EESTI NSV TEADUSTE AKADEEMIA TOIMETISED. XVIII KÖIDE
FÜÜSIKA * MATEMAATIKA. 1969, NR. 1

ИЗВЕСТИЯ АКАДЕМИИ НАУК ЭСТОНСКОЙ ССР. ТОМ XVIII
ФИЗИКА * МАТЕМАТИКА. 1969, № 1

https://doi.org/10.3176/phys.math.1969.1.01

С. УЛЬМ

ОБ ОПТИМИЗАЦИИ НА ДВУХ УРОВНЯХ

В последнее время значительно возрос интерес к разработке методов для решения нелинейных многомерных задач оптимизации. Одним из таких подходов является оптимизация на двух уровнях, когда исходная сложная задача оптимизации разбивается на ряд более мелких подзадач, зависящих от некоторых параметров (декомпозиция). Нахождение оптимальных параметров (часто множителей Лагранжа) происходит на втором уровне оптимизации (координация). При этом для нахождения оптимального решения между первым и вторым уровнями происходит некоторый итерационный процесс с помощью обмена информацией (см., напр., [1–4]).

Здесь разрабатывается схема оптимизации на двух уровнях для одного класса задач нелинейного программирования, рассмотренного раньше в [5–6]. Если в последних работах для решения упомянутой задачи был обобщен метод разложения Данцига—Вульфа [7], то в данной работе процесс координации проводится более простым образом — посредством метода градиентов. Задача координации при этом состоит в нахождении безусловного минимума некоторой функции. В статье дается обоснование применения метода градиентов для рассматриваемого случая. Для ускорения процесса координации используются также и так наз. интерполяционные методы. Рассматриваются возможности применения полученных результатов к решению задач дискретного оптимального управления. Дается и итеративный алгоритм для решения на двух уровнях некоторого другого класса задач нелинейного программирования, рассмотренного в [8].

1. Рассмотрим задачу нелинейного программирования: максимизировать функцию

$$f(x) \tag{1}$$

при условиях

$$Ax = b, \tag{2}$$

$$x \in R, \tag{3}$$

где $x = (x_1, \ldots, x_n)$; $A = (a_{ij})_{mn}$; $b = (b_1, \ldots, b_m)$. Предполагается, что

а) R — выпуклое замкнутое ограниченное множество;

б) $f(x)$ — вогнутая непрерывная функция на множестве R;

в) рангом матрицы A является m.

Обозначим $p = (p_1, \ldots, p_n)$ и рассмотрим функцию

$$\tilde{f}(p) = \max_{x \in R} [f(x) - (x, p)]. \tag{4}$$

Через $\tilde{R}$ обозначим множество точек, для которых $\tilde{f}(p) < \infty$. Функция $\tilde{f}(p)$, определенная на множестве $\tilde{R}$, называется сопряженной к $f(x)$ функцией [9]. Можно показать, что $\tilde{R}$ является непустым выпуклым множеством и $\tilde{f}(p)$ — выпуклой функцией [9].

Двойственную для (1)—(3) задачу можно сформулировать с помощью сопряженной функции $\tilde{f}(p)$ [10]. Для этого образуем функцию Лагранжа

$$F(x, \lambda) = f(x) - (\lambda, Ax) + (b, \lambda),$$

где $\lambda = (\lambda_1, \ldots, \lambda_m)$ является вектором множителей Лагранжа. Поскольку

$$\psi(\lambda) = \max_{x \in R} F(x, \lambda) = \max_{x \in R} [f(x) - (x, A^* \lambda)] +$$

$$+ (b, \lambda) = \tilde{f}(A^* \lambda) + (b, \lambda),$$

то двойственная для (1)—(3) задача состоит в минимизации функции

$$\psi(\lambda) = \tilde{f}(A^* \lambda) + (b, \lambda) \tag{5}$$

в пространстве векторов λ [10].

Из [10] также следует, что при сделанных предположениях справедливо соотношение двойственности

$$\min_{\lambda} \psi(\lambda) = \max_{\substack{Ax = b \\ x \in R}} f(x). \tag{6}$$

2. Обозначим максимизирующую функцию в выражении (4) через $x(p)$, т. е.

$$\tilde{f}(p) = f(x(p)) - (x(p), p).$$

Дифференцируя (5), получим

$$\operatorname{grad} \psi(\lambda) = A \operatorname{grad} \tilde{f}(A^* \lambda) + b.$$

Поскольку [9]

$$\operatorname{grad} \tilde{f}(p) = -x(p),$$

то окончательно

$$\operatorname{grad} \psi(\lambda) = -A\, x(A^* \lambda) + b. \tag{7}$$

Отсюда ясно, что если $x(p)$ — непрерывная функция, то $\psi(\lambda)$ — непрерывно дифференцируемая функция.

Допустим теперь дополнительно, что функция $f(x)$ на множестве R является строго вогнутой. Обозначим $\Phi(x, p) = f(x) - (x, p)$. Поскольку $f(x)$ — строго вогнутая, то при каждом фиксированном p функция $\Phi(x, p)$ также строго вогнутая. Из соотношения $\tilde{f}(p) = \max_{x \in R} \Phi(x, p)$ следует, что при каждом p функция $x(p)$ является единственной.

Покажем непрерывность $x(p)$. Для этого рассмотрим произвольную последовательность $\{p^{(k)}\}$, имеющую пределом $\bar{p}$. Допустим, что

$x(p^{(k)}) \not\to x(\bar{p})$. Поскольку из каждой ограниченной последовательности можно выделить сходящуюся подпоследовательность, то существует подпоследовательность $\{p^{(k_l)}\}$ такая, что $x(p^{(k_l)}) \to \tilde{x} \neq x(\bar{p})$ $(\tilde{x} \in R)$. Исходя из этого, допустим без ограничения общности, что первоначальная последовательность была выбрана так, что $p^{(k)} \to \bar{p}$ и $x(p^{(k)}) \to \to \tilde{x} \neq x(\bar{p})$. Так как $\Phi(x(p^{(k)}), p^{(k)}) \geqslant \Phi(x, p^{(k)})$ для каждого $x \in R$ и $\Phi(x, p)$ является непрерывной, то

$$\Phi(\tilde{x}, \bar{p}) \geqslant \Phi(x(\bar{p}), \bar{p}).$$

Отсюда по определению $x(\bar{p})$ следует, что $\tilde{x} = x(\bar{p})$. Полученное противоречие доказывает непрерывность функции $x(p)$. Итак, при сделанных допущениях $\psi(\lambda)$ имеет непрерывный градиент.

3. Рассмотрим теперь задачу (ср. [5, 6]): максимизировать функцию

$$f(x) = \sum_{i=1}^{n} f_i(x_i) \tag{8}$$

при условиях

$$\sum_{k=1}^{n} A_k x_k = b, \tag{9}$$

$$x_i \in R_i \quad (i = 1, \dots, n). \tag{10}$$

Здесь $x = (x_1, \dots, x_n)$; $x_i = (x_{i1}, \dots, x_{in_i})$, причем при $i \neq j$ в общем $n_i \neq n_j$; $A_k = (a_{ij}^k)_{mn_k}$; функции $f_i(x_i)$ являются непрерывными строго вогнутыми соответственно на выпуклых ограниченных замкнутых множествах R_i; ранг матрицы $A = (A_1, \dots, A_n)$ равен m; $b = (b_1, \dots, b_m)$.

Обозначим $p = (p_1, \dots, p_n)$, где p_i являются n_i-мерными векторами, т. е.

$$p_i = (p_{i1}, \dots, p_{in_i}).$$

В данном случае сопряженная к $f(x)$ функция выражается в виде

$$\tilde{f}(p) = \max_{x \in R} [f(x) - (x, p)] =$$

$$= \max_{x \in R} [\sum_{i=1}^{n} f_i(x_i) - \sum_{i=1}^{n} (x_i, p_i)] =$$

$$= \sum_{i=1}^{n} \{\max_{x_i \in R_i} [f_i(x_i) - (x_i, p_i)]\} =$$

$$= \sum_{i=1}^{n} \tilde{f}_i(p_i),$$

где $\tilde{f}_i(p_i) = \max_{x_i \in R_i} [f_i(x_i) - (x_i, p_i)]$ является сопряженной к функции $f_i(x_i)$.

Обозначим вектор множителей Лагранжа через $\lambda = (\lambda_1, \ldots, \lambda_m)$. Поскольку

$$A^* = \left\| \begin{matrix} A_1^* \\ \vdots \\ A_n^* \end{matrix} \right\|,$$

то

$$A^*\lambda = (A_1^*\lambda, \ldots, A_n^*\lambda).$$

Итак,

$$\psi(\lambda) = \max_{x \in R} [f(x) - (x, A^*\lambda)] + (b, \lambda) =$$

$$= \max_{x \in R} [\sum_{i=1}^{n} f_i(x_i) - \sum_{i=1}^{n} (x_i, A_i^*\lambda)] + (b, \lambda) =$$

$$= \sum_{i=1}^{n} \{\max_{x_i \in R_i} [f_i(x_i) - (x_i, A_i^*\lambda)]\} + (b, \lambda) =$$

$$= \sum_{i=1}^{n} \tilde{f}_i(A_i^*\lambda) + (b, \lambda).$$

Двойственная для (8)—(10) задача состоит в минимизации функции

$$\psi(\lambda) = \sum_{i=1}^{n} \tilde{f}_i(A_i^*\lambda) + (b, \lambda) \tag{11}$$

в пространстве векторов λ.

Так как максимизирующие $x_i(p_i)$ определяются из выражений $\max\limits_{x_i \in R_i} [f_i(x_i) - (x_i, p_i)]$ единственно и непрерывно, то существует непрерывный градиент функции $\psi(\lambda)$, причем (ср. (7))

$$\operatorname{grad} \psi(\lambda) = -\sum_{i=1}^{n} A_i x_i(A_i^*\lambda) + b. \tag{12}$$

4. Даем схему решения задачи (8)—(10) на двух уровнях.

А. Пусть выбрано начальное приближение

$$\lambda^{(0)} = (\lambda_1^{(0)}, \ldots, \lambda_m^{(0)}).$$

Б. Если построено приближение $\lambda^{(l)}$, то $\lambda^{(l+1)}$ вычисляется следующим образом:

а) подсистема i $(i = 1, \ldots, n)$ вычисляет максимизирующую $x_i(A_i^*\lambda^{(l)})$ из задачи

$$\max_{x_i \in R_i} [f_i(x_i) - (x_i, A_i^*\lambda^{(l)})]; \tag{13}$$

б) второй уровень (центр), задачей которой является минимизация функции (11), вычисляет на основании найденных $x_i(A_i^*\lambda^{(l)})$ новое приближение $\lambda^{(l+1)}$ по методу градиентов, т. е.

$$\lambda^{(l+1)} = \lambda^{(l)} + \varepsilon_l [\sum_{i=1}^{n} A_i x_i(A_i^*\lambda^{(l)}) - b], \tag{14}$$

где $\varepsilon_l > 0$ $(l = 0, 1, \ldots)$.

Если $\varepsilon_l \to 0$ и $\sum_{l=0}^{\infty} \varepsilon_l = \infty$, то на основании выпуклости и непрерывной дифференцируемости функции $\psi(\lambda)$ последовательность $\{\lambda^{(l)}\}$ сходится к точке минимума $\bar{\lambda}$ функции $\psi(\lambda)$* (см. [11]). По формуле (6) вектор

$$\bar{x} = (x_1(A_1^*\bar{\lambda}), \ldots, x_n(A_n^*\bar{\lambda}))$$

решает задачу (8)—(10). Действительно,

$$\max_{\substack{\sum_{k=1}^{n} A_k x_k = b \\ x_i \in R_i}} \sum_{i=1}^{n} f_i(x_i) = \min_{\lambda} \psi(\lambda) =$$

$$= \sum_{i=1}^{n} \tilde{f}_i(A_i^*\bar{\lambda}) + (b, \bar{\lambda}) =$$

$$= \sum_{i=1}^{n} [f_i(x_i(A_i^*\bar{\lambda})) - (x_i(A_i^*\bar{\lambda}), A_i^*\bar{\lambda})] + (b, \bar{\lambda}) =$$

$$= \sum_{i=1}^{n} f_i(x_i(A_i^*\bar{\lambda})) - \sum_{i=1}^{m} (A_i x_i(A_i^*\bar{\lambda}), \bar{\lambda}) + (b, \bar{\lambda}) =$$

$$= \sum_{i=1}^{n} f_i(x_i(A_i^*\bar{\lambda})),$$

поскольку $\operatorname{grad}\psi(\bar{\lambda}) = 0$.

Приведем пример, заимствованный из [6]: максимизировать функцию

$$f(x) = \sum_{i=1}^{3} f_i(x_i),$$

где

$$\begin{aligned}
f_1(x_1) &= -(x_{11}-1)^2 - (x_{12}-1)^2, \\
f_2(x_2) &= -(x_{21}-1)^2 - (x_{22}-1)^2 - (x_{23}-1)^2, \\
f_3(x_3) &= -(x_{31}-1)^2 - (x_{32}-1)^2
\end{aligned}$$

при условиях

$$\begin{aligned}
x_{11} + 2x_{12} + 4x_{21} + 2x_{22} + 4x_{23} + 2x_{31} + x_{32} &= 5, \\
-x_{11} - 2x_{12} + x_{21} + 3x_{22} + x_{23} &= 1, \\
x_{21} + 3x_{22} + x_{23} - x_{31} - 2x_{32} &= 1,
\end{aligned}$$

$$0 \leqslant x_{ij} \leqslant 1.$$

Выбирая $\lambda^{(0)} = (11; 1; 1)$ и $\varepsilon_l = \frac{1}{l+1}$, после 21-й итерации по методу (13) получаем следующие результаты:

$\lambda = (0{,}525011;\ -0{,}071654;\ -0{,}149830)$

$x_{11} = 0{,}701666$	$x_{31} = 0{,}400073$
$x_{12} = 0{,}403333$	$x_{32} = 0{,}587663$
$x_{21} = 0{,}060719$	$f(x) = -2{,}77661$
$x_{22} = 0{,}807215$	$\psi(\lambda) = -2{,}77703.$
$x_{23} = 0{,}060719$	

* Если $\psi(\lambda^{(l+1)}) \geqslant \psi(\lambda^{(l)})$, то выбираем $\lambda^{(l+1)} = \lambda^{(l)}$ (ср. [11]).

5. Рассмотрим применение полученных результатов для решения задачи дискретного оптимального управления.

Пусть требуется максимизировать функцию

$$f(x, u) = \sum_{i=0}^{N-1} f_i(x^i, u^i) + f_N(x^N) \tag{15}$$

при условиях

$$x^{i+1} = C_i x^i + D_i u^i + b^i \quad (i = 0, \ldots, N-1), \tag{16}$$

$$x^0 \text{ — задана}, \tag{17}$$

$$u^i \in U_i \quad (i = 0, \ldots, N-1), \tag{18}$$

$$x^i \in X_i \quad (i = 1, \ldots, N). \tag{19}$$

Здесь $x^i = (x_1^i, \ldots, x_n^i)$ $(i = 0, \ldots, N)$ — n-мерные векторы фазовых координат; $u^i = (u_1^i, \ldots, u_m^i)$ $(i = 0, \ldots, N-1)$ — m-мерные векторы управлений; $C_i = (c_{kl}^i)_{nn}$; $D_i = (d_{kl}^i)_{nm}$; $b^i = (b_1^i, \ldots, b_n^i)$ $(i = 0, \ldots, N-1)$.

Обозначим $z = (z^0, \ldots, z^N)$, где $z^i = (x^i, u^i)$ $(i = 0, \ldots, N-1)$; $z^N = (x^N, 0)$.

Проблему (15)—(19) можно рассматривать в виде (8)—(10): максимизировать функцию

$$f(z) = \sum_{i=0}^{N} f_i(z^i)$$

при условиях

$$\sum_{k=0}^{N} A_k z^k = b,$$

$$z^i \in R_i \ (i = 0, \ldots, N).$$

В данном случае

$$f_i(z^i) = f_i(x^i, u^i) \quad (i = 0, \ldots, N-1);$$
$$f_N(z^N) = f_N(x^N);$$

$$A_k = \begin{array}{c} \left\| \begin{array}{cc} 0 & 0 \\ \cdot & \cdot \\ 0 & 0 \\ E & 0 \\ -C_k & -D_k \\ 0 & 0 \\ \cdot & \cdot \\ 0 & 0 \end{array} \right\| \end{array} \begin{array}{l} \ldots 1 \\ \\ \\ \ldots k \\ \ldots k+1 \\ \\ \\ \ldots N \end{array} \qquad (k = 0, \ldots, N).$$

Здесь E — n-мерная единичная матрица;

$b = (b^0, \ldots, b^{N-1})$; $R_0 = \{x_0, u_0) / u^0 \in U_0\}$;

$R_i = \{(x^i, u^i) / x^i \in X_i;\ u_i \in U_i\}$ $(i = 1, \ldots, N-1)$;

$R_N = \{(x^N, 0) / x^N \in X_N\}$.

Допустим, что $f_i(z^i)$ — непрерывные строго вогнутые функции на множествах R_i и X_i, U_i — выпуклые замкнутые ограниченные множества.

Обозначим вектор множителей Лагранжа через $\lambda = (\lambda^1, \ldots, \lambda^N)$ (λ^i — n-мерные векторы) и l-е его приближение через

$$\lambda^{(l)} = (\lambda^{1(l)}, \ldots, \lambda^{N(l)}).$$

Поскольку

$$A_i^* \lambda^{(l)} = (\lambda^{i(l)} - C_i^* \lambda^{i+1(l)}; \; -D_i^* \lambda^{i+1(l)}),$$

то подзадачи, соответствующие (13), приобретают вид

$$\max_{u^0 \in U_0} [f_i(x^0, u^0) + (u^0, D_0^* \lambda^{1(l)})];$$

$$\max_{\substack{x^i \in X_i \\ u^i \in U_i}} [f_i(x^i, u^i) - (x^i, \lambda^{i(l)} - C_i^* \lambda^{i+1(l)}) + (u^i, D_i^* \lambda^{i+1(l)})] \quad (i = 1, \ldots, N-1);$$

$$\max_{x^N \in X_N} [f_N(x^N) - (x^N, \lambda^{N(l)})].$$

Соответствующие решения подзадач обозначим через $z^i(A_i^* \lambda^{(l)})$ $(i = 0, \ldots, N)$.

Градиентный метод для координации выражается в виде (ср. (14))

$$\lambda^{(l+1)} = \lambda^{(l)} + \varepsilon_l [\sum_{i=0}^{N} A_i z^i (A_i^* \lambda^{(l)}) - b]$$

или более подробно

$$\lambda^{i(l+1)} = \lambda^{i(l)} + \varepsilon_l [x^i(A_i^* \lambda^{(l)}) - C_{i-1} x^{i-1}(A_{i-1}^* \lambda^{(l)}) -$$
$$- D_{i-1} u^{i-1}(A_{i-1}^* \lambda^{(l)}) - b_{i-1}]$$
$$(i = 1, \ldots, N; \; l = 0, 1, \ldots).$$

6. Очевидно, что при решении задачи (8)—(10) на двух уровнях для координации можно решить систему нелинейных уравнений

$$P(\lambda) = \sum_{i=1}^{n} A_i x_i (A_i^* \lambda) - b = 0. \tag{20}$$

Поскольку на практике мы часто не в состоянии найти аналитические выражения для $x_i(A_i^* \lambda)$, но зато можем вычислить их значения при фиксированных λ (задачи подсистем), то в качестве координирующих процессов часто удобно применять интерполяционные методы решения систем нелинейных уравнений (методы хорд [12], Стеффенсена [13] и т. д.), использующие только значения функций x_i.

Для простоты рассмотрим случай, когда x_k в формуле (14) одномерны.

Метод хорд для решения системы (20) выражается в виде

$$P(\lambda^{(l)}\lambda^{(l-1)})(\lambda^{(l+1)}-\lambda^{(l)})=-P(\lambda^{(l)}) \quad (l=1,2,\ldots), \tag{21}$$

где $\lambda^{(0)}$, $\lambda^{(1)}$ — начальные приближения, $P(\lambda^{(l)}\lambda^{(l-1)})$ — разделенная разность оператора $P(\lambda)$.

Пусть

$$A_i=\begin{pmatrix} a_{1i} \\ \vdots \\ a_{mi} \end{pmatrix},$$

тогда

$$P(\lambda^{(l)}\lambda^{(l-1)})=(p_{ij}^{(l)})_{mm},$$

где

$$p_{ij}^{(l)}=(\lambda_j^{(l)}-\lambda_j^{(l-1)})^{-1}\{\sum_{s=1}^{m}a_{is}[x_s(\sum_{k=1}^{j-1}a_{ks}\lambda_k^{(l-1)}+$$

$$+\sum_{k=j}^{m}a_{ks}\lambda_k^{(l)})-x_s(\sum_{k=1}^{j}a_{ks}\lambda_k^{(l-1)}+\sum_{k=j+1}^{m}a_{ks}\lambda_k^{(l)})]\}.$$

Итак, на каждом итерационном шаге ($l=1, 2, \ldots$) приходится решать линейную m-мерную систему, причем нужные для проведения итераций значения функций $x_s(\cdot)$ вычисляются при решении соответствующих подзадач для фиксированных значений λ.

Между первым и вторым уровнями управления происходит следующий итерационный процесс:

а) второй уровень дает подсистемам l-е приближение для вектора множителей Лагранжа

$$\lambda^{(l)}=(\lambda_1^{(l)};\ldots,\lambda_m^{(l)})$$

(на первом шаге, $l=1$, — два приближения: $\lambda^{(0)}$ и $\lambda^{(1)}$);

б) подсистемы первого уровня ($i=1,\ldots,n$) решают каждый по m задач оптимизации: найти

$$x_i(\sum_{k=1}^{r-1}a_{ki}\lambda_k^{(l-1)}+\sum_{k=r}^{m}a_{ki}\lambda_k^{(l)})$$

($r=1,\ldots,m$; если $l=1$, то $r=1,\ldots,m+1$) из условия максимума функций (13);

в) на основании найденных x_i второй уровень вычисляет из системы (21) новое приближение $\lambda^{(l+1)}$;

г) процесс повторяется, пока уравнения (20) не удовлетворяются с заданной степенью точности.

Для исследования сходимости процесса координации можно применить общие теоремы о сходимости метода хорд (см., напр., [14]).

Приведем (в несколько иных обозначениях) пример, заимствованный из статьи [2]: Максимизировать функцию

$$f = -v_1^2 - 2v_2^2 - 4[3(u_1-2)^2 + 4(u_1-2)(v_1-3) + 2(v_1-3)^2] + 12 \tag{22}$$

при условиях

$$u_1 - 5 + 2[(v_2-1)^2 + (v_2-1)(u_2-2) + (u_2-2)^2] = 0, \tag{23}$$

$$u_2 - 2 + 4(u_1-2)^2 + 2(u_1-2)(v_1-3) + (v_1-3)^2 = 0, \tag{24}$$

$$u_2 \leqslant 0{,}8, \tag{25}$$

$$u_1 + v_1 \leqslant 5. \tag{26}$$

В данном случае ограничения (23), (24) нелинейные. Разбив задачу (22)—(26) по [2] на подзадачи, получаем для нахождения оптимального λ задачу решения нелинейной системы (23), (24), где $u_i = u_i(\lambda)$ и $v_i = v_i(\lambda)$ определяются решениями подзадач. В качестве начальных приближений для λ были выбраны

$$\lambda^{(0)} = (0{,}005;\ 1{,}9),$$

$$\lambda^{(1)} = (0{,}001;\ 2{,}0).$$

После 14-й итерации были получены следующие результаты:

$$\begin{array}{lll} u_1 = 2{,}607144; & v_1 = 2{,}086050; & \lambda = (3{,}689115;\ \ -1{,}200749) \\ u_2 = 0{,}8; & v_2 = 1{,}258783; & f = 2{,}251787. \end{array}$$

Для решения задачи, приведенной в пункте 4, были выбраны

$$\lambda^{(0)} = (1{,}059817;\ \ -0{,}270712;\ \ -0{,}258189);$$
$$\lambda^{(1)} = (0{,}830102;\ \ -0{,}204106;\ \ -0{,}198435).$$

После восьми итераций были получены следующие результаты:

$$\lambda = (0{,}525122;\ \ -0{,}058179;\ \ -0{,}163203);$$

$$\begin{array}{lll} x_{11} = 0{,}708349 & x_{22} = 0{,}806951 & x_{32} = \ \ 0{,}574234 \\ x_{12} = 0{,}416698 & x_{23} = 0{,}060445 & \psi(\lambda) = -2{,}777483 \\ x_{21} = 0{,}060445 & x_{31} = 0{,}393275 & f(x) = -2{,}777483. \end{array}$$

7. Рассмотрим, наконец, задачу нелинейного программирования [8]:

$$\min_{x \in X} f(x), \tag{27}$$

где
$$x = (x_1, \ldots, x_m, \bar{x});\ \ f(x) = \sum_{i=1}^{m} f_i(x_i, \bar{x});$$

$$x_i \in X_i;\ \ \bar{x} \in \bar{X};\ \ \ X = X_1 \times \ldots \times X_m \times \bar{X}.$$

В [8] предлагается следующий метод решения задачи (27) на двух уровнях:

а) проблемами подсистем ($i = 1, \ldots, m$) являются проблемы параметрической оптимизации

$$\min_{x_i \in X_i} f_i(x_i, \bar{x}) = f_i^*(\bar{x}); \tag{28}$$

б) задача второго уровня состоит в нахождении

$$\min_{\bar{x} \in \bar{X}} \sum_{i=1}^{m} f_i^* (\bar{x}) = \min_{x \in X} f(x). \tag{29}$$

Так как задачи (28), (29) в общем аналитически трудно разрешимы, то для их совместного решения можно применить метод покомпонентного спуска [11, 15], который в данном случае дает следующий итерационный процесс между подсистемами и вторым уровнем.

1) Пусть выбрано начальное приближение *

$$x^{(0)} = (x_1^{(0)}, \ldots, x_m^{(0)}, \bar{x}^{(0)}) \quad (x_i^{(0)} \in X_i; \bar{x}^{(0)} \in \bar{X});$$

2) если построено приближение $x^{(k)}$, то следующее приближение $x^{(k+1)}$ строится по правилам:

а) подсистемы решают задачи

$$\min_{x_i \in X_i} f_i(x_i, \bar{x}^{(k)}) \quad (i = 1, \ldots, m),$$

решения которых выбираются в качестве $x_i^{(k+1)}$;

б) второй уровень (центр) решает задачу

$$\min_{\bar{x} \in \bar{X}} \sum_{i=1}^{m} f_i(x_i^{(k+1)}, \bar{x}),$$

решение которой выбирается в качестве $\bar{x}^{(k+1)}$ **.

Для сходимости последовательности $\{x^{(k)}\}$ к точке минимума функции $f(x)$ на множестве X достаточны следующие условия [15]:

1) $X_i, \bar{X}$ — замкнутые выпуклые ограниченные множества;

2) $f(x)$ — выпуклая дифференцируемая функция.

Автор выражает признательность Э. Талкоп за вычисление примеров на ЭВМ.

ЛИТЕРАТУРА

1. Everett H., Oper Res., 3, 399 (1963).
2. Lasdon L. S., Schoeffler J. D., ISA Trans., 5, No. 2, 175 (1966).
3. Mesarovic M. D., Pearson J. D., Takahara Y., Joint Automat Contr. Conf., Rensselan Polytech. Inst. (Troy, N. Y.) Preprints papers, 93 (1965).
4. Pearson J. D., J. SIAM Control, 4, No. 1, 165 (1966).
5. Sekine Y., IEEE Trans. on Circuit Theory, CT-10, No. 2, 161 (1963).
6. Fukunaga K., Yoshida D., Papers on IFAC Symp. on Systems Eng. for Control System Design, Tokyo (1965).

* Часто в выборе $x_i^{(0)}$ $(i = 1, \ldots, n)$ нет необходимости.

** Более подробно эти вопросы рассматриваются в статье «Покомпонентный спуск и иерархическая оптимизация», которая будет опубликована в этом же журнале, № 2, 1969 г.

7. Danzig G. B., Wolfe P., Econometrica, 29, 767 (1961).
8. Findeisen W., Arch. Autom. i. Telemech., 12, No. 4, 391 (1967).
9. Карлин С., Математические методы в теории игр, программировании и экономике, М., 1964.
10. Гольштейн Е. Г., Экон. и мат. методы, 1, вып. 3, 410 (1965).
11. Ермольев Ю. М., Кибернетика, АН Укр. ССР, № 4, 1 (1966).
12. Schmidt J. W., Z. angew. Math. Mech., 43, No. 1 u 2, 1 (1963); 43, No. 3, 97 (1963).
13. Ульм С. Ю., Ж. вычисл. матем. и матем. физ., 4, № 6, 1093 (1964).
14. Ульм С., Изв. АН ЭССР, Сер. физ.-матем. и техн. наук, 13, № 3, 217 (1964).
15. Гольштейн Е. Г., Юдин Д. Б., Авт. и телемех., 24, № 12, 1643 (1963).

Институт кибернетики
Академии наук Эстонской ССР

Поступила в редакцию
25/III 1968

S. ULM

KAHENIVOOLISEST OPTIMISEERIMISEST

Ülesande (8)—(10) lahendamiseks esitatakse kahenivooline skeem. Sellele vastavalt jaotatakse ülesanne teatavaks arvuks parameetritest sõltuvateks alamülesanneteks, kusjuures optimaalsete parameetrite määramiseks (koordineerimiseks) rakendatakse gradientide meetodit. Antakse piisavad tingimused vastava kahenivoolise iteratsiooniprotsessi koonduvuseks. Koordineerimisprotsessi kiirendamiseks kasutatakse ka interpolatsioonimeetodeid. Skeemi rakendatakse ka diskreetsete optimaalse juhtimise ülesannete (15)—(19) kahenivooliseks lahendamiseks. Antakse kahenivooline iteratsiooniprotsess ja selle koonduvuse piisavad tingimused ülesande (27) puhul. Esitatakse kaks arvulist näidet.

S. ULM

ON TWO-LEVEL OPTIMIZATION

In the paper a two-level optimization scheme for problem (8)—(10) is put forward. The initial problem is decomposed into a number of parameter-dependent subproblems. The optimal values of the parameters are found in the process of an unconstrained minimization of some function via the gradient method (coordination). Sufficient conditions for the convergence of such two-level iteration process are given. Interpolation methods are used for the acceleration of the coordination process. The proposed two-level iteration scheme is applied for solving discrete optimal control problem (15)—(19). A two-level iteration process and sufficient conditions of the convergence for problem (27) are described. Two numerical examples are given.