Двойственный симплекс-метод (обоснование и реализация)

А. О. Махорин*

Август 2016 г.

Аннотация

В данной пояснительной записке рассмотрено математическое обоснование модифицированного двойственного симплекс-метода, предназначенного для решения задач линейного программирования, а также приведено описание программных модулей, входящих в состав пакета GLPK, которые реализуют указанный метод на языке программирования ANSI C 89.

Данная работа является логическим продолжением работы [1]. Здесь используются те же основные понятия и обозначения, и кроме того, данная работа заимствует из [1] вывод многих формул и соотношений.

*Кафедра прикладной информатики, Московский авиационный институт. E-mail: `<mao@gnu.org>`.

Содержание

1 Общие сведения

1.1 Рабочий формат ЛП-задачи

В рассматриваемой реализации используется тот же *рабочий формат* задачи линейного программирования (ЛП-задачи), что и в [1]:

$$z = c^T x + c_0 \to \min \tag{1.1}$$

$$Ax = b \tag{1.2}$$

$$l \le x \le u \tag{1.3}$$

где $x = (x_k)$ — вектор переменных, z — целевая функция, $c = (c_k)$ — вектор коэффициентов целевой функции, c_0 — постоянный член целевой функции, $A = (a_{ik})$ — матрица коэффициентов ограничений, $b = (b_i)$ — вектор правых частей ограничений, $l = (l_k)$ — вектор нижних границ переменных, $u = (u_k)$ — вектор верхних границ переменных.

Если нижняя (верхняя) граница переменной x_k отсутствует, то формально считается, что $l_k = -\infty$ ($u_k = +\infty$). Переменная, у которой отсутствуют обе границы, называется *свободной (неограниченной) переменной*. Если $l_k = u_k$, то переменная x_k считается *фиксированной*.

В дальнейшем предполагается, что число переменных равно n, а число ограничений-равенств (1.2) равно m, поэтому матрица A имеет m строк и n столбцов. Также предполагается, что $\operatorname{rank}(A) = m$, т. е. что A имеет полный строчный ранг (все строки указанной матрицы линейно независимы), откуда, в частности, следует, что $m \le n$.

1.2 Базисные решения прямой задачи

Базисное решение ЛП-задачи (1.1)—(1.3) однозначно определено, если указано, какие переменные являются базисными, а какие небазисными, и для каждой небазисной переменной, имеющей обе нижнюю и верхнюю границы, дополнительно указано, какая из этих двух границ активна.[1]

Пусть Π — подходящая перестановочная матрица такая, что:

$$\Pi x = \begin{pmatrix} x_B \\ x_N \end{pmatrix}, \ \Pi c = \begin{pmatrix} c_B \\ c_N \end{pmatrix}, \ \Pi l = \begin{pmatrix} l_B \\ l_N \end{pmatrix}, \ \Pi u = \begin{pmatrix} u_B \\ u_N \end{pmatrix}, \tag{1.4}$$

где $x_B = [(x_B)_i]$ — вектор базисных переменных, $x_N = [(x_N)_j]$ — вектор небазисных переменных, $c_B = [(c_B)_i]$ и $c_N = [(c_N)_j]$ — векторы коэффициентов целевой функции, $l_B = [(l_B)_i]$ и $l_N = [(l_N)_j]$ — векторы нижних границ, $u_B = [(u_B)_i]$ и $u_N = [(u_N)_j]$ — векторы верхних границ, соответственно, базисных и небазисных переменных.

Так как столбцы матрицы коэффициентов ограничений A находятся во взаимно однозначном соответствии с переменными $x = (x_k)$, то разбиение переменных на базисные и небазисные посредством матрицы Π порождает соответствующее разбиение столбцов матрицы A:

$$(B \,|\, N) = A\Pi^T, \tag{1.5}$$

[1]См. более подробное изложение материала данного подраздела в [1, подразд. 1.2 и 1.3].

где B — невырожденная матрица порядка m, составленная из столбцов A, соответствующих базисным переменным, и называемая *базисной матрицей* (или просто *базисом*), N — $m \times (n-m)$-матрица, составленная из столбцов A, соответствующих небазисным переменным.

Базисная матрица B предполагается невырожденной, что позволяет явно выразить целевую функцию и базисные переменные через небазисные переменные и тем самым записать исходную ЛП-задачу (1.1)—(1.3) в *преобразованном виде* для заданного базиса:

$$z = d^T x_N + d_0 \to \min \tag{1.6}$$

$$x_B = \Xi x_N + g \tag{1.7}$$

$$l_B \le x_B \le u_B \tag{1.8}$$

$$l_N \le x_N \le u_N \tag{1.9}$$

где

$$\Xi = -B^{-1}N, \tag{1.10}$$

$$g = B^{-1}b, \tag{1.11}$$

$$d = c_N + \Xi^T c_B = c_N - N^T \pi, \tag{1.12}$$

$$\pi = B^{-T} c_B, \tag{1.13}$$

$$d_0 = \pi^T b + c_0. \tag{1.14}$$

Здесь $\Xi = (\xi_{ij})$ — $m \times (n-m)$-матрица, называемая здесь *симплекс-таблицей*, $g = (g_i)$ — m-вектор, называемый *преобразованным вектором правых частей ограничений*, $d = (d_j)$ — $(n-m)$-вектор *относительных оценок небазисных переменных*, $\pi = (\pi_i)$ — m-вектор *симплексных множителей*, d_0 — постоянный член преобразованной целевой функции.

Пусть $f = (f_j)$ — $(n-m)$-вектор значений активных границ небазисных переменных, т. е. $f_j = (l_N)_j$ или $f_j = (u_N)_j$. Тогда значения переменных в базисном решении равны:

$$x = \Pi^T \begin{pmatrix} x_B \\ x_N \end{pmatrix} = \Pi^T \begin{pmatrix} \beta \\ f \end{pmatrix}, \tag{1.15}$$

где

$$\beta = (\beta_i) = -B^{-1}Nf + B^{-1}b = \Xi f + g \tag{1.16}$$

есть m-вектор значений базисных переменных.

1.3 Условия оптимальности

Необходимые и достаточные условия оптимальности для ЛП-задачи в формате (1.1)—(1.3) имеют следующий вид:

$$Ax = b \tag{1.17}$$

$$l \le x \le u \tag{1.18}$$

$$A^T \pi + \lambda^+ + \lambda^- = c \tag{1.19}$$

$$\lambda^+ \ge 0, \quad \lambda^- \le 0 \tag{1.20}$$

$$(x_k - l_k)\lambda_k^+ = 0 \tag{1.21}$$

$$(x_k - u_k)\lambda_k^- = 0 \tag{1.22}$$

где $\pi = (\pi_i)$ — вектор множителей Лагранжа для ограничений-равенств (1.2), $\lambda^+ = (\lambda_k^+)$ и $\lambda^- = (\lambda_k^-)$ — векторы множителей Лагранжа для нижних и верхних границ переменных (ограничений-неравенств) (1.3), соответственно.

Условия (1.17) и (1.18) соответствуют ограничениям (1.2) и (1.3) исходной ЛП-задачи и означают, что оптимальное решение должно быть допустимым.

Условие (1.19) означает, что в оптимальной точке градиент целевой функции должен принадлежать ортогональному дополнению аффинного подпространства, образованного активными ограничениями из числа (1.17) и (1.18), а значит, должен быть линейной комбинацией нормалей к гиперплоскостям, соответствующим отдельным активным ограничениям, где множители Лагранжа суть коэффициенты этой линейной комбинации. Нарушение этого условия означает, что проекция градиента, и следовательно, проекция антиградиента целевой функции на аффинное подпространство активных ограничений отличны от нуля, поэтому возможно смещение точки под острым углом к антиградиенту целевой функции, при котором точка не покидает указанное аффинное подпространство, а значит, остается допустимой.

Условия (1.20) означают, что в оптимальной точке не существует допустимых направлений,[2] которые образуют острый угол с антиградиентом целевой функции. Нарушение этого условия означает, что возможно смещение точки под острым углом к антиградиенту, при котором точка смещается «вглубь» множества допустимых решений, т. е. когда одно или несколько активных ограничений перестают быть активными.

Условия (1.21) и (1.22), которые называются *условиями дополняющей нежесткости*, означают, что в оптимальной точке ненулевые множители Лагранжа могут иметь только активные ограничения.

1.4 Двойственная ЛП-задача

Задача, *двойственная* к заданной (*прямой*) задаче — это задача, которая имеет те же условия оптимальности, что и прямая задача, но в которой переменными являются множители Лагранжа, называемые в этом случае *двойственными переменными*.

Чтобы упростить выкладки, ограничимся случаем, когда все переменные прямой ЛП-задачи (1.1)—(1.3) имеют обе нижнюю и верхнюю границы (общий случай будет рассмотрен при изложении двойственного симплекс-метода). В этом случае соответствующая двойственная задача также является задачей линейного программирования и имеет следующий вид:

$$\zeta = b^T\pi + l^T\lambda^+ + u^T\lambda^- + c_0 \to \max \tag{1.23}$$

$$A^T\pi + \lambda^+ + \lambda^- = c \tag{1.24}$$

$$\lambda^+ \geq 0, \quad \lambda^- \leq 0 \tag{1.25}$$

где ζ — двойственная целевая функция, π, λ^+, λ^- — векторы двойственных переменных (множителей Лагранжа). Заметим, что переменные π_i не ограничены по знаку.[3]

Исходя из практических соображений определим вектор *комбинированных двойственных переменных* (множителей Лагранжа) $\lambda = (\lambda_k)$ как сумму:

$$\lambda = \lambda^+ + \lambda^-. \tag{1.26}$$

[2]Допустимые направления образуют выпуклый многогранный конус с вершиной в заданной допустимой точке (не обязательно вершине многогранного множества допустимых решений), гранями которого являются гиперплоскости активных ограничений.

[3]Так как они соответствуют ограничениям-равенствам прямой задачи.

Поскольку нижняя и верхняя границы переменной x_k прямой задачи не могут быть активны одновременно,[4] то из условий дополняющей нежесткости (1.21) и (1.22) следует, что $\lambda_k^+\lambda_k^- = 0$. Таким образом, с учетом ограничений на знаки (1.25) значение комбинированной двойственной переменной λ_k однозначно определяет значения двойственных переменных λ_k^+ и λ_k^-:

$$\begin{cases} \text{если } \lambda_k = 0, \text{ то } \lambda_k^+ = \lambda_k^- = 0; \\ \text{если } \lambda_k > 0, \text{ то } \lambda_k^+ = \lambda_k, \ \lambda_k^- = 0; \\ \text{если } \lambda_k < 0, \text{ то } \lambda_k^+ = 0, \ \lambda_k^- = \lambda_k. \end{cases} \tag{1.27}$$

Соотношения (1.26) и (1.27) позволяют записать двойственную ЛП-задачу (1.23)—(1.25) в эквивалентном виде:

$$\zeta = b^T\pi + l^T\lambda^+ + u^T\lambda^- + c_0 \to \max \tag{1.28}$$

$$A^T\pi + \lambda = c \tag{1.29}$$

где (комбинированные) двойственные переменные λ_k не ограничены по знаку.

1.5 Базисные решения двойственной задачи

Поскольку двойственная задача (1.23)—(1.25) является обычной задачей линейного программирования, то для нее можно определить понятие базиса и соответствующего базисного решения аналогично тому, как это сделано в подразд. 1.2 для прямой задачи (1.1)—(1.3). При этом важнейшим обстоятельством, лежащим в основе двойственного симплекс-метода, является то, что условия дополняющей нежесткости (1.21) и (1.22) позволяют установить естественное взаимно однозначное соответствие между всевозможными базисными решениями (как допустимыми, так и недопустимыми) прямой и двойственной задач.

Допустим, что с точки зрения прямой задачи переменная x_k базисная. В этом случае ее границы l_k и u_k являются неактивными ограничениями, и тогда из условий дополняющей нежесткости (1.21) и (1.22) следует, что соответствующие двойственные переменные λ_k^+ и λ_k^- должны быть равны нулю. Но с точки зрения двойственной задачи это означает, что нулевые границы (1.25) двойственных переменных λ_k^+ и λ_k^- являются активными ограничениями, и поэтому обе эти переменные небазисные. Допустим далее, что переменная x_k небазисная с активной нижней границей l_k. Тогда двойственная переменная λ_k^+ базисная, поскольку она может принимать ненулевые значения, а λ_k^- небазисная, так как верхняя граница u_k является неактивным ограничением. Аналогично, если переменная x_k небазисная с активной верхней границей u_k, то базисной будет двойственная переменная λ_k^-, а λ_k^+ будет небазисной, так как нижняя граница l_k является неактивным ограничением.

Допустим теперь, что для прямой задачи определен некоторый базис, т. е. указана подходящая перестановочная матрица Π, определяющая базисные x_B и небазисные x_N переменные (1.4), а также указано, какие именно границы (нижняя или верхняя) небазисных переменных являются активными. Заметим, что каждой переменной x_k прямой задачи соответствует (комбинированная) двойственная переменная $\lambda_k = \lambda_k^+ + \lambda_k^-$. Это означает, что для вектора λ (1.26) матрица Π определяет разбиение, идентичное разбиению (1.4):

$$\Pi\lambda = \Pi(\lambda_k^+ + \lambda_k^-) = \begin{pmatrix} \lambda_B \\ \lambda_N \end{pmatrix} = \begin{pmatrix} \lambda_B^+ + \lambda_B^- \\ \lambda_N^+ + \lambda_N^- \end{pmatrix}, \tag{1.30}$$

[4] Случай фиксированной переменной, т. е. когда $l_k = u_k$, можно рассматривать как предельный случай двустороннего ограничения $l_k - \varepsilon \le x_k \le u_k + \varepsilon$ при $\varepsilon \to 0$.

где двойственная переменная $(\lambda_B)_i = (\lambda_B^+)_i + (\lambda_B^-)_i$ соответствует базисной переменной $(x_B)_i$ и поэтому, как это было отмечено выше, является небазисной (в том смысле, что обе переменные $(\lambda_B^+)_i$ и $(\lambda_B^-)_i$ небазисные), а двойственная переменная $(\lambda_N)_j = (\lambda_N^+)_j + (\lambda_N^-)_j$ соответствует небазисной переменной $(x_N)_j$ и поэтому является базисной (в том смысле, что одна из переменных $(\lambda_B^+)_i$ и $(\lambda_B^-)_i$ базисная, а другая небазисная).

Итак, базис прямой ЛП-задачи (1.1)—(1.3) однозначно определяет соответствующий базис двойственной ЛП-задачи (1.28)—(1.29), в котором базисными являются двойственные переменные $(\lambda_N)_j$ и π_i,[5] а небазисными — двойственные переменные $(\lambda_B)_i$. Это позволяет явно выразить двойственную целевую функцию и двойственные базисные переменные через двойственные небазисные переменные и тем самым определить компоненты двойственного базисного решения подобно тому, как это сделано для прямого базисного решения (см. подразд. 1.2).

Выполним преобразование двойственной системы ограничений-равенств (1.29), для чего умножим матрицу Π слева на обе части этой системы:

$$\Pi A^T \pi + \Pi \lambda = \Pi c,$$

откуда с учетом (1.4), (1.5) и (1.30) получим:

$$\begin{pmatrix} B^T \\ N^T \end{pmatrix} \pi + \begin{pmatrix} \lambda_B \\ \lambda_N \end{pmatrix} = \begin{pmatrix} c_B \\ c_N \end{pmatrix},$$

что эквивалентно следующей системе:

$$\begin{cases} B^T \pi + \lambda_B = c_B \\ N^T \pi + \lambda_N = c_N \end{cases} \tag{1.31}$$

Базисная матрица B по определению базисного решения прямой ЛП-задачи является невырожденной, что позволяет разрешить первую подсистему (1.31) относительно вектора π:

$$\pi = B^{-T}(c_B - \lambda_B) = -B^{-T}\lambda_B + B^{-T}c_B. \tag{1.32}$$

Дальнейшая подстановка выражения для π из (1.32) во вторую подсистему (1.31) позволяет разрешить эту подсистему относительно λ_N:

$$\lambda_N = -N^T \pi + c_N = N^T B^{-T} \lambda_B - N^T B^{-T} c_B + c_N,$$

или окончательно:

$$\lambda_N = -\Xi^T \lambda_B + d, \tag{1.33}$$

где $\Xi = -B^{-1}N$, $d = c_N + \Xi^T c_B$.

В двойственном базисном решении значения всех небазисных двойственных переменных $(\lambda_B)_i$ равны нулю (так как соответствующие двойственные переменные $(\lambda_B^+)_i$ и $(\lambda_B^-)_i$ находятся на своих активных нулевых границах), а значения базисных переменных π_i и $(\lambda_N)_j$ (а значит, $(\lambda_N^+)_j$ и $(\lambda_N^-)_j$) определяются формулами (1.32) и (1.33). Таким образом, объединенный вектор значений всех двойственных переменных в двойственном базисном решении с учетом (1.30) равен:

$$\begin{pmatrix} \pi \\ \lambda \end{pmatrix} = \begin{pmatrix} I & \\ & \Pi^T \end{pmatrix} \begin{pmatrix} \pi \\ \lambda_B \\ \lambda_N \end{pmatrix} = \begin{pmatrix} I & \\ & \Pi^T \end{pmatrix} \begin{pmatrix} B^{-T} c_B \\ 0 \\ d \end{pmatrix}. \tag{1.34}$$

[5]Поскольку π_i соответствуют ограничениям-равенствам прямой ЛП-задачи, которые активны в любом базисном решении.

Сопоставляя преобразованную систему ограничений-равенств (1.33) и значения переменных (1.34) для двойственной задачи с преобразованной системой ограничений-равенств (1.7) и компонентами (1.10), (1.12) и (1.13) соответствующей прямой задачи, можно отметить, что:

двойственная симплекс-таблица $(-\Xi^T)$ есть транспонированная прямая симплекс-таблица Ξ (1.10), взятая с противоположным знаком;

значения базисных переменных π_i двойственной задачи равны значениям симплексных множителей (1.13) прямой задачи;

значения базисных переменных $(\lambda_N)_j$ двойственной задачи равны относительным оценкам d_j (1.12) прямой задачи.

Соответствующее преобразование целевой функции двойственной задачи рассмотрено в следующем подразделе.

1.6 Двойственный симплекс-метод

Двойственный симплекс-метод — это численный метод решения ЛП-задач, предложенный К. Лемке (C. E. Lemke) [2] в 1953 г. (см. также [3]). Данный метод является вариантом симплекс-метода, в основе которого лежит теория двойственности.

Уместно отметить, что двойственный симплекс-метод по существу представляет собой прямой симплекс-метод, примененный к двойственной задаче, но сформулированный в терминах прямой задачи.

Как и в случае прямой ЛП-задачи, множество допустимых решений двойственной ЛП-задачи представляет собой выпуклое многогранное множество, вершины которого соответствуют допустимым базисным решениям двойственной ЛП-задачи. Если движение вдоль некоторого ребра приводит к улучшению двойственной целевой функции, то выполняется переход в смежную вершину с лучшим значением двойственной целевой функции. Указанный процесс повторяется до тех пор, пока не будет найдена вершина, соответствующая оптимальному решению двойственной ЛП-задачи. А так как условия оптимальности прямой и двойственных ЛП-задач совпадают, то оптимальное решение двойственной ЛП-задачи будет определять соответствующее оптимальное решение прямой ЛП-задачи.

Перейдем к алгебраическому описанию двойственного симплекс-метода применительно к ЛП-задаче в формате (1.28)—(1.29) (См. аналогичное описание прямого симплекс-метода в [1].)

Допустим, что у нас имеется допустимое двойственное базисное решение (1.34), определяемое подходящей перестановочной матрицей Π. Тогда двойственную целевую функцию (1.28) можно записать в следующем виде:

$$\begin{aligned} \zeta &= b^T\pi + l^T\Pi^T\Pi\lambda^+ + u^T\Pi^T\Pi\lambda^- + c_0 = \\ &= b^T\pi + l_B\lambda_B^+ + l_N\lambda_N^+ + u_B\lambda_B^- + u_N\lambda_N^- + c_0 \end{aligned} \tag{1.35}$$

Движение вдоль ребра множества допустимых решений двойственной ЛП-задачи означает, что соответствующая этому ребру небазисная двойственная переменная начинает изменяться в допустимом направлении: увеличиваться, если в текущем базисе активной является ее нижняя граница, или уменьшаться, если в текущем базисе активной является ее верхняя граница. При этом нас интересуют те небазисные двойственные переменные, изменение которых в допустимом направлении приводит к улучшению (увеличению) двойственной целевой функции.

В разд. 1.5 было отмечено, что из двух переменных $(\lambda_N^+)_j$ и $(\lambda_N^-)_j$ одна из них всегда базисная, а другая небазисная. Допустим, что небазисной является переменная $(\lambda_N^+)_j$, которая начинает возрастать, так как в текущем базисе активной является ее нижняя (нулевая) граница. Из (1.32) и (1.33) следует, что остальные двойственные переменные не зависят от $(\lambda_N^+)_j$ и $(\lambda_N^-)_j$, поэтому вследствие сохранения равенства (1.24) при движении вдоль ребра другая (базисная) переменная $(\lambda_N^-)_j$ начнет пропорционально убывать. А поскольку $(l_N)_j \leq (u_N)_j$, то двойственная целевая функция при таком изменении переменных будет либо убывать (если $(l_N)_j < (u_N)_j$), либо оставаться неизменной (если $(l_N)_j = (u_N)_j$). Совершенно аналогично можно показать, что если небазисной переменной является $(\lambda_N^-)_j$, то ее изменение в допустимом направлении (уменьшение) также не может привести к увеличению двойственной целевой функции. Из сказанного следует, что для выбора подходящей небазисной двойственной переменной достаточно ограничиться небазисными переменными $(\lambda_B^+)_i$ и $(\lambda_B^-)_i$.

Заметим далее, что при движении вдоль ребра вплоть до смежной вершины состав базисных и небазисных переменных не изменяется. Это, в частности, означает, что при изменении какой-либо небазисной переменной $(\lambda_B^+)_i$ или $(\lambda_B^-)_i$ если переменная $(\lambda_N^+)_j$ базисная, то $(\lambda_N^-)_j$ будет оставаться небазисной (равной нулю), и наоборот. Полагая состав базисных и небазисных переменных неизменным и используя соотношение (1.26), представим двойственную целевую функцию (1.35) в виде:

$$\zeta = b^T\pi + f^T\lambda_N + l_B\lambda_B^+ + u_B\lambda_B^- + c_0, \tag{1.36}$$

где $f_j = (l_N)_j$, если $(\lambda_N^+)_j$ базисная, а $(\lambda_N^-)_j$ небазисная, и $f_j = (u_N)_j$, если $(\lambda_N^+)_j$ небазисная, а $(\lambda_N^-)_j$ базисная. Подставим теперь в (1.36) выражение для π из (1.32) и выражение для λ_N из (1.33):

$$\begin{aligned}\zeta &= b^T(-B^{-T}\lambda_B + B^{-T}c_B) + f^T(-\Xi^T\lambda_B + d) + l_B\lambda_B^+ + u_B\lambda_B^- + c_0 = \\ &= -(f^T\Xi^T + b^TB^{-T})\lambda_B + l_B\lambda_B^+ + u_B\lambda_B^- + b^TB^{-T}c_B + f^Td + c_0 = \\ &= -(\Xi f + B^{-1}b)^T\lambda_B + l_B\lambda_B^+ + u_B\lambda_B^- + (B^{-1}b)^Tc_B + f^Td + c_0 = \\ &= -\beta^T\lambda_B + l_B\lambda_B^+ + u_B\lambda_B^- + g^Tc_B + f^Td + c_0\end{aligned}$$

откуда, выполняя подстановку $\lambda_B = \lambda_B^+ + \lambda_B^-$, окончательно получим:

$$\zeta = (l_B - \beta)^T\lambda_B^+ + (u_B - \beta)^T\lambda_B^- + \zeta_0, \tag{1.37}$$

где $\beta = \Xi f + g$, $g = B^{-1}b$, $\zeta_0 = g^Tc_B + f^Td + c_0$.

В формуле (1.37) коэффициенты $(l_B)_i - \beta_i$ и $(u_B)_i - \beta_i$ представляют собой *относительные оценки* соответствующих двойственных небазисных переменных $(\lambda_B^+)_i$ и $(\lambda_B^+)_i$. Эти оценки показывают суммарное (т. е. прямое и косвенное через двойственные базисные переменные) влияние указанных двойственных небазисных переменных на двойственную целевую функцию.

В случае максимизации движение вдоль выбранного ребра должно приводить к увеличению (двойственной) целевой функции. Так как выбор ребра соответствует выбору (двойственной) небазисной переменный, это означает, что мы можем выбрать любую переменную $(\lambda_B^+)_i$, которая имеет положительную оценку $(l_B)_i - \beta_i$, так как изменение этой переменной в допустимом направлении соответствует ее возрастанию, либо любую переменную $(\lambda_B^-)_i$, которая имеет отрицательную оценку $(u_B)_i - \beta_i$, так как изменение этой переменной в допустимом направлении соответствует ее убыванию. Заметим теперь, что с точки зрения прямой задачи β есть вектор значений базисных переменных x_B в текущем базисе (1.16). Поэтому выбор двойственной небазисной переменной $(\lambda_B^+)_i$, для которой $(l_B)_i - \beta_i > 0$, с точки зрения прямой задачи соответствует выбору базисной переменной $(x_B)_i$, для которой имеет место нарушение нижней границы $\beta_i < (l_B)_i$. Аналогично, выбор двойственной небазисной переменной $(\lambda_B^-)_i$, для которой $(u_B)_i - \beta_i < 0$, с точки зрения прямой задачи соответствует выбору базисной переменной $(x_B)_i$, для которой имеет место нарушение верхней границы $\beta_i > (u_B)_i$.

Следует отметить, что рассмотренное правило выбора двойственной небазисной переменной автоматически учитывает общий случай, когда соответствующая базисная переменная прямой задачи $(x_B)_i$ не ограничена снизу и/или сверху. Действительно, если $(l_B)_i = -\infty$ (соответственно $(u_B)_i = +\infty$), то формально $(l_B)_i - \beta_i = -\infty < 0$ (соответственно $(u_B)_i - \beta_i = +\infty > 0$), поэтому $(\lambda_B^+)_i$ (соответственно $(\lambda_B^-)_i$) не будет подходящей двойственной небазисной переменной, обеспечивающей увеличение двойственной целевой функции.

Допустим, что выбрана подходящая двойственная небазисная переменная $(\lambda_B^+)_p$ или $(\lambda_B^-)_p$, изменение которой в допустимом направлении приводит к увеличению двойственной целевой функции. В соответствии с (1.26) в случае выбора $(\lambda_B^+)_p$ мы можем считать, что соответствующая комбинированная небазисная переменная $(\lambda_B)_p = (\lambda_B^+)_p$ начинает возрастать, а в случае выбора $(\lambda_B^-)_p$ — что $(\lambda_B)_p = (\lambda_B^-)_p$ начинает убывать. При этом изменение $(\lambda_B)_p$ в выбранном направлении приведет к изменению двойственных базисных переменных π_i и $(\lambda_N)_j$ в соответствии с равенством (1.33).

Текущее базисное решение предполагается двойственно допустимым. Это означает, что если $(\lambda_N)_j = (\lambda_N^+)_j$ (небазисная переменная прямой ЛП-задачи $(x_N)_j$ имеет активную нижнюю границу), то $d_j \geq 0$, а если $(\lambda_N)_j = (\lambda_N^-)_j$ (небазисная переменная прямой ЛП-задачи $(x_N)_j$ имеет активную верхнюю границу), то $d_j \leq 0$.[6] Понятно, что если в процессе изменения двойственной небазисной переменной $(\lambda_B)_p$ в выбранном направлении некоторая двойственная базисная переменная $(\lambda_N)_q = (\lambda_N^+)_q$ или $(\lambda_N)_q = (\lambda_N^-)_q$ первой достигнет своей нулевой границы, то это будет означать, что мы достигли смежной вершины, и дальнейшее изменение $(\lambda_B)_p$ в выбранном направлении невозможно, так как это приведет к нарушению нулевой границы соответствующей двойственной базисной переменной, а значит, к выходу из множества допустимых решений двойственной ЛП-задачи.

Чтобы найти двойственную базисную переменную, определяющую смежное двойственное базисное решение, положим

$$(\lambda_B)_p = s\theta, \tag{1.38}$$

где $s = +1$ (соответственно, $s = -1$), если $(\lambda_B)_p = (\lambda_B^+)_p$ возрастает (соответственно, если $(\lambda_B)_p = (\lambda_B^-)_p$ убывает), $\theta \geq 0$ — скалярный параметр луча, который исходит из текущей вершины и направление которого совпадает с направлением ребра, соответствующего выбранной небазисной переменной $(\lambda_B)_p$. Заметим, что увеличение параметра θ, начиная с нуля, соответствует изменению $(\lambda_B)_p$ в выбранном направлении, начиная с ее активной нулевой границы. Поскольку остальные двойственные небазисные переменные не изменяются, то с учетом (1.33) отдельная двойственная базисная переменная $(\lambda_N)_j$ будет зависеть от θ следующим образом:

$$(\lambda_N)_j = d_j - \xi_{pj} s\theta, \tag{1.39}$$

где d_j — значение $(\lambda_N)_j$ в текущем базисном решении, ξ_{pj} — элемент текущей симплекс-таблицы $\Xi = -B^{-1}N$. Очевидно, что если $\xi_{pj}s < 0$ (соответственно, если $\xi_{pj}s > 0$), то переменная

[6] Таким образом, с точки зрения прямой ЛП-задачи двойственная допустимость означает *«сверхоптимальность»* в том смысле, что относительные оценки всех небазисных переменных имеют правильные знаки, и если бы при этом все базисные переменные находились внутри своих границ (условие прямой допустимости), то решение было бы оптимальным.

$(\lambda_N)_j$ убывает (соответственно, возрастает) вдоль луча. Поэтому $(\lambda_N)_j = (\lambda_N^+)_j$ (с текущим значением $d_j \geq 0$) достигнет своей нижней нулевой границы, либо $(\lambda_N)_j = (\lambda_N^-)_j$ (с текущим значением $d_j \leq 0$) достигнет своей верхней нулевой границы при $\theta = \theta_j$, где

$$\theta_j = \frac{d_j}{\xi_{pj} s}, \tag{1.40}$$

а значит, при возрастании θ, начиная с нуля, *первой* достигнет своей нулевой границы та двойственная базисная переменная $(\lambda_N)_q$, для которой

$$\theta_q = \theta_{\min} = \min_{\theta_i \geq 0} \theta_i. \tag{1.41}$$

Заметим, что возможен случай, когда $\theta_{\min} = \infty$, который означает, что двойственная ЛП-задача имеет неограниченный оптимум (максимум), откуда в соответствии с теорией двойственности следует, что прямая ЛП-задача не имеет допустимых решений.

Если $\theta_{\min} = \theta_q$, то в смежном двойственном базисном решении та нижняя (соответственно верхняя) нулевая граница двойственной базисной переменной $(\lambda_N)_q = (\lambda_N^+)_q$ (соответственно $(\lambda_N)_q = (\lambda_N^-)_q$), которая препятствует дальнейшему изменению двойственной небазисной переменной $(\lambda_B)_p$ в выбранном направлении, т. е. которая определяет величину θ_q, становится активной, а текущая активная (нулевая) граница переменной $(\lambda_B)_p = (\lambda_B^+)_p$ или $(\lambda_B)_p = (\lambda_B^-)_p$, наоборот, перестает быть активной. Другими словами, переменные $(\lambda_B)_p$ и $(\lambda_N)_q$ в смежном двойственном базисе меняются ролями: $(\lambda_B)_p$ становится базисной (входит в базис), а $(\lambda_N)_q$ становится небазисной (выходит из базиса).

Если сравнить свойства прямого и двойственного допустимых базисных решений с точки зрения условий оптимальности (см. разд. 1.3), то можно заметить, что прямое допустимое, но неоптимальное базисное решение удовлетворяет всем условиям, за исключением условия (1.20), а двойственное допустимое, но неоптимальное базисное решение удовлетворяет всем условиям, за исключением условия (1.18). Таким образом, прямой симплекс-метод поддерживает прямую допустимость базисного решения и пытается сделать его двойственно допустимым, а двойственный симплекс-метод, наоборот, поддерживает двойственную допустимость базисного решения и пытается сделать его прямо допустимым.

В заключение рассмотрим принципиальную схему двойственного симплекс-метода, выраженного в терминах базисных решений *прямой* ЛП-задачи (1.1)—(1.3),[7] полагая, что задано некоторое начальное двойственно допустимое базисное решение (указано, какие переменные являются базисными, какие небазисными, и для небазисных переменных указаны их активные границы).

ДВОЙСТВЕННЫЙ СИМПЛЕКС-МЕТОД

Шаг 1. Вычислить $\beta = (\beta_i)$ — вектор текущих значений базисных переменных $x_B = [(x_B)_i]$ (1.16), $d = (d_j)$ — вектор относительных оценок небазисных переменных $x_N = [(x_N)_j]$ (1.12).

Шаг 2. Выбрать базисную переменную $(x_B)_p$, которая нарушает свою нижнюю или верхнюю границу, т. е. для которой имеет место $\beta_p < (l_B)_p$ или $\beta_p > (u_B)_p$. Если выбор невозможен, то СТОП — текущее базисное решение является оптимальным.

Шаг 3. Выбрать небазисную переменную $(x_N)_q$, которая соответствует величине $\theta_q = \theta_{\min}$ (1.41). Если $\theta_{\min} = \infty$, т. е. выбор невозможен, то СТОП — ЛП-задача не имеет допустимых решений.

[7] Ранее уже было отмечено, что двойственный симплекс-метод, выраженный в терминах двойственной задачи, ничем не отличается от прямого симплекс-метода.

Шаг 4. Перейти к смежному базису: если $\beta_p < (l_B)_p$ (соответственно, если $\beta_p > (u_B)_p$), то пометить переменную $(x_B)_p$ как небазисную с активной нижней (соответственно, верхней) границей, а переменную $(x_N)_q$ пометить как базисную.

Вернуться к шагу 1. ■

Как и в случае прямого симплекс-метода, возможны два варианта двойственного симплекс-метода — *табличный* [1, подразд. 1.4], в котором симплекс-таблица Ξ (1.10) хранится и пересчитывается в явном виде как двумерный массив, и *модифицированный* [1, разд. 2], в котором для вычисления всех необходимых величин используется подходящее представление базисной матрицы B (1.5).

2 Двойственный выбор базисной переменной

В соответствии с принципиальной схемой двойственного симплекс-метода, рассмотренной в конце подразд. 1.6, на шаге 2 требуется выбрать подходящую базисную переменную $(x_B)_p$, которая нарушает свою нижнюю или верхнюю границу в текущем базисном решении.[8] С двойственной точки зрения указанный выбор эквивалентен выбору подходящей двойственной небазисной переменной $(\lambda_B)_p = (\lambda_B)_p^+$ или $(\lambda_B)_p = (\lambda_B)_p^-$ (в зависимости от того, какая граница переменной $(x_B)_p$ нарушена), изменение которой в допустимом направлении приводит к улучшению (увеличению) двойственной целевой функции. Как и в случае прямого симплекс-метода, геометрически выбор подходящей двойственной переменной соответствует выбору ребра многогранного множества допустимых решений двойственной ЛП-задачи, исходящего из текущей вершины, вдоль которого (ребра) имеет место улучшение двойственной целевой функции.

Пусть $I \subseteq \{1, \ldots, m\}$ — множество индексов базисных переменных $(x_B)_i$, для которых в текущем базисе либо $\beta_i < (l_B)_i$ (нарушена нижняя граница), либо $\beta_i > (u_B)_i$ (нарушена верхняя граница). Если $I \neq \varnothing$, то текущее базисное решение не является оптимальным, и в этом случае в качестве базисной переменной можно, в принципе, выбрать любую переменную из I. Таким образом, проблема выбора состоит в принятии решения, какую именно базисную переменную следует выбирать на каждой итерации двойственного симплекс-метода.

В данном разделе рассмотрены основные стратегии выбора базисной переменной, получившие практическое применение.

2.1 Двойственное правило Данцига

Двойственное правило Данцига представляет собой аналог обычного правила Данцига (см. [1, подразд. 3.1]) применительно к двойственному симплекс-методу.

Определим нарушение границы базисной переменной $(x_B)_i$ следующим образом:

$$r_i = \begin{cases} (l_B)_i - \beta_i, & \text{если } \beta_i < (l_B)_i \\ 0, & \text{если } (l_B)_i \leq \beta_i \leq (u_B)_i \\ (u_B)_i - \beta_i, & \text{если } \beta_i > (u_B)_i \end{cases} \tag{2.1}$$

Заметим, что r_i является относительной оценкой переменной $(\lambda_B)_i$ для двойственной целевой функции (1.37). Поэтому двойственное правило Данцига состоит в выборе базисной переменной $(x_B)_p$ с максимальным по абсолютной величине нарушением границы:

$$|r_p| = \max_{i \in I} |r_i|, \tag{2.2}$$

так как в этом случае соответствующая двойственная небазисная переменная $(\lambda_B)_p$ будет иметь максимальную по абсолютной величине относительную оценку. При этом, очевидно, если $r_p > 0$, то следует увеличивать $(\lambda_B)_p = (\lambda_B)_p^+$, а если $r_p < 0$, то следует уменьшать $(\lambda_B)_p = (\lambda_B)_p^-$.

Двойственное правило Данцига обладает тем же недостатком, что и прямое правило Данцига, а именно, его асимметрия является во многих случаях причиной медленной сходимости двойственного симплекс-метода (подробнее об этом см. [1, подразд. 3.1 и 3.2]).

[8] Поскольку базисные переменные соответствуют строкам текущей симплекс-таблицы (1.10), то данную операцию часто называют *выбором строки*.

2.2 Двойственный метод наиболее крутого ребра

Двойственный метод наиболее крутого ребра является аналогом обычного метода наиболее крутого ребра (см. [1, подразд. 3.2]) применительно к двойственному симплекс-методу. В соответствии с этим методом предлагается выбирать ту подходящую базисную переменную $(x_B)_p$, для которой направление ребра, определяемого соответствующей двойственной переменной $(\lambda_B)_p$, образует наиболее острый угол с направлением градиента двойственной целевой функции в пространстве *всех* двойственных переменных π, λ^+ и λ^- двойственной ЛП-задачи (1.23)—(1.25).

Чтобы получить двойственный аналог метода наиболее крутого ребра, найдем двойственную симплекс-таблицу для двойственной ЛП-задачи (1.23)—(1.25). По определению симплекс-таблица показывает явную зависимость базисных переменных от небазисных для текущего базиса. Как уже было отмечено в подразд. 1.5, все двойственные переменные π_i являются базисными, $(\lambda_B^+)_i$ и $(\lambda_B^-)_i$ — небазисными, и среди каждой пары двойственных переменных $(\lambda_N^+)_j$ и $(\lambda_N^-)_j$ одна переменная базисная, а другая небазисная. Так как отдельные переменные λ_j^+ и λ_j^- имеют идентичные столбцы в матрице коэффициентов ограничений-равенств (1.24), то без ограничения общности можно считать, что все переменные $(\lambda_N^+)_j$ являются базисными, а $(\lambda_N^-)_j$ — небазисными, поскольку взаимная замена этих переменных не влияет на двойственную симплекс-таблицу. Тогда из (1.30), (1.32) и (1.33) имеем:

$$\begin{array}{rcl} \pi & = & -B^{-T}\lambda_B^+ - B^{-T}\lambda_B^- + B^{-T}c_B \\ \lambda_N^+ & = & -\Xi^T\lambda_B^+ - \Xi^T\lambda_B^- - \lambda_N^- + d \end{array}$$

или в матричной записи:

$$\begin{pmatrix} \pi \\ \lambda_N^+ \end{pmatrix} = \begin{pmatrix} -B^{-T} & -B^{-T} & 0 \\ -\Xi^T & -\Xi^T & -I \end{pmatrix} \begin{pmatrix} \lambda_B^+ \\ \lambda_B^- \\ \lambda_N^- \end{pmatrix} + \begin{pmatrix} B^{-T}c_B \\ d \end{pmatrix},$$

откуда следует, что *двойственная симплекс-таблица* представляет собой следующую матрицу:

$$\widetilde{\Xi} = \begin{pmatrix} -B^{-T} & -B^{-T} & 0 \\ -\Xi^T & -\Xi^T & -I \end{pmatrix}. \tag{2.3}$$

В соответствии с методом наиболее крутого ребра [1, подразд. 3.2] при выборе небазисной переменной ее относительная оценка масштабируется нормой соответствующего столбца симплекс-таблицы. В рассматриваемом случае, как это было отмечено в подразд. 1.6, достаточно ограничиться небазисными переменными λ_B^+ и λ_B^-. А так как обе переменные $(\lambda_B^+)_i$ и $(\lambda_B^-)_i$ имеют идентичные столбцы в двойственной симплекс-таблице (2.3), то выбор одной из этих двух переменных эквивалентен выбору соответствующей комбинированной двойственной переменной $(\lambda_B)_i$. Таким образом, мы получаем следующее правило выбора переменной $(\lambda_B)_p$, а значит, и переменной $(x_B)_p$, которое является двойственным аналогом правила [1, подразд. 3.2, формула (3.15)]:

$$\frac{r_p^2}{\gamma_p} = \max_{i\in I} \frac{r_i^2}{\gamma_i}. \tag{2.4}$$

Здесь r_i есть относительная оценка переменной $(\lambda_B)_i$ для двойственной целевой функции, совпадающая с нарушением границы переменной $(x_B)_i$ (см. (2.1)). Величина γ_i представляет собой квадрат длины вектора-приращения в пространстве всех двойственных переменных, соответствующего единичному приращению переменной $(\lambda_B)_i$, и является двойственным аналогом для [1, подразд. 3.2, формула (3.10)]:

$$\gamma_i = 1 + \widetilde{\Xi}_i^T \widetilde{\Xi}_i = 1 + \rho_i^T \rho_i + \xi_i^T \xi_i = 1 + \sum_{j=1}^{m} \rho_{ij}^2 + \sum_{j=1}^{n-m} \xi_{ij}^2, \tag{2.5}$$

где $\widetilde{\Xi}_i$ — i-й столбец двойственной симплекс-таблицы (2.3), $\rho_i = (\rho_{ij})$ — i-я строка матрицы B^{-1} и $\xi_i = (\xi_{ij})$ — i-я строка симплекс-таблицы Ξ для текущего базиса прямой ЛП-задачи.

Непосредственное вычисление одной величины γ_i требует двух операций обратного преобразования (BTRAN), поэтому при практическом применении метода наиболее крутого ребра всякий раз при переходе к смежному базису следует использовать *пересчет* величин $\gamma = (\gamma_i)$ для смежного базиса. Однако, как было указано в [1, подразд. 3.2], метод наиболее крутого ребра имеет существенный недостаток, связанный с тем, что для него в общем случае не существует эффективного способа *инициализации* величин $\gamma = (\gamma_i)$ для заданного (начального) базиса. Поэтому соответствующие формулы пересчета здесь не рассматриваются.

2.3 Двойственный проекционный метод наиболее крутого ребра

Двойственный проекционный метод наиболее крутого ребра является аналогом обычного проекционного метода наиболее крутого ребра (см. [1, подразд. 3.3]) применительно к двойственному симплекс-методу. Таким образом, в рассматриваемом случае это означает, что углы между направлениями ребер двойственного многогранного множества допустимых решений и градиентом двойственной целевой функции измеряются не во всем пространстве двойственных переменных π, λ^+ и λ^-, а в некотором *эталонном подпространстве*, которое является координатным подпространством, определяемым двойственными небазисными переменными на некоторой произвольно выбранной итерации (двойственного) симплекс-метода. При этом правило выбора (2.4) остается тем же, но величины $\gamma = (\gamma_i)$, введенные в предыдущем подразделе, определяются иначе, чтобы теперь они соответствовали измерению углов в эталонном подпространстве.

Пусть V — эталонное координатное подпространство двойственных переменных, которое первоначально (т. е. на некоторой произвольно выбранной итерации симплекс-метода) совпадает с подпространством небазисных переменных, и допустим, что формула [1, подразд. 3.3, (3.27)] применяется к двойственной симплекс-таблице $\widetilde{\Xi}$ (2.3), где γ_i относится к i-му столбцу этой симплекс-таблицы, $1 \leq i \leq m + m + (n - m)$. Как было отмечено в подразд. 1.6, выбор небазисных переменных ограничивается только переменными вида $(\lambda_B^+)_i$ и $(\lambda_B^-)_i$, поэтому величины γ_i для столбцов $m + m + 1 \leq i \leq m + m + (n - m)$ не представляют интереса. Кроме того, переменные $(\lambda_B^+)_i$ и $(\lambda_B^-)_i$ имеют идентичные столбцы в двойственной симплекс-таблице, следовательно, $\gamma_i = \gamma_{m+i}$, поэтому достаточно иметь величины γ_i лишь для $1 \leq i \leq m$, т. е. для первых m столбцов двойственной симплекс-таблицы. Наконец, заметим, что двойственные переменные $\pi = (\pi_i)$ остаются базисными в любом двойственном базисном решении, поэтому V может включать только переменные вида λ_j^+ и λ_j^-, а значит, в выражение для γ_i не войдут элементы первых m строк двойственной симплекс-таблицы, содержащих матрицу B^{-T}.

Определим следующие вспомогательные величины для текущего базиса:

$$\eta_j = \begin{cases} 1, & \text{если } (\lambda_N)_j \in V \\ 0, & \text{если } (\lambda_N)_j \notin V \end{cases} \tag{2.6}$$

$$\delta_i = \begin{cases} 1, & \text{если } (\lambda_B)_i \in V \\ 0, & \text{если } (\lambda_B)_i \notin V \end{cases} \tag{2.7}$$

где $(\lambda_N)_j \notin V$ (соответственно, $(\lambda_B)_i \notin V$) означает, что либо $(\lambda_N^+)_j \notin V$, либо $(\lambda_N^+)_j \notin V$ (соответственно, либо $(\lambda_B^+)_j \notin V$, либо $(\lambda_B^+)_j \notin V$), так как обе переменные λ_j^+ и λ_j^- не могут быть базисными одновременно. Тогда, с учетом сделанных выше замечаний, формула [1, подразд. 3.3, (3.27)] применительно к двойственной симплекс-таблице (2.3) будет следующей:[9]

$$\gamma_i = \delta_i + \sum_{j \in C} \xi_{ij}^2, \ i = 1, \dots, m, \tag{2.8}$$

где $C = \{j : \eta_j = 1\} = \{j : (\lambda_N)_j \in V\}$ — множество индексов двойственных базисных переменных, принадлежащих эталонному подпространству, ξ_{ij} — элементы прямой (не двойственной!) симплекс-таблицы $\Xi = -B^{-1}N$.

Допустим, что для текущего базиса величины γ_i (2.8) известны. Допустим также, что в смежном базисе переменная $(x_B)_p$ становится небазисной, а $(x_N)_q$ — базисной. С точки зрения двойственного симплекс-метода это будет означать, что в смежном базисе небазисная переменная $(\lambda_B)_p = (\lambda_B^+)_p$ или $(\lambda_B)_p = (\lambda_B^-)_p$ становится базисной, а базисная переменная $(\lambda_N)_q = (\lambda_N^+)_q$ или $(\lambda_N)_q = (\lambda_N^-)_q$ становится небазисной. Нас интересуют формулы пересчета величин γ_i при переходе к смежному базису.

Для начала рассмотрим пересчет величины γ_p для ведущей строки ($i = p$).[10] В соответствии с (2.8) имеем:

$$\overline{\gamma}_p = \overline{\delta}_p + \sum_{j \in \overline{C}} \overline{\xi}_{pj}^2.$$

Так как в смежном базисе $(\lambda_B)_p$ и $(\lambda_N)_q$ меняются местами, то $\overline{\delta}_p = \eta_q$, $\overline{\eta}_q = \delta_p$ и $\overline{C} \backslash \{q\} = C \backslash \{q\}$. Поэтому:

$$\overline{\gamma}_p = \overline{\delta}_p + \overline{\eta}_q \overline{\xi}_{pq} + \sum_{j \in \overline{C} \backslash \{q\}} \overline{\xi}_{pj}^2 = \eta_q + \delta_p \overline{\xi}_{pq} + \sum_{j \in C \backslash \{q\}} \overline{\xi}_{pj}^2.$$

Используя формулы пересчета элементов симплекс-таблицы для смежного базиса [1, подразд. 1.4, (1.33)—(1.36)], получим:

$$\overline{\gamma}_p = \eta_q + \delta_p \left(\frac{1}{\xi_{pq}} \right)^2 + \sum_{j \in C \backslash \{q\}} \left(\frac{\xi_{pj}}{\xi_{pq}} \right)^2.$$

Заметим, что:

$$\sum_{j \in C} \left(\frac{\xi_{pj}}{\xi_{pq}} \right)^2 = \eta_q \left(\frac{\xi_{pq}}{\xi_{pq}} \right)^2 + \sum_{j \in C \backslash \{q\}} \left(\frac{\xi_{pj}}{\xi_{pq}} \right)^2 = \eta_q + \sum_{j \in C \backslash \{q\}} \left(\frac{\xi_{pj}}{\xi_{pq}} \right)^2,$$

[9] Эту же формулу можно получить и более коротким путем, если в качестве двойственной симплекс-таблицы взять матрицу зависимости λ_N от λ_B (1.33).

[10] Здесь имеются в виду строки прямой симплекс-таблицы $\Xi = -B^{-1}N$.

поэтому:

$$\overline{\gamma}_p = \delta_p \left(\frac{1}{\xi_{pq}}\right)^2 + \sum_{j\in C} \left(\frac{\xi_{pj}}{\xi_{pq}}\right)^2 = \frac{1}{\xi_{pq}^2} \left(\delta_p + \sum_{j\in C} \xi_{pj}^2\right),$$

откуда окончательно получим:

$$\overline{\gamma}_p = \frac{1}{\xi_{pq}^2}\gamma_p. \tag{2.9}$$

Рассмотрим теперь пересчет величин γ_i для остальных строк ($i \neq p$). В соответствии с (2.8) имеем:

$$\overline{\gamma}_i = \overline{\delta}_i + \sum_{j\in\overline{C}} \overline{\xi}_{ij}^2.$$

Поскольку $i \neq p$, то $\overline{\delta}_i = \delta_i$. Кроме того, как уже было отмечено выше, $\overline{\eta}_q = \delta_p$ и $\overline{C}\backslash\{q\} = C\backslash\{q\}$. Поэтому:

$$\overline{\gamma}_i = \overline{\delta}_i + \overline{\eta}_q\overline{\xi}_{iq}^2 + \sum_{j\in\overline{C}\backslash\{q\}} \overline{\xi}_{ij}^2 = \delta_i + \delta_p\overline{\xi}_{iq}^2 + \sum_{j\in C\backslash\{q\}} \overline{\xi}_{ij}^2.$$

Используя формулы пересчета элементов симплекс-таблицы для смежного базиса [1, подразд. 1.4, (1.33)—(1.36)], получим:

$$\overline{\gamma}_i = \delta_i + \delta_p \left(\frac{\xi_{iq}}{\xi_{pq}}\right)^2 + \sum_{j\in C\backslash\{q\}} \left(\xi_{ij} - \frac{\xi_{iq}\xi_{pj}}{\xi_{pq}}\right)^2.$$

Заметим, что:

$$\sum_{j\in C} \left(\xi_{ij} - \frac{\xi_{iq}\xi_{pj}}{\xi_{pq}}\right)^2 = \eta_q \left(\xi_{iq} - \frac{\xi_{iq}\xi_{pq}}{\xi_{pq}}\right)^2 + \sum_{j\in C\backslash\{q\}} \left(\xi_{ij} - \frac{\xi_{iq}\xi_{pj}}{\xi_{pq}}\right)^2 =$$

$$= \sum_{j\in C\backslash\{q\}} \left(\xi_{ij} - \frac{\xi_{iq}\xi_{pj}}{\xi_{pq}}\right)^2,$$

поэтому:

$$\begin{aligned}
\overline{\gamma}_i &= \delta_i + \delta_p \left(\frac{\xi_{iq}}{\xi_{pq}}\right)^2 + \sum_{j\in C} \left(\xi_{ij} - \frac{\xi_{iq}\xi_{pj}}{\xi_{pq}}\right)^2 = \\
&= \delta_i + \delta_p \left(\frac{\xi_{iq}}{\xi_{pq}}\right)^2 + \sum_{j\in C} \xi_{ij}^2 + \sum_{j\in C} \left(\frac{\xi_{iq}\xi_{pj}}{\xi_{pq}}\right)^2 - 2\sum_{j\in C} \frac{\xi_{ij}\xi_{iq}\xi_{pj}}{\xi_{pq}} = \\
&= \left(\delta_i + \sum_{j\in C} \xi_{ij}^2\right) + \left(\frac{\xi_{iq}}{\xi_{pq}}\right)^2 \left(\delta_p + \sum_{j\in C} \xi_{pj}^2\right) - 2\left(\frac{\xi_{iq}}{\xi_{pq}}\right) \sum_{j\in C} \xi_{ij}\xi_{pj},
\end{aligned}$$

откуда следует, что:

$$\overline{\gamma}_i = \gamma_i + \left(\frac{\xi_{iq}}{\xi_{pq}}\right)^2 \gamma_p - 2\left(\frac{\xi_{iq}}{\xi_{pq}}\right) \sum_{j\in C} \xi_{ij}\xi_{pj}.$$

Сумма $\sum_{j\in C}\xi_{ij}\xi_{pj} = \sum_{j=1}^{n-m}\xi_{ij}\eta_j\xi_{pj}$ есть скалярное произведение i-й строки симплекс-таблицы $\xi_i = (\xi_{ij}) = \Xi^T e_i = -N^T B^{-T} e_i$ и вектора $H\xi_p$, где $\xi_p = (\xi_{pj}) = \Xi^T e_p = -N^T B^{-T} e_p$ — p-я (ведущая) строка симплекс-таблицы, $H = \mathrm{diag}(\eta_j)$ — диагональная матрица, составленная из величин η_j (2.6). Поэтому:

$$\sum_{j\in C}\xi_{ij}\xi_{pj} = (H\xi_p)^T\xi_i = -(H\xi_p)^T N^T B^{-T} e_i = -(B^{-1}NH\xi_p)^T e_i.$$

Таким образом, окончательно имеем:

$$\overline{\gamma}_i = \gamma_i + \left(\frac{\xi_{iq}}{\xi_{pq}}\right)^2 \gamma_p + 2\left(\frac{\xi_{iq}}{\xi_{pq}}\right)(B^{-1}NH\xi_p)^T e_i. \tag{2.10}$$

Как и в случае обычного проекционного метода наиболее крутого ребра, многократный пересчет величин γ_i по формулам (2.9) и (2.10) может привести к отклонению этих величин от их точных значений вследствие ошибок округления. Поэтому, чтобы уменьшить влияние этих ошибок, в формуле (2.9) рекомендуется, во-первых, использовать не текущее значение γ_p, а более точное значение, вычисленное непосредственно по формуле (2.8):

$$\gamma_p = \delta_p + \sum_{j\in C}\xi_{pj}^2. \tag{2.11}$$

Кроме того, можно заметить, что для $i \neq p$:

$$\begin{aligned}\overline{\gamma}_i = \overline{\delta}_i + \sum_{j\in\overline{C}}\overline{\xi}_{ij}^2 = \delta_i + \delta_p\overline{\xi}_{iq}^2 + \sum_{j\in C\setminus\{q\}}\overline{\xi}_{ij}^2 \geq \\ \geq \delta_i + \delta_p\overline{\xi}_{iq}^2 = \delta_i + \delta_p\left(\frac{\xi_{iq}}{\xi_{pq}}\right)^2.\end{aligned} \tag{2.12}$$

Поэтому если вследствие ошибок округления значение $\overline{\gamma_i}$, вычисленное по формуле (2.10), оказывается меньше правой части неравенства (2.12), то в качестве более точного значения $\overline{\gamma_i}$ имеет смысл использовать величину в правой части указанного неравенства.

Допустим, что p-я ведущая строка $\xi_p = (\xi_{pj})$ и q-й (ведущий) столбец $\Xi_q = (\xi_{iq})$ текущей симплекс-таблицы уже вычислены. Тогда практическая схема пересчета величин γ_i для смежного базиса в соответствии с двойственным проекционным методом наиболее крутого ребра может быть следующей.

1. Вычислить более точное значение γ_p для текущего базиса:

$$\gamma_p = \delta_p + \sum_{j\in C}\xi_{pj}^2.$$

2. Вычислить вспомогательный вектор, используя операцию FTRAN:

$$u = B^{-1}NH\xi_p,$$

где $H = \mathrm{diag}(\eta_j)$ — диагональная матрица, составленная из величин η_j.

3. Выполнить пп. 4—5 для $i = 1, \dots, p-1, p+1, \dots, m$.

4. Вычислить вспомогательную величину:

$$r_i = \xi_{iq}/\xi_{pq}.$$

5. Вычислить величину $\overline{\gamma}_i$ для смежного базиса:

$$\overline{\gamma}_i = \max(\gamma_i + r_i^2\gamma_p + 2r_i u_i, \delta_i + \delta_p r_i^2),$$

где $u_i = u^T e_i$ — i-я компонента вектора u.

6. Вычислить величину $\overline{\gamma}_p$ для смежного базиса:

$$\overline{\gamma}_p = \gamma_p/\xi_{pq}^2. \qquad \blacksquare$$

Начальное определение (инициализация) и переопределение эталонного подпространства возможны на любой итерации двойственного симплекс-метода и состоят в том, что эталонным подпространством становится текущее подпространство двойственных небазисных переменных $\lambda_B = (\lambda_B)_i$, соответствующих базисным переменным $x_B = (x_B)_i$ с точки зрения прямой ЛП-задачи. В соответствии с формулами (2.6) и (2.7) это означает, что $\eta_j = 0$ для всех $j = 1, \dots, n-m$ (а значит, $C = \varnothing$) и $\delta_i = 1$ для всех $i = 1, \dots, m$. Таким образом, как это следует из (2.8), сразу после инициализации или переопределения эталонного подпространства имеет место $\gamma_i = 1$ для всех $i = 1, \dots, m$.

Поскольку двойственный проекционный метод наиболее крутого ребра основан на измерении углов в подпространстве, а не во всем пространстве (двойственных) переменных, то время от времени (например, через каждые 500 или 1000 итераций двойственного симплекс-метода) рекомендуется выполнять переопределение эталонного подпространства.

В заключение еще раз отметим, что двойственный проекционный метод наиболее крутого ребра в точности совпадает с обычным проекционным методом наиболее крутого ребра, если прямой симплекс-метод применяется для решения двойственной ЛП-задачи (1.23)—(1.25).

2.4 Двойственный метод оценивания Devex

Двойственный метод оценивания Devex является аналогом обычного метода оценивания Devex (см. [1, подразд. 3.4]) применительно к двойственному симплекс-методу. Как прямой, так и двойственный варианты данного метода были предложены П. Харрис (P. M. J. Harris) [4], кто первой обратила внимание на асимметрию правила Данцига и исходя из эмпирических соображений предложила измерять угол между выбираемым направлением и градиентом целевой функции не в текущем подпространстве небазисных переменных, которое изменяется на каждой итерации, а в фиксированном (эталонном) подпространстве.

Двойственный метод оценивания Devex по существу представляет собой двойственный проекционный метод наиболее крутого ребра (см. подразд. 2.3), где единственное отличие состоит в том, что для пересчета масштабирующих величин γ_i (2.8) используется приближенная формула, которую можно получить из формулы (2.10), если отбросить последнее слагаемое:

$$\overline{\gamma}_i \approx \gamma_i + \left(\frac{\xi_{iq}}{\xi_{pq}}\right)^2 \gamma_p. \qquad (2.13)$$

В данном случае требуется лишь одна операция прямого преобразования (FTRAN) на каждой итерации двойственного симплекс-метода для вычисления элементов ведущего столбца симплекс-таблицы $\Xi_q = (\xi_{iq})$,[11] поэтому использование двойственного метода оценивания Devex позволяет сэкономить одну операцию FTRAN на каждой итерации по сравнению с двойственным проекционным методом наиболее крутого ребра.

Допустим, что p-я (ведущая) строка $\xi_p = (\xi_{pj})$ и q-й (ведущий) столбец $\Xi_q = (\xi_{iq})$ текущей симплекс-таблицы уже вычислены. Тогда практическая схема пересчета величин γ_i для смежного базиса в соответствии с двойственным методом оценивания Devex может быть следующей. (Здесь считается, что эталонное подпространство V двойственных переменных определено посредством величин η_j (2.6) и δ_i (2.7) точно так же, как и в случае двойственного метода наиболее крутого ребра.)

1. Вычислить точное значение γ_p для текущего базиса:

$$\gamma_p = \delta_p + \sum_{j \in C} \xi_{pj}^2,$$

где $C = \{j : \eta_j = 1\} = \{j : (\lambda_N)_j \in V\}$ — множество индексов двойственных базисных переменных, принадлежащий эталонному подпространству. (Заметим, что в соответствии с формулой (2.13) величины γ_i монотонно возрастают. Поэтому если вычисленное точное значение γ_p значительно отличается от его текущего приближенного значения, следует переопределить эталонное подпространство.)

2. Вычислить величины $\overline{\gamma}_i$, $i = 1, \ldots, p-1, p+1, \ldots, m$ для смежного базиса:

$$\overline{\gamma}_i = \gamma_i + (\xi_{iq}/\xi_{pq})^2 \gamma_p.$$

3. Вычислить величину $\overline{\gamma}_p$ для смежного базиса:

$$\overline{\gamma}_p = \gamma_p / \xi_{pq}^2. \qquad \blacksquare$$

Выбор базисной переменной $(x_B)_p$, соответствующий выбору двойственной небазисной переменной $(\lambda_B)_p$, выполняется по формуле (2.4) аналогично тому, как это делается при использовании двойственного проекционного метода наиболее крутого ребра. Начальное определение (инициализация) и переопределение эталонного подпространства также выполняются аналогичным образом.

[11] Этот столбец также используется для пересчета значений базисных переменных прямой ЛП-задачи (см. [1, подразд. 2.3]).

3 Двойственный выбор небазисной переменной

В соответствии с принципиальной схемой двойственного симплекс-метода, рассмотренной в конце подразд. 1.6, на шаге 3 требуется определить $\theta_{\min}$ — абсолютную величину приращения двойственной небазисной переменной $(\lambda_B)_p = (\lambda_B^+)_p$ или $(\lambda_B)_p = (\lambda_B^-)_p$, соответствующей выбранной базисной переменной $(x_B)_p$, и тем самым выбрать небазисную переменную $(x_N)_q$,[12] соответствующую двойственной базисной переменной $(\lambda_N)_q = (\lambda_N^+)_q$ или $(\lambda_N)_q = (\lambda_N^-)_q$, которая первой достигает своей (нулевой) границы при изменении $(\lambda_B)_p$ в допустимом направлении.

Как и в случае прямого симплекс-метода, геометрически выбор двойственной базисной переменной представляет собой определение грани (гиперплоскости), соответствующей (нижней или верхней) нулевой границе двойственной базисной переменной, которая (грань) встречается первой при движении вдоль ребра многогранного множества допустимых решений двойственной ЛП-задачи, исходящего из текущей вершины и соответствующего выбранной двойственной небазисной переменной, и которая (грань), таким образом, определяет смежную вершину, где эта грань становится активным ограничением.

3.1 «Стандартный» двойственный метод

Рассмотрим вначале *«стандартный» двойственный метод* выбора небазисной переменной (textbook dual ratio test). Этот метод соответствует буквальному следованию формулам (1.38)—(1.41) и служит, таким образом, основой для других методов выбора.

Представим множество индексов $J = \{1, \ldots, n-m\}$ всех небазисных переменных $(x_N)_j$ в виде следующего разбиения:

$$J = J_F \cup J_L \cup J_U \cup J_S, \tag{3.1}$$

где J_F — свободные (неограниченные по знаку) небазисные переменные, J_L — небазисные переменные с активной нижней границей, J_U — небазисные переменные с активной верхней границей, J_S — фиксированные небазисные переменные. Тогда из определения двойственной ЛП-задачи и комбинированных двойственных переменных (см. подразд. 1.4) следует, что для текущего базиса:

если $(x_N)_j \in J_F$, то $(\lambda_N)_j = (\lambda_N^+)_j = (\lambda_N^-)_j = 0$, т. е. $(\lambda_N)_j$ является фиксированной в нуле;

если $(x_N)_j \in J_L$, то $(\lambda_N)_j = (\lambda_N^+)_j \geq 0$, т. е. $(\lambda_N)_j$ имеет нижнюю нулевую границу;

если $(x_N)_j \in J_U$, то $(\lambda_N)_j = (\lambda_N^-)_j \leq 0$, т. е. $(\lambda_N)_j$ имеет верхнюю нулевую границу;

если $(x_N)_j \in J_S$, то либо $(\lambda_N)_j = (\lambda_N^+)_j \geq 0$, либо $(\lambda_N)_j = (\lambda_N^-)_j \leq 0$, т. е. $(\lambda_N)_j$ является свободной (не ограниченной по знаку).

Следовательно, поскольку текущий базис предполагается двойственно допустимым, при изменении выбранной двойственной небазисной переменной $(\lambda_B)_p$ в допустимом направлении двойственная базисная переменная $(\lambda_N)_j$ может достичь своей нулевой границы в двух случаях:

если $(x_N)_j \in J_F \cup J_L$ и $(\lambda_N)_j = (\lambda_N^+)_j$ убывает, так как в текущем базисе $(\lambda_N)_j = d_j \geq 0$;

если $(x_N)_j \in J_F \cup J_L$ и $(\lambda_N)_j = (\lambda_N^-)_j$ возрастает, так как в текущем базисе $(\lambda_N)_j = d_j \leq 0$.

«Стандартный» двойственный метод выбора небазисной переменной показан на рис. 3.1.

[12] Поскольку небазисные переменные прямой задачи соответствуют столбцам текущей симплекс-таблицы (1.10), то данную операцию часто называют *выбором столбца*.

```
«СТАНДАРТНЫЙ» ДВОЙСТВЕННЫЙ МЕТОД
/* установить s, чтобы обеспечить θ = s(λ_B)_p ≥ 0 */
s := if (λ_B)_p возрастает then +1 else −1;
/* вначале ничего не выбрано */
q := 0, θ_min := ∞;
/* основной цикл */
for j := 1, ..., n − m do
  if sξ_pj > 0 AND (j ∈ J_F OR j ∈ J_L) then
    /* (λ_N)_j = d_j ≥ 0 убывает и ограничена снизу нулем */
  else if sξ_pj < 0 AND (j ∈ J_F OR j ∈ J_U) then
    /* (λ_N)_j = d_j ≤ 0 возрастает и ограничена сверху нулем */
  else
    /* (λ_N)_j не может достичь нуля при возрастании θ */
    continue;
  end if
  /* (λ_N)_j достигает нуля при θ = θ_j ≥ 0 */
  θ_j := s(d_j/ξ_pj);
  if θ_min > θ_j then
    q := j, θ_min := θ_j;
  end if
end for
if q = 0 then STOP; /* θ_min = ∞ */
```

Рис. 3.1. «Стандартный» двойственный метод выбора небазисной переменной $(x_N)_q$.

Как и в случае прямого симплекс-метода, рассмотренный «стандартный» двойственный метод работоспособен лишь при отсутствии ошибок округления. В условиях приближенных вычислений непосредственное использование данного метода сопровождается следующими проблемами (см. также [1, подразд. 4.1 и 4.2]):

1. *Неприемлемый ведущий элемент.* Ведущий элемент — это элемент ξ_{pq} текущей симплекс-таблицы (1.10), соответствующий выбранным базисной $(x_B)_p$ и небазисной $(x_N)_q$ переменным. Если не учитывать абсолютную величину ξ_{pq} при выборе $(x_N)_q$, то слишком малый по абсолютной величине ведущий элемент может привести к очень плохой обусловленности базисной матрицы для смежного базиса и, как следствие, к существенному ухудшению точности всех компонент базисного решения.

2. *Двойственная вырожденность.* Базисное решение является двойственно вырожденным, если для некоторых двойственных базисных переменных $(\lambda_N)_j$ имеет место $\theta_j = \theta_{\min} = 0$. Двойственная вырожденность замедляет сходимость и может привести к зацикливанию двойственного симплекс-метода.

3. *Нарушение двойственной допустимости.* Так как в текущем базисном решении вектор относительных оценок небазисных переменных $d = (d_j)$, которые являются значениями соответствующих двойственных базисных переменных, вычисляется приближенно, то при переходе к очередному смежному базису может случиться, что некоторые из указанных величин будут иметь неправильный знак ($d_j < 0$ для $(x_N)_j$ с активной нижней границей или $d_j > 0$ для $(x_N)_j$ с активной верхней границей).

Чтобы частично избавиться от перечисленных проблем, «стандартный» двойственный метод можно несколько улучшить, используя те же приемы, что и для обычного «стандартного» метода [1, подразд. 4.2], следующим образом.

1. Перед выполнением очередной итерации выполнять проверку, что $d_j \geq -\delta_1$ (соответственно, $d_j \leq +\delta_1$) для всех небазисных переменных $(x_N)_j$ с активной нижней (соответственно, верхней) границей, т. е. что текущее базисное решение является двойственно допустимым в пределах заданного допуска $\delta_1 > 0$ (например, для $\delta_1 = 10^{-8}$. Если же указанные условия нарушается, то следует прекратить поиск оптимального решения и перейти к поиску двойственно допустимого базисного решения для восстановления двойственной допустимости базиса.

2. Если $(-\delta_1 \leq) d_j < 0$ (соответственно, $0 < d_j (\leq +\delta_1)$) для $(x_N)_j$ с активной нижней (соответственно, верхней) границей, то единственный выход в рамках «стандартного» двойственного метода — считать, что точное значение d_j в этих случаях равно нулю.

3. Если $d_j \leq +\delta$ (соответственно, $d_j \geq -\delta$) для $(x_N)_j$ с активной нижней (соответственно, верхней) границей (например, для $\delta = 10^{-9}$), то по аналогии с предыдущим пунктом можно считать, что точное значение d_j в этих случаях также равно нулю. Это позволяет искусственно увеличить двойственную вырожденность базисного решения и тем самым обеспечить бо́льшую свободу выбора ведущего элемента.

4. Если $|\xi_{pj}| < \varepsilon$ (например, для $\varepsilon = 10^{-10}$), т. е. если элемент симплекс-таблицы является очень малым по абсолютной величине, то считать, что в этом случае $\xi_{pj} = 0$, т. е. что двойственная базисная переменная $(\lambda_N)_j$ не зависит от выбранной двойственной небазисной переменной $(\lambda_B)_p$. (Заметим, что если точное значение ξ_{pj} на самом деле отлично от нуля, то пропуск $(\lambda_N)_j$ может привести к нарушению двойственной допустимости в смежном базисе.)

Схема улучшенного «стандартного» двойственного метода показана на рис. 3.2.

Улучшенный «стандартный» двойственный метод дает возможность частично разрешить проблему ведущего элемента [1, подразд. 4.1], а также проблему нарушения двойственной допустимости (если это нарушение не превышает заданного допуска). Однако, как и его прямой аналог [1, подразд. 4.2], данный метод имеет ряд недостатков, связанных в основном с трудностью выбора подходящих допусков.

3.2 Двойственный метод Харрис

Двойственный метод Харрис является аналогом метода Харрис для прямого симплекс-метода (см. [1, подразд. 4.3]). Оба этих метода были предложены П. Харрис (P. M. J. Harris) в ее работе [4]. Основная идея двойственного варианта основана на конструктивном использовании возможного нарушения (нулевых) границ двойственных базисных переменных $\lambda_N = [(\lambda_N)_j]$ в текущем базисном решении с тем, чтобы увеличить свободу выбора и тем самым обеспечить выбор двойственной базисной переменной с более подходящим ведущим элементом симплекс-таблицы.

Как и в случае прямого симплекс-метода, двойственный метод Харрис (рис. 3.3) включает два этапа (просмотра двойственных базисных переменных $\lambda_N = [(\lambda_N)_j]$). Первый этап служит для определения величины $\widetilde{\theta}_{\min}$, причем в данном случае, в отличие от «стандартного» метода (см. подразд. 3.1), используются не исходные нулевые границы двойственных базисных переменных $(\lambda_N^+)_j \geq 0$ и $(\lambda_N^-)_j \leq 0$, а ослабленные границы $(\lambda_N^+)_j \geq -\delta$ и $(\lambda_N^-)_j \leq +\delta$, где $\delta > 0$ — заданный допуск на нарушение двойственных границ. Это означает, что для $\theta \leq \widetilde{\theta}_{\min}$ выход

```
УЛУЧШЕННЫЙ «СТАНДАРТНЫЙ» ДВОЙСТВЕННЫЙ МЕТОД
/* определить s, чтобы обеспечить θ = s(λ_B)_p ≥ 0 */
s := if (λ_B)_p возрастает then +1 else −1;
/* вначале ничего не выбрано */
q := 0, θ_min := ∞;
/* основной цикл */
for j := 1, ..., n − m do
  if sξ_pj ≥ +ε AND (j ∈ J_F OR j ∈ J_L) then
    /* (λ_N)_j = d_j ≥ 0 убывает и ограничена снизу нулем */
    if d_j < −δ_1 then
      STOP; /* чрезмерное нарушение границы */
    else if d_j < +δ then
      d̃_j := 0;
    else
      d̃_j := d_j;
    end if
  else if sξ_pj ≤ −ε AND (j ∈ J_F OR j ∈ J_U) then
    /* (λ_N)_j = d_j ≤ 0 возрастает и ограничена сверху нулем */
    if d_j > +δ_1 then
      STOP; /* чрезмерное нарушение границы */
    else if d_j > −δ then
      d̃_j := 0;
    else
      d̃_j := d_j;
    end if
  else
    /* (λ_N)_j не может достичь нуля при возрастании θ */
    continue;
  end if
  /* (λ_N)_j достигает нуля при θ = θ_j ≥ 0 */
  θ_j := s(d̃_j/ξ_pj);
  if θ_min > θ_j OR (θ_min = θ_j) AND |ξ_pq| < |ξ_pj| then
    q := j, θ_min := θ_j;
  end if
end for
if q = 0 then STOP; /* возрастание θ не ограничено */
```

Рис. 3.2. Улучшенный «стандартный» двойственный метод выбора небазисной переменной $(x_N)_q$.

значений двойственных базисных переменных за нулевую границу не может превысить заданного допуска δ. Затем на втором этапе выполняется выбор двойственной базисной переменной $(\lambda_N)_q$, для которой

$$|\xi_{pq}| = \max_{j \in J^+ \cup J^-,\ \theta_j \le \widetilde{\theta}_{\min}} |\xi_{pj}|, \tag{3.2}$$

где $J^+ = \{j \in J_F \cup J_L : \xi_{pj} > 0\}$ — двойственные базисные переменные $(\lambda_N)_j = (\lambda_N^+)_j$, которые убывают при возрастании параметра θ, $J^- = \{j \in J_F \cup J_U : \xi_{pj} < 0\}$ — двойственные базисные

ДВОЙСТВЕННЫЙ МЕТОД ХАРРИС
/* определить s, чтобы обеспечить $\theta = s(\lambda_B)_p \geq 0$ */
$s :=$ **if** $(\lambda_B)_p$ возрастает **then** $+1$ **else** -1;
/* первый просмотр */
$\widetilde{\theta}_{\min} := \infty$;
for $j := 1, \ldots, n-m$ **do**
  if $s\xi_{pj} \geq +\varepsilon$ AND ($j \in J_F$ OR $j \in J_L$) **then**
    /* $(\lambda_N)_j = d_j \geq 0$ убывает и ограничена снизу нулем */
    if $d_j < -\delta_1$ **then**
      STOP; /* чрезмерное нарушение границы */
    else if $d_j < 0$ **then**
      $\widetilde{d}_j := 0 + \delta$;
    else
      $\widetilde{d}_j := d_j + \delta$;
    end if
  else if $s\xi_{pj} \leq -\varepsilon$ AND ($j \in J_F$ OR $j \in J_U$) **then**
    /* $(\lambda_N)_j = d_j \leq 0$ возрастает и ограничена сверху нулем */
    if $d_j > +\delta_1$ **then**
      STOP; /* чрезмерное нарушение границы */
    else if $d_j > 0$ **then**
      $\widetilde{d}_j := 0 - \delta$;
    else
      $\widetilde{d}_j := d_j - \delta$;
    end if
  else
    /* $(\lambda_N)_j$ не может достичь нуля при возрастании θ */
    continue;
  end if
  /* $(\lambda_N)_j$ достигает $\mp\delta$ при $\theta = \widetilde{\theta}_j \geq 0$ */
  $\widetilde{\theta}_j := s(\widetilde{d}_j / \xi_{pj})$;
  $\widetilde{\theta}_{\min} := \min(\widetilde{\theta}_{\min}, \widetilde{\theta}_j)$;
end for
if $\widetilde{\theta}_{\min} = \infty$ **then** STOP; /* возрастание θ не ограничено */
/* второй просмотр */
$q := 0$;
for $j := 1, \ldots, n-m$ **do**
  if $s\xi_{pj} \geq +\varepsilon$ AND ($j \in J_F$ OR $j \in J_L$) **then**
    /* $(\lambda_N)_j = d_j \geq 0$ убывает и ограничена снизу нулем */
    nop;
  else if $s\xi_{pj} \leq -\varepsilon$ AND ($j \in J_F$ OR $j \in J_U$) **then**
    /* $(\lambda_N)_j = d_j \leq 0$ возрастает и ограничена сверху нулем */
    nop;
  else
    /* $(\lambda_N)_j$ не может достичь нуля при возрастании θ */
    continue;
  end if
  /* $(\lambda_N)_j$ достигает нуля при $\theta = \theta_j \geq 0$ */

```
    θ_j := s(d_j/ξ_pj);
    if θ_j ≤ θ̃_min AND (q = 0 OR ξ_pq < ξ_pj) then
      q := j;
    end if
  end for
  assert q ≠ 0;
```

Рис. 3.3. Двойственный метод Харрис выбора небазисной переменной $(x_N)_q$.

переменные $(\lambda_N)_j = (\lambda_N^-)_j$, которые возрастают при возрастании параметра θ, $\theta_j = d_j/\xi_{pj}$ — значение θ, при котором переменная $(\lambda_N)_j$ достигает своей нулевой границы.

Вследствие ослабления нулевых границ двойственных базисных переменных имеет место неравенство $\widetilde{\theta}_{\min} \geq \theta_{\min}$, где $\theta_{\min}$ — максимальное приращение параметра луча θ, которое было бы определено «стандартным» методом для неослабленных (нулевых) границ. Вследствие этого в методе Харрис появляется дополнительная свобода, которая используется для выбора двойственной базисной переменной $(\lambda_N)_q$ с наиболее подходящим (максимальным по абсолютной величине) коэффициентом ξ_{pq} в соответствии с условием (3.2). А так как $\theta_q \leq \widetilde{\theta}_{\min}$, то нарушение (нулевых) границ двойственных базисных переменных в смежном базисном решении не превысит величины допуска δ.

Метод Харрис основан на неявном предположении, что в текущем базисном решении двойственные базисные переменные не нарушают своих границ. Однако может случиться (хотя бы в результате применения самого метода Харрис), что (в пределах допуска δ_1) такое нарушение имеет место, т. е. когда $(-\delta_1 \leq) d_j < 0$ для $(\lambda_N)_j = (\lambda_N^+)_j$ или $0 < d_j (\leq +\delta_1)$ для $(\lambda_N)_j = (\lambda_N^-)_j$, где d_j — значение $(\lambda_N)_j$ в текущем базисе (относительная оценка $(x_N)_j$). Если при этом такая двойственная базисная переменная препятствует увеличению параметра θ, то Харрис [4] предлагает брать $\widetilde{\theta}_{\min} = 0$. Но, как уже было отмечено в [1, подразд. 4.3] в аналогичной ситуации для прямого симплекс-метода, нарушение двойственных границ может быть следствием ошибок округления, поэтому имеется больше оснований считать, что «точное» значение d_j для указанных $(\lambda_N)_j$ равно нулю. Такое изменение по сравнению с оригинальным методом Харрис (отраженное на рис. 3.3), дает бóльшую свободу выбора, так как обеспечивает $\widetilde{\theta}_{\min} > 0$ даже в случае нарушения двойственных границ. При этом можно ожидать, что возможное дополнительное нарушение границ двойственных базисных переменных будет оставаться в пределах разрешенного допуска, тем более, что на последующих итерациях изменения значений d_j могут привести к уменьшению общей степени нарушения границ.

В заключение можно отметить, что если двойственный метод Харрис приводит к выбору двойственной базисной переменной, отличной от той, которая была бы выбрана с использованием «стандартного» двойственного метода, то этот выбор всегда соответствует большему приращению параметра луча θ, что существенно с точки зрения предотвращения зацикливания двойственного симплекс-метода, и более подходящему (большему по абсолютной величине) ведущему элементу ξ_{pq}, что существенно с точки зрения численной устойчивости.

3.3 Длинношаговый метод

3.3.1 Общая идея

Допустим, что все переменные исходной ЛП-задачи (1.1)—(1.3) имеют обе нижнюю и верхнюю границы, и рассмотрим соответствующую двойственную ЛП-задачу (1.28)—(1.29). Заметим, что система ограничений-равенств двойственной задачи (1.29) включает только комбинированные двойственные переменные $\lambda = (\lambda_k)$. Чтобы выразить двойственную целевую функцию (1.28) также через эти комбинированные переменные, определим следующие вспомогательные функции для каждой комбинированной переменной λ_k:

$$\psi_k(\lambda_k) = \begin{cases} l_k\lambda_k, & \text{если } \lambda_k = \lambda_k^+ \geq 0 \\ u_k\lambda_k, & \text{если } \lambda_k = \lambda_k^- \leq 0 \end{cases} \tag{3.3}$$

Поскольку (см. подразд. 1.4) $\lambda_k = \lambda_k^+ + \lambda_k^-$ и $\lambda_k^+\lambda_k^- = 0$, т. е. хотя бы одна из переменных λ_k^+ и λ_k^- равна нулю, то двойственную ЛП-задачу (1.28)—(1.29) можно записать в следующем виде, используя только двойственные переменные $\pi = (\pi_i)$ и $\lambda = (\lambda_k)$:

$$\zeta = b^T\pi + \sum_k \psi_k(\lambda_k) + c_0 \to \max \tag{3.4}$$

$$A^T\pi + \lambda = c \tag{3.5}$$

По определению каждая функция $\psi_k(\lambda_k)$ является непрерывной кусочно-линейной функцией с единственной точкой излома в нуле. При этом, так как $l_k \leq u_k$ для каждой переменной x_k прямой ЛП-задачи[13], то каждая такая функция, кроме того, является выпуклой вверх. Пример функции $\psi_k(\lambda_k)$ показан на рис. 3.4.

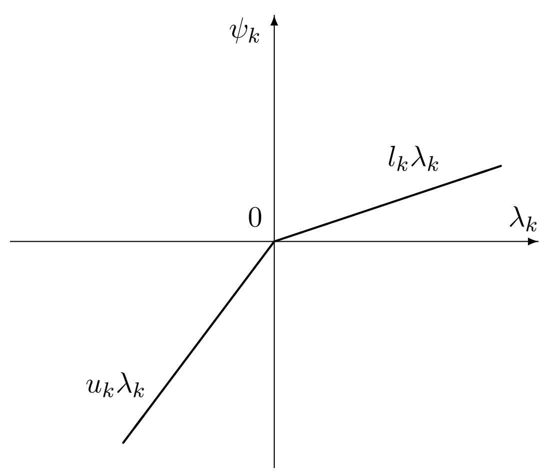

Рис. 3.4. Пример функции $\psi_k(\lambda_k)$.

Таким образом, целевая функция (3.4), будучи суммой выпуклых вверх функций,[14] сама есть выпуклая вверх сеперабельная[15] кусочно-линейная функция. С другой стороны, множество допустимых решений задачи (3.4)—(3.5), т. е. множество точек, удовлетворяющих системе

[13]В противном случае прямая ЛП-задача не имела бы допустимых решений.

[14]Линейные функции $b_i\pi_i$ являются частным случаем выпуклых вверх функций.

[15]Имеющая вид суммы функций, каждая из которых зависит только от одной переменной.

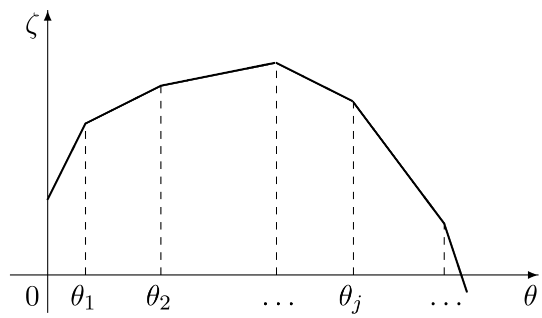

Рис. 3.5. Изменение двойственной целевой функции (3.4)
вдоль луча.

ограничений-равенств (3.5), является плоским, а значит, выпуклым. Следовательно, с учетом максимизации двойственную задачу в формате (3.4)—(3.5) можно классифицировать как задачу выпуклого программирования с сепарабельной кусочно-линейной целевой функцией и линейными ограничениями-равенствами.

В принципе, если не принимать во внимание эффективность, любую задачу выпуклого программирования с системой линейных ограничений в виде равенств можно решить, используя следующую простую технику. Допустим, что у нас имеется некоторая текущая допустимая точка. Выберем допустимое направление, вдоль которого имеет место улучшение целевой функции. Геометрически это направление представляет собой луч, исходящий из допустимой точки и принадлежащий плоскому множеству, определяемому системой ограничений-равенств. Если такого направления не существует, то текущая точка является искомым (причем, глобальным) оптимальным решением. В противном случае сделаем шаг в выбранном направлении в точку с лучшим значением целевой функции и повторим поиск, используя новую точку в качестве текущей.

Применительно к задаче (3.4)—(3.5) рассмотренная техника означает следующее. Пусть (π^0, λ^0) — текущая допустимая точка, (π', λ') — произвольная допустимая точка, определяющее выбранное направление. Тогда луч, исходящий из текущей точки и идущий в выбранном направлении есть следующее множество (допустимых) точек:

$$\left(\begin{array}{c} \pi \\ \lambda \end{array} \right) = \left(\begin{array}{c} \pi' - \pi^0 \\ \lambda' - \lambda^0 \end{array} \right) \theta + \left(\begin{array}{c} \pi^0 \\ \lambda^0 \end{array} \right), \tag{3.6}$$

где $\theta \geq 0$ — параметр луча. Поскольку луч есть выпуклое множество, то на этом множестве (т. е. вдоль луча) целевая функция $\zeta = \zeta(\theta)$ также будет непрерывной кусочно-линейной и выпуклой вверх (рис. 3.5). При этом точки излома целевой функции θ_1, θ_2, $\ldots$ суть точки, где переменные λ_k пересекают нулевую точку при возрастании параметра луча θ в соответствии с (3.6).

Решение двойственной ЛП-задачи (1.23)—(1.25) симплекс-методом по существу есть решение ЛП-задачи (1.28)—(1.29), а следовательно, и задачи (3.4)—(3.5), так как все эти три задачи эквивалентны. Отметим, однако, что симплекс-метод решает ЛП-задачу (1.23)—(1.25) в пространстве двойственных переменных $(\pi, \lambda^+, \lambda^-)$, где целевая функция является линейной, а переменные λ^+ и λ^- имеют нулевые границы. В то же время, задача (3.4)—(3.5) сформулирована в пространстве двойственных переменных (π, λ), где целевая функция является нелинейной, а переменные не ограничены по знаку.[16] Поэтому имеет смысл рассмотреть принципиальную

[16]При условии, что все переменные исходной ЛП-задачи (1.1)—(1.3) имеют обе нижнюю и верхнюю границы.

схему симплекс-метода (см. подразд. 1.6) с точки зрения изложенной выше общей техники решения задачи (3.4)—(3.5), что позволяет установить следующее.

1. Текущей точкой является двойственное базисное решение (1.34).

2. Допустимое направление определяется выбором двойственной небазисной переменной $(\lambda_B^+)_p$ или $(\lambda_B^-)_p$, изменение (увеличение или, соответственно, уменьшение) которой приводит к улучшению (увеличению) двойственной целевой функции. Геометрически в пространстве двойственных переменных $(\pi, \lambda^+, \lambda^-)$ это направление есть направление соответствующего ребра многогранного множества, исходящего из текущей вершины. С точки зрения задачи (3.4)—(3.5) такой выбор означает, что $\theta = +(\lambda_B)_p$, если $(\lambda_B)_p = (\lambda_B^+)_p$, или $\theta = -(\lambda_B)_p$, если $(\lambda_B)_p = (\lambda_B^-)_p$, где $(\lambda_B)_p = (\lambda_B^+)_p + (\lambda_B^-)_p$ — комбинированная двойственная небазисная переменная, $\theta \geq 0$ — параметр луча.

3. Величина шага при возрастании параметра θ определяется выбором двойственной базисной переменной $(\lambda_N^+)_q$ или $(\lambda_N^-)_q$, которая (убывая или, соответственно, возрастая) первой достигает своей нулевой границы. Геометрически в пространстве $(\pi, \lambda^+, \lambda^-)$ это соответствует достижению смежной вершины многогранного множества при движении вдоль выбранного ребра. Достаточно очевидно, что с точки зрения задачи (3.4)—(3.5) такой выбор соответствует *самой первой* точке излома ($\theta = \theta_1$ на рис. 3.5), где соответствующая комбинированная базисная переменная $(\lambda_N)_q = (\lambda_N^+)_q + (\lambda_N^-)_q$ *самой первой* пересекает нуль.

Основная идея *длинношагового метода* выбора двойственной базисной переменной $(\lambda_N)_q$ (и, соответственно, небазисной переменной $(x_N)_q$ прямой ЛП-задачи) состоит в том, чтобы при увеличении параметра луча θ не останавливаться в самой первой точке излома $\theta = \theta_1$, а *продолжать увеличивать* θ до тех пор, пока это приводит к улучшению (увеличению) двойственной целевой функции.

С точки зрения решения двойственной ЛП-задачи (1.23)—(1.25) такой длинношаговый метод можно интерпретировать следующим образом. Пусть параметр луча θ пересекает точку излома двойственной целевой функции (3.4). Это означает, что некоторая комбинированная двойственная базисная переменная $(\lambda_N)_j = (\lambda_N^+)_j + (\lambda_N^-)_j$ пересекает нуль, что, в свою очередь, означает, что либо базисная переменная $(\lambda_N^+)_j$ (убывая) достигает своей нижней нулевой границы, либо базисная переменная $(\lambda_N^-)_j$ (возрастая) достигает своей верхней нулевой границы. Пусть, например, такой базисной переменной является $(\lambda_N^+)_j$. Тогда, чтобы при дальнейшем увеличении θ предотвратить выход $(\lambda_N^+)_j$ в недопустимую (отрицательную) область, достаточно вывести ее из базиса, а вместо нее ввести в базис двойственную переменную $(\lambda_N^-)_j$.[17] Здесь также важно заметить, что переменные $(\lambda_N^+)_j$ и $(\lambda_N^-)_j$ имеют идентичные (единичные) столбцы в матрице коэффициентов ограничений (1.24) двойственной ЛП-задачи, поэтому указанное изменение базиса *не изменяет* текущую базисную матрицу B (1.5).

Если рассматривать длинношаговый вариант двойственного симплекс-метода с точки зрения исходной (прямой) ЛП-задачи (1.1)—(1.3), то переход через точку излома двойственной целевой функции, где двойственные переменные $(\lambda_N^+)_j$ и $(\lambda_N^-)_j$ меняются местами в базисе, означает, что соответствующая небазисная переменная $(x_N)_j$ переходит со своей текущей активной границы на противоположную, оставаясь при этом небазисной.[18]

[17] Как уже было отмечено в подразд. 1.5, если одна из переменных $(\lambda_N^+)_j$ и $(\lambda_N^-)_j$ базисная, то другая обязательно небазисная.

[18] Поэтому в англоязычной литературе длинношаговый выбор небазисной переменной $(x_N)_q$ в двойственном симплекс-методе также называют «flipping ratio test» или «flip-flop ratio test» (от англ. flip — щелкать, flip-flop — щелкающий звук), разумея под этим скачкообразный переход («перещелкивание») отдельных небазисных переменных с одной активной границы на другую.

Длинношаговый вариант двойственного симплекс-метода имеет два существенных преимущества по сравнению с другими вариантами, рассмотренными в предыдущих подразделах:

1) длинношаговая итерация, как правило, обеспечивает большее приращение двойственной целевой функции, что, в конечном итоге, может привести к значительному сокращению общего числа итераций;

2) в качестве точки, определяющей величину шага (а значит, и смежный базис), можно брать любую точку излома θ_j двойственной целевой функции, для которой $\zeta(\theta_j) > \zeta(0)$, где $\zeta(0)$ — значение двойственной целевой функции в текущей точке $\theta = 0$. Поскольку таких точек, как правило, несколько, то это дает бóльшую свободу выбора, что существенно с точки зрения обеспечения численной устойчивости.

По всей видимости, впервые длинношаговый вариант двойственного симплекс-метода был предложен в работе [5]. Более современное изложение этого метода вместе с результатами численных экспериментов дано в работе [6]. В работе [7] рассмотрены различные аспекты реализации этого метода.

3.3.2 Описание метода

Применение длинношагового метода начинается с определения всех точек излома двойственной целевой функции (см. рис. 3.5).

Пусть (см. (1.38))

$$(\lambda_B)_p = s\theta, \tag{3.7}$$

где $(\lambda_B)_p$ — выбранная подходящая двойственная небазисная переменная, $s = +1$ ($s = -1$), если $(\lambda_B)_p$ возрастает (соответственно убывает), $\theta \geq 0$ — скалярный параметр луча. Поскольку остальные двойственные небазисные переменные не изменяются, оставаясь равными нулю, то в соответствии с (1.33) отдельная двойственная базисная переменная $(\lambda_N)_j$ зависит от θ следующим образом (см. (1.39)):

$$(\lambda_N)_j = d_j - s\xi_{pj}\theta, \tag{3.8}$$

где d_j — значение $(\lambda_N)_j$ в текущем базисном решении, ξ_{pj} — элемент симплекс-таблицы $\Xi = -B^{-1}N$. Тогда из (3.8) следует, что точка излома $\theta = \theta_j$ двойственной целевой функции, т. е. точка, где $(\lambda_N)_j$ пересекает нуль, определяется условием (см. (1.40))

$$\theta_j = s(d_j/\xi_{pj}). \tag{3.9}$$

Заметим, что двойственные базисные переменные $(\lambda_N)_j$, для которых $\theta_j < 0$, не принимаются во внимание, так как они уходят от нуля при движении вдоль луча.

Далее необходимо определить величину наклона двойственной целевой функции справа от каждой точки излома (между точками излома эта величина остается постоянной):

$$v_j = \lim_{\theta \to \theta_j + 0} \frac{d\zeta}{d\theta}. \tag{3.10}$$

Из (1.37) следует, что в начальной точке $\theta = \theta_0 = 0$ величина наклона равна

$$v_0 = |r_p|, \tag{3.11}$$

где r_p — относительная оценка выбранной двойственной небазисной переменной $(\lambda_B)_p$ (2.1).

Без ограничения общности можно считать, что $\theta_0 = 0 \le \theta_1 \le \theta_2 \ldots$. Тогда, зная величину наклона v_{j-1} в точке $\theta = \theta_{j-1}$, можно определить величину наклона v_j в точке $\theta = \theta_j$ следующим образом.

Допустим для определенности, что $(\lambda_B)_p = (\lambda_B^+)_p = \theta$, т. е. что выбранная двойственная небазисная переменная возрастает. Понятно, что в областях, не содержащих точек излома, двойственная целевая функция линейно зависит от двойственных переменных, а значит, там справедливо соотношение (1.37). В частности, это соотношение справедливо для луча, соответствующего возрастанию $(\lambda_B)_p$ в диапазоне $\theta_{j-1} \le \theta \le \theta_j$. Тогда, поскольку все остальные двойственные небазисные переменные остаются нулевыми, из (1.37) следует, что в указанном диапазоне:

$$\zeta(\theta) = [(l_B)_p - \beta_p](\lambda_B^+)_p = v_{j-1}\theta, \tag{3.12}$$

где $(l_B)_p$ и $\beta_p = e_p^T(\Xi f + B^{-1}b)$ — соответственно нижняя граница и текущее (в указанном диапазоне) значение базисной переменной $(x_B)_p$ прямой ЛП-задачи, f — вектор активных границ (в указанном диапазоне) небазисных переменных x_N прямой ЛП-задачи, e_p — p-й столбец единичной матрицы. Пусть теперь параметр θ, возрастая, переходит через точку θ_j, вследствие чего двойственная базисная переменная $(\lambda_N)_j$ пересекает свою нулевую границу. С точки зрения прямой ЛП-задачи это означает, что соответствующая небазисная переменная $(x_N)_j$ переходит со своей текущей активной границы на противоположную, которая становится активной, скажем, с нижней границы $(l_N)_j$ на верхнюю $(u_N)_j$, что соответствует убыванию $(\lambda_N)_j$. Это, в свою очередь, приводит к изменению вектора активных границ f небазисных переменных прямой ЛП-задачи, который получает приращение $\Delta f = [(u_N)_j - (l_N)_j]e_j$ (где e_j — j-й столбец единичной матрицы), и как следствие, к изменению величины β_p, которая получает приращение $\Delta\beta_p = e_p^T \Xi \Delta f = \xi_{pj}[(u_N)_j - (l_N)_j]$, где ξ_{pj} — элемент симплекс-таблицы $\Xi = -B^{-1}N$. Но из (3.12) следует, что $\Delta\beta_p$ как раз и есть приращение величины наклона (с противоположным знаком) при пересечении точки излома θ_j, т. е. $v_j = v_{j-1} - \Delta\beta_p$.

Достаточно просто показать, что в общем случае искомые величины наклонов определяются следующим рекуррентным соотношением:

$$v_j = v_{j-1} - |\xi_{pj}|[(u_N)_j - (l_N)_j]. \tag{3.13}$$

Необходимо заметить, что если в прямой ЛП-задаче имеются небазисные переменные с односторонними границами, то соответствующая двойственная базисная переменная, которая *первой* достигает своей нулевой границы, определяет максимальное значение параметра луча. Например, если $(x_N)_j$ находится на своей нижней активной границе $(l_N)_j$ и не имеет верхней границы (формально, $(u_N)_j = +\infty$), и при этом соответствующая двойственная базисная переменная $(\lambda_N)_j$, уменьшаясь, достигает своей нижней нулевой границы при $\theta = \theta_j$, то дальнейшее увеличение θ становится невозможным, так как из (3.13) следует, что $v_j = -\infty$. С точки зрения двойственной ЛП-задачи (1.23)—(1.25) это обусловлено тем, что $(\lambda_N^-)_j = 0$, и поэтому уменьшение комбинированной переменной $(\lambda_N)_j$ ограничено снизу нулем.[19]

[19]Таким образом, если в прямой ЛП-задаче нет переменных с двусторонними ограничениями, то преимущества длинношагового метода теряются, так как он становится эквивалентным обычному методу выбора. Однако эту ситуацию в большинстве случаев удается исправить, поскольку вместо отсутствующих границ переменных можно использовать их искусственные неявные границы, которые являются следствием других ограничений прямой ЛП-задачи.

Получение информации о поведении двойственной целевой функции вдоль луча завершается определением ее приращений относительно начального значения $\zeta(\theta_0) = \zeta(0)$ в точках излома $\theta = \theta_j$ с помощью следующих рекуррентных формул:

$$\Delta\zeta(\theta_0) = 0, \tag{3.14}$$

$$\Delta\zeta(\theta_j) = \Delta\zeta(\theta_{j-1}) + v_{j-1}(\theta_j - \theta_{j-1}). \tag{3.15}$$

Как уже было отмечено в предыдущем подразделе, в качестве двойственной базисной переменной $(\lambda_N)_q$ можно выбрать любую переменную $(\lambda_N)_j$, для которой $\Delta\zeta(\theta_j) > 0$, т. е. которая обеспечивает увеличение двойственной целевой функции. При этом $(\lambda_N)_q$ определяет соответствующую небазисную переменную исходной ЛП-задачи $(x_N)_q$, которую следует ввести в базис.

4 Отыскание начального двойственно допустимого базиса

Как и в случае прямого симплекс-метода [1, разд. 5], для решения ЛП-задачи двойственным симплекс-методом необходимо иметь некоторый начальный базис, который является двойственно допустимым. Поэтому если заданный базис не является двойственно допустимым, то возникает проблема отыскания начального двойственно допустимого базиса.

Так как двойственный симплекс-метод есть прямой симплекс-метод, примененный к двойственной ЛП-задаче, то для отыскания начального двойственно допустимого базиса можно использовать решение вспомогательной задачи, которая является двойственным аналогом задачи [1, (5.18)—(5.19)] и состоит в минимизации кусочно-линейной штрафной функции, выражающей сумму нарушений двойственных ограничений. Примечательным является тот факт, что *кусочно-линейная целевая функция в исходной задаче приводит к двусторонним ограничениям в двойственной задаче и наоборот*, поэтому вспомогательная задача получается в рабочем формате (1.1)—(1.3), что позволяет применить для ее решения двойственный симплекс-метод без каких-либо специальных добавлений.[20]

Заметим, что как и в случае прямого симплекс-метода, формирование начального (не обязательно двойственно допустимого) базиса представляет собой самостоятельную проблему, которая здесь не рассматривается.

4.1 Кусочно-линейная целевая функция

Рассмотрим вначале ЛП-задачу в стандартном формате:

$$\begin{aligned} &z = c^T x \to \min, \\ &Ax = b, \\ &x \geq 0. \end{aligned} \tag{4.1}$$

Соответствующая двойственная ЛП-задача имеет вид:

$$\begin{aligned} &\zeta = b^T \pi \to \max, \\ &A^T \pi + \lambda = c, \\ &\lambda \geq 0, \end{aligned} \tag{4.2}$$

где $\pi = (\pi_i)$ — вектор множителей Лагранжа для ограничений-равенств $Ax = b$, $\lambda = (\lambda_j)$ — вектор множителей Лагранжа для условий неотрицательности $x \geq 0$.

Рассмотрим теперь следующую задачу с сепарабельной кусочно-линейной целевой функцией:

$$\begin{aligned} &z = \sum_j \varphi_j(x_j) \to \min, \\ &Ax = b, \end{aligned} \tag{4.3}$$

где

$$\varphi_j(x_j) = \begin{cases} +\alpha_j x_j, & \text{если } x_j \geq 0, \\ -\beta_j x_j, & \text{если } x_j \leq 0. \end{cases} \tag{4.4}$$

График функции $\varphi_j(x_j)$ для $\alpha_j, \beta_j \geq 0$ показан на рис. 4.1.

[20] В отличие от прямого симплекс-метода, где при решении вспомогательной задачи требуется учитывать наличие неявных вспомогательных переменных.

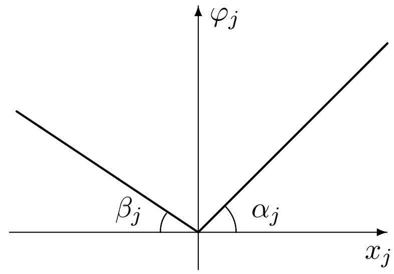

Рис. 4.1. График функции $\varphi_j(x_j)$ для $\alpha_j, \beta_j \geq 0$.

Чтобы преобразовать задачу (4.3) к стандартному формату (4.1), заменим свободные переменные в этой задаче разностью неотрицательных переменных:

$$x = (x_j) = x^+ - x^- = (x_j^+ - x_j^-), \quad x^+, x^- \geq 0.$$

В любом базисном решении задачи (4.3) переменные x_j^+ и x_j^- не могут быть базисными одновременно, поскольку в противном случае в базисной матрице были бы два линейно зависимых столбца. Следовательно, хотя бы одна переменная x_j^+ или x_j^- будет небазисной, а значит, равной нулю. Поэтому:

$$x_j = \begin{cases} +x_j^+, & \text{если } x_j \geq 0, \\ -x_j^-, & \text{если } x_j \leq 0, \end{cases}$$

откуда следует, что:

$$\varphi(x_j) = \alpha_j x_j^+ + \beta_j x_j^-.$$

Таким образом, задача (4.3) принимает стандартный вид:

$$\begin{aligned} & z = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}^T \begin{pmatrix} x^+ \\ x^- \end{pmatrix} \to \min, \\ & (A \mid -A) \begin{pmatrix} x^+ \\ x^- \end{pmatrix} = b, \\ & \begin{pmatrix} x^+ \\ x^- \end{pmatrix} \geq 0, \end{aligned} \tag{4.5}$$

где $\alpha = (\alpha_j)$, $\beta = (\beta_j)$. Двойственную задачу к (4.5) можно получить из (4.2), подставив компоненты задачи (4.5) в (4.1):

$$\begin{aligned} & \zeta = \beta^T \pi \to \max, \\ & (A \mid -A)^T \pi + \begin{pmatrix} \lambda^+ \\ \lambda^- \end{pmatrix} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}, \\ & \begin{pmatrix} \lambda^+ \\ \lambda^- \end{pmatrix} \geq 0, \end{aligned} \tag{4.6}$$

где $\lambda^+ = (\lambda_j^+)$ и $\lambda^- = (\lambda_j^-)$ — множители Лагранжа для условий $x^+ \geq 0$ и $x^- \geq 0$, соответственно. Двойственную систему ограничений-равенств задачи (4.6) можно записать в виде:

$$\begin{aligned} A^T \pi + \lambda^+ &= \alpha, \\ -A^T \pi + \lambda^- &= \beta, \end{aligned}$$

откуда следует, что:

$$\lambda^+ + \lambda^- = \alpha + \beta,$$

а поскольку $\lambda^- \geq 0$, то:

$$\lambda^+ = \alpha + \beta - \lambda^- \leq \alpha + \beta.$$

В результате двойственная задача (4.6) принимает следующий вид:

$$\begin{aligned} &\zeta = b^T\pi \to \max, \\ &A^T\pi + \lambda = \alpha, \\ &0 \leq \lambda \leq \alpha + \beta, \end{aligned}$$

где $\lambda = \lambda^+$. Заметим далее, что:

$$\lambda = -A^T\pi + \alpha,$$

поэтому:

$$0 \leq -A^T\pi + \alpha \leq \alpha + \beta \;\Leftrightarrow\; -\beta \leq A^T\pi \leq \alpha.$$

Это позволяет записать двойственную задачу (4.6) в окончательном виде:

$$\begin{aligned} &\zeta = b^T\pi \to \max, \\ &-\beta \leq A^T\pi \leq \alpha. \end{aligned} \tag{4.7}$$

Обратим внимание, что кусочно-линейная целевая функция в исходной задаче (4.3) приводит к двусторонним ограничениям в соответствующей двойственной задаче (4.7).

4.2 Простой случай поиска допустимого решения

Рассмотрим в качестве примера задачу отыскания допустимого решения простой системы ограничений:

$$\begin{aligned} &Ax = b, \\ &x \geq 0. \end{aligned} \tag{4.8}$$

Указанная задача эквивалентна задаче (4.3), в которой целевая функция имеет смысл суммы невязок (штрафов) для условий неотрицательности $x_j \geq 0$ (рис. 4.2):

$$\varphi_j(x_j) = \begin{cases} 0, & \text{если } x_j \geq 0, \\ -\beta_j x_j, & \text{если } x_j < 0, \end{cases} \tag{4.9}$$

где $\beta = (\beta_j) > 0$ можно рассматривать как вектор весовых коэффициентов невязок для отдельных переменных. Подставляя $\alpha = 0$ в (7), получим соответствующую двойственную задачу:

$$\begin{aligned} &\zeta = b^T\pi \to \max, \\ &-\beta \leq A^T\pi \leq 0, \end{aligned} \tag{4.10}$$

оптимальное базисное решение которой однозначно определяет некоторое базисное допустимое решение исходной системы ограничений (4.8) при условии, что сумма невязок ζ в оптимальной точке равна нулю.

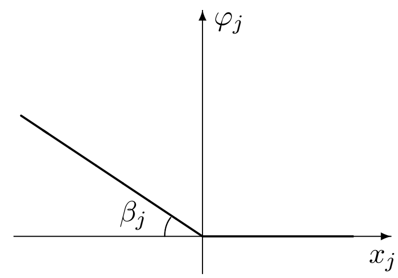

Рис. 4.2. Штраф $\varphi_j(x_j)$ для условия $x_j \geq 0$.

4.3 Поиск допустимого решения в общем случае

Рассмотрим теперь задачу отыскания допустимого решения более общей системы ограничений:

$$\begin{aligned} (A_F\ A_P\ A_N\ A_Z) \begin{pmatrix} x_F \\ x_P \\ x_N \\ x_Z \end{pmatrix} &= b, \\ -\infty < x_F &< +\infty, \\ 0 \leq x_P &< +\infty, \\ -\infty < x_N &\leq 0, \\ x_Z &= 0, \end{aligned} \tag{4.11}$$

где x_F, x_P, x_N, x_Z — подвекторы соответственно свободных (неограниченных по знаку), неотрицательных, неположительных, нулевых переменных. Как и в простом случае (см. предыдущий подраздел) указанная задача эквивалентна задаче (4.3), в которой целевая функция представляет собой сумму невязок. Однако теперь возможны четыре типа невязок $\varphi_j(x_j)$ в зависимости от типа границ переменной x_j (рис. 4.3). При этом соответствующая двойственная задача (4.7) будет следующей:

$$\zeta = \beta^T \pi \to \max,$$
$$\begin{pmatrix} 0 \\ -\beta_P \\ 0 \\ -\beta_Z \end{pmatrix} \leq (A_F\ A_P\ A_N\ A_Z)^T \pi \leq \begin{pmatrix} 0 \\ 0 \\ \alpha_N \\ \alpha_Z \end{pmatrix},$$

или в окончательном виде:

$$\begin{aligned} \zeta = b^T \pi &\to \max, \\ A_F^T \pi &= 0, \\ -\beta_P \leq A_P^T \pi &\leq 0, \\ 0 \leq A_N^T \pi &\leq \alpha_N, \\ -\beta_Z \leq A_Z^T \pi &\leq \alpha_Z. \end{aligned} \tag{4.12}$$

Заметим, что каждой переменной x_j исходной системы ограничений (4.11) соответствует ограничение двойственной задачи (4.12), при этом тип границ переменной x_j однозначно определяет границы соответствующего двойственного ограничения.

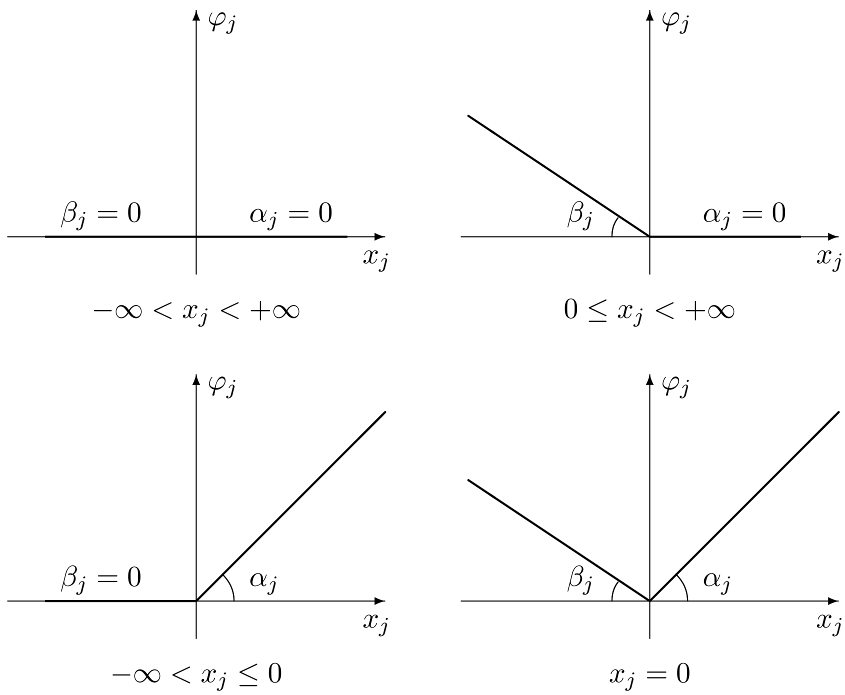

Рис. 4.3. Штрафы $\varphi_j(x_j)$ для различных типов границ x_j.

4.4 Поиск двойственно допустимого решения ЛП-задачи в рабочем формате

Запишем ЛП-задачу в рабочем формате (1.1)—(1.3) следующим образом:[21]

$$
\begin{aligned}
& z = c^T x = c_F^T x_F + c_L^T x_L + c_U^T x_U + c_D^T x_D \to \min, \\
& Ax = (A_F \ A_L \ A_U \ A_D) \begin{pmatrix} x_F \\ x_L \\ x_U \\ x_D \end{pmatrix} = b, \\
& -\infty < x_F < +\infty, \\
& \quad l \le x_L < +\infty, \\
& -\infty < x_U \le u, \\
& \quad l_D \le x_D \le u_D,
\end{aligned}
\tag{4.13}
$$

где x_F, x_L, x_U, x_D — подвекторы соответственно свободных (неограниченных по знаку), ограниченных снизу, ограниченных сверху, ограниченных с двух сторон переменных.[22]

[21] Постоянный член c_0 целевой функции опущен, чтобы не загромождать формулы.

[22] Фиксированные переменные можно рассматривать как переменные, ограниченные с двух сторон, у которых нижняя и верхняя границы совпадают.

Из условий оптимальности (1.17)—(1.22) следует, что соответствующая двойственная система ограничений для задачи (4.13) имеет следующий вид:

$$
\begin{aligned}
&A^T\pi + \lambda = (A_F\ A_L\ A_U\ A_D)^T\pi + \begin{pmatrix}\lambda_F\\ \lambda_L\\ \lambda_U\\ \lambda_D\end{pmatrix} = c = \begin{pmatrix}c_F\\ c_L\\ c_U\\ c_D\end{pmatrix}, \\
&-\infty < \ \pi \ < +\infty, \\
&\qquad \lambda_F = 0, \\
&\qquad \lambda_L \geq 0, \\
&\qquad \lambda_U \leq 0, \\
&-\infty < \lambda_D < +\infty,
\end{aligned}
\tag{4.14}
$$

где π — вектор множителей Лагранжа для ограничений-равенств $Ax = b$, λ_F, λ_L, λ_U, λ_D — подвекторы множителей Лагранжа для границ переменных x_F, x_L, x_U, x_D соответственно.

Заметим, что система ограничений (4.14) совпадает с системой ограничений (4.11), поэтому допустимое решение системы (4.14) можно найти в результате решения соответствующей двойственной ЛП-задачи. Чтобы сделать матрицу системы (4.14) более наглядной, объединим все двойственные переменные в один вектор:

$$
\begin{gathered}
\begin{pmatrix}A_F^T & I & & & \\ A_L^T & & I & & \\ A_U^T & & & I & \\ A_D^T & & & & I\end{pmatrix}\begin{pmatrix}\pi\\ \lambda_F\\ \lambda_L\\ \lambda_U\\ \lambda_D\end{pmatrix} = \begin{pmatrix}c_F\\ c_L\\ c_U\\ c_D\end{pmatrix}, \\
\begin{pmatrix}-\infty\\ 0\\ 0\\ -\infty\\ -\infty\end{pmatrix} \leq \begin{pmatrix}\pi\\ \lambda_F\\ \lambda_L\\ \lambda_U\\ \lambda_D\end{pmatrix} \leq \begin{pmatrix}+\infty\\ 0\\ +\infty\\ 0\\ +\infty\end{pmatrix}.
\end{gathered}
\tag{4.15}
$$

Рассматривая систему (4.15) с точки зрения системы (4.11) и используя вектор переменных $(-y) = (-y_F, -y_L, -y_U, -y_D)$ в качестве вектора двойственных переменных для ограничений-равенств системы (4.15), можно построить соответствующую двойственную ЛП-задачу (4.12), которая в обозначениях системы (4.15) будет следующей:

$$
\begin{aligned}
&-c_F^T y_F - c_L^T y_L - c_U^T y_U - c_D^T y_D \to \max, \\
&-A_F y_F - A_L y_L - A_U y_U - A_D y_D = 0, \\
&-\beta_F \leq -y_F \leq \alpha_F, \\
&-\beta_L \leq -y_L \leq 0, \\
&\quad\ 0 \leq -y_U \leq \alpha_U, \\
&\qquad\quad -y_D = 0,
\end{aligned}
$$

или, если умножить все компоненты на (-1):

$$
\begin{aligned}
&c_F^T y_F + c_L^T y_L + c_U^T y_U + c_D^T y_D = c^T y \to \min,\\
&A_F y_F + A_L y_L + A_U y_U + A_D y_D = Ay = 0,\\
&-\alpha_F \le y_F \le \beta_F,\\
&\quad\ \ 0 \le y_L \le \beta_L,\\
&-\alpha_U \le y_U \le 0,\\
&\qquad\quad\ y_D = 0,
\end{aligned}
\tag{4.16}
$$

где α_j и β_j — весовые коэффициенты невязок для переменных λ_j.

Таким образом, базисное оптимальное решение ЛП-задачи (4.16) будет двойственно допустимым для исходной ЛП-задачи (4.13) при условии, что сумма невязок $c^T y$ в оптимальной точке равна нулю.

С точки зрения реализации рассмотренного метода поиска двойственно допустимого решения существенными являются следующие два обстоятельства.

1. Исходная ЛП-задача (4.13) и вспомогательная ЛП-задача (4.16) имеют один и тот же вектор коэффициентов целевой функции и матрицу коэффициентов ограничений и отличаются лишь правыми частями ограничений-равенств и границами переменных. Поэтому любой вариант двойственного симплекс-метода, предназначенного для решения исходной задачи (4.13), можно использовать без каких-либо изменений и для решения вспомогательной задачи (4.16).

2. Для задачи (4.16) двойственно допустимым является любое базисное решение, поскольку все переменные в этой задаче имеют конечные нижние и верхние границы. Поэтому решение этой задачи можно начинать с любого заданного базиса.

Литература

[1] А. О. Махорин. Прямой симплекс-метод (обоснование и реализация).

[2] C. E. Lemke. The Dual Method of Solving the Linear Programming Problem. Carnegie Institute of Technology, Department of Mathematics, Technical Report No. 29, March 4, 1953.

[3] G. B. Dantzig. The Dual Simplex Algorithm. RAND RM-1270, 1954.

[4] P. M. J. Harris. Pivot Selection Methods of the Devex LP Code. Mathematical Programming 12, 1973, pp. 1-28.

[5] Р. Габасов, Ф. М. Кириллова, О. И. Костюкова. К методам решения общей задачи линейного программирования. Препринт Ин-та математики АН БССР, 14 (30).— Минск: 1977.

[6] E. Kostina. The long step rule in the bounded-variable dual simplex method: Numerical experiments. Math. Meth. Op. Res. 55, 2002, pp. 413–29.

[7] A. Koberstein. The Dual Simplex Method: Techniques for a fast and stable implementation. Dissertation at the University of Paderborn, Fakultät für Wirtschaftswissenschaften, Department Wirtschaftsinformatik, 2005.

Приложение

Программная реализация

П.1 Модуль SPYCHUZR — двойственный выбор базисной переменной

П.1.1 spy_chuzr_sel — отбор подходящих переменных

```
int spy_chuzr_sel(SPXLP *lp, const double beta[/*1+m*/], double tol,
   double tol1, int list[/*1+m*/]);
```

Данная подпрограмма выполняет отбор подходящих базисных переменных, т. е. базисных переменных $(x_B)_i$, значения β_i которых в текущем базисном решении нарушают соответствующую нижнюю $(l_B)_i$ или верхнюю $(u_B)_i$ границу.

Компоненты вектора значений базисных переменных $\beta = (\beta_i)$ должны находиться в ячейках массива `beta[1]`, ..., `beta[m]`.

Базисная переменная $(x_B)_i$ считается подходящей, если нарушена ее нижняя граница:

$$\beta_i < (l_B)_i - \delta_i \text{ для } \delta_i = tol + tol_1|(l_B)_i|,$$

или верхняя граница:

$$\beta_i > (u_B)_i + \delta_i \text{ для } \delta_i = tol + tol_1|(u_B)_i|,$$

где `tol` и `tol1` — заданные допуски.

На выходе подпрограмма записывает номера i подходящих базисных переменных $(x_B)_i$ в ячейки массива `list[1]`, ..., `list[num]`, где общее число таких переменных $0 \leq$ `num` $\leq m$ возвращается в качестве результата. (Если указатель на массив `list` задан как `NULL`, то номера базисных переменных не записываются.)

П.1.2 spy_chuzr_std — выбор базисной переменной по двойственному правилу Данцига

```
int spy_chuzr_std(SPXLP *lp, const double beta[/*1+m*/], int num, const int
   list[]);
```

Данная подпрограмма выполняет «стандартный» выбор наиболее подходящей базисной переменной $(x_B)_p$ в соответствии с двойственным правилом Данцига (см. подразд. 2.1):

$$|r_p| = \max_{i \in I} |r_i|,$$

где $I \subseteq \{1, \ldots, m\}$ — множество номеров подходящих базисных переменных, r_i — величина нарушения границы переменной $(x_B)_i$, определяемая следующим образом:

$$r_i = \begin{cases} (l_B)_i - \beta_i, & \text{если } \beta_i < (l_B)_i \\ 0, & \text{если } (l_B)_i \leq \beta_i \leq (u_B)_i \\ (u_B)_i - \beta_i, & \text{если } \beta_i > (u_B)_i \end{cases}$$

Компоненты вектора значений базисных переменных $\beta = (\beta_i)$ должны находиться в ячейках массива `beta[1]`, ..., `beta[m]`.

Номера подходящих базисных переменных $i \in I$ должны находиться в ячейках массива `list[1]`, ..., `list[num]`, где `num` $= |I| > 0$ — общее число таких переменных.

На выходе подпрограмма возвращает номер p выбранной базисной переменной $(x_B)_p$.

П.1.3 Структурный тип SPYSE

Структурный тип данных `SPYSE` используется подпрограммами, реализующими двойственный проекционный метод наиболее крутого ребра (см. подразд. 2.3) и двойственный метод оценивания Devex (см. подразд. 2.4).

Структура `SPYSE` имеет следующие поля:

`int valid` — если данный признак установлен, то содержимое структуры является действительным. В противном случае, если данный признак сброшен, содержимое структуры не действительно.

`char *refsp` — указатель на массив длины $1+n$. Ячейка `refsp[0]` не используется. Ячейка `refsp[k]`, $1 \leq k \leq n$, содержит признак того, что переменная x_k принадлежит эталонному пространству переменных $V \subseteq E^n$.

`double *gamma` — указатель на массив длины $1 + m$. Ячейка `gamma[0]` не используется. Ячейка $\gamma[i]$, $1 \leq i \leq m$, содержит весовой множитель γ_i (2.8) для нарушения границы r_i (2.1) базисной переменной $(x_B)_i$ в текущем базисе.

`double *work` — указатель на рабочий массив длины $1 + m$.

П.1.4 spy_alloc_se — размещение массивов

```
void spy_alloc_se(SPXLP *lp, SPYSE *se);
```

Данная подпрограмма выполняет отведение памяти для всех массивов, входящих в состав объекта `se` (блок данных двойственного проекционного метода наиболее крутого ребра и двойственного метода оценивания Devex), используя для этого информацию о размерности рабочей ЛП-задачи.

П.1.5 spy_reset_refsp — переопределение эталонного пространства

```
void spy_reset_refsp(SPXLP *lp, SPYSE *se);
```

Данная подпрограмма выполняет переопределение эталонного пространства, включая в него переменные, которые являются базисными в текущем базисе, а также устанавливает все весовые множители γ_i равными единице.

П.1.6 spy_eval_gamma_i — прямое вычисление весового множителя

```
double spy_eval_gamma_i(SPXLP *lp, SPYSE *se, int i);
```

Данная подпрограмма выполняет непосредственное вычисление весового множителя γ_i, $1 \leq i \leq m$, для текущего базиса по формуле (2.8).

П.1.7 spy_chuzr_pse — выбор базисной переменной по двойственному методу наиболее крутого ребра

```
int spy_chuzr_pse(SPXLP *lp, SPYSE *se, const double beta[/*1+m*/], int num,
    const int list[]);
```

Данная подпрограмма выполняет выбор наиболее подходящей базисной переменной $(x_B)_p$ в соответствии с двойственным методом наиболее крутого ребра (2.4):

$$\frac{r_p^2}{\gamma_p} = \max_{i \in I} \frac{r_i^2}{\gamma_i},$$

где $I \subseteq \{1, \dots, m\}$ — множество номеров подходящих базисных переменных, r_i — величина нарушения границы переменной $(x_B)_i$ (см. подразд. П.1.2), γ_i — весовой множитель нарушения границы.

Компоненты вектора значений базисных переменных $\beta = (\beta_i)$ должны находиться в ячейках массива `beta[1]`, ..., `beta[m]`.

Номера подходящих базисных переменных $i \in I$ должны находиться в ячейках массива `list[1]`, ..., `list[num]`, где `num` $= |I| > 0$ — общее число таких переменных.

На выходе подпрограмма возвращает номер p выбранной базисной переменной $(x_B)_p$.

П.1.8 spy_update_gamma — пересчет весовых множителей для смежного базиса

```
double spy_update_gamma(SPXLP *lp, SPXSE *se, int p, int q, const double
    trow[/*1+n-m*/], const double tcol[/*1+m*/]);
```

Данная подпрограмма выполняет точный пересчет вектора весовых множителей $\gamma = (\gamma_i)$ (2.8), используемых в двойственном проекционном методе наиболее крутого ребра, для смежного базиса.

Содержимое объекта `se` должно быть действительным и должно соответствовать текущему базису.

Параметр $1 \le p \le m$ задает номер базисной переменной $(x_B)_p$, которая становится небазисной переменной $(\overline{x_N})_q$ в смежном базисе.

Параметр $1 \le q \le n - m$ задает номер небазисной переменной $(x_N)_q$, которая становится базисной переменной $(\overline{x_B})_p$ в смежном базисе.

Массив `trow` должен содержать элементы p-й (ведущей) строки симплекс-таблицы $\xi_p = (\xi_{pj})$ в ячейках `trow[1]`, ..., `trow[n-m]`. (Эта строка должна соответствовать текущему базису.)

Массив `tcol` должен содержать элементы q-го (ведущего) столбца симплекс-таблицы $\Xi_q = (\xi_{iq})$ в ячейках `tcol[1]`, ..., `tcol[m]`. (Этот столбец должен соответствовать текущему базису.)

Для пересчета весовых множителей подпрограмма использует алгоритм, приведенный в подразд. 2.3.

Подпрограмма также вычисляет относительную ошибку:

$$e = \frac{|\gamma_p - \gamma'_p|}{1 + |\gamma_p|},$$

где γ'_p — менее точное значение весового множителя γ_p на входе в подпрограмму, и возвращает e в качестве результата. (Вызывающая программа может переопределить эталонное пространство, если величина e оказывается слишком большой, так как в этом случае другие весовые множители вероятно тоже являются недостаточно точными.)

П.1.9 spy_free_se — освобождение массивов

```
void spy_free_se(SPXLP *lp, SPXSE *se);
```

Данная подпрограмма выполняет освобождение памяти, занятой всеми массивами, входящими в состав объекта `se`.

П.2 Модуль SPYCHUZC — двойственный выбор небазисной переменной

П.2.1 spy_chuzc_std — «стандартный» двойственный выбор небазисной переменной

```
int spy_chuzc_std(SPXLP *lp, const double d[/*1+n-m*/], double r,
   const double trow[/*1+n-m*/], double tol_piv, double tol, double tol1);
```

Данная подпрограмма реализует улучшенный «стандартный» двойственный метод выбора небазисной переменной $(x_N)_q$ (см. подразд. 3.1) применительно к ЛП-задаче в рабочем формате (1.1)—(1.3).

Компоненты вектора относительных оценок небазисных переменных $d = (d_j)$ должны находиться в ячейках массива `d[1]`, ..., `d[n-m]`. Предполагается, что в пределах допуска эти оценки имеют «правильные» знаки.

Параметр `r` задает нарушение границы выбранной базисной переменной $(x_B)_p$. Если $r = (l_B)_p - \beta_p > 0$, то $(x_B)_p$ нарушает свою нижнюю границу, и в этом случае двойственная небазисная переменная $(\lambda_B)_p = (\lambda_B^+)_p$ *возрастает*. Если же $r = (u_B)_p - \beta_p < 0$, то $(x_B)_p$ нарушает свою нижнюю границу, и в этом случае двойственная небазисная переменная $(\lambda_B)_p = (\lambda_B^-)_p$ *убывает*. (Заметим, что r есть относительная оценка двойственной переменной $(\lambda_B)_p$.)

Элементы строки текущей симплекс-таблицы $\xi_p = (\xi_{pj})$, соответствующей выбранной базисной переменной $(x_B)_p$, должны находиться в ячейках массива `trow[1]`, ..., `trow[n-m]`.

Параметр `tol_piv` задает допуск на абсолютную величину элементов симплекс-таблицы. Если $|\xi_p j| <$ `tol_piv`, то двойственная базисная переменная $(\lambda_N)_j$ пропускается, т. е. считается, что эта переменная не зависит от параметра луча θ.

Параметры `tol` и `tol1` задают допуски, которые используются для создания искусственной двойственной вырожденности с целью увеличения свободы выбора ведущего элемента. Если $(\lambda_N)_j = (\lambda_N^+)_j \geq 0$ и $d_j \leq +\delta_j$ или если $(\lambda_N)_j = (\lambda_N^-)_j \leq 0$ и $d_j \geq -\delta_j$, где $\delta_j = tol + tol_1|(c_N)_j|$ и $(c_N)_j$ — коэффициент целевой функции при небазисной переменной $(x_N)_j$, то считается, что в пределах допуска относительная оценка d_j равна нулю.

Подпрограмма определяет номер $1 \leq q \leq n - m$ небазисной переменной $(x_N)_q$, для которой соответствующая двойственная базисная переменная $(\lambda_N)_q = (\lambda_N^+)_q$ или $(\lambda_N)_q = (\lambda_N^-)_q$, которая первой достигает своей нулевой границы при возрастании параметра двойственного луча $\theta \geq 0$, и возвращает номер q в качестве результата. В том случае, когда возможно неограниченное возрастание параметра θ, подпрограмма возвращает нуль.

П.2.2 spy_chuzc_harris — выбор небазисной переменной двойственным методом Харрис

```
int spy_chuzc_harris(SPXLP *lp, const double d[/*1+n-m*/], double r,
   const double trow[/*1+n-m*/], double tol_piv, double tol, double tol1);
```

Данная подпрограмма реализует двойственный метод Харрис выбора небазисной переменной $(x_N)_q$ (см. подразд. 3.2) применительно к ЛП-задаче в рабочем формате (1.1)—(1.3).

Все параметры подпрограммы, за исключением `tol` и `tol1`, а также возвращаемое значение имеют тот же смысл, что и для подпрограммы `spy_chuzc_std` (см. выше).

Параметры `tol` и `tol1` задают допуски на нарушение нулевых границ относительных оценок небазисных переменных. Для относительной оценки d_j такой допуск принимается равным $\delta_j = tol + tol_1|(c_N)_j|$, где $(c_N)_j$ — коэффициент целевой функции при небазисной переменной $(x_N)_j$.

П.2.3 Структурный тип SPYBP

Структурный тип данных `SPYBP` определяет отдельную точку излома двойственной целевой функции. Этот тип данных используется подпрограммами, реализующими двойственный длинношаговый выбор небазисной переменной (см. подразд. 3.3).

Структура `SPYBP` имеет следующие поля:

`int j` — номер двойственной базисной переменной $(\lambda_N)_j$, которая пересекает нуль в соответствующей точке излома при возрастании параметра луча θ.

`int teta` — значение параметра луча $\theta = \theta_j \geq 0$ в соответствующей точке излома (см. (3.9)).

`double dz` — приращение двойственной целевой функции $\Delta\xi(\theta_j)$ в соответствующей точке излома (см. (3.15)).