УДК 512.25/26

ОБ ОДНОМ ДЕКОМПОЗИЦИОННОМ АЛГОРИТМЕ ЛИНЕЙНОГО ПРОГРАММИРОВАНИЯ

Н.В.Шмырёва

Часто в задачах линейного программирования большой размерности фиксация определенной части переменных (будем называть их управляющими) позволяет расчленить задачу на ряд более простых. В этом случае исходная задача сводится к минимизации кусочно-линейной выпуклой функции от управляющих переменных. При этом для вычисления значения функции в конкретной точке необходимо решить одну или несколько вспомогательных задач линейного программирования более простой структуры, чем исходная задача.

Этот прием рассматривался многими авторами, причем для минимизации возникающей кусочно-линейной функции применялись итеративные процедуры теории матричных игр, обобщенного градиентного спуска (см. [2 , 6]) или алгоритмы эвристического характера. В частности, в работе [3] для специальной задачи распределения централизованных ресурсов была сделана попытка применить схему метода возможных направлений. Однако при этом в процедуре определения возможного направления требовалось перечисление всех оптимальных базисов вспомогательных задач, что практически осуществимо лишь в простейших случаях. Для более сложных случаев авторы предлагают приближенную эвристическую процедуру.

В настоящей работе предлагается метод, относящийся к классу методов возможных направлений с квадратичной нормализацией. Для нахождения подходящего направления используется идея, близ-

кая к описанной в [4], где в качестве направления выбирается элемент с минимальной нормой из выпуклой оболочки множества субградиентов минимизируемой функции и нормалей ограничений, проходящих через данную точку.

Трудность непосредственного применения этого подхода в данном случае заключается в том, что множество субградиентов функции управляющих переменных, вообще говоря, не ограничено. Ввиду этого, процедура поиска элемента с минимальной нормой, используемая в [4], в данном случае в чистом виде не применима. Для преодоления указанных затруднений осуществляется переход от неограниченного множества субградиентов к некоторому многограннику. Для этого многогранника излагается итеративная процедура, которая хотя и не всегда позволяет получить элемент с минимальной нормой, однако позволяет за конечное число шагов определить требуемое подходящее направление, если такое существует. Важно отметить, что на каждом шаге этой процедуры не требуется перечисление всех вершин многогранника, а используется лишь одна из них.

§ I. Формулировка задачи

Рассмотрим задачу линейного программирования вида:
минимизировать $(c,x)+(d,z)$
при ограничениях

$$Ax+Bz=b, \quad x\geq 0,\ z\geq 0. \qquad (1)$$

Здесь $c,x\in R^{n_1}$, $d,z\in R^{n_2}$; $b\in R^{m}$; A и B – матрицы соответствующих размерностей. Предполагая, что матрица A имеет простую структуру, принимаем переменные z в качестве управляющих. Тем самым переходим к задаче минимизации выпуклой кусочно-линейной функции

$$F(z)=(d,z)+\min_{\substack{Ax+Bz=b\\ x\geq 0}}(c,x) \qquad (2)$$

на множестве неотрицательных z, при которых функция $F(z)$ определена. Множество таких z будем обозначать через Ω.

Если $z^o \in \Omega$, то для вычисления значения $F(z^o)$ необходимо при $z = z^o$ решить следующую задачу линейного программирования: минимизировать (c, x) при условиях

$$A x = b - Bz, \qquad (3)$$
$$x \geqslant 0.$$

Пусть y^o – оптимальный вектор задачи, двойственной к (3), определяемый каким-либо из оптимальных базисов. Если область оптимальности этого базиса не вырождается в точку, то при изменении $z = z^o + \Delta z$ в пределах этой области для функции $F(z)$ будем иметь:

$$F(z^o + \Delta z) = F(z^o) + (d - y^o B, \Delta z). \qquad (4)$$

Если же $z^o + \Delta z$ выходит за пределы этой области, будем иметь неравенство

$$F(z^o + \Delta z) > F(z^o) + (d - y^o B, \Delta z). \qquad (5)$$

Тем самым вектор $d - y^o B$ является одним из субградиентов функции F в точке z^o.

Пусть x^o – какое-либо оптимальное решение задачи (3), соответствующее точке $z^o \in \Omega$.

Покажем, что субградиентом функции F в точке z^o является любой вектор вида $d - yB$, если y – решение следующей системы:

$$y A^j \leqslant c_j, \quad j \in J^o(x^o) = \{ j \mid x_j^o = 0 \},$$
$$y A^j = c_j, \quad j \bar{\in} J^o(x^o). \qquad (6)$$

Действительно, рассмотрим функцию F в точке $z^o + \Delta z \in \Omega$. Этой точке соответствуют решение задачи (3) – вектор $\bar{x}$ и решение задачи, двойственной к (3),– вектор $\bar{y}$. Любой вектор y, удовлетворяющий (6), представляет собой оптимальное решение задачи, двойственной к (3) при $z = z^o$ и допустимое решение той же задачи при $z = z^o + \Delta z$. А тогда справедливо следующее:

$$F(z^o + \Delta z) = (c, \bar{x}) + (d, z^o + \Delta z) = (\bar{y}, b - B(z^o + \Delta z)) +$$

$$+(d, z^o+\Delta z) \geq (y, b-B(z^o+\Delta z)) + (d, z^o+\Delta z) =$$
$$=(y, b-Bz^o)+(d, z^o)+(d-yB, \Delta z) = (c, x^o)+(d, z^o)+$$
$$+(d-yB, \Delta z) = F(z^o)+(d-yB, \Delta z),$$

что и требовалось доказать.

ЗАМЕЧАНИЕ. Несложно показать, что множество решений системы (6) не зависит от выбора оптимального вектора x^o задачи (3) и определяется лишь точкой z^o.

§ 2. Признаки оптимальности

Описанное выше множество субградиентов функции F позволяет переформулировать признак оптимальности для задачи (I), а именно: если 0 принадлежит множеству

$$S(z^o) = \{h \mid h = d - yB - \sum_{i \in I^o(z^o)} t_i e^i,\ t_i \geq 0\},$$

где $I^o(z^o) = \{i \mid z_i^o = 0\}$, а y удовлетворяет системе (6), то x^o, z^o образуют решение исходной задачи.

Действительно, в этом случае $yB^i \leq d_i$ (B^i - i-й столбец матрицы B, а d_i - соответствующая компонента вектора d), и при этом $yB^i = d_i$, если $i \bar{\in} I^o(z^o)$, что вместе с условиями (6) образует условия критерия оптимальности для задачи (I).

Сформулированный признак оптимальности приводит к следующей задаче квадратичного программирования:

$$\|h\|^2 \longrightarrow min!$$
$$h \in S(z^o). \tag{7}$$

Решение этой задачи позволяет либо установить оптимальность имеющихся векторов x^o и z^o, либо (в случае, когда эти векторы не образуют решения исходной задачи) получить подходящее направление вариации векторов x^o и z^o. Действительно, если $\bar{y}$ и $\bar{t}_i$ $(i \in I^o(z^o))$ - решение задачи (7), то в последнем случае $\|\bar{h}\| > 0$. Из критерия оптимальности для

задачи (7) следует существование такого вектора w, что выполняются соотношения:

$$Aw = -B\bar{h}, \tag{8}$$

$$w_j \leq 0, \quad j \in J^0(x^0), \tag{9}$$

$$(\bar{y}A - c, w) = 0, \tag{10}$$

$$\bar{h}_i \leq 0, \quad i \in I^0(z^0), \tag{11}$$

$$\bar{t}_i \bar{h}_i = 0, \quad i \in I^0(z^0). \tag{12}$$

Из этих соотношений следует, что векторы $x^0 - \lambda w$, $z^0 - \lambda \bar{h}$ при достаточно малых $\lambda > 0$ образуют допустимое решение исходной задачи (1). При этом с ростом λ целевая функция $(c,x) + (d,z) = (c,x^0) + (d,z^0) + \lambda((c,w) + (d,h))$ убывает, ибо

$$\|\bar{h}\|^2 = (d,\bar{h}) - (\bar{y}B,\bar{h}) - \sum_{i \in I^0(z^0)} \bar{t}_i \bar{h}_i ,$$

и учитывая (8), (10) и (12), получаем

$$\|\bar{h}\|^2 = (d,\bar{h}) + (c,w) > 0 .$$

Непосредственное решение задачи (7) конечными методами хотя и позволяет построить процедуру метода возможных направлений, однако приводит к неоправданному усложнению общего алгоритма. Применение же итеративных алгоритмов для решения этой задачи в общем случае не приводит к определению подходящего направления за конечное число шагов. Это легко видеть из таких рассуждений.

Множество решений системы (6) в общем случае представляет собой сумму подпространства, многогранника и конечно-порожденного конуса. Однако если предполагать, что ранг этой системы равен m, то подпространство будет нулевым, и любой вектор y, удовлетворяющий системе (6), будет иметь вид:

$$y = \sum_{\ell=1}^{L} \alpha_\ell y^\ell + \sum_{\kappa=1}^{K} \beta_\kappa \eta^\kappa , \tag{13}$$

где α_ℓ, $\beta_\kappa \geq 0$ и $\sum_\ell \alpha_\ell = 1$.

Тем самым для $h \in S(z^0)$ имеем:

$$h = \sum_{\ell=1}^{\mathcal{L}} \alpha_\ell (d - y^\ell B) - \sum_{\kappa=1}^{K} \beta_\kappa q^\kappa B - \sum_{i \in I^\circ(x^\circ)} t_i e^i . \qquad (14)$$

Для того чтобы вектор $-h$ определял подходящее направление вариации точки x°, необходимо, чтобы было $q^\kappa B h \le 0$ для всех $\kappa = 1, 2, \dots K$. Несложно показать, что в противном случае замена x° на $x^\circ - \lambda h$ приведет к тому, что задача (3) не будет иметь допустимых решений ни при каком $\lambda > 0$. С другой стороны, если для вектора $\bar{h}$ в представлении (14) некоторый коэффициент β_κ положителен, то $(q^\kappa B, h) = 0$. (Это следует из признака оптимальности для задачи (7).) Незначительно уменьшая β_κ, мы по-прежнему будем получать векторы h из $S(x^\circ)$, однако они уже не будут определять подходящих направлений, ибо для них будет $(q^\kappa B, h) > 0$. Тем самым в любой окрестности вектора $\bar{h}$ существуют векторы $h \in S(x^\circ)$, не определяющие подходящих направлений. Поэтому мы не можем гарантировать, что применяя итеративные алгоритмы для решения квадратичной задачи (7), мы сможем получить подходящее направление за конечное число шагов.

Для преодоления указанного затруднения мы введем в рассмотрение многогранник $P(x^\circ)$, вершинами которого являются точки $d - y^\ell B$, $\ell = 1, \dots, \mathcal{L}$; $-q^\kappa B$, $\kappa = 1, \dots, K$; $-e^i$, $i \in I^\circ(x^\circ)$.

Легко видеть, что из $0 \in S(x^\circ)$ следует $0 \in P(x^\circ)$. Покажем, что при условии телесности множества Ω справедливо и обратное утверждение.

ЛЕММА I. Если Ω телесно и $0 \in P(x^\circ)$, то $0 \in S(x^\circ)$.

ДОКАЗАТЕЛЬСТВО. Так как $0 \in P(x^\circ)$, то найдутся такие неотрицательные числа α_ℓ, $\ell = 1, \dots, \mathcal{L}$; β_κ, $\kappa = 1, \dots, K$; t_i, $i \in I^\circ(x^\circ)$, в сумме равные единице, что $0 = \sum_\ell \alpha_\ell (d - y^\ell B) - \sum_\kappa \beta_\kappa q^\kappa B - \sum_{i \in I^\circ(x^\circ)} t_i e^i$. Если $\sum \alpha_\ell > 0$, то, разделив это равенство на $\sum \alpha_\ell$, получим $0 \in S(x^\circ)$. Пусть $\sum \alpha_\ell = 0$. Тогда

$$0 = -\sum_\kappa \beta_\kappa q^\kappa B - \sum_{i \in I^\circ(x^\circ)} t_i e^i , \qquad (15)$$

причем среди β_k и t_i есть ненулевые. Покажем, что в этом случае множество Ω нетелесно. Векторы η^k являются решениями однородной системы

$$\eta^k A^j \leqslant 0, \quad j \in J^0(x^0),$$
$$\eta^k A^j = 0, \quad j \bar{\in} J^0(x^0). \qquad (16)$$

Для того чтобы для некоторого z существовало решение задачи (3), необходимо, чтобы для каждого η^k выполнялось неравенство

$$(b - Bz, \eta^k) \leqslant 0$$

(в противном случае в задаче, двойственной к (3), целевая функция не ограничена). Таким образом, z удовлетворяет системе неравенств

$$\eta^k Bz \geqslant \eta^k b, \quad k = 1, \dots, K.$$

Для точки z^0 все эти неравенства обращаются в равенства. Действительно, если умножить условия (16) на соответствующие компоненты вектора x^0 и сложить, получим требуемое.

Таким образом, векторы $-\eta^k B$ и орты $-e^i$, $i \in I^0(z^0)$, являются нормалями гиперплоскостей, задающих границу множества Ω и проходящих через точку z^0. Теперь ясно, что из (15) следует нетелесность множества Ω.

Таким образом, если ранг матрицы A равен m и множество Ω телесно, то для оптимальности векторов x^0 и z^0 в задаче (1) необходимо и достаточно, чтобы $0 \in P(z^0)$.

§ 3. Определение подходящего направления

Замена множества $S(z^0)$ многогранником $P(z^0)$ в критерии оптимальности позволяет отказаться от точного решения квадратичной задачи (7): если $0 \bar{\in} P(z^0)$, то для элемента с минимальной нормой $\bar{h} \in P(z^0)$ будет

$$(d - y^l B, \bar{h}) > 0, \quad l = 1, \dots, L;$$

$$(-\eta^k B, \bar{h}) > 0, \quad k = 1, \dots, K;$$

$$(-e^i, \bar{h}) > 0, \qquad i \in I^0(z^0).$$

Но тогда эти неравенства будут выполняться и для всех $h \in P(z^0)$, достаточно близких к $\bar{h}$. Поэтому применение любой итеративной процедуры для поиска $\bar{h}$ даст подходящее направление за конечное число шагов.

Однако построение итеративных процедур для поиска $\bar{h}$ в рассматриваемом случае оказывается затруднительным, ввиду того, что не все вершины многогранника $P(z^0)$ заданы явно, и могут быть получены лишь путем минимизации некоторой линейной функции (f, y) на множестве решений системы (6). Если такая задача окажется разрешимой, мы получим одну из вершин $d - y^\ell B$, если же оказывается, что минимизируемая функция не ограничена снизу, то мы получим одну из вершин $(-\eta^\kappa B)$. Существенно, что последнее будет иметь место лишь в случае, когда $(-\eta^\kappa B, f) < 0$. Ввиду этого процедура метода условного градиента, используемая в [4] для получения элемента с минимальной нормой в многограннике, к многограннику $P(z^0)$, вообще говоря, неприменима.

Мы рассмотрим некоторое видоизменение метода условного градиента, которое хотя и не позволяет в общем случае получить вектор из $P(z^0)$ с минимальной нормой, но дает требуемое подходящее направление (если такие существуют) за конечное число шагов.

Для простоты изложения переобозначим вершины многогранника $P(z^0)$. Пусть z^s, $s = 1, \dots, \mathcal{L}$, – вершины типа $d - y^\ell B$; z^s, $s = \mathcal{L}+1, \dots, R$ – остальные вершины. Предлагаемая процедура состоит в следующем:

Пусть $\theta \in (0, 1)$ и $h^n \in P(z^0)$ – вектор, полученный к n-му шагу процесса (в качестве h^1 можно взять любую из вершин).

I. Для имеющегося h^n нужно проверить выполнение неравенств

$$(h^n, z^s) \geqslant 0, \qquad s = \mathcal{L}+1, \dots, R. \tag{17}$$

Для этого проверяем неравенства

$$(-e^i, h^n) \geqslant 0, \qquad i \in I^0(z^0), \tag{18}$$

и решаем задачу линейного программирования: минимизировать

$(d - yB, h^n)$ при условиях

$$y A^j \leq c_j, \quad j \in J^0(x^0),$$
$$y A^j = c_j, \quad j \overline{\in} J^0(x^0). \tag{19}$$

Если не выполнено одно из условий (18), или же целевая функция задачи (19) не ограничена снизу, то тем самым мы получаем некоторый вектор z^{s_0} такой, что $(h^n, z^{s_0}) < 0$, $s_0 > \mathcal{L}$. В этом случае переходим к п. 3.

Если же неравенства (18) выполнены и задача (19) разрешима, то в качестве z^{s_0} принимаем вектор $d - yB$, где y – оптимальное базисное решение этой задачи, и переходим к п. 2.

2. Проверяем выполнение неравенства

$$(h^n, z^{s_0}) > \theta \|h^n\|^2. \tag{20}$$

Если это неравенство выполняется, то процесс окончен. В противном случае переходим к п. 3.

3. В качестве h^{n+1} принимаем точку отрезка $[h^n, z^{s_0}]$, ближайшую к нулю и возвращаемся к п. I.

ЛЕММА 2. Если $0 \overline{\in} P(x^0)$, то описанная процедура заканчивается через конечное число шагов.

ДОКАЗАТЕЛЬСТВО. Ясно, что в процессе $\|h^n\|$ убывает. Действительно, для $h(\alpha) = h^n + \alpha(z^{s_0} - h^n)$ имеем

$$\|h(\alpha)\|^2 = \|h^n\|^2 + \alpha^2 \|z^{s_0} - h^n\|^2 + 2\alpha(h^n, z^{s_0} - h^n)$$

и

$$\frac{d}{d\alpha}\|h(\alpha)\|^2 = 2\alpha \|z^{s_0} - h^n\|^2 + 2(h^n, z^{s_0} - h^n).$$

Ясно, что при нарушении неравенства (20) эта производная отрицательная при достаточно малых α. Приравнивая производную к нулю, находим

$$\alpha_0 = \frac{(h^n, h^n - z^{s_0})}{\|h^n - z^{s_0}\|^2}.$$

Если оказалось, что $\alpha > 1$, то $h^{n+1} = z^{s_0}$, иначе $h^{n+1} = h(\alpha_0)$.

Ввиду монотонности убывания $|h^n|$ случай $\alpha_0 > 1$ может реализоваться лишь конечное число раз. Поэтому, если процесс продолжается достаточно долго, то окончательно будет $h^{n+1} = h(\alpha_0)$, т.е. h^{n+1} ортогонально к разности $h^n - z^{s_0}$, а значит, и к $h^{n+1} - h^n$. Поэтому $|h^{n+1} - h^n|^2 = |h^n|^2 - |h^{n+1}|^2$. Теперь имеем

$$|h^n|^2 - |h^{n+1}|^2 = \alpha_0^2 \; |h^n - z^{s_0}|^2 = \frac{(h^n, h^n - z^{s_0})^2}{|h^n - z^{s_0}|^2} .$$

Левая часть равенства стремится к нулю, а знаменатель правой части ограничен сверху, значит, числитель правой части стремится к нулю, т.е.

$$|h^n|^2 - (h^n, z^{s_0}) \xrightarrow[n\to\infty]{} 0 . \tag{21}$$

Так как $0 \bar{\in} P(z^0)$, то $|h^n| \underset{n\to\infty}{\not\to} 0$, а поэтому (h^n, z^{s_0}) для достаточно больших значений n будет положительным и, следовательно, будет $s_0 \leq \mathcal{L}$, т.е. будет выполняться п. 2.

Так как $\theta \in (0,1)$, то, ввиду (21), при некотором n непременно выполнится неравенство (20), и процесс прекратится. Лемма доказана.

ЛЕММА 3. Если $0 \in P(z^0)$ и Ω телесно, то $|h^n| \xrightarrow[n\to\infty]{} 0$.

ДОКАЗАТЕЛЬСТВО. Ввиду леммы 1 будем иметь $0 \in S(z^0)$, т.е. найдутся неотрицательные α_s, $s = 1,\dots,R$, такие, что $\sum\limits_{s=1}^{\mathcal{L}} \alpha_s = 1$ и $0 = \sum\limits_{s=1}^{R} \alpha_s z^s$. Покажем, что в этом случае на каждом шаге процесса будет $(h^n, z^{s_0}) \leq 0$. Это так, если на данном шаге процесса нарушается одно из неравенств (17). Пусть все неравенства (17) выполнены. В этом случае из $0 = \sum\limits_{s=1}^{R} \alpha_s (h^n, z^s)$ следует

$$\sum_{s=1}^{\mathcal{L}} \alpha_s (h^n, z^s) = -\sum_{s=\mathcal{L}+1}^{R} \alpha_s (h^n, z^s) \leq 0 .$$

Значит, среди (h^n, z^s), $s = 1,\dots,\mathcal{L}$, есть неположительные и при решении задачи (19) будет получен вектор z^{s_0}, для которого $(h^n, z^{s_0}) = \min\limits_{s=1,\dots,\mathcal{L}} (h^n, z^s) \leq 0$.

Таким образом, либо процесс закончится за конечное число шагов, и при этом из (20) следует $\|h^n\| = 0$, либо процесс бесконечен и из (21) будем иметь

$$\lim_{n\to\infty} \|h^n\|^2 = \lim_{n\to\infty} (h^n, z^{k}) \leq 0,$$

и стало быть $\lim_{n\to\infty} \|h^n\|^2 = 0$, что и требовалось.

Покажем теперь, что если $0 \bar{\in} P(x^o)$, то вектор $-h^n$, на котором прекратится выполнение изложенной процедуры, действительно будет определять подходящее направление. В самом деле, для полученного вектора h^n будут выполняться неравенства (18), т.е.

$$h_i^n \leq 0, \qquad i \in I^o(x^o),$$

и будет разрешима задача (19), причем для вектора $\hat{y}$, решающего эту задачу, выполняются неравенства

$$(d - \hat{y}B, h^n) \geq \theta \|h^n\|^2 > 0.$$

Пусть $\hat{w}$ – решение задачи, двойственной к (19). Для него будет:

$$A\hat{w} + Bh^n = 0,$$

$$w_j \leq 0, \quad j \in J^o(x^o).$$

Следовательно, $x(\lambda) = x^o - \lambda\hat{w}$ и $z(\lambda) = z^o - \lambda h^n$ будут допустимыми в исходной задаче (I). При этом с ростом λ целевая функция будет убывать, ибо

$$(c, \hat{w}) = (-\hat{y}B, h^n)$$

и

$$(c, \hat{w}) + (d, h^n) = (d - \hat{y}B, h^n) > 0.$$

Из сказанного следует, что при достаточно малых $\lambda > 0$ функция $F(z(\lambda))$ определена и

$$F(z(\lambda)) \leq F(z^o) - \lambda(d - \hat{y}B, h^n). \tag{22}$$

§ 4. Описание алгоритма и доказательство сходимости

Опишем теперь вычислительную схему процесса возможных направлений, основанную на изложенной в предыдущем параграфе процедуре получения подходящего направления. Для обеспечения сходимости к оптимальному решению воспользуемся известным приемом введения параметров ε и μ , регулирующих ход процесса.

К началу k -го шага процесса имеется некоторый вектор z^k, для которого задача (3) разрешима. Пусть x^k - базисное решение этой задачи и $\mathcal{D}^{-1}$ - обратная матрица, соответствующая оптимальному базису; y - соответствующий вектор двойственных переменных. Имеются также некоторые значения параметров ε , $\mu > 0$ и $\theta \in (0,1)$.

I. Формируем множества:

$$\mathcal{J}^{\varepsilon}(x^k) = \{ j \mid x_j^k \leqslant \varepsilon \},$$

$$I^{\varepsilon}(z^k) = \{ i \mid z_i^k \leqslant \varepsilon \}.$$

Полагаем $h = d - yB$.

2. Если $h_{i_0} > 0$ для некоторого $i_0 \in I^{\varepsilon}(z^k)$, то принимаем $z = -e^{i_0}$ и переходим к п. 5.

3. Пытаемся решить задачу (19), определяемую множеством $\mathcal{J}^{\varepsilon}(x^k)$ и вектором h . Для этой цели удобно воспользоваться двойственным симплекс-методом, применяя его к соответствующей двойственной задаче:

$$(c, w) \longrightarrow max,$$

$$A w = -Bh, \tag{23}$$

$$w_j \leqslant 0, \quad j \in \mathcal{J}^{\varepsilon}(x^k).$$

При реализации двойственного метода в качестве начальной для каждой задачи (23) можно брать ту базисную матрицу $\mathcal{D}$, которой закончено решение предыдущей задачи (23) или, если эта задача решается впервые для данного x^k - базисную матрицу, которой закончилось решение задачи (3). Базисные компоненты вектора w будут образовывать вектор $-\mathcal{D}^{-1}Bh$.

Если задача (23) имеет решение, то в результате мы получим вектор $\hat{w}$, новую матрицу $\mathcal{D}^{-1}$ и вектор $\hat{y}$, решающий задачу (19).

Проверяем неравенство

$$(d-\hat{y}B,h) \geq \theta\|h\|^2. \qquad (24)$$

Если оно выполняется, то переходим к п. 7; если нет, полагаем $z = d - \hat{y}B$ и переходим к п. 5.

4. Если оказалось, что задача (23) неразрешима, то при ее решении на очередном шаге двойственного метода для некоторого k будет $-(\mathcal{D}^{-1}Bh)_k > 0$ и $(\mathcal{D}^{-1}A^j)_k > 0$ для всех $j = 1,\dots,n_1$. В то же время для соответствующей компоненты вектора w имеется требование неположительности.

В этом случае полагаем $z = (\mathcal{D}^{-1}B)_k$ и переходим к п. 5.

5. В качестве нового значения h принимаем ближайшую к нулю точку отрезка $[h, z]$.

6. Если $\|h\| \leq \mu$, то уменьшаем значения ε и μ с последующим переходом к п. I (либо прекращаем решение задачи, если ε и μ уже достаточно малы). Если $\|h\| > \mu$, возвращаемся к п. 2.

7. Имеющийся вектор $-h$ определяет подходящее направление, и мы полагаем

$$z^{k+1} = z^k - \lambda_k h,$$

$$x^{k+1} = x^k - \lambda_k \hat{w},$$

определяя λ_k из условий неотрицательности компонент векторов z^{k+1} и x^{k+1}. Вектор x^{k+1} при этом может оказаться небазисным. Досчитав, если необходимо, задачу (3) до оптимального базисного решения, возвращаемся к п. I.

Введение параметра ε привело к замене множеств $J^0(x)$ и $I^0(z)$ множествами $J^\varepsilon(x)$ и $I^\varepsilon(z)$. Это в свою очередь приводит к замене множества $S(z)$ множеством $S^\varepsilon(x,z)$ и соответствующего многогранника $P(z)$ многогранником $P^\varepsilon(x,z)$.

ЗАМЕЧАНИЕ. Заменив множество $J^0(x)$ на $J^\varepsilon(x)$, мы уже не можем утверждать, что множество решений системы (6) не будет зависеть от выбора конкретного оптимального решения задачи (3) для данного z. Тем самым множества $S^\varepsilon(x,z)$ при одном

и том же z, но разных x могут оказаться разными.

Доказательство сходимости описанного алгоритма проводится по схеме, аналогичной доказательству метода возможных направлений.

ОПРЕДЕЛЕНИЕ. Точка (x, z) будет называться (ε, μ)-стационарной, если в многограннике $P^{\varepsilon}(x, z)$ имеются точки h такие, что $\|h\| \leq \mu$.

ЛЕММА 4. Если исходная задача (I) разрешима, то при любых $\varepsilon>0$, $\mu>0$ за конечное число шагов процесса будет получена (ε, μ)-стационарная точка.

ДОКАЗАТЕЛЬСТВО. Обозначим векторы $\hat{y}$ и h, фигурирующие при выполнении п. 7 на k-м шаге процесса через $\hat{y}^{k}$ и h^{k}. Ввиду неравенства (22) будет

$$F(z^{k+1}) \leq F(z^{k}) - \lambda_{k}(d - \hat{y}^{k}B, h^{k}).$$

Так как $\varepsilon > 0$, то $\lambda_{k} \not\to 0$ в силу выбора множеств $I^{\varepsilon}(z^{k})$ и $J^{\varepsilon}(x^{k})$. Поэтому, ввиду ограниченности последовательности $F(z^{k})$, будем иметь $(d - \hat{y}^{k}B, h^{k}) \xrightarrow[k\to\infty]{} 0$. Заметим, что п. 7 выполняется лишь тогда, когда справедливо неравенство (24), следовательно, $\|h^{k}\| \xrightarrow[k\to\infty]{} 0$, что и требуется.

ЛЕММА 5. Если $\varepsilon_{k} \to 0$, $\mu_{k} \to 0$, то любая сходящаяся последовательность (ε, μ)-стационарных точек сходится к оптимальному решению исходной задачи.

ДОКАЗАТЕЛЬСТВО. Пусть (x^{k}, z^{k}) - последовательность $(\varepsilon_{k}, \mu_{k})$-стационарных точек, сходящаяся к (x', z'). Покажем, что (x', z') - решение задачи (I).

Каждой точке (x^{k}, z^{k}) соответствует пара множеств $J^{\varepsilon_{k}}(x^{k})$, $I^{\varepsilon_{k}}(z^{k})$. Некоторые из этих пар будут повторяться бесконечное число раз. Пусть такой является пара $\hat{J}$, $\hat{I}$. Соответствующая этой паре подпоследовательность будет сходиться к той же точке (x', z'). Чтобы не усложнять обозначения, будем считать, что уже каждой точке исходной последовательности (x^{k}, z^{k}) соответствует одна и та же пара множеств $\hat{J}$, $\hat{I}$.

Точке (x', z') соответствуют множества $\mathcal{J}' = \mathcal{J}^0(x')$ и $I' = I^0(z')$. Так как $\varepsilon_k \longrightarrow 0$, то из того, что $j \in \hat{\mathcal{J}}$, т.е. что $x_j^k < \varepsilon_k$ при всех k, получаем $x'_j = 0$, иначе говоря, $j \in \mathcal{J}'$, следовательно, $\hat{\mathcal{J}} \subset \mathcal{J}'$. Аналогично $\hat{I} \subset I'$. Пара $\hat{\mathcal{J}}$, $\hat{I}$ определяет некоторый многогранник $\hat{P}$, соответствующий множеству $\hat{S}$ векторов вида $d - yB - \sum_{i \in \hat{I}} t_i e^i$, где $t_i \geqslant 0$, а y удовлетворяет системе

$$y A^j \leqslant c_j, \quad j \in \hat{\mathcal{J}},$$
$$y A^j = c_j, \quad j \bar{\in} \hat{\mathcal{J}}.$$

Аналогично пара $\mathcal{J}'$, I' определяет многогранник $P' = P(z')$. Несложно показать, что из $\hat{\mathcal{J}} \subset \mathcal{J}'$, $\hat{I} \subset I'$ следует $\hat{P} \subset P'$. Многогранник $\hat{P}$ содержит точки, сколь угодно близкие к нулю, ибо $\mu_k \longrightarrow 0$. Ввиду замкнутости многогранника $\hat{P}$ он содержит нулевой вектор, и тем более $0 \in P'$. А тогда, как следует из признака оптимальности, сформулированного в § 2, точка (x', z') оптимальна.

Из лемм 4 и 5 следует справедливость утверждения:

ТЕОРЕМА. *Если множество Q телесно, ранг матрицы A равен m и задача (I) разрешима, то любая предельная точка последовательности, порождаемой описанным алгоритмом, является оптимальным решением этой задачи.*

Литература

1. КОРНАИ И., ЛИПТАК Т. Планирование на двух уровнях. - В кн.: Применение математики в экономических исследованиях, "Мысль", 1965, т.3, с.107-136.
2. ШОР Н.З. Применение обобщенного градиентного спуска в блочном программировании. - "Кибернетика", 1967, т.3, с. 53-55.
3. ПЕРВОЗВАНСКАЯ Т.Н., ПЕРВОЗВАНСКИЙ А.А. Алгоритм поиска оптимального распределения централизованных ресурсов. - Изв. АН СССР, "Техническая кибернетика", 1966, № 3, с.16-19.

4. ДЕМЬЯНОВ В.Ф., МАЛОЗЁМОВ В.Н. Введение в минимакс. М.,"Наука", 1972, с. 368.

5. ШМЫРЕВА Н.В. Использование параметрического программирования для построения алгоритма метода возможных направлений при решении больших задач линейного программирования. Тезисы докладов Четвертого Всесоюзного симпозиума "Системы програмного обеспечения решения задач оптимального планирования", г. Нарва-Йыэссуу М., 1976, с. 128-130.

6. ЕРМОЛЬЕВ Ю.М., ЕРМОЛЬЕВА Л.Г. Метод параметрической декомпозиции. - "Кибернетика", 1973, № 2, с. 66-69.

Поступила в ред.-изд. отд.
20. УП. 1976 г.