Дифференциальные уравнения, т. 31, No.11, 1995, с. 1786–1795.

УДК 517.977

А.С. АНТИПИН

О ДИФФЕРЕНЦИАЛЬНЫХ ГРАДИЕНТНЫХ МЕТОДАХ ПРОГНОЗНОГО ТИПА ДЛЯ ВЫЧИСЛЕНИЯ НЕПОДВИЖНЫХ ТОЧЕК ЭКСТРЕМАЛЬНЫХ ОТОБРАЖЕНИЙ[1]

(Пересмотрено 1 декабря 2003)

1. Постановка проблемы. Рассмотрим задачу вычисления неподвижной точки экстремального отображения

$$v^* \in \operatorname{Argmin}\{\Phi(v^*, w) \mid w \in \Omega\}, \quad v^* \in \Omega^*, \tag{1.1}$$

где функция $\Phi(v, w)$ определена на произведении пространств $R^n \times R^n$ и $\Omega \subset R^n$. Предполагается, что $\Phi(v, w)$ выпуклая по переменной $w \in \Omega$ при каждом фиксированном $v \in \Omega$, экстремальное (маргинальное) отображение $G(v) \equiv \operatorname{argmin}\{\Phi(v, w) \mid \mid w \in \Omega\}$ определено для всех $v \in \Omega$, а множество решений $\Omega^* \subset \Omega$ исходной задачи непусто. Каждая точка этого множества согласно (1.1) удовлетворяет неравенству

$$\Phi(v^*, v^*) \leq \Phi(v^*, w) \quad \forall w \in \Omega, \quad v^* \in \Omega^*. \tag{1.2}$$

Если функция $\Phi(v, w)$ дифференцируема по второй переменной, то (1.2) эквивалентна вариационному неравенству [1]

$$\langle \nabla\Phi_w(v^*, v^*), w - v^* \rangle \geq 0 \quad \forall w \in \Omega, \quad v^* \in \Omega^*. \tag{1.3}$$

где $\nabla\Phi_w(v, w)$ — вектор-градиент функции относительно второй переменной.

Неравенство (1.2) связано с фактом существования неподвижной точки и никак не связано с ее устойчивостью, т.е. с возможностью попасть в малую окрестность этой точки с помощью какого-нибудь метода. Для характеристики устойчивости точки предлагаем использовать другое неравенство, а именно [2, 3]

$$\Phi(w, v^*) \leq \Phi(w, w) \quad \forall w \in \Omega, \quad v^* \in \Omega^*, \tag{1.4}$$

которое совместно с (1.2) обобщает понятие седловой точки [4, 5]. Действительно, если для всех $w \in \Omega$ $\Phi(w, w) \equiv \Phi(v^*, v^*) = const$, т.е. на диагонали квадрата $\Omega \times \Omega$ функция $\Phi(v, w)$ тождественно равна константе $\Phi(v^*, v^*)$, то неравенство (1.4) принимает вид

$$\Phi(w, v^*) \leq \Phi(v^*, v^*) \quad \forall w \in \Omega, \quad v^* \in \Omega^*. \tag{1.5}$$

Система неравенств (1.2) и (1.5) определяет седловую точку, причем ее первая компонента (в равной мере, и вторая) — неподвижная точка.

Неравенство (1.4) можно рассматривать как следствие более ограничивающего неравенства вида

$$\Phi(w, w) - \Phi(w, v^*) - \Phi(v^*, w) + \Phi(v^*, v^*) \geq 0 \quad \forall w \in \Omega, \quad v^* \in \Omega^*, \tag{1.6}$$

---

[1]Работа поддержана Российским Фондом Фундаментальных Исследований.
Грант No 94-01-00005

которое верно для всех $w \in \Omega$ и всех $w^* \in \Omega^*$. Действительно, в силу (1.2) из (1.6) немедленно следует (1.4).

Цель работы состоит в том, чтобы доказать сходимость градиентного прогнозного метода к решению задачи (1.1), а также получить оценки скорости сходимости этого метода при выполнении условий (1.4) и (1.6) и их различных модификаций.

2. Задачи с кососимметричной целевой функцией. Неравенства (1.4) и (1.6), как станет ясным из дальнейшего изложения, позволяют устанавливать свойства сходимости процессов, рассмотренных в этой работе. Однако эти неравенства носят неконструктивный характер, так как в их формальное определение входит неизвестный вектор v^* . Это значит, что любую конкретную задачу мы не можем проверить на предмет выполнения для нее неравенств (1.4) или (1.6). Поэтому возникает важная проблема описания классов функций $\Phi(v, w)$, для которых указанные выше неравенства выполняются автоматически. Такие классы функций существуют. Перейдем к описанию одного из них. Рассмотрим класс функций, подчиненный соотношению вида

$$\Phi(w, v) + \Phi(v, w) \leq 0 \quad \forall w \in \Omega, \quad \forall v \in \Omega. \tag{2.1}$$

Напомним, что точки v, w и w, v расположены симметрично относительно диагонали квадрата $w = v$. Особенно отметим случай, когда (2.1) выполняется со знаком равенства

$$\Phi(w, v) + \Phi(v, w) = 0 \quad \forall w \in \Omega, \quad \forall v \in \Omega. \tag{2.2}$$

При $w = v$ из (2.2) следует, что на диагонали квадрата функция $\Phi(w, w) = 0$. Если $\Phi(v, w)$ определена на конечном наборе значений v_i, w_j, $i = 1, 2, \dots, n$, $j = 1, 2, \dots, m$, то ее можно рассматривать как матрицу $\Phi_{i,j}$, при этом (2.2) переходит в хорошо известное определение кососимметричной матрицы $\Phi_{i,j} + \Phi_{j,i} = 0 \quad \forall i, j$. Функцию $\Phi(v, w)$ со свойством (2.2) естественно назвать кососимметричной функцией. Соответственно с помощью условия

$$\Phi(w, v) - \Phi(v, w) = 0 \quad \forall w \in \Omega, \quad \forall v \in \Omega \tag{2.3}$$

можно ввести понятие симметричной функции $\Phi(v, w)$. Если точке v, w поставить в соответствие значение функции $\Phi(\cdot, \cdot)$ в точке w, v, то получим функцию, которую можно назвать транспонированной $\Phi^\top(v, w)$. В терминах этой функции условия (2.2) и (2.3) имеют вид $\Phi(v, w) = -\Phi^\top(v, w)$ и $\Phi(v, w) = \Phi^\top(v, w)$. Нетрудно проверить, что вещественную функцию $\Phi(v, w)$ всегда можно представить в виде суммы

$$\Phi(v, w) = \Psi_1(v, w) + \Psi_2(v, w),$$

где функция $\Psi_1(v, w)$ — симметричная, а $\Psi_2(v, w)$ — кососимметричная. Это разложение единственно, причем

$$\Psi_1(v, w) = 0.5(\Phi(v, w) + \Phi^\top(v, w)), \qquad \Psi_2(v, w) = 0.5(\Phi(v, w) - \Phi^\top(v, w)).$$

В данной работе мы ограничимся случаем кососимметричной $\Phi(v, w)$, причем под свойством кососимметричности мы будем понимать условие несколько более общее, чем (2.1), а именно

$$\Phi(w, w) - \Phi(w, v) - \Phi(v, w) + \Phi(v, v) \geq 0 \quad \forall w \in \Omega, \quad \forall v \in \Omega. \tag{2.4}$$

Если $\Phi(w,w)=0$, то (2.4) переходит в (2.1).

Из дальнейшего изложения станет ясно, что за понятием кососимметричности скрываются равновесные задачи и равновесные решения, которые унаследовали свойства задач оптимизации и седловых задач.

Неравенство (2.4) верно для всех $v\in\Omega$ и $w\in\Omega$, в частности, оно верно и при $v=v^*\in\Omega^*$. С учетом (1.2) из него немедленно следует (1.4). Таким образом, если целевая функция задачи (1.1) обладает свойством кососимметричности (2.4), то равновесное решение задачи (1.1) удовлетворяет условию (1.4).

Покажем, что нормализованная функция седловой задачи обладает свойством кососимметричности. Седловая задача представляет собой решение системы неравенств

$$L(x^*,p)\le L(x^*,p^*)\le L(x,p^*),\qquad x\in Q\subset R^n,\quad p\in P\subset R^m, \tag{2.5}$$

где $L(x,p)$ — функция, выпуклая по x и вогнутая по p. С помощью нормализованной функции вида [6] $\Phi(v,w)=L(z,p)-L(x,y)$, $w=(z,y)$, $v=(x,p)$ система (2.5) может быть сведена к вычислению неподвижной точки экстремального включения (1.1). Поскольку функция $\Phi(v,w)$ сепарабельна по переменным z и y, а множество $\Omega=Q\times P$ имеет блочную структуру, задача (1.1) эквивалентна задаче (2.4), а множества решений обеих задач совпадают.

Убедимся, что нормализованная функция $\Phi(v,w)$ седловой задачи (2.5) удовлетворяет следующим условиям [4, 5]:

$$\Phi(w,w)=0\quad\forall w\in\Omega, \tag{2.6}$$

$$\Phi(w,v)+\Phi(v,w)=0\quad\forall w\in\Omega,\quad\forall v\in\Omega. \tag{2.7}$$

Первое свойство означает, что на диагонали квадрата, т.е. при $v=w$, функция $\Phi(v,w)$ равна нулю. Выполнимость его для функции $\Phi(v,w)=L(z,p)-L(x,y)$, где $w=(z,y)$, $v=(x,p)$, очевидна, так как при $v=w$ имеем $\Phi(w,w)=L(z,y)-L(z,y)=$ $=0$.

В справедливости второго свойства убедиться также нетрудно. Действительно, пусть $\Phi(v,w)=L(z,p)-L(x,y)$. Так как область изменения переменных $w\in\Omega$ и $v\in\Omega$ одна и та же, положим $v=w$, а $w=v$ в $\Phi(v,w)$, тогда $\Phi(w,v)=L(x,y)-L(z,p)$. Отсюда $\Phi(w,v)+\Phi(v,w)=L(x,y)-L(z,p)+L(z,p)-L(x,y)=0$.

Сложив (2.6) и (2.7), получим равенство

$$\Phi(w,w)-\Phi(w,v)-\Phi(v,w)+\Phi(v,v)=0\quad\forall w\in\Omega,\quad\forall v\in\Omega. \tag{2.8}$$

С учетом (1.2) при $v=v^*$ из (2.8) немедленно следует (1.4). Так как $\Phi(w,w)=0$, $\Phi(v^*,v^*)=0$, то имеем

$$\Phi(w,v^*)\le\Phi(v^*,v^*)\le\Phi(v^*,w),\qquad w\in\Omega,\quad v^*\in\Omega^*. \tag{2.9}$$

Таким образом, v^*,v^* — седловая точка для нормализованной функции $\Phi(v,w)$, причем $\Phi(v,v)=0$, где v — неподвижная точка. Левое неравенство системы (2.9) представляет собой условие (1.4).

Соотношения (2.6), (2.7) могут быть обобщены до уровня неравенств

$$\Phi(w,w)\ge 0\quad\forall w\in\Omega. \tag{2.10}$$

$$\Phi(w,v)+\Phi(v,w)\le 0 \quad \forall w\in\Omega,\quad \forall v\in\Omega. \tag{2.11}$$

Такое обобщение расширяет класс решаемых равновесных задач. Отметим, что система неравенств (2.10), (2.11) была использована в [7] для доказательства неравенства Ки Фаня, справедливость которого в свою очередь эквивалентна теореме Какутани о существовании неподвижной точки непрерывного точечно-множественного отображения на выпуклом замкнутом ограниченном множестве [8].

Если функция $\Phi(v,w)$ дифференцируема по второй переменной, то в силу неравенств выпуклости

$$\langle\nabla f(x),y-x\rangle\le f(y)-f(x)\le\langle\nabla f(y),y-x\rangle, \tag{2.12}$$

для всех y и x для некоторого множества из (1.6) имеем неравенство

$$\langle\nabla\Phi_w(w,w)-\nabla\Phi_w(v^*,v^*),w-v^*\rangle\ge 0 \quad \forall w\in\Omega,\quad v^*\in\Omega^*,$$

которое означает, что частный градиент $\nabla\Phi_w(v,w)$ является монотонным оператором относительно равновесия v^*.

Существуют классы равновесных задач, для которых неравенства (1.4) и (1.6) можно усилить. Эти задачи являются аналогами задач оптимизации с сильно выпуклыми целевыми функциями и их решения удовлетворяют следующим неравенствам [2, 3]:

1) решение задачи единственно, тогда

$$\Phi(w,w)-\Phi(w,v^*)\ge\gamma|w-v^*|^{1+\nu} \quad \forall w\in\Omega, \tag{2.13}$$

константа $\gamma>0$, параметр $\nu\in[0,\infty]$, при $\nu=0$ имеем острое изолированное равновесие.

2) совокупность решений задачи представляют собой замкнутое связное множество, тогда

$$\Phi(w,w)-\Phi(w,\pi_{\Omega^*}(w))\ge\gamma|w-\pi_{\Omega^*}(w)|^{1+\nu} \quad \forall w\in\Omega, \tag{2.14}$$

$\pi_{\Omega^*}(w)$ — оператор проектирования вектора w на множество Ω^*. Если множество Ω^* состоит из одной точки, то (2.14) переходит в (2.13). Отметим также, что из (2.14) следует (1.4).

Аналог (1.6) имеет вид

$$\begin{aligned}(\Phi(w,w)-\Phi(w,\pi_{\Omega^*}(w))) \;&-\; (\Phi(\pi_{\Omega^*}(w),w)-\Phi(\pi_{\Omega^*}(w),\pi_{\Omega^*}(w)))\ge\\ &\ge\; \gamma|w-\pi_{\Omega^*}(w)|^{1+\nu} \quad \forall w\in\Omega.\end{aligned} \tag{2.15}$$

Если функция $\Phi(w,v)$ дифференцируемая и выпуклая по второй переменной, то неравенство (2.15) можно представить в форме [2, 3]

$$\langle\nabla\Phi_w(w,w)-\nabla\Phi_w(\pi_{\Omega^*}(w),\pi_{\Omega^*}(w)),w-\pi_{\Omega^*}(w)\rangle\ge\gamma|w-\pi_{\Omega^*}(w)|^{1+\nu} \quad \forall w\in\Omega.$$

3. Линейные и квадратичные классы задач. В этом разделе в качестве примеров исследуем два популярных класса задач.

3.1. Квадратичное равновесие. Рассмотрим задачу отыскания неподвижной точки квадратичного экстремального включения

$$v^*\in\mathrm{Argmin}\{0.5\langle Nw,w\rangle+\langle Mv^*+m,w\rangle\mid w\in\Omega\}, \tag{3.1}$$

где N и M — неотрицательные матрицы, т.е. $\langle Nv, v\rangle \geq 0$ и $\langle Mv, v\rangle \geq 0$ для всех $v \in R^n$. Кроме того, будем предполагать, что N — симметричная матрица. В частности, если $\Omega = R^n$, то задача (3.1) сводится к решению системы линейных уравнений $(N + M)w = -m$.

Убедимся, что целевая функция задачи (3.1) удовлетворяет условию кососимметричности (2.4), а в точках решений выполняется условие (1.6). Рассмотрим

$$\begin{gathered}\Phi(w, w) - \Phi(w, v) - \Phi(v, w) + \Phi(v, v) = \\ = 0.5\langle Nw, w\rangle + \langle Mw, w\rangle + \langle m, w\rangle - 0.5\langle Nv, v\rangle - \langle Mw, v\rangle - \langle m, v\rangle - \\ -0.5\langle Nw, w\rangle - \langle Mv, w\rangle - \langle m, w\rangle + 0.5\langle Nv, v\rangle + \langle Mv, v\rangle + \langle m, v\rangle = \\ = \langle M(w - v), w - v\rangle \geq 0 \qquad \forall w \in \Omega, \quad \forall v \in \Omega.\end{gathered} \tag{3.2}$$

Отсюда следует, что если матрица M неотрицательна, то $\Phi(w, v)$ из (3.1) кососимметрична, и, следовательно, решение этой задачи удовлетворяет условию (2.13) или (2.14) в зависимости от поведения собственных чисел матрицы M.

Действительно, пусть $v = v^*$ в (3.2), т.е.

$$\Phi(w, w) - \Phi(w, v^*) - \Phi(v^*, w) + \Phi(v^*, v^*) = \langle M(w - v^*), w - v^*\rangle \geq 0. \tag{3.3}$$

Пусть матрица M вырожденная, тогда рассмотрим разложение пространства R^n в прямую сумму $R^n = H_1 + H_2$, где H_1 — ядро матрицы M, а H_2 — ортогональное дополнение к H_1. В этом случае любой вектор $w - v^* \in R^n$ имеет представление $w - v^* = h_1 + h_2$, где $h_1 = \pi_{H_1}(w - v^*)$ и $h_2 = \pi_{H_2}(w - v^*)$, кроме того $Mh_1 = 0$, $Mh_2 \in H_2$. Учитывая все это, продолжим вычисления:

$$\begin{gathered}\langle M(w - v^*), w - v^*\rangle = \langle M^{1/2}(w - v^*), M^{1/2}(w - v^*)\rangle = \\ = \langle M^{1/2}(h_1 + h_2), M^{1/2}(h_1 + h_2)\rangle = \langle M^{1/2}h_2, M^{1/2}h_2\rangle = \\ = \langle Mh_2, h_2\rangle \geq \mu|h_2|^2 = \mu|w - v^* - h_1|^2 = \mu|w - v^* - \pi_{H_1}(w - v^*)|^2 = \\ = \mu|w - v^* - \pi_{H_1}(w) - \pi_{H_1}(v^*)|^2 = \mu|w - \pi_{H_1}(w)|^2.\end{gathered} \tag{3.4}$$

Здесь μ — минимальное ненулевое собственное число вырожденной матрицы M. В этой цепочке рассуждений были дополнительно использованы существование квадратного корня симметричной матрицы M и свойство линейности оператора проектирования $\pi_{H_1}(w - v^*)$, причем $\pi_{H_1}(v^*) = v^*$. Сопоставляя (3.3) и (3.4), получим (2.15) при $n = 1$, а учитывая (1.2), и (2.14).

3.2. Острое равновесие. Пусть в задаче (3.1) матрицы N и M имеют вид $N = 0$ и $M = \begin{pmatrix} 0 & -L^\top \\ L & 0 \end{pmatrix}$, где L — некоторая подматрица, а Ω имеет структуру многогранного множества

$$\Omega = \left\{ \begin{pmatrix} z \\ y \end{pmatrix} \mid \begin{pmatrix} 0 & -B \\ A & 0 \end{pmatrix} \begin{pmatrix} z \\ y \end{pmatrix} \leq \begin{pmatrix} -c \\ b \end{pmatrix} \begin{pmatrix} z \\ y \end{pmatrix} \geq \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\}.$$

Введем обозначения для векторов $v = (x, p)$, $w = (z, y)$, $m = (-c, b)$ и перепишем задачу (3.1) в векторно-матричном виде

$$(x^*, p^*) \in \operatorname{Argmin} \left\{ \begin{array}{c} \begin{pmatrix} 0 & -L^\top \\ L & 0 \end{pmatrix} \begin{pmatrix} z \\ y \end{pmatrix} + (-c, b) \begin{pmatrix} z \\ y \end{pmatrix} \mid \\ \begin{pmatrix} 0 & -B \\ A & 0 \end{pmatrix} \begin{pmatrix} z \\ y \end{pmatrix} \leq \begin{pmatrix} -c \\ b \end{pmatrix} \begin{pmatrix} z \\ y \end{pmatrix} \geq \begin{pmatrix} 0 \\ 0 \end{pmatrix} \end{array} \right\} \tag{3.5}$$

или в обозначениях (3.1)

$$v^* \in \text{Argmin}\{\langle Mv^* + m, w\rangle \mid w \in \Omega\}. \tag{3.6}$$

Используя линейный характер задачи и блочную структуру ее ограничений, представим (3.5) в виде игры двух лиц с ненулевой суммой

$$\begin{array}{lcl} x^* & \in & \text{Argmin}\{-\langle c, z\rangle + \langle Lz, p^*\rangle \mid Az \leq b,\ z \geq 0\}, \\ p^* & \in & \text{Argmin}\{\langle b, y\rangle + \langle Lx^*, y\rangle \mid By \geq c,\ y \geq 0\}. \end{array} \tag{3.7}$$

Предполагается, что размерности матриц L, A и B согласованы с размерностями переменных x и y и векторов c и b. Если $L = 0$ и $B = A^\top$, то (3.7) совпадает с прямой и двойственной задачами линейного программирования.

Каждая из задач (3.7) представляет собой задачу линейного программирования. В работе [9] показано, что множество оптимумов любой из них является множеством острых минимумов, т.е. удовлетворяет условию

$$\begin{array}{l} \gamma|z - x^*| \leq (-\langle c, z\rangle + \langle Lz, p^*\rangle) - (-\langle c, x^*\rangle + \langle Lx^*, p^*\rangle), \\ \gamma|y - p^*| \leq (\langle b, y\rangle - \langle Lx^*, y\rangle) - (\langle b, p^*\rangle - \langle Lx^*, p^*\rangle) \end{array} \tag{3.8}$$

для всех $z \in \{z \mid Az \leq b,\ z \geq 0\}$ и всех $y \in \{y \mid By \geq c,\ y \geq 0\}$, где $\gamma > 0$. Сложим неравенства (3.8) и полученное суммарное неравенство представим в обозначениях задачи (3.6)

$$\langle Mv^*, w\rangle + \langle m, w\rangle - \langle Mv^*, v^*\rangle - \langle m, v^*\rangle \geq \gamma|w - v^*|. \tag{3.9}$$

Неравенство справедливо для всех $w \in \Omega$. Поскольку матрица M в задаче (3.6) неотрицательна, т.е. $\langle Mv, v\rangle \geq 0$ для всех $v \in R^n$, то целевая функция этой задачи удовлетворяет условию кососимметричности (1.6). Комбинация (1.6) и (3.9) приводит к основному условию (1.4): $\langle Mw, w\rangle + \langle m, w\rangle - \langle Mw, v^*\rangle - \langle m, v^*\rangle \geq |w - v^*|\ \forall w \in \Omega$. Таким образом, неподвижная точка линейной игры двух лиц (3.7) является острым равновесием.

4. Управляемые методы. Геометрически понятие сходимости любого метода на фазовом портрете динамической системы ассоциируется с особенностью типа “устойчивый узел”, к которому сходятся все траектории из любой точки некоторой окрестности равновесия. В связи с этим возникает естественная идея о такой деформации фазового портрета динамической системы, при которой неустойчивая равновесная точка преобразовалась бы в асимптотически устойчивый узел, не изменив при этом своих координат. В данной работе эта концепция реализуется с помощью идеи управления в виде обратных связей [10].

Перейдем к описанию этого подхода. Предположим, что функция $\Phi(v, w)$ дифференцируема по второй переменной, и, следовательно, необходимое (и достаточное) условие минимума задачи (1.1) может быть выражено в форме уравнения

$$v^* = \pi_\Omega(v^* - \alpha\nabla\Phi_w(v^*, v^*)), \tag{4.1}$$

где $\pi_\Omega(\ldots)$ — оператор проектирования вектора на множество Ω, а $\alpha > 0$ — параметр типа длины шага, $\nabla\Phi_w(v, w)$ — вектор-градиент функции $\Phi(v, w)$ по второй переменной w.

Точка v^* является неподвижной точкой или точкой равновесия. Если из точки v^* сделать шаг по частному антиградиенту функции $\Phi(v, w)$, то после проектирования процесс снова окажется в точке v^*. Задачи (1.1) и (4.1) эквивалентны.

Невязка, т.е. разность между левой и правой частями уравнения (4.1), равная нулю в точке v^* и соответственно не равная нулю в произвольной точке v, задает преобразование R^n в R^n. Образ этого преобразования можно рассматривать как векторное поле, неподвижная точка которого v^*. Поставим задачу о проведении траектории такой, чтобы ее касательная совпала с заданным направлением поля в этой точке. Формально задача описывается системой дифференциальных уравнений

$$\frac{dv}{dt} + v = \pi_\Omega(v - \alpha\nabla\Phi_w(v, v)), \qquad v(t_0) = v^0. \tag{4.2}$$

Поскольку правая часть этой системы удовлетворяет всем условиям теоремы существования и единственности, то система (4.2) для всех $v(t_0) = v^0$ порождает траекторию $v(t)$ для всех $t \geq t_0$.

Градиент по второй переменной $\nabla\Phi_w(v, v)$ не является потенциальным оператором, т.к., вообще говоря, не существует функции, градиент которой совпадает с $\nabla\Phi_w(v, v)$. Как показывают примеры [4, 5], процесс не сходится к равновесию, для обеспечения сходимости в правую часть динамической системы (4.2) введем аддитивное управление [4, 5]

$$\frac{dv}{dt} + v = \pi_\Omega(v - \alpha\nabla\Phi_w(v + u, v + u)), \qquad v(t_0) = v^0, \tag{4.3}$$

и поставим следующую задачу. В некотором классе обратных связей требуется выбрать управление $u = u(v, \dot{v})$ (где $\dot{v} = dv/dt$) как функцию состояния динамической системы, которое обеспечило бы сходимость траектории к равновесию. В точке равновесия объект неподвижен и его скорости равны нулю, поэтому $u = u(v^*, \dot{v}^*) = 0$.

Рассматриваются управления двух видов: по производной $u = \dot{v}$ и по невязке

$$u = \pi_\Omega(v - \alpha\nabla\Phi_w(v, v)) - v. \tag{4.4}$$

Замыкание системы (4.3) управлением по производной $u = \dot{v}$ приводит к неявной (не разрешенной относительно производной) дифференциальной системе

$$\frac{dv}{dt} + v = \pi_\Omega(v - \alpha\nabla\Phi_w(v + \dot{v}, v + \dot{v})), \qquad v(t_0) = v^0, \tag{4.5}$$

итеративный аналог которой представляет собой неявный итеративный процесс

$$v^{n+1} = \pi_\Omega(v^n - \alpha\nabla\Phi_w(v^{n+1}, v^{n+1})), \tag{4.6}$$

v^n — ранее найденное приближение, а уравнение (4.6) должно быть разрешено относительно переменных v^{n+1}. Для решения уравнения в свою очередь нужны другие итеративные подпроцессы.

Отдавая должное преимуществам процесса (4.6), необходимо отметить его недостаток, который проявляется в его неявности (или неразрешимости) относительно производной. После замыкания управлением (4.4) системы (4.3) получим

$$\frac{dv}{dt} + v = \pi_\Omega(v - \alpha\nabla\Phi_w(\bar{u}, \bar{u})), \qquad \bar{u} = \pi_\Omega(v - \alpha\nabla\Phi_w(v, v)). \tag{4.7}$$

Система (4.7) является явной, что особенно проявляется в ее итеративном аналоге:

$$\bar{u}^n = \pi_\Omega(v^n - \alpha\nabla\Phi_w(v^n, v^n)), \qquad v^{n+1} = \pi_\Omega(v^n - \alpha\nabla\Phi_w(\bar{u}^n, \bar{u}^n)),$$

представляющим собой явную итеративную схему с предварительным (или прогнозным) шагом, в результате которого вычисляется сначала прогноз $\bar{u}^n$, а затем следующее приближение v^{n+1}.

5. Сходимость к острому равновесию. Рассмотрим поведение прогнозного метода проекции градиента (4.7) в случае острого равновесия [2, 11].

Представим процесс (4.7) в соответствии с определением оператора проектирования в форме следующих вариационных неравенств:

$$\langle \dot{v} + v - v + \alpha\nabla\Phi_w(\bar{u}, \bar{u}), w - v - \dot{v}\rangle \geq 0 \tag{5.1}$$

и

$$\langle \bar{u} - v + \alpha\nabla\Phi_w(v, v), w - \bar{u}\rangle \geq 0 \tag{5.2}$$

для всех $v \in \Omega$. Кроме того, для доказательства сходимости метода необходимо условие Липшица в форме

$$|\nabla\Phi_w(w, w) - \nabla\Phi_w(v, v)| \leq |\Phi|\,|w - v| \qquad \forall w, v \in \Omega, \tag{5.3}$$

где $|\Phi|$ — константа Липшица для оператора $\nabla\Phi_w(v, w)$. Это условие заведомо будет иметь место, если градиент функции $\Phi(v, w)$ удовлетворяет условию Липшица по совокупности переменных.

С учетом (5.3) получим оценку величины отклонения двух векторов $v + \dot{v}$ и $\bar{u}$ из (4.7):

$$\begin{aligned} |v + \dot{v} - \bar{u}| &\leq |\pi_\Omega(v - \alpha\nabla\Phi_w(\bar{u}, \bar{u})) - \pi_\Omega(v - \alpha\nabla\Phi_w(v, v))| \leq \\ &\leq |\nabla\Phi_w(\bar{u}, \bar{u}) - \nabla\Phi_w(v, v)| \leq \alpha|\Phi|\,|\bar{u} - v|, \qquad \forall w, v \in \Omega, \end{aligned} \tag{5.4}$$

где $|\Phi|$ — константа Липшица векторной функции $\Phi(v, v)$.

Отсутствующее условие монотонности градиента $\nabla\Phi_w(v, w)$ заменим на условие остроты (2.13) с $\nu = 0$. При этом предполагаем, что рассматривается случай, когда исходная задача имеет изолированное равновесие

$$\Phi(w, w) - \Phi(w, v^*) \geq \gamma|w - v^*|, \quad \forall w \in \Omega. \tag{5.5}$$

Процесс (4.7) в этом случае сходится за конечное время.

Теорема 1. *Если множество решений задачи (1.1) непусто и удовлетворяет условию остроты (5.5), целевая функция $\Phi(v, w)$ — выпуклая по переменной w при каждом фиксированном значении v, Ω — выпуклое замкнутое множество, кроме того, частный градиент функции $\Phi(v, w)$ удовлетворяет условию Липшица (5.3), то траектория $v(t)$ процесса (4.7) с параметром $0 < \alpha < 1/(\sqrt{2}|\Phi|)$, где $|\Phi|$ — константа из (5.3), сходится за конечное время, т.е. существует номер n_f такой, что $\bar{u}^{n_f} = v^*$.*

Д о к а з а т е л ь с т в о . Положим $w = v^*$ в (5.1) и представим полученное выражение в виде

$$\langle \dot v, v^* - v\rangle - |\dot v|^2 + \alpha\langle \nabla\Phi_w(\bar u, \bar u), v^* - \bar u\rangle + \alpha\langle \nabla\Phi_w(\bar u, \bar u), \bar u - v - \dot v\rangle \geq 0.$$

Отсюда с учетом выпуклости (2.12) имеем

$$\begin{aligned}\langle \dot v, v^* - v\rangle - |\dot v|^2 &+ \alpha(\Phi(\bar u, v^*) - \Phi(\bar u, \bar u)) - \alpha\langle \nabla\Phi_w(v, v) - \nabla\Phi_w(\bar u, \bar u), \bar u - v - \dot v\rangle + \\ &+ \alpha\langle \nabla\Phi_w(v, v), \bar u - v - \dot v\rangle \geq 0.\end{aligned}$$

Положим $w = v + \dot v$ в (5.2) и сложим полученное неравенство с предыдущим. Учитывая (5.3) и (5.4), имеем

$$\langle \dot v, v^* - v\rangle - |\dot v|^2 + \alpha(\Phi(\bar u, v^*) - \Phi(\bar u, \bar u)) + \alpha^2|\Phi|^2|\bar u - v|^2 + \langle \bar u - v, v + \dot v - \bar u\rangle \geq 0. \quad (5.6)$$

Последнее скалярное произведение в правой части (5.6) преобразуем с помощью тождества

$$|v_1 - v_2|^2 = |v_1 - v_3|^2 + 2\langle v_1 - v_3, v_3 - v_2\rangle + |v_3 - v_2|^2, \quad (5.7)$$

тогда

$$\langle \bar u - v, v + \dot v - \bar u\rangle = 0.5|\dot v|^2 - 0.5|\bar u - v|^2 - 0.5|\bar u - v - \dot v|^2.$$

Теперь (5.6) можно записать в форме

$$\langle \dot v, v - v^*\rangle + |\dot v|^2 + \alpha(\Phi(\bar u, \bar u) - \Phi(\bar u, v^*)) - \alpha^2|\Phi|^2|\bar u - v|^2 - 0.5|\dot v|^2 + 0.5|\bar u - v|^2 + 0.5|\bar u - v - \dot v|^2 \leq 0$$

или

$$\langle \dot v, v - v^*\rangle + 0.5|\dot v|^2 + \alpha(\Phi(\bar u, \bar u) - \Phi(\bar u, v^*)) + (0.5 - \alpha^2|\Phi|^2)|\bar u - v|^2 \leq 0. \quad (5.8)$$

Согласно (5.5) третье слагаемое в этом неравенстве неотрицательное, тогда

$$0.5\frac{d}{dt}|v - v^*|^2 + 0.5|\dot v|^2 + (0.5 - \alpha^2|\Phi|^2)|\bar u - v|^2 \leq 0. \quad (5.9)$$

Так как $0 < \alpha < 1/(\sqrt{2}|\Phi|)$, то $d = 0.5 - \alpha^2|\Phi|^2 > 0$. Проинтегрируем неравенство (5.9) от t_0 до t:

$$|v - v^*|^2 + \int_{t_0}^{t} |\dot v|^2 d\tau + (1 - 2\alpha^2|\Phi|^2)\int_{t_0}^{t} |\bar u - v|^2 d\tau \leq |v_0 - v^*|^2.$$

Отсюда вытекает монотонное убывание величины $|v(t) - v^*|$, а также ограниченность траектории $v(t)$ и, кроме того, сходимость интегралов:

$$\int_{t_0}^{t} |\dot v|^2 d\tau < \infty, \qquad \int_{t_0}^{t} |\bar u - v|^2 d\tau < \infty$$

при $t \to \infty$ и, следовательно,

$$\liminf |\dot v|^2 = 0, \qquad \liminf |\bar u - v|^2 = 0, \qquad t \to \infty. \quad (5.10)$$

Получим еще одно утверждение типа (5.10). Для этого рассмотрим вытекающее из (5.9) неравенство

$$0.5\frac{d}{dt}|v-v^*|^2+0.5|\dot v|^2\le 0,$$

дифференцирование которого дает

$$|v-v^*|\frac{d}{dt}|v-v^*|+0.5|\dot v|^2\le 0.$$

Предполагая, что $|v(t)-v^*|\ne 0$ для всех $t\ge t_0$, имеем

$$\frac{d}{dt}|v-v^*|^2+0.5\frac{|\dot v|^2}{|v-v^*|}\le 0.$$

Проинтегрировав это неравенство от t_0 до t, получим оценку

$$|v-v^*|^2+\frac{1}{2}\int\limits_{t_0}^{t}\frac{|\dot v|^2}{|v-v^*|}d\tau\le|v_0-v^*|^2,$$

из которой вытекает сходимость интеграла и, следовательно, $\liminf(|\dot v|^2/|v-v^*|)=0$, $t\to\infty$.

Поскольку имеет место оценка $|\dot v|^2/(|v-\bar u|+|\bar u-v^*|)\le|\dot v|^2/|v-v^*|$, то в силу $|v-\bar u|\to 0$, $t\to\infty$ получим из нее

$$\liminf(|\dot v|^2/|\bar u-v^*|)=0,\qquad t\to\infty. \tag{5.11}$$

Вернемся к неравенству (5.8) и с учетом условия остроты (5.5) представим его в виде

$$\langle\dot v,\bar u-v^*\rangle+\langle\dot v,v-\bar u\rangle+0.5|\dot v|^2+\alpha\gamma|\bar u-v^*|+d|\bar u-v|^2\le 0.$$

Из второго и пятого слагаемых выделим полный квадрат

$$\langle\dot v,\bar u-v^*\rangle+|(0.5/\sqrt{d})\dot v+\sqrt{d}(v-\bar u)|^2-(0.25/d)|\bar u|^2+0.5|\dot v|^2+\alpha\gamma|\bar u-v^*|\le 0.$$

Отсюда $\alpha\gamma|\bar u-v^*|\le 0.25(1/d-2)|\dot v|^2+|\dot v|\,|\bar u-v^*|$. Предполагая, что $|\bar u-v^*|\ne 0$ для всех $t\ge t_0$, получим

$$\alpha\gamma\le 0.25(1/d-2)|\dot v|^2/|\bar u-v^*|+|\dot v|^2. \tag{5.12}$$

Сопоставляя неравенства (5.12) с (5.10) и (5.11), получаем противоречие. Следовательно, предположение, что $|\bar u-v^*|\ne 0$ для всех $t\ge t_0$ неверно, поэтому существует такое $t=t_f$, что $\bar u(t_f)=v^*$. $\square$

Аналогично можно доказать, что в условиях теоремы 1 сходится к равновесию за конечное время также неявный процесс (4.5).

6. Сходимость к квадратичному равновесию. В этом разделе получим экспоненциальную оценку скорости сходимости процесса (4.7) в предположении, что исходная задача (1.1) имеет изолированное квадратичное равновесие. Итак, пусть выполняется условие (2.13) с $n=1$

$$\Phi(w,w)-\Phi(w,v^*)\ge\gamma|w-v^*|^2,\quad\forall w\in\Omega. \tag{6.1}$$

Т е о р е м а 2. *Если в условиях теоремы 1 вместо (5.5) используется неравенство (6.1), то траектория $v(t)$ сходится к равновесному решению с экспоненциальной оценкой, т.е.*

$$|v(t) - v^*|^2 \le |v^0 - v^*|^2 \exp(2s(\alpha)(t_0 - t)), \tag{6.2}$$

где $s(\alpha) = \alpha\gamma(d/d_1) > 0$ и $d = 0.5 - \alpha^2|\Phi|^2$, $d_1 = 0.5 + \alpha\gamma - \alpha^2|\Phi|^2$.

Д о к а з а т е л ь с т в о . В неравенстве (5.8) из теоремы 1, используя (6.1), оценим третье слагаемое; получим

$$\langle \dot{v}, v - v^* \rangle + 0.5|\dot{v}|^2 + \alpha\gamma|\bar{u} - v^*| + d|\bar{u} - v|^2 \le 0. \tag{6.3}$$

С помощью тождества (5.7), где $v_1 = \bar{u}$, $v_2 = v^*$, $v_3 = v$, преобразуем третье слагаемое из левой части (6.3), тогда (6.3) примет вид

$$\langle \dot{v}, v - v^* \rangle + 0.5|\dot{v}|^2 + d_1|\bar{u} - v|^2 + 2\alpha\gamma\langle \bar{u} - v, v + v^* \rangle + \alpha\gamma|v - v^*|^2 \le 0. \tag{6.4}$$

Из суммы третьего и четвертого слагаемых (6.4) выделим полный квадрат $\langle \dot{v}, v - v^* \rangle + 0.5|\dot{v}|^2 + |\sqrt{d_1}(\bar{u} - v^*) + (\alpha\gamma/\sqrt{d_1})(v - v^*)|^2 - ((\alpha\gamma)^2/d_1)|v - v^*|^2 + \alpha\gamma|v - v^*| \le 0$, откуда

$$\langle \dot{v}, v - v^* \rangle + 0.5|\dot{v}|^2 + s(\alpha)|v - v^*| \le 0, \tag{6.5}$$

где $s(\alpha) = \alpha\gamma(d/d_1) > 0$. По предположению v^* — единственный минимум. Следовательно, (6.5) можно представить в виде $d|v - v^*|^2/dt + 2s(\alpha)|v - v^*|^2 \le 0$. Перепишем полученное неравенство в форме $\exp(-2s(\alpha)t)d(\exp(2s(\alpha)t)|v - v^*|^2)/dt \le 0$ или $d(\exp(2s(\alpha)t)|v - v^*|^2)/dt \le 0$, откуда и следует справедливость оценки (6.2). Нетрудно видеть, что при $0 < \alpha < 1/(\sqrt{2}|\Phi|)$ величина $s(\alpha) = \alpha\gamma(d/d_1) > 0$. $\square$

Таким образом, если значение параметра $\alpha > 0$ не слишком велико, то траектория $v(t)$ сходится к равновесному решению с экспоненциальной оценкой. Неявный процесс (4.5) также имеет экспоненциальную оценку скорости сходимости, если выполнены условия теоремы 2.

7. Сходимость к вырожденному равновесию. Сформулируем теорему о сходимости процесса (4.7) в вырожденном случае, т.е. когда выполняются условия (1.4)

Т е о р е м а 3. *Если множество решений задачи (1.1) непусто и удовлетворяет условию (1.4), целевая функция $\Phi(v, w)$ — выпуклая по переменной w при каждом фиксированном значении v, Ω — выпуклое замкнутое множество, кроме того, векторная функция $\nabla\Phi_w(v, w)$ удовлетворяет условию Липшица (5.3), то траектория $v(t)$ процесса (4.7) с параметром $0 < \alpha < 1(\sqrt{2}|\Phi|)$, где $|\Phi|$ — константа из (5.3), сходится монотонно по норме к одному из равновесных состояний, т.е. $v(t) \to v^* \in \Omega^*$, при $t \to \infty$ $\forall v^0 \in R^n$.*

Д о к а з а т е л ь с т в о . Проводя преобразования так, как в доказательстве теоремы 1, устанавливаем справедливость утверждения (5.10).

Пусть последовательность $t_i \to \infty$ такая, что $v(t_i) \to v'$, $\bar{u}(t_i) \to v'$, $\dot{v}(t_i) \to 0$. Рассмотрим (5.1) или (5.2) для всех $t \to \infty$. Переходя к пределу, получим предельное неравенство вида $\langle \nabla\Phi_w(v', v'), w - v' \rangle \geq 0 \quad \forall w \in \Omega$, совпадающее с неравенством (1.3). Последнее означает, что v' — равновесное решение задачи.

Таким образом, каждая предельная точка траектории $v(t)$ является решением исходной задачи. Этот факт вместе с монотонностью убывания величины $|v(t) - v^*|$ при $t \to \infty$ означает, что $v(t)$ имеет только одну предельную точку, т.е. траектория $v(t)$ сходится монотонно по норме к одному из решений исходной задачи, $v(t) \to v^* \in \in \Omega^*$ при $t \to \infty$ и любого $v^0 = v(t_0) \in R^n$. $\square$

В условиях теоремы 3 аналогично доказывается сходимость неявного процесса (4.5).

Таким образом, показано, что в случае острого равновесия рассматриваемые в работе процессы сходятся за конечное время, для квадратичного равновесия методы имеют экспоненциальную оценку скорости сходимости, и в случае вырожденного равновесия процессы могут сходиться сколь угодно медленно.

Литература

1. *Васильев Ф.П.* Численные методы решения экстремальных задач. М., 1988.

2. *Antipin A.S.* // Operations Research Proceedings / Eds.: U. Derigs, A. Bachem, A. Drexl. 1994. Berlin; Heidelberg. P. 11–15.

3. *Антипин А.С.* // Журн. вычислит. математики и мат. физики. 1995. Т. 35, No.5. С. 688–704.

4. *Антипин А.С.* // Дифференциальные уравнения. 1992. Т.28. No.11. С. 1846–1861.

5. *Антипин А.С.* // Автоматика и телемеханика. 1994. No.3. С. 12–23.

6. *Nikaido H., Isoda K.* Note on noncooperative convex games.// Pacific Journal of Math. 1955. V.5. Suppl.1. P. 807–815.

7. *Aubin J.-P., Ekeland I.* Applied Nonlinear Analysis. John Wiley & Sons. N-Y. 1984.

8. *Aubin J.-P., Frankowska H.* Set Valued Analysis. Birkhauser. Boston; Basel; Berlin. 1990.

9. *Поляк Б.Т.* Введение в оптимизацию. М., 1983.

10. *Емельянов С.В., Коровин С.К.* // Динамика неоднородных систем: Труды сем. ВНИИСИ. М., 1982. С. 5–27.

11. *Антипин А.С.* // Докл. РАН. 1995. Т.342. No.3. С. 1–4.