБ.

Перейдем к рассмотрению численных методов.

§1. Метод полного перебора дополнительных базисов

Как было сказано выше, решение задачи ВКП можно свести к поиску экстремального базиса системы КТ. Этот поиск можно осуществить путем перебора всех дополнительных базисов системы КТ и проверки каждого на экстремальность. Если после полного перебора выяснилось, что экстремального базиса нет, то согласно теореме 5.1 гл. 1 исходная задача ВКП не имеет решения. Если же найден экстремальный базис, то часть найденного базисного вектора, а именно $w[N]$, является решением исходной задачи ВКП.

Разумеется, полный перебор базисов практически можно использовать лишь для задач малого размера.

Проиллюстрируем на примере работу описанного метода. Решим задачу

$$\frac{1}{2}x[1]^2 + \frac{1}{2}x[2]^2 + x[1] - 2x[2] \rightarrow \min$$

при ограничениях

$$x[1] + x[2] = 1, \quad x[1] \geq 0, \quad x[2] \geq 0.$$

Запишем КТ-таблицу (см. §4 гл.1), заметив, что в данном случае $N = \{1, 2\}$, $M = \{3\}$:

$w[1]$	$w[2]$	$w[3]$	$w[-1]$	$w[-2]$	$w[-3]$	$d[P]$
1	2	3	-1	-2	-3	-1
1	0	-1	-1	0	0	-1
0	1	-1	0	-1	0	2
1	1	0	0	0	-1	1
+	+	*	+	+	0	

Очевидный дополнительный базис $Q_0 = \{-1, -2, -3\}$ не подходит, так как $w_0[Q_0] = (1, -2, -1)$. Заменяя индекс (-2) на 2, получаем базис $Q_1 = (-1, 2, -3)$ с базисным решением $w_1[Q_1] = (1, 2, 1)$. Этот базис не экстремальный, так как $w_1[-3]$ отлична от нуля. Заменяем индекс (-3) на 3. Базис $Q_2 = \{-1, 2, 3\}$ является экстремальным, так как базисное решение $w_2[Q_2] = (2, 1, -1)$ удовлетворяет всем условиям КТ. Очевидно, что решением задачи будет вектор $x_* = (0, 1)$.

§2. Решение линейной системы со знаковыми ограничениями на переменные

Прежде чем переходить к описанию базисных методов, остановимся на способе решения системы

$$\begin{cases} A[M, N] \cdot x[N] = b[M], \\ x[N_1] \geq 0, \quad N_1 \subset N, \\ \text{rank } A = |M|. \end{cases} \quad (2.1)$$

Предполагается, что известен базис Γ матрицы $A[M, N]$, у которого соответствующий базисный план $x_0[\Gamma] = (A[M, \Gamma])^{-1} \cdot b[M]$, $x_0[N \setminus \Gamma] = 0$ не удовлетворяет знаковому ограничению, т. е.

$$\min_{j \in \Gamma \cap N_1} x_0[j] = x_0[p] < 0. \quad (2.2)$$

Расширим матрицу $A[M, N]$, добавив к ней один искусственный столбец

$$A[M, 0] = A[M, \Gamma] \cdot h[\Gamma], \quad (2.3)$$

где $h[N]$ определяется по формуле

$$h[i] = \begin{cases} -1, & \text{если } x_0[i] < 0 \text{ и } i \in \Gamma \cap N_1, \\ 0 & \text{для остальных компонент.} \end{cases}$$

Введем новую переменную $x[0]$ и поставим вспомогательную задачу

$$\begin{cases} x[0] \rightarrow \min, \\ A[M, N] \cdot x[N] + A[M, 0] \cdot x[0] = b[M], \\ x[N_1] \geq 0, \quad x[0] \geq 0. \end{cases} \quad (2.4)$$

Искусственный столбец подобран так, чтобы $\Gamma_1 := \Gamma \setminus \{p\} \cup \{0\}$ был допустимым базисом для задачи (2.4)⁵. Действительно, из (2.3) следует, что $h[\Gamma]$ — коэффициенты разложения $A[M, 0]$ по базису Γ . Так как $h[p]$ отличен от нуля, то Γ_1 есть базис. Далее, по известным формулам пересчета плана (гл. 1, (8.4)) имеем

$$x_1[\Gamma] = x_0[\Gamma] - h[\Gamma] \cdot t_0, \quad \text{где } t_0 = -x_0[p].$$

Учитывая (2.2), получаем

$$\left. \begin{cases} x_1[0] = -x_0[p] > 0, \\ x_1[i] = x_0[i] + h[i]x_0[p] \geq 0 \quad \forall i \in \Gamma \cap N_1, \end{cases} \right\} \quad (2.5)$$

т. е. $x_1[N_1] \geq 0$.

Таким образом, вспомогательная задача разрешима, так как множество планов не пусто и целевая функция ограничена снизу на множестве планов. Для решения этой задачи воспользуемся симплекс-методом (см. §6 гл. 1). Пусть Γ_* — оптимальный базис вспомогательной задачи и $\{x_*[N], x_*[0]\}$ — ее решение. Очевидно, что если $x_*[0] > 0$, то система (2.1) несовместна. Если $x_*[0] = 0$, то $x_*[N]$ является базисным решением системы (2.1).

Может оказаться, что $x_*[0] = 0$, а Γ_* содержит индекс 0, т. е. искусственный столбец остался в оптимальном базисе. Так как ранг матрицы A равен $|M|$, то существует небазисный столбец матрицы A , которым можно заменить искусственный столбец. Иными словами, если потребуется, то для данного базисного плана $x_*[N]$ можно найти другой базис $\Gamma_{**} \subset N$.

Эта идея добавки лишь одного искусственного столбца при известном недопустимом базисе широко используется в методах ВКП.

⁵Базис допустимый, если базисное решение удовлетворяет условиям $x[N_1] \geq 0$, $x[0] \geq 0$.

§3. Метод дополнительного базиса

3.1. Общая схема метода

Для задач ВКП с общими ограничениями разработан метод поиска экстремального базиса, который назван методом дополнительного базиса (метод ДБ) [6, 8]. Приведем общую схему метода.

Рассмотрим задачу

$$F(x) := \frac{1}{2} \langle x, Dx \rangle + \langle c, x \rangle \rightarrow \min_{x \in \Omega} \quad (3.1)$$

где D — симметричная неотрицательно определенная матрица и Ω — множество векторов $x[N]$, которые удовлетворяют ограничениям:

$$\begin{aligned} A[M_1, N] \cdot x[N] &\geq b[M_1], \\ A[M_2, N] \cdot x[N] &= b[M_2], \\ x[N_1] &\geq 0, \quad N_1 \subset N. \end{aligned} \quad (3.2)$$

Метод ДБ основан на решении системы КТ, которую запишем в виде (см. гл. 1, формулы (4.8) и (4.9))

$$T[P, R] \cdot w[R] = H[P, P] \cdot w[P] - E[P, P^-] \cdot w[P^-] = d[P], \quad (3.3)$$

$$w[P_2^-] = 0, \quad (3.4)$$

$$w[P_1] \geq 0, \quad w[P_1^-] \geq 0, \quad (3.5)$$

$$w[i] \cdot w[-i] = 0 \quad \forall i \in P_1. \quad (3.6)$$

Напомним, что *дополнительным* базисом системы КТ называется базис системы (3.3), в который не входят столбцы с номерами, отличающимися только знаком, и тем самым базисный план на дополнительном базисе удовлетворяет условию (3.6). *Экстремальным* базисом называется дополнительный базис, базисное решение на котором удовлетворяет условиям (3.4), (3.5).

Метод ДБ заключается в построении *экстремального* базиса системы КТ (3.3)–(3.6) и состоит из трех частей: алгоритмы А0, А1, А2.

Алгоритм А0. Начиная с дополнительного базиса $Q_0 = P^-$ по определенному правилу строится последовательность базисов Q_1, Q_2, \dots , пока не будет построен дополнительный базис Q_r системы (3.3), базисное решение w_r на котором удовлетворяет условию

(3.4), при этом $r = |P_2^-|$. Если оказалось, что w_r удовлетворяет и ограничениям (3.5), то Q_r — экстремальный базис. Если же условие (3.5) для w_r не выполнено, переходим к следующему алгоритму.

Алгоритм А1. С помощью базиса Q_r формируется вспомогательная задача типа (2.4) и для нее строится дополнительный базис Q_{r+1} с базисным планом w_{r+1} , удовлетворяющим условиям (3.5). Затем переходим к последнему алгоритму.

Алгоритм А2. Начиная с дополнительного базиса Q_{r+1} путем перебора дополнительных базисов решается вспомогательная задача.

Таким образом, метод ДБ заключается в переборе базисов системы (3.3), а затем системы с добавленным искусственным столбцом. Переход от одного базиса к другому будем называть *итерацией* метода. На всех итерациях метода следующий базис Q_{k+1} отличается от предыдущего Q_k одним индексом, т. е.

$$Q_{k+1} = Q_k \setminus \{q_k\} \cup \{l_k\}.$$

Правило выбора индекса q_k , который исключается из текущего базиса Q_k , и индекса l_k , который включается в базис вместо q_k , свое для каждого алгоритма: А0, А1, А2.

Общее для всех итераций метода то, что для определения индекса q_k , подлежащего исключению из базиса Q_k , вычисляется $s_k[Q]$ — вектор коэффициентов разложения вводимого в базис столбца $T[P, l_k]$ по базису Q_k . Этот вектор находится из системы

$$T[P, Q_k] \cdot s[Q_k] = T[P, l_k]. \quad (3.7)$$

Кроме того, на каждой итерации производится пересчет базисного решения по формулам (8.4) гл. 1:

$$\left. \begin{aligned} w_{k+1}[l_k] &= w_k[q_k] / s_k[q_k], \\ w_{k+1}[j] &= w_k[j] - w_{k+1}[l_k] \cdot s_k[j] \quad \forall j \in Q_k \setminus \{q_k\}. \end{aligned} \right\} \quad (3.8)$$

Таким образом, для описания метода достаточно описать правила выбора индексов q_k и l_k в каждом из алгоритмов А0, А1, А2.

3.2. Описание метода

В этом разделе сделаем подробное описание метода ДБ с пояснениями. Далее, в разделе 3.3, будет приведен алгоритм метода ДБ с использованием обратных матриц при решении системы (3.7).

А0. За $r = |P_2|$ шагов строится дополнительный базис системы (3.3), базисное решение на котором удовлетворяет лишь условию (3.4).

В качестве начального дополнительного базиса берем $Q_0 = P^-$. Начальный базисный план очевиден:

$$w_0[j] = 0, \quad w_0[-j] = -d[j] \quad \forall j \in P.$$

Пусть $P_2 = \{l_0, l_1, \dots, l_{r-1}\}$, $P_2^- = \{-l_0, -l_1, \dots, -l_{r-1}\}$. Чтобы переменную $w[-l_k]$ сделать равной нулю, достаточно исключить индекс $(-l_k)$ из базиса, а чтобы сохранить дополнителность базиса, в него нужно включить индекс l_k . Поэтому *попытаемся* по очереди заменить все индексы из P_2^- на индексы из P_2 (т.е. столбцы $-E[P, -l_k]$ на столбцы $H[P, l_k]$.)

Прежде чем переходить к описанию алгоритма, заметим, что в процессе исключения индекса $(-l_k)$ из базиса может попутно исключиться еще не рассмотренный индекс $(-l_{k+i}) \in P_2^-$. Поэтому в начале k -го шага следует выяснить, не исключился ли уже индекс $(-l_k)$ из базиса. Опишем k -й шаг ($k = 0, 1, \dots, r-1$), считая что известны дополнительный базис Q_k и базисное решение $w_k[Q_k]$. Переход к очередному дополнительному базису Q_{k+1} происходит следующим образом.

Если $(-l_k) \notin Q_k$, то полагаем $Q_{k+1} = Q_k$, $w_{k+1} = w_k$, и на этом k -й шаг заканчивается.

Если $(-l_k) \in Q_k$, то находим $s_k[Q_k]$ (коэффициенты разложения вводимого в базис столбца $H[P, l_k]$ по столбцам базисной матрицы) из системы

$$T[P, Q_k] \cdot s[Q_k] = H[P, l_k].$$

Возможны четыре случая.

1) $s_k[-l_k] \neq 0$. Тогда замену индекса $(-l_k)$ на l_k произвести можно, т.е. новый *дополнительный* базис Q_{k+1} находится за одну итерацию:

$$Q_{k+1} = Q_k \setminus \{-l_k\} \cup \{l_k\}.$$

Полагая $q_k = -l_k$, по формулам (3.8) находим новый базисный план w_{k+1} . На этом k -й шаг заканчивается.

2) $s_k[-l_k] = 0$, по существу индекс $j \in Q_k \setminus P_2$, на котором

$$s_k[j] \neq 0. \quad (3.9)$$

В этом случае замена индекса $(-l_k)$ на l_k проводится за две итерации. Сначала строим *почти дополнительный* базис

$$Q_{k1} = Q_k \setminus \{j\} \cup \{l_k\}.$$

В базисе Q_{k1} содержатся индексы l_k и $(-l_k)$ и отсутствуют индексы j и $(-j)$. Вычисляем w_{k1} по формулам (3.8), полагая $q_k = j$ и заменяя индекс $k+1$ на индекс $k1$.

На следующей итерации вводим в базис индекс $(-j)$. Для этого, как обычно, находим s_{k1} из системы

$$T[P, Q_{k1}] \cdot s[Q_k] = T[P, -j].$$

Ниже будет показано (см. теорему 3.1), что

$$s_{k1}[-l_k] = -s_k[j],$$

а тогда в силу (3.9) $s_{k1}[-l_k] \neq 0$. Это дает возможность заменить в Q_{k1} индекс $(-l_k)$ на индекс $(-j)$. Получаем

$$\begin{aligned} Q_{k+1} &= Q_{k1} \setminus \{-l_k\} \cup \{-j\} = \\ &= Q_k \setminus \{j\} \cup \{-j\} \cup \{l_k\} \setminus \{-l_k\}. \end{aligned}$$

Очевидно, что базис Q_{k+1} дополнительный и не содержит индекса $(-l_k)$. Вычисляем w_{k+1} через w_{k1} , полагая в (3.8) $q_k = -l_k$, а вместо l_k подставляя $(-j)$. На этом k -й шаг алгоритма А0 заканчивается.

Заметим, что если $j \in P_2^-$, то попутно произойдет замена индекса j на индекс $(-j) \in P_2$, что сократит объем вычислений.

3) $s_k[Q_k \setminus P_2] = \emptyset$ и при этом $w_k[-l_k] \neq 0$. В этом случае исходная задача (3.1), (3.2) *неразрешима* (см. ниже теорему 3.2). На этом вычислительный процесс по методу ДБ прекращается.

4) $s_k[Q_k \setminus P_2] = \emptyset$ и при этом $w_k[-l_k] = 0$. Так как для компоненты $w_k[-l_k]$ выполнено условие (3.4), то столбец $-E[P, -l_k]$ оставлен в базисе. Ниже будет показано, что компонента $w[-l_k]$ сохранит нулевое значение до конца вычислений (см. ниже теорему 3.3). В этом случае $Q_{k+1} = Q_k$, и на этом k -й шаг алгоритма А0 заканчивается.

Итак, если мы не попали в п. 3, то за $r = |P_2^-|$ шагов просматривается все множество P_2^- и строится дополнительный базис Q_r системы (3.3), такой, что соответствующее базисное решение w_r

удовлетворяет условию (3.4). Заметим, что если P_2^- пусто, то $r = 0$ и $Q_r = Q_0 = P^-$.

Если w_r удовлетворяет условиям (3.5), то вектор $w_r[R]$ является решением системы КТ, а $w_r[N]$ — решением исходной задачи (3.1). На этом вычисления по методу ДБ завершаются.

Если же w_r не удовлетворяет условию (3.5), то переходим к алгоритму А1.

А1. Этот алгоритм реализуется за одну итерацию. Формируем вспомогательную задачу, как в §2. Напомним, что $R_1 = P_1 \cup P_1^-$, и положим

$$L_r = \{j \in Q_r \cap R_1 \mid w_r[j] < 0\}.$$

Пусть $q_r \in L_r$ — тот индекс, на котором $w_r[q_r] = \min_{j \in L_r} w_r[j]$.

Расширим систему (3.3), добавив к матрице $T[P, R]$ искусственный столбец $T[P, 0]$. Положив $R_0 = R \cup \{0\}$, запишем *расширенную систему*

$$T[P, R_0] \cdot w[R_0] = d[P].$$

Искусственный столбец зададим явным разложением по столбцам базисной матрицы $T[P, Q_r]$ в виде

$$T[P, 0] := T[P, Q_r] \cdot s_r[Q_r],$$

где

$$s_r[j] = \begin{cases} -1, & \text{если } j \in L_r, \\ 0, & \text{если } j \notin L_r. \end{cases}$$

Индексное множество Q_r остается базисом расширенной системы, а вектор $w_r[R_0]$ — базисным решением с $w_r[0] = 0$. Заменяем в базисе столбец $T[P, q_r]$ на искусственный столбец $T[P, 0]$, т.е. положим

$$Q_{r+1} = Q_r \setminus \{q_r\} \cup \{0\}.$$

Очевидно, что Q_{r+1} является *дополнительным* базисом расширенной системы, т.е. не содержит одновременно индексов, отличающихся лишь знаком. Вычислим вектор w_{r+1} согласно формулам (2.5):

$$\begin{aligned} w_{r+1}[j] &= w_r[j] - w_r[q_r] \geq 0 \quad \forall j \in L_r, \\ w_{r+1}[0] &= -w_r[q_r] > 0. \end{aligned}$$

Итак, базис Q_{r+1} является дополнительным базисом расширенной системы, а базисный план w_{r+1} удовлетворяет условиям (3.4)–(3.6) и условию $w_{r+1}[0] > 0$, но не удовлетворяет системе (3.3). Переходим к последней части метода.

А2. Этот алгоритм заключается в переборе базисов расширенной системы начиная с дополнительного базиса Q_{r+1} . При этом изменение базисов производится так, чтобы базисы были дополнительными, базисные планы w_k удовлетворяли условию (3.5) и $w_k[0] \geq 0$. Вычисления закончатся, когда выполнится равенство $w_k[0] = 0$. Заметим, что условие (3.6) по-прежнему выполняется автоматически. Условие (3.4) тоже выполняется автоматически (см. ниже теорему 3.3).

Опишем правило построения нового базиса. Заметим, что все рассматриваемые дополнительные базисы расширенной системы содержат индекс 0, поэтому каждый из них имеет одну и только одну пару индексов $\{j, -j\}$, такую, что ни j , ни $(-j)$ не принадлежат базису. У базиса Q_{r+1} такой парой является $\{q_r, -q_r\}$, при этом индекс q_r исключился из предыдущего базиса Q_r . Для сохранения дополнительности нового базиса вводим в него индекс $l_{r+1} = -q_r$. На следующих итерациях поступаем аналогично: чтобы базис Q_{k+1} был дополнительным, вводим в него индекс, который исключился на предыдущей итерации, но с противоположным знаком, т.е.

$$Q_{k+1} = Q_k \setminus \{q_k\} \cup \{l_k\}, \quad l_k = -q_{k-1}.$$

Заметим, что теперь индекс l_k может быть и отрицательным.

Индекс q_k , подлежащий исключению из базиса, определяем в соответствии со знаковыми ограничениями

$$w[P_1] \geq 0, \quad w[P_1^-] \geq 0, \quad w[0] \geq 0$$

следующим образом. Находим вектор $s_k[Q_k]$ из системы

$$T[P, Q_k] \cdot s[Q_k] = T[P, l_k]. \quad (3.10)$$

Ниже (см. теорему 3.4) будет доказано, что $s_k[0] \geq 0$.

Если

$$s_k[0] = 0 \text{ и } s_k[Q_k \cap R_1] \leq 0,$$

то исходная задача неразрешима (см. далее теорему 3.5). На этом вычислительный процесс заканчивается.

Иначе — вычисляем величину

$$t_k = \min \left\{ \frac{w_k[j]}{s_k[j]} \mid j \in Q_k \cap R_1 \cup \{0\} : s_k[j] > 0 \right\}.$$

Пусть L_k — множество тех индексов, на которых реализуется минимум при вычислении t_k . Если $0 \in L_k$, то полагаем $q_k = 0$. В противном случае в качестве q_k берем любой индекс из L_k .

Далее строим новый базис $Q_{k+1} = Q_k \setminus \{q_k\} \cup \{l_k\}$ и пересчитываем базисный план по формулам (3.8).

Если $q_k = 0$, то базис Q_{k+1} является экстремальным, а вектор $x_*[N] = w_{k+1}[N]$ — решение исходной задачи. На этом вычисления заканчиваются.

Если $q_k \neq 0$, то полагаем $l_{k+1} = -q_k$, $k = k+1$ и снова переходим к решению системы (3.10).

Описание алгоритма метода ДБ закончено.

Заметим, что так как $t_k \geq 0$ и $s_k[0] \geq 0$, то

$$w_{k+1}[0] := w_k[0] - t_k \cdot s_k[0] \leq w_k[0],$$

т. е. переменная при искусственном векторе не возрастает, оставаясь положительной. Практически вычисления можно прекратить как только искусственная переменная $w_{k+1}[0]$ обратится в ноль (или будет достаточно мала), не дожидаясь вывода индекса 0 из базиса.

С помощью метода ДБ результат будет получен за конечное число шагов, если в алгоритме А2 не произойдет повторение базиса, приводящее к закликиванию. Для устранения закликивания можно воспользоваться известным правилом минимального индекса (см. гл. 3, §6).

3.3. Алгоритм метода с обратными матрицами

При описании метода не определялся способ решения системы (3.7), которую приходится решать почти на всех итерациях. В зависимости от структуры матрицы $T[P, R]$ можно выбирать разные

способы решения этой системы. Кроме того, как и в линейном программировании, можно использовать обратные базисные матрицы $B_k[Q_k, P] = (T[P, Q_k])^{-1}$. Так как на каждой итерации метода новый базис отличается от предыдущего одним индексом, то новая обратная базисная матрица легко пересчитывается через предыдущую по формулам (8.3) гл. 1.

Приведем алгоритм метода ДБ с использованием обратных базисных матриц.

А0. Имеем дополнительный базис $Q_0 = P^-$, базисный план

$$w_0[j] = 0, \quad w_0[-j] = -d[j] \quad \forall j \in P,$$

обратную базисную матрицу $B_0[Q_0, P]$, которая равна единичной со знаком минус, и множество индексов

$$P_2^- = \{-l_0, -l_1, \dots, -l_k, \dots, -l_{r-1}\},$$

которые желательно исключить из базиса. Полагаем $k = 0$ и продолжаем r следующих шагов.

1. Если $k = r$, то переходим к п. 2, иначе выполняем следующие операции.

Если $(-l_k) \notin Q_k$, то полагаем $Q_{k+1} = Q_k$, $w_{k+1} = w_k$, $k = k+1$ и возвращаемся к началу п. 1.

Если $(-l_k) \in Q_k$, то вычисляем

$$s_k[Q_k] = B_k[Q_k, P] \cdot H[P, l_k]$$

и исследуем четыре случая.

а) Если $s_k[-l_k] \neq 0$, то формируем дополнительный базис $Q_{k+1} = Q_k \setminus \{-l_k\} \cup \{l_k\}$. Полагаем $q_k = -l_k$ и пересчитываем базисное решение w_{k+1} по формулам (3.8) и обратную базисную матрицу $B_{k+1}[Q_{k+1}, P]$ по формулам (см. (8.3) гл. 1):

$$B_{k+1}[l_k, P] = \frac{1}{s_k[q_k]} B_k[q_k, P],$$

$$B_{k+1}[i, P] = B_k[i, P] - s_k[i] \cdot B_{k+1}[l_k, P] \quad \forall i \in Q_k \setminus \{q_k\}. \quad (3.11)$$

Полагаем $k = k+1$ и возвращаемся к началу п. 1.

б) Если $s_k[-l_k] = 0$ и $s_k[j] \neq 0$, где $j \in Q_k \setminus P_2$, то производим две итерации. На первой итерации строим почти дополнительный базис $Q_{k1} = Q_k \setminus \{j\} \cup \{l_k\}$. Полагаем $q_k = j$ и вычисляем w_{k1} и

B_{k1} по формулам (3.8) и (3.11) с заменой индекса $k+1$ на $k1$. На второй итерации в базис вводится $(-j)$, поэтому находим

$$s_{k1}[Q_{k1}] = B_{k1}[Q_{k1}, P] \cdot T[P, -j],$$

строим дополнительный базис $Q_{k+1} = Q_{k1} \setminus (-l_k) \cup \{-j\}$, пересчитываем w_{k+1} и B_{k+1} , заменяя в формулах (3.8) и (3.11) индекс k на $k1$, l_k на $(-j)$ и q_k на $(-l_k)$. Полагаем $k = k+1$ и возвращаемся к началу п. 1.

в) Если $s_k[Q_k \setminus P_2] = \mathbb{O}$ и при этом $w_k[-l_k] = 0$, то полагаем $Q_{k+1} = Q_k$, $k = k+1$ и переходим к началу п. 1.

г) Если $s_k[Q_k \setminus P_2] = \mathbb{O}$ и при этом $w_k[-l_k] \neq 0$, то исходная задача (3.1), (3.2) неразрешима. На этом вычислительный процесс метода ДБ прекращается.

2. Если $w_r[R_1] \geq \mathbb{O}$, то вектор $w_r[R]$ является решением системы КТ, а $w_r[N]$ — решением исходной задачи (3.1). На этом вычисления по методу ДБ завершаются. В противном случае переходим к А1.

А1. Находим $L_r = \{j \in Q_r \cap R_1 \mid w_r[j] < 0\}$ и $w_r[q_r] = \min_{j \in L_r} w_r[j]$.

Меняем базис $Q_{r+1} := Q_r \setminus \{q_r\} \cup \{0\}$ и пересчитываем базисный план:

$$\begin{aligned} w_{r+1}[0] &= -w_r[q_r], \\ w_{r+1}[j] &= w_r[j] - w_r[q_r] \quad \forall j \in L_r. \end{aligned}$$

Пересчитываем строки обратной базисной матрицы с индексами i только из L_r :

$$\begin{aligned} B_{r+1}[i, P] &= B_r[i, P] - B_r[q_r, P], \quad i \neq q_r, \\ B_{r+1}[0, P] &= -B_r[q_r, P]. \end{aligned}$$

Увеличиваем k на единицу, т.е. $k := r+1$, и переходим к А2.

А2. Полагаем $l_k = -q_{k-1}$ и находим

$$s_k[Q_k] = B_k[Q_k, P] \cdot T[P, l_k].$$

Если $s_k[0] = 0$ и $s_k[Q_k \cap R_1] \leq \mathbb{O}$, то исходная задача неразрешима и на этом вычислительный процесс заканчивается. Иначе — вычисляем

$$t_k = \min \left\{ \frac{w_k[j]}{s_k[j]} \mid j \in Q_k \cap R_1 \cup \{0\} : s_k[j] > 0 \right\} = \frac{w_k[q_k]}{s_k[q_k]}.$$

Меняем базис $Q_{k+1} := Q_k \setminus \{q_k\} \cup \{l_k\}$ и пересчитываем базисный план по формулам (3.8).

Если $q_k = 0$, то базис Q_{k+1} является экстремальным, а вектор $x_*[N] = w_{k+1}[N]$ — решением исходной задачи. На этом вычисления заканчиваются.

Если $q_k \neq 0$, то пересчитываем обратную базисную матрицу по формулам (3.11), полагаем $k = k+1$ и переходим к началу А2.

3.4. Обоснование метода

Рассмотрим k -ю итерацию метода. Напомним, что через l_k обозначается индекс, вводимый в базис. Обоснование метода базируется на свойствах $s_k[Q_k]$ — вектора коэффициентов разложения вводимого в базис столбца $T[P, l_k]$ по базису Q_k :

$$T[P, l_k] = T[P, Q_k] \cdot s_k[Q_k]. \quad (3.12)$$

Прежде всего доопределим вектор s_k на все индексное множество $R_0 = P \cup P^- \cup \{0\}$, положив $s_k[l_k] = -1$, а все остальные небазисные компоненты — равными нулю. Напомним, что число шагов в алгоритме А0 равно $r = |P_2^-|$. Перепишем (3.12) для разных k :

$$\begin{aligned} H[P, P] \cdot s_k[P] - E[P, P^-] \cdot s_k[P^-] &= \mathbb{O}, & k < r, \\ H[P, P] \cdot s_k[P] - E[P, P^-] \cdot s_k[P^-] + T[P, 0] \cdot s_k[0] &= \mathbb{O}, & k \geq r. \end{aligned} \quad (3.13)$$

Напомним, что матрица $H[P, P]$ неотрицательно определена, но не симметрична, поэтому

$$\text{если } \langle Hz, z \rangle = 0, \text{ то } (H + H^T) \cdot z = \mathbb{O}. \quad (3.14)$$

Начнем с обоснования алгоритма А0. Введем величину

$$\alpha_k := s_k[P] \cdot s_k[P^-].$$

Лемма 3.1. Если базис Q_k дополнительный, то для $k < r$

$$\alpha_k = -s_k[-l_k] = s_k[P] \cdot H[P, P] \cdot s_k[P] \geq 0.$$

Доказательство. Вычислим α_k непосредственно. Так как базис Q_k дополнительный, то у векторов $s_k[P]$ и $s_k[P^-]$ может быть

лишь одна общая ненулевая компонента с индексами l_k и $-l_k$. Но $s_k[l_k] = -1$, поэтому $\alpha_k = -s_k[-l_k]$.

Далее, домножая обе части первого равенства (3.13) слева на $s_k[P]$, имеем

$$\alpha_k = s_k[P] \cdot H[P, P] \cdot s_k[P].$$

Учитывая, что матрица H неотрицательно определена, получаем требуемое. ■

Обратимся к п. 2 алгоритма А0.

Теорема 3.1. Если $s_k[-l_k] = 0$ и существует индекс $j \in Q_k \setminus P_2$, такой, что $s_k[j] \neq 0$, то $s_{k1}[-l_k] = -s_k[j]$.

Доказательство. Рассмотрим сумму скалярных произведений

$$\eta = s_k[P] \cdot s_{k1}[P^-] + s_k[P^-] \cdot s_{k1}[P].$$

Базис Q_k — дополнительный, а $Q_{k1} = Q_k \setminus \{j\} \cup \{l_k\}$ — почти дополнительный, он содержит индексы l_k и $(-l_k)$. Поэтому ненулевые пары в скалярных произведениях для η могут быть лишь на индексах $l_k, (-l_k), j, (-j)$. Учитывая, что индекс l_k вводится в базис Q_k , а индекс $(-j)$ вводится в базис Q_{k1} , имеем

$$s_k[l_k] = -1, \quad s_{k1}[-j] = -1.$$

Заметим, что $j \notin Q_{k1}$, поэтому $s_{k1}[j] = 0$. Так как по условию теоремы $s_k[-l_k] = 0$, то независимо от знака индекса j получаем

$$\eta = s_k[j] \cdot s_{k1}[-j] + s_k[l_k] \cdot s_{k1}[-l_k] = -s_k[j] - s_{k1}[-l_k].$$

Осталось показать, что $\eta = 0$.

Из (3.13) при $k < r$ имеем

$$s_k[P^-] = H[P, P] \cdot s_k[P], \quad s_{k1}[P^-] = H[P, P] \cdot s_{k1}[P].$$

Домножим слева первое равенство на вектор $s_{k1}[P]$, а второе — на $s_k[P]$ и сложим. Тогда

$$\eta = s_{k1}[P] \cdot (H[P, P] + H^T[P, P]) \cdot s_k[P].$$

Так как $s_k[-l_k] = 0$, то по лемме 3.1 $s_k[P] \cdot H[P, P] \cdot s_k[P] = 0$. Тогда в силу (3.14) имеем $(H[P, P] + H^T[P, P]) \cdot s_k[P] = 0$, значит, $\eta = 0$. ■

Обратимся к п. 3 алгоритма А0. Сначала докажем вспомогательное утверждение, которое касается также и п. 4.

Лемма 3.2. Если $s_\mu[Q_\mu \setminus P_2] = \mathbb{O}$ при $\mu < r$, то

$$s_\mu[P_1] = \mathbb{O}, \quad (3.15)$$

$$s_\mu[P] \cdot H[P, P] = \mathbb{O}. \quad (3.16)$$

Доказательство. Заметим, что все небазисные компоненты $s_\mu[R]$ равны нулю, кроме $s_\mu[l_\mu] = -1$, при этом $l_\mu \in P_2$. Значит, из условия леммы следует, что ненулевыми компонентами $s_\mu[R]$ могут быть компоненты с индексами из множества $Q_\mu \cap P_2 \cup \{l_\mu\}$. В этом множестве нет индексов из P_1 и отрицательных индексов, поэтому

$$s_\mu[P_1] = \mathbb{O}, \quad s_\mu[P^-] = \mathbb{O},$$

т. е. вектор $s_\mu[P]$ удовлетворяет (3.15).

Теперь из (3.13) имеем

$$\mathbb{O} = s_\mu[P^-] = H[P, P] \cdot s_\mu[P].$$

Тогда в силу (3.14) $s_\mu[P] \cdot H[P, P] = \mathbb{O}$, т. е. вектор $s_\mu[P]$ удовлетворяет (3.16). ■

Теорема 3.2. Пусть для $\mu < r$ выполнены условия

$$s_\mu[Q_\mu \setminus P_2] = \mathbb{O}, \quad w_\mu[-l_\mu] \neq 0, \quad (3.17)$$

тогда система КТ несовместна.

Доказательство. Несовместность всей системы КТ следует из несовместности системы (3.3), (3.4), которую можно переписать в виде

$$H[P, P] \cdot w[P] - E[P, P_1^-] \cdot w[P_1^-] = d[P].$$

Известно [3], что эта система несовместна тогда и только тогда, когда существует вектор $\xi[P]$ со свойствами

$$\begin{aligned} \xi[P] \cdot H[P, P] &= \mathbb{O}, \\ \xi[P] \cdot E[P, P_1^-] &= \xi[P_1^-] = \mathbb{O}, \\ \langle \xi, d \rangle &\neq 0. \end{aligned} \quad (3.18)$$

Покажем, что такими свойствами обладает вектор $s_\mu[P]$.

Выполнение первых двух равенств для $s_\mu[P]$ следует из леммы 3.2. Проверим справедливость последнего неравенства из (3.18). Ясно, что вектор w_μ удовлетворяет равенству

$$H[P, P] \cdot w_\mu[P] - E[P, P^-] \cdot w_\mu[P^-] = d[P].$$

Умножим это равенство слева на вектор $s_\mu[P]$ и учтем доказанное первое равенство (3.18). Тогда

$$s_\mu[P] \cdot d[P] = -s_\mu[P] \cdot w_\mu[P^-] = -s_\mu[l_\mu] \cdot w_\mu[-l_\mu] = w_\mu[-l_\mu].$$

В силу второго неравенства (3.17) для вектора $s_\mu[P]$ выполнено последнее неравенство (3.18). ■

Обратимся к п.4 алгоритма A0. В этом случае попытка исключения из базиса индекса $(-l_\mu) \in P_2^-$ не удалась, и в базисе остался искусственный вектор $-E[P, -l_\mu]$ с нулевой компонентой $w_\mu[-l_\mu]$. Покажем, что на всех следующих итерациях метода ДБ $w_k[-l_\mu] = 0$.

Теорема 3.3. Пусть для $\mu < r$ выполнены равенства

$$s_\mu[Q_\mu \setminus P_2] = \mathbb{O}, \quad w_\mu[-l_\mu] = 0,$$

тогда на всех следующих итерациях метода ДБ $w_k[-l_\mu] = 0$.

Доказательство. Заметим, что по алгоритму

$$w_{k+1}[-l_\mu] = w_k[-l_\mu] - t_k \cdot s_k[-l_\mu].$$

По условию теоремы $w_\mu[-l_\mu] = 0$ и $s_\mu[-l_\mu] = 0$. Поэтому теорема будет доказана, если показать, что $s_k[-l_\mu] = 0$ при $k > \mu$.

Пусть $\Gamma = Q_\mu \cap P_2$. Согласно алгоритму $\Gamma \subset Q_k$ при всех $k \geq \mu$, поэтому индексы $i \in \Gamma^-$ всегда небазисные. Значит,

$$s_k[\Gamma^-] = \mathbb{O} \quad \forall k \geq \mu.$$

По условию теоремы ненулевыми компонентами вектора $s_\mu[P]$ могут быть только компоненты с индексами из Γ и еще небазисная компонента $s_\mu[l_\mu] = -1$. Тогда

$$\eta_k := s_\mu[P] \cdot s_k[P^-] = s_\mu[\Gamma] \cdot s_k[\Gamma^-] - s_k[-l_\mu] = -s_k[-l_\mu]. \quad (3.19)$$

Вычислим η_k при разных k .

Пусть $k < r$. Из первого равенства (3.13) имеем

$$s_k[P^-] = H[P, P] \cdot s_k[P].$$

Подставляя это выражение в представление η_k и учитывая (3.16), получаем

$$\eta_k = s_\mu[P] \cdot H[P, P] \cdot s_k[P] = 0.$$

Теперь из (3.19) следует, что $s_k[-l_\mu] = 0$.

Пусть $k \geq r$. Умножим слева обе части второго равенства (3.13) на $s_\mu[P]$. Учитывая (3.16), имеем

$$\eta_k = \lambda_\mu \cdot s_k[0], \quad \text{где } \lambda_\mu = s_\mu[P] \cdot T[P, 0].$$

Покажем, что $\lambda_\mu = 0$. Согласно алгоритму A1

$$T[P, 0] = T[P, Q_r] \cdot s_r[Q_r],$$

все компоненты вектора $s_r[L_r]$ равны -1 , а остальные нули. Доопределим вектор s_r на все множество R нулями. Тогда

$$\lambda_\mu = s_\mu[P] \cdot (H[P, P] \cdot s_r[P] - s_r[P^-]).$$

В силу (3.16) $\lambda_\mu = -s_\mu[P] \cdot s_r[P^-] = -s_\mu[P_1] \cdot s_r[P_1^-]$. Учитывая (3.15), получаем $\lambda_\mu = 0$, т.е. $\eta_k = 0$. Теперь из (3.19) следует, что $s_k[-l_\mu] = 0$. ■

Перейдем к обоснованию алгоритма A2, поэтому далее будем рассматривать $k > r$.

Алгоритм A2 можно интерпретировать как некоторый вариант симплекс-метода для следующей задачи ЛП:

$$\begin{aligned} w[0] &\rightarrow \min, \\ T[P, R] \cdot w[R] + T[P, 0] \cdot w[0] &= d[P], \\ w[P_1 \cup P_1^-] &\geq \mathbb{O}, \quad w[0] \geq 0. \end{aligned} \quad (3.20)$$

Заметим, что эта задача разрешима, так как вектор w_{r+1} является планом задачи, а целевая функция ограничена снизу нулем.

Сравним алгоритм A2 с симплекс-методом. Согласно алгоритму ДБ каждый вектор w_k при $k \geq r+1$ является планом задачи (3.20). Кроме того, процесс исключения индекса q_k из базиса в методе ДБ

такой же как, в симплекс-методе. Осталось сравнить способы выбора индекса l_k , подлежащего включению в базис. Как известно, в симплекс-методе в очередной базис включается какой-либо столбец $T[P, l_k]$ с отрицательной оценкой. При этом выбор такого столбца среди всех, имеющих отрицательную оценку, может осуществляться разными способами.

В методе ДБ оценки не вычисляются, а для сохранения дополнителности следующего базиса в текущий базис вводится единственно возможный столбец с индексом $l_k = -q_{k-1}$. Спрашивается, какова оценка столбца $T[P, l_k]$. Ниже будет показано, что в методе ДБ оценка вводимого вектора всегда неположительна, а если она оказалась равной нулю (целевая функция не меняется), то на следующей итерации она обязательно отрицательна. Это означает, что целевая функция задачи (3.20) не возрастает: $w_{k+1}[0] \leq w_k[0]$.

Перейдем к доказательству утверждений, обосновывающих алгоритм А2. Обозначим $B_k[Q_k, P] := (T[P, Q_k])^{-1}$ и введем в рассмотрение *вектор оценок* $\Delta_k[R]$, определив его следующим образом:

$$\Delta_k[R] = -B_k[0, P] \cdot T[P, R]. \quad (3.21)$$

Заметим, что строка обратной базисной матрицы $B_k[0, P]$ является двойственным вектором для задачи (3.20), поэтому определение вектора оценок полностью соответствует определению оценок в симплекс-методе для этой задачи. Из (3.21) и представления матрицы $T[P, R]$ (см.(3.3)) следует

$$\Delta_k[P^-] = B_k[0, P], \quad \Delta_k[P] = -\Delta_k[P^-] \cdot H[P, P]. \quad (3.22)$$

Заметим, что при $k > r$ все базисы Q_k *дополнительные* и имеют структуру

$$Q_k = I_k \cup (-J_k) \cup \{0\}, \quad I_k \cap J_k = \emptyset.$$

Так как $B_k[0, P]$ является строкой обратной базисной матрицы, то оценки базисных столбцов таковы:

$$\Delta_k[I_k] = \mathbb{O}, \quad \Delta_k[-J_k] = \mathbb{O}, \quad \Delta_k[P^-] \cdot T[P, 0] = 1. \quad (3.23)$$

Напомним, что $s_k[Q_k] = B_k[Q_k, P] \cdot T[P, l_k]$, где l_k — вводимый в базис индекс. Согласно (3.21) для вводимого в базис столбца имеем оценку

$$\Delta_k[l_k] = -s_k[0]. \quad (3.24)$$

Обозначим $\delta_k := \Delta_k[-l_k]$. Очевидно, что пара индексов $\{l_k, -l_k\}$ не принадлежит базису Q_k , и так как l_k может быть любого знака, то

$$P = I_k \cup J_k \cup \{l_k\}. \quad (3.25)$$

Отсюда следует, что $\Delta_k[P] \cdot \Delta_k[P^-] = -s_k[0] \cdot \delta_k$. Из представлений (3.22) получаем равенство

$$s_k[0] \cdot \delta_k = \Delta_k[P^-] \cdot H[P, P] \cdot \Delta_k[P^-] \geq 0. \quad (3.26)$$

Установим связь между вектором оценок $\Delta_k[R]$ и вектором $s_k[R]$ в случае, когда $s_k[0] = 0$.

Лемма 3.3. *Если $s_k[0] = 0$ при $k > r$, то*

$$\delta_k \neq 0, \quad \Delta_k[P] = -\delta_k \cdot s_k[P^-], \quad \Delta_k[P^-] = -\delta_k \cdot s_k[P].$$

Доказательство. Из (3.26) при $s_k[0] = 0$ следует, что

$$\Delta_k[P^-] \cdot H[P, P] \cdot \Delta_k[P^-] = 0.$$

Тогда в силу (3.14) имеем

$$\Delta_k[P^-] \cdot H[P, P] = -H[P, P] \cdot \Delta_k[P^-].$$

Заменяя правое выражение в (3.22), получаем

$$H[P, P] \cdot \Delta_k[P^-] - E[P, P^-] \cdot \Delta_k[P] = \mathbb{O}.$$

Учитывая равенства (3.25), (3.24) и то, что $s_k[0] = 0$, последнее равенство перепишем в виде

$$H[P, I_k] \cdot \Delta_k[-I_k] - E[P, -J_k] \cdot \Delta_k[J_k] + \delta_k \cdot T[P, l_k] = \mathbb{O} \quad (3.27)$$

(если $l_k > 0$, то коэффициент δ_k — при векторе $H[P, l_k]$, в противном случае — при $(-E[P, l_k])$, т. е. всегда при $T[P, l_k]$). Покажем сначала, что $\delta_k \neq 0$. Допустим противное, т. е. $\delta_k = 0$. Заметим, что в силу (3.23) вектор $\Delta_k[P^-]$ ненулевой и его ненулевые компоненты могут быть лишь с индексами из $(-I_k)$, так как $\Delta_k[l_k] = \Delta_k[-l_k] = 0$. Значит, $\Delta_k[-I_k] \neq 0$. А тогда равенство (3.27) означает линейную зависимость базисных столбцов, что невозможно.

Сравнивая (3.27) с разложением столбца $T[P, l_k]$ по базису Q_k и учитывая, что $\delta_k \neq 0$, убеждаемся в справедливости леммы. ■

Установим связь между $s_k[0]$ и δ_{k+1} . Напомним, что через q_k обозначается индекс, выводимый из базиса Q_k . Пусть

$$c_k := s_k[q_k].$$

Согласно методу ДБ $c_k > 0$ на всех итерациях алгоритма А2.

Лемма 3.4. *Имеют место равенства*

$$\delta_{r+1} = 1, \quad s_k[0] = c_k \cdot \delta_{k+1} \quad \forall k > r.$$

Доказательство. Учитывая, что $l_{k+1} = -q_k$, из (3.21) имеем

$$\delta_{k+1} = \Delta_{k+1}[-l_{k+1}] = -B_{k+1}[0, P] \cdot T[P, q_k] \quad \forall k \geq r. \quad (3.28)$$

Покажем, что $\delta_{r+1} = 1$. Напомним, что на r -й итерации (алгоритм А1) в базис вводится искусственный вектор $T[P, 0]$ и $c_r = s_r[q_r] = -1$, поэтому согласно формулам пересчета обратной базисной матрицы (см. (8.3) гл. 1) имеем ($l_r = 0$)

$$B_{r+1}[0, P] = -B_r[q_r, P].$$

Так как столбец $T[P, q_r]$ при $k = r$ является базисным, то

$$B_r[q_r, P] \cdot T[P, q_r] = 1.$$

Теперь из (3.28) при $k = r$ получаем $\delta_{r+1} = 1$.

Пусть $k > r$. Воспользуемся формулами преобразования строки $B_k[0, P]$, замечая, что она уже не является ведущей строкой:

$$B_{k+1}[0, P] = B_k[0, P] - \frac{s_k[0]}{c_k} B_k[q_k, P] \quad \forall k \geq r + 1.$$

Подставим это выражение в (3.28), учитывая, что

$$B_k[0, P] \cdot T[P, q_k] = 0, \quad B_k[q_k, P] \cdot T[P, q_k] = 1.$$

Тогда получим требуемую формулу. ■

Перейдем к доказательству основных утверждений.

Теорема 3.4. *Если $k > r$, то $s_k[0] \geq 0$. Если же $s_k[0] = 0$, то $\delta_k > 0$ и $s_{k+1}[0] > 0$.*

Доказательство. Из (3.26) имеем

$$s_k[0] \cdot \delta_k \geq 0 \quad \forall k > r. \quad (3.29)$$

Так как $\delta_{r+1} = 1$, то

$$s_{r+1}[0] \geq 0. \quad (3.30)$$

Далее, подставляя в (3.29) выражение для δ_k из леммы 3.4 и учитывая, что $c_{k-1} > 0$ для $k > r + 1$, получаем

$$s_k[0] \cdot s_{k-1}[0] \geq 0 \quad \forall k > r + 1. \quad (3.31)$$

Рассмотрим числовую последовательность

$$s_{r+1}[0], s_{r+2}[0], \dots, s_k[0], \dots$$

Если все члены этой последовательности отличны от нуля, то согласно (3.30), (3.31) все они положительны. Допустим, что на k -й итерации $s_k[0] = 0$, причем $s_{k-1}[0] > 0$. Покажем, что $s_{k+1}[0] > 0$.

Так как $s_k[0] = 0$, то согласно формулам пересчета обратной базисной матрицы имеем $B_{k+1}[0, P] = B_k[0, P]$. Из определения вектора оценок $\Delta_{k+1}[R] = \Delta_k[R]$, поэтому в силу (3.24)

$$s_{k+1}[0] = -\Delta_{k+1}[l_{k+1}] = -\Delta_k[l_{k+1}] = -\Delta_k[-q_k].$$

Так как $s_k[0] = 0$, то по лемме 3.3 $\Delta_k[-q_k] = -\delta_k \cdot s_k[q_k] = -\delta_k \cdot c_k$, т. е.

$$s_{k+1}[0] = \delta_k \cdot c_k. \quad (3.32)$$

Если $k = r + 1$, то $\delta_k = 1$, и тогда $s_{r+2}[0] = c_{r+1} > 0$.

Если $k > r + 1$, то по лемме 3.4 $\delta_k = \frac{s_{k-1}[0]}{c_{k-1}}$, значит, $\delta_k > 0$. Подставляя это выражение в (3.32) и учитывая, что $s_{k-1}[0] > 0$, получаем

$$s_{k+1}[0] = \frac{c_k}{c_{k-1}} s_{k-1}[0] > 0. \quad \blacksquare$$

Обратимся теперь к вопросу о неразрешимости исходной задачи, которая может проявиться в алгоритме А2.

Теорема 3.5. Если при $k > r$ оказалось, что

$$s_k[Q_k \setminus P_2] \leq 0, \quad w_k[0] > 0, \quad (3.33)$$

то система КТ несовместна.

Доказательство. Обозначим через V^- множество тех индексов из P_2^- , которые остались в базисе после завершения алгоритма А0. Положим $K^- = P_1^- \cup V^-$. Рассмотрим задачу ЛП

$$\begin{aligned} H[P, P] \cdot w[P] - E[P, K^-] \cdot w[K^-] + T[P, 0] \cdot w[0] &= d[P], \\ w[P_1] \geq 0, \quad w[P_1^-] \geq 0 \quad w[0] \geq 0. \end{aligned} \quad (3.34)$$

Покажем, что базис Q_k является для нее оптимальным, т. е. оценки удовлетворяют следующим условиям:

$$\Delta_k[P_1 \cup P_1^-] \geq 0, \quad \Delta_k[P_2 \cup V^-] = 0. \quad (3.35)$$

Напомним, что $\Delta_k[Q_k] = 0$, а индексы из V^- и 0 принадлежат базису Q_k . Далее, в силу (3.33) $s_k[0] \leq 0$, поэтому из теоремы 3.4 имеем $s_k[0] = 0$, $\delta_k > 0$. Из леммы 3.3 следует, что в этом случае оценки определяются по формулам

$$\Delta_k[P] = -\delta_k \cdot s_k[P^-], \quad \Delta_k[P^-] = -\delta_k \cdot s_k[P].$$

Так как $s_k[l_k] = -1$, а остальные небазисные компоненты вектора s_k равны нулю, то из (3.33) вытекает, что $s_k[P_1 \cup P_1^-] \leq 0$, т. е. выполнено первое равенство (3.35). Замечаем, что $\Delta_k[V] = -\delta_k \cdot s_k[V^-] = 0$ в силу теоремы 3.3, а $P_2 \setminus V \subset Q_k$. Итак, получаем, что план w_k является оптимальным для задачи (3.34), но $w_k[0] > 0$, поэтому система КТ неразрешима. \blacksquare

3.5. Численные примеры

Пример 3.1. Методом ДБ с использованием обратных базисных матриц решим следующую задачу:

$$\begin{aligned} \frac{1}{2}x[1]^2 + \frac{1}{2}x[2]^2 - \frac{3}{2}x[1] - 2x[2] \rightarrow \min, \\ (3) \quad x[1] + x[2] = 1, \\ x[1] \geq 0, \quad x[2] \geq 0. \end{aligned}$$

Здесь $N = \{1, 2\}$, $M = \{3\}$, $P_2^- = \{-3\}$. КТ-таблица имеет вид (см. §1 гл. 2)

1	2	3	-1	-2	-3	$d[P]$
1	0	-1	-1	0	0	3/2
0	1	-1	0	-1	0	2
1	1	0	0	0	-1	1
+	+	*	+	+	0	

Вычисления удобно записывать и проводить с помощью таблиц. Каждой итерации метода соответствует своя таблица, которая заполняется по следующей схеме:

k	l_k	$T^T[l_k, P]$		
	Q_k	$B_k[Q_k, P]$	$w_k[Q_k]$	$s_k[Q_k]$

Здесь k — номер итерации; l_k — вводимый в базис индекс; $T[P, l_k]$ — столбец с индексом l_k ; Q_k — текущий базис; $B_k[Q_k, P]$ — обратная базисная матрица; $w_k[Q_k]$ — текущий базисный план, равный $B_k[Q_k, P] \cdot d[P]$ (все небазисные компоненты равны нулю); $s_k[Q_k]$ — коэффициенты разложения $T[P, l_k]$ по базису Q_k , так что $s_k[Q_k] = B_k[Q_k, P] \cdot T[P, l_k]$. В первом столбце будем обозначать звездочкой тот индекс из Q_k , для которого знак переменной безразличен, а признак 0 будет означать, что переменную нужно сделать нулевой.

Начальный базис в методе ДБ всегда одинаков:

$$Q_0 = \{-1, -2, -3\},$$

а $B_0[Q_0, P]$ — единичная матрица со знаком минус.

Алгоритм А0 осуществляется за две итерации (две следующие таблицы). Так как $-l_0 = -3$, вводим в базис индекс $l_0 = 3$. Начальная таблица имеет вид

0	3	-1	-1	0		
	-1	-1	0	0	-3/2	1
	-2	0	-1	0	-2	1
0	-3	0	0	-1	-1	0

Так как $s_0[-3] = 0$, то индекс (-3) из базиса не исключать, поэтому исключаем первый возможный, т.е. (-1) , а вводим в базис индекс 1. Пользуясь предыдущей таблицей, пересчитываем обратную матрицу и базисный план и вносим их в очередную таблицу:

01	1	1	0	1		
*	3	-1	0	0	-3/2	-1
	-2	1	-1	0	-1/2	1
0	-3	0	0	-1	-1	-1

Следующая таблица относится к концу алгоритма А0 и началу А1:

1	0					
*	3	-1	0	1	-1/2	0
	-2	1	-1	-1	-3/2	-1
	1	0	0	1	1	0

А1: Из таблицы видим, что $L = \{-2\}$. Исключаем из базиса Q_1 индекс (-2) и включаем индекс искусственного столбца 0. Следующая таблица относится к концу алгоритма А1 и началу А2. Исключению из базиса Q_2 подлежит индекс 0, так что следующий базис $Q_3 = \{3, 2, 1\}$ экстремальный. Остается пересчитать базисный план:

$w_3[3] = -\frac{5}{4}$, $w_3[2] = \frac{3}{4}$, $w_3[1] = \frac{1}{4}$. Очевидно, что решением задачи является вектор $x_* = (\frac{1}{4}, \frac{3}{4})$. Для проверки результата можно подставить решение w_3 в систему КТ:

2	2	0	1	1		
*	3	-1	0	1	-1/2	1
	0	-1	1	1	3/2	2
	1	0	0	1	1	1

Пример 3.2. Приведем задачу с пустым множеством планов. Покажем, что пустота множества может проявиться лишь в последней части метода.

Рассмотрим следующую задачу:

$$\begin{aligned} \frac{1}{2}x[1]^2 + \frac{1}{2}x[2]^2 + x[1] - 2x[2] &\rightarrow \min, \\ x[1] + x[2] &= -1, \\ x[1] \geq 0, \quad x[2] &\geq 0. \end{aligned}$$

Здесь $N = \{1, 2\}$, $M = \{3\}$, $P_2^- = \{-3\}$. КТ-таблица имеет вид

1	2	3	-1	-2	-3	$d[P]$
1	0	-1	-1	0	0	-1
0	1	-1	0	-1	0	2
1	1	0	0	0	-1	-1
+	+	*	+	+	0	

Алгоритм А0 осуществляется за две итерации (две следующие таблицы). Так как $-l_0 = -3$, вводим в базис индекс $l_0 = 3$:

0	3	-1	-1	0		
	-1	-1	0	0	1	1
	-2	0	-1	0	-2	1
0	-3	0	0	-1	1	0

Так как $s_0[-3] = 0$, то индекс (-3) из базиса не исключать, поэтому исключаем первый возможный, т. е. (-1) , а вводим в базис индекс 1:

01	1	1	0	1		
*	3	-1	0	0	1	-1
	-2	1	-1	0	-3	1
0	-3	0	0	-1	1	-1

Следующая таблица относится к концу алгоритма A0 и началу A1.

Алгоритм A1: $L = \{-2, 1\}$. Исключаем из базиса Q_1 индекс $q_r = -2$ и включаем индекс искусственного столбца 0:

1	0					
*	3	-1	0	1	0	0
	-2	1	-1	-1	-2	-1
	1	0	0	1	-1	-1

Следующий базис Q_2 является результатом алгоритма A1. Далее работает алгоритм A2. В базис вводится индекс $(-q_r) = 2$:

2	2	0	1	1		
*	3	-1	0	1	0	1
	0	-1	1	1	2	2
	1	-1	1	2	1	3

Исключается из базиса индекс 1, так как $t_2 = \frac{w_2[1]}{s_2[1]}$, а включается (-1) :

3	-1	-1	0	0		
*	3	-2/3	-1/3	1/3	-1/3	2/3
	0	-1/3	1/3	-1/3	4/3	1/3
	2	-1/3	1/3	2/3	1/3	1/3

Теперь из базиса исключается индекс 2 и вводится индекс (-2) :

4	-2	0	-1	0		
*	3		-1			1
	0		0			0
	-1	-1	1	2		-1

В последней таблице нет надобности пересчитывать остальные компоненты матрицы $B_4[Q_4, P]$ и план $w_4[Q_4]$, так как уже видно, что $s_4[0]$ и $s_4[-1]$ неположительны, поэтому исходная задача неразрешима.

§4. Базисные методы Вулфа и Лемке

Метод Вулфа [5, 11, 22] разработан для задачи ВКП в первой канонической форме

$$F(x) := \frac{1}{2} \langle Dx, x \rangle + \langle c, x \rangle \rightarrow \min, \quad (4.1)$$

$$A[M, N] \cdot x[N] = b[M], \quad x[N] \geq \Phi. \quad (4.2)$$

Предполагаем, что $\text{rank } A[M, N] = |M|$. Сначала, как в симплекс-методе для задач ЛП, находятся базис $N_0 \subset N$ и базисный план $x_0[N]$ системы (4.2). Для этого можно воспользоваться методом искусственного базиса, описанным в §7 гл. 1.

Обратимся к системе КТ (3.3). Для задачи (4.1), (4.2) $P_2^- = M^-$. Можно показать, что индексное множество $Q_0 = N_0 \cup M \cup (-J_0)$, где $J_0 = N \setminus N_0$, является дополнительным базисом системы КТ для задачи (4.1), (4.2), базисное решение на котором удовлетворяет условию $w_0[M^-] = \Phi$. Дальнейший процесс совпадает с методом ДБ, начиная с алгоритма A1. При этом в качестве Q_r используется построенный дополнительный базис Q_0 . По существу, алгоритм A0 метода ДБ заменяется описанным алгоритмом построения базиса Q_0 .

Как было показано на примере 3.2, информация о пустоте планов задачи (4.1), (4.2) может проявиться лишь в алгоритме A2 метода ДБ. Если есть подозрения, что задача не имеет планов, лучше воспользоваться методом Вулфа.

Метод Лемке [1] разработан для задачи ВКП с ограничениями в симметричной форме

$$A[M, N] \cdot x[N] \geq b[M],$$

$$x[N] \geq \mathbb{0}.$$

Сравнивая эти ограничения с общими (см. (3.1) из гл. 1), видим, что здесь $M_2 = \emptyset$, $N_2 = \emptyset$. Метод Лемке полностью совпадает с методом ДБ без алгоритма А0. Вычисления начинаются с алгоритма А1, а в качестве Q_r используется дополнительный базис $Q_0 = P^-$. Именно метод Лемке послужил стимулом для создания метода ДБ.

§5. Метод Данцига

5.1. Постановка задачи и схема метода

Метод Данцига является естественным обобщением известного в линейном программировании симплекс-метода на задачу квадратичного программирования. Метод разработан для задачи, записанной в первой канонической форме:

$$F(x) := \frac{1}{2} \langle Dx, x \rangle + \langle c, x \rangle \rightarrow \min_{x \in \Omega}, \quad (5.1)$$

где

$$\Omega = \{x[N] \mid A[M, N] \cdot x[N] = b[M], x[N] \geq \mathbb{0}\}.$$

При этом никаких ограничений на матрицу A и векторы c , b не накладывается, а симметричная матрица D предполагается неотрицательно определенной, так что метод пригоден для случая $D = \mathbb{0}$ и в этом случае он совпадает с симплекс-методом (см. разд. 5.7).

Метод Данцига начинается с поиска начального базисного плана. Для этого можно воспользоваться методом Булавского (см. §7 гл. 1). Тогда либо выяснится, что Ω пусто, либо будет найден базис и базисный план. При этом, если $\text{rank} A < |M|$, то найденный базис будет содержать искусственные столбцы-орты, поэтому множество Ω следует заменить на Ω_* вида (7.4) (гл. 1), а матрицу D и вектор c следует дополнить нулями по числу оставшихся в базисе искусственных переменных. В связи со сказанным будем далее считать, что

$$\text{rank} A[M, N] = |M|. \quad (5.2)$$

Заметим, что рассмотренный ранее метод ДБ (§3) универсален, но если есть подозрение, что множество Ω может быть пустым, то целесообразно воспользоваться методом Данцига, так как в методе ДБ пустота множества планов может проявиться лишь в последней части алгоритма (см. пример 3.2).

Обратимся к условиям КТ (см. (4.6) в гл. 1). Для задачи (5.1) $N_1 = N$, $M_2 = M$, поэтому вектор $v[M] = \mathbb{0}$. Исключив вектор $v[M]$ из системы КТ, получим систему КТ для задачи (5.1) в виде

$$\begin{aligned} D[N, N] \cdot x[N] - A^T[N, M] \cdot u[M] - E[N, N^-] \cdot y[N] &= -c[N], \\ A[M, N] \cdot x[N] &= b[M], \end{aligned} \quad (5.3a)$$

$$x[N] \geq \mathbb{0}, \quad (5.3b)$$

$$y[N] \geq \mathbb{0}, \quad (5.3в)$$

$$x[i] \cdot y[i] = 0 \quad \forall i \in N. \quad (5.3r)$$

В силу (5.2) матрица системы (5.3a) полного ранга.

Выявим структуру дополнительного базиса. Любой базис Q системы (5.3a) имеет вид $Q = \Gamma \cup \overline{M} \cup (-J)$, где

$$\Gamma, J \subset N, \quad \overline{M} \subset M, \quad |\Gamma| + |\overline{M}| + |J| = n + m.$$

При этом $|\Gamma| \geq |M|$, что следует из линейной независимости последних m строк базисной матрицы $T[P, Q]$. Согласно (5.3r) базис является *дополнительным*, если $\Gamma \cap J = \emptyset$. Для такого базиса $|\Gamma| + |J| \leq n$, и так как $n + m = |Q| \leq n + |\overline{M}|$, то $\overline{M} = M$. Значит, дополнительный базис Q системы (5.3a) имеет вид

$$Q = \Gamma \cup M \cup (-J), \quad \Gamma \cup J = N, \quad |\Gamma| \geq |M|.$$

Как и в методе ДБ, введем еще один тип базиса.

Базис Q назовем *почти дополнительным*, если

$$Q = \Gamma \cup M \cup (-J), \quad \Gamma \cap J = \{l\},$$

так что при некотором $q \in N$ пара индексов $\{q, -q\}$ является небазисной.

Метод Данцига состоит из двух этапов: *начального* и *основного*. На начальном этапе осуществляется поиск дополнительного базиса Q_0 системы (5.3a), (5.3б). Если множество планов Ω непусто, то такой базис найдется.

На основном этапе, начиная с базиса Q_0 , по определенному правилу строится последовательность *дополнительных* базисов Q_k системы (5.3а), базисные решения на которых удовлетворяют условиям (5.3б), и $F(x_{k+1}) \leq F(x_k)$. Процесс закончится, когда либо выяснится, что система КТ несовместна, либо выполнится условие (5.3в), т. е. будет получено решение задачи.

Прежде чем переходить к детальному описанию метода, докажем вспомогательные утверждения, необходимые для пояснения метода.

5.2. Вспомогательные утверждения

Всякое изменение базиса в методе Данцига связано с уменьшением (или сохранением) значения целевой функции $F(x)$ на соответствующем базисном плане. Чтобы проследить за изменением $F(x)$ при описании метода, выведем некоторые формулы.

Как обычно, матрицу системы (5.3а) обозначим $T[P, R]$. Здесь $P = N \cup M$, $R = P \cup N^-$. Пусть $Q = \Gamma \cup M \cup (-J)$ — базис системы (5.3а) (он либо дополнительный, либо почти дополнительный), а $w[R] := (x[N], u[M], y[N])$ — базисный план, причем $w[N^-] = y[N]$. Пусть $s[Q]$ — вектор коэффициентов разложения небазисного столбца $T[P, r]$, $r \in N \cup N^-$, по базису Q :

$$T[P, r] = T[P, Q] \cdot s[Q].$$

Доопределим вектор $s[Q]$ на множество $R = N \cup M \cup (N^-)$, полагая $s[r] = -1$, а остальные небазисные компоненты — равными нулю. Тогда вектор $s[R]$ является решением однородной системы

$$\begin{aligned} D[N, N] \cdot s[N] - A^T[N, M] \cdot s[M] - E[N, N^-] \cdot s[N^-] &= \mathbb{O}, \\ A[M, N] \cdot s[N] &= \mathbb{O}. \end{aligned} \quad (5.4)$$

Введем обозначения:

$$s_+ := s_+[N] = s[N], \quad s_- := s_-[N] = s[N^-].$$

Рассмотрим луч $x(t) = x - t \cdot s_+$, $t \geq 0$. Пользуясь разложением квадратичной функции $F(x)$, имеем

$$F(x(t)) = F(x) - t \langle s_+, (Dx + c) \rangle + \frac{1}{2} t^2 \langle s_+, Ds_+ \rangle.$$

Домножим слева обе части первого равенства (5.4) на s_+ . Тогда с учетом второго равенства получим

$$\langle Ds_+, s_+ \rangle = \langle s_+, s_- \rangle \geq 0. \quad (5.5)$$

Неотрицательность правого скалярного произведения в (5.5) следует из неотрицательной определенности матрицы D . Домножив первое равенство (5.3а) на s_+ слева, получим

$$\langle s_+, (Dx + c) \rangle = \langle s_+, y \rangle.$$

Тогда

$$F(x(t)) = F(x) - t \cdot \langle s_+, y \rangle + \frac{1}{2} t^2 \langle s_+, s_- \rangle. \quad (5.6)$$

Проследим за изменением значения целевой функции при вводе индекса $l \in N$ в дополнительный базис.

Утверждение 5.1. Пусть базис Q дополнительный и пусть $s[R]$ — коэффициенты разложения по этому базису небазисного столбца $T[P, l]$, $l \in N$. Тогда $s[-l] \leq 0$ и

$$F(x - ts_+) = F(x) + t \cdot y[l] - \frac{1}{2} t^2 s[-l],$$

так что при отрицательном $y[l]$ вектор $(-s_+)$ — направление строгого убывания $F(x)$ в точке x .

Доказательство. Вычислим скалярное произведение $\langle s_+, y \rangle$, заметив, что $y[N \setminus J] = \mathbb{O}$. Так как базис Q дополнительный, то $l \in J$ и $s[J \setminus \{l\}] = \mathbb{O}$. Учитывая, что $s[l] = -1$, получаем

$$\langle s_+, y \rangle = s[J] \cdot y[J] = s[l] \cdot y[l] = -y[l].$$

Заметим, что у векторов s_+ и s_- может быть только одна общая ненулевая компонента с индексом l , поэтому согласно (5.5)

$$0 \leq \langle s_+, s_- \rangle = s[l] \cdot s[-l] = -s[-l].$$

Подставив эти выражения в (5.6), получим требуемое. ■

Теперь проследим за поведением целевой функции при вводе отрицательного индекса в почти дополнительный базис.

Утверждение 5.2. Пусть базис $Q = \Gamma \cup M \cup (-J)$ почти дополнительный с общей базисной парой $\{l, -l\}$ и небазисной $\{r, -r\}$.

Пусть $s[R]$ — вектор коэффициентов разложения по этому базису столбца $-E[P, -r]$. Тогда

$$s[-l] \cdot s[l] \geq 0, \text{ а если } s[\Gamma] \neq \emptyset, \text{ то } s[-l] \cdot s[l] > 0.$$

Кроме того,

$$F(x - ts_+) - F(x) = \frac{1}{2}t \cdot s[l] \cdot (t \cdot s[-l] - 2y[l]).$$

Доказательство. Заметим, что $s[r]$ — небазисная переменная, поэтому $s[r] = 0$. Тогда из (5.5) следует

$$0 \leq \langle s_+, s_- \rangle = s[l] \cdot s[-l].$$

Пусть теперь $s[\Gamma] \neq \emptyset$. Если допустить, что $s[l] \cdot s[-l] = 0$, то в силу неотрицательной определенности матрицы D из (5.5) получим $D[N, N] \cdot s[N] = \emptyset$. При этом из (5.4) следует, что и $A[M, N] \cdot s[N] = \emptyset$. Так как $s[N \setminus \Gamma] = \emptyset$, то полученные равенства в совокупности означают линейную зависимость базисных столбцов $T[P, \Gamma]$, что невозможно. Тем самым первая часть утверждения доказана полностью.

Очевидно, что $\langle s_+, y \rangle = s[l] \cdot y[l]$. Подставляя это в (5.6), получаем требуемое равенство для целевой функции. ■

5.3. Описание метода

В этом разделе приведем подробное описание метода Данцига с пояснениями. В разд. 5.4 будет описан алгоритм метода с использованием обратных базисных матриц.

Начальный этап. Определим базисный план исходной задачи. Воспользуемся методом Булавского (см. гл. 1, §7). Если оказалось, что Ω пусто, то задача (5.1) не имеет решения и на этом вычисления заканчиваются. Если Ω не пусто, то найдется $\Gamma_0 \subset N$, такое, что

$$|\Gamma_0| = |M|, \quad \text{rank } A[M, \Gamma_0] = |M|, \quad x_0[\Gamma_0] \geq \emptyset, \quad x_0[N \setminus \Gamma_0] = \emptyset,$$

где $x_0[\Gamma_0]$ — решение системы $A[M, \Gamma_0] \cdot x[\Gamma_0] = b[M]$. Положим $J_0 = N \setminus \Gamma_0$. Индексное множество

$$Q_0 = \Gamma_0 \cup M \cup (-J_0)$$

образует дополнительный базис системы (5.3а). Действительно, достаточно убедиться, что матрица $T[P, Q_0]$ неособенная. В результате перестановки строк из P в порядке $P' = (M, \Gamma_0, J_0)$ матрица $T[P', Q_0]$ приобретет следующую блочную структуру:

$$\begin{pmatrix} A[M, \Gamma_0] & \emptyset[M, M] & \emptyset[M, -J_0] \\ D[\Gamma_0, \Gamma_0] & -A^T[\Gamma_0, M] & \emptyset[\Gamma_0, -J_0] \\ D[J_0, \Gamma_0] & -A^T[J_0, M] & -E[J_0, -J_0] \end{pmatrix}.$$

Легко видеть, что определитель этой матрицы равен квадрату определителя матрицы $A[M, \Gamma_0]$, поэтому он отличен от нуля. Положим $C_0[\Gamma_0, M] = (A[M, \Gamma_0])^{-1}$ и найдем базисное решение системы (5.3а). Нетрудно видеть, что

$$\begin{aligned} u_0[M] &= C_0^T[M, \Gamma_0] \cdot (D[\Gamma_0, \Gamma_0] \cdot x_0[\Gamma_0] + c[\Gamma_0]), \\ y_0[J_0] &= D[J_0, \Gamma_0] \cdot x_0[\Gamma_0] - A^T[J_0, M] \cdot u_0[M]. \end{aligned}$$

Таким образом, в результате выполнения начального этапа мы либо убедимся в пустоте множества Ω , либо получим дополнительный базис Q_0 системы (5.3а) и базисный план $w_0 = (x_0, u_0, y_0)$, удовлетворяющий условию (5.3б). Полагаем $k = 0$ и переходим к основному этапу.

Основной этап.

1. *Проверка базиса Q_k на оптимальность.* Положим

$$Y_k = \{j \in J_k \mid y_k[j] < 0\}.$$

Если $Y_k = \emptyset$, то базис Q_k экстремальный и $x_k[N]$ — решение задачи (5.1). В противном случае выберем какое-либо $l_k \in Y_k$ и перейдем к п. 2.

2. *Построение нового дополнительного базиса.* Чтобы отрицательную компоненту $y_k[l_k]$ сделать нулевой, исключим индекс $(-l_k)$ из базиса Q_k . Но чтобы новый базис был дополнительным, необходимо ввести в него индекс l_k . Процесс исключения индекса $(-l_k)$ из базиса связан с ограничениями

$$x[N] \geq \emptyset, \quad -y[l_k] \geq 0, \tag{5.7}$$

поэтому при вводе в базис индекса l_k исключиться из базиса может не $(-l_k)$, а какой-либо индекс из Γ_k . В связи с этим процесс замены

в базисе индекса $(-l_k)$ на индекс l_k может произойти не сразу, а за несколько шагов, на каждом из которых какой-то индекс $j \in N$ будет заменяться индексом $(-j)$, пока наконец не исключится индекс $(-l_k)$.

Таким образом, замена индекса $(-l_k)$ на l_k сводится к решению следующей локальной задачи:

$$\begin{aligned} -y[l_k] \rightarrow \min, \\ D[N, N] \cdot x[N] - A^T[N, M] \cdot u[M] - E[N, N^-] \cdot y[N] &= -c[N], \\ A[M, N] \cdot x[N] &= b[M], \\ x[N] \geq 0, \quad -y[l_k] \geq 0, \end{aligned}$$

которую в принятых обозначениях (см. разд. 5.2) можно записать в виде

$$\begin{aligned} -w[-l_k] \rightarrow \min, \\ T[P, R] \cdot w[R] = d[P], \\ w[N] \geq 0, \quad -w[-l_k] \geq 0. \end{aligned}$$

Напомним, что $y[N] = w[N^-]$, $x[N] = w[N]$.

Решаем локальную задачу, начиная с базиса Q_k . Сначала вводим в базис столбец $T[P, l_k]$. Для этого находим его коэффициенты разложения $s_k[Q_k]$ по базису Q_k из системы

$$T[P, Q_k] \cdot s[Q_k] = T[P, l_k].$$

Согласно утверждению 5.1 вектор $(-s_k)$ является направлением строгого убывания $F(x)$ в точке x_k . Определим максимально возможный шаг в этом направлении. Так как

$$y(t)[l_k] = y_k[l_k] - t \cdot s_k[-l_k],$$

а в силу утверждения 5.1 $s_k[-l_k] \leq 0$, то максимально возможный шаг t'_k для соблюдения второго ограничения (5.7) определяется по формуле

$$t'_k = \begin{cases} \infty, & \text{если } s_k[-l_k] = 0, \\ \frac{y_k[l_k]}{s_k[-l_k]} > 0, & \text{если } s_k[-l_k] < 0. \end{cases}$$

Вместе с тем $x(t)[i] = x_k[i] - t \cdot s_k[i] \quad \forall i \in \Gamma_k$. Тогда для соблюдения первого ограничения (5.7) шаг следует вычислять по формуле

$$t''_k = \begin{cases} \infty, & \text{если } s_k[\Gamma_k] \leq 0, \\ \min\left\{\frac{x_k[i]}{s_k[i]} \mid i \in \Gamma_k, s_k[i] > 0\right\} & \text{иначе.} \end{cases}$$

Окончательно, чтобы не нарушить условия (5.7), полагаем

$$t_k = \min\{t'_k, t''_k\}.$$

Возможны три случая.

а) Шаг бесконечен, $t_k = \infty$, тогда *исходная задача неразрешима*. Действительно, весь луч $x(t) = x_k - t \cdot s_k$, $t \geq 0$, принадлежит Ω , $s_k[-l_k] = 0$, $y_k[l_k] < 0$, поэтому согласно утверждению 5.1 вдоль этого луча $F(x)$ бесконечно убывает.

б) Шаг конечен и $0 < t_k = t'_k \leq t''_k$, так что из базиса исключается индекс $(-l_k)$. Новый базис $Q_{k+1} := Q_k \setminus \{-l_k\} \cup \{l_k\}$ будет *дополнительным*. Пересчитываем базисное решение:

$$w_{k+1}[l_k] = t_k, \quad w_{k+1}[j] := w_k[j] - t_k \cdot s_k[j] \quad \forall j \in Q_k \setminus \{-l_k\}.$$

В силу утверждения 5.1 при $t_k = t'_k$ получаем

$$F(x_{k+1}) - F(x_k) = \frac{1}{2} t_k \cdot y_k[l_k] < 0.$$

Полагаем $k = k + 1$ и возвращаемся к п. 1.

в) Шаг конечен и $t_k = t''_k < t'_k$. В этом случае для построения нового дополнительного базиса требуется проделать конечное число итераций путем перебора почти дополнительных базисов $Q_{k\nu}$, $\nu = 1, 2, \dots$, пока из базиса не исключится индекс $(-l_k)$.

Обозначим через P_k множество индексов из Γ_k , на котором реализуется минимум при вычислении t''_k . Из базиса исключается любой индекс $r_1 \in P_k$. Новый базис

$$Q_{k1} := Q_k \setminus \{r_1\} \cup \{l_k\}$$

является *почти дополнительным*, так как он содержит индексы $\{l_k, -l_k\}$. При этом небазисной парой является $\{r_1, -r_1\}$. Вычисляем новое базисное решение:

$$w_{k1}[l_k] = t_k, \quad w_{k1}[j] := w_k[j] - t_k \cdot s_k[j] \quad \forall j \in Q_k \setminus \{r_1\}.$$

Согласно утверждению 5.1 и условию $t_k = t_k'' < t_k'$ получаем оценку

$$F(x_{k+1}) - F(x_k) \leq \frac{1}{2} t_k \cdot y_k[l_k] \leq 0.$$

Заметим, что $y_{k+1}[l_k]$, может быть, и увеличился, но по-прежнему остался отрицательным.

Положим $\nu = 1$ и продолжим усилия по исключению индекса $(-l_k)$ из базиса. Раз из базиса исключился индекс $r_\nu \in N$, то чтобы не испортить почти-дополнительность нового базиса, будем включать в базис отрицательный индекс $(-r_\nu)$. Найдем $s_{k\nu}[Q_{k\nu}]$ — коэффициенты разложения вводимого в базис вектора $-E[P, -r_\nu]$ по базису $Q_{k\nu}$, решая систему

$$T[P, Q_{k\nu}] \cdot s[Q_{k\nu}] = -E[P, -r_\nu]. \quad (5.8)$$

Ниже будет показано (см. утверждение 5.3), что

$$s_{k\nu}[-l_k] < 0. \quad (5.9)$$

Согласно знаковому ограничению (5.7) и неравенству (5.9) индекс, подлежащий исключению из базиса, определяется по следующему правилу:

$$t_{k\nu} = \min\{t_{k\nu}', t_{k\nu}''\}, \quad t_{k\nu}' = \frac{y_{k\nu}[l_k]}{s_{k\nu}[-l_k]} > 0,$$

$$t_{k\nu}'' = \begin{cases} \infty, & \text{если } s_{k\nu}[\Gamma_{k\nu}] \leq \mathbb{O}, \\ \min\left\{\frac{x_{k\nu}[i]}{s_{k\nu}[i]} \mid i \in \Gamma_{k\nu}, s_{k\nu}[i] > 0\right\} & \text{иначе.} \end{cases}$$

Очевидно, что шаг $t_{k\nu}$ конечен.

Если $t_{k\nu} = t_{k\nu}' \leq t_{k\nu}''$, то индекс $(-l_k)$ исключится из базиса, так что базис $Q_{k+1} = Q_{k\nu} \setminus \{-l_k\} \cup \{-r_\nu\}$ будет *дополнительным*. Пересчитываем базисное решение:

$$w_{k+1}[-r_\nu] = t_{k\nu}, \quad w_{k+1}[j] := w_{k\nu}[j] - t_{k\nu} \cdot s_{k\nu}[j] \quad \forall j \in Q_{k\nu} \setminus \{-l_k\}.$$

Так как $t_{k\nu} \cdot s_{k\nu}[-l_k] = y_{k\nu}[l_k]$, то в силу утверждения 5.2 имеем

$$F(x_{k+1}) - F(x_{k\nu}) = -\frac{1}{2} t_{k\nu} \cdot s_{k\nu}[l_k] \cdot y_{k\nu}[l_k] \leq 0,$$

причем равенство нулю возможно только если $s_{k\nu}[l_k] = 0$. Полагаем теперь $k = k + 1$ и переходим к п. 1.

Если $t_{k\nu} = t_{k\nu}'' < t_{k\nu}'$, то из базиса исключается индекс q_ν из $\Gamma_{k\nu}$, на котором реализовался минимум при вычислении $t_{k\nu}''$. Получаем новый *почти дополнительный базис*

$$Q_{k, \nu+1} = Q_{k\nu} \setminus \{q_\nu\} \cup \{-r_\nu\}$$

и находим базисное решение $w_{k, \nu+1}$. При этом в силу утверждения 5.2 и неравенства (5.9) имеем

$$F(x_{k, \nu+1}) - F(x_{k\nu}) = \frac{1}{2} t_{k\nu} \cdot s_{k\nu}[l_k] \cdot s_{k\nu}[-l_k] \cdot (-2t_{k\nu}' + t_{k\nu}) \leq 0.$$

Заметим, что $q_\nu \neq l_k$. Действительно, из (5.9) и утверждения 5.2 следует, что $s_{k\nu}[l_k] \leq 0$, т. е. индекс l_k из базиса $Q_{k\nu}$ исключиться не может.

Полагаем $r_{\nu+1} = q_\nu$, $\nu = \nu + 1$ и переходим к определению $s_{k\nu}$ по формуле (5.8).

На этом описание метода Данцига заканчивается.

Заметим, что итеративный процесс п. в) конечен, так как в нем происходит пересылка индексов из конечного множества $\Gamma_{k\nu}$ в множество $J_{k\nu}$.

5.4. Метод Данцига с обратными матрицами

На каждой итерации метода Данцига приходится решать систему линейных уравнений с базисной матрицей для нахождения $s[Q_k]$ — коэффициентов разложения вводимого в базис столбца по текущему базису Q_k :

$$T[P, Q_k] \cdot s[Q_k] = T[P, l_k].$$

Обратим внимание, что каждый следующий базис Q_{k+1} отличается от предыдущего Q_k одним индексом. Значит, если известна начальная обратная базисная матрица $B_0[Q_0, P]$, то все следующие обратные матрицы легко пересчитываются через предыдущую по формулам (8.3) гл. 1. Пусть $Q_{k+1} = Q_k \setminus \{q_k\} \cup \{l_k\}$. Напомним формулы пересчета обратной базисной матрицы и базисного плана:

$$\left. \begin{aligned} B_{k+1}[l_k, P] &= \frac{1}{s_k[q_k]} B_k[q_k, P], \\ B_{k+1}[i, P] &= B_k[i, P] - s_k[i] \cdot B_{k+1}[l_k, P], \quad i \in Q_k \setminus \{q_k\}, \\ w_{k+1}[l_k] &= t_k, \quad w_{k+1}[i] = w_k[i] - s_k[i] \cdot t_k, \quad i \in Q_k \setminus \{q_k\}. \end{aligned} \right\} \quad (5.10)$$

Начальный этап метода Данцига с обратными матрицами та-
кой же, как в разд. 5.3. Дополнительно нужно лишь построить
матрицу $B_0[Q_0, P']$ — обратную к базисной матрице $(T[P', Q_0])^{-1}$.
Легко проверить, что она имеет вид

$$\begin{pmatrix} C_0[\Gamma_0, M] & \mathbb{O}[\Gamma_0, \Gamma_0] & \mathbb{O}[\Gamma_0, J_0] \\ K[M, M] & -C_0^T[M, \Gamma_0] & \mathbb{O}[M, J_0] \\ K_1[J_0, M] & K_2[J_0, \Gamma_0] & -E[J_0, J_0] \end{pmatrix},$$

где

$$\begin{aligned} C_0[\Gamma_0, M] &= (A[M, \Gamma_0])^{-1}, \\ K[M, M] &= C_0^T[M, \Gamma_0] \cdot D[\Gamma_0, \Gamma_0] \cdot C_0[\Gamma_0, M], \\ K_1[J_0, M] &= D[J_0, \Gamma_0] \cdot C_0[\Gamma_0, M] - A^T[J_0, M] \cdot K[M, M], \\ K_2[J_0, \Gamma_0] &= A^T[J_0, M] \cdot C_0^T[M, \Gamma_0]. \end{aligned}$$

Сделаем краткое описание одного шага **основного этапа** ме-
тода Данцига с обратными матрицами.

Имеем: дополнительный базис $Q_k = \Gamma_k \cup M \cup (-J_k)$, базисный
план $w_k[Q_k] = (x_k[\Gamma_k], u_k[M], y_k[J_k])$ и обратную базисную матри-
цу $B_k[Q_k, P]$.

Если $y_k \geq \mathbb{O}$, то x_k — решение исходной задачи. На этом вычис-
ления заканчиваются.

В противном случае выбираем индекс l_k , для которого

$$y_k[l_k] < 0.$$

Вычисляем $s_k[Q_k] = B_k[Q_k, P] \cdot T[P, l_k]$ и шаг $t_k = \min\{t'_k, t''_k\}$, где

$$t'_k = \begin{cases} \infty, & \text{если } s_k[-l_k] = 0, \\ \frac{y_k[l_k]}{s_k[-l_k]} > 0, & \text{если } s_k[-l_k] < 0, \end{cases}$$

$$t''_k = \begin{cases} \infty, & \text{если } s_k[\Gamma_k] \leq \mathbb{O}, \\ \min\{\frac{x_k[i]}{s_k[i]} \mid i \in \Gamma_k, s_k[i] > 0\} & \text{иначе.} \end{cases}$$

Возможны три случая.

1. Если $t_k = \infty$, то задача неразрешима и вычисления заканчи-
ваются.

2. Если $t_k = t'_k \leq t''_k$, то получаем *новый дополнительный базис*

$$Q_{k+1} = Q_k \setminus \{-l_k\} \cup \{l_k\}.$$

Пересчитываем обратную базисную матрицу и базисный план по
формулам (5.10), полагая $q_k = -l_k$. На этом один шаг основного
этапа заканчивается.

3. Если $t_k = t''_k < t'_k$ и r_1 — индекс, на котором достигается мини-
мум при вычислении t''_k , то получаем почти дополнительный базис

$$Q_{k1} = Q_k \setminus \{r_1\} \cup \{l_k\},$$

пересчитываем обратную базисную матрицу и базисный план по
формулам (5.10), подставляя вместо q_k индекс r_1 , а вместо индекса
($k+1$) индекс $k1$. Далее полагаем $\nu = 1$ и переходим к п. 3'.

3'. Продолжаем итерации для исключения $(-l_k)$ из базиса. Так
как в базис вводится отрицательный столбец-орт с индексом r_ν , то
коэффициенты разложения его по базису известны, а именно

$$s_{k\nu}[Q_{k\nu}] = -B_{k\nu}[Q_{k\nu}, r_\nu].$$

Находим шаг $t_{k\nu} = \min\{t'_{k\nu}, t''_{k\nu}\}$, где

$$t'_{k\nu} = \frac{y_{k\nu}[l_k]}{s_{k\nu}[-l_k]} > 0,$$

$$t''_{k\nu} = \begin{cases} \infty, & \text{если } s_{k\nu}[\Gamma_{k\nu}] \leq \mathbb{O}, \\ \min\{\frac{x_{k\nu}[i]}{s_{k\nu}[i]} \mid i \in \Gamma_{k\nu}, s_{k\nu}[i] > 0\} & \text{в противном случае.} \end{cases}$$

Если $t_{k\nu} = t''_{k\nu} < t'_{k\nu}$, то из базиса исключается индекс q_ν из $\Gamma_{k\nu}$ и
получается почти дополнительный базис

$$Q_{k\nu+1} = Q_{k\nu} \setminus \{q_\nu\} \cup \{-r_\nu\}.$$

Пересчитываем базисную матрицу и базисный план, используя
формулы (5.10). Далее полагаем $r_{\nu+1} = q_\nu$, $\nu = \nu + 1$ и возвра-
щаемся в начало п. 3'.

Если $t_{k\nu} = t'_{k\nu} \leq t''_{k\nu}$, то получаем *новый дополнительный базис*

$$Q_{k+1} = Q_{k\nu} \setminus \{-l_k\} \cup \{-r_\nu\}.$$

Используя формулы (5.10), пересчитываем базисную матрицу и ба-
зисный план. На этом один шаг основного этапа заканчивается.

5.5. Обоснование метода

Установим справедливость неравенства (5.9). Предварительно докажем лемму.

Лемма 5.1. Пусть $Q = \Gamma \cup M \cup (-J)$ — почти дополнительный базис с общим базисным индексом $l \in N$ и небазисным индексом $r \in N$. Пусть $s[Q]$ и $\omega[Q]$ — векторы коэффициентов разложения небазисных столбцов $T[P, -r]$ и $T[P, r]$ по базису Q . Тогда имеет место неравенство

$$s[l] \cdot \omega[l] \leq 0.$$

Доказательство. Положим $B[Q, P] = (T[P, Q])^{-1}$ и введем в рассмотрение матрицу коэффициентов разложения всех столбцов матрицы системы КТ по базису Q :

$$Z[Q, R] = B[Q, P] \cdot T[P, R]. \quad (5.11)$$

Пользуясь структурой матрицы T , перепишем равенство (5.11) для строки $i \in Q$ в виде⁶

$$Z[i, N] = B[i, N] \cdot D[N, N] + B[i, M] \cdot A[M, N], \quad (5.12)$$

$$Z[i, N^-] = -B[i, N] \cdot E[N, N^-] = -B[i, N], \quad (5.13)$$

$$Z[i, M] = -B[i, N] \cdot A^T[N, M].$$

Из последнего равенства следует, что

$$Z[\Gamma, M] = \mathbb{O}, \quad (5.14)$$

так как $Z[i, M]$ для $i \in \Gamma$ есть произведение строки обратной базисной матрицы на базисные столбцы прямой матрицы с индексами из M . Из таких же соображений следуют равенства

$$Z[l, -J] = \mathbb{O}, \quad Z[l, \Gamma \setminus \{l\}] = \mathbb{O}. \quad (5.15)$$

Из (5.12)–(5.14) получаем

$$\eta := Z[l, N] \cdot Z[l, N^-] = -B[l, N] \cdot D[N, N] \cdot B[l, N].$$

Так как матрица D неотрицательно определена, то $\eta \leq 0$.

⁶Равенство $Z[i, N^-] = -B[i, N]$ означает, что $Z[i, -j] = -B[i, j]$ для всех $j \in N$.

Теперь вычислим η непосредственно, пользуясь почти-дополнительностью базиса Q . Напомним, что $N = \Gamma \cup J \cup \{r\}$, $l \in \Gamma$, $l \in J$. Учитывая равенства (5.15), получаем

$$\eta = Z[l, r] \cdot Z[l, -r] = \omega[l] \cdot s[l]. \quad \blacksquare$$

Рассмотрим два произвольных базиса системы (5.3а) вида

$$Q_k = \Gamma_k \cup M \cup (-J_k), \quad Q_p = \Gamma_p \cup M \cup (-J_p).$$

Пусть $s_k[Q_k]$ и $s_p[Q_p]$ — векторы коэффициентов разложения небазисных столбцов $T[P, r_k]$ по базису Q_k и $T[P, r_p]$ по базису Q_p .

Доопределим обычным способом векторы $s_k[Q_k]$ и $s_p[Q_p]$ на множество $R = N \cup M \cup (N^-)$. Как мы уже знаем, векторы $s_k[R]$, $s_p[R]$ удовлетворяют однородной системе (5.4). В частности,

$$\begin{aligned} D[N, N] \cdot s_k[N] - A^T[N, M] \cdot s_k[M] - s_k[N^-] &= \mathbb{O}, \\ A[M, N] \cdot s_p[N] &= \mathbb{O}. \end{aligned}$$

Домножим первое равенство слева на $s_p[N]$ и учтем второе. Тогда получим $s_p[N] \cdot s_k[N^-] = s_p[N] \cdot D[N, N] \cdot s_k[N]$. Аналогично, поменяв местами $s_k[N]$ и $s_p[N]$, получим

$$s_k[N] \cdot s_p[N^-] = s_k[N] \cdot D[N, N] \cdot s_p[N].$$

Так как матрица D симметрична, то

$$s_p[N] \cdot s_k[N^-] = s_k[N] \cdot s_p[N^-]. \quad (5.16)$$

Утверждение 5.3. В п. в) основного этапа описания метода

$$s_{k\nu}[-l_k] < 0 \quad \forall \nu \geq 1.$$

Доказательство. Как обычно, через $s_{k\nu}$ обозначим коэффициенты разложения по базису $Q_{k\nu}$ вводимого в базис столбца $T[P, -r_\nu]$, а через $\omega_{k\nu}$ — коэффициенты разложения по этому же базису столбца $T[P, r_\nu]$. Так как этот столбец на предыдущем шаге исключился из базиса, то его коэффициенты разложения легко получаются из $s_{k\nu-1}[N]$, а именно:

$$\omega_{k\nu}[R] = -\frac{1}{s_{k\nu-1}[r_\nu]} \cdot s_{k\nu-1}[R]. \quad (5.17)$$

Заметим, что согласно алгоритму $s_{k\nu-1}[r_\nu] > 0$. В силу леммы 5.1 имеем $s_{k\nu}[l_k] \cdot \omega_{k\nu}[l_k] \leq 0$. Подставляя сюда выражение из (5.17), получаем

$$s_{k\nu}[l_k] \cdot s_{k\nu-1}[l_k] \geq 0. \quad (5.18)$$

Заметим, что из равенства (5.16) следует

$$s_{k\nu}[N] \cdot s_{k\nu-1}[N^-] = s_{k\nu}[N^-] \cdot s_{k\nu-1}[N]. \quad (5.19)$$

Доказательство леммы будем проводить по индукции.

Пусть $\nu = 1$. Тогда предыдущий базис $Q_{k\nu-1}$ совпадает с дополнительным базисом Q_k . Пользуясь соотношением (5.19), получаем равенство

$$s_{k1}[N] \cdot s_k[N^-] = s_{k1}[N^-] \cdot s_k[N].$$

Заметим, что в этих произведениях одновременно ненулевыми могут быть компоненты только с индексами l_k , r_1 , так что последнее равенство можно переписать в виде

$$\begin{aligned} s_{k1}[l_k] s_k[-l_k] + s_{k1}[r_1] s_k[-r_1] &= \\ &= s_{k1}[-l_k] s_k[l_k] + s_{k1}[-r_1] s_k[r_1]. \end{aligned} \quad (5.20)$$

Напомним, что индекс l_k вводится в базис Q_k , а индекс r_1 исключается из базиса Q_k , поэтому $s_k[l_k] = -1$, $s_k[r_1] > 0$, и, так как $(-r_1)$ не принадлежит базису Q_k , то $s_k[-r_1] = 0$. Кроме того, согласно алгоритму $s_k[-l_k] \leq 0$, $s_{k1}[-r_1] = -1$. Подставляя эти значения в (5.20), получаем

$$s_k[r_1] + s_{k1}[-l_k] = -s_{k1}[l_k] \cdot s_k[-l_k].$$

В силу (5.18) $s_{k1}[l_k] \cdot s_k[l_k] = -s_{k1}[l_k] \geq 0$, тогда из предыдущего равенства и следует, что $s_{k1}[-l_k] < 0$.

Заметим, что если процесс построения дополнительного базиса не закончился, т.е. из базиса исключился индекс $r_2 \in N$, то $s_{k1}[N] \neq 0$ и из утверждения 5.2 следует $s_{k1}[-l_k] \cdot s_{k1}[l_k] > 0$, значит, $s_{k1}[l_k] < 0$.

Пусть для $(\nu-1)$ итерации утверждение леммы верно. Докажем его для итерации ν .

Имеем

$$s_{k\nu-1}[-l_k] < 0, \quad s_{k\nu-1}[l_k] < 0. \quad (5.21)$$

Из (5.18) следует, что $s_{k\nu}[l_k] \leq 0$. Аналогичным образом из (5.19) получаем равенство

$$\begin{aligned} s_{k\nu}[l_k] \cdot s_{k\nu-1}[-l_k] + s_{k\nu}[r_\nu] \cdot s_{k\nu-1}[-r_\nu] &= \\ &= s_{k\nu}[-l_k] \cdot s_{k\nu-1}[l_k] + s_{k\nu}[-r_\nu] \cdot s_{k\nu-1}[r_\nu]. \end{aligned}$$

Замечаем, что $s_{k\nu}[-r_\nu] = -1$, $s_{k\nu-1}[-r_\nu] = 0$, поэтому последнее равенство переписывается в виде

$$s_{k\nu}[-l_k] \cdot s_{k\nu-1}[l_k] = s_{k\nu-1}[r_\nu] + s_{k\nu}[l_k] \cdot s_{k\nu-1}[-l_k].$$

Отсюда и из неравенств (2.1) с учетом, что $s_{k\nu-1}[r_\nu] > 0$, получаем, что $s_{k\nu}[-l_k] < 0$. Кроме того, если процесс построения дополнительного базиса не закончился, т.е. из базиса исключился индекс $q_\nu \in N$, то $s_{k\nu}[N] \neq 0$ и из утверждения 5.2 имеем $s_{k\nu}[-l_k] \cdot s_{k\nu}[l_k] > 0$, значит, $s_{k\nu}[l_k] < 0$. Утверждение доказано. ■

5.6. О конечности метода

Описанный алгоритм строит последовательность дополнительных базисов Q_k системы (5.3а) и соответствующих им векторов $x_k \in \Omega$, на которых

$$F(x_{k+1}) \leq F(x_k).$$

Если не произойдет повторения базиса, то описанный алгоритм конечен. Теоретически повторение базиса возможно, так как целевая функция убывает не строго. Можно воспользоваться правилом наименьшего индекса, а именно в п. 1 метода Данцига выбирать *минимальный* индекс из множества Y_k , т.е.

$$l_k = \min\{j | j \in Y_k\}.$$

Кроме того, каждый раз при вычислении шага, если неоднозначен выбор индекса, на котором реализуется минимум, следует выбирать *минимальный* индекс из кандидатов. Метод Данцига с применением правила наименьшего индекса конечен [7].

5.7. Метод Данцига в линейном случае

Покажем, что описанный метод Данцига при $D = \mathbb{O}$ полностью совпадает с известным симплекс-методом (см. §1).

Обратимся к системе (5.3а) при $D = \mathbb{O}$ и выясним вид матрицы системы на дополнительном базисе $Q_k = \Gamma_k \cup M \cup (-J_k)$. Напомним, что $N = \Gamma_k \cup J_k$, $\Gamma_k \cap J_k = \emptyset$. Переставив строки матрицы $T[P, Q_k]$ в порядке $P' = \Gamma_k \cup J_k \cup M$, получим матрицу следующей структуры:

$$T[P', Q_k] = \begin{pmatrix} \mathbb{O} & -A^T[\Gamma_k, M] & \mathbb{O} \\ \mathbb{O} & -A^T[J_k, M] & -E[J_k, -J_k] \\ A[M, \Gamma_k] & \mathbb{O} & \mathbb{O} \end{pmatrix}. \quad (5.22)$$

Так как базисная матрица $T[P, Q_k]$ не вырождена, то из ее вида следует, что матрица $A[M, \Gamma_k]$ квадратная и невырожденная.

Таким образом, на дополнительном базисе система

$$T[P', Q_k] \cdot w[Q_k] = d[P']$$

распадается на две независимые системы с одной и той же невырожденной матрицей $A[M, \Gamma_k]$:

$$A^T[\Gamma_k, M] \cdot u[M] = c[\Gamma_k],$$

$$A[M, \Gamma_k] \cdot x[\Gamma_k] = b[M]$$

и равенство для $y[J_k]$:

$$y[J_k] = c[J_k] - A^T[J_k, M] \cdot u[M].$$

Нетрудно видеть, что Γ_k — базис матрицы $A[M, N]$, а вектор $y[J_k]$ совпадает с оценками небазисных столбцов в симплекс-методе. Так что в п. 1 метода Данцига определяется индекс с отрицательной оценкой, который требуется ввести в следующий базис.

Обратимся к п. 2 метода Данцига. Ясно, что столбцы матрицы $T[P, l_k]$, $l_k \in N$, раскладываются только по первым n столбцам базисной матрицы (5.22), так что вектор $s_k[\Gamma_k]$ в п. 2 получается из системы

$$A[M, \Gamma_k] \cdot s_k[\Gamma_k] = A[M, l_k],$$

а $s_k[M \cup N^-] = \mathbb{O}$. В частности, $s_k[-l_k] = 0$, поэтому в п. 2 реализуются только случаи а) (нет решения) или в), так что

$$t_k = \begin{cases} \infty, & \text{если } s_k[\Gamma_k] \leq \mathbb{O}, \\ \min \left\{ \frac{x_k[i]}{s_k[i]} \mid i \in \Gamma_k, s_k[i] > 0 \right\} & \text{иначе.} \end{cases}$$

Пусть реализуется п. в) и $t_k = \frac{x_k[r_1]}{s_k[r_1]}$. Тогда

$$\Gamma_{k1} = \Gamma_k \cup \{l_k\} \setminus \{r_1\}, \quad J_{k1} = J_k, \\ x_{k1} = x_k - t_k \cdot s_k, \quad y_{k1} = y_k,$$

а индексное множество $Q_{k1} = \Gamma_{k1} \cup M \cup (-J_k)$ является почти дополнительным базисом с общей парой индексов $\{l_k, -l_k\}$. Покажем, что для построения дополнительного базиса требуется еще лишь одна итерация.

Действительно, в базис должен вводиться отрицательный индекс $(-r_1)$, которому соответствует вектор-орт $-e_{r_1}[P]$. Очевидно, что коэффициенты разложения этого орта по базису Q_{k1} таковы, что $s_{k1}[\Gamma_{k1}] = \mathbb{O}$, т. е. $t'_{k1} = \infty$. Значит, $t_{k1} = t'_k = y_{k1}[l_k]/s_{k1}[-l_k] > 0$ и из базиса исключается индекс $(-l_k)$. Очевидно, что

$$\Gamma_{k+1} = \Gamma_{k1} = \Gamma_k \cup \{l_k\} \setminus \{r_1\}.$$

Практически п. в) сводится к построению Γ_{k1} и пересчету оценок y_{k+1} .

Проследив теперь за изложенным здесь алгоритмом метода Данцига, видим, что он полностью совпадает с традиционным симплекс-методом.

5.8. Численные примеры

Пример 5.1. Методом Данцига с использованием обратных матриц решить следующую задачу:

$$\frac{1}{2}x[1]^2 + \frac{1}{2}x[2]^2 + x[1] - 2x[2] \rightarrow \min, \\ x[1] + x[2] = 1, \\ x[1] \geq 0, \quad x[2] \geq 0.$$

Здесь $N = \{1, 2\}$, $M = \{3\}$. КТ-таблица имеет вид

1	2	3	-1	-2	$d[P]$
1	0	-1	-1	0	-1
0	1	-1	0	-1	2
1	1	0	0	0	1
+	+	*	+	+	

Начальный базисный план очевиден: $x_0[1] = 1$, $x_0[2] = 0$. Его базис $\Gamma_0 = \{1\}$. Полагаем $J_0 = \{2\}$, тогда исходный дополнительный базис $Q_0 = \{1, 3, -2\}$. Согласно (5.7) построим матрицу B_0 , замечая, что $C_0 = (1)$:

$$B_0 = \begin{pmatrix} 0 & 0 & 1 \\ -1 & 0 & 1 \\ 1 & -1 & -1 \end{pmatrix}.$$

Находим базисное решение $w_0 = B_0 d = (1, 2, -4)$.

Как и в примерах на метод дополнительного базиса, вычисления удобно записывать и проводить с помощью таблиц. Напомним, что каждой итерации метода соответствует своя таблица, которая заполняется по следующей схеме:

k	l_k	$T^T[l_k, P]$		
	Q_k	$B_k[Q_k, P]$	$w_k[Q_k]$	$s_k[Q_k]$

Здесь k — номер итерации; l_k — вводимый в базис индекс; $T[P, l_k]$ — столбец с индексом l_k ; Q_k — текущий базис; $B_k[Q_k, P]$ — обратная базисная матрица; $w_k[Q_k]$ — текущий базисный план, равный $B_k[Q_k, P] \cdot d[P]$ (все небазисные компоненты равны нулю); $s_k[Q_k]$ — коэффициенты разложения $T[P, l_k]$ по базису Q_k , так что $s_k[Q_k] = B_k[Q_k, P] \cdot T[P, l_k]$. Звездочкой в первом столбце таблиц обозначен индекс из $M = \{3\}$, который из базиса не выводится.

Базисный план пересчитывается по формуле

$$w_{k+1}[l_k] = t_k, \quad w_{k+1}[i] = w_k[i] - t_k \cdot s_k[i], \quad i \in Q_k \setminus \{l_k\}. \quad (5.23)$$

Так как $w_0[-2] = -4$ отрицательна, то вводим в базис индекс 2 и пытаемся исключить (-2) :

0	2	0	1	1		
	1	0	0	1	1	1
*	3	-1	0	1	2	1
	-2	1	-1	-1	-4	-2

Согласно алгоритму из базиса исключается не (-2) , а индекс 1. Вводу в базис подлежит (-1) :

01	-1	-1	0	0		
	2	0	0	1	1	0
*	3	-1	0	0	1	1
	-2	1	-1	1	-2	-1

Теперь из базиса исключается индекс (-2) , так что следующий базис $Q_1 = \{2, 3, -1\}$. Этот базис экстремальный, так как базисное решение $w_1[Q_1] = (1, -1, 2)$ удовлетворяет условиям КТ. Очевидно, что решением задачи является вектор $x_1 = (0, 1)$.

Пример 5.2. Методом Данцига (без использования обратных матриц) решить следующую задачу:

$$\begin{aligned} \frac{1}{2}(x[1]^2 + x[2]^2 + x[3]^2) + x[3] &\rightarrow \min, \\ x[1] + 2x[2] + x[3] &\geq 2, \\ x[1] - x[2] &= 1, \\ x[1] \geq 0, \quad x[2] &\geq 0. \end{aligned}$$

Сначала приведем задачу к первой канонической форме.

Заменим переменную $x[3]$ на разность $x[3] - x[4]$, где $x[3] \geq 0$, $x[4] \geq 0$, и в первое ограничение введем искусственную неотрицательную переменную $x[5]$. Тогда матрица квадратичной формы станет матрицей пятого порядка и преобразуется в матрицу (все нулевые компоненты, кроме диагональных, не выписаны)

$$D = \begin{pmatrix} 1 & & & & \\ & 1 & & & \\ & & 1 & -1 & \\ & & -1 & 1 & \\ & & & & 0 \end{pmatrix}, \quad c = \begin{pmatrix} 0 \\ 0 \\ 1 \\ -1 \\ 0 \end{pmatrix}.$$

Задача принимает канонический вид:

$$\begin{aligned} \frac{1}{2}\langle Dx, x \rangle + \langle c, x \rangle &\rightarrow \min, \\ x[1] + 2x[2] + x[3] - x[4] - x[5] &= 2, \\ x[1] - x[2] &= 1, \\ x[i] \geq 0, \quad i &= 1, 2, \dots, 5. \end{aligned}$$

Запишем КТ-таблицу. Заметим, что все переменные, кроме $u[1]$ и $u[2]$, должны быть неотрицательными, поэтому последнюю знаковую строку таблицы опустим:

1	2	3	4	5	6	7	-1	-2	-3	-4	-5	d
1	0	0	0	0	-1	-1	-1	0	0	0	0	0
0	1	0	0	0	-2	1	0	-1	0	0	0	0
0	0	1	-1	0	-1	0	0	0	-1	0	0	-1
0	0	-1	1	0	1	0	0	0	0	-1	0	1
0	0	0	0	0	1	0	0	0	0	0	-1	0
1	2	1	-1	-1	0	0	0	0	0	0	0	2
1	-1	0	0	0	0	0	0	0	0	0	0	1

Нетрудно подобрать начальный базисный план: $x_0 = (1, 0, 1, 0, 0)$ с базисом $\Gamma_0 = \{1, 3\}$. Полагая $J_0 = \{2, 4, 5\}$, в качестве начального дополнительного базиса системы КТ берем

$$Q_0 = \{1, 3, 6, 7, -2, -4, -5\}.$$

Из КТ-таблицы находим базисный план $w_0 = (1, 1, 2, -1, -5, 0, 2)$. Так как $y[2] = -5 < 0$, то базис Q_0 не экстремальный.

Далее на каждой k -й итерации заполняется таблица следующего вида:

k	l_k		t_k
Γ_k	$w_k[\Gamma_k]$	$s_k[\Gamma_k]$	t_k''
$-J_k$	$w_k[-J_k]$	$s_k[-J_k]$	t_k'

Здесь k — номер итерации; l_k — вводимый в базис индекс; $Q_k = \Gamma_k \cup (-J_k)$ — текущий базис; $w_k[Q_k]$ — текущий базисный план (все небазисные компоненты равны нулю); $s_k[Q_k]$ — коэффициенты разложения $T[P, l_k]$ по базису Q_k , которые находятся из системы

$$T[P, Q_k] \cdot s[Q_k] = T[P, l_k].$$

Базисный план пересчитывается по формуле (5.23).

Сначала, начиная с Q_0 , за две итерации строится дополнительный базис Q_1 :

0	2		1/3
1	1	-1	
3	1	3	1/3
-2	-5	-11	5/11
-4	0	0	
-5	2	3	

01	-3		4/9
1	4/3	0	
2	1/3	0	∞
-2	-4/3	-3	4/9
-4	0	1	
-5	1	1	

Теперь за одну итерацию из Q_1 находится Q_2 :

1	4		4/11
1	4/3	-1/3	
2	1/3	-1/3	∞
-3	4/9	11/9	
-4	-4/9	-11/9	4/11
-5	5/9	-2/9	

Отсюда, пользуясь формулой (5.24), получаем $\Gamma_2 = \{1, 2, 4\}$, $J_2 = \{3, 5\}$,

$$w_2[\Gamma_2] = \left(\frac{16}{11}, \frac{5}{11}, \frac{4}{11} \right), \quad w_2[-J_2] = \left(0, \frac{7}{11} \right).$$

Нетрудно видеть, что базис Q_2 экстремальный и решением исходной задачи является вектор $x = \left(\frac{16}{11}, \frac{5}{11}, 0, \frac{4}{11} \right)$.

МЕТОД ПЕРЕБОРА ГРАНЕЙ

Описанный в гл. 2 метод дополнительного базиса (ДБ) имеет ряд достоинств, в частности он годится для любой задачи выпуклого квадратичного программирования без перевода ее к какой-либо стандартной форме и не требует автономного построения начального плана (в некотором смысле это построение осуществляется на первом этапе метода). Однако на каждом шаге метода ДБ (а также и других базисных методов) приходится решать систему порядка $p = n + m$ (здесь n — число переменных, а m — число ограничений). Поэтому при больших значениях параметра p методом пользоваться затруднительно.

Если задача имеет много ограничений (m) и мало переменных (n), то удобно воспользоваться методом перебора граней (ПГ). Забегая вперед, скажем, что в методе ПГ на каждой итерации решается система, порядок которой меняется, но не превосходит $2n$, т. е. не зависит от числа ограничений.

Разумеется, выбор метода для решения конкретной задачи ВПК определяется ее параметрами и имеющимися техническими возможностями.

§1. Идея метода перебора граней

Рассмотрим задачу с ограничениями-неравенствами:

$$F(x) := \frac{1}{2} \langle Dx, x \rangle + \langle c, x \rangle \rightarrow \min_{x \in \Omega}, \quad (1.1)$$

где

$$\Omega = \left\{ x[N] \mid A[M, N] \cdot x[N] \geq b[M] \right\} \quad (1.2)$$

и матрица D симметрична и неотрицательно определена.

Предположим, что Ω непусто, и напомним представление множества Ω в виде объединения граней.

Грань — множество точек $x \in \Omega$, у которых активны одни и те же ограничения. Формально это определение выглядит следующим образом.

Определение. Пусть $J \subseteq M$. Гранью Ω называется множество

$$\Omega(J) = \left\{ x[N] \mid \begin{array}{l} A[J, N] \cdot x[N] = b[J], \\ A[M \setminus J, N] \cdot x[N] > b[M \setminus J] \end{array} \right\}.$$

Заметим, что индексное множество J может быть пустым, тогда грань совпадает с внутренностью множества Ω . Сама грань тоже может быть пустой. Очевидно, что непустые грани с разными индексными множествами J не пересекаются. Граней конечное число и $\Omega(J) \subseteq \Omega$. Нетрудно видеть, что любая точка x из Ω принадлежит единственной грани $\Omega(I(x))$, где

$$I(x) := \left\{ i \in M \mid A[i, N] \cdot x[N] = b[i] \right\}.$$

Из сказанного следует, что множество Ω есть объединение всех его граней:

$$\Omega = \bigcup_{J \subseteq M} \Omega(J).$$

Важную роль в дальнейшем будет играть *аффинная оболочка грани*, т. е. множество

$$\omega(J) := \left\{ x[N] \mid A[J, N] \cdot x[N] = b[J] \right\}.$$

Очевидно, что если $J = \emptyset$, то аффинная оболочка грани совпадает со всем пространством $\mathbb{R}^{|N|}$.

Основной характеристикой грани является ее размерность. Размерность грани определяется как размерность ее аффинной оболочки и равна $|N| - \text{rank } A[J, N]$.

Грань $\Omega(J)$ нулевой размерности состоит из одной точки x^* и при этом $J = I(x^*)$. Эта точка (и сама грань) называется *вершиной* Ω . Вершина называется *невыврожденной*, если $|I(x^*)| = |N|$. Если $|I(x^*)| > |N|$, то вершина называется *вырожденной*. Любое подмножество $J \subseteq I(x^*)$, такое, что

$$|J| = |N|, \quad \text{rank } A[J, N] = |N|,$$

называется *строчным базисом* вершины x^* .⁷

⁷Подробнее о гранях можно посмотреть в работе [3].

Введем понятия *квазистационарной грани* и *квазистационарной точки* задачи (1.1). Для этого рассмотрим задачу минимизации квадратичной функции $F(x)$ на аффинной оболочке грани $\Omega(J)$:

$$F(x) := \frac{1}{2}(Dx, x) + \langle c, x \rangle \rightarrow \inf, \quad (1.3)$$

$$A[J, N] \cdot x[N] = b[J].$$

Определение. Грань $\Omega(J)$ называется *квазистационарной*, если точка минимума функции $F(x)$ на аффинной оболочке этой грани существует и принадлежит Ω . Сама точка минимума при этом называется *квазистационарной точкой*, а значение целевой функции $F(x)$ в квазистационарной точке — *оценкой грани*.

Нетрудно видеть, что любая вершина множества Ω является квазистационарной точкой и одновременно квазистационарной гранью.

Пример. Рассмотрим задачу минимизации

$$F(x) = (x[1] - 2)^2 + x[2]^2 \rightarrow \min_{x \in \Omega},$$

$$\Omega = \{x \in \mathbb{R}^2 \mid |x[1]| \leq 1, |x[2]| \leq 1\}.$$

На рис. 1 изображены множество Ω и линии уровня целевой функции (окружности с центром в точке (2.0)). Все вершины и два

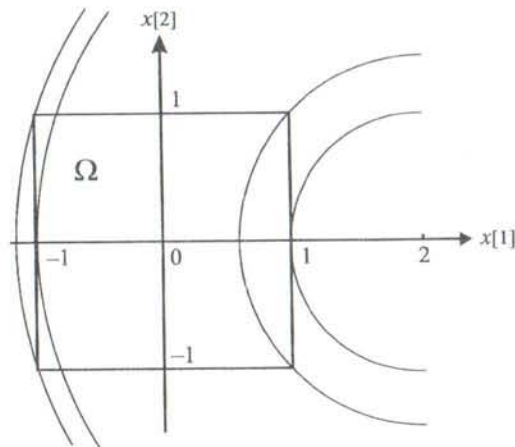


Рис. 1.

вертикальных ребра квадрата Ω являются квазистационарными гранями, а два горизонтальных ребра и внутренность квадрата ($J = \emptyset$) не являются квазистационарными гранями, так как точка минимума $F(x)$ на аффинных оболочках этих граней не принадлежит Ω . Очевидно, что среди квазистационарных граней наименьшую оценку имеет правое вертикальное ребро, а соответствующей квазистационарной точкой является точка (1, 0), которая и является решением рассматриваемой задачи. ■

Заметим, что квазистационарных точек может быть бесконечно много. Например, минимизируя $x[1]$ на плоскости при ограничении $x[1] \geq 1$, видим, что вся прямая $x[1] = 1$ (рис. 2) состоит из квазистационарных точек и в то же время эта прямая является единственной квазистационарной гранью.

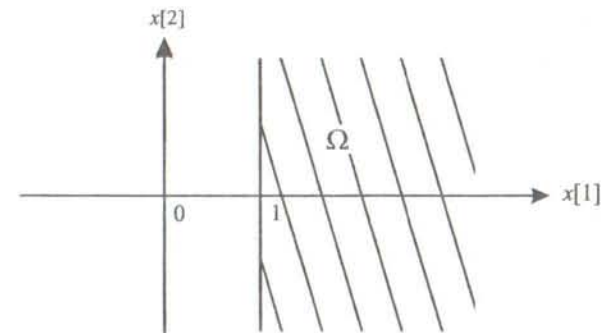


Рис. 2.

Таким образом, квазистационарных граней конечное число и, если исходная задача разрешима, то квазистационарная грань с наименьшей оценкой является оптимальной, а соответствующая ей квазистационарная точка (любая, если их много) является решением задачи. Метод ПГ заключается в *поиске квазистационарной грани с наименьшей оценкой*.

Для вывода признаков оптимальности квазистационарной точки приведем ее аналитическую характеристику.

§2. Характеристика квазистационарной точки

Заметим, что ограничения задачи (1.3) состоят только из равенств, поэтому ее решение равносильно решению линейной системы КТ относительно неизвестных $x[N], u[J]$:

$$\begin{aligned} D[N, N] \cdot x[N] - A^T[N, J] \cdot u[J] &= -c[N], \\ A[J, N] \cdot x[N] &= b[J]. \end{aligned} \quad (2.1)$$

Введем в рассмотрение вектор — градиент целевой функции $F(x)$

$$g(x)[N] := D[N, N] \cdot x[N] + c[N].$$

Теорема 2.1. Точка $x^* \in \Omega$ квазистационарна тогда и только тогда, когда градиент целевой функции в этой точке есть линейная комбинация строк матрицы ограничений, активных в точке x^* , т. е.

$$g(x^*)[N] = A^T[N, I(x^*)] \cdot u^*[I(x^*)]. \quad (2.2)$$

Доказательство. Если $x^* \in \Omega$ квазистационарна, то для $J = I(x^*)$ разрешима система (2.1) и ее решением являются x^* и $u^*[J]$. Первое равенство (2.1) совпадает с (2.2).

Пусть теперь выполнено (2.2) для $x^* \in \Omega$. Положим $J = I(x^*)$. Очевидно, что тогда x^* и $u^*[J]$ удовлетворяют системе (2.1), следовательно, x^* является квазистационарной точкой. ■

Теперь дадим характеристику решения задачи (1.1), (1.2) через квазистационарную точку.

Теорема 2.2. Точка $x^* \in \Omega$ является решением задачи (1.1), (1.2) тогда и только тогда, когда x^* квазистационарна и коэффициенты в представлении градиента (2.2) неотрицательны, т. е. $u^*[I(x^*)] \geq \mathbb{O}$.

Доказательство. Точка x^* является решением задачи (1.1), (1.2) тогда и только тогда, когда в точке x^* выполнены условия КТ (см. гл. 1), которые для задачи (1.1), (1.2) имеют вид

$$\begin{aligned} Dx^* - A^T u^* &= -c, \\ Ax^* - v^* &= b, \\ u^* &\geq \mathbb{O}, \quad v^* \geq \mathbb{O}, \\ u[i] \cdot v[i] &= 0, \quad i \in M. \end{aligned}$$

Нетрудно видеть, что эти условия равносильны следующим:

$$x^* \in \Omega, \quad g(x^*)[N] = A^T[N, I(x^*)] \cdot u[I(x^*)],$$

$$u^*[I(x^*)] \geq \mathbb{O}.$$

Теорема доказана. ■

Геометрически условия теоремы 2.2 означают, что градиент в точке $x^* \in \Omega$ принадлежит конической оболочке, натянутой на строки матрицы ограничений, активных в точке x^* .

В заключение приведем достаточный признак квазистационарности точки.

Теорема 2.3. Пусть при некотором $J \subseteq M$ разрешима система (2.1) и пара $\{x_0, u_0\}$ — ее решение. Если точка x_0 принадлежит Ω , то она квазистационарна.

Доказательство. Из (2.1) следует, что

$$J \subset I(x_0), \quad g(x_0)[N] = A^T[N, J] \cdot u_0[J].$$

Положив $u_0[I(x_0) \setminus J] = \mathbb{O}$, можно представить градиент в виде $g(x_0)[N] = A^T[N, I(x_0)] \cdot u_0[I(x_0)]$. Так как $x_0 \in \Omega$, то по теореме 2.1 точка x_0 квазистационарна. ■

§3. Описание и обоснование метода перебора граней

Существует несколько методов [13, 16, 19–21], реализующих изложенную здесь идею перебора граней. В работе Бэста [16] представлен некоторый алгоритм ПГ и показано, что все выше перечисленные методы описываются этим же алгоритмом, а различаются лишь способом минимизации $F(x)$ на аффинной оболочке грани, т. е. способом решения системы (2.1).

Приведем описание метода Бэста без указания способа решения системы (2.1).

3.1. Подходящая пара

Метод ПГ требует предварительного поиска точки из Ω , которую будем называть *подходящей*. Для определения этой точки обратимся к системе (2.1), положив $w[J] = -u[J]$, чтобы матрица системы $G(J)$ стала симметричной:

$$G(J) = \begin{pmatrix} D[N, N] & A^T[N, J] \\ A[J, N] & \mathbb{O}[J, J] \end{pmatrix}.$$

Определение. Точка $y \in \Omega$ называется *подходящей*, если существует множество $J \subseteq I(y)$, на котором матрица $G(J)$ невырождена. Пару $\{y, J\}$ будем называть *подходящей парой*, индексное множество J будем называть *подходящим* для y .

Выясним, когда матрица $G(J)$ будет невырожденной.

Утверждение 3.1. Матрица $G(J)$ невырождена тогда и только тогда, когда выполнены два условия:

$$\text{rang } A[J, N] = |J| \quad (3.1)$$

и матрица D положительно определена на множестве

$$S(J) := \left\{ s[N] \mid A[J, N] \cdot s[N] = \mathbb{O} \right\}.$$

Доказательство. Необходимость. Равенство (3.1) очевидно, иначе была бы линейная зависимость последних строк матрицы $G(J)$. Допустим, что (3.1) выполнено, но существует ненулевой вектор $s \in S(J)$, такой, что $\langle Ds, s \rangle = 0$. В силу неотрицательной определенности матрицы D вектор s из $S(J)$ удовлетворяет равенствам

$$\begin{aligned} D[N, N] \cdot s[N] &= \mathbb{O}, \\ A[J, N] \cdot s[N] &= \mathbb{O}. \end{aligned}$$

Эти равенства означают линейную зависимость столбцов матрицы $G(J)$, что невозможно.

Достаточность. Допустим, что для столбцов матрицы $G(J)$ имеют место равенства

$$\begin{aligned} D[N, N] \cdot \xi[N] + A^T[N, J] \cdot \xi[J] &= \mathbb{O}, \\ A[J, N] \cdot \xi[N] &= \mathbb{O}. \end{aligned}$$

Покажем, что $\xi[N \cup J] = \mathbb{O}$. Из второго равенства следует, что $\xi[N] \in S(J)$. Умножим слева обе части первого равенства на $\xi[N]$ и учтем второе равенство. Тогда имеем $\xi[N] \cdot D[N, N] \cdot \xi[N] = 0$. В силу положительной определенности матрицы D на S отсюда

следует, что $\xi[N] = \mathbb{O}$. Теперь из первого равенства и (3.1) получаем $\xi[J] = \mathbb{O}$. ■

В дальнейшем множество J придется расширять. Выясним, при каком условии расширенная матрица $G(J)$ будет невырожденной.

Следствие. Пусть строка $A[l, N]$ линейно независима со строками $A[J, N]$. Если матрица $G(J)$ невырожденная, то невырожденной является расширенная матрица $G(J_1)$, где $J_1 = J \cup \{l\}$.

Доказательство. Условия утверждения 3.1 выполнены, так как $S(J_1) \subset S(J)$. ■

Замечание 3.1. Если $|J| = |N|$ или матрица D положительно определена, то для невырожденности матрицы $G(J)$ достаточно выполнения условия (3.1).

Общий алгоритм для нахождения подходящей точки сложен. В каждом отдельном случае подходящая точка находится своим способом. В частности:

- подходящей точкой является вершина множества Ω ;
- для положительно определенной матрицы D подходящей является любая точка $x \in \Omega$ (в качестве J можно взять пустое множество);
- для $D = \mathbb{O}$ подходящей точкой является только вершина, а подходящим для нее множеством J может быть только строчный базис вершины (см. определение вершины в §1).

Часто удается найти подходящую точку очень просто, не прибегая к построению вершины.

Заметим, что разрешимая задача ВКП может не иметь подходящей точки. Например, для следующей задачи на плоскости ($n = 2$):

$$\frac{1}{2}x[1] \rightarrow \inf_{x[1] \geq 1}$$

матрица $G(J)$ при любом J содержит нулевую строку и нулевой столбец. В таких случаях подходящую точку можно создать искусственно. Добавив в рассматриваемом примере нейтральное ограничение $x[2] \leq 1$, получим вершину, которая и является подходящей точкой.

3.2. Выбор направления спуска

Выясним, как перейти от одной грани к другой с меньшей оценкой. Обратимся к системе (2.1). Будем искать ее решение в виде

$$x[N] = z[N] + s[N], \quad z \in \Omega.$$

Так как $g(z) := Dz + c$, то система (2.1) относительно новых переменных $s[N]$ и $w[J]$ ($w[J] = -u[J]$) переписывается следующим образом:

$$\begin{aligned} D[N, N] \cdot s[N] + A^T[N, J] \cdot w[J] &= -g(z)[N], \\ A[J, N] \cdot s[N] &= \mathbb{O}. \end{aligned} \quad (3.2)$$

Если пара $\{z, J\}$ подходящая, то матрица $G(J)$ невырожденная, поэтому система (3.2) для подходящей пары разрешима. Выясним смысл вектора $s[N]$.

Утверждение 3.2. Пусть пара $\{z, J\}$ является подходящей и вектор $s[N]$ — решение системы (3.2). Если точка z не квазистационарна, то вектор s определяет направление строгого убывания функции $F(x)$ в точке z , а именно: при $t \in (0, 2)$ имеем

$$F(z + ts) - F(z) = -t\left(1 - \frac{t}{2}\right) \langle Ds, s \rangle < 0.$$

Доказательство. Так как точка z не является квазистационарной, то $s[N] \neq \mathbb{O}$. Воспользуемся разложением квадратичной функции $F(x)$ в точке $z + ts$:

$$F(z + ts) - F(z) = t \langle g(z), s \rangle + \frac{1}{2} t^2 \langle Ds, s \rangle.$$

Умножим первое равенство (3.2) слева на s и учтем второе. Тогда получим $\langle g(z), s \rangle = -\langle Ds, s \rangle$, т. е.

$$F(z + ts) - F(z) = -t \left(1 - \frac{t}{2}\right) \langle Ds, s \rangle.$$

Так как пара $\{z, J\}$ подходящая, то матрица $G(J)$ невырожденна. В силу (3.2) $s \in S(J)$, поэтому из утверждения 3.1 следует, что $\langle Ds, s \rangle > 0$. Остальное очевидно. ■

3.3. Описание метода

Приведем сначала общую схему метода ПГ, разбив его на три части.

Пусть $\{z_0, J_0\}$ — начальная подходящая пара и $k = 1$.

A1. Построение квазистационарной точки x_k . Начиная с подходящей пары $\{z_0, J_0\}$, путем добавления ограничений строим грани, пока не получим квазистационарную грань. В результате будут найдены квазистационарная точка x_k , индексное множество $I_k \subseteq I(x_k)$ и получено представление градиента

$$g(x_k) = A^T[N, I_k] \cdot u_k[I_k]. \quad (3.3)$$

При этом $F(x_k) \leq F(z_0)$.

A2. Проверка квазистационарной точки на оптимальность. Если $u_k[I_k] \geq \mathbb{O}$, то согласно теореме 2.2 точка x_k — решение исходной задачи. На этом процесс заканчивается.

Если условие оптимальности нарушено, т. е. не пусто множество

$$P_k = \{j \in I_k \mid u_k[j] < 0\},$$

то выберем любой индекс $p_k \in P_k$ и положим $J'_k = I_k \setminus \{p_k\}$. Построим матрицу $G(J'_k)$. Если матрица $G(J'_k)$ невырожденная, то пара $\{x_k, J'_k\}$ является подходящей. Переходим к A1 на построение следующей квазистационарной пары, полагая $z_0 = x_k$, $J_0 = J'_k$, $k = k + 1$.

Если матрица $G(J'_k)$ вырождена, то переходим к A3.

A3. Формирование подходящей пары в вырожденном случае. Из пары $\{x_k, I_k\}$ путем замены индекса p_k на некоторый индекс l_k строим новую подходящую пару $\{z_0, J_0\}$, для которой $F(z_0) \leq F(x_k)$. Затем переходим к A1, полагая $k = k + 1$. Заметим, что требуемый индекс l_k может не существовать. Это означает, что исходная задача неразрешима.

Сделаем подробное описание алгоритмов A1 и A3 с пояснением.

A1. Имеем подходящую пару $\{z_0, J_0\}$. Положим $\nu = 0$ и проведем следующие операции.

A1.1. *Нахождение вектора s_ν , определяющего направление спуска целевой функции в точке z_ν .* Для этого решаем систему

$$\begin{aligned} D[N, N] \cdot s[N] + A^T[N, J_\nu] \cdot w[J_\nu] &= -g(z_\nu)[N], \\ A[J_\nu, N] \cdot s[N] &= \mathbb{O}. \end{aligned} \quad (3.4)$$

Заметим, что система разрешима, так как матрица этой системы $G(J_\nu)$ невырожденная. Обозначим через $\{s_\nu, w_\nu\}$ решение системы и переходим к А1.2.

А1.2. *Определение допустимого шага σ_ν в направлении спуска.* Для этого подставляем в ограничения (1.2) точки луча

$$z(t) = z_\nu + ts_\nu, \quad t \geq 0,$$

и находим σ_ν — максимальное значение t , при котором $z(t) \in \Omega$. Очевидно, что

$$\sigma_\nu = \infty, \quad \text{если } As_\nu \geq \mathbb{O},$$

в противном случае

$$\sigma_\nu = \min \left\{ \frac{b[j] - A[j, N] \cdot z_\nu[N]}{A[j, N] \cdot s_\nu[N]} \mid j \in M : A[j, N] \cdot s_\nu[N] < 0 \right\},$$

так что $z(t) = z_\nu + ts_\nu \in \Omega \quad \forall t \in [0, \sigma_\nu]$.

Если $\sigma_\nu > 1$, переходим к А1.4. Если $\sigma_\nu \leq 1$, переходим к А1.3.

А1.3. *Формирование новой подходящей пары.* Нетрудно видеть, что при $\sigma_\nu \leq 1$ вектор s_ν ненулевой и точка

$$z_{\nu+1} = z_\nu + \sigma_\nu s_\nu$$

принадлежит Ω . Пусть L'_ν — множество индексов, на которых реализуется минимум при вычислении σ_ν , и пусть индекс $l_\nu \in L'_\nu$. Полагаем

$$J_{\nu+1} = J_\nu \cup \{l_\nu\}.$$

Ясно, что множество $J_{\nu+1}$ содержится в $I(z_{\nu+1})$ и согласно следствию к утверждению 3.1 пара $\{z_{\nu+1}, J_{\nu+1}\}$ является подходящей. Заметим, что согласно утверждению 3.2 (при $z = z_\nu, J = J_\nu, t = \sigma_\nu \leq 1$) имеем

$$F(z_{\nu+1}) \leq F(z_\nu) \quad \text{и, если } \sigma_\nu > 0, \quad \text{то } F(z_{\nu+1}) < F(z_\nu).$$

Полагаем $\nu = \nu + 1$ и возвращаемся к А1.1.

А1.4. *Формирование квазистационарной пары.* В этом случае точка $z(1) = z_\nu + s_\nu$ принадлежит Ω и по теореме 2.3 является квазистационарной. Полагаем

$$x_k = z(1), \quad I_k := J_\nu, \quad u_k[I_k] := -w_\nu[J_\nu].$$

Из утверждения 3.2 (при $z = z_\nu, J = J_\nu, t = 1$) следует, что

$$F(x_k) \leq F(z_\nu) \quad \text{и, если } s_\nu \neq \mathbb{O}, \quad \text{то } F(x_k) < F(z_\nu). \quad (3.5)$$

Переходим к А2.

Заметим, что если $\sigma_\nu = 1$, то $z(1) = z_\nu + s_\nu \in \Omega$, значит, точка $z(1)$ является квазистационарной, но ограничение с индексом l_ν не принадлежит J_ν , хотя и является активным в точке $z(1)$. Чтобы добавить индекс l_ν в активное индексное множество, при $\sigma_\nu = 1$ переходим к А1.3.

Очевидно, что алгоритм А1 для построения квазистационарной точки *копечен*. Действительно, на каждой итерации множество J_ν расширяется. Если же при некотором ν оказалось $|J_\nu| = |N|$, то из (3.4) следует, что $s_\nu[N] = \mathbb{O}$, т.е. точка z_ν квазистационарна.

А3. Эта часть алгоритма состоит из одной итерации. Находим вектор $r_k[N]$, который удовлетворяет равенствам

$$\begin{aligned} D[N, N] \cdot r_k[N] &= \mathbb{O}, \\ A[J'_k, N] \cdot r_k[N] &= \mathbb{O}, \\ A[p_k, N] \cdot r_k[N] &= 1. \end{aligned} \quad (3.6)$$

Нетрудно показать, что $r_k[N] = [G(I_k)]^{-1}[N, p_k]$ (см. ниже утверждение 4.1). Такой вектор r_k определяет направление строгого убывания $F(x)$ в квазистационарной точке x_k .

Действительно, из первого условия (3.6) при $t > 0$ имеем

$$F(x_k + tr_k) - F(x_k) = t \langle r_k, g(x_k) \rangle.$$

Из условия (3.3) $g(x_k)[N] = A^T[N, I_k] \cdot u_k[I_k]$, поэтому в силу (3.6)

$$\begin{aligned} F(x_k + tr_k) - F(x_k) &= t \cdot r_k[N] \cdot A^T[N, I_k] \cdot u_k[I_k] = \\ &= t \cdot u_k[p_k] < 0. \end{aligned} \quad (3.7)$$

Находим t_k — максимальное значение параметра t , при котором точки $x(t) = x_k + tr_k$ принадлежат Ω . Если $Ar_k \geq \mathbb{O}$, то $t_k = \infty$. В противном случае

$$t_k = \min \left\{ \frac{b[j] - A[j, N] \cdot x_k[N]}{A[j, N] \cdot r_k[N]} \mid j \in M : A[j, N] \cdot r_k[N] < 0 \right\}.$$

Пусть t_k конечно. Обозначим через L_k множество индексов из M , на которых реализуется минимум при вычислении шага t_k , и пусть $l_k \in L_k$. Полагаем

$$z_0 = x_k + t_k \cdot r_k, \quad J_0 = I_k \setminus \{p_k\} \cup \{l_k\}.$$

Полученная пара $\{z_0, J_0\}$ является подходящей (см. ниже утверждение 3.3). При этом согласно (3.7) $F(z_0) \leq F(x_k)$, а если $t_k > 0$, то $F(z_0) < F(x_k)$. Далее, полагая $k = k + 1$, переходим к А1.

Если $t_k = \infty$, то в силу (3.7) исходная задача не имеет решения, так как луч $x(t) = x_k + t \cdot r_k$, $t \geq 0$, целиком принадлежит Ω и целевая функция $F(x)$ на этом луче монотонно убывает. На этом процесс заканчивается. ■

Теперь метод описан полностью.

3.4. Обоснование метода

Для полного обоснования алгоритма осталось доказать

Утверждение 3.3. Построенная в АЗ пара $\{z_0, J_0\}$ является подходящей.

Доказательство. Напомним, что $J_0 = I_k \setminus \{p_k\} \cup \{l_k\}$. Достаточно показать, что матрица $G(J_0)$ невырожденная.

Согласно алгоритму

$$A[l_k, N] \cdot r_k[N] < 0, \quad (3.8)$$

так что из (3.6) следует, что строки матрицы $A[J_0, N]$ линейно независимы.

Допустим теперь противное, т. е. матрица $G(J_0)$ вырожденная. По утверждению 3.1 тогда существует ненулевой вектор $\xi[N]$, такой, что

$$\begin{aligned} D[N, N] \cdot \xi[N] &= \mathbb{O}, \\ A[J_0, N] \cdot \xi[N] &= \mathbb{O}. \end{aligned}$$

Так как предыдущая матрица $G(I_k)$ невырожденная, то

$$q := A[p_k, N] \cdot \xi[N] \neq 0.$$

Введем в рассмотрение вектор $\eta[N] = \xi[N] - q \cdot r_k[N]$. Из последнего равенства (3.6) следует, что $A[p_k, N] \cdot \eta[N] = 0$. Учитывая (3.8), видим, что вектор η ненулевой, так как

$$A[l_k, N] \cdot \eta[N] = -q \cdot A[l_k, N] \cdot r_k[N] \neq 0.$$

Учитывая (3.6), получаем

$$\begin{aligned} D[N, N] \cdot \eta[N] &= \mathbb{O}, \\ A[J'_k, N] \cdot \eta[N] &= \mathbb{O}, \\ A[p_k, N] \cdot \eta[N] &= 0. \end{aligned}$$

Но эти равенства невозможны, так как $J'_k \cup \{p_k\} = I_k$, а $G(I_k)$ — невырожденная матрица. ■

Замечание 3.2. Если матрица D положительно определена, то алгоритм АЗ лишний, он подключается, только если матрица D вырожденная. Если же матрица $D = \mathbb{O}$, то алгоритм АЗ реализуется на каждом шаге.

§4. Численная реализация метода перебора граней

Прежде уже говорилось, что разными авторами в разное время были разработаны методы решения задач ВКП [13, 16, 19 – 21], которые по своему идейному содержанию являются методами перебора граней, а различаются способом минимизации квадратичной функции на аффинной оболочке грани, т. е. способом решения задачи (1.3). На это обратил внимание Бэст в работе [16]. Его наблюдение сводится к следующему: *если все перечисленные методы и метод, описанный в §3, применить к одной и той же начальной подходящей паре $\{x_0, J_0\}$ и во всех местах, где выбор индекса не однозначен, выбрать его одинаково, то все указанные методы построят идентичную последовательность точек x_k .*

Приведем один из способов решения задачи (1.3), а именно, будем решать систему КТ для этой задачи, т. е. систему (2.1), с использованием и пересчетом обратной матрицы системы.

Пусть в начале алгоритма метода ПГ имеется не только подходящая пара $\{y_0, J_0\}$, но и обратная матрица $G^{-1}(J_0)$. Тогда все следующие обратные матрицы получаются путем пересчета предыдущей, а решение системы (2.1) сводится к умножению обратной матрицы на правую часть системы.

Напомним, что матрица $G(J)$ симметрична и имеет следующую блочную структуру:

$$G(J) = \begin{pmatrix} D[N, N] & A^T[N, J] \\ A[J, N] & \mathbb{O}[J, J] \end{pmatrix}.$$

Обозначим $B(J) = G^{-1}(J)$. Выведем формулы для пересчета обратных матриц в разных ситуациях.

Когда в очередной раз мы попадаем в А1.1, где решается система (2.1)⁸, то индексное множество J отличается от предыдущего лишь одним индексом: либо оно увеличивается на один индекс, либо уменьшается, либо происходит замена одного индекса другим. Рассмотрим эти три ситуации.

1) В А1.1 попадаем из А1.3. Тогда новое индексное множество становится шире: $J_1 = J \cup \{l\}$. Обозначим

$$v^T := (A[l, N], \mathbb{O}[J]).$$

К невырожденной матрице $G(J)$ добавляется строка и столбец и получается невырожденная матрица

$$G(J_1) = \begin{pmatrix} G(J) & v \\ v^T & 0 \end{pmatrix}.$$

Пусть $\alpha = v^T \cdot B(J) \cdot v$, $z = \frac{1}{\alpha} B(J) \cdot v$. Тогда

$$B(J_1) = \begin{pmatrix} B(J) - \alpha \cdot z \cdot z^T & z \\ z^T & -\frac{1}{\alpha} \end{pmatrix}. \quad (4.1)$$

Действительно, нетрудно проверить, что $G(J_1) \cdot B(J_1) = E$. Заметим, что так как матрица $B(J_1)$ существует, то $\alpha \neq 0$.

2) Обратимся к п. А2. Имеем обратную матрицу $B := G^{-1}(I)$ и индексное множество $J = I \setminus \{p\}$ (для упрощения записи опустим итерационный индекс k и положим $J = J'$). В данном случае индексное множество J стало уже. Положим

$$b := B[p, p], \quad N_p := N \cup J, \quad r := r[N_p] = B[N_p, p].$$

Утверждение 4.1. Если $b \neq 0$, то матрица $G(J)$ невырожденная и компоненты обратной к ней матрицы $B_0[N_p, N_p]$ определяются по формулам

$$B_0[i, j] = B[i, j] - \frac{1}{b} r[i] \cdot r[j], \quad i, j \in N_p. \quad (4.2)$$

Если $b = 0$, то матрица $G(J)$ вырожденная и вектор $r[N]$ удовлетворяет системе (3.6).

⁸В А1.1 эта система имеет номер (3.4) и $J = J_v$.

Доказательство. Переставим p -ю строку и p -й столбец матрицы $G(I)$ на последнее место. Получим новую матрицу $C = P \cdot G(I) \cdot P^T$, где P — матрица перестановок. Заметим, что

$$C^{-1} = (P \cdot G(I) \cdot P^T)^{-1} = P \cdot B \cdot P^T,$$

т. е. матрица C^{-1} получается из матрицы B в результате такой же перестановки строк и столбцов. Положим

$$v^T[N_p] := (A[p, N], \mathbb{O}[J]), \quad B_p = B[N_p, N_p].$$

Тогда матрицы C и C^{-1} имеют следующие блочные представления:

$$C = \begin{pmatrix} G(J) & v \\ v^T & 0 \end{pmatrix}, \quad C^{-1} = \begin{pmatrix} B_p & r \\ r^T & b \end{pmatrix}.$$

Перемножая матрицы C и C^{-1} , получаем равенства

$$G(J) \cdot B_p + v \cdot r^T = E, \quad (4.3)$$

$$G(J) \cdot r + b \cdot v = \mathbb{O}, \quad (4.4)$$

$$v^T \cdot r = 1. \quad (4.5)$$

Пусть $b \neq 0$, тогда из (4.4) имеем $v = -\frac{1}{b} \cdot G(J) \cdot r$. Подставляя это выражение в (4.3), видим, что

$$G(J) \cdot \left(B_p - \frac{1}{b} r \cdot r^T \right) = E.$$

Так как справа в этом выражении стоит единичная матрица, то матрица $G(J)$ невырожденная и справедливо (4.2).

Пусть теперь $b = 0$. Из (4.5) следует, что

$$A[p, N] \cdot r[N] = 1,$$

а из (4.4) получаем линейную зависимость столбцов матрицы $G(J)$, т. е.

$$\begin{aligned} D[N, N] \cdot r[N] + A^T[N, J] \cdot r[J] &= \mathbb{O}, \\ A[J, N] \cdot r[N] &= \mathbb{O}. \end{aligned} \quad (4.6)$$

Умножим первое равенство слева на $r[N]$. Учитывая второе равенство, получаем $r[N] \cdot D[N, N] \cdot r[N] = 0$. Так как матрица D неотрицательно определена, то $D[N, N] \cdot r[N] = \mathbb{O}$. Присоединяя к этому

равенству второе равенство (4.6), видим, что $r[N]$ удовлетворяет условиям (3.6):

$$\begin{aligned} D[N, N] \cdot r[N] &= \mathbb{O}, \\ A[J, N] \cdot r[N] &= \mathbb{O}, \\ A[p, N] \cdot r[N] &= 1. \quad \blacksquare \end{aligned}$$

3) Пусть в A2 матрица $G(J')$ вырожденная и согласно алгоритму A3 построена невырожденная матрица $G(J_0)$, которая получена из невырожденной матрицы $G(I)$ заменой $A[p, N]$ на $A[l, N]$ в строке и в столбце, т. е. $J_0 = I \setminus \{p\} \cup \{l\}$ (нижний индекс k у индексов p, l и у множества I опущен). Требуется построить матрицу $B_0 := G^{-1}(J_0)$, используя матрицу $B := G^{-1}(I)$.

Утверждение 4.2. Пусть

$$s[i] = \sum_{\nu \in N} A[l, \nu] \cdot B[i, \nu], \quad \tau[i] = \frac{1}{s[p]} B[p, i], \quad i \in N \cup I,$$

$$\gamma = \sum_{\nu \in N} A[l, \nu] \cdot s[\nu].$$

Тогда для всех $i, j \neq l$

$$\begin{aligned} B_0[j, i] &= B[j, i] - \tau[i] \cdot s[j] - \tau[j] \cdot s[i] + \gamma \cdot \tau[i] \cdot \tau[j], \\ B_0[l, i] &= B_0[i, l] = \tau[i]. \end{aligned} \quad (4.7)$$

Доказательство. Сначала в матрице $G(I)$ заменяем только столбец с индексом p на столбец с индексом l , т. е. строим матрицу

$$G_1 = \begin{pmatrix} D[N, N] & A^T[N, J_0] \\ A[I, N] & \mathbb{O} \end{pmatrix}.$$

Для того чтобы выяснить, является ли эта матрица невырожденной, находим коэффициенты разложения вектора

$$\begin{pmatrix} A^T[N, l] \\ \mathbb{O} \end{pmatrix}$$

по столбцам матрицы $G(I)$, т. е. вычисляем

$$s[i] = B[i, N] \cdot A^T[N, l], \quad i \in N \cup I. \quad (4.8)$$

Как известно, матрица G_1 невырожденная, если отлична от нуля компонента $s[p]$. По утверждению 4.1 $B[p, N] = r[N]$, поэтому

$$s[p] = r[N] \cdot A^T[N, l].$$

Согласно алгоритму индекс l выбирался так, чтобы $s[p] < 0$. Значит, матрица G_1 невырожденная и ее обратная B_1 строится по формулам (8.3) гл. 1 следующим образом:

$$B_1[l, j] = B[p, j]/s[p], \quad j \in N \cup I, \quad (4.9)$$

$$B_1[i, j] = B[i, j] - B_1[l, j] \cdot s[i], \quad i \neq p, \quad j \in N \cup I. \quad (4.10)$$

Отметим одно свойство построенной матрицы B_1 . Напомним, что матрица $G(I \setminus \{p\})$ вырожденная, так что согласно утверждению 4.1 $B[p, p] = 0$, поэтому $B_1[l, p] = 0$ и p -й столбец матрицы B не меняется, т. е.

$$B_1[i, p] = B[i, p], \quad i \in I \cup N. \quad (4.11)$$

Теперь заменим в матрице G_1 строку $(A[p, N], \mathbb{O})$ новой строкой $h[N] := (A[l, N], \mathbb{O})$, в результате чего получим матрицу $G(J_0)$. Для построения $B_0 := G^{-1}(J_0)$ произведем преобразование матрицы B_1^T по формулам (8.3) гл. 1. Для этого найдем вектор коэффициентов разложения строки $h[N]$ по строкам матрицы G_1 :

$$s_1[i] = A[l, N] \cdot B_1[N, i], \quad i \in I \cup N. \quad (4.12)$$

В силу (4.11) $s_1[p] = s[p]$.

Обозначим $\tau[j] := B[j, p]/s[p]$, тогда в силу (4.11) по формулам преобразования обратной матрицы имеем

$$B_0[j, l] = B_1[j, p]/s_1[p] = \tau[j],$$

поэтому

$$B_0[j, i] = B_1[j, i] - \tau[j] \cdot s_1[i], \quad i \neq l. \quad (4.13)$$

Выведем формулу выражения B_0 непосредственно через B . Для этого выразим s_1 через s , подставив в (4.12) выражения (4.9) и (4.10):

$$s_1[i] = \sum_{\nu \in N} A[l, \nu] \cdot (B[\nu, i] - \tau[i] \cdot s[\nu]) = s[i] - \tau[i] \cdot \gamma,$$

где $\gamma = \sum_{\nu \in N} A[l, \nu] \cdot s[\nu]$. Подставив это выражение в (4.13), получим требуемые формулы пересчета. ■

Итак, в методе обратных матриц для пересчета текущей обратной матрицы используются формулы: либо (4.1), либо (4.2), либо (4.7).

§5. Алгоритм перебора граней с обратными матрицами

Приведем краткую запись алгоритма Бэста с использованием обратных матриц.

Считаем, что известны подходящая пара $\{z_0, J_0\}$ и матрица $B_0 = [G(J_0)]^{-1}$.

1. Пусть ν — номер текущей итерации в п. 1. Положим $\nu = 0$.

1.1. Находим $s[N] = -B_\nu[N, N] \cdot g(z_\nu)[N]$.

Вычисляем шаг σ . Если $As \geq \mathbb{O}$, то $\sigma = \infty$. В противном случае

$$\sigma = \min \left\{ \frac{b[j] - A[j, N] \cdot z_\nu[N]}{A[j, N] \cdot s[N]} \mid j \in M : A[j, N] \cdot s[N] < 0 \right\}.$$

Пусть l — тот индекс, на котором реализуется минимум при вычислении σ . Если $\sigma \leq 1$, то полагаем $z_{\nu+1} := z_\nu + \sigma \cdot s$, $J_{\nu+1} := J_\nu \cup \{l\}$. Затем строим матрицу $B_{\nu+1}$ из B_ν методом окаймления по формулам (4.1), в которых $J = J_\nu$. Затем полагаем $\nu = \nu + 1$ и возвращаемся к п. 1.1.

Если $\sigma > 1$, то точка $x := z_\nu + s$ квазистационарна. Вычисляем

$$u[J_\nu] := B_\nu[J_\nu, N] \cdot g(z_\nu)[N], \quad u[p] := \min_{j \in J_\nu} u[j].$$

Если $u[p] \geq 0$, то точка x — решение исходной задачи. На этом процесс заканчивается.

Если $u[p] < 0$, то полагаем $J' = J_\nu \setminus \{p\}$. Если $b := B_\nu[p, p] \neq 0$, то матрица $G(J')$ невырожденная. В этом случае полагаем $z_0 = x$, $J_0 = J'$ и находим обратную матрицу B_0 через B_ν по формулам (4.2) согласно обозначениям $B := B_\nu$, $J := J_0$. Затем переходим к п. 1.

Если $b = 0$, то полагаем $r[N] = B_\nu[N, p]$. Если $Ar \geq \mathbb{O}$, то исходная задача неразрешима. Иначе находим допустимый шаг в направлении r :

$$t = \min \left\{ \frac{b[j] - A[j, N] \cdot x_k[N]}{A[j, N] \cdot r[N]} \mid j \in M : A[j, N] \cdot r[N] < 0 \right\}.$$

Пусть l — тот индекс, на котором достигается этот минимум. Полагаем $z_0 = x + t \cdot r$, $J_0 = J' \cup \{l\}$ и находим B_0 из B_ν по формулам (4.7), в которых $B := B_\nu$. Затем переходим к п. 1. ■

§6. О конечности метода перебора граней

Метод ПГ строит последовательность квазистационарных пар

$$\{x_k, I_k\}, \quad (6.1)$$

где x_k — квазистационарная точка, $I_k \subseteq I(x_k)$ — подходящее для нее множество, при этом $F(x_{k+1}) \leq F(x_k)$. Если в этой последовательности множества I_k попарно различны, то описанный алгоритм *конечен*. Если же в указанной последовательности найдется $I_{k+\tau} = I_k$, то вычисления от I_k до $I_{k+\tau}$ бесконечное число раз повторяются. Это явление называется *защелкиванием*.

Существует много разных способов устранения защелкивания. Остановимся на правиле выбора наименьшего индекса [17]. Покажем, что если в A1.2, A2, A3 индексы l_ν , p_k , l_k выбирать по правилу *наименьшего индекса*:

$$\begin{aligned} l_\nu &= \min \{ j \mid j \in L'_\nu \}, \\ p_k &= \min \{ j \mid j \in P_k \}, \\ l_k &= \min \{ j \mid j \in L_k \}, \end{aligned}$$

то защелкивание невозможно.⁹

Предварительно докажем две леммы.

Лемма 6.1. Пусть на k -й итерации реализуется алгоритм A2 → A1. Подходящая пара $\{z_0 = x_k, J_0 = J'_k\}$, сформированная в A2, такова, что для нее в A1

$$s_0 \neq \mathbb{O}, \quad A[p_k, N] \cdot s_0[N] > 0.$$

⁹Вопрос о применении правила наименьшего индекса к задаче ВКП был рассмотрен Келлером [18]. Здесь же приводится независимое доказательство правомерности этого правила.

Доказательство. Вектор s_0 удовлетворяет линейной системе (3.4), поэтому

$$\begin{aligned} D[N, N] \cdot s_0[N] + A^T[N, J'_k] \cdot w_0[J'_k] &= -g(x_k)[N], \\ A[J'_k, N] \cdot s_0[N] &= \mathbb{O}. \end{aligned} \quad (6.2)$$

Напомним, что точка x_k квазистационарна и для нее выполнено (3.3), т.е. $g(x_k)[N] = A^T[N, I_k] \cdot u_k[I_k]$.

Допустим, что $s_0 = \mathbb{O}$. Подставляя в правую часть (6.2) выражение градиента (3.3), получаем равенство

$$A^T[N, J'_k] \cdot (w_0[J'_k] + u_k[J'_k]) + A^T[N, p_k] \cdot u_k[p_k] = \mathbb{O},$$

которое означает линейную зависимость строк матрицы $A[I_k, N]$, что неверно. Значит, $s_0 \neq \mathbb{O}$.

Далее, так как матрица $G(J'_k)$ невырожденная, то $\langle Ds_0, s_0 \rangle > 0$. Из (6.2) и представления градиента находим

$$\begin{aligned} \langle Ds_0, s_0 \rangle &= -\langle s_0, g(x_k) \rangle = \\ &= -s_0[N] \cdot A^T[N, I_k] \cdot u_k[I_k] = -s_0[N] \cdot A^T[N, p_k] \cdot u_k[p_k] > 0. \end{aligned}$$

Так как $u_k[p_k] < 0$, то $A[p_k, N] \cdot s_0[N] > 0$. Лемма доказана. ■

Лемма 6.2. Если для двух соседних квазистационарных пар $\{x_{k-1}, I_{k-1}\}$ и $\{x_k, I_k\}$ последовательности (6.1) выполнено неравенство

$$|I_k| < |I_{k-1}|, \quad (6.3)$$

то $F(x_k) < F(x_{k-1})$.

Доказательство. Так как попадание из A2 в A1.3 расширяет подходящее индексное множество, а попадание из A2 в A3 не меняет его мощности, то неравенство (6.3) может иметь место, только если после A2 реализуется путь A1.1 → A1.2 → A1.4 → A2. Как было показано в лемме 6.1, тогда $s_0 \neq \mathbb{O}$ и из (3.5) следует, что $F(x_k) < F(x_{k-1})$. ■

Пусть в последовательности (6.1) образовался цикл, состоящий из попарно различных индексных множеств

$$I_j, I_{j+1}, \dots, I_{j+\tau-1}, \text{ а } I_{j+\tau} = I_j.$$

Лемма 6.3. Цикл характеризуется тем, что для одной и той же квазистационарной точки $x = x_j$ на k -й итерации цикла происходит изменение подходящего для нее индексного множества $I_k \subset I(x)$ путем замены индекса p_k на индекс $l_k \in I(x)$. При этом $\tau > 1$.

Доказательство. Заметим, что согласно алгоритму вектор x_k по I_k определяется однозначно, поэтому $x_j = x_{j+\tau}$, следовательно, $F(x_j) = F(x_{j+\tau})$. А так как последовательность $\{F(x_k)\}$ невозрастающая, то

$$F(x_j) = F(x_{j+1}) = \dots = F(x_{j+\tau-1}). \quad (6.4)$$

Из этого равенства и леммы 6.2 следует, что мощность индексного множества I_k в цикле не может уменьшиться. А тогда она не может и увеличиться, так как $I_j = I_{j+\tau}$. Следовательно,

$$|I_j| = |I_{j+1}| = \dots = |I_{j+\tau-1}|. \quad (6.5)$$

Покажем теперь, что $x_j = x_{j+1} = \dots = x_{j+\tau-1}$. Попутно покажем, что $\tau > 1$. Рассмотрим k -ю итерацию цикла. Точка x_k — квазистационарна и из A2 следует, что $J'_k = I_k \setminus \{p_k\}$. Возможны две ситуации.

1) На k -й итерации цикла реализуется переход A2 → A1. Согласно лемме 6.1 в A1.1 $s_0 \neq \mathbb{O}$. В силу (6.4) должен реализоваться алгоритм A1.3 с $\sigma_0 = 0$, т.е. $z_1 = z_0 = x_k$. Кроме того, по лемме 6.1 выбранный в A1.2 индекс l_0 отличен от p_k . Далее, в силу (6.5) и (6.4) должен реализоваться алгоритм A1.4 с $s_1 = \mathbb{O}$. Значит, $x_{k+1} = z_1 = x_k$ и $I_{k+1} = J'_k \cup \{l_0\} \neq I_k$.

2) На k -й итерации цикла реализуется алгоритм A3. Тогда в силу (6.4) шаг t_k в выбранном направлении r_k равен нулю, поэтому $z_0 = x_k$ и $l_k \neq p_k$. Далее по алгоритму выполняется A1.1. В силу (6.4) и (6.5) вектор s_0 должен равняться нулю. Значит, сразу реализуется A1.4 и $x_{k+1} = z_0 = x_k$, а $I_{k+1} = J'_k \cup \{l_k\} \neq I_k$.

В обоих случаях $I_{k+1} \neq I_k$, поэтому $\tau > 1$. ■

Теорема 6.1. Метод перебора граней с использованием правила наименьшего индекса конечен.

Доказательство. Введем индексные множества

$$\Pi = \bigcup_{k \in 0:\tau-1} I_{k+j}, \quad \Pi_0 = \bigcap_{k \in 0:\tau-1} I_{k+j}, \quad Q = \Pi \setminus \Pi_0.$$

Очевидно, что в Q содержатся «подвижные» индексы цикла. Положим

$$\lambda = \max\{k \mid k \in Q\}.$$

Каждый «подвижный» индекс должен хотя бы один раз исключиться из некоторого множества I_k и хотя бы один раз включиться в другое множество I_k из цикла. Напомним, что цикл содержит не менее двух пар.

Пусть индекс λ исключился из множества I_q . Согласно правилу минимального индекса $P_q = \{\lambda\}$, т. е.

$$u_q[\lambda] < 0, \quad u_q[I_q \setminus \{\lambda\}] \geq 0.$$

Действительно, если бы были еще отрицательные компоненты $u_q[i]$, то так как $i < \lambda$, то выбрался бы индекс i , что не так.

Пусть индекс λ включился в $I_{\mu+1}$. Рассмотрим две возможности.

1. Реализуется АЗ. Тогда $L_\mu = \{\lambda\}$. Действительно, так как $t_\mu = 0$, то $L_\mu \subset I(x)$. Если бы нашлся отличный от λ индекс $i \in I(x)$, для которого $A[i, N] \cdot r_\mu[N] < 0$, то согласно правилу минимального индекса выбрался бы не индекс λ , а другой индекс. Далее, так как $I_q \subset I(x)$, то

$$A[I_q \setminus \{\lambda\}, N] \cdot r_\mu[N] \geq 0, \quad A[\lambda, N] \cdot r_\mu[N] < 0. \quad (6.6)$$

Из приведенных неравенств и представления градиента (3.3) следует, что

$$g(x)[N] \cdot r_\mu[N] = u_q[I_q] \cdot A[I_q, N] \cdot r_\mu[N] > 0.$$

Вместе с тем r_μ — направление строгого убывания F в точке x , поэтому $g(x)[N] \cdot r_\mu[N] < 0$. Получили противоречие.

2. На μ -й итерации цикла реализуется А2→А1. Тогда из тех же рассуждений, что и в первом случае, получаем, что в А1.2 $L'_0 = \{\lambda\}$ и имеют место неравенства (6.6) с заменой r_μ на вектор $s_0[N]$. Повторяя предыдущие выкладки, получаем аналогичное противоречие, которое и доказывает отсутствие циклов. ■

§7. Метод перебора вершин

Применим метод ПГ к задаче линейного программирования

$$F(x) := \langle c, x \rangle \rightarrow \min_{x \in \Omega}, \quad (7.1)$$

$$\Omega = \{x[N] \mid A[M, N] \cdot x[N] \geq b[M]\}.$$

Будем считать, что множество Ω не содержит прямых, т. е. содержит хотя бы одну вершину.

Сразу заметим, что описанный ниже метод перебора вершин полностью совпадает с методом последовательного улучшения плана (ПУП), изложенным в работе [3]. Он же совпадает с двойственным симплекс-методом ([2]) для задачи ЛП в первой канонической форме:

$$F(x) := \langle c, x \rangle \rightarrow \min, \quad (7.2)$$

$$Ax = b, \quad x \geq 0.$$

Действительно, если для задачи (7.2) записать двойственную задачу и к последней применить метод ПУП, то вычисления полностью совпадают с двойственным симплекс-методом, хотя терминология в двойственном симплекс-методе другая.

Для задачи ЛП метод ПГ сильно упрощается. Сначала покажем, что подходящей точкой может быть только вершина x , при этом подходящую пару составляют $\{x, J\}$, где J — строчный базис вершины. Действительно, пусть $\{x, J\}$ — подходящая пара для задачи (7.1), т. е. $x \in \Omega$, $J \subseteq I(x)$ и матрица

$$G(J) = \begin{pmatrix} \mathbb{O}[N, N] & A^T[N, J] \\ A[J, N] & \mathbb{O}[J, J] \end{pmatrix}$$

является невырожденной. Отсюда следует, что матрица $A[J, N]$ квадратная невырожденная, поэтому x — вершина Ω , а J — ее строчный базис (см. §1).

Как известно, вершина является квазистационарной точкой, поэтому первая часть алгоритма ПГ А1, а именно, построение квазистационарной точки из подходящей сводится лишь к определению коэффициентов $u[J_k]$ из системы (3.4):

$$A^T[N, J_k] \cdot u[J_k] = c[N].$$

Если же известна обратная матрица $B_k[N, J_k] = (A[J_k, N])^{-1}$, то А1 сводится к вычислению $u_k[J_k] = c[N] \cdot B[N, J_k]$.

Обратимся к А2. Так как $D = \emptyset$, то матрица $G(J')$ вырожденная, поэтому далее работает алгоритм А3.

Таким образом, для ЛП метод ПГ заключается в переходе от одной вершины к другой с меньшим или равным значением целевой функции. Заметим, что на всех итерациях $|J_k| = |N|$.

Учитывая сказанное выше, опишем метод ПГ для задачи ЛП, используя матрицу $B_k[N, J_k] = (A[J_k, N])^{-1}$.

Пусть x_0 — вершина Ω , J_0 — ее строчный базис и известна $B_0[N, J_0] = (A[J_0, N])^{-1}$. Положим $k = 0$ и опишем переход от вершины x_k со строчным базисом J_k к вершине x_{k+1} со строчным базисом J_{k+1} . Он состоит из следующих операций.

А1. Вычисление двойственного вектора:

$$u_k[J_k] = c[N] \cdot B_k[N, J_k].$$

А2. Проверка вершины x_k на оптимальность. Если

$$u_k[J_k] \geq \emptyset,$$

то вершина x_k оптимальна. В противном случае определяем множество

$$P_k = \{i \in J_k \mid u_k[i] < 0\}$$

и выбираем любой индекс $p_k \in P_k$.

А3. Построение новой вершины. Спуск из вершины x_k производим в направлении $r_k[N] = B_k[N, p_k]$. Находим шаг в направлении $r_k[N]$. Он равен ∞ , если $Ar_k \geq \emptyset$, а иначе

$$t_k = \min \left\{ \frac{b[i] - A[i, N] \cdot x_k[N]}{A[i, N] \cdot r_k[N]} \mid i \in M : A[i, N] \cdot r_k[N] < 0 \right\}.$$

Если $t_k = \infty$, то согласно общему алгоритму исходная задача не имеет решения. На этом процесс заканчивается.

Если t_k конечно, то строим новую точку

$$x_{k+1} = x_k + t_k \cdot r_k.$$

Обозначим через L_k множество индексов из M , на которых реализуется минимум при вычислении шага t_k , и возьмем любое l_k из L_k . Полагаем $J_{k+1} = J_k \setminus \{p_k\} \cup \{l_k\}$. Очевидно, что

$$J_{k+1} \subseteq I(x_{k+1}), \quad |J_{k+1}| = |N|, \quad \text{rank } A[J_{k+1}, N] = |N|,$$

так что точка x_{k+1} — вершина, а J_{k+1} — ее строчный базис. При этом

$$F(x_{k+1}) \leq F(x_k) \quad \text{и, если } t_k > 0, \quad \text{то } F(x_{k+1}) < F(x_k).$$

На этом одна итерация заканчивается.

В силу теоремы 6.1 метод перебора вершин конечен, если индексы p_k и l_k выбирать по правилу наименьшего индекса:

$$p_k = \min\{i \mid i \in P_k\}, \quad l_k = \min\{j \mid j \in L_k\}.$$

§8. Численные примеры

Пример 1. Методом перебора граней решить задачу

$$\frac{1}{2}(x[1]^2 + x[2]^2) \rightarrow \min$$

при ограничениях

$$\begin{aligned} (1) \quad & x[1] \geq -2, \\ (2) \quad & -x[2] \geq -1, \\ (3) \quad & -3x[1] + 2x[2] \geq 5. \end{aligned}$$

Здесь $D = E$, $g(x) = x$. В качестве начальной подходящей пары возьмем вершину $z_0 = (-2, 1)$ и индексное множество $J_0 = \{1\}$.

А1. Решаем систему (3.4):

$$\begin{aligned} s[1] + w[1] &= 2, \\ s[2] &= -1, \\ s[1] &= 0, \end{aligned}$$

откуда получаем $s_0 = (0, -1)$ и точку минимума $x_1 = z_0 + s_0 = (-2, 0)$, которая удовлетворяет ограничениям, поэтому x_1 квази-стационарна. Переходим к А2.

A2. Так как $u[1] = -w[1] = -2$, то $x_1 = (-2, 0)$ не оптимальна, $P_1 = \{1\}$. Исключаем из J_0 индекс $p = 1$ и получаем новое J_0 , равное пустому множеству. Так как матрица $G(J_0) = E$ невырожденная, то пара $\{z_0, J_0\}$, где $z_0 = x_1 = (-2, 0)$, является подходящей. Снова переходим к A1.

A1. Решаем систему (3.4):

$$\begin{aligned} s[1] &= 2, \\ s[2] &= 0, \end{aligned}$$

откуда следует, что $s_0 = (2, 0)$, так что точка минимума равна $(0, 0)$. Эта точка не принадлежит Ω . Тогда в выбранном направлении s_0 определяем шаг: $\sigma_0 = \frac{1}{6}$. Получаем точку

$$z_1 = z_0 + \sigma_0 \cdot s_0 = \left(-\frac{5}{3}, 0\right)$$

и индексное множество $J_1 = \{3\}$. Это подходящая пара. Возвращаемся к A1.1, полагая $z_0 = \left(-\frac{5}{3}, 0\right)$, $J_0 = \{3\}$.

A1.1. Решаем систему (3.4):

$$\begin{aligned} s[1] - 3w[1] &= \frac{5}{3}, \\ s[2] + 2w[1] &= 0, \\ -3s[1] + 2s[2] &= 0, \end{aligned}$$

откуда получаем $s_0 = \left(\frac{20}{39}, \frac{10}{13}\right)$, $u_0[1] = -w_0[1] = \frac{5}{13}$. Точка минимума $z_0 + s_0 = \left(-\frac{15}{13}, \frac{10}{13}\right)$ удовлетворяет ограничениям, поэтому она квазистационарна. Переходим к A2.

A2. Так как $u[1] > 0$, то $x_* = \left(-\frac{15}{13}, \frac{10}{13}\right)$ является решением задачи.

Пример 2. Методом перебора граней решить задачу

$$\frac{1}{2}x[1]^2 + x[2] \rightarrow \min$$

при ограничениях

$$\begin{aligned} (1) \quad -x[1] - x[2] &\geq -1, \\ (2) \quad -x[2] &\geq 0. \end{aligned}$$

Здесь

$$D = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad c = (0, 1).$$

При любых x

$$g(x) = Dx + c = \begin{pmatrix} x[1] \\ 1 \end{pmatrix}.$$

В качестве начальной подходящей пары возьмем вершину $x_0 = (1, 0)$ и индексное множество $J_0 = \{1, 2\}$.

A1. Решаем систему (3.4):

$$\begin{aligned} s[1] - w[1] &= -1, \\ -w[1] - w[2] &= -1, \\ -s[1] - s[2] &= 0, \\ -s[2] &= 0, \end{aligned}$$

откуда получаем $s_0 = (0, 0)$, $w[1] = 1$, $w[2] = 0$ и точку минимума $x_0 = (1, 0)$. Так как точка x_0 удовлетворяет ограничениям, то она квазистационарна.

A2. Так как $u[1] = -w[1] = -1$, то x_0 не оптимальна. Исключаем из J_0 индекс $p = 1$ и получаем $J_1 = \{2\}$. Полагаем $z_0 = (1, 0)$. Пара $\{z_0, J_1\}$ является подходящей. Снова переходим к A1.

A1. Решаем систему (3.4):

$$\begin{aligned} s[1] &= -1, \\ -w[2] &= -1, \\ -s[2] &= 0, \end{aligned}$$

откуда следует, что $s_0 = (-1, 0)$, $w[2] = 1$, так что точка минимума равна $(0, 0)$. Эта точка принадлежит Ω , поэтому $x_1 = (0, 0)$ квазистационарна.

A2. Так как $u[2] = -1 < 0$, то x_1 не оптимальна. Исключаем из J_1 индекс $p = 2$ и получаем J_2 , равное пустому множеству. Так как матрица $G(J_2) = D$ вырожденная, то переходим к A3.

A3. Ищем направление спуска из системы

$$D[N, N] \cdot r[N] = \mathbb{O}, \quad A[2, N] \cdot r[N] = 1.$$

Получаем $r[1] = 0$, $r[2] = -1$. Ясно, что весь луч $x(t) = t \cdot r$, $t \geq 0$, удовлетворяет ограничениям, поэтому исходная задача не имеет решения.

ИТЕРАТИВНЫЕ МЕТОДЫ

В этой главе приводятся описания и обоснования некоторых методов, базирующихся на решении системы КТ, которые не вкладываются в ранее рассмотренные классы методов. Вопросы сходимости этих методов детально не обсуждаются, но приводятся ссылки на соответствующую литературу.

§1. Метод Дикина

1.1. Круг задач

Метод Дикина [9, 10] разработан для задачи ВКП с ограничениями, записанными в первой канонической форме:

$$F(x) := \frac{1}{2} \langle Dx, x \rangle + \langle c, x \rangle \rightarrow \min_{x \in \Omega} \quad (1.1)$$

$$\Omega = \{x \in \mathbb{R}^n \mid Ax = b, x \geq \mathbb{O}\}.$$

Метод годится для задач, относительно которых предполагается, что

а) множество планов содержит относительно внутреннюю точку, т. е. не пусто множество

$$\Omega_0 = \{x \in \mathbb{R}^n \mid Ax = b, x > \mathbb{O}\};$$

б) ранг матрицы $A[M, N]$ равен $|M|$;

в) целевая функция ограничена снизу на множестве планов.

Из предположений а) и в) следует, в частности, что задача (1.1) разрешима.

1.2. Условия оптимальности

Запишем условия КТ для задачи (1.1), введя новые переменные $u[M]$, $y[N]$:

$$Dx - A^T u - y = -c, \quad (1.2)$$

$$Ax = b, \quad (1.3)$$

$$x \geq \mathbb{O}, \quad y \geq \mathbb{O}, \quad (1.4)$$

$$x[i] \cdot y[i] = 0 \quad \forall i \in N. \quad (1.5)$$

Несколько преобразуем эти условия. Обозначим $g(x) := Dx + c$ и из равенства (1.2) выделим $y[N]$:

$$y[N] = g(x)[N] - u[M] \cdot A[M, N].$$

Введем в рассмотрение функцию

$$\Phi(x, u) := \sum_{j \in N} (x[j] \cdot y[j])^2.$$

Нетрудно видеть, что условие ортогональности (1.5) равносильно равенству нулю функции $\Phi(x, u)$. Так как условие (1.3) и первое неравенство (1.4) равносильны принадлежности вектора x множеству Ω , то справедливо следующее утверждение.

Теорема 1.1. Вектор $x_* \in \Omega$ является решением задачи (1.1) тогда и только тогда, когда существует вектор $u_*[M]$, такой, что

$$y_*[N] := g(x_*)[N] - u_*[M] \cdot A[M, N] \geq \mathbb{O}, \quad (1.6)$$

$$\Phi(x_*, u_*) = 0. \quad \blacksquare$$

1.3. Идея метода и его обоснование

Предложенный Дикиным метод заключается в минимизации функции $\Phi(x, u)$ на Ω при ограничении (1.6).

Схема одной итерации метода следующая. В фиксированной точке $x_k \in \Omega_0$ ищется u_k — точка минимума $\Phi(x_k, u)$ по u . Если $\Phi(x_k, u_k) = 0$, то в нуль обращается вектор

$$y_k[N] = g(x_k)[N] - u_k[M] \cdot A[M, N],$$

значит, x_k — решение исходной задачи. В противном случае строим новую точку $x_{k+1} = x_k + \lambda_k s_k$, где $s_k[N]$ — направление убывания $F(x)$ в точке x_k , а λ_k — шаг в направлении убывания, который

выбирается так, чтобы точка x_{k+1} принадлежала Ω_0 и значение $F(x_{k+1})$ было строго меньше $F(x_k)$.

Сначала выясним, как находить точку минимума $\Phi(x_k, u)$. Покажем, что $\Phi(x_k, u)$ является квадратичной функцией по u . Введем диагональную матрицу $\Lambda_k := \text{diag}(x_k[1], \dots, x_k[n])$ и положим $y_k(u) := g(x_k) - A^T u$. Очевидно, что

$$\Phi(x_k, u) = \langle \Lambda_k y_k(u), \Lambda_k y_k(u) \rangle.$$

Подставляя сюда выражение для $y_k(u)$ и пользуясь обозначением $g_k := g(x_k)$, получаем

$$\Phi(x_k, u) = \langle \Lambda_k A^T u, \Lambda_k A^T u \rangle - 2 \langle \Lambda_k A^T u, \Lambda_k g_k \rangle + \|\Lambda_k g_k\|^2.$$

Обозначим $B_k := A \Lambda_k^2 A^T$, $d_k := A \Lambda_k^2 g_k$, $\alpha_k := \|\Lambda_k g_k\|^2$. В этих обозначениях

$$\Phi(x_k, u) = \langle B_k u, u \rangle - 2 \langle d_k, u \rangle + \alpha_k.$$

Из вида матрицы B_k следует, что она неотрицательно определена. Более того, в силу предположения б) и условия $x_k \in \Omega_0$ матрица B_k положительно определена.

Таким образом, $\Phi(x_k, u)$ — строго выпуклая квадратичная функция от u , поэтому ее единственная точка минимума u_k находится из линейной системы

$$B_k \cdot u = d_k. \quad (1.7)$$

Обратимся к выбору направления $s_k[N]$. Напомним, что $y_k = g_k - A^T u_k$. Покажем, что если $\Phi_k > 0$, то вектор

$$s_k := -\Lambda_k^2 \cdot y_k \quad (1.8)$$

является направлением строгого убывания функции $F(x)$ в точке x_k . Действительно, пользуясь разложением квадратичной функции, имеем

$$F(x_k + \lambda s_k) - F(x_k) = \lambda \langle g_k, s_k \rangle + \frac{\lambda^2}{2} \langle D s_k, s_k \rangle.$$

Заметим, что

$$A s_k = -A \Lambda_k^2 (g_k - A^T u_k) = B_k \cdot u_k - d_k = \mathbb{0},$$

поэтому $\langle s_k, g_k \rangle = \langle s_k, y_k + A^T u_k \rangle = \langle s_k, y_k \rangle = -\Phi_k$. Теперь из разложения получаем

$$F(x_k + \lambda s_k) - F(x_k) = -\lambda \Phi_k + \frac{\lambda^2}{2} \langle D s_k, s_k \rangle. \quad (1.9)$$

Отсюда ясно, что $s_k[N]$ — направление строгого убывания целевой функции исходной задачи в точке x_k .

Займемся теперь выбором шага λ_k при $\Phi_k > 0$.

Утверждение 1.1. Если шаг выбирается по формуле

$$\lambda_k = \min \left(\frac{\alpha}{\sqrt{\Phi_k}}, \frac{\Phi_k}{\langle D s_k, s_k \rangle} \right), \quad (1.10)$$

где параметр $\alpha < 1$ и достаточно близок к единице, то

$$x_{k+1} := x_k + \lambda_k \cdot s_k \in \Omega_0, \quad (1.11)$$

$$F(x_{k+1}) - F(x_k) < 0. \quad (1.12)$$

Доказательство. Чтобы удовлетворить (1.12), найдем точку минимума $F(x)$ на луче $x(\lambda) = x_k + \lambda s_k$, $\lambda > 0$. Из представления (1.9) следует, что точка минимума равна

$$\lambda'_k = \frac{\Phi_k}{\langle D s_k, s_k \rangle}.$$

Если же $\langle D s_k, s_k \rangle = 0$, то $\lambda'_k = \infty$. Значит, для выполнения (1.12) необходимо, чтобы $\lambda_k \leq \lambda'_k$.

Обратимся к условию (1.11). Так как $A s_k = \mathbb{0}$, то

$$A x_{k+1} = b,$$

поэтому шаг подбирают так, чтобы $x_{k+1} > \mathbb{0}$. Согласно (1.8) $s_k[j] = -x_k^2[j] \cdot y_k[j]$, тогда

$$x_{k+1}[j] = x_k[j] (1 - \lambda_k x_k[j] \cdot y_k[j]) \quad \forall j \in N.$$

Отсюда

$$\frac{x_{k+1}[j]}{x_k[j]} - 1 = -\lambda_k x_k[j] \cdot y_k[j] \quad \forall j \in N.$$

Возведя обе части в квадрат и суммируя по j , получаем

$$\eta_k := \sum_{j \in N} \left(\frac{x_{k+1}[j]}{x_k[j]} - 1 \right)^2 = \lambda_k^2 \Phi_k.$$

Заметим, что если $x_{k+1}[j] \leq 0$ при некотором $j \in N$, то $\eta_k \geq 1$. Значит, для того чтобы $x_{k+1}[N] > 0$, достаточно выполнения неравенства $\eta_k < 1$. В связи с этим выбираем шаг λ_k так, чтобы

$$\lambda_k \leq \frac{\alpha}{\sqrt{\Phi_k}}, \quad \alpha < 1.$$

При этом параметр α полагаем достаточно близким к единице.

Таким образом, в качестве шага в выбранном направлении s_k следует выбрать наименьшее из чисел: $\lambda'_k, \frac{\alpha}{\sqrt{\Phi_k}}$. ■

1.4. Описание алгоритма

Метод заключается в построении последовательности точек $x_k \in \Omega_0$ и u_k до тех пор, пока $\Phi(x_k, u_k)$ не станет равной нулю (или достаточно малой).

Опишем k -й шаг алгоритма. Пусть $x_k \in \Omega_0$.

1. Вычисляем $g_k = Dx_k + c$ и решаем систему (1.7):

$$A\Lambda_k^2 A^T \cdot u = A\Lambda_k^2 g_k.$$

Пусть u_k — решение системы.

2. Находим $y_k = g_k - A^T u_k$, $\Phi_k = \sum_{i \in N} x_k^2[i] y_k^2[i]$. Если $\Phi_k = 0$, то x_k — решение исходной задачи. В противном случае переходим к п. 3.

3. Определяем направление спуска по формуле (1.8):

$$s_k[i] = -x_k^2[i] y_k[i] \quad \forall i \in N.$$

4. Вычисляем шаг λ_k по формуле (1.10).

5. Строим новое приближение $x_{k+1} = x_k + \lambda_k s_k$.

На этом описание одного шага алгоритма заканчивается.

Отметим ряд особенностей алгоритма.

Параметр α в процессе вычислений можно менять, т.е. устремлять к единице.

Наиболее трудоемкой частью алгоритма является решение системы в п. 1, поэтому метод Дикина удобно применять к задачам с большим числом переменных и с малым числом ограничений, так как порядок матрицы B_k равен числу ограничений.

Накопления ошибок в методе нет, так как каждое новое приближение можно рассматривать как начальное. Метод удобно использовать, когда известно какое-то приближение к решению из Ω_0 , которое и можно принять за начальное.

Из алгоритма следует, что $x_k \in \Omega_0$ при всех k и последовательность $\{F(x_k)\}$ строго убывает. В силу предположения в) $F(x_k)$ стремится к некоторому пределу. В линейном случае, при $D = 0$, нетрудно доказать утверждение о сходимости последовательности $\{\Phi_k\}$ к нулю. Действительно, в этом случае $\lambda_k = \frac{\alpha}{\sqrt{\Phi_k}}$ и в силу (1.9) $F(x_k) - F(x_{k+1}) = \alpha\sqrt{\Phi_k}$. Перейдя к пределу в последнем равенстве, получим требуемое.

Описание метода и некоторые результаты о сходимости имеются в работах [9, 10].

§2. Метод Хилдрета и д'Эзопо

2.1. Круг задач

Пусть матрица $C[M, M]$ неотрицательно определена и при этом ее диагональные элементы положительны: $C[i, i] > 0$. Пусть $M_1 \subset M := \{1, 2, \dots, m\}$. Рассмотрим задачу ВКП с ограничениями только на знаки некоторых переменных:

$$\Phi(u) := \frac{1}{2} \langle Cu, u \rangle + \langle d, u \rangle \rightarrow \min_{u[M_1] \geq 0}. \quad (2.1)$$

Такого вида задачи часто возникают на практике (далее об этом будет сказано подробнее).

Система КТ для этой задачи имеет вид

$$C[M, M] \cdot u[M] - v[M] = -d[M], \quad (2.2)$$

$$u[M_1] \geq 0, \quad v[M_1] \geq 0, \quad v[M_2] = 0, \quad (2.3)$$

$$u[i] \cdot v[i] = 0, \quad i \in M_1. \quad (2.4)$$

Здесь будет описан еще один подход к решению системы КТ, а именно, для решения линейной системы (2.2) Хилдрет и д'Эзопо [11] предложили использовать метод Зейделя.

2.2. Описание метода

Распишем i -ю строку системы (2.2) в виде

$$C[i, i] \cdot u[i] - v[i] = - \sum_{j=1}^{i-1} C[i, j] \cdot u[j] - \sum_{j=i+1}^m C[i, j] \cdot u[j] - d[i]. \quad (2.5)$$

Рассмотрим k -й шаг метода, т.е. уже найден вектор $u_k[M]$. Начиная с $i = 1$, из системы (2.5) будем определять компоненты $u_{k+1}[i]$, $v_{k+1}[i]$ следующим образом. Пусть уже построены

$$u_{k+1}[1], u_{k+1}[2], \dots, u_{k+1}[i-1], \quad v_{k+1}[1], v_{k+1}[2], \dots, v_{k+1}[i-1].$$

Находим $u_{k+1}[i]$ и $v_{k+1}[i]$ из (2.5), подставляя в первую сумму уже вычисленные $u_{k+1}[j]$, а во вторую сумму $u_k[j]$. Тогда для $u_{k+1}[i]$ и $v_{k+1}[i]$ получаем формулу

$$C[i, i] \cdot u_{k+1}[i] - v_{k+1}[i] = -\xi_k[i], \quad (2.6)$$

где

$$\xi_k[i] := \sum_{j=1}^{i-1} C[i, j] \cdot u_{k+1}[j] + \sum_{j=i+1}^m C[i, j] \cdot u_k[j] + d[i].$$

Значения $u_{k+1}[i]$ и $v_{k+1}[i]$ выбираются так, чтобы выполнялись знаковые условия (2.3) и условие ортогональности (2.4). Для этого полагаем:

$$\begin{cases} \text{при } i \in M_1 \\ \text{если } \xi_k < 0, & \text{то } u_{k+1}[i] = -\xi_k[i]/C[i, i], \quad v_{k+1}[i] = 0, \\ \text{если } \xi_k \geq 0, & \text{то } u_{k+1}[i] = 0, \quad v_{k+1}[i] = \xi_k[i]; \end{cases} \quad (2.7)$$

при $i \in M_2$

$$v_{k+1}[i] = 0, \quad u_{k+1}[i] = -\xi_k[i]/C[i, i]. \quad (2.8)$$

Проделав такую процедуру m раз, получим пару векторов

$$\{u_{k+1}, v_{k+1}\},$$

которые вместе с u_k удовлетворяют равенствам (2.6) при всех $i \in M$. Если оказалось, что $u_{k+1} = u_k$, то пара $\{u_{k+1}, v_{k+1}\}$ удовлетворяет системе КТ (2.2) – (2.4), т.е. u_k – решение задачи (2.1). На этом вычисления заканчиваются.

Если же $u_{k+1} \neq u_k$, то процесс повторяется. Вообще говоря, этот процесс *бесконечен*.

Отметим одно свойство последовательности $\{u_k\}$, построенной этим методом.

Теорема 2.1. *Последовательность $\{\Phi(u_k)\}$ строго убывает при условии $u_{k+1} \neq u_k$.*

Доказательство. Рассмотрим k -й шаг метода. В результате выполнения i -й операции по формуле (2.6) имеем вектор

$$z_i := (u_{k+1}[1], u_{k+1}[2], \dots, u_{k+1}[i], u_k[i+1], u_k[i+2], \dots, u_k[m]).$$

Ясно, что $z_0 = u_k$, $z_m = u_{k+1}$. Покажем, что $\Phi(z_i) \leq \Phi(z_{i-1})$. Заметим, что у вектора $h = z_i - z_{i-1}$ имеется лишь одна ненулевая компонента $h[i] = u_{k+1}[i] - u_k[i]$. Пользуясь разложением квадратичной функции, имеем

$$\Phi(z_i) = \Phi(z_{i-1} + h) = \Phi(z_{i-1}) + \langle Cz_{i-1} + d, h \rangle + \frac{1}{2} \langle Ch, h \rangle.$$

Заметим, что

$$\frac{1}{2} \langle Ch, h \rangle = \frac{1}{2} h^2[i] \cdot C[i, i]. \quad (2.9)$$

Нетрудно видеть, что

$$\begin{aligned} \eta &:= C[i, M] \cdot z_{i-1}[M] + d[i] = \\ &= \sum_{j=1}^{i-1} C[i, j] \cdot u_{k+1}[j] + C[i, i] \cdot u_k[i] + \sum_{j=i+1}^m C[i, j] \cdot u_k[j] + d[i] = \\ &= \xi_k[i] + C[i, i] \cdot u_k[i]. \end{aligned}$$

Тогда из (2.6) следует $\eta = v_{k+1}[i] - C[i, i](u_{k+1}[i] - u_k[i])$, т.е. $\langle Cz_{i-1} + d, h \rangle = h[i] \cdot v_{k+1}[i] - C[i, i] \cdot h^2[i]$. Теперь, учитывая (2.9), получаем

$$\begin{aligned} \Phi(z_i) - \Phi(z_{i-1}) &= h[i] \cdot v_{k+1}[i] - \frac{1}{2} C[i, i] \cdot h^2[i] = \\ &= u_{k+1}[i] \cdot v_{k+1}[i] - u_k[i] \cdot v_{k+1}[i] - \frac{1}{2} C[i, i] \cdot h^2[i]. \quad (1) \end{aligned}$$

Согласно (2.7) и (2.8) $u_{k+1}[i] \cdot v_{k+1}[i] = 0$, $u_k[i] \cdot v_{k+1}[i] \geq 0$, а так как $C[i, i] > 0$, то $\Phi(z_i) - \Phi(z_{i-1}) \leq 0$, причем равенство возможно лишь тогда, когда $u_{k+1}[i] = u_k[i]$. Теперь ясно, что

$$\Phi(u_{k+1}) = \Phi(z_m) \leq \Phi(z_{m-1}) \leq \dots \leq \Phi(z_0) = \Phi(u_k).$$

Заметим, что из равенства $\Phi(u_{k+1}) = \Phi(u_k)$ следует равенство $u_{k+1} = u_k$. ■

В работе [11] рассматривается задача, в которой M_2 пусто и множество Ω содержит внутреннюю точку. Для такой задачи доказано, что

$$\lim_{k \rightarrow \infty} \Phi(u_k) = \inf_{u \geq \Phi} \Phi(u).$$

2.3. Применение метода к задачам с положительно определенной матрицей D

Круг задач, которые приводятся к задаче (2.1), достаточно широк. Рассмотрим задачу ВКП

$$F(x) := \frac{1}{2} \langle Dx, x \rangle + \langle c, x \rangle \rightarrow \min_{x \in \Omega}, \quad (2.10)$$

в которой матрица D положительно определена, а множество Ω определяется системой неравенств и равенств:

$$\begin{aligned} A[M_1, N] \cdot x[N] &\geq b[M_1], \\ A[M_2, N] \cdot x[N] &= b[M_2]. \end{aligned} \quad (2.11)$$

Такая задача выпуклого квадратичного программирования относится к строго выпуклым.

Упражнение. Привести пример строго выпуклой задачи ВКП, у которой матрица D является вырожденной.

Заметим, что отсутствие ограничений на знак переменной не умаляет общности задачи (2.10), так как эти ограничения включаются в ограничения неравенства (2.11).

Далее от матрицы $A[M, N]$ ($M = M_1 \cup M_2$) потребуются, чтобы она не содержала чисто нулевых строк.

Так как матрица D обратима, то задачу (2.10) можно привести к задаче ВКП с ограничениями только на знак переменных. Это есть

известное преобразование Хилдрета (см., например, [3, 11]). Покажем, как это сведение сделать совсем просто из условий Куна — Таккера.

Запишем условия КТ для задачи (2.10):

$$D[N, N] \cdot x[N] - A^T[N, M] \cdot u[M] = -c[N], \quad (2.12)$$

$$A[M, N] \cdot x[N] - v[M] = b[M], \quad (2.13)$$

$$u[M_1] \geq 0, \quad v[M_1] \geq 0, \quad v[M_2] = 0, \quad \langle u, v \rangle = 0.$$

Исключим переменную x из системы (2.12), (2.13). Согласно (2.12)

$$x[N] = D^{-1}[N, N](A^T[N, M] \cdot u[M] - c[N]). \quad (2.14)$$

Подставив вектор x в (2.13), получим

$$AD^{-1}A^T u - v = b + AD^{-1}c.$$

Введем обозначения $C := AD^{-1}A^T$, $d = -b - AD^{-1}c$. В этих обозначениях система КТ примет вид

$$C[M, M] \cdot u[M] - v[M] = -d[M],$$

$$u[M_1] \geq 0, \quad v[M_1] \geq 0, \quad v[M_2] = 0,$$

$$\langle u, v \rangle = 0.$$

Нетрудно видеть, что эта система совпадает с (2.2) — (2.4), т.е. является системой КТ для задачи (2.1).

Осталось заметить, что если у матрицы $A[M, N]$ нет чисто нулевых строк, то диагональные элементы матрицы C положительны. Действительно, матрица D^{-1} положительно определена, поэтому $C[i, i] = A[i, N] \cdot D^{-1}[N, N] \cdot A^T[N, i] > 0$.

Таким образом, для решения задачи (2.10) требуется обратить матрицу D , решить задачу (2.1) и воспользоваться формулой (2.14).

Есть ряд задач, для которых этот способ решения удобен. К таким задачам относятся, в частности, те, у которых матрица D легко обрабатывается и без потери точности, т.е. матрица D достаточно хорошо обусловлена.

Особо интересен случай, когда исходная задача имеет мало ограничений, так как размерность задачи (2.1) определяется числом ограничений в исходной задаче (2.10).

Заметим, что задачу (2.1) можно решать и другим методом, например методом дополнительного базиса (см. гл. 2).

Приведем ответы к упражнениям из главы 1.

2.1. Нет! Контрпримером является функция $F(x) = e^{-x}$ на $(-\infty, +\infty)$.

2.2. Рассмотрим экстремальную задачу

$$F(x) = \frac{1}{2} \langle Dx, x \rangle \rightarrow \min_{x \in \mathbb{R}^n}.$$

Ясно, что решением задачи является вектор x_0 , для которого $\langle Dx_0, x_0 \rangle = 0$. Для такого вектора $F'(x_0) = Dx_0 = 0$.

3.1. Все $c[i]$ равны между собой.

3.2. Пусть J — множество индексов минимальных компонент вектора $c[N]$. Решением является любой план x_* , у которого $x_*[N \setminus J] = 0$.

4.3. Запишем для этой задачи систему Куна — Таккера:

$$\begin{aligned} 8x[1] - u &= -3, \\ \alpha^2 x[2] - \alpha u &= -2, \\ x[1] + \alpha x[2] &= 1. \end{aligned}$$

Очевидно, что при $\alpha = 0$ эта система несовместна. При $\alpha \neq 0$ получаем

$$x[1] = \frac{2(1-\alpha)}{9\alpha}, \quad x[2] = \frac{1-x[1]}{\alpha}, \quad u = 8x[1] + 3.$$

Итак, исходная задача ВКП разрешима при всех $\alpha \neq 0$.

4.4. Задача имеет вид

$$F(x) := x[1]^2 + \frac{1}{2}x[2]^2 + x[3]^2 - x[1] + 2x[2] \rightarrow \min$$

при ограничениях

$$\begin{aligned} x[1] - 3x[2] - x[3] &= 3, \\ -2x[1] + x[2] - 3x[3] &\geq -1, \\ x[2] &\geq 0. \end{aligned}$$

1. Базара М., Шетти К. Нелинейное программирование. М.: Мир, 1982. 584 с.
2. Булавский В. А., Звягина Р. А., Яковлева М. А. Численные методы линейного программирования. М.: Наука, 1977. 368 с.
3. Гавурич М. К., Малоземов В. Н. Экстремальные задачи с линейными ограничениями. Л.: Изд-во Ленингр. ун-та, 1984. 176 с.
4. Данциг Дж. Б. Линейное программирование, его обобщение и применение. М.: Прогресс, 1966. 600 с.
5. Даугавет В. А. Модификация метода Вулфа // Журн. вычисл. математики и мат. физики. 1981. Т. 21. № 2. С. 504–508.
6. Даугавет В. А. Метод дополнительного базиса в квадратичном программировании // Вестн. С.-Петербург. ун-та. 1992. № 15. С. 19–24.
7. Даугавет В. А., Лазарев А. В. Развитие метода Данцига в квадратичном программировании // Журн. вычисл. математики и мат. физики. 1986. Т. 26. № 3. С. 430–438.
8. Даугавет В. А., Салих Т. М. О методе дополнительного базиса для задач выпуклого квадратичного программирования // Журн. вычисл. математики и мат. физики. 1992. Т. 32. № 12. С. 1981–1992.
9. Дикин И. И. Итеративное решение задач линейного и квадратичного программирования // Докл. АН СССР. 1967. Т. 174. № 4. С. 747–748.
10. Дикин И. И., Зоржальцев В. Итеративное решение задач математического программирования. Новосибирск: Наука, 1980. 142 с.
11. Кюнц Г. П., Крелле В. Нелинейное программирование. М.: Советское радио, 1965. 304 с.
12. Пшеничный Б. Н. Метод линеаризации. М.: Наука, 1983. 136 с.
13. Пшеничный Б. Н., Данилин Ю. М. Численные методы в экстремальных задачах. М.: Наука, 1975. 320 с.
14. Романовский И. В. Алгоритмы решения экстремальных задач. М.: Наука, 1977. 352 с.
15. Фаддеев Д. К., Фаддеева В. Н. Вычислительные методы линейной алгебры. М.: Физматгиз, 1960. 656 с.
16. Best M. J. Equivalence of some quadratic programming algorithms // Math. Progr. 1984. Vol. 30. P. 71–87.
17. Bland R. G. New finite pivoting rules for the simplex method // Math. Operat. Res. 1977. Vol. 2. N 2. P. 103–107.
18. Chang Y., Cottle W. Least-index resolution of degeneracy in quadratic programming // Math. Progr. 1980. Vol. 18. P. 127–137.
19. Fletcher R. A general quadratic programming algorithm // J. Inst. Math. Appl. 1971. Vol. 7. P. 76–91.

20. Gill P. E., Murray W. Numerically stable methods for quadratic programming // Math. Progr. 1978. Vol. 14. P. 349–372.

21. Van de Panne C., Whinston A. The symmetric formulation of simplex method for quadratic programming // Econometrica. 1969. Vol. 37. P. 507–527.

22. Wolfe P. The simplex method for quadratic programming // Econometrica. 1959. Vol. 27. P. 382–398.

ОГЛАВЛЕНИЕ

Основные обозначения	3
Введение	5
ГЛАВА 1. Необходимые сведения из теории линейного и квадратичного программирования	11
§1. Стандартные формы записи задач квадратичного программирования	—
§2. Признаки разрешимости задачи выпуклого квадратичного программирования	14
§3. Двойственность в линейном программировании	15
§4. Условия оптимальности в форме Куна–Таккера	17
§5. Дополнительный и экстремальный базисы системы Куна–Таккера	22
§6. Обобщение симплекс-метода	27
§7. Способ построения начального базисного плана	31
§8. Пересчет обратной базисной матрицы и базисного решения	34
ГЛАВА 2. Базисные методы	37
§1. Метод полного перебора дополнительных базисов	—
§2. Решение линейной системы со знаковыми ограничениями на переменные	38
§3. Метод дополнительного базиса	40
3.1. Общая схема метода	—
3.2. Описание метода	41
3.3. Алгоритм метода с обратными матрицами	46
3.4. Обоснование метода	49
3.5. Численные примеры	59
§4. Базисные методы Вулфа и Лемке	63
§5. Метод Данцига	64
5.1. Постановка задачи и схема метода	—
5.2. Вспомогательные утверждения	66
5.3. Описание метода	68
5.4. Метод Данцига с обратными матрицами	73
5.5. Обоснование метода	76
5.6. О конечности метода	79
5.7. Метод Данцига в линейном случае	—
5.8. Численные примеры	81
ГЛАВА 3. Метод перебора граней	86
§1. Идея метода перебора граней	—
§2. Характеристика квазистационарной точки	89
§3. Описание и обоснование метода перебора граней	91
3.1. Подходящая пара	—
3.2. Выбор направления спуска	94
3.3. Описание метода	95
3.4. Обоснование метода	98
§4. Численная реализация метода перебора граней	99
§5. Алгоритм перебора граней с обратными матрицами	104

299 =

§6. О конечности метода перебора граней	105
§7. Метод перебора вершин	109
§8. Численные примеры	111
ГЛАВА 4. Итеративные методы	114
§1. Метод Дикина	-
1.1. Круг задач	-
1.2. Условия оптимальности	-
1.3. Идея метода и его обоснование	115
1.4. Описание алгоритма	118
§2. Метод Хилдрета и д"Эзопо	119
2.1. Круг задач	-
2.2. Описание метода	120
2.3. Применение метода к задачам с положительно определенной матрицей D	122
Ответы к упражнениям	124
Указатель литературы	125

1059

175/48655

902

Учебное издание

Валентина Александровна Даугавет

**ЧИСЛЕННЫЕ МЕТОДЫ
КВАДРАТИЧНОГО ПРОГРАММИРОВАНИЯ**

Учебное пособие

Редактор *И. Н. Рязанова*

Корректор *Т. Ф. Шпагина*

Художественный редактор *Е. И. Егорова*

Лицензия ИД № 05679 от 24.08.2001

Подписано в печать 18.05.2004. Формат 60 × 84¹/₁₆.

Печать офсетная. Усл. печ. л. 7,44. Заказ **247**

Издательство СПбГУ. 199034, Санкт-Петербург, Университетская наб., 7/9.

Тел. (812)328-77-63; факс (812)328-44-22

E-mail: books@dk2478.spb.edu

www.unipress.ru

ЦОП типографии Издательства СПбГУ.

199061, С.-Петербург, Средний пр., 41.