

ФЕЙЕРОВСКИЕ ИТЕРАЦИОННЫЕ ПРОЦЕССЫ ДЛЯ НЕСОБСТВЕННЫХ ЗАДАЧ ЛИНЕЙНОГО ПРОГРАММИРОВАНИЯ

И.И. Еремин, И.М. Соколинская

In this article, we consider the Fejer's methods of solving the intrinsic (having a solution) and non-intrinsic (having no solution) linear optimization problems of 1-st, 2-nd and 3-rd types. The method is based on the reduction of mentioned problems to consistent and inconsistent systems of linear inequalities. For these systems, we construct the various variants of Fejerian processes which are converging to a solution or quasi-solution of the linear inequalities system. We discuss some aspects concerned to constructing of the iterative Fejerian mappings and the iterative sequences generated by them.

1. Введение

Основным методом решения задач линейного программирования (ЛП) является симплекс-метод. Он показал свою эффективность для задач ЛП средней размерности (несколько сотен и даже тысяч переменных и ограничений), решаемых на k -процессорных вычислительных машинах при небольших k , например при $k \leq 10$, и достаточно мощных процессорах. При больших размерностях задач ЛП (десятки и более тысяч переменных и ограничений) возникает необходимость их решения на многопроцессорных машинах (при значениях для k – сотни и тысячи процессоров) на основе принципа распараллеливания. Однако опыт решения больших задач показывает (это подкрепляется и теоретическим анализом [2]), что эффективность конечных методов, в частности симплекс-метода, падает с ростом числа задействованных процессоров в силу нарастания числа обменных операций. Это обстоятельство и порождает ограничения на использование симплекс-метода для решения больших задач ЛП. Определенный оптимизм можно связывать с использованием итерационных методов, дающих решение как предел итерационной последовательности, генерируемой тем или иным итерационным оператором. Известен широкий класс релаксационных методов фейеровского типа [1, 3], которые являются хорошими кандидатами для использования в указанных целях. Фейеровские методы отличаются простотой итерации, задаваемой формульно, удобны для массового распараллеливания

© 2002 И.И. Еремин, И.М. Соколинская

E-mail: ermii@imm.uran.ru, isok@csu.ru

Институт математики и механики УрО РАН, Челябинский государственный университет

Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований (проекты 00–15–96041, 01–01–00563)

вычислительного процесса, дают возможность реализации счета при эволюционном состоянии исходных данных задачи, т.е. ее информационной составляющей. Некоторый экспериментальный материал на этот счет имеется в [4].

В статье рассматриваются фейеровские методы для решения задач ЛП с акцентом на несобственные задачи, применительно к которым метод превращается в аппроксимационно-оптимизационный.

В работе будут использованы следующие обозначения:

\mathbf{R}^n – n -мерное евклидово пространство;

$\mathbf{R}_+^n := \{x \in \mathbf{R}^n \mid x^T = [x_1, \dots, x_n] \geq 0\}$;

$\alpha^+ = \max\{\alpha, 0\}$, $\alpha \in \mathbf{R}$ – поле действительных чисел;

$x^+ = [x_1^+, \dots, x_n^+]^T$ – положительная срезка вектора x ;

$\text{Arg} \dots$ – оптимальное множество выписанной вслед Arg оптимизационной задачи или ее номера; если " \dots " – система неравенств, то $\text{Arg} \dots$ – множество ее решений;

$\text{arg} \dots$ – конкретный элемент из $\text{Arg} \dots$;

$\text{opt} \dots$ – оптимальное значение оптимизационной задачи;

$|x|$ – евклидова норма вектора $x \in \mathbf{R}^n$;

F_M – класс M -фейеровских отображений;

\bar{F}_M – класс непрерывных M -фейеровских отображений;

$\text{rank } A$ – ранг матрицы A ;

$\{\varphi^k(x_0)\}_k$ – процесс (или последовательность), порождаемый отображением $\varphi(x) \in \{\mathbf{R}^n \rightarrow \mathbf{R}^n\}$ согласно рекуррентному соотношению: $x_{k+1} = \varphi(x_k)$, x_0 – начальный элемент процесса.

2. Краткие сведения о фейеровских и слабо фейеровских отображениях ([1], гл. II)

Пусть $\varphi \in \{\mathbf{R}^n \rightarrow \mathbf{R}^n\}$, $M \subset \mathbf{R}$. Отображение φ называется M -фейеровским, если

$$\varphi(y) = y, \quad |\varphi(x) - y| < |x - y|, \quad \forall y \in M, \quad \forall x \notin M.$$

Класс M -фейеровских отображений обозначим через F_M . Отображение $\varphi(\cdot) \in \{\mathbf{R}^n \rightarrow \mathbf{R}^n\}$ называется *слабо M -фейеровским*, если в выше приведенном определении неравенство " $<$ " заменить на " \leq ". Некоторые свойства фейеровских отображений:

1. Если $\varphi(\cdot) \in F_M$ и оператор $\varphi(x)$ непрерывен, т.е. $\varphi(\cdot) \in \bar{F}_M$, то

$$\{\varphi^k(x_0) =: x_k\}_k \rightarrow \bar{x} \in M \quad (1)$$

при произвольном начальном $x_0 \in \mathbf{R}^n$.

2. Если $\varphi_j(\cdot) \in F_{M_j}$, $j = 1, \dots, m$ и $M_0 = \bigcap_{j=1}^m M_j \neq \emptyset$, то

$$\varphi(x) := \sum_{j=1}^m \alpha_j \varphi_j(x) \in F_{M_0}, \quad \alpha_j > 0, \quad \sum_{j=1}^m \alpha_j = 1. \quad (2)$$

Следовательно, если $\forall j: M_j = M$, то $\varphi(x) \in F_M$.

3. Если $\varphi_j(\cdot) \in F_{M_j}$, $j = 1, \dots, m$, то

$$\varphi_{j_1} \varphi_{j_2} \cdots \varphi_{j_m}(x) \in F_{M_0}, \quad M_0 = \bigcap_{j=1}^m M_j \neq \emptyset; \quad (3)$$

(j_1, \dots, j_m) – любая перестановка индексов $\{1, \dots, m\}$.

4. Если $F_M \neq \emptyset$, то M автоматически выпукло и замкнуто.

5. Если $\varphi(\cdot)$ – слабо фейеровское отображение с множеством неподвижности $M \neq \emptyset$, то $\varphi_\alpha(x) := \alpha x + (1 - \alpha)\varphi(x) \in F_M$, $\alpha \in (0, 1)$. В частности, если отображение $\varphi(\cdot)$ является нерасширяющим, т.е. $|\varphi(x) - \varphi(y)| \leq |x - y|$, $\forall x, y \in \mathbf{R}^n$, и M – множество его точек неподвижности, то $\alpha x + (1 - \alpha)\varphi(x) \in \bar{F}_M$. Примером нерасширяющего отображения может служить операция проектирования на выпуклое замкнутое множество $M \subset \mathbf{R}^n$.

Примеры M -фейеровских отображений:

1⁰. Пусть M – выпуклое замкнутое множество из \mathbf{R}^n и $\text{Pr}_M(x)$ – оператор метрического проектирования на M . Тогда $\varphi(x) := \text{Pr}_M(x) \in \bar{F}_M$. Как уже было отмечено, $\varphi(x)$ является нерасширяющим.

2⁰. Пусть $H = \{x \mid l(x) := (a, x) - \alpha = 0\}$, $a \neq 0$ и $\varphi(x) = x - \lambda \frac{l^+(x)}{\|a\|^2} \cdot a$, $\lambda \in (0, 2)$. Тогда $\varphi(\cdot) \in \bar{F}_H$. Оператор $\varphi(x)$ реализует проектирование на H с релаксацией λ .

3⁰. Пусть $P = \{x \geq 0 \mid l(x) := (a, x) - \alpha \leq 0\}$, тогда $\varphi(x) := \left[x - \lambda \frac{l^+(x)}{\|a\|^2} a \right]^+ \in \bar{F}_P$.

4⁰. Пусть $M = \{x \mid l(x) := (a_j, x) - b_j \leq 0, j = 1, \dots, m\}$, $\forall j: a_j \neq 0$;
 $\varphi(x) = x - (\lambda / \delta) \sum_{j=1}^m l_j^+(x) a_j$, $\delta := \sum_{j=1}^m \|a_j\|^2$, $\lambda \in (0, 2)$. Тогда:

если $M \neq \emptyset$, то $\varphi(\cdot) \in \bar{F}_M$;

если $M = \emptyset$, то $\varphi(\cdot) \in F_{\tilde{M}}$, $\tilde{M} = \{x \mid \nabla d(x) = 0\}$, где $d(x) = \sum_{j=1}^m l_j^{+2}(x)$.

5⁰. Пусть $H := \{x \mid Ax = b\}$ и строчки матрицы A линейно независимы. Тогда

$$\varphi(x) := \text{Pr}_H(x) = x - A^T(AA^T)^{-1}(b - Ax) \quad (4)$$

и $\varphi(\cdot) \in \bar{F}_H$. Формула (4) хорошо известна из линейной алгебры.

3. Базовый фейеровский процесс

Если M – множество решений некоторой системы линейных неравенств, то относительно этой системы формально можно конструировать отображения $\varphi(\cdot) \in F_M$, но сходимость процесса $\{\varphi^k(x_0)\}_k$ гарантируется только при условии $M \neq \emptyset$. Мы хотим выделить из всех способов конструирования $\varphi(\cdot)$ базовый

способ, нацеленный на возможность использования его для построения сходящихся итерационных процессов в ситуации несовместности системы, т.е. когда $M = \emptyset$. Предел же последовательности, порожденной базовым отображением, должен давать *квазирешение* этой системы. Этот же процесс может быть приспособлен к нахождению квазирешений несобственных задач ЛП (посредством сведения задачи ЛП к системе линейных неравенств).

Будем исходить из системы линейных неравенств и уравнений, а также одного включения:

$$\left. \begin{aligned} l_j(x) &:= (a_j, x) - b_j \leq 0, & j \in J_{\leq}; \\ l_i(x) &:= (a_i, x) - b_i = 0, & i \in J_{=} \\ x &\in M_0. \end{aligned} \right\} \quad (5)$$

Система может быть как совместной, так и несовместной. Множество M_0 предполагается выпуклым и замкнутым. Положим

$$\varphi(x) = x - (\lambda / \delta) \left[\sum_{j \in J_{\leq}} R_j l_j^+(x) a_j + \sum_{i \in J_{=}} R_i l_i(x) a_i \right], \quad (6)$$

$\lambda \in (0, 2)$, $l_k(x) = (a_k, x) - b_k$, $k \in J := J_{\leq} \cup J_{=}$, $\delta = \sum_{k \in J} R_k |a_k|^2$, $\{R_k > 0\}_{k \in J}$

– параметры. Пусть

$$\psi(x) := \text{Pr}_{M_0} \varphi(x), \quad \psi_{\alpha}(x) := (1 - \alpha)\psi(x) + \alpha x, \quad (7)$$

$\alpha \in (0, 1)$. Под *базовой* конструкцией фейеровского оператора и *базовым* фейеровским процессом будем понимать $\psi_{\alpha}(x)$ и

$$\{\psi_{\alpha}^k(x_0)\}_k, \quad (8)$$

соответственно. Повсюду ниже x_0 – произвольный начальный элемент процесса.

Процесс (8), порождаемый оператором $\psi_{\alpha}(x)$ при произвольном начальном $x_0 \in \mathbf{R}^n$, сходится к некоторой точке \bar{x} , являющейся решением системы (5) в случае ее совместности; в случае же несовместности \bar{x} является некоторым *аппроксимационным квазирешением*. Вопросы сходимости процесса (8) и содержательного смысла предельной точки этого процесса будут рассмотрены ниже.

Пусть $d(x) = \sum_{j \in J_{\leq}} R_j l_j^{+2}(x) + \sum_{i \in J_{=}} R_i l_i^2(x)$, $\widetilde{M} := \text{Arg} \min_{x \in M_0} d(x) \neq \emptyset$, $\varphi(x)$ – согласно (6).

Теорема 1. Пусть отображение $\psi_{\alpha}(x)$ задано согласно (7) и $\widetilde{M} \neq \emptyset$. Тогда $\psi_{\alpha}(\cdot) \in \bar{F}_{\widetilde{M}}$. ■

Следствие 1. Так как $\psi_{\alpha}(x)$ непрерывно, то согласно п. 1 раздела 2 процесс (8) сходится к некоторой точке $\bar{x} \in \widetilde{M}$. В частности, если $M = \mathbf{R}^n$ и система (5) совместна, то процесс (8) сходится к одному из ее решений. ■

Доказательству теоремы 1 предпшлем ряд лемм.

Лемма 1. (Она хорошо известна; см., например, [3], стр. 188). Пусть M – выпуклое замкнутое множество из \mathbf{R}^n , $p \notin M$ и $\bar{p} := \text{Pr}_M(p)$. Тогда $(p - \bar{p}, x - \bar{p}) \leq 0$, $\forall x \in M$. Верно и обратное утверждение. ■

Лемма 2. Для любых x и y из \mathbf{R}^n справедливо неравенство

$$|\text{Pr}_M(x) - \text{Pr}_M(y)| \leq |x - y|, \quad (9)$$

т.е. $\text{Pr}_M(x)$ является оператором нерастяжения (см., например, [1], гл. II). ■

Лемма 3. Пусть $P_j := \{x \mid l_j(x) := (a_j, x) - b_j \leq 0\}$, $a_j \neq 0$, $\varphi_j(x) = x - \lambda_j \frac{l_j^+(x)}{|a_j|^2} \cdot a_j$, $\lambda_j \in (0, 2)$. Тогда $\varphi_j(\cdot) \in F_{P_j}$ (см. п. 3⁰ раздела 2) и

$$|\varphi_j(x) - \varphi_j(y)| \leq |x - y| \quad (10)$$

для любых x и y из \mathbf{R}^n . Проверяется непосредственно. ■

Замечание. Если $P_i = \{x \mid l_i(x) = 0\}$ и $\bar{\varphi}_i(x) = x - \lambda \frac{l_i(x)}{|a_i|^2} a_i$, то свойство нерастяжения для $\bar{\varphi}_i$ также выполняется.

Пусть M – выпуклое замкнутое множество из \mathbf{R}^n .

Лемма 4. Отображение $\psi(x)$, заданное согласно $\psi(x) := \text{Pr}_M \varphi(x)$, $\varphi(x)$ – согласно (6), является оператором нерастяжения.

Доказательство. Отображение (6) можно записать в виде:

$$\varphi(x) = \sum_{j \in J_{\leq}} \alpha_j \left(x - \lambda \frac{l_j^+(x)}{|a_j|^2} \cdot a_j \right) + \sum_{i \in J_{=}} \alpha_i \left(x - \lambda \frac{l_i(x)}{|a_i|^2} \cdot a_i \right),$$

где $\alpha_k = \frac{R_k |a_k|^2}{\delta}$, $\delta = \sum_{j \in J_{\leq}} R_j |a_j|^2 + \sum_{i \in J_{=}} R_i |a_i|^2$, $k \in J_{\leq} \cup J_{=}$. Следовательно,

$\sum_{k \in J_{\leq} \cup J_{=}} \alpha_k = 1$. Опираясь последовательно на леммы 2 и 3, будем иметь:

$$\begin{aligned} & |\text{Pr}_M \varphi(x) - \text{Pr}_M \varphi(y)| \leq |\varphi(x) - \varphi(y)| = \\ & = \left| \sum_{j \in J_{\leq}} \alpha_j (\varphi_j(x) - \varphi_j(y)) + \sum_{i \in J_{=}} \alpha_i (\bar{\varphi}_i(x) - \bar{\varphi}_i(y)) \right| \leq \\ & \leq \sum_{j \in J_{\leq}} \alpha_j |\varphi_j(x) - \varphi_j(y)| + \sum_{i \in J_{=}} \alpha_i |\bar{\varphi}_i(x) - \bar{\varphi}_i(y)| \leq \\ & \leq \sum_{j \in J_{\leq}} \alpha_j |x - y| + \sum_{i \in J_{=}} \alpha_i |x - y| = \left(\sum_{j \in J_{\leq}} \alpha_j + \sum_{i \in J_{=}} \alpha_i \right) |x - y| = |x - y|, \end{aligned}$$

что и требовалось. ■

Отображение (7) перепишем в следующей форме

$$\varphi(x) = x - (\lambda/2\delta) \nabla d(x), \quad (11)$$

где $d(x)$ и δ уже были введены, ∇ – символ градиента.

Лемма 5. Пусть $f(x)$ – выпуклая дифференцируемая функция и $\nabla f(\bar{x}) \neq 0$. Тогда при достаточно малом $\alpha > 0$: $f(\bar{x} - \alpha \nabla f(\bar{x})) < f(\bar{x})$ (известный факт из математического анализа). ■

Лемма 6. Пусть $\bar{x} \in M_0$ и $\widetilde{M} = \text{Arg} \min_{x \in M_0} f(x) \neq \emptyset$. Тогда

$$\bar{x} \in \widetilde{M} \iff \text{Pr}_M(\bar{x} - \gamma \nabla f(\bar{x})) = \bar{x}, \quad \gamma > 0. \quad (12)$$

Доказательство. Докажем " \Leftarrow ", т.е. достаточность правого соотношения в (12). Если $\nabla d(\bar{x}) = 0$, то $\bar{x} \in \text{Arg} \min_{(x)} d(x)$. Но так как $\bar{x} \in M_0$, то $\bar{x} \in$

$\text{Arg} \min_{x \in M_0} d(x)$, т.е. $\bar{x} \in \widetilde{M}$. Пусть теперь $\nabla d(\bar{x}) \neq 0$. Обозначим $p := \bar{x} - \gamma \nabla f(\bar{x})$. По лемме 1: $(p - \text{Pr}_M(p), x - \text{Pr}_M(p)) \leq 0, \forall x \in M_0$, или $(\bar{x} - \gamma \nabla f(\bar{x}) - \bar{x}, x - \bar{x}) = -\gamma(\nabla f(\bar{x}), x - \bar{x}) \leq 0, \forall x \in M_0$. Так как $f(x)$ выпуклая функция, то из полученного неравенства вытекает $0 \leq (\nabla f(\bar{x}), x - \bar{x}) \leq f(x) - f(\bar{x}), \forall x \in M_0$, или $f(x) \geq f(\bar{x}), \forall x \in M_0$, т.е. $\bar{x} \in \text{Arg} \min_{x \in M_0} f(x)$.

Докажем " \Rightarrow ", т.е. необходимость правого соотношения в (12). Если $\nabla d(\bar{x}) = 0$, то доказываемое соотношение выполняется очевидным образом, ибо $\bar{x} \in M_0$. Пусть $\nabla f(\bar{x}) \neq 0$. Тогда $p := \bar{x} - \gamma \nabla f(\bar{x}) \notin M_0$. Если это не так, то при любом $\alpha \in (0, 1)$: $p_\alpha := (1 - \alpha)\bar{x} + \alpha p = \bar{x} - \alpha\gamma \nabla f(\bar{x}) \in M_0$. Но тогда в силу леммы 5 при достаточно малом α : $f(p_\alpha) < f(\bar{x})$. Так как $p_\alpha \in M_0$, то получилось противоречие с тем, что $\bar{x} \in \text{Arg} \min_{x \in M_0} f(x)$. Итак, $p \notin M_0$.

Рассмотрим множество $\mathcal{N} := \{x \mid f(x) \leq f(\bar{x})\}$. Так как $M_0 \cap \mathcal{N}^0 = \emptyset$ ($\mathcal{N}^0 := \{x \mid f(x) < f(\bar{x})\}$), то по теореме отделимости существует гиперплоскость $H := \{x \mid (h, x - \bar{x}) = 0\}$ с $h \neq 0$, разделяющая M и \mathcal{N} , при этом можно считать, что h – внешняя нормаль к \mathcal{N} в точке \bar{x} , поэтому $(h, x - \bar{x}) \geq 0, \forall x \in M_0$. Так как $f(x)$ – выпуклая гладкая функция, то $h = \gamma_0 \nabla f(\bar{x}), \gamma_0 > 0$. Предыдущее неравенство можно переписать в виде: $((\underbrace{\bar{x} - \gamma_0 \nabla f(\bar{x})}_p) - \bar{x}, x - \bar{x}) \leq 0, \forall x \in M_0$.

По лемме 1: \bar{x} – проекция точки p на M_0 , т.е. $\text{Pr}_{M_0}(\bar{x} - \gamma_0 \nabla f(\bar{x})) = \bar{x}$. Заметим, что выполнимость или невыполнимость этого соотношения не зависит от значения величины $\gamma_0 > 0$.

Лемма 6 доказана полностью. ■

Перейдем к доказательству теоремы 1. По существу, оно вытекает из леммы 6 и п. 5 раздела 2. Действительно, пусть $\bar{x} \in \widetilde{M}$ и $x \notin M$. Согласно лемме 6: $\psi(\bar{x}) = \bar{x}$, а потому и $\psi_\alpha(\bar{x}) = \bar{x}$ (напомним: $\psi(x) = \text{Pr}_M \varphi(x), \varphi(x)$ – согласно (6)).

Далее, $|\psi(x) - \bar{x}| = |\psi(x) - \psi(\bar{x})| \stackrel{\text{лемма 4}}{\leq} |x - \bar{x}|$. Отсюда вытекает (согласно п. 5 раздела 2): $|\psi_\alpha(x) - \bar{x}| < |x - \bar{x}|$.

Теорема 1 доказана.

Замечание. Если $M_0 := \mathbf{R}^n$, то, очевидно, $\psi_\alpha(x) = \varphi_\alpha(x) := (1-\alpha)\varphi(x) + \alpha x$, причем $\psi_\alpha(\cdot) \in F_{\widetilde{M}}$. На самом деле уже само отображение $\varphi(x)$ является \widetilde{M} -фейеровским, т.е. $\varphi(\cdot) \in F_{\widetilde{M}}$. Действительно, положив $\lambda' = \lambda / (1-\alpha)$, будем иметь: $\varphi(x) = (1-\alpha) \underbrace{[x - (\lambda' / \delta) \nabla d(x)]}_{\varphi_0(x)} + \alpha x$, при этом $\lambda' \in (0, 2)$ при $\alpha < 1 - \lambda / 2$.

Так как оператор $\varphi_0(x)$ – нерастягивающий, то опять-таки в силу п. 5 раздела 2: $\varphi(x) = (1-\alpha)\varphi_0(\cdot) + \alpha x \in F_{\widetilde{M}}$.

4. Фейеровский процесс для решения задачи линейного программирования

Пусть

$$L : \max \{(c, x) \mid Ax \leq b, x \geq 0\} \quad (13)$$

– разрешимая задача ЛП;

$$L^* : \min \{(b, u) \mid A^T u \geq c, u \geq 0\} \quad (14)$$

– двойственная к L .

Задачам L и L^* поставим в соответствие систему линейных неравенств

$$\left. \begin{aligned} Ax \leq b, \quad x \geq 0 & \quad ; & (15)_1 \\ S : \quad A^T u \geq c, \quad u \geq 0 & \quad ; & (15)_2 \\ (c, x) = (b, u) & \quad , & (15)_3 \end{aligned} \right\} \quad (15)$$

называемую *симметрической*.

В теории ЛП хорошо известен следующий факт:

$$\boxed{\text{Arg } S = \text{Arg } L \times \text{Arg } L^*} \quad (16)$$

В (16): $\text{Arg } S$ – множество решений системы S .

Если

$$\begin{aligned} \varphi_1(x) &\in F_{M_1}, & M_1 &:= \text{Arg } (15)_1; \\ \varphi_2(u) &\in F_{M_2}, & M_2 &:= \text{Arg } (15)_2; \\ \varphi_3(x, u) &:= \text{Pr}_H(x, u), & H &= \text{Arg } (15)_3 \end{aligned}$$

и

$$\varphi(x, u) := \text{Pr}_H(\varphi_1^+(x), \varphi_2^+(u)), \quad (17)$$

то $\varphi(x, u) \in \widetilde{F}_{\widetilde{M}}$, где $\widetilde{M} = \text{Arg } S$, т.е. $\varphi(x, u)$ является непрерывным фейеровским отображением относительно множества $\text{Arg } L \times \text{Arg } L^*$. Следовательно (в силу п. 1 раздела 3):

$$\{\varphi^k(x_0, u_0)\}_k \rightarrow [\bar{x}, \bar{u}] \in \text{Arg } S. \quad (18)$$

Что касается выбора отображений $\varphi_1(x)$ и $\varphi_2(x)$, то это можно делать разными способами, в частности *базовым* способом, изложенными в разделе 3, а именно:

$$\left. \begin{aligned} \varphi_1(x) &= x - (\lambda / \delta_1) \sum_{j=1}^m l_j^+(x) a_j, \\ \varphi_2(u) &= u - (\lambda / \delta_2) \sum_{i=1}^m h_i^+(u) h_i, \\ \varphi_3(x, u) &= [x, u]^T - \frac{(c, x) - (b, u)}{|c|^2 + |b|^2} [c, -b]^T; \end{aligned} \right\} \quad (19)$$

выше $\{h_i\}$ – столбцы матрицы A , $h_i(u) = c_i - (h_i, u)$, $\delta_1 = \sum_{j=1}^m |a_j|^2$, $\delta_2 = \sum_{i=1}^n |h_i|^2$,

$$\lambda \in (0, 2), \quad [x, u]^T = \begin{bmatrix} x \\ u \end{bmatrix}, \quad [c, -b]^T = \begin{bmatrix} c \\ -b \end{bmatrix}.$$

Заметим, что вид отображений $\varphi_1(\cdot)$ и $\varphi_2(\cdot)$ будет меняться в зависимости от того, в какой форме записана исходная задача линейного программирования. Если, например, задача L задана в канонической форме

$$\min \{(c, x) \mid Ax = b, x \geq 0\}, \quad (20)$$

то двойственная к ней будет иметь вид

$$\max \{(b, u) \mid A^T u \geq 0\}, \quad (21)$$

а отображения $\varphi_1(\cdot)$, $\varphi_2(\cdot)$ и $\varphi(x, u)$ соответственно примут вид:

$$\left. \begin{aligned} \varphi_1(x) &= x - (\lambda / \delta_1) \sum_{j=1}^m l_j(x) a_j, \\ \varphi_2(u) &= u - (\lambda / \delta_2) \sum [(h_i, u) - c_i]^+ h_i, \\ \varphi_3(x, u) &= \text{Pr}_H(\varphi_1(x), \varphi_2^+(u)). \end{aligned} \right\} \quad (22)$$

Вид $\varphi_3(x, u)$ формулы проектирования на H сохраняется.

В связи с тем, что основные ограничения в (20) записаны в форме системы линейных уравнений $Ax = b$, пусть S – множество решений H_1 , то $\varphi_1(\cdot)$ можно сконструировать, исходя из операции проектирования на H_1 , аналогично (4):

$$\text{Pr}_{H_1}(x) = x - A^T(AA^T)^{-1}(b - Ax) \quad (23)$$

(в предположении, что $\text{rank } A = n$). Полагая $\varphi_1(x) = \text{Pr}_{H_1}(x)$, заключительный итерационный оператор $\varphi(x, u)$ будет иметь ту же форму, что и в (22).

Случай задания отображения $\varphi(x, u)$ применительно к задаче (20) в форме

$$\varphi(x, u, v) := \text{Pr}_{H_0}(x, u^+, v^+), \quad (24)$$

где H_0 – множество решений системы

$$Ax = b, \quad A^T u + v = c, \quad (c, x) = (b, u), \quad (25)$$

рассматривался в работе [4], посвященной численным экспериментам решения задач ЛП большой размерности на многопроцессорных вычислительных машинах.

5. Базовый аппроксимационно-фейеровский процесс для несобственной задачи ЛП 1-го рода

Если исходная задача ЛП неразрешима, т.е. по принятой терминологии является *несобственной*, то ее соответствующая симметрическая система S является несовместной (верно и обратное). Для этой задачи в этом случае можно ввести понятие *квазирешения* через определение *квазирешения* системы S . Поясним это для случая задания задачи ЛП в форме

$$L : \max \{(c, x) \mid Ax \leq b\}. \quad (26)$$

Двойственной к ней будет

$$L^* : \min \{(b, u) \mid A^T u = c, u \geq 0\}. \quad (27)$$

Симметрической системой S является

$$\begin{aligned} Ax \leq b, \quad A^T u = c, \quad u \geq 0, \\ (c, x) = (b, u). \end{aligned} \quad (28)$$

Если L – несобственная задача ЛП 1-го рода, т.е.

$$M := \{x \mid Ax \leq b\} = \emptyset, \quad M^* := \{u \mid A^T u = c, u \geq 0\} \neq \emptyset,$$

то, введя

$$H = \left\{ \begin{bmatrix} x \\ u \end{bmatrix} \mid A^T u = c, (c, x) = (b, u) \right\}, \quad d(x, u) = |(Ax - b)^+|^2 + |(-u)^+|^2,$$

можно сформировать задачу

$$\min \{d(x, u) \mid u \in H\} \quad (29)$$

– как аппроксимационную для несовместной системы (28).

Если $\begin{bmatrix} \bar{x} \\ \bar{u} \end{bmatrix} \in \text{Arg}(29)$, то \bar{x} и будем называть квазирешением системы (28), а вместе с тем и квазирешением задачи (26). Если последняя разрешима, то \bar{x} – обычное ее решение. Для (29) можно записать фейеровское отображение относительно множества $\text{Arg}(29)$ по стандарту из раздела 3, но применительно к системе (28):

$$\begin{aligned} \varphi(x, u) &= \begin{bmatrix} x \\ u \end{bmatrix} - (\lambda / 2\delta) \nabla_{x,u} d(x, u), \\ \psi(x, u) &= \text{Pr}_H \varphi(x, u), \end{aligned}$$

$$\psi_\alpha(x, u) = (1 - \alpha) \Psi(x, u) + \alpha \begin{bmatrix} x \\ u \end{bmatrix}, \quad (30)$$

где $\lambda \in (0, 2)$, $\delta = \sum_{j=1}^m |a_j|^2 + m$, $\alpha \in (0, 1)$.

В соответствии с теоремой 1 $\widetilde{\psi}_\alpha(x, u)$ является \widetilde{M} -фейеровским непрерывным отображением относительно $\widetilde{M} = \text{Arg} (5.4)$. Следовательно, процесс

$$\{\psi_\alpha^k(x_0, u_0)\}_k$$

сходится к некоторому вектору $\begin{bmatrix} \bar{x} \\ \bar{u} \end{bmatrix}$, при этом по определению

\bar{x} – квазирешение задачи (26),

\bar{u} – квазирешение задачи (27).

Операция $\text{Pr}_H(\cdot)$, фигурирующая в определении $\psi(x, u)$ и входящая, следовательно, в формулу (30), может быть записана согласно соотношению (23) с заменой матрицы A на матрицу

$$\bar{A} := \begin{bmatrix} 0 & A^T \\ c^T & -b^T \end{bmatrix},$$

а вектор b (из (23)) – на $\begin{bmatrix} c \\ 0 \end{bmatrix}$.

6. Аппроксимационно-фейеровский процесс для несобственной задачи ЛП 2-го рода

Рассмотрение вопроса соотнесем к задачам (26)–(28). Ситуация несобственности 2-го рода соответствует тому, что $M \neq \emptyset$, $M^* = \emptyset$, где смысл символов M и M^* тот же, что и в разделе 5. Систему (28) разобьем на две части:

$$A^T u = c, \quad u \geq 0; \quad (31)$$

$$Ax + v = b, \quad v \geq 0, \quad (c, x) = (b, u); \quad (32)$$

выше система $Ax \leq b$ заменена на $Ax + v = b$, $v \geq 0$ – обычная перезапись системы неравенств, если в этом есть необходимость. Положим, $d(x, u, v) = |A^T u - c|^2 + |(-u)^+|^2 + |(-v)^+|^2$, H_0 – множество решений системы (32). В силу предположения о непустоте M имеем: $H_0 \neq \emptyset$. В согласии с базовой конструкцией фейеровского отображения $\psi_\alpha(\cdot)$ из раздела 3, реализованного для НЗ ЛП 1-го рода в разделе 5, соответствующим отображением для рассматриваемого случая НЗ ЛП 2-го рода будет отображение

$$\psi_\alpha(x, u, v) = \text{Pr}_{H_0} \left(\begin{bmatrix} x \\ u \\ v \end{bmatrix} - (\lambda / 2\delta) \nabla_{x,u,v} d(x, u, v) \right); \quad (33)$$

здесь $\lambda \in (0, 2)$, $\delta = \sum_{i=1}^n |h_i|^2 + 2m$.

Процесс, порожденный отображением (33), будет сходиться к некоторому вектору $[\bar{x}, \bar{u}, \bar{v}]^T$, при этом \bar{x} – квазирешение задачи (26), \bar{u} – квазирешение задачи (27).

7. Фейеровский процесс для НЗ ЛП 3-го рода

В предыдущих разделах 5 и 6 мы апеллировали к несовместности симметрической системы S , поставленной в соответствие исходной задаче ЛП, которая предполагалась либо НЗ ЛП 1-го рода, либо 2-го. Аппроксимационная суть, соотношенная системе S , заключалась в формировании ее совместной подсистемы из уравнений и квадратичной функции невязки $d(x)$ для оставшихся ограничений этой системы. Понятно, что эта операция неоднозначна, что является весьма полезным обстоятельством, ибо во всяком конкретном случае, т.е. в случае конкретного формата задачи (например, формата транспортной задачи, блочного формата и т.д.), указанная операция выделения подсистемы может быть увязана с потребностью простой и эффективной численной реализации итерационного шага, порождаемого итерационным оператором. К этому следует еще добавить, что построение итогового фейеровского итерационного оператора будет зависеть от конкретного вида исходной задачи ЛП. Всякий раз при построении нужного фейеровского отображения необходимо все это учитывать.

Для рассмотрения НЗ ЛП 3-го рода возьмем задачу ЛП, ей двойственную, и систему S в формате (13)–(15).

Так как итоговое итерационное отображение $\psi_\alpha(\cdot)$ формируется из фрагментов $\varphi(x)$, $\psi(x) = \text{Pr}_M \varphi(x)$, а они в свою очередь – из M , $d(\cdot)$ и δ , то, формируя аналоги отображений $\psi_\alpha(\cdot)$ применительно к системе (15), мы в рассматриваемых ниже вариантах будем приводить лишь вид множества M , функции невязки $d(\cdot)$ и числа δ .

$$\text{Вариант 1. } M_1 := \left\{ \begin{bmatrix} x \\ u \end{bmatrix} \mid (c, x) = (b, u) \right\}, \quad d_1(x, u) = |(Ax - b)^+|^2 + |(c - A^T u)^+|^2 + |(-x)^+|^2 + |(-u)^+|^2, \quad \delta_1 = \sum_{j=1}^m |a_j|^2 + \sum_{i=1}^n |h_i|^2 + m + n.$$

$$\text{Вариант 2. } M_2 := \left\{ \begin{bmatrix} x \\ u \\ v \end{bmatrix} \mid A^T u - v = c, (c, x) = (b, u) \right\}, \quad d_2(x, u, v) = |(Ax - b)^+|^2 + |(-x)^+|^2 + |(-u)^+|^2 + |(-v)^+|^2, \quad \delta_2 = \sum_{j=1}^m |a_j|^2 + m + 2n.$$

$$\text{Вариант 3. } M_3 := \left\{ \begin{bmatrix} x \\ u \\ v \end{bmatrix} \mid Ax - w = b, (c, x) = (b, u) \right\}, \quad d_3(x, u, w) = |(c - A^T u)^+|^2 + |(-x)^+|^2 + |(-u)^+|^2 + |(-w)^+|^2, \quad \delta_3 = \sum_{i=1}^n |h_i|^2 + 2m + n.$$

Вариант 4. $M_4 := \mathbf{R}^n$, $d_4(x, u) = |(Ax - b)^+|^2 + |(c - A^T u)^+|^2 + |(-x)^+|^2 + |(-u)^+|^2$,

$$\delta_4 = \sum_{j=1}^m |a_j|^2 + \sum_{i=1}^n |h_i|^2 + m + n.$$

Последний вариант годится как для разрешимой задачи, так и для неразрешимой любого рода неразрешимости (1-го, 2-го или 3-го).

8. Приближенный метод решения разрешимой задачи ЛП, основанный на ее редукции к задаче минимизации кусочно-квадратичной выпуклой функции

Нам безразлично, в какой форме представить исходную задачу ЛП, но для простоты последующего формульного текста запишем ее в виде

$$L : \min \{(c, x) \mid Ax \leq b\}. \quad (34)$$

Известен результат по методам штрафных функций [3, теорема 20.2]: *Задача (34) асимптотически по $R = [R_1, \dots, R_m] > 0$ эквивалентна задаче*

$$\min \left\{ (c, x) + \sum_{j=1}^m R_j l_j^{+2}(x) \right\}, \quad (35)$$

т.е. задача (35) приближает (как по аргументу, так и по значению) задачу (34), при этом тем точнее, чем больше число $\min_{(j)} R_j$. На самом деле имеются оценки для соответствующих уклонений, но приводить их не будем. Результат верен и для нелинейной целевой функции в задаче (34).

Пусть γ_0 – некоторая нижняя оценка для opt (8.1), т.е. $(c, x) \geq \gamma_0$, $\forall x \in M := \{x \mid Ax \leq b\}$. В конкретных задачах такую оценку получить не составляет труда. Именно в силу того, что $(c, x) - \gamma_0 \geq 0$ для всех допустимых x , задачу (34) можно записать в эквивалентном виде:

$$\min \{[(c, x) - \gamma_0]^2 \mid x \in M\}. \quad (36)$$

Применив к ней приведенный выше метод квадратичных штрафных функций, получим задачу

$$\min_{(x)} \{l_0^2(x) + \sum_{j=1}^m R_j l_j^{+2}(x)\}, \quad (37)$$

где $l_0(x) = (c, x) - \gamma_0$. Так что задача (37) приближенно решает исходную задачу (34). Но минимизируемая функция в задаче (37) является частным случаем для функции $d(x)$ из раздела 3, так что для поиска точки из $\text{Arg} \min_{(x)} d(x)$ можно

применить фейеровский процесс

$$\{\varphi^k(x_0)\}_k, \quad (38)$$

в котором

$$\varphi(x) = x - (\lambda / \delta) \nabla d(x), \quad \lambda \in (0, 2), \quad \delta = |c|^2 + \sum_{j=1}^m |a_j|^2,$$

$$d(x) = [(c, x) - \gamma_0]^2 + \sum_{j=1}^m R_j l_j^{+2}(x),$$

$$\nabla d(x) = 2 \left\{ [(c, x) - \gamma_0]c + \sum_{j=1}^m R_j l_j^+(x) a_j \right\}.$$

Если \bar{x} – предел сходящейся последовательности, порождаемой процессом (37) при произвольном $x_0 \in \mathbf{R}^n$, то \bar{x} решает приближенно задачу (34), и тем точнее, чем больше числа $R_j > 0$, $j = 1, \dots, m$. Поскольку $\bar{x} \in \underset{(x)}{\text{Arg min}} d(x)$, то $\nabla d(\bar{x}) = 0$, а потому признаком останова вычислительного процесса может служить выполнимость неравенства $|\nabla d(x_k)| \leq \varepsilon$, ε – наперед заданное положительное число.

9. Декомпозиционные методы построения фейеровских отображений применительно к системам линейных неравенств

Как уже отмечалось (см. раздел 4), решение задачи ЛП сводится к решению системы линейных неравенств S с матрицей коэффициентов, имеющей блочную структуру. Если взять в качестве исходной задачи задачу (13), то системой S является (15) с матрицей коэффициентов перед переменными x_i и y_j :

$$\bar{A} := \begin{bmatrix} \boxed{A} & & \\ & \boxed{-A^T} & \\ \boxed{c^T} & & \boxed{-b^T} \end{bmatrix}.$$

Построение отображения (17) было основано на декомпозиционном принципе, учитывающем блочную структуру матрицы \bar{A} . Ниже этот вопрос рассмотрим подробнее применительно к канонической системе линейных уравнений и неравенств

$$Ax = b, \quad x \geq 0 \tag{39}$$

при разных вариантах блочности матрицы A (в скобках заметим, что символ A в данном контексте не следует связывать с символом A в предыдущих записях).

Вариант 1. Пусть матрица A разбита на горизонтальные подматрицы A_i :

$$\left[\begin{array}{c} \boxed{A_1} \\ \vdots \\ \boxed{A_n} \end{array} \right], \quad (40)$$

что соответствует разбиению системы (39) на подсистемы

$$A_i x = b^i, \quad x \geq 0, \quad i = 1, \dots, n. \quad (41)$$

Пусть $H_i = \{x \mid A_i x = b^i\}$. Тогда

$$\varphi(x) = \left[\sum_{i=1}^n \alpha \Pr_{H_i}(x) \right]^+ \in F_M,$$

где $M = \{x \geq 0 \mid Ax \leq b\}$ (см. п. 2 и 5⁰ из раздела 2). На самом деле $\varphi(x)$ непрерывно, поэтому $\varphi(\cdot) \in \bar{F}_M$, что дает

$$\{\varphi^k(x_0)\}_k \rightarrow \bar{x} \in M.$$

Целесообразность разбиения системы (39) на подсистемы можно связать с проблемой преодоления трудностей, связанных с обращением матрицы AA^T (см. формулу проектирования (4) на $H = \{x \mid Ax = b\}$). При разбиении системы (39) на подсистемы (41) нужно будет обращать матрицы $A_i A_i^T$ меньших размеров. К тому же само разбиение, учитывающее структуру матрицы A , может быть нацелено именно на эффективность указанных обращений.

Вариант 2. Пусть

$$A := \left[\begin{array}{ccc} \boxed{A_1} & & \\ & \ddots & \boxed{A_n} \\ \boxed{A_0} & & \end{array} \right],$$

тогда система (39) имеет вид

$$\left. \begin{array}{l} A_i x^i = b^i, \quad i = 1, \dots, n, \\ A_0 x = b^0, \quad x^T = [x^1, \dots, x^n] \geq 0 \end{array} \right\} \quad (42)$$

и $H_i := \{x^i \mid A_i x^i = b^i\} \subset \mathbf{R}^{n_i}$, $H_0 = \{x \mid A_0 x = b^0\}$.

Если подсистемам $A_i x^i = b^i$, $i = 1, \dots, n$, $A_0 x = b^0$ соотнести $\varphi_i(\cdot) \in \bar{F}_{H_i}$ и $\varphi_0(x) \in \bar{F}_{H_0}$, то

$$\varphi(x) := \varphi_0^+(\varphi_1(x^1), \dots, \varphi_n(x^n)) \in \bar{F}_M,$$

причем выбор отображений $\{\varphi_i\}_{i=0}^n$ может осуществляться разными способами, например, $\varphi_i(x^i) = \Pr_{H_i}(x^i)$, $i = 1, \dots, n$; $\varphi_0(x) = \Pr_{H_0}(x)$.

Вариант 3. Рассмотрим матрицу

$$\left[\begin{array}{ccc} \boxed{A_1} & & \\ & \ddots & \\ & & \boxed{A_n} \end{array} \begin{array}{c} \boxed{B_1} \\ \vdots \\ \boxed{B_n} \end{array} \right]$$

и ей соответствующую систему

$$\left. \begin{array}{l} A_i x^i + B_i x^0 = b^i, \quad i = 1, \dots, n, \\ x^T = [x^1, \dots, x^n, x^0] \geq 0. \end{array} \right\} \quad (43)$$

Подсистеме $A_i x^i + B_i x^0 = b^i$ поставим в соответствие $\varphi_i(x^i, x^0) \in \bar{F}_{H_i}(x^i, x^0)$, где $H_i = \{[x^i, x^0] \mid A_i x^i + B_i x^0 = b^i\}$. Обозначим через \bar{x}^i алгебраический след вектора $\varphi_i(x^i, x^0)$ в \mathbf{R}^{n_i} – как подпространство пространства $(\prod_{i=1}^n \mathbf{R}^{n_i}) \times \mathbf{R}^{n_0}$ исходной переменной $x^T = [x^1, \dots, x^n, x^0]$, $(\bar{x}^0)_i$ – аналогичный след в \mathbf{R}^{n_0} . образуем отображение

$$\varphi(x) = \left[\bar{x}^1, \dots, \bar{x}^n, \frac{1}{n} \sum_{i=1}^n (\bar{x}^0)_i \right]^+.$$

Оно является непрерывным M -фейеровским [5].

10. Вычислительные аспекты применения фейеровских отображений

Обсудим некоторые вопросы эффективных реализаций фейеровских процессов, соотнесенных задачам поиска решения, либо системы линейных неравенств, либо задачи линейного программирования. Остановимся просто на системе линейных неравенств

$$Ax \leq b \sim l_j(x) := (a_j, x) - b_j \leq 0, \quad j = 1, \dots, m, \quad (44)$$

поскольку задачи ЛП сводятся к таким системам. Будем предполагать $M := \{Ax \leq b\} \neq \emptyset$. В разделе 9 мы уже обсуждали вопрос об учете структуры матрицы A при построении фейеровского итерационного отображения $\varphi(\cdot) \in \bar{F}_M$. Несомненно, эффективность вычислительного процесса, порождаемого отображением $\varphi(\cdot)$, зависит от разумного учета особенностей структуры матрицы A . Но это только одна сторона дела. Само генерирование итерационной последовательности допускает большое число модификаций, основанных на иных соображениях. Пусть выбрано для системы (44) базовое отображение

$$\varphi(x) = x - (\lambda / \delta) \sum_{j=1}^m l_j^+(x) a_j \in \bar{F}_M.$$

Для $k < m$ введем

$$\varphi_k(x) = x - (\lambda / \delta_k) \sum_{j=1}^k l_j^+(x) a_j, \quad \delta_k = \sum_{j=1}^k |a_j|^k$$

– это отображение, соотнесенное подсистеме $l_j(x) \leq 0$, $j = 1, \dots, k$. При больших значениях m (тысячи и десятки тысяч) целесообразней счет вести согласно соотношению

$$x_{k+1} = \varphi_k^{t_k}(x_k), \quad (45)$$

при этом $\varphi_k(x) = \varphi(x)$ и $t_k = 1$ при $k \geq m$.

В процессе (45) предусмотрены по крайней мере два обстоятельства:

1) на шаге с номером k (при $k < m$) итерация $\varphi_k(x_k)$ выполняется для подсистемы из k неравенств, что сокращает число арифметических и логических операций, необходимых для подсчета $\varphi_k(x_k)$;

2) так вычисляемые $\varphi_k(\cdot)$ можно повторить несколько, пусть t_k , раз, до обеспечения, например, условия $\sum_{j=1}^k l_j^+(x_k) \leq \varepsilon$. После этого срабатывает интегрированная итерация (45).

При такой тактике перехода от x_k к x_{k+1} подсчет δ_{k+1} ведется рекуррентно: $\delta_{k+1} = \delta_k + |a_{k+2}|^2$.

При описанной организации счета факт сходимости итерационной последовательности $\{x_k\}$ к решению системы (44) сохраняется.

Заметим, что рекуррентное соотношение (45) пригодно и для решения бесконечных систем линейных неравенств ([1], стр. 106).

Что касается скорости сходимости рассмотренных фейеровских процессов применительно к совместным системам линейных неравенств, пусть это будет (44), то ее можно охарактеризовать *скоростью* по геометрической прогрессии, понимаемой в смысле

$$|x_{k+1} - M| \leq \theta |x_k - M|, \quad \forall k, \quad \theta \in (0, 1);$$

здесь $|x - M| := \inf_{y \in M} |x - y|$. Оценки для θ были даны в работе одного из авторов ([1], стр. 68).

По фейеровским методам проводились вычислительные эксперименты на многопроцессорной вычислительной машине МВС–100/1000 с применением методов распараллеливания (см. [2–4]). Решались задачи ЛП с числом переменных $n \leq 4000$ и числом ограничений $m \leq 1000$. В основу метода был положен итерационный процесс $z_{k+1} = \text{Pr}_M^+(z_k)$ для системы $Az = b$, $z \geq 0$, к которой редуцировалась задача ЛП. Результаты экспериментов были удовлетворительными. Эксперименты планируются продолжить.

ЛИТЕРАТУРА

1. Еремин И.И., Мазуров В.Д. *Нестационарные процессы математического программирования*. М.: Наука, 1979. –205 с. (гл. II).
2. Попов Л.Д. *Вопросы реализации методов ЛП в траспьютерных сетях* / Сб.: «Алгоритмы и программные средства параллельных вычислений». Екатеринбург: Изд-во УрО РАН, 1995. N.1. С.148-156.
3. Еремин И.И. *Теория линейной оптимизации*. Екатеринбург: Изд-во УрО РАН, 1999. –312 с.
4. Бердникова Л.Д., Попов Л.Д. *О применении декомпозиции при реализации фейеровских методов решения больших систем линейных неравенств на МВС-100* / Сб.: «Алгоритмы и программные средства параллельных вычислений». Екатеринбург: Изд-во УрО РАН, 2000. N.4. С.51-62.
5. Еремин И.И. *Синтез фейеровских отображений с несовпадающими пространствами их образов* // ДАН. 2001. Т.378, N.1. С.11-13.