ТРУДЫ ИНСТИТУТА МАТЕМАТИКИ И МЕХАНИКИ УрО РАН

Том 22 № 2 2016

УДК 517.977

ЭКСТРАГРАДИЕНТНЫЙ МЕТОД ПОИСКА ТОЧКИ РАВНОВЕСИЯ В МНОГОКРИТЕРИАЛЬНОЙ ЗАДАЧЕ С ДИНАМИКОЙ[1]

Ф. П. Васильев, А. С. Антипин, Л. А. Артемьева

Рассматривается задача оптимального управления для линейной системы обыкновенных дифференциальных уравнений с граничным условием, заданным неявно и связанным с многокритериальной задачей. Такие задачи возникают, например, при изучении управляемых объектов, когда под влиянием внешних возмущений они теряют устойчивость и требуется вернуть их в состояние равновесия с помощью подходящего выбора управления. В статье приводится описание одной из возможных математических моделей такого рода, предлагается экстраградиентный метод восстановления равновесия, исследуется его сходимость.

Ключевые слова: задача оптимального управления, задача Коши, многокритериальная задача, седловая точка, сходимость.

F. P. Vasil'ev, A. S. Antipin, L. A. Artem'eva. Extragradient method for finding a saddle point in a multicriteria problem with dynamics.

We consider an optimal control problem for a linear system of ordinary differential equations with an implicitly given boundary condition connected with a multicriteria problem. Such problems arise, for example, in the study of controlled objects that lose their stability under the influence of external perturbations, where it is required to return the object to stability by means of an appropriate choice of the control. We describe a possible mathematical model of this kind, propose an extragradient method for recovering the stability, and investigate its convergence.

Keywords: optimal control problem, Cauchy problem, multicriteria problem, saddle point, convergence.

MSC: 37N40, 49J15, 49M30, 93C15, 90C29

DOI: 10.21538/0134-4889-2016-22-2-71-78

1. Постановка задачи

Пусть E^n — евклидово пространство размерности n; $\langle a,b\rangle = \sum_{i=1}^n a^i b^i$ — скалярное произведение векторов $a=(a^1,\dots,a^n)$, $b=(b^1,\dots,b^n)$; $\|a\| = \left(\sum_{i=1}^n (a^i)^2\right)^{1/2}$ — норма вектора a; $E^n_+ = \{a\in E^n\colon a^i\geqslant 0,\ i=1,\dots,n\}$ — неотрицательный ортант в E^n; $L_2^r[t_0,t_1]$ — Лебегово пространство вектор-функций $u=u(t)=(u^1(t),\dots,u^r(t))$, $t_0\leqslant t\leqslant t_1$, с координатами $u^i(t)$, интегрируемыми по Лебегу вместе со своим квадратом; $\langle u,v\rangle = \displaystyle\int_{t_0}^{t_1}\sum_{i=1}^n u^i(t)v^i(t)dt$ — скалярное произведение; $\|u(t)\| = \left(\displaystyle\int_{t_0}^{t_1}\sum_{i=1}^n (u^i(t))^2 dt\right)^{1/2}$ — норма в $L_2^r[t_0,t_1]$.

Пусть управляемый процесс описывается следующей задачей Коши:

$$\dot x(t) = D(t)x(t) + B(t)u(t) + g(t),\quad t_0\leqslant t\leqslant t_1,\quad x(t_0)=x_0; \tag{1.1}$$

где $D(t)$, $B(t)$, $g(t)$ — матрицы размера $n\times n$, $n\times r$, $n\times 1$ соответственно с кусочно-непрерывными элементами, t_0, t_1 — заданные моменты времени, $x_0\in E^n$ — заданная точка, $u=u(t)\in L_2^r[t_0,t_1]$ — управление, $x=x(t;u)=(x^1(t),\dots,x^n(t))$, $t_0\leqslant t\leqslant t_1$, — траектория системы (1.1),

[1]Работа выполнена при поддержке РФФИ (проект 15-01-06045-а) и Министерства образования и науки РФ по программе повышения конкурентоспособности НИТУ "МИСиС" среди ведущих мировых научно-образовательных центров (соглашение №02.A03.21.0004).

соответствующая управлению $u(t)$. Как известно (см., например, [1, кн. 1, гл. 6, § 1, теорема 2]), для каждого управления $u = u(t) \in L_2^r[t_0, t_1]$ задача (1.1) имеет единственное решение $x = x(t; u)$, являющееся абсолютно непрерывной функцией на отрезке $[t_0, t_1]$, почти всюду удовлетворяющее уравнению (1.1).

Будем говорить, что управление $u = u(t)$ *переводит точку* x_0 *в точку* x_1 в момент времени t_1, если $x(t_1; u) = x_1$. Пусть U — выпуклое замкнутое множество из $L_2^r[t_0, t_1]$. Обозначим $X = \{x \in E^n \colon x = x(t_1; u),\ u \in U\}$ — *множество достижимости* системы (1.1) в момент t_1 для управлений $u = u(t)$ из множества U. Заметим, что множество X является выпуклым и замкнутым.

Далее будем рассматривать следующую задачу терминального оптимального управления — найти управление $u \in U$ и вектор $\lambda \in E_+^m$ из условий

$$\langle \lambda, f(x(t_1; u)) \rangle \to \inf, \tag{1.2}$$

$$\langle \mu - \lambda, f(x(t_1; u)) - \lambda \rangle \leqslant 0 \quad \forall\, \mu \in E_+^m; \tag{1.3}$$

где $f(x) = (f^1(x), \ldots, f^m(x))$, $x \in E^n$ — заданная вектор-функция с выпуклыми, дифференцируемыми координатами $f^i(x)$, $i = 1, \ldots, m$.

Следуя работе [2], кратко поясним прикладной смысл задачи (1.1)–(1.3). Пусть группа из m участников образует сообщество (экономический союз) для реализации общей цели (проекта). Цели и интересы каждого из участников описываются стоимостными целевыми функциями $f^i(x)$, $i = 1, \ldots, m$, которые определены на множестве X. Проблема, стоящая перед участниками, заключается в том, как выбрать распределение поставок ресурсов так, чтобы, с одной стороны, проект был реализован с наименьшими суммарными расходами, а, с другой стороны, каждый из участников минимизировал свой вклад в реализацию общего проекта. Предполагается, что участники проекта договорились в качестве своего вклада в проект выбирать точки Парето множества $\{f(x),\ x \in X\}$. Как известно [3], если функции $f^i(x)$, $i = 1, \ldots, m$, выпуклы на X, то точки Парето являются точками минимума функции $\langle \lambda, f(x) \rangle$ на множестве X при условии $\lambda \in E_+^m,\ \sum_{i=1}^m \lambda_i = 1$. Так мы и приходим к задаче (1.2).

Заметим, что задача (1.2), как правило, порождает обширное множество точек Парето. Для разных Парето оптимальных точек, которые задают конфигурацию параметров будущего проекта, его стоимость, вообще говоря, разная. Естественно выбрать параметры, при которых проект имеет минимальную стоимость. Такой выбор можно реализовать, рассматривая игру двух лиц с равновесием по Нэшу [4]. Здесь одно лицо — это группа участников проекта, второе лицо — игрок, распоряжающийся параметрами (ценами) $\lambda \in E_+^m$ и учитывающий групповые интересы, что отражено в (1.3).

В сформулированной задаче (1.1)–(1.3) наличие управляемой динамической системы (1.1) позволяет удерживать модель вблизи состояния равновесия, стабилизируя ситуацию. Реальный прообраз данной задачи можно найти во многих сферах человеческой деятельности.

Задачу (1.1)–(1.3) удобно представить в седловой форме. Введем функцию, аналогичную функции Лагранжа:

$$\mathcal{L}(u, \lambda) = \langle \lambda, f(x(t_1; u)) - \lambda/2 \rangle, \quad u \in U, \quad \lambda \in E_+^m. \tag{1.4}$$

Седловой точкой этой функции будем называть точку $(u_*, \lambda_*) \in U \times E_+^m$, удовлетворяющую неравенствам

$$\mathcal{L}(u_*, \lambda) \leqslant \mathcal{L}(u_*, \lambda_*) \leqslant \mathcal{L}(u, \lambda_*) \quad \forall\, u \in U, \quad \lambda \in E_+^m. \tag{1.5}$$

Убедимся, что точка $(u_*, \lambda_*) \in U \times E_+^m$ будет решением задачи (1.1)–(1.3) тогда и только тогда, когда она является седловой точкой функции (1.4). В самом деле, правое неравенство (1.5) с учетом определения функции (1.4) означает, что

$$\langle \lambda_*, f(x(t_1; u_*)) \rangle \leqslant \langle \lambda_*, f(x(t_1; u)) \rangle \quad \forall\, u \in U;$$

это равносильно (1.2) при $\lambda = \lambda_*$. Из левого неравенства (1.5) следует, что вогнутая функция $\mathcal{L}(u_*, \lambda)$ достигает максимума при $\lambda = \lambda_*$; это равносильно неравенству [1, кн. 1, гл. 4, § 2, теорема 3]

$$\Big\langle \frac{\partial \mathcal{L}(u_*, \lambda_*)}{\partial \lambda}, \mu - \lambda_* \Big\rangle = \langle f(x(t_1; u_*)) - \lambda_*, \mu - \lambda_* \rangle \leqslant 0 \quad \forall\, \mu \in E^m_+,$$

совпадающему с (1.3) при $u = u_*$, $\lambda = \lambda_*$. Таким образом, чтобы получить решение задачи (1.1)–(1.3) достаточно найти седловую точку функции (1.4).

Заметим, что при сделанных выше предположениях функция $\mathcal{L}(u, \lambda)$ выпукла и слабо-полунепрерывна снизу по переменной u на выпуклом замкнутом множестве U при каждом фиксированном $\lambda \in E^m_+$, сильно вогнута и непрерывна по λ на E^m_+ при каждом $u \in U$.

2. Описание метода

Опишем экстраградиентный метод поиска седловой точки функции (1.4). Эта функция имеет производные по каждой из переменных u, λ, которые представимы в виде [1, кн. 2, гл. 8, § 3, пример 7]

$$\frac{\partial \mathcal{L}(u, \lambda)}{\partial u} = B^\top(t)\psi(t, u, \lambda), \quad t_0 \leqslant t \leqslant t_1, \tag{2.1}$$

где $\psi(t) = \psi(t, u, \lambda)$ — решение задачи Коши

$$\dot{\psi}(t) = -D^\top(t)\psi(t), \quad t_0 \leqslant t \leqslant t_1, \quad \psi(t_1) = \lambda^\top f'(x)\Big|_{x = x(t_1; u)}, \tag{2.2}$$

а $f'(x)$ — матрица Якоби вектор-функции $f(x) = (f^1(x), \ldots, f^m(x))$, $\top$ — знак транспонирования, и

$$\frac{\partial \mathcal{L}(u, \lambda)}{\partial \lambda} = f(x)|_{x = x(t_1; u)} - \lambda. \tag{2.3}$$

Пусть $(u_0, \lambda_0) \in U \times E^m_+$ — какое-либо начальное приближение. Пусть при некотором $k \geqslant 0$ известно k-е приближение $(u_k, \lambda_k) \in U \times E^m_+$. Тогда следующее приближение $(u_{k+1}, \lambda_{k+1}) \in U \times E^m_+$ определяем так: сначала находим вспомогательное прогнозное приближение $(\bar{u}_k, \bar{\lambda}_k)$:

$$\bar{u}_k = \pi_U\Big(u_k - \beta_k \frac{\partial \mathcal{L}(u_k, \lambda_k)}{\partial u}\Big), \quad \bar{\lambda}_k = \pi_+\Big(\lambda_k + \beta_k \frac{\partial \mathcal{L}(u_k, \lambda_k)}{\partial \lambda}\Big), \tag{2.4}$$

затем определяем основное приближение:

$$u_{k+1} = \pi_U\Big(u_k - \beta_k \frac{\partial \mathcal{L}(\bar{u}_k, \bar{\lambda}_k)}{\partial u}\Big), \quad \lambda_{k+1} = \pi_+\Big(\lambda_k + \beta_k \frac{\partial \mathcal{L}(\bar{u}_k, \bar{\lambda}_k)}{\partial \lambda}\Big); \tag{2.5}$$

где $\pi_U u$, $\pi_+\lambda$ — проекции точек $u \in L^r_2[t_0, T]$, $\lambda \in E^m$, на множества U, E^m_+ соответственно [1, кн. 1, гл. 4, § 4, с. 213], $\beta_k > 0$ — длина шага метода. Отметим, что другой, более сложный, вариант экстраградиентного метода для задачи (1.1)–(1.3), основанный на функции вида (1.4), учитывающей систему (1.1) явно, рассмотрен в [2].

Запишем метод (2.4), (2.5) в краткой форме. Обозначим

$$z = (u, \lambda), \quad \bar{z} = (\bar{u}, \bar{\lambda}), \quad Z = U \times E^m_+,$$
$$\frac{\partial \mathcal{L}(z)}{\partial z} = \Big(\frac{\partial \mathcal{L}(z)}{\partial u}, \frac{\partial \mathcal{L}(z)}{\partial \lambda}\Big), \quad Az = \Big(\frac{\partial \mathcal{L}(z)}{\partial u}, -\frac{\partial \mathcal{L}(z)}{\partial \lambda}\Big). \tag{2.6}$$

Учитывая, что проекция точки z на множество Z имеет вид $\pi_Z z = (\pi_U u, \pi_+ \lambda)$, метод (2.4), (2.5) можем кратко записать следующим образом:

$$\bar{z}_k = \pi_Z(z_k - \beta_k A z_k), \quad z_{k+1} = \pi_Z(z_k - \beta_k A \bar{z}_k), \quad k = 0, 1, \ldots . \tag{2.7}$$

3. Сходимость метода

Исследуем сходимость метода (2.7), предполагая, что функция Лагранжа (1.4) имеет хотя бы одну седловую точку $z_* = (u_*, \lambda_*)$. Пользуясь характеристическим свойством проекции (см. [1, кн. 1, гл. 4, § 4, теорема 1])

$$\langle \pi_Z z - z, y - \pi_Z z\rangle \geqslant 0 \quad \forall\, y \in Z,$$

запишем метод (2.7) в форме вариационных неравенств:

$$\langle \bar z_k - z_k + \beta_k A z_k, z - \bar z_k\rangle \geqslant 0 \quad \forall\, z \in Z, \tag{3.1}$$

$$\langle z_{k+1} - z_k + \beta_k A \bar z_k, z - z_{k+1}\rangle \geqslant 0 \quad \forall\, z \in Z, \quad k = 0, 1, \dots . \tag{3.2}$$

Возьмем произвольную седловую точку $z_* = (u_*, \lambda_*)$ функции (1.4). Положим в (3.1) $z = z_{k+1}$, в (3.2) $z = z_*$ и сложим получившиеся неравенства:

$$\begin{gathered}\langle \bar z_k - z_k, z_{k+1} - \bar z_k\rangle + \langle z_{k+1} - z_k, z_* - z_{k+1}\rangle \\ + \beta_k\left(\langle A z_k, z_{k+1} - \bar z_k\rangle + \langle A \bar z_k, z_* - z_{k+1}\rangle\right) \geqslant 0, \quad k = 0, 1, \dots .\end{gathered} \tag{3.3}$$

Первые два слагаемых из левой части неравенства (3.3) преобразуем, пользуясь тождеством

$$2\langle a - c, c - b\rangle = \|a - b\|^2 - \|a - c\|^2 - \|c - b\|^2 \quad \forall\, a, b, c \in L_2^r[t_0, t_1] \times E^m.$$

После несложных преобразований из (3.3) получим

$$\begin{gathered}\|z_{k+1} - z_*\|^2 + \|z_{k+1} - \bar z_k\|^2 + \|z_k - \bar z_k\|^2 \leqslant \|z_k - z_*\|^2 \\ + 2\beta_k\left(\langle A z_k - A \bar z_k, z_{k+1} - \bar z_k\rangle + \langle A \bar z_k - A z_*, z_* - \bar z_k\rangle + \langle A z_*, z_* - \bar z_k\rangle\right) \geqslant 0, \quad k = 0, 1, \dots .\end{gathered} \tag{3.4}$$

Так как оператор A, определенный согласно (2.6), является монотонным на Z, и z_* — седловая точка функции (1.4) [1, кн. 1, гл. 4, § 2, теоремы 3, 4], то справедливы неравенства

$$\langle A \bar z_k - A z_*, z_* - \bar z_k\rangle \leqslant 0, \quad \langle A z_*, z_* - \bar z_k\rangle \leqslant 0.$$

Отсюда и из (3.4) имеем

$$\|z_{k+1} - z_*\|^2 + \|z_{k+1} - \bar z_k\|^2 + \|z_k - \bar z_k\|^2 \leqslant \|z_k - z_*\|^2 + 2\beta_k\langle A z_k - A \bar z_k, z_{k+1} - \bar z_k\rangle, \quad k = 0, 1, \dots . \tag{3.5}$$

Для оценки последнего слагаемого из правой части неравенства (3.5) достаточно было бы потребовать выполнения условия Липшица для оператора A на множестве $Z = U \times E_+^m$. Как видно из (2.6), это возможно, если условию Липшица удовлетворяют производные $\partial\mathcal{L}(z)/\partial u$, $\partial\mathcal{L}(z)/\partial\lambda$. Посмотрим, насколько такое требование естественно в рассматриваемой задаче (1.1)–(1.3). Для траектории $x(t; u)$ задачи Коши (1.1) справедливы оценки [1, кн. 1, гл. 6, § 3, теорема 1]

$$\begin{gathered}\max_{t \in [t_0, t_1]} \|x(t; u_1) - x(t; u_2)\| \leqslant C\|u_1 - u_2\| \quad \forall\, u_1, u_2 \in L_2^r[t_0, t_1], \quad C = \mathrm{const} > 0, \\ \max_{[t \in t_0, t_1]} \|x(t; u)\| \leq C(1 + \|u\|)\ \forall u \in L_2^r[t_0, t_1].\end{gathered} \tag{3.6}$$

Далее будем предполагать, что вектор-функция $f(x)$ и матрица Якоби $f'(x)$ удовлетворяют условию Липшица на ограниченном множестве из E^n. Тогда из (2.3), (3.6) следует, что

$$\left\|\frac{\partial\mathcal{L}(u_1, \lambda_1)}{\partial\lambda} - \frac{\partial\mathcal{L}(u_2, \lambda_2)}{\partial\lambda}\right\| \leqslant L_0\|u_1 - u_2\| \quad \forall\, (u_1, \lambda_1), (u_2, \lambda_2) \in Z, \quad L = \mathrm{const} > 0. \tag{3.7}$$

Константы Липшица и другие константы, конкретные значения которых в наших теоретических рассуждениях не важны, ниже мы будем обозначать одними и теми же буквами C, L.

Выясним, когда производная $\partial\mathcal{L}(u,\lambda)/\partial u$ будет удовлетворять условию Липшица. С этой целью воспользуемся формулами (2.1), (2.2). Для решения $\psi(t)=\psi(t,u,\lambda)$ задачи (2.2) справедлива оценка [1, кн. 1, гл. 6, § 3, теорема 1]

$$\max_{t\in[t_0,t_1]}\|\psi(t,u_1,\lambda_1)-\psi(t,u_2,\lambda_2)\|\leqslant C\|\lambda_1\|\|f'(x(t_1;u_1))-f'(x(t_1;u_2))\|$$

$$+\|\lambda_1-\lambda_2\|\|f'(x(t_1;u_2))\| \;\;\forall\,(u_1,\lambda_1),\;(u_2,\lambda_2)\in Z. \tag{3.8}$$

Пусть $\sup_{u\in U}\|f'(x(t_1;u))\|<\infty$. Тогда из (3.8) и (2.1) следует

$$\left\|\frac{\partial\mathcal{L}(u_1,\lambda_1)}{\partial\lambda}-\frac{\partial\mathcal{L}(u_2,\lambda_2)}{\partial\lambda}\right\|\leqslant L\left(\|\lambda_1\|\|f'(x(t_1;u_1))-f'(x(t_1;u_2))\|+\|\lambda_1-\lambda_2\|\right). \tag{3.9}$$

Так как λ_1 принадлежит неограниченному множеству E^m_+, то условие Липшица для $\partial\mathcal{L}(u,\lambda)/\partial u$ на Z можно получить, если первое слагаемое в правой части неравенства (3.9) равно нулю. Это возможно, если $f^i(x)=\langle a^i,x\rangle-b^i$, $i=1,\dots,m$, — афинные функции. В общем случае из-за неограниченности множества E^m_+ выполнения условия Липшица на Z для оператора $\partial\mathcal{L}(u,\lambda)/\partial u$ ожидать не приходится.

Однако внимательный анализ экстраградиентного метода (2.7) [5] показывает, что при естественных ограничениях на задачу (1.1)–(1.3) последовательности $\{\bar z_k\}$, $\{z_k\}$, порождаемые этим методом, принадлежат некоторому ограниченному подмножеству S множества Z и можно говорить об условии Липшица для оператора A на S.

Покажем, что в качестве такого подмножества S в задаче (1.1)–(1.3) можно, например, взять

$$S=\left\{z=(u,\lambda)\in Z\colon\|z-z_*\|\leqslant 4\left(\|z_0\|+\|z_*\|\right)+\|Az_0\|\overset{def}{=}R\right\}, \tag{3.10}$$

где $z_0=(u_0,\lambda_0)\in Z$ — начальное приближение метода (2.7), z_* — произвольная седловая точка функции (1.4). При сделанных выше предположениях множество S (3.10) выпукло, замкнуто, ограничено в $L^r_2[t_0,t_1]\times E^m_+$. Отсюда и из (2.1), (3.6)–(3.8) следует, что оператор A на множестве S удовлетворяет условию Липшица

$$\|Ay_1-Ay_2\|\leqslant L\|y_1-y_2\|\quad\forall\,y_1,y_2\in S. \tag{3.11}$$

Выберем в методе (2.7) длину шага β_k из условий

$$0<\beta_{\min}\leqslant\beta_k\leqslant\min\left\{1;\frac{1}{2L}\right\},\quad k=0,1,\dots, \tag{3.12}$$

где постоянная L взята из (3.11). Покажем, что тогда последовательности $\bar z_k=(\bar u_k,\bar\lambda_k)$, $z_k=(u_k,\lambda_k)$, $k=0,1,\dots$, полученные методом (2.7), (3.12), таковы, что

$$z_k\in S,\quad\bar z_k\in S,\quad\|z_k-z_*\|\leqslant\|z_0-z_*\|,\quad k=0,1,\dots, \tag{3.13}$$

при любом выборе точки z_*. Воспользуемся методом математической индукции. Очевидно, $z_0\in S$. Кроме того, $\|z_*-z_0\|\leqslant\|z_*\|+\|z_0\|\leqslant R$, так что $z_*\in S$. Далее, учитывая, что оператор проектирования нерастягивающий [1, кн. 1, гл. 4, § 4, теорема 2], из первого равенства (2.7) и условия (3.12) при $k=0$ имеем

$$\|\bar z_0-z_*\|=\|\pi_Z(z_0-\beta_0Az_0)-\pi_Zz_0\|\leqslant\beta_0\|Az_0\|\leqslant R_0,$$

т. е. $\bar z_0\in S$. Заметим, что неравенство в (3.13) при $k=0$ превращается в равенство. Таким образом, при $k=0$ соотношения (3.13) выполняются. Предположим, что соотношения (3.13) выполнены при некотором $k \geqslant 0$. Тогда можем воспользоваться условием Липшица (3.11) при $y_1 = z_k \in S$, $y_2 = \bar{z}_k \in S$:

$$\|Az_k - A\bar{z}_k\| \leqslant L\|z_k - \bar{z}_k\|.$$

Поэтому для последнего слагаемого из правой части оценки (3.5) имеем

$$2\beta_k\langle Az_k - A\bar{z}_k, z_{k+1} - \bar{z}_k\rangle \leqslant 2\beta_k L\|z_k - \bar{z}_k\|\|z_{k+1} - \bar{z}_k\| \leqslant \beta_k L(\|z_k - \bar{z}_k\|^2 + \|z_{k+1} - \bar{z}_k\|^2),$$

и, продолжая оценку (3.5), получаем

$$\|z_{k+1} - z_*\|^2 + (1 - \beta_k L)(\|z_{k+1} - \bar{z}_k\|^2 + \|z_k - \bar{z}_k\|^2) \leqslant \|z_k - z_*\|^2.$$

В силу (3.12) $(1 - \beta_k L) \geqslant 1/2$. Отсюда и из (3.13) следует

$$\|z_{k+1} - z_*\|^2 \leqslant \|z_{k+1} - z_*\|^2 + \frac{1}{2}(\|z_{k+1} - \bar{z}_k\|^2 + \|z_k - \bar{z}_k\|^2) \leqslant \|z_k - z_*\|^2 \leqslant \|z_0 - z_*\|^2, \quad (3.14)$$

иначе говоря,

$$\|z_{k+1} - z_*\|^2 \leqslant \|z_0 - z_*\|^2.$$

Тогда

$$\|z_{k+1} - z_0\| \leqslant \|z_{k+1} - z_*\| + \|z_* - z_0\| \leqslant 2\|z_* - z_0\| \leqslant 2(\|z_*\| + \|z_0\|) \leqslant R, \quad (3.15)$$

т. е. $z_{k+1} \in S$. Покажем, что $\bar{z}_{k+1} \in S$. Из первого равенства (2.7) для номера $k+1$ имеем

$$\|\bar{z}_{k+1} - z_0\| = \|\pi_Z(z_{k+1} - \beta_{k+1}Az_{k+1}) - \pi_Z z_0\| \leqslant \|z_{k+1} - z_0\| + \beta_{k+1}\left(\|Az_{k+1} - Az_0\| + \|Az_0\|\right). \quad (3.16)$$

Так как $z_{k+1} \in S$, $z_0 \in S$, то можем воспользоваться условием Липшица (3.11) при $y_1 = z_{k+1}$, $y_2 = z_0$:

$$\|Az_{k+1} - Az_0\| \leqslant L\|z_{k+1} - z_0\|.$$

Отсюда и из (3.16) с учетом (3.12), (3.15) получаем

$$\|\bar{z}_{k+1} - z_0\| \leqslant (1 + \beta_{k+1}L)\|z_{k+1} - z_0\| + \beta_{k+1}\|Az_0\|$$

$$\leqslant 2\|z_{k+1} - z_0\| + \|Az_0\| \leqslant 4(\|z_*\| + \|z_0\|) + \|Az_0\| = R,$$

т. е. $\bar{z}_{k+1} \in S$. Индуктивные рассуждения закончены, соотношения (3.13) доказаны.

Продолжим доказательство сходимости метода (2.7), (3.12) к решению задачи (1.1)–(1.3). Суммируя (3.14) по k, от $k = 0$ до некоторого номера N, получим

$$\|z_{N+1} - z_*\|^2 + \frac{1}{2}\sum_{k=0}^{N}(\|z_{k+1} - \bar{z}_k\|^2 + \|z_k - \bar{z}_k\|^2) \leqslant \|z_0 - z_*\|^2, \quad N = 0, 1, \ldots. \quad (3.17)$$

Это означает, что ряды $\sum_{k=0}^{\infty}\|z_{k+1} - \bar{z}_k\|^2$, $\sum_{k=0}^{\infty}\|z_k - \bar{z}_k\|^2$ сходятся. Поэтому

$$\lim_{k\to\infty}\|z_{k+1} - \bar{z}_k\| = \lim_{k\to\infty}\|z_k - \bar{z}_k\| = \lim_{k\to\infty}\|z_{k+1} - z_k\| = 0. \quad (3.18)$$

Из ограниченности $\{z_k\}$, вытекающей из (3.13), (3.17), и из (3.18) следует существование подпоследовательностей $\{z_{k_l}\}$, $\{\bar{z}_{k_l}\}$, $\{z_{k_l+1}\}$, $\{\bar{z}_{k_l+1}\}$, сходящихся к некоторой точке $z' = (u', \lambda') \in S$ слабо в $L_2^r[t_0, T] \times E^m$. Это означает, что u_{k_l}, $\bar{u}_{k_l}$, $\bar{u}_{k_l+1}$ сходятся к точке $u' \in U$ слабо в $L_2^r[t_0, t_1]$, $\{\lambda_{k_l}\}$, $\{\bar{\lambda}_{k_l}\}$, $\{\bar{\lambda}_{k_l+1}\}$ сходятся к $\lambda' \in E_+^m$ в норме E^m. Несложно доказать (см. [1, кн. 2, гл. 10, § 1, с. 910]), что тогда траектории $x(t; u_{k_l})$, $\psi(t, u_{k_l}, \lambda_{k_l})$ задач (1.1), (2.2) сходятся равномерно к траекториям, соответствующим $(u', \lambda') = z'$. В частности, из формул (2.1), (2.3) тогда следует, что при $k = k_l \to \infty$

$$\frac{\partial\mathcal{L}(u_{k_l}, \lambda_{k_l})}{\partial u} \to \frac{\partial\mathcal{L}(u', \lambda')}{\partial u}, \quad \frac{\partial\mathcal{L}(u_{k_l}, \lambda_{k_l})}{\partial \lambda} \to \frac{\partial\mathcal{L}(u', \lambda')}{\partial \lambda},$$

или, иначе говоря,

$$Az_{k_l} \to Az'$$

равномерно на отрезке $[t_0, t_1]$. Кроме того,

$$|\langle \bar{z}_{k_l} - z_{k_l}, z - z_{k_l}\rangle| \leqslant \|\bar{z}_{k_l} - z_{k_l}\|\|z - z_{k_l}\| \to 0, \quad k = k_l \to \infty,$$

в силу (3.18) и ограниченности по норме слабо сходящейся в $L_2^r[t_0, t_1] \times E^m$ последовательности. Поэтому, совершив в (3.1) предельный переход при $k = k_l \to \infty$, получим

$$\varliminf_{k_l\to\infty} \beta_{k_l}\langle Az', z - z'\rangle \geqslant 0, \quad \varliminf_{k_l\to\infty} \beta_{k_l} \geqslant \beta_{\min} > 0,$$

что соответствует двум неравенствам

$$\Big\langle \frac{\partial \mathcal{L}(u', \lambda')}{\partial u}, u - u' \Big\rangle \geqslant 0 \ \ \forall\, u \in U, \qquad \Big\langle \frac{\partial \mathcal{L}(u', \lambda')}{\partial \lambda}, \lambda - \lambda' \Big\rangle \leqslant 0 \ \ \forall\, \lambda \in E^m_+,$$

первое из которых равносильно правому неравенству седла $L(u', \lambda') \leqslant L(u, \lambda')$ $\forall\, u \in U$, второе — левому неравенству седла $L(u', \lambda) \leqslant L(u', \lambda')$ $\forall\, \lambda \in E^m_+$ [1, кн. 1, гл. 4, § 2, теорема 3]. Таким образом, (u', λ') — решение задачи (1.1)–(1.3). Тем самым доказана следующая теорема.

Теорема. *Пусть U — выпуклое замкнутое множество из $L_2^r[t_0, t_1]$, функции $f^i(x)$, $i = 0, 1, \ldots, m$, выпуклы и непрерывно дифференцируемы на E^n, их градиенты удовлетворяют условию Липшица на ограниченном множестве из E^n, функция Лагранжа* (1.4) *имеет хотя бы одну седловую точку, $z_0 = (u_0, \lambda_0) \in Z = L_2^r[t_0, t_1] \times E^m_+$ — произвольная начальная точка метода* (2.7); *длина шага β_k в* (2.7) *выбирается из условия* (3.12), *где L — константа Липшица из* (3.11) *для оператора A, соответствующая множеству* (3.10). *Тогда последовательность $z_k = (u_k, \lambda_k)$, $k = 0, 1, \ldots$, порожденная методом* (2.7), *ограничена в норме пространства $L_2^r[t_0, t_1] \times E^m$ и любая ее слабая предельная точка (u', λ') является решением задачи* (1.1)–(1.3).

4. Заключение

Необходимо заметить, что правило (3.12) выбора шага β_k в методе (2.7) не является конструктивным, так как точка z_*, присутствующая в определении (3.10) множества S, нам неизвестна и, следовательно, константа L в (3.11) теоретически существует, но мы ее явно не знаем. Однако, если известна дополнительная информация о задаче (1.1)–(1.3), правило выбора длины шага β_k можно сделать более конструктивным. Например, если функция $f(x)$ афинна, то, как следует из неравенства (3.9), в качестве множества (3.10) можно взять исходное, известное нам множество $S = Z$ и вычислить явно константу L в (3.11), (3.12). В общем случае, когда функция $f(x)$ нелинейна, допустим, что нам удалось получить оценку

$$\|z_*\| \leqslant M \tag{4.1}$$

для какой-либо седловой точки функции (1.4). Тогда вместо (3.10) можно взять, например, множество

$$S_1 = \big\{z \in Z : \ \|z - z_0\| \leqslant 4(\|z_0\| + M) + \sup_{\|z\|\leqslant M} \|Az\| \stackrel{def}{=} R_1\big\}. \tag{4.2}$$

Это множество получено из множества (3.10) заменой величины $\|z_*\|$ ее оценкой M из (4.1), и константа R_1 здесь конструктивно вычислима. Если теперь в (3.11), (3.12) вместо L взять константу Липшица, соответствующую множеству (4.2), то, рассуждая по индукции, также несложно доказать соотношения (3.13) для множества (4.2) и вытекающую отсюда сходимость метода (2.7).

Если оценки (4.1) у нас нет, то остается выбирать длину шага β_k в методе (2.7) эмпирически (например, дроблением некоторого фиксированного $\beta > 0$), руководствуясь соображениями сходимости, монотонности и т. п. Ценность доказанной выше теоремы — в том, что при решении задачи (1.1)–(1.3) методом (2.7) такой эмпирический подход также применим.

СПИСОК ЛИТЕРАТУРЫ

1. **Васильев Ф.П.** Методы оптимизации: в 2 кн. М.: МЦНМО, 2011. 1056 с.
2. **Антипин А.С., Хорошилова Е.В.** Многокритериальная краевая задача в динамике // Тр. Ин-та математики и механики УрО РАН. 2015. Т. 21, № 3. С. 20–29.
3. **Подиновский В.В., Ногин В.Д.** Парето-оптимальные решения многокритериальных задач. М.: Физматлит, 2007. 256 с.
4. **Антипин А.С.** О двух постановках равновесных задач // Оптимизация и приложения. М.: Изд-во ВЦ РАН, 2011. Вып. 2. С. 13–41.
5. **Артемьева Л.А.** Экстраградиентный метод поиска точки равновесия в седловых играх двух лиц // Журн. вычисл. математики и мат. физики. 2011. Т. 51, № 12. С. 2143–2157.

Поступила 24.02.2016

Васильев Федор Павлович
д-р физ.-мат. наук, профессор
Московский государственный университет им. М.В. Ломоносова, ВМиК
e-mail: vasiliev.fp@gmail.com

Антипин Анатолий Сергеевич
д-р физ.-мат. наук, профессор
главный науч. сотрудник
ВЦ РАН
e-mail: asantip@yandex.ru

Артемьева Людмила Анатольевна
канд. физ.-мат. наук
ассистент
Московский государственный университет им. М.В. Ломоносова, ВМиК
e-mail: artemieva.luda@gmail.com