ТРУДЫ ИНСТИТУТА МАТЕМАТИКИ И МЕХАНИКИ УрО РАН

Том 23 № 3 2017

УДК 517.988.68

О МЕТОДАХ ОПТИМИЗАЦИИ ФУНКЦИИ ЧУВСТВИТЕЛЬНОСТИ ПРИ ОГРАНИЧЕНИЯХ[1]

А. С. Антипин

Рассматривается параметрическое семейство задач выпуклого программирования. В качестве параметра выступает вектор правых частей функциональных ограничений задачи. Каждому векторному значению параметра, взятому из неотрицательного ортанта, отвечает регулярная (условие Слейтера) задача выпуклого программирования и ее минимальное значение целевой функции. Это значение, зависящее от параметра ограничений, порождает функцию чувствительности. Наряду с этой функцией априори задается выпуклое множество (геометрически или функционально заданное). Ставится задача минимизации неявно заданной функции чувствительности на этом множестве. Такая задача имеет содержательную интерпретацию как задача выпуклого программирования, когда вместо заданного вектора правых частей функциональных ограничений указывается только множество, которому этот вектор принадлежит. В результате получаем двухуровневую задачу. В отличие от классических двухуровневых иерархических задач, где неявно задаются ограничения, в нашем случае неявно задаются целевые функции. Никакой иерархии в этой задаче нет. Как правило функции чувствительности обсуждаются в научной литературе в более общем контексте как функции оптимального значения. Автору не известны оптимизационные постановки этих задач как самостоятельных исследований и, тем более, не известны предлагаемые методы их решения. В работе предлагается оригинальный седловой подход к решению задач с функциями чувствительности. Доказывается монотонная сходимость метода к решению задачи по переменным пространства, в котором рассматривается задача.

Ключевые слова: функция чувствительности, параметрическая оптимизация, параметрическая функция Лагранжа, седловая точка, экстрапроксимальные методы, сходимость.

A. S. Antipin. Optimization methods for the sensitivity function with constraints.

We consider a parametric family of convex programming problems. The parameter is the vector of the right-hand sides in the functional constraints of the problem. Each vector value of the parameter taken from the nonnegative orthant corresponds to a regular (Slater's condition) convex programming problem and the minimum value of its objective function. This value depends on the constraint parameter and generates the sensitivity function. Along with this function, a convex set is given geometrically or functionally. The problem of minimization of the implicit sensitivity function on this set is posed. It can be interpreted as a convex programming problem in which, instead of a given vector of the right-hand sides of functional constraints, only a set to which this vector belongs is specified. As a result, we obtain a two-level problem. In contrast to the classical two-level hierarchical problems with implicitly given constraints, it is objective functions that are given implicitly in out case. There is no hierarchy in this problem. As a rule, sensitivity functions are discussed in the literature in a more general context as functions of the optimal value. The author does not know optimization statements of these problems as independent studies or, even more so, solution methods for them. A new saddle approach to the solution of problems with sensitivity functions is proposed. The monotone convergence of the method is proved with respect to the variables of the space in which the problem is considered.

Keywords: sensitivity function, parametric optimization, parametric Lagrangian, saddle point, extraproximal methods, convergence.

MSC: 90C25, 90C31, 90C46, 90C90, 49K40

DOI: 10.21538/0134-4889-2017-23-3-33-42

1. Постановка задачи

Задача минимизации функции чувствительности представляет собой систему двух задач. Одна из них является параметрической задачей выпуклого программирования, которая, собственно говоря, относительно параметра порождает функцию чувствительности. Другая задача — это задача оптимизации функции чувствительности на выпуклых множествах:

$$\varphi(y) = f(x^*(y)) = \mathrm{Min}\{f(x) \mid g(x) \le y,\ x \in X \subseteq \mathbb{R}^n\}, \quad y \in \mathbb{R}^m_+, \tag{1.1}$$

[1]Работа выполнена при финансовой поддержке Российского научного фонда (проект 17-11-01353).

$$y^* \in \operatorname{Arg}\min\{\varphi(y) \mid y \in Y \subseteq \mathbb{R}^m_+\}. \tag{1.2}$$

Здесь $f(x)$ — выпуклая скалярная, $g(x)$ — векторная функции, причем каждая компонента векторной функции также выпуклая функция, $y \geq 0$ — параметр, $\mathbb{R}^m_+$ — положительный ортант, $X \subset \mathbb{R}^n$, $Y \subset \mathbb{R}^m_+$ — выпуклые замкнутые множества. Совокупности решений задач (1.1), (1.2) представляют собой выпуклые замкнутые множества X^*, Y^*, тогда $x^* \in X^*$, $y^* \in Y^*$. Фиксированное множество Y в этой работе будет выступать в двух формах: в геометрической форме как выпуклое замкнутое множество и в форме, заданной системой функциональных неравенств $Y = \{y \mid g_1(y) \leq y_1\}$, $y_1 \in \mathbb{R}^q_+$. Все компоненты векторной функции $g_1(y)$ выпуклы. В частности, функциональные ограничения могут порождать многогранники.

Первая задача из системы (1.1), (1.2) порождает функцию чувствительности. Это происходит следующим образом: когда параметр $y \in \mathbb{R}^m_+$ пробегает положительный ортант, внутренняя задача по переменной $x \in X$ порождает оптимальное значение $f(x^*(y))$, которое присваивается функции $\varphi(y)$, вычисленной в точке $y \in \mathbb{R}^m_+$. В системе (1.1), (1.2) требуется найти минимум функции чувствительности на множестве $y \in Y$, при этом целевая функция задана неявно.

Задачу (1.1) можно интерпретировать как модель производства, в которой требуется выбрать вектор ресурсов $y \in Y$ и отвечающий ему вектор интенсивностей $x \in X$ работы предприятия так, чтобы обеспечить выпуск продукции с минимальными затратами. Эффективность предприятия, зависящая от ресурсов, оценивается величиной $f(x^*(y))$ в точке минимума $x^*(y)$ из задачи (1.1). Очевидно, при разных наборах ресурсов эффективность предприятия, вообще говоря, будет разная. В этой ситуации возникает задача выбора вектора ресурсов из некоторого фиксированного геометрического множества $Y \in \mathbb{R}^m_+$ или функционально заданного множества $Y = \{y \mid g_1(y) \leq y_1\}$, $y_1 \in \mathbb{R}^q_+$, так, чтобы обеспечить наилучшую эффективность предприятия. Формальное описание этой ситуации приводит нас к задаче (1.1), (1.2). Если в этой задаче операцию min заменить на операцию max, то можно говорить о максимальной выгоде предприятия при выборе того или иного набора ресурсов.

Функция $\varphi(y)$, $y \in \mathbb{R}^m_+$, известна как функция чувствительности. Перечислим основные свойства этой функции (см., например, [1]), где приведена соответствующая библиография.

1) Функция чувствительности является монотонно убывающей.

2) Надграфик функции чувствительности — выпуклое замкнутое множество.

3) Функция чувствительности является субдифференцируемой. Прокомментируем утверждение более детально. Для каждого $y \in Y = \{y \mid g_1(y) \leq y_1\}$, $y_1 \in \mathbb{R}^q_+$, функция чувствительности тесно связана с задачей выпуклого программирования (1.1), которая в свою очередь тесно связана с функцией Лагранжа

$$L(x,p) = f(x) + \langle p, g(x) - y\rangle, \tag{1.3}$$

определенной для всех $x \in X$, $p \in \mathbb{R}^q_+$ и любого фиксированного $y \in \mathbb{R}^q_+$. Связь эта состоит в том, что прямое и двойственное решения регулярной задачи выпуклого программирования образуют седловую точку функции Лагранжа, т. е. седловая точка (x^y, p^y) по определению удовлетворяет системе неравенств

$$f(x^y) + \langle p, g(x^y) - y\rangle \leq f(x^y) + \langle p^y, g(x^y) - y\rangle \leq f(x) + \langle p^y, g(x) - y\rangle \tag{1.4}$$

для всех допустимых $x \in \mathbb{R}^n$, $p \in \mathbb{R}^q_+$ и фиксированного $y \geq 0$. Из правого неравенства системы (1.4) видно, что вторая компонента седловой точки (вместе с координатой -1) является нормалью $(-1, p^n)$ опорной функции относительно образа отображения прямых переменных $(f(x), g(x))$. Обозначим эту компоненту нормали через $\nabla\varphi(y) = -p^y$ и будем называть ее *субградиентом.* В общем случае пара $(-1, \nabla\varphi(y))$ порождает опорную плоскость в точке $(f(x^y), g(x^y) - y)$. Опорных плоскостей может быть много, как правило, это целый конус. Конус нормалей опорных плоскостей будем называть *субдифференциалом* и обозначать его через $\dfrac{\partial\varphi(y)}{\partial y}$, тогда отдельный элемент конуса — *субградиент* $\nabla\varphi(y)$. В частности, конус может содержать один элемент, который порождает касательную плоскость. Проведенные рассуждения верны для любого значения параметра $y \in \mathbb{R}^q_+$, и в частности для некоторого $y_0 \in \mathbb{R}^q_+$. С учетом введенных обозначений можно написать:

$$\nabla\varphi(y) \in \frac{\partial\varphi(y)}{\partial y}, \quad \nabla\varphi(y_0) \in \frac{\partial\varphi(y)}{\partial y}\Big|_{y=y_0}.$$

Анализируя систему (1.4), нетрудно получить известные неравенства выпуклости функции чувствительности

$$\langle\nabla\varphi(y_0), y - y_0\rangle \le \varphi(y) - \varphi(y_0) \le \nabla\langle\varphi(y), y - y_0\rangle \tag{1.5}$$

для всех $y \ge 0$ и $y_0 \ge 0$. Подробное и формальное изложение этого пункта представлено в [1].

4) Функция чувствительности $\varphi(y)$ выпукла в смысле неравенства Йенссена [2], т. е. удовлетворяет условиям

$$\varphi(\alpha y + (1-\alpha)y_0) \le \alpha\varphi(y) + (1-\alpha)\varphi(y_0), \quad y \ge 0, \quad y_0 \ge 0, \quad 0 \le \alpha \le 1. \tag{1.6}$$

5) Субдифференциал функции чувствительности является многозначным выпуклозначным замкнутым отображением [1].

6) Субдифференциал функции чувствительности есть ограниченное отображение.

7) Субдифференциал функции чувствительности — многозначное монотонное отображение

$$\Big\langle \frac{\partial\varphi(y)}{\partial y} - \frac{\partial\varphi(y_0)}{\partial y}, y - y_0 \Big\rangle \ge 0$$

для всех $y \ge 0$ и $y_0 \ge 0$. Это неравенство очевидно следует из (1.5).

Отметим также, что график функции чувствительности задачи выпуклого программирования и совокупность парето-оптимальных векторных оценок векторной функции $(f(x), g(x))$ задачи многокритериальной оптимизации совпадают для выпуклого случая как множества [1].

2. Минимизация функции чувствительности на геометрическом множестве

Функция чувствительности не задана в явном виде, но, тем не менее, для любого $y \ge 0$ всегда можно вычислить ее значение в любой точке и любой ее субградиент (т. е. вектор множителей Лагранжа задачи (1.1)). Это значит, что для решения задачи минимизации $\varphi(y)$ на множестве $Y \in \mathbb{R}^m_+$ можно сформулировать проксимальные и градиентные итеративные процессы и доказать их сходимость к решению задачи (1.2), т. е. вычислить значение функции и ее градиент.

Пусть точка $y^* \in Y$ является точкой минимума функции чувствительности $\varphi(y)$ на множестве $Y \subseteq \mathbb{R}^m_+$. Тогда $y^* \in Y$ — неподвижная точка экстремального (проксимального) однозначного отображения, т. е.

$$y^* = \arg\min\big\{1/2\,|y - y^*|^2 + \alpha\varphi(y) \mid y \in Y\big\}, \quad \alpha > 0. \tag{2.1}$$

Это уравнение одновременно является необходимым и достаточным условием для задачи оптимизации (1.2). Метод простой итерации (проксимальный метод) для решения этого уравнения представляется естественным [3; 4]:

$$y^{n+1} = \arg\min\big\{1/2\,|y - y^n|^2 + \alpha\varphi(y) \mid y \in Y\big\}. \tag{2.2}$$

Процесс является неявным аналогом градиентного (в случае дифференцируемости $\varphi(y)$) метода, и сходимость его к точке минимума почти очевидна. Точка y^{n+1} из (2.2) — единственная точка минимума и по определению удовлетворяет неравенству

$$1/2\,|y^{n+1} - y^n|^2 + \alpha\varphi(y^{n+1}) \le 1/2\,|y - y^n|^2 + \alpha\varphi(y)$$

для всех $y \in Y$. Однако рассматриваемая функция является сильно выпуклой, и, как показано в [5], для точки y^{n+1} выполняется более сильное неравенство, а именно:

$$1/2\,|y^{n+1}-y^n|^2+\alpha\varphi(y^{n+1}) \le 1/2\,|y-y^n|^2+\alpha\varphi(y)-1/2\,|y-y^{n+1}|^2 \tag{2.3}$$

для всех $y \in Y$. Аналогично, уравнение (2.1) можем записать в форме усиленного неравенства

$$1/2\,|y^*-y^*|^2+\alpha\varphi(y^*) \le 1/2\,|y-y^*|^2+\alpha\varphi(y)-1/2\,|y-y^*|^2 \tag{2.4}$$

для всех $y \in Y$. Приведем кратко доказательство теоремы, которая нам понадобится в дальнейшем.

Теорема 1. *Если множество решений задачи* (1.1), (1.2) *не пусто, целевая функция подчинена условию* (1.6), *множество* $Y \subseteq \mathbb{R}^m_+$ *выпукло, замкнуто, то процесс* (2.2) *при любом значении параметра* $\alpha > 0$ *сходится монотонно по норме к одному из решений задачи* (1.1), (1.2).

Доказательство. Неравенства (2.3) и (2.4) выполняются для всех $y \in Y$, поэтому можно принять в первом неравенстве $y = y^*$, во втором — $y = y^{n+1}$, сложить оба неравенства и получить

$$|y^{n+1}-y^*|^2+|y^{n+1}-y^n|^2+2\alpha(\varphi(y^{n+1})-\varphi(y^*)) \le |y^n-y^*|^2. \tag{2.5}$$

Поскольку $(\varphi(y^{n+1})-\varphi(y^*)) \ge 0$, то отсюда следует ограниченность последовательности $|y^{n+1}-y^*|^2$ и ее монотонное по норме убывание. Просуммируем неравенство (2.5) от $n=0$ до $n=N$ и получим

$$|y^{N+1}-y^*|^2+\sum_{k=0}^{N}|y^{n+1}-y^n|^2+2\alpha\sum_{k=0}^{N}(\varphi(y^{n+1})-\varphi(y^*)) \le |y^0-y^*|^2.$$

Из этого неравенства следуют ограниченность последовательности $|y^{N+1}-y^*|^2 \le |y^0-y^*|^2$, сходимость ряда $\sum_{k=0}^{\infty}|y^{n+1}-y^n|^2 < \infty$ и соответственно стремление к нулю величины $|y^{n+1}-y^n| \to 0,\ n \to \infty$. Так как последовательность y^n ограничена, то существуют подпоследовательность y^{n_i} и элемент $y' \in Y$ такие, что $y^{n_i} \to y'$ при $n_i \to \infty$, причем $|y^{n_i+1}-y^{n_i}|^2 \to 0$.

Рассмотрим неравенство (2.3) на элементах подпоследовательности $n_i \to \infty$ и, перейдя к пределу, выпишем предельное неравенство $\varphi(y') \le \varphi(y)$ для всех $y \in Y$. Отсюда следует, что $y' = y^* \in Y$. Таким образом, любая предельная точка последовательности y^n является решением задачи, при этом величина $|y^n-y^*|$ монотонно убывает. В совокупности данные два фактора означают, что исходная последовательность имеет только одну предельную точку, т. е. y^n монотонно по норме сходится к одному из решений $y^* \in Y$ задачи.

Теорема доказана.

Для численной реализации процесса (2.2) требуется явное задание функции чувствительности. Однако в нашей ситуации эта функция задана неявно, хотя в каждой точке $y \in Y$ всегда можно вычислить ее значение и любой ее субградиент. Воспользуемся последним обстоятельством и выпишем для каждой итерации (2.2) и уравнения (2.1) необходимые и достаточные условия минимумов:

$$\langle y^{n+1}-y^n+\alpha\nabla\varphi(y^{n+1}),\, y-y^{n+1}\rangle \ge 0, \quad y \in Y; \tag{2.6}$$

$$\langle \nabla\varphi(y^*),\, y-y^*\rangle \ge 0, \quad y \in Y. \tag{2.7}$$

Полученная итеративная последовательность вариационных неравенств и предельное неравенство эквивалентны проксимальному процессу (2.2), (2.1). Эквивалентность понимается в том смысле, что они порождают одну и ту же последовательность решений, которая сходится к одному и тому же решению $y^* \in Y$. Поэтому теорема о сходимости в равной мере относится к итеративной системе вариационных неравенств и ее предельному случаю (2.6), (2.7).

При практическом применении подхода, основанного на использовании вариационных неравенств (2.6), конечно, нужно записать процесс в форме

$$\left\langle y^{n+1} - y^n + \alpha \frac{\partial \varphi(y^{n+1})}{\partial y}, y - y^{n+1} \right\rangle \geq 0, \quad y \in Y, \tag{2.8}$$

где $\dfrac{\partial \varphi(y)}{\partial y}$ — многозначное выпуклозначное ограниченное монотонное отображение. В рассматриваемом случае оператор вариационного неравенства регуляризован прибавлением к нему единичного оператора, тогда обратный оператор к вариационному неравенству в целом будет однозначным. Другими словами, любое вариационное неравенство (2.8) имеет единственное решение при любом n. Это хорошо согласуется с (2.2) (методы решения вариационных неравенств и другие детали см. в [6]).

3. Минимизация функции чувствительности на множестве функциональных ограничений

В этом разделе вместо задачи (1.1), (1.2) рассмотрим задачу с функциональными ограничениями [7]:

$$\varphi(y) = f(x^*) = \mathrm{Min}\{f(x) \mid g(x) \leq y,\ x \in X \subseteq \mathbb{R}^n\}, \quad y \in \mathbb{R}^m_+, \tag{3.1}$$

$$y^* \in \mathrm{Arg\,min}\{\varphi(y) \mid y \in Y\}, \tag{3.2}$$

где $y \in Y = \{y \mid g_1(y) \leq y_1 \in \mathbb{R}^m_+\}$. Целевая функция (3.1) характеризует уровень чувствительности задачи к изменению значения ее целевой функции при сдвиге i-го ограничения в направлении его градиента. Если оптимум $x^* \in X$ задачи (3.1) или (1.1) сильно лимитируется i-м ограничением, то i-й множитель Лагранжа p_i^* достаточно велик, а это значит, что целевая функция по этому направлению будет быстро меняться; при малом же множителе Лагранжа сдвиг ограничения в направлении его i-го градиента оказывает малое возмущение на задачу. Если же точка $x^* \in X$ лежит строго “внутри” ограничения, то i-й множитель Лагранжа вообще равен нулю. Вектор $y \in \mathbb{R}^m_+$ в этой задаче имеет определенную смысловую и содержательную нагрузку как вектор ресурсов.

Рассмотрим эту ситуацию более подробно. Введем функцию Лагранжа для задачи (3.1)

$$x^* \in \mathrm{Arg\,min}\{f(x) \mid g(x) - y^* \leq 0,\ x \in X,\ y \in Y\}$$

при фиксированном векторе $y = y^*$:

$$L(x, y, p) = f(x) + \langle p, g(x) - y^* \rangle.$$

Эта функция определена для всех $x \in \mathbb{R}^n,\ p \in \mathbb{R}^m_+,\ y^* \in \mathbb{R}^m_+$. Седловая точка (x^*, p^*) функции удовлетворяет системе седловых неравенств

$$f(x^*) + \langle p, g(x^*) - y^* \rangle \leq f(x^*) + \langle p^*, g(x^*) - y^* \rangle \leq f(x) + \langle p^*, g(x) - y^* \rangle \tag{3.3}$$

для всех допустимых $x \in X$, $p \in \mathbb{R}^m_+$. Перепишем систему (3.3) в виде

$$x^* \in \mathrm{Arg\,min}\{f(x) + \langle p^*, g(x) - y^* \rangle \mid x \in X\}, \quad y^* \in Y,$$

$$\langle p - p^*, g(x^*) - y^* \rangle \leq 0, \quad p \geq 0$$

и добавим к ней необходимое и достаточное условие минимума задачи (3.2)

$$\langle \nabla \varphi(y^*), y - y^* \rangle \geq 0, \quad y \in Y,$$

что эквивалентно $\varphi(y^*) \leq \varphi(y),\ y \in Y$.

Согласно (1.4) в п. 3) имеем $\nabla\varphi(y^*) = -p^*$, тогда последнюю систему неравенств можно представить как

$$\begin{gathered} x^* \in \operatorname{Arg}\min\{f(x) + \langle p^*, g(x) - y^*\rangle \mid x \in X\}, \quad y^* \in Y, \\ \langle p - p^*, g(x^*) - y^*\rangle \le 0, \quad p \ge 0, \\ \langle p^*, y - y^*\rangle \le 0, \quad y \in Y. \end{gathered} \tag{3.4}$$

Отсюда следует, что решение системы (3.4) удовлетворяет системе (3.1), (3.2) и наоборот. В свою очередь тройка векторов $x^*,\ p^*,\ y^*$ является решением системы экстремальных задач:

$$x^* \in \operatorname{Arg}\min\{f(x) \mid g(x) \le y^*,\ x \in X\}, \tag{3.5}$$

$$p^* \in \operatorname{Arg}\max\{\langle p, g(x^*) - y^*\rangle \mid p \ge 0\}, \tag{3.6}$$

$$y^* \in \operatorname{Arg}\min\{\langle \nabla\varphi(y^*), y\rangle \mid y \in Y\}. \tag{3.7}$$

Теперь можно сформулировать

Утверждение. *Система соотношений* (3.4) *является необходимым и достаточным условием оптимальности для системы задач* (3.5)–(3.7).

4. Прямой и двойственный экстрапроксимальные методы

Седловую структуру задачи (3.5)–(3.7) представим в форме

$$x^* \in \operatorname{Arg}\min\{f(x) + \langle p^*, g(x) - y^*)\rangle \mid x \in X\},$$

$$\begin{gathered} p^* = \pi_+(p^* + \alpha(g(x^*) - y^*)), \\ y^* = \pi_Y(y^* + \alpha p^*). \end{gathered}$$

Чтобы придать экстремальному отображению свойство “быть нерасширяющимся оператором” в области его определения, разумно его регуляризовать и записать в эквивалентной форме проксимального оператора; тогда система приобретает вид:

$$\begin{gathered} x^* \in \operatorname{Arg}\min\big\{1/2\,|x - x^*|^2 + \alpha(f(x) + \langle p^*, g(x) - y^*)\rangle) \mid x \in X\big\}, \\ p^* = \pi_+(p^* + \alpha(g(x^*) - y^*)), \\ y^* = \pi_Y(y^* + \alpha p^*). \end{gathered} \tag{4.1}$$

Полученная система уравнений относится к системам седлового типа, поэтому для ее решения естественно использовать седловые подходы. В работах [7; 8] седловые подходы применялись для решения разного типа задач и назывались *экстрапроксимальными методами* (см. также [9; 10]).

Здесь представляется разумным сопоставить постановку задачи, упомянутой в работе [7], с рассматриваемой здесь задачей (3.1), (3.2). В указанной работе рассматривается задача вычисления неподвижной точки некоторого отображения множества $Y \subset \mathbb{R}^q$ в себя. Эта постановка в неявном виде включена в задачу (3.1), (3.2), поэтому последнюю постановку можно рассматривать как обобщение первой. Содержательная интерпретация у этих задач может быть разная, но внутренняя логика одна и та же, поэтому методы их решения совпадают.

Что касается различия между прямыми и двойственными методами, то это различие в первую очередь связано со структурой векторных полей седловых задач. Эти векторные поля являются полями вращения с центром вращения в седловой точке. Такая структура векторного поля порождает замкнутые траектории. Чтобы обеспечить движение вычислительного процесса из произвольной точки к центру вращения, необходимо на каждой итерации “перескакивать” с одной замкнутой кривой на другую, которая ближе к центру вращения. Это приводит к расщеплению итеративного шага на два полушага. В зависимости от того, по каким переменным (прямым или двойственным) происходит расщепление итеративного шага, получаем прямой или двойственный методы.

Задача минимизации неявно заданной функции чувствительности при ограничениях представляет собой очень эластичную конструкцию, которая легко сочетается с другими задачами, включая игровые, сетевые задачи, вариационные неравенства, различные экономические рынки, а также динамические управляемые системы. Перечисленные связки дают возможность строить сложные сети задач, которые в свою очередь позволят создавать сложные математические модели для описания сложных равновесных состояний больших прикладных систем.

В настоящей работе рассмотрим два варианта (прямой и двойственный) экстрапроксимальных методов для решения системы (3.5)–(3.7). Первый из них будем называть прямой метод, а второй — соответственно двойственный метод.

Прямой метод:

$$\begin{gathered}\bar{y}^n = \pi_Y(y^n + \alpha p^n),\\ \bar{x}^n \in \arg\min\big\{1/2\,|x - x^n|^2 + \alpha(f(x) + \langle p^n, g(x)\rangle) \mid x \in X\big\};\\ p^{n+1} = \pi_+(p^n + \alpha(g(\bar{x}^n) - \bar{y}^n)),\\ y^{n+1} = \pi_Y(y^n + \alpha p^{n+1}),\\ x^{n+1} \in \arg\min\big\{1/2\,|x - x^n|^2 + \alpha(f(x) + \langle p^{n+1}, g(x)\rangle) \mid x \in X\big\}.\end{gathered} \tag{4.2}$$

Поскольку целевые функции процесса (4.2) имеют структуру функции (2.3), то этот процесс может быть записан в эквивалентной форме:

$$|\bar{x}^n - x^n|^2 + 2\alpha f(\bar{x}^n) + 2\alpha\langle p^n, g(\bar{x}^n)\rangle$$
$$\le |x - x^n|^2 + 2\alpha f(x) + 2\alpha\langle p^n, g(x)\rangle - |x - \bar{x}^n|^2; \tag{4.3}$$
$$|x^{n+1} - x^n|^2 + 2\alpha f(x^{n+1}) + 2\alpha\langle p^{n+1}, g(x^{n+1})\rangle$$
$$\le |x - x^n|^2 + 2\alpha f(x) + 2\alpha\langle p^{n+1}, g(x)\rangle - |x - x^{n+1}|^2. \tag{4.4}$$

Операторные уравнения процесса (4.3) согласно [5] представим в форме вариационных неравенств

$$\langle \bar{y}^n - y^n - \alpha p^n, y - \bar{y}^n\rangle \ge 0, \quad y \in Y,$$
$$\langle y^{n+1} - y^n - \alpha p^{n+1}, y - y^{n+1}\rangle \ge 0, \tag{4.5}$$
$$\langle p^{n+1} - p^n - \alpha(g(\bar{x}^n) - \bar{y}^n)), p - p^{n+1}\rangle \ge 0, \quad p \ge 0. \tag{4.6}$$

Для д о к а з а т е л ь с т в а сходимости рассматриваемого метода понадобится условие Липшица для векторной функции $g(x)$, которое используем в форме

$$|g(x + h) - g(h)| \le |g|\,|h| \tag{4.7}$$

для всех $x + h \in X$, $h \in \mathbb{R}^n$, где $|g|$ — константа Липшица.

Оценим отклонение векторов $\bar{x}^n$ и x^{n+1} на каждом шаге процесса (4.3), (4.4). С этой целью положим в неравенствах (4.5) и (4.6) значения $x = x^{n+1}$ и $x = x^n$ соответственно; тогда

$$|\bar{x}^n - x^n|^2 + 2\alpha f(\bar{x}^n) + 2\alpha\langle p^n, g(\bar{x}^n)\rangle \le |x^{n+1} - x^n|^2 + 2\alpha f(x^{n+1}) + 2\alpha\langle p^n, g(x^{n+1})\rangle - |x^{n+1} - \bar{x}^n|^2,$$

$$|x^{n+1} - x^n|^2 + 2\alpha f(x^{n+1}) + 2\alpha\langle p^{n+1}, g(x^{n+1})\rangle \le |\bar{x}^n - x^n|^2 + 2\alpha f(\bar{x}^n) + 2\alpha\langle p^{n+1}, g(\bar{x}^n)\rangle - |\bar{x}^n - x^{n+1}|^2.$$

Сложим полученные неравенства:

$$|\bar{x}^n - x^{n+1}|^2 \le \alpha\langle p^{n+1} - p^n, g(\bar{x}^n) - g(x^{n+1}\rangle.$$

С учетом (4.7) окончательно получим

$$|\bar{x}^n - x^{n+1}| \le \alpha\, |g|\, |p^{n+1} - p^n|. \tag{4.8}$$

Оценим отклонение векторов $\bar{y}^n$ и y^{n+1} на каждом шаге процесса (4.3):

$$|\bar{y}^n - y^{n+1}| \le \alpha\, |p^{n+1} - p^n|.$$

Сходимость итеративного процесса (4.2) устанавливает следующая теорема.

Теорема 2. *Если решение равновесной задачи* (3.1), (3.2) *существует, функции* $f(x)$ *и* $g(x)$ *— выпуклые, функция* $g(x)$ *дополнительно подчинена условию Липшица* (4.8), X, Y *— выпуклые замкнутые множества, то последовательность* p^n, x^n, y^n *прямого экстрапроксимального метода* (4.2) *с параметром* α*, удовлетворяющим условию* $0 < \alpha < 1/\sqrt{2(|g|^2+1)}$*, сходится монотонно по норме к одному из решений задачи* (4.1)*, т. е.* $p^n, x^n, y^n \to p^*, x^*, y^*$ *при* $n \to \infty$ *для всех* $p^0,\ x^0,\ p^0$*.*

Схема доказательства метода близка к доказательству, представленному в работе [7]. Доказательство легко может быть восстановлено читателем, поэтому здесь не приводится.

Наряду с прямым методом, который только что был рассмотрен, для решения системы задач (3.5)–(3.7) можно использовать двойственный экстрапроксимальный подход [7; 8]. Формулы этого метода имеют следующий вид.

Двойственный метод:

$$\begin{gathered}
\bar{p}^n = \pi_+(p^n + \alpha(g(x^n) - y^n));\\
y^{n+1} = \pi_Y(y^n + \alpha \bar{p}^n),\\
x^{n+1} \in \arg\min\bigl\{1/2\, |x - x^n|^2 + \alpha(f(x) + \langle \bar{p}^n, g(x)\rangle) \mid x \in X\bigr\},\\
p^{n+1} = \pi_+(p^n + \alpha(g(x^{n+1}) - y^{n+1})).
\end{gathered} \tag{4.9}$$

Если в исходной задаче (3.1), (3.2) вектор $y^* \in Y$ — константа, т. е. множество Y содержит одну точку, то итеративные формулы по переменой y отсутствуют. В этом случае получаютя формулы процесса для решения задачи выпуклого программирования (3.1) или вычисления седловой точки функции (1.3). Если же, наоборот, задача выпуклого программирования вырождается и отсутствует, то процесс (4.9) содержит формулы только по переменной y и этот подпроцесс сходится к решению задачи (3.2), т. е. к вычислению граничной точки множества Y, которая является опорной для линейного функционала $\langle p^*, y\rangle,\ y \in Y$, где p^* — априори заданный вектор.

Представим процесс (4.9) в форме неравенств:

$$\begin{gathered}
\langle \bar{p}^n - p^n - \alpha(g(x^n) - y^n)), p - \bar{p}^n\rangle \ge 0, \quad p \ge 0,\\
\langle y^{n+1} - y^n - \alpha \bar{p}^n, y - y^{n+1}\rangle \ge 0,\\
|x^{n+1} - x^n|^2 + 2\alpha f(x^{n+1}) + 2\alpha \langle \bar{p}^n, g(x^{n+1})\rangle\\
\le |x - x^n|^2 + 2\alpha f(x) + 2\alpha\langle \bar{p}^n, g(x)\rangle - |x - x^{n+1}|^2, \quad x \in X,\\
\langle p^{n+1} - p^n - \alpha(g(x^{n+1}) - y^{n+1})), p - p^{n+1}\rangle \ge 0, \quad p \ge 0.
\end{gathered}$$

Получим оценки отклонения векторов $\bar{p}^n$ и p^{n+1} друг от друга. Сопоставляя первое и последнее уравнения из (4.9), имеем $|\bar{p}^n - p^{n+1}| = \alpha\, |g(x^n) - y^n - g(x^{n+1}) + y^{n+1}|$.

Относительно сходимости метода (4.9) справедлива теорема.

Теорема 3. *Если решение равновесной задачи* (3.1), (3.2) *существует, функции* $f(x)$ *и* $g(x)$ *выпуклы, функция* $g(x)$ *подчинена условию Липшица* (4.8)*,* X, Y *— выпуклые замкнутые множества, то последовательность* p^n, x^n, y^n *двойственного экстрапроксимального метода* (4.9) *с параметром* α*, удовлетворяющим условию* $0 < \alpha < \min\{1/(2|g|), 1/2\}$*, сходится монотонно по норме к одному из решений задачи* (4.9)*, т. е.* $p^n, x^n, y^n \to p^*, x^*, y^*$ *при* $n \to \infty$ *для всех* $p^0,\ x^0,\ p^0$*.*

Теорема также приводится без доказательства (см. [7]).

5. Заключение

В работе рассматривалась задача минимизации функции чувствительности на выпуклом замкнутом множестве. Эта задача, несмотря на свою традиционную постановку, носит оригинальный характер в силу того обстоятельства, что представляет собой систему двух задач оптимизации. Целевая функция задачи задана неявно, и эта функция как функция чувствительности позволяет описывать и оценивать изменение оптимального значения целевой функции при сдвиге ограничений в направлении своего градиента. Кроме того, данная задача часто является самостоятельной компонентой сложных систем задач оптимизации. Например, в настоящей работе рассматривался случай, когда функция чувствительности явилась результатом задачи первого уровня. В свою очередь функция чувствительности, рассматриваемая на параметрическом множестве, порождает функцию чувствительности второго уровня. Такой процесс создает связанную систему из конечного числа задач, каждая из которых обусловлена своими функциями чувствительности. Развитие методов решения такого сорта сложных систем имеет особое значение для приложений, поскольку такие системы описывают равновесные состояния реальных сложных объектов.

СПИСОК ЛИТЕРАТУРЫ

1. **Антипин А. С., Голиков А. И., Хорошилова Е. В.** Функция чувствительности, ее свойства и приложения // Журн. вычисл. математики и мат. физики. 2011. Т. 51. № 12. С. 2126–2142.
2. **Васильев Ф. П.** Методы оптимизации. Т. 1,2. М.: Изд-во МЦНМО, 2011. 1056 р.
3. **Rockafellar R. T.** Monotone operators and the proximal point algorithm // SIAM J. Control Optim. 1976. Vol. 14, № 5. С. 877–898. doi: 10.1137/0314056 .
4. **Антипин А. С.** О методе выпуклого программирования, использующем симметрическую модификацию функции Лагранжа // Экономика и мат. методы. 1976. XII, вып. 6. С. 1164–1173.
5. **Антипин А. С.** О сходимости и оценках скорости сходимости проксимальных методов к неподвижным точкам экстремальных отображений // Журн. вычисл. математики и мат. физики. 1995. Т. 35, № 5. С. 688–704.
6. **Коннов И. В.** Нелинейная оптимизация и вариационные неравенства. Казань: Изд-во Казан. ун-та, 2013. 508 р.
7. **Антипин А. С.** Седловая задача и задача оптимизации как единая система // Тр. Ин-та математики и механики УрО РАН. 2008. Т. 14, № 2. С. 5–15.
8. **Антипин А. С.** Экстрапроксимальный метод решения равновесных и игровых задач со связанными переменными // Журн. вычисл. математики и мат. физики. 2005. Т. 45, № 12. С. 2102–2111.
9. **Clempner J., Poznyak A. S.** Using the extraproximal method for computing the shotest-path mixeв Lyapunov equilibrium in stackelberg security games // Inernational J. on Artificial Intelligence Tools. 2014. Vol. 11. P. 3–23.
10. **Trejo K. K., Clempner J. B., Poznyak A. S.** A stackelberg security game with random strategic based on the extraproximal theoretic approach // Engineering Applications of Artificial Intelligence. 2015. Vol. 37. P. 145–153. doi: 10.1016/j.engappai.2014.09.002 .

Антипин Анатолий Сергеевич
д-р физ.-мат. наук, профессор
главный науч. сотрудник
Федеральный исследовательский центр “Информатика и управление” РАН,
г. Москва
e-mail: asantip@yandex.ru

Поступила 13.06.2017

REFERENCES

1. Antipin A.S., Golikov A.I., Khoroshilova E.V. Sensitivity function: Properties and applications. *Comput. Math. Math. Phys.*, 2011, vol. 51, no. 12, pp. 2000–2016. doi: 10.1134/S0965542511120049 .
2. Vasil’ev F.P. *Metody optimizatsii:* T. 1,2. [Optimization methods: Vol. 1,2]. Moscow, MTsNMO Publ., 2011, 1056 p.

3. Rockafellar R.T. Monotone operators and the proximal point algorithm. *SIAM J. Control and Optimization*, 1976, vol. 14, no. 5, pp. 877–898. doi: 10.1137/0314056 .
4. Antipin A.C. On the method of convex programming using the symmetric modification of the Lagrange function. *Economy and Mat. Methods*, 1976, vol. 12, no. 6, pp. 1164–1173 (in Russian).
5. Antipin A.S. The convergence of proximal methods to fixed points of extremal mappings and estimates of their rate of convergence. *Comput. Math. Math. Phys.*, 1995, vol. 35, no. 5, pp. 539–551.
6. Konnov I.V. *Nelineinaya optimizatsiya i variatsionnye neravenstva* [Nonlinear optimization and variational inequalities]. Kazan: Kazan Federal University Publ., 2013, 508 p. ISBN: 978-5-00019-059-3 .
7. Antipin A.S. Saddle problem and optimization problem as an integrated system. *Proc. Steklov Inst. Math.* (*Suppl.*), 2008, vol. 263 (suppl. 2), pp. 3–14. doi: 10.1134/S0081543808060023 .
8. Antipin A.S. An extraproximal method for solving equilibrium programming problems and games with coupled variables. *Comput. Math. Math. Phys.*, 2005, vol. 45, no. 12, pp. 2020–2029.
9. Clempner J., Poznyak A.S. Using the extraproximal method for computing the shortest-path mixed Lyapunov equilibrium in Stackelberg security games. *Mathematics and Computers in Simulation*, 2017, vol. 138, pp. 14–30. doi: 10.1016/j.matcom.2016.12.010 .
10. Trejo K.K., Clempner J.B., Poznyak A.S. A stackelberg security game with random strategic based on the extraproximal theoretic approach. *Engineering Applications of Artificial Intelligence*, 2015, vol. 37, pp. 145–153. doi: 10.1016/j.engappai.2014.09.002 .

The paper was received by the Editorial Office on June 13, 2017.

Anatolii Sergeevich Antipin, Dr. Phys.-Math. Sci., Prof., Federal Research Center “Computer Science and Control” of Russian Academy of Science, Moscow, 119333 Russian, e-mail: asantip@yandex.ru .