Выпуск 37(54) ОПТИМИЗАЦИЯ 1986 г.
Численные методы

УДК 512.25/26

МЕТОД ОТЫСКАНИЯ СЕДЛОВОЙ ТОЧКИ ПРИ ДВОЙНОЙ ДЕКОМПОЗИЦИИ В ЗАДАЧАХ ЛИНЕЙНОГО ПРОГРАММИРОВАНИЯ

Н.В.Шмырева

Предлагаемый в настоящей работе алгоритм связан с двусторонней декомпозицией в задачах линейного программирования. Этот подход состоит в том, что на каждом шаге процесса фиксируется часть переменных (управляющих) как в прямой, так и в двойственной задачах, а относительно оставшихся переменных решается задача линейного программирования меньшего размера. Относительно управляющих переменных возникает задача поиска седловой точки некоторой выпукло-вогнутой функции, кусочно-линейной по каждой из переменных. Для решения этой координирующей задачи предлагается применять два различных варианта градиентного процесса: относительно "далеко" от седловой точки применяется вариант метода Эрроу - Гурвица [1], а в относительно "близкой" окрестности седловой точки - вариант метода Попова [2]. Указываются условия применимости первого из этих методов.

Предлагаемый способ решения задач линейного программирования может оказаться удобным в тех случаях, когда матрица уменьшенной задачи имеет специальную структуру.

§1. Двусторонняя декомпозиция и процедура Эрроу - Гурвица

Рассматривается следующая пара двойственных задач:

$$\begin{cases} c^T x + d^T z \to min! \\ Ax + Bz = b, \\ Gx + Dz = g, \\ x \geq 0,\ z \geq 0. \end{cases} \tag{1.1}$$

$$\begin{cases} b^T y + g^T w \to max! \\ A^T y + G^T w \le c, \\ B^T y + \mathcal{D}^T w \le d. \end{cases} \qquad (I.2)$$

Здесь $c, x \in R^{n_1}$; $d, z \in R^{n_2}$; $y, b \in R^{m_1}$; $w, g \in R^{m_2}$ - вектор-столбцы, а $A, B, G, \mathcal{D}$ - матрицы соответствующих размерностей. Выбирая в качестве управляющих переменные z прямой задачи и переменные w двойственной задачи, сведем решение этой пары задач к нахождению седловой точки функции $\Phi(z, w)$, заданной на множестве $Z \times W$. При этом

$$\Phi(z, w) = d^T z + g^T w - w^T \mathcal{D} z + f(z, w), \qquad (I.3)$$

где

$$f(z, w) = \min_{x \in X(z)} (c - G^T w)^T x = \max_{y \in Y(w)} (b - Bz)^T y,$$

$$X(z) = \{x \,/\, Ax = b - Bz,\ x \ge 0\}; \quad Y(w) = \{y \,/\, A^T y \le c - G^T w\},$$

а множества Z и W определяются условием непустоты множеств $X(z)$ и $Y(w)$. Таким образом, для вычисления функции Φ в точке (z, w) необходимо решить одну из пары двойственных задач:

Задача $U(z, w)$: $(c - G^T w)^T x \to min!$

$$Ax = b - Bz,$$
$$x \ge 0.$$

Задача $W(z, w)$: $(b - Bz)^T y \to max!$

$$A^T y \le c - G^T w.$$

Функция $\Phi(z, w)$ - кусочно-линейная по каждой из переменных, выпукла по z и вогнута по w.

Для отыскания седловой точки выпукло-вогнутых функций известен градиентный процесс Эрроу - Гурвица [I]. В применении к дифференцируемой выпукло-вогнутой функции $\varphi(z, w)$ этот процесс описывается формулами:

$$z_{k+1} = z_k - \lambda_k h_k, \qquad (I.4)$$

$$w_{k+1} = w_k + \lambda_k p_k,$$

где $h_k = \varphi'_z(z_k, w_k)$, $p_k = \varphi'_w(z_k, w_k)$. Доказательство сходимости такого процесса основывается на убывании функции расстояния от текущей точки (z_k, w_k) до седловой точки $(\bar{z}, \bar{w})$:

Здесь $\|\cdot\|$ – евклидова норма в соответствующих пространствах.

Легко видеть, что

$$D(z_{k+1}, w_{k+1}) = \|z^{k+1} - \bar{z}\|^2 + \|w^{k+1} - \bar{w}\|^2 =$$

$$= D(z_k, w_k) - 2\lambda_k ((z_k - \bar{z})^T h_k - (w_k - \bar{w})^T p_k) + \lambda_k^2 (\|h_k\|^2 + \|p_k\|^2).$$

Отсюда ясно, что уменьшение значения функции D при малых $\lambda_k > 0$ зависит от величины $\rho_k = (z_k - \bar{z})^T h_k - (w_k - \bar{w})^T p_k$. Эта величина всегда неотрицательна ввиду выпукло-вогнутости функции φ. Действительно, для любых (z, w) имеем

$$\varphi(\bar{z}, w) \geq \varphi(z, w) + \varphi_z'(z, w)(\bar{z} - z), \tag{I.5}$$

$$\varphi(z, \bar{w}) \leq \varphi(z, w) + \varphi_w'(z, w)(\bar{w} - w). \tag{I.6}$$

Кроме того, так как $(\bar{z}, \bar{w})$ – седловая точка, то

$$\varphi(\bar{z}, w) \leq \varphi(\bar{z}, \bar{w}) \leq \varphi(z, \bar{w}). \tag{I.7}$$

Комбинируя (I.5)-(I.7), получим

$$\varphi_z'(z, w)(z - \bar{z}) - \varphi_w'(z, w)(w - \bar{w}) \geq 0. \tag{I.8}$$

Заметим, что если хоть одно из условий (I.5) или (I.6) выполняется как строгое (что верно, если функция φ строго выпукла по z или строго вогнута по w), то строгим будет и неравенство (I.8), т.е. $\rho_k > 0$. Это и используется для доказательства того, что $D(z_k, w_k) \xrightarrow[k\to\infty]{} 0$ при подходящем выборе последовательности $\{\lambda_k\}$.

Рассмотрим теперь применение этой схемы к введенной выше функции $\Phi(z, w)$. Хотя функция $\Phi(z, w)$ и не является дифференцируемой, однако неравенства (I.5) и (I.6) будут также выполняться, но под $\Phi_z'(z, w)$ следует понимать субградиент функции Φ по z, а под $\Phi_w'(z, w)$ – суперградиент функции Φ по w. А тогда, если рассматривать процесс Эрроу – Гурвица, описываемый формулами (I.4) с $h_k = \Phi_z'(z_k, w_k)$ и $p_k = \Phi_w'(z_k, w_k)$, по-прежнему будем иметь $\rho_k \geq 0$.

Оказывается, что $\rho_k = 0$ лишь тогда, когда текущее решение имеет уже в определенном смысле оптимальную структуру. Для простоты изложения докажем это утверждение, предполагая,

что функция Φ определена при всех z и w, тем самым предполагается, что для переменных z нет требования неотрицательности. При таком предположении множество субградиентов Φ'_z в точке (z,w) будет ограниченным выпуклым многогранником, каждая точка которого имеет вид:

$$h = d - D^T w - B^T y_{zw}, \tag{I.9}$$

где y_{zw} - какое-либо решение задачи $W(z,w)$. Аналогично, множество суперградиентов Φ'_w будет ограниченным выпуклым многогранником, и для любого ρ из этого многогранника верно представление:

$$\rho = g - \mathcal{D} z - G x_{zw}, \tag{I.10}$$

где x_{zw} - решение задачи $V(z,w)$. Если среди векторов h и ρ вида (I.9),(I.10), соответствующих некоторой точке $(\bar z,\bar w)$, нашлись нулевые, это означает, что $(\bar z,\bar w)$ - седловая точка функции Φ.

Для удобства и краткости изложения введем множество $M(z,w)$ всевозможных векторов $\gamma = (h,\rho) \in R^{n_2+m_2}$, где h и ρ получаются в соответствии с (I.9) и (I.10). Тогда признак седловой точки сведется к условию $0 \in M(z,w)$.

Введем следующие обозначения: $J = \{1,2,\dots,n_1\}$, $K = \{j_1, j_2, \dots, j_{m_1}\} \subset J$ - базисное подмножество множества J; $A(K)$ - соответствующая базисная матрица, $c(K) = (c_{j_1}, c_{j_2}, \dots, c_{j_{m_1}})$; $G(K)$ - соответствующая подматрица матрицы G. Перенумеруем все базисные подмножества множества J: $K_1, K_2, \dots, K_s$ и обозначим для краткости через A_ℓ матрицу $A(K_\ell)$, а через $x(A_\ell)$ и $y(A_\ell)$ - решения систем $A_\ell x = b - Bz$, $A_\ell^T y = c(K_\ell) - G(K_\ell)^T w$. Тогда $\mathcal{B}(z,w) = \{\ell \mid x(A_\ell) \geq 0,\ A^T y(A_\ell) \leq c - G^T w\}$ является множеством номеров всех оптимальных базисов пары задач $V(z,w)$ и $W(z,w)$. Для $\ell \in \mathcal{B}(z,w)$ через x^ℓ_{zw} и y^ℓ_{zw} будем обозначать соответствующие базисные оптимальные решения этих задач. Тем самым любое оптимальное решение задачи $V(z,w)$ представимо в виде:

$$x_{zw} = \sum_{\ell \in \mathcal{B}(z,w)} \alpha_\ell x^\ell_{z,w}, \quad \alpha_\ell \geq 0, \quad \sum_{\ell \in \mathcal{B}(z,w)} \alpha_\ell = 1. \tag{I.11}$$

Аналогично, любое оптимальное решение задачи $W(z,w)$ представимо в виде:

$$y_{zw} = \sum_{l \in \mathcal{B}(z,w)} \beta_l y^l_{zw}, \quad \beta_l \geq 0, \quad \sum_{l \in \mathcal{B}(z,w)} \beta_l = 1. \qquad (I.12)$$

В результате множество $M(z,w)$ является многогранником с вершинами F_l вида

$$F_l = (d - \mathcal{D}^T w - B^T y^l_{zw},\ g - \mathcal{D}z - G x^l_{zw}),\ l \in \mathcal{B}(z,w).$$

ОПРЕДЕЛЕНИЕ. Будем говорить, что точки (z_1, w_1) и (z_2, w_2) принадлежат одной зоне, если $\mathcal{B}(z_1, w_1) \cap \mathcal{B}(z_2, w_2) \neq \emptyset$.

ТЕОРЕМА. Пусть $(\bar{z}, \bar{w})$ – седловая точка функции $\Phi(z,w)$ такая, что ей соответствует единственный оптимальный базис K_l в задачах $V(\bar{z},\bar{w})$ и $W(\bar{z},\bar{w})$. Если для некоторой точки (z,w) будем иметь $0 \notin M(z,w)$ и для $(h,p) \in M(z,w)$ оказалось $\rho = (z-\bar{z})^T h - (w-\bar{w})^T p = 0$, то $l \in \mathcal{B}(z,w)$, т.е. точки (z,w) и $(\bar{z},\bar{w})$ принадлежат одной зоне.

ДОКАЗАТЕЛЬСТВО. Любой вектор из многогранника $M(z,w)$ представляет собой выпуклую линейную комбинацию его вершин F_l, $l \in \mathcal{B}(z,w)$. Компоненты же каждого вектора F_l представляют собой субградиент по z и суперградиент по w функции Φ, а тогда и выпуклая линейная комбинация этих векторов в качестве компонент содержит субградиент по z и суперградиент по w функции Φ в точке (z,w). Поэтому из (I.8) следует, что $\rho = (z-\bar{z})^T h - (w-\bar{w})^T p \geq 0$, причем из вывода неравенства (I.8) видно, что если $\rho = 0$, то необходимо

$$\Phi(\bar{z}, w) = \Phi(\bar{z}, \bar{w}) = \Phi(z, \bar{w}). \qquad (I.13)$$

I. Запишем подробно левое из равенств (I.13):

$$d^T\bar{z} + g^T w - w^T \mathcal{D}\bar{z} + (c - G^T w)^T x_{\bar{z}w} - d^T\bar{z} - g^T\bar{w} + \bar{w}^T\mathcal{D}\bar{z}$$

$$-(c - G^T\bar{w})^T x_{\bar{z}\bar{w}} = (w-\bar{w})^T(g - \mathcal{D}\bar{z} - G x_{\bar{z}\bar{w}}) +$$

$$+(c - G^T w)^T x_{\bar{z}w} - (c - G^T w)^T x_{\bar{z}\bar{w}} = 0. \qquad (I.14)$$

Так $(\bar{z},\bar{w})$ – седловая точка, то $0 \in M(\bar{z},\bar{w})$, а в силу единственности оптимального базиса в задачах $V(\bar{z},\bar{w})$ и

$\mathcal{W}(\bar z,\bar w)$ имеем $M(\bar z,\bar w)=\{0\}$. Тем самым $p=g-$ $-\mathcal{D}(\bar z-Gx_{\bar z\bar w})=0$. Тогда из (I.14) получаем

$$(c-G^T w)^T x_{\bar z w}=(c-G^T w)^T x_{\bar z\bar w}. \tag{I.15}$$

Но так как $x_{\bar z\bar w}$ является допустимым решением задачи $\mathcal{V}(\bar z,w)$, то из (I.15) и невырожденности $x_{\bar z\bar w}$ (что следует из единственности оптимального базиса K_ℓ в точке $(\bar z,\bar w)$) получаем, что базис K_ℓ является оптимальным и в задаче $\mathcal{V}(\bar z,w)$.

2. Аналогично, записывая правое равенство из (I.13), получим

$$d^T\bar z+g^T\bar w-\bar w^T\mathcal{D}\bar z+(b-B\bar z)^T y_{\bar z\bar w}-d^T z-g^T\bar w+\bar w^T\mathcal{D}z-$$

$$-(b-Bz)^T y_{z\bar w}=(\bar z-z)^T(d-\mathcal{D}^T\bar w-B^T y_{\bar z\bar w})+$$

$$+(b-Bz)^T y_{\bar z\bar w}-(b-Bz)^T y_{z\bar w}=0. \tag{I.16}$$

Используя отмеченный выше факт, что $M(\bar z,\bar w)=\{0\}$, имеем $h=(d-\mathcal{D}^T\bar w-B^T y_{\bar z\bar w})=0$ и из (I.16) получаем

$$(b-Bz)^T y_{z\bar w}=(b-Bz)^T y_{\bar z\bar w}. \tag{I.17}$$

Как и выше, отсюда получаем, что базис K_ℓ является оптимальным в задаче $\mathcal{W}(z,\bar w)$.

3. Поскольку $\ell\in\mathcal{B}(\bar z,w)$, то базис K_ℓ является допустимым в задаче $\mathcal{W}(z,w)$, а из $\ell\in\mathcal{B}(z,\bar w)$ заключаем, что K_ℓ допустим в задаче $\mathcal{V}(z,w)$. А тогда K_ℓ - оптимальный базис в точке (z,w), т.е. $\ell\in\mathcal{B}(z,w)$.

Используя те же рассуждения, что и при доказательстве сходимости метода Эрроу - Гурвица, можно показать, что в условиях теоремы описанный процесс при надлежащем выборе последовательности λ_k обеспечивает попадание текущей точки (z_k,w_k) в сколь угодно малую окрестность зоны оптимальности базиса K_ℓ. Например, в качестве λ_k можно взять любую последовательность положительных λ_k, лишь бы выполнялось

$$\lambda_k \xrightarrow[k\to\infty]{} 0,\quad \sum_{k=1}^{\infty}\lambda_k=\infty.$$

§2. Модифицированный процесс Эрроу - Гурвица

В этом параграфе рассмотрим модификацию процесса Эрроу - Гурвица, аналогичную рассмотренной для случая дифференцируемых функций Л.Д.Поповым [2]. Идея этой модификации состоит в том, что наряду с последовательностью $\{(z_k, w_k)\}$ строится вспомогательная последовательность точек $\{(\hat{z}_k, \hat{w}_k)\}$, которая и определяет направление сдвига из точек (z_k, w_k) Формулы процесса таковы:

$$\begin{cases} \hat{z}_k = z_k - \lambda_k h_k, \\ \hat{w}_k = w_k + \lambda_k p_k; \end{cases} \tag{2.1}$$

$$\begin{cases} z_{k+1} = z_k - \lambda_k \hat{h}_k, \\ w_{k+1} = w_k + \lambda_k \hat{p}_k, \end{cases} \tag{2.2}$$

где $(h_k, p_k) \in M(z_k, w_k)$; $(\hat{h}_k, \hat{p}_k) \in M(\hat{z}_k, \hat{w}_k)$, но не произвольные, как это было выше, а выбираемые специальным образом, о чем мы подробнее скажем в дальнейшем.

Анализ сходимости процесса основан, как и прежде, на рассмотрении функции $D(z, w)$. Выясним вначале, при каких условиях на шаг λ_k можно гарантировать убывание значения этой функции в течение рассматриваемого процесса.

Используя (2.2), запишем

$$D(z_{k+1}, w_{k+1}) = \| z_{k+1} - \bar{z} \|^2 + \| w_{k+1} - \bar{w} \|^2 =$$

$$= D(z_k, w_k) - 2\lambda_k ((z_k - \bar{z})^T \hat{h}_k - (w_k - \bar{w})^T \hat{p}_k) + \lambda_k^2 (\|\hat{h}_k\|^2 + \|\hat{p}_k\|^2).$$

Для точки $(\hat{z}_k, \hat{w}_k)$ и вектора $(\hat{h}_k, \hat{p}_k)$ верно неравенство (I.8), откуда следует

$$(\hat{z}_k - \bar{z})^T \hat{h}_k - (\hat{w}_k - \bar{w})^T \hat{p}_k = (z_k - \bar{z})^T \hat{h}_k - (w^k - \bar{w})^T \hat{p}_k -$$

$$- \lambda (h_k^T \hat{h}_k + p_k^T \hat{p}_k) \geq 0,$$

и для функции D можно записать

$$D(z_{k+1}, w_{k+1}) \leq D(z_k, w_k) - \lambda_k^2 (2 h_k^T \hat{h}_k + 2 p_k^T \hat{p}_k -$$

$$- \|\hat{h}_k\|^2 - \|\hat{p}_k\|^2. \tag{2.3}$$

Требуется оценить коэффициент при λ_k^2 .

Конкретизируем теперь способ выбора векторов (h_k, p_k) и $(\hat{h}_k, \hat{p}_k)$. В качестве вектора (h_k, p_k) будем выбирать вектор из многогранника $M(z_k, w_k)$, имеющий минимальную евклидову норму.

Обозначим через $\mathcal{B}^0(z_k, w_k)$ множество номеров "существенных" вершин многогранника $M(z_k, w_k)$, участвующих в представлении этого вектора, т.е.

$$(h_k, p_k) = \sum_{\ell \in \mathcal{B}^0(z_k, w_k)} \alpha_\ell (h_k^\ell, p_k^\ell), \sum_{\ell \in \mathcal{B}^0(z_k, w_k)} \alpha_\ell = 1, \alpha_\ell > 0. \quad (2.4)$$

Можно показать, что такой выбор приводит к тому, что при малых λ_k выполняется $\mathcal{B}^0(z_k, w_k) \subset \mathcal{B}(\hat{z}_k, \hat{w}_k)$. Предполагая, что такое требование на выбор шага λ_k соблюдается, в качестве вектора $(\hat{h}_k, \hat{p}_k)$ будем принимать вектор с минимальной нормой из выпуклой оболочки тех вершин многогранника $M(\hat{z}_k, \hat{w}_k)$ номера которых содержатся в множестве $\mathcal{B}^0(z_k, w_k)$, тем самым для вектора $(\hat{h}_k, \hat{p}_k)$ справедливо представление

$$(\hat{h}_k, \hat{p}_k) = \sum_{\ell \in \mathcal{B}^0(z_k, w_k)} \beta_\ell (\hat{h}_k^\ell, \hat{p}_k^\ell), \sum_{\ell \in \mathcal{B}^0(z_k, w_k)} \beta_\ell = 1, \beta_\ell > 0. \quad (2.5)$$

Прежде чем переходить к выводу необходимых оценок, установим связь между векторами (h_k^ℓ, p_k^ℓ) и $(\hat{h}_k^\ell, \hat{p}_k^\ell)$.

Пусть $\mathcal{A}_\ell$ - матрица, соответствующая базисному множеству K_ℓ, $\bar{G} = G(K_\ell)$, $\bar{c} = c(K_\ell)$. Тогда решения задач $\mathcal{V}(z_k, w_k)$ и $\mathcal{W}(z_k, w_k)$ имеют вид:

$$x^\ell_{z_k w_k} = \mathcal{A}_\ell^{-1}(b - Bz_k); \quad y^\ell_{z_k w_k} = (\bar{c} - \bar{G}^T w_k)^T \mathcal{A}_\ell^{-1}, \quad (2.6)$$

и векторы (h_k^ℓ, p_k^ℓ) определяются формулами:

$$h_k^\ell = d - \mathcal{D}^T w_k - B^T y^\ell_{z_k w_k}; \quad p_k^\ell = g - \mathcal{D} z_k - \bar{G} x^\ell_{z_k w_k}. \quad (2.7)$$

Векторы $(\hat{h}_k^\ell, \hat{p}_k^\ell)$ определяются подобными формулами, и, учитывая (2.I) и (2.6), можно записать

$$\hat{h}_k^\ell = d - \mathcal{D}^T \hat{w}_k - B^T \hat{y}^\ell_{\hat{z}_k \hat{w}_k} = h_k^\ell - \lambda_k (\mathcal{D} - \bar{G}\mathcal{A}_\ell^{-1} B)^T p_k,$$

$$\hat{p}_k^\ell = g - \mathcal{D}\hat{z}_k - \bar{G}\hat{x}^\ell_{\hat{z}_k \hat{w}_k} = p_k^\ell + \lambda_k (\mathcal{D} - \bar{G}\mathcal{A}_\ell^{-1} B) h_k,$$

или, обозначая через H^ℓ матрицу $\mathcal{D} - \bar{G}\mathcal{A}_\ell^{-1} B$, имеем

$$\hat{h}_k^\ell = h_k^\ell - \lambda_k p_k^T H^\ell,$$
$$\hat{p}_k^\ell = p_k^\ell + \lambda_k H^\ell h_k. \qquad (2.8)$$

Тогда, используя представление (2.5), для вектора $(\hat{h}_k, \hat{p}_k)$ получаем

$$\hat{h}_k = \tilde{h}_k - \lambda_k \sum_\ell \beta_\ell p_k^T H^\ell,$$
$$\hat{p}_k = \tilde{p}_k + \lambda_k \sum_\ell \beta_\ell H^\ell h_k, \qquad (2.9)$$

где $(\tilde{h}_k, \tilde{p}_k) = \sum_{\ell \in \mathcal{B}^\circ(z_k, w_k)} \beta_\ell (h_k^\ell, p_k^\ell)$ – некоторый вектор из многогранника $M(z_k, w_k)$. Так как вектор $(h_k, p_k) \in M(z_k, w_k)$ и имеет минимальную норму, то

$$\|\tilde{h}_k\|^2 + \|\tilde{p}_k\|^2 \geqslant \|h_k\|^2 + \|p_k\|^2. \qquad (2.10)$$

Теперь, используя формулы (2.9), неравенство (2.10), а также неравенства

$$(x_1^T Q y_1 - x_2^T Q y_2) \leqslant \|Q\|(\|x_1\| \cdot \|y_1\| + \|x_2\| \cdot \|y_2\|), \qquad (2.11)$$

$$2ab \leqslant a^2 + b^2, \qquad (2.12)$$

получим вначале оценку для величины $\|\hat{h}_k\|^2 + \|\hat{p}_k\|^2$:

$$\|\hat{h}_k\|^2 + \|\hat{p}_k\|^2 = \|\tilde{h}_k\|^2 + \|\tilde{p}_k\|^2 - 2\lambda_k \sum_\ell \beta_\ell p_k^T H^\ell \tilde{h}_k +$$
$$+ 2\lambda_k \sum_\ell \beta_\ell \tilde{p}_k^T H^\ell h_k + \lambda_k^2 \|\sum_\ell \beta_\ell p_k^T H^\ell\|^2 + \lambda_k^2 \|\sum_\ell \beta_\ell H^\ell h_k\|^2 \geqslant$$
$$\geqslant \|\tilde{h}_k\|^2 + \|\tilde{p}_k\|^2 - 2\lambda_k \sum_\ell \beta_\ell \|H^\ell\|(\|p_k\| \cdot \|\tilde{h}_k\| + \|\tilde{p}_k\|^2 \cdot \|h_k\|^2) \geqslant$$
$$\geqslant \|\tilde{h}_k\|^2 + \|\tilde{p}_k\|^2 - \lambda_k \|H\|(\|p_k\|^2 + \|\tilde{h}_k\|^2 + \|\tilde{p}_k\|^2 + \|h_k\|^2) \geqslant$$
$$\geqslant (1 - 2\lambda_k \|H\|)(\|h_k\|^2 + \|p_k\|^2). \qquad (2.13)$$

Здесь $\|H\|$ – максимальная из норм $\|H^\ell\|$ по всем $\ell = 1, 2, \ldots, s$.

Пусть теперь $(\bar{h}_k, \bar{p}_k) = \sum_\ell \alpha_\ell (\hat{h}_k, \hat{p}_k^\ell)$, где суммирование идет по $\ell \in \mathcal{B}^\circ(z_k, w_k)$, а коэффициенты α_ℓ взяты такими же, как и в представлении (2.4) вектора (h_k, p_k) . Поскольку вектор $(\bar{h}_k, \bar{p}_k)$ принадлежит многограннику $M(\hat{z}_k, \hat{w}_k)$, но мо-

жет не совпадать с минимальным вектором $(\hat{h}_k, \hat{\rho}_k)$, верно неравенство

$$\hat{h}_k^T \bar{h}_k + \hat{\rho}_k^T \bar{\rho}_k = \sum_\ell \alpha_\ell \hat{h}_k^T \hat{h}_k^\ell + \sum_\ell \alpha_\ell \hat{\rho}_k^T \hat{\rho}_k^\ell \geq |\hat{h}_k|^2 + |\hat{\rho}_k|^2. \quad (2.14)$$

Учитывая (2.8) и неравенства (2.11),(2.12),(2.14), а также представление (2.4), получаем следующую оценку:

$$2 h_k^T \hat{h}_k + 2\rho_k^T \hat{\rho}_k = 2\sum_\ell \alpha_\ell \hat{h}_k^T h_k^\ell + 2 \sum_\ell \hat{\rho}_k^T \rho_k^\ell =$$

$$= 2\sum_\ell \alpha_\ell \hat{h}_k^T \hat{h}_k^\ell + 2\sum_\ell \alpha_\ell \hat{\rho}_k^T \hat{\rho}_k^\ell + 2\lambda_k \sum_\ell \alpha_\ell \rho_k^T H^\ell \hat{h}_k -$$

$$-2\lambda_k \sum_\ell \alpha_\ell \hat{\rho}_k H^\ell h_k \geq 2(\|\hat{h}_k\|^2 + \|\hat{\rho}_k\|^2) - 2\lambda_k \sum \alpha_\ell (\hat{\rho}_k^T H^\ell h_k - \rho_k^T H^\ell \hat{h}_k) \geq$$

$$\geq (2 - \lambda_k \|H\|)(\|\hat{h}_k\|^2 + \|\hat{\rho}_k\|^2) - \lambda_k \|H\|(\|h_k\|^2 + \|\rho_k\|^2). \quad (2.15)$$

Объединяя (2.13) и (2.15), получим оценку для коэффициента при λ_k^2 в (2.3):

$$2h_k^T \hat{h}_k + 2\rho_k^T \hat{\rho}_k - \|\hat{h}_k\|^2 - \|\rho_k\|^2 \geq$$

$$\geq (1 - 4\lambda_k \|H\| + \lambda_k^2 \|H\|^2)(\|h_k\|^2 + \|\rho_k\|^2). \quad (2.16)$$

Отсюда видно, что при $0 < \lambda_k \leq \frac{1}{4} \cdot \frac{1}{\|H\|}$ правая часть последнего неравенства будет оцениваться снизу величиной $\frac{1}{16}(\|h_k\|^2 + \|\rho_k\|^2)$.

Возвращаясь к неравенству (2.3), в результате получаем

$$D(z_{k+1}, w_{k+1}) \leq D(z_k, w_k) - \frac{1}{16}\lambda_k^2 (\|h_k\|^2 + \|\rho_k\|^2). \quad (2.17)$$

Таким образом, при достаточно малой величине шага λ_k значение функции D в процессе будет строго убывать. Однако напомним, что длина шага λ_k регламентируется также условием $\mathcal{B}^\circ(z_k, w_k) \subset \mathcal{B}(\hat{z}_k, \hat{w}_k)$. Поэтому для того, чтобы доказать сходимость процесса, следует воспользоваться известным ε-приемом, предотвращающим "заедание" в процессах методов возможных направлений. Применительно к рассматриваемому процессу это сведется к следующим дополнительным правилам: вместо множеств $\mathcal{B}(z, w)$ следует вводить более широкие множества $\mathcal{B}_\varepsilon(z; w)$, в которые включаются номера не только оптимальных, но и "поч-

ти оптимальных" базисов, т.е. таких, для которых условия допустимости в прямой и двойственной задачах выполнены с точностью до $\mathcal{E}$.

При таком дополнении несложно показать, что процесс будет осуществим с малым постоянным ненулевым шагом $\lambda_k = \lambda(\mathcal{E}) < 0$, обеспечивающим строгое убывание функции $\mathcal{D}$. Ввиду доказанной оценки (2.17) получаем $\lim_{k\to\infty} (\|h_k\|^2 + \|\rho_k\|^2) = 0$.

Тем самым за конечное число шагов будет получена $\mathcal{E}$-стационарная точка, т.е. такая точка $(\tilde{z}, \tilde{w})$, для которой в многограннике $M_{\mathcal{E}}(\tilde{z}, \tilde{w})$, определяемом множеством $\mathcal{B}_{\mathcal{E}}(\tilde{z}, \tilde{w})$, найдутся точки (h, ρ) такие, что $\|h\|^2 + \|\rho\|^2 < \mathcal{E}$. Несложно показать, что при $\mathcal{E} \to 0$ последовательность $\mathcal{E}$-стационарных точек сходится к седловой точке функции Φ.

Относительно последовательности λ_k заметим, что требуемым условием, обеспечивающим сходимость процесса, удовлетворяет любая последовательность, для которой выполняется $\lambda_k \xrightarrow[k\to\infty]{} 0$, $\sum_{k=1}^{\infty} \lambda_k^2 = +\infty$.

§3. Заключительные замечания

Рассмотренная модификация процесса Эрроу - Гурвица значительно более трудоемка, чем его исходный вариант, рассмотренный в §I. Поэтому при практической реализации целесообразно использовать комбинированный вариант, когда вначале применяется процедура, описанная в §I, а после того, как возникнет повторяемость множеств $\mathcal{B}(z_k, w_k)$, осуществляется переход к модифицированному процессу.

Реализуемость предложенной модификации процесса Эрроу - Гурвица в большой степени зависит от того, насколько эффективно мы можем решать задачу об отыскании вектора с минимальной нормой в многограннике $M(z, w)$. Как и в случае односторонней декомпозиции [3], здесь оказывается возможным применение метода условного градиента. При этом мы избавляемся от необходимости иметь в явном виде все вершины многогранника $M(z, w)$ одновременно. Очередная вершина получается при выполнении шага условного градиента путем некоторых преобразований симплексного типа для задачи с матрицей A.

Следует отметить, что переход от точного решения задачи нахождения вектора с минимальной нормой к итеративной процедуре метода условного градиента вызывает дополнительные вопросы согласования точности получаемого приближенно вектора (h,ρ) с $\mathcal{E}$ - процедурой модифицированного процесса Эрроу - Гурвица.

В заключение заметим, что хотя основные рассмотрения мы провели в предположении, что задачи $\mathcal{V}(z,w)$ и $\mathcal{W}(z,w)$ разрешимы при любых z и w, однако это ограничение можно снять. При этом множество субградиентов функции $\mathcal{P}(z,w)$ будет представлять собой многогранное множество, уже не всегда ограниченное. Однако, как и в [3], при определении направления сдвига (h,ρ) его можно заменить ограниченным многогранником, что позволяет получить подходящее направление за конечное число шагов процедуры условного градиента. Заметим также, что если функция $\mathcal{P}$ определена не при всех (z,w), то и в основной процедуре Эрроу - Гурвица приходится в качестве направления (h,ρ) брать не произвольный элемент из $M(z,w)$, а также переходить к процедуре условного градиента для отыскания направления, не выводящего из области задания функции $\mathcal{P}$.

ЛИТЕРАТУРА

1. ЭРРОУ К.Дж., ГУРВИЦ Л., УДЗАВА Х. Исследования по линейному и нелинейному программированию. - М.: ИЛ, 1962.
2. ПОПОВ Л.Д. Модификация метода Эрроу - Гурвица поиска седловых точек. - Мат. заметки, 1980, т.28, № 5, с.777-784.
3. ШМЫРЕВА Н.В. Об одном декомпозиционном алгоритме линейного программирования. - Оптимизация, 1977, вып. 19(36), с. 100-115.

Поступила в ред.-изд. отдел
27.11.1985 г.