УДК 517.957+517.988+517.977.56

© *А. В. Чернов*

О ПРИМЕНИМОСТИ ТЕХНИКИ ПАРАМЕТРИЗАЦИИ УПРАВЛЕНИЯ К РЕШЕНИЮ РАСПРЕДЕЛЕННЫХ ЗАДАЧ ОПТИМИЗАЦИИ[1]

Изучаются аппроксимирующие конечномерные задачи математического программирования, возникающие в результате кусочно-постоянной дискретизации управления (в рамках техники параметризации управления) при оптимизации распределенных систем достаточно широкого класса. Устанавливается непрерывность по Липшицу градиентов функций аппроксимирующих задач; приводятся соответствующие формулы градиентов, использующие аналитическое решение исходной управляемой системы и сопряженной к ней системы и тем самым обеспечивающие возможность алгоритмического разделения проблемы оптимизации и проблемы решения управляемой начально-краевой задачи. Применение к численному решению задач оптимизации иллюстрируется на примере задачи Коши–Дарбу, управляемой по интегральному критерию. Приводятся результаты численного решения соответствующей аппроксимирующей задачи в системе MatLab с помощью программы `fmincon`, а также авторской программы, реализующей метод условного градиента. Кроме того, рассматривается задача безусловной минимизации, получаемая из аппроксимирующей задачи с ограничениями методом синус-параметризации. Приводятся результаты численного решения указанной задачи в системе MatLab с помощью программы `fminunc`, а также авторских программ, реализующих методы наискорейшего спуска и BFGS. Результаты численных экспериментов подробно анализируются.

Ключевые слова: оптимизация систем с распределенными параметрами, дифференцирование функционала, кусочно-постоянная аппроксимация управления, техника параметризации управления.

Введение

В работе [1] (см. также [2, глава IV, § 8]) для одной нелинейной управляемой системы уравнений в частных производных первого порядка был предложен способ сведения задачи оптимизации к конечномерной задаче математического программирования путем конечномерной аппроксимации только управления. Такой способ позволяет довольно естественным образом алгоритмически разделять проблему решения управляемой системы и собственно проблему оптимизации. Там же было предложено для вычисления градиента целевой функции (или функции, задающей ограничение) *аппроксимирующей задачи* (АЗ) использовать решение сопряженной системы; в результате трудоемкость вычисления градиента оказывается сопоставимой с трудоемкостью вычисления самой функции[2]. Позднее под названием *техника параметризации управления* (*control parametrization technique*) получил распространение подобный способу [1] эффективный подход к численному решению сосредоточенных задач оптимального управления (см., например, [4]). Для распределенных систем зачастую использовался метод *пространственно модального разложения*[3] (*spatial modal expansion*) с последующим применением «сосредоточенной» техники параметризации управления (теоретические основания которой на данный момент достаточно хорошо разработаны), см., например, [5]. В определенном смысле нечто похожее мы наблюдаем и в [1]: левые части системы уравнений там были частными производными первого порядка неизвестных функций; в численном отношении решение такой системы аналогично решению системы обыкновенных дифференциальных уравнений.

[1]Работа выполнена при финансовой поддержке Минобрнауки РФ в рамках государственного задания на оказание услуг в 2012–2014 гг. подведомственными высшими учебными заведениями (шифр заявки 1.1907.2011).

[2]Сейчас подобные алгоритмы иногда называют *быстрыми* [3].

[3]Метод заключается в сведении распределенной системы к системе обыкновенных дифференциальных уравнений за счет дискретизации по пространственным переменным. В отечественной литературе более распространены термины *метод пространственного квантования* и *метод прямых* [2, глава IV, §§ 2, 4].

Вместе с тем учет специфики управляемой *начально-краевой задачи* (НКЗ) именно как распределенной системы может весьма существенно повысить скорость численного решения. Так, например, если при решении задачи Коши–Дарбу методом конечных разностей раскладывать сетку на слои, ориентированные вдоль кривой начальных данных, и решать задачу послойно, то, скажем, в системе MatLab можно использовать матричные операции вместо двойных циклов. В результате скорость вычислений возрастает в 5–10 раз.

В связи с вышесказанным представляется актуальным всестороннее изучение применимости техники параметризации управления непосредственно к распределенным системам.

Отметим, что использование численных методов первого порядка условной конечномерной оптимизации для решения АЗ в рамках техники параметризации управления может быть оправдано только в том случае, когда функции АЗ обладают определенными свойствами гладкости. Соответствующие теоремы сходимости доказываются, как правило, в предположении непрерывности по Липшицу градиентов указанных функций. Для распределенных систем, ввиду их большого разнообразия, при доказательстве упомянутого свойства удобным оказывается описание управляемых систем с помощью тех или иных функционально-операторных уравнений (см., например, [6, с. 265–266], [7, 8], [9, с. 169–170], [10–13]).

В работе [14] изучалась аппроксимирующая конечномерная задача математического программирования, возникающая в результате простейшей кусочно-постоянной дискретизации управления при оптимизации нелинейных распределенных систем, представимых операторным уравнением типа Гаммерштейна[4] с нелинейной частью в виде зависящего от управления оператора внутренней суперпозиции. Как показывают примеры [10–13], указанный класс распределенных систем достаточно широк — к операторным уравнениям такого типа естественным образом сводятся разнообразные НКЗ для управляемых полулинейных гиперболических и параболических уравнений в частных производных. В качестве множества допустимых управлений в [14] принималось множество измеримых существенно ограниченных s-вектор-функций, принимающих значения в заданном параллелепипеде. Рассматривался способ дискретизации, превращающий функционалы исходной задачи оптимизации в обычные функции многих переменных, определенные на указанном параллелепипеде. В описанной ситуации в [14] были получены достаточные условия гладкости функций АЗ. Вместе с тем распространенным требованием, обеспечивающим сходимость численных методов оптимизации первого порядка и позволяющим оценивать скорость сходимости, является липшицевость градиента целевой функции. В связи с этим в данной статье мы продолжаем исследование, начатое в [14], и для той же АЗ устанавливаем условия, обеспечивающие непрерывность по Липшицу градиентов функций конечномерной АЗ математического программирования и приводим формулы указанных градиентов, использующие аналитическое решение исходной управляемой системы и сопряженной к ней. Применение к численному решению задач оптимизации иллюстрируется на примере задачи Коши–Дарбу, управляемой по интегральному критерию; подробно анализируются результаты численных экспериментов[5]. Отметим, что использование аналитического решения исходной и сопряженной систем в формулах градиентов позволяет в ходе вычислений алгоритмически разделить проблему решения АЗ оптимизации и проблему численного решения управляемой системы. А именно, управляемую НКЗ численно можно решать, вообще говоря, любым подходящим способом: программа, вычисляющая значение функции АЗ и ее производных по дискретизированному управлению, вызывает подпрограмму, решающую исходную НКЗ для данного управления, а также сопряженную к ней. Конкретная реализация этой подпрограммы может быть разной. В частности, можно использовать готовые математические пакеты. Мы использовали собственную подпрограмму на языке MatLab, реализующую метод конечных разностей. Достоинством этой подпрограммы (важным с учетом необходимости ее

[4] См., например, [15, глава VI, п. 19.2].

[5] Аналогичные численные эксперименты проводились автором и для оптимизационных задач, связанных с полулинейными уравнениями колебаний струны, теплопроводности, системы Гурса–Дарбу и др. Из-за ограниченного объема статьи мы останавливаемся для иллюстрации лишь на оптимизационной задаче для системы Коши–Дарбу, которая представляет, на наш взгляд, определенный самостоятельный интерес.

многократного вызова) явилась достаточно высокая скорость вычислений, которой удалось добиться за счет применения матричных операций MatLab вместо двойных циклов. В свою очередь, возможность применения матричных операций была основана на последовательном решении задачи Коши–Дарбу по наклонным («диагональным») слоям, ориентированным вдоль кривой начальных данных.

Автор выражает искреннюю признательность В. И. Сумину за неоднократные полезные обсуждения материала статьи.

§ 1. Постановка задачи оптимизации

Далее, $n, m, \ell, s \in \mathbb{N}$, $p \in [1, \infty)$, $q \in [1, \infty]$ — заданные числа, $q \geqslant p$; $\Pi \subset \mathbb{R}^n$ — измеримое (в смысле Лебега) ограниченное множество; $\mathcal{U} = L_\infty(\Pi)$, $\mathcal{D}_\infty \equiv \Big\{ u \in \mathcal{U}^s : u_i(t) \in \in [\alpha_i; \beta_i]$ для п.в. $t \in \Pi$, $i = \overline{1, s} \Big\}$ — множество допустимых управлений в исходной бесконечномерной задаче; $\mathcal{D} \subset \mathcal{D}_\infty$ — множество допустимых аппроксимаций управления (подробнее об этом см. ниже); $\mathcal{Z} = L_p(\Pi)$, $\mathcal{X} = L_q(\Pi)$, $\mathcal{Z}_\mathcal{X} = L_\sigma(\Pi)$, $q^{-1} + \sigma^{-1} = p^{-1}$ (в частности, при $q = p$ имеем $\sigma = \infty$). Определим класс $\mathbb{F}(\ell, s, m; \mathcal{X}, \mathcal{Z})$ всех функций $f(t, y, u) : \Pi \times \mathbb{R}^\ell \times \mathbb{R}^s \to \mathbb{R}^m$, каждая из которых непрерывно дифференцируема по переменным $y \in \mathbb{R}^\ell$, $u \in \mathbb{R}^s$ и вместе с производными измерима по $t \in \Pi$ и непрерывна по $\{y; u\} \in \mathbb{R}^\ell \times \mathbb{R}^s$ и такова, что:

F$_1$) $f\big(., y(.), u(.)\big) \in \mathcal{Z}^m$ $\forall y \in \mathcal{X}^\ell$, $u \in L^s_\infty(\Pi)$;

F$_2$) $f'_y\big(., y(.), u(.)\big) \in \mathcal{Z}^{m \times \ell}_\mathcal{X}$, $f'_u\big(., y(.), u(.)\big) \in \mathcal{Z}^{m \times s}$ $\forall \{y, u\} \in \mathcal{X}^\ell \times \mathcal{U}^s$;

F$_3$) существует неубывающая функция $\mathcal{N} : \mathbb{R}^+ \to \mathbb{R}^+$ такая, что

$$\max\Big\{ \big\| f'_y(., y, u) - f'_y(., z, u) \big\|_{\mathcal{Z}^{m \times \ell}_\mathcal{X}}, \big\| f'_v(., y, u) - f'_v(., z, u) \big\|_{\mathcal{Z}^{m \times s}} \Big\} \leqslant \mathcal{N}(M)\, \|y - z\|_{\mathcal{X}^\ell}$$

$$\forall y, z \in \mathcal{X}^\ell,\ u \in \mathcal{D},\ \|y\|, \|z\| \leqslant M;$$

F$_4$) существует неубывающая функция $\mathcal{N}_1 : \mathbb{R}^+ \to \mathbb{R}^+$ такая, что

$$\max\Big\{ \big\| f'_y(., y, u) - f'_y(., y, \overline{u}) \big\|_{\mathcal{Z}^{m \times \ell}_\mathcal{X}}, \big\| f'_v(., y, u) - f'_v(., y, \overline{u}) \big\|_{\mathcal{Z}^{m \times s}} \Big\} \leqslant \mathcal{N}_1(M)\, \|u - \overline{u}\|_{\mathcal{U}^s}$$

$$\forall y \in \mathcal{X}^\ell,\ u, \overline{u} \in \mathcal{D},\ \|y\| \leqslant M.$$

Через $\mathcal{X}^+$ обозначаем конус неотрицательных функций в пространстве $\mathcal{X}$. Для векторов $a, b \in \mathbb{R}^\ell$, $a \leqslant b$ (векторное неравенство понимаем покомпонентно), используем обозначение $[a; b] \equiv [a_1; b_1] \times \cdots \times [a_\ell; b_\ell]$. Модуль вектора понимаем как сумму модулей компонент.

Будем рассматривать следующее функционально-операторное уравнение, являющееся удобной формой описания широкого класса управляемых распределенных систем, см. [10–13]:

$$x(t) = \theta(t) + A\Big[f\big(., x(.), u(.)\big) \Big](t), \quad t \in \Pi, \quad x \in \mathcal{X}^\ell. \tag{1.1}$$

Здесь $u \in \mathcal{D}_\infty$ — управление; $\theta \in \mathcal{X}^\ell$ — заданный элемент; $A : \mathcal{Z}^m \to \mathcal{X}^\ell$ — заданный *линейный ограниченный оператор* (ЛОО); $f \in \mathbb{F}(\ell, s, m; \mathcal{X}, \mathcal{Z})$ — заданная функция. Относительно ЛОО A сделаем следующие предположения, часто выполняющиеся в приложениях, связанных с эволюционными уравнениями (см., например, [10–12, 16]):

A$_1$) для всякого $y \in \mathcal{Z}^{m \times \ell}_\mathcal{X}$ ЛОО $A_{\sim y} : \mathcal{X}^\ell \to \mathcal{X}^\ell$ и $A_{y\sim} : \mathcal{Z}^m \to \mathcal{Z}^m$, определяемые соответственно формулами $A_{\sim y}[x] = A[yx]$, $x \in \mathcal{X}^\ell$; $A_{y\sim}[z] = yA[z]$, $z \in \mathcal{Z}^m$, являются *квазинильпотентными*, то есть их спектральные радиусы равны нулю: $\rho\big(A_{\sim y}\big) = \rho\big(A_{y\sim}\big) = 0$;

A$_2$) ЛОО $A : \mathcal{Z}^m \to \mathcal{X}^\ell$ имеет положительную мажоранту[6] $B : \mathcal{Z} \to \mathcal{X}$ такую, что для всякого $y \in \mathcal{Z}_{\mathcal{X}}$ оператор $B_y : \mathcal{X} \to \mathcal{X}$, определяемый формулой $B_y[x] = B[yx]$, $x \in \mathcal{X}$, квазинильпотентен.

Относительно уравнения (1.1) предполагаем следующее:

H$_1$) уравнение (1.1) имеет единственное решение $x_u \in \mathcal{X}^\ell$ $\forall u \in \mathcal{D}_\infty$;

H$_2$) существуют функции $u_* \in \mathcal{U}$, $x_* \in \mathcal{X}$ такие, что: $|u(t)| \leqslant u_*(t)$, $|x_u(t)| \leqslant x_*(t)$ для всех $u \in \mathcal{D}_\infty$.

Укажем простое достаточное условие выполнения предположений **H**$_1$), **H**$_2$) (см. [10, теорема 1.1][7]):

H) существует функция $\varphi(t,\xi) : \Pi \times \mathbb{R}^+ \to \mathbb{R}^+$, измеримая по $t \in \Pi$, непрерывная и не убывающая по $\xi \in \mathbb{R}^+$ и такая, что имеет место оценка $\Big|f\big(.,y(.),u(.)\big)\Big| \leqslant \varphi\big(.,|y(.)|\big) \in \mathcal{Z}$ $\forall y \in \mathcal{X}^\ell$, $u \in \mathcal{D}_\infty$, и, кроме того, разрешимо *мажорантное уравнение*

$$x(t) = |\theta(t)| + B\Big[\varphi\big(.,x(.)\big)\Big](t), \quad t \in \Pi, \quad x \in \mathcal{X}^+.$$

При выполнении предположения **H**$_1$) для каждого $u \in \mathcal{D}_\infty$ однозначно определено $x_u \in \mathcal{X}^\ell$ — решение уравнения (1.1), отвечающее управлению $u \in \mathcal{D}_\infty$. Конструктивные достаточные условия разрешимости мажорантного уравнения получены в [11, § 4]. Будем рассматривать целевой функционал вида $J[u] = \Phi\Big[F\big(.,x_u(.),u(.)\big)\Big]$, где $\Phi \in \big(\widehat{\mathcal{Z}}^{\hat m}\big)^*$, $\widehat{\mathcal{Z}}$ — лебегово пространство функций на Π с индексом суммируемости из $[1;+\infty)$ и такое, что $\mathcal{X} \subset \widehat{\mathcal{Z}}$; $F \in \mathbb{F}(\ell, s, \hat m; \mathcal{X}, \widehat{\mathcal{Z}})$.

§ 2. Формулировка основного результата

Пусть $\kappa \in \mathbb{N}$ — заданное число, причем множество Π представлено в виде дизъюнктного объединения $\Pi = \bigsqcup\limits_{j=1}^{\kappa} \Pi_j$. Положим $\mu = s\,\kappa$. В качестве множества допустимых аппроксимаций управления (иначе говоря, допустимого множества аппроксимирующей задачи) примем

$$\mathcal{D} = \Big\{u \in L^s_\infty(\Pi) : u_i(t) \equiv w_{ij} \in [\alpha_i;\beta_i],\, t \in \Pi_j,\, i = \overline{1,s},\, j = \overline{1,\kappa}\Big\}. \tag{2.1}$$

Обозначим через W множество μ-мерных векторов $w = \big\{w_{11},\ldots,w_{1\kappa};\ldots;w_{s1},\ldots,w_{s\kappa}\big\}$ со свойствами $w_{ij} \in [\alpha_i;\beta_i]$, $i = \overline{1,s}$, $j = \overline{1,\kappa}$. Формула (2.1) устанавливает взаимно однозначное соответствие между множествами $\mathcal{D}$ и W. Таким образом, на множестве $\mathcal{D}$ функционал $J[u]$ можно рассматривать как функцию μ переменных, которую мы условимся обозначать $J\{w\}$, $w \in W$.

Представленная далее формула для градиента $\nabla J\{w\}$ (лемма 2) связана с сопряженным уравнением

$$\Psi - A^*\Lambda_u^*\Psi = A^*\widehat{\Lambda}_u^*\Phi, \quad \Psi \in \Big(\mathcal{Z}^m\Big)^*, \tag{2.2}$$

в котором $\Lambda_u : \mathcal{X}^\ell \to \mathcal{Z}^m$ — оператор умножения на функцию $f'_x(.,x_u,u)$; $\widehat{\Lambda}_u : \mathcal{X}^\ell \to \widehat{\mathcal{Z}}^{\hat m}$ — оператор умножения на функцию $F'_x(.,x_u,u)$.

Для более общего случая банаховых идеальных пространств в [18, теорема 1.3] доказано следующее утверждение.

Лемма 1. *При любом $u \in \mathcal{D}_\infty$ уравнение (2.2) имеет единственное в $\Big(\mathcal{Z}^m\Big)^*$ решение. Это решение дается формулой $\Psi_u = \Phi\widehat{\Lambda}_u R(A\Lambda_u)A$, где $R(A\Lambda_u) = \sum\limits_{k=0}^{\infty}(A\Lambda_u)^k = (I - A\Lambda_u)^{-1}$.*

---

[6] То есть $B : \mathcal{Z} \to \mathcal{X}$ — ЛОО, $B[\mathcal{Z}^+] \subset \mathcal{X}^+$ и $\big|A[z]\big| \leqslant B\big[|z|\big]$ для всех $z \in \mathcal{Z}^m$.

[7] Строго говоря, единственность решения из результатов [10] не следует, однако устанавливается аналогично [10, теорема 2.2]. Отличие заключается только в том, что на последнем этапе доказательства вместо серии оценок на вольтерровой цепочке нужно воспользоваться условием **A**$_2$) и обобщенной леммой Гронуолла [17, теорема 1.9.3].

В качестве очевидного следствия теоремы 1.2 из [18] в совокупности с теоремой Рисса о представлении линейного непрерывного функционала в лебеговом пространстве получаем следующее утверждение.

Лемма 2. *Функция $J\{w\}$ имеет на W производную по любому вектору $h \in \mathbb{R}^{\mu}$ при условии $w + h \in W$, которая описывается формулами*

$$\frac{\partial}{\partial h} J\{w\} = \Big(\nabla J\{w\}, h\Big), \quad \nabla J\{w\} = \left\{\frac{\partial}{\partial w_{11}} J\{w\}, \ldots, \frac{\partial}{\partial w_{s\kappa}} J\{w\}\right\},$$

$$\frac{\partial}{\partial w_{ij}} J\{w\} = \Psi_u\Big[f'_{v_i}(.,x_u,u)\chi_j\Big] + \Phi\Big[F'_{v_i}(.,x_u,u)\chi_j\Big], \tag{2.3}$$

где $i = \overline{1,s}$, $j = \overline{1,\kappa}$; χ_j — характеристическая функция множества Π_j.

Более того, существует функция $\gamma : \mathbb{R}^+ \to \mathbb{R}^+$, бесконечно малая в окрестности нуля и такая, что для остатка в формуле приращения

$$J\{w + \lambda h\} - J\{w\} = \lambda\Big(\nabla J\{w\}, h\Big) + r_{\lambda}(w,h) \tag{2.4}$$

справедлива равномерная оценка $\Big|r_{\lambda}(w,h)\Big| \leqslant \lambda \cdot \gamma(\lambda)$ для всех $\lambda \in [0;1]$, w, $w + h \in W$.

Теперь, наконец, можем сформулировать основной результат данной статьи.

Теорема 1. *Пусть выполнены все сделанные выше предположения. Тогда градиент $\nabla J\{w\}$ удовлетворяет условию Липшица на множестве W.*

§ 3. Вспомогательные утверждения

Лемма 3. *Существуют функции $x_* \in \mathcal{X}^+$, $u_* \in \mathcal{U}^+$, $z_* \in \mathcal{Z}^+$, $\xi_* \in \mathcal{Z}^+_{\mathcal{X}}$, $\eta_* \in \mathcal{Z}^+$ такие, что для всякого $u \in \mathcal{D}_{\infty}$ справедливы оценки*

$$|x_u| \leqslant x_*, \quad |u| \leqslant u_*, \quad \big|f(.,x_u,u)\big| \leqslant z_*, \quad \big|f'_x(.,x_u,u)\big| \leqslant \xi_*, \quad \big|f'_v(.,x_u,u)\big| \leqslant \eta_*.$$

Аналогично, существуют функции $\widehat{z}_* \in \widehat{\mathcal{Z}}^+$, $\widehat{\xi}_* \in \widehat{\mathcal{Z}}^+_{\mathcal{X}}$, $\widehat{\eta}_* \in \widehat{\mathcal{Z}}^+$ такие, что $\big|F(.,x_u,u)\big| \leqslant \widehat{z}_*$, $\big|F'_x(.,x_u,u)\big| \leqslant \widehat{\xi}_*$, $\big|F'_v(.,x_u,u)\big| \leqslant \widehat{\eta}_*$ $\forall\, u \in \mathcal{D}_{\infty}$.

Д о к а з а т е л ь с т в о. Существование функций u_* и x_* — непосредственное следствие предположений $\mathbf{H}_1$), $\mathbf{H}_2$). Существование функции z_* очевидным образом следует из уже доказанных двух неравенств (с x_* и u_*), условий относительно функции f (Каратеодори и $\mathbf{F}_1$)) [18, лемма 3.1] и идеальности пространств $\mathcal{X}$ и $\mathcal{U}$. Существование функций ξ_* и η_* доказывается аналогично с использованием условия $\mathbf{F}_2$). Вторая часть утверждения леммы доказывается аналогично первой части (просто f заменяется на F). $\square$

Лемма 4. *Для любого $u \in \mathcal{D}_{\infty}$ оператор $R(A\Lambda_u) = \sum\limits_{k=0}^{\infty} (A\Lambda_u)^k$ является ЛОО в пространстве $\mathcal{X}^{\ell}$. Более того, существует число $R_* > 0$ такое, что $\big\|R(A\Lambda_u)\big\|_{\mathcal{X}^{\ell}\to\mathcal{X}^{\ell}} \leqslant R_*$ для всех $u \in \mathcal{D}_{\infty}$.*

Д о к а з а т е л ь с т в о. По лемме 3 и условию $\mathbf{A}_2$) имеем $\big\|R(A\Lambda_u)\big\|_{\mathcal{X}^{\ell}\to\mathcal{X}^{\ell}} \leqslant R_*$, где $R_* \equiv \sum\limits_{k=0}^{\infty} \|B^k_{\xi_*}\|_{\mathcal{X}\to\mathcal{X}}$. И так как спектральный радиус $\rho(B_{\xi_*}) = 0$, то число R_* конечно. $\square$

Лемма 5. *Существует число $\beta > 0$ такое, что для любых $u, \overline{u} \in \mathcal{D}_\infty$ справедлива оценка*

$$\big\|R(A\Lambda_u) - R(A\Lambda_{\overline{u}})\big\|_{\mathcal{X}^\ell \to \mathcal{X}^\ell} \leqslant \|B\|_{\mathcal{Z}\to\mathcal{X}}\, \beta S\, d_x(u, \overline{u}), \tag{3.1}$$

где $d_x(u,\overline{u}) \equiv \big\|f'_x(.,x_u,u) - f'_x(.,x_{\overline{u}},\overline{u})\big\|_{\mathcal{Z}_{\mathcal{X}}^{m\times\ell}}$, $S = \sum\limits_{k=1}^{\infty} 4^k \|B_{\xi_*}^{r(k)}\|_{\mathcal{X}\to\mathcal{X}}$ — *сумма сходящегося ряда,* $r(k) = \left[\dfrac{k-1}{2}\right]$.

Д о к а з а т е л ь с т в о. Очевидно, что

$$\big\|R(A\Lambda_u) - R(A\Lambda_{\overline{u}})\big\|_{\mathcal{X}^\ell \to \mathcal{X}^\ell} \leqslant \sum_{k=1}^{\infty}\big\|(A\Lambda_u)^k - (A\Lambda_{\overline{u}})^k\big\|_{\mathcal{X}^\ell \to \mathcal{X}^\ell}\,. \tag{3.2}$$

Выберем любое число $k \in \mathbb{N}$. Заметим, что $(A\Lambda_u)^k = (A\Lambda_{\overline{u}} + A(\Lambda_u - \Lambda_{\overline{u}}))^k$. Отсюда получаем $\big\|(A\Lambda_u)^k - (A\Lambda_{\overline{u}})^k\big\|_{\mathcal{X}^\ell\to\mathcal{X}^\ell} \leqslant \sum\limits_{\nu=1}^{k} C_k^\nu \delta_{k,\nu}$, где $\delta_{k,\nu}$ — оценка нормы всевозможных операторных произведений, в которые множитель $A(\Lambda_u - \Lambda_{\overline{u}})$ входит ν раз, а множитель $A\Lambda_{\overline{u}}$ соответственно $(k-\nu)$ раз. Согласно условию $\mathbf{A}_2$), для любого $y \in \mathcal{X}^\ell$ справедлива оценка

$$\big|A(\Lambda_u - \Lambda_{\overline{u}})[y]\big| \leqslant B\Big[\big|f'_x(.,x_u,u) - f'_x(.,x_{\overline{u}},\overline{u})\big|\,|y|\Big].$$

По лемме 3 $\big|A(\Lambda_u - \Lambda_{\overline{u}})[y]\big| \leqslant 2\,B_{\xi_*}\big[|y|\big]$, $\big|A\Lambda_{\overline{u}}[y]\big| \leqslant B_{\xi_*}\big[|y|\big]$. Поэтому можно считать, что $\delta_{k,\nu} = \|B\|_{\mathcal{Z}\to\mathcal{X}} d_x(u,\overline{u})\, \sigma_k\, 2^\nu$, где величина $\sigma_k = \max\limits_{\alpha=\overline{1,k}} \|B_{\xi_*}^{\alpha-1}\|_{\mathcal{X}\to\mathcal{X}} \|B_{\xi_*}^{k-\alpha}\|_{\mathcal{X}\to\mathcal{X}}$. По условию $\mathbf{A}_2$) последовательность $\|B_{\xi_*}^k\|_{\mathcal{X}\to\mathcal{X}} \to 0$ при $k\to\infty$, следовательно, ограничена некоторым числом $\sqrt{\beta} > 0$. Заметив, что $\sigma_k \leqslant \beta\, \|B_{\xi_*}^{r(k)}\|_{\mathcal{X}\to\mathcal{X}}$, получаем

$$\big\|(A\Lambda_u)^k - (A\Lambda_{\overline{u}})^k\big\|_{\mathcal{X}^\ell\to\mathcal{X}^\ell} \leqslant \|B\|_{\mathcal{Z}\to\mathcal{X}}\, d_x(u,\overline{u})\, \beta\, \|B_{\xi_*}^{r(k)}\|_{\mathcal{X}\to\mathcal{X}}\, 2^k \sum_{\nu=0}^{k} C_k^\nu, \quad k \in \mathbb{N}.$$

С учетом (3.2) и равенства $\sum\limits_{\nu=0}^{k} C_k^\nu = 2^k$ приходим к (3.1). Заметим, что

$$\lim_{k\to\infty} \sqrt[k]{4^k\, \|B_{\xi_*}^{r(k)}\|_{\mathcal{X}\to\mathcal{X}}} = 4 \lim_{k\to\infty} \sqrt[k]{\|B_{\xi_*}^{r(k)}\|_{\mathcal{X}\to\mathcal{X}}} = 4\sqrt{\rho(B_{\xi_*})} = 0 < 1$$

в силу условия $\mathbf{A}_2$). Поэтому согласно признаку Коши сходимости ряда сумма S конечна. $\square$

Лемма 6. *Найдется $\Psi_* > 0$ такое, что $\|\Psi_u\|_{(\mathcal{Z}^m)^*} \leqslant \Psi_*$ $\forall u \in \mathcal{D}_\infty$.*

Д о к а з а т е л ь с т в о. Из лемм 3, 4 и неравенства Гёльдера получаем

$$\|\Psi_u\|_{(\mathcal{Z}^m)^*} \leqslant \|\Phi\|\, \|\widehat{\xi}_*\|_{\widehat{\mathcal{Z}}_{\mathcal{X}}}\, \big\|R(A\Lambda_u)\big\|_{\mathcal{X}^\ell\to\mathcal{X}^\ell}\, \|A\|_{\mathcal{Z}^\ell\to\mathcal{X}^\ell} \leqslant \Psi_*\,,$$

где $\Psi_* \equiv \|\Phi\|\, \|\widehat{\xi}_*\|_{\widehat{\mathcal{Z}}_{\mathcal{X}}}\, R_*\, \|A\|_{\mathcal{Z}^\ell\to\mathcal{X}^\ell}$. $\square$

Лемма 7. *Существует константа $\sigma_* > 0$ такая, что*

$$\|\Psi_u - \Psi_{\overline{u}}\|_{(\mathcal{Z}^m)^*} \leqslant \sigma_*\Big[d_x(u,\overline{u}) + D_x(u,\overline{u})\Big] \quad \forall u, \overline{u} \in \mathcal{D}_\infty,$$

где $D_x(u,\overline{u}) \equiv \big\|F'_x(.,x_u,u) - F'_x(.,x_{\overline{u}},\overline{u})\big\|_{\widehat{\mathcal{Z}}_{\mathcal{X}}^{\widehat{m}\times\ell}}$.

Доказательство. Оценим

$$\|\Psi_u - \Psi_{\overline{u}}\|_{(\mathcal{Z}^m)^*} = \left\|\Phi\widehat{\Lambda}_u R(A\Lambda_u)A - \Phi\widehat{\Lambda}_{\overline{u}} R(A\Lambda_{\overline{u}})A\right\|_{(\mathcal{Z}^m)^*} \leqslant$$

$$\leqslant \left\|\Phi\widehat{\Lambda}_u\big\{R(A\Lambda_u) - R(A\Lambda_{\overline{u}})\big\}A\right\|_{(\mathcal{Z}^m)^*} + \left\|\Phi\big\{\widehat{\Lambda}_u - \widehat{\Lambda}_{\overline{u}}\big\}R(A\Lambda_{\overline{u}})A\right\|_{(\mathcal{Z}^m)^*}.$$

Дальнейшее очевидно в силу лемм 3, 5 и 4. □

Лемма 8. *Существует положительный ЛОО $B_* : \mathcal{Z} \to \mathcal{X}$ такой, что для всех управлений $u, \overline{u} \in \mathcal{D}_\infty$ и отвечающих им решений $x_u, x_{\overline{u}}$ уравнения* (1.1) *справедлива оценка* $\Big|x_u - x_{\overline{u}}\Big| \leqslant B_*\Big[\big|\Delta_u f(x_{\overline{u}})\big|\Big]$, *где* $\Delta_u f(x_{\overline{u}}) = f\Big(.,x_{\overline{u}},u\Big) - f\Big(.,x_{\overline{u}},\overline{u}\Big)$.

Доказательство следует непосредственно из [18, теорема 2.1]. □

Лемма 9. *Существует положительный ЛОО $G_* : \mathcal{U} \to \mathcal{X}$ такой, что для всех управлений $u, \overline{u} \in \mathcal{D}_\infty$ и отвечающих им решений $x_u, x_{\overline{u}}$ уравнения* (1.1) *справедлива оценка* $\Big|x_u - x_{\overline{u}}\Big| \leqslant G_*\big[|u - \overline{u}|\big]$, *и, следовательно,* $\Big\|x_u - x_{\overline{u}}\Big\|_{\mathcal{X}^\ell} \leqslant L_*\,\|u - \overline{u}\|_{\mathcal{U}^s}$ *при* $L_* = \|G_*\|$.

Доказательство. Обозначим $\Delta u \equiv u - \overline{u}$. В соответствии с леммой Адамара и леммой 3 можем оценить

$$\big|f(.,x_{\overline{u}},u) - f(.,x_{\overline{u}},\overline{u})\big| \leqslant \int_0^1 \big|f'_v(.,x_{\overline{u}},\overline{u} + \theta\Delta u)\big|\,d\theta\,|\Delta u| \leqslant \eta_*\,|\Delta u|\,.$$

Остается применить лемму 8 и определить ЛОО $G_* : \mathcal{U} \to \mathcal{X}$ формулой: $G_*[h] = B_*[\eta_*\,h]$, $h \in \mathcal{U}$. □

Лемма 10. *Существует число $K_* > 0$ такое, что для любых векторов $w, \overline{w} \in W$ и отвечающих им управлений $u, \overline{u} \in \mathcal{D}$ имеем $d_x(u,\overline{u})$, $D_x(u,\overline{u})$, $d_v(u,\overline{u})$, $D_v(u,\overline{u}) \leqslant K_*\,\|u - \overline{u}\|_{\mathcal{U}^s}$.*

Доказательство. Докажаем первую оценку, остальные доказываются аналогично. Оценим

$$d_x(u,\overline{u}) \equiv \big\|f'_x(.,x_u,u) - f'_x(.,x_{\overline{u}},\overline{u})\big\|_{\mathcal{Z}_{\mathcal{X}}^{m\times\ell}} \leqslant d^1_x(u,\overline{u}) + d^2_x(u,\overline{u}),$$

$$d^1_x(u,\overline{u}) \equiv \big\|f'_x(.,x_u,u) - f'_x(.,x_{\overline{u}},u)\big\|,\ d^2_x(u,\overline{u}) \equiv \big\|f'_x(.,x_{\overline{u}},u) - f'_x(.,x_{\overline{u}},\overline{u})\big\|.$$

В соответствии с леммой 3 выберем $M = \|x_*\|_{\mathcal{X}}$. Тогда согласно условию $\mathbf{F}_3$) и лемме 9 можем оценить

$$d^1_x(u,\overline{u}) \leqslant \mathcal{N}(M)\,\|x_u - x_{\overline{u}}\|_{\mathcal{X}^\ell} \leqslant \mathcal{N}(M)\,L_*\,\|u - \overline{u}\|_{\mathcal{U}^s}.$$

Кроме того, из условия $\mathbf{F}_4$) получаем $d^2_x(u,\overline{u}) \leqslant \mathcal{N}_1(M)\,\|u - \overline{u}\|_{\mathcal{U}^s}$. Дальнейшее очевидно. □

§ 4. Доказательство основного утверждения

В соответствии с леммой 2 имеем

$$\partial J_{ij} \equiv \left|\frac{\partial}{\partial w_{ij}}J\{w\} - \frac{\partial}{\partial w_{ij}}J\{\overline{w}\}\right| \leqslant \left|\Phi\Big[\big\{F'_{v_i}(.,x_u,u) - F'_{v_i}(.,x_{\overline{u}},\overline{u})\big\}\chi_j\Big]\right| +$$

$$+ \left|\Psi_u\Big[f'_{v_i}(.,x_u,u)\chi_j\Big] - \Psi_{\overline{u}}\Big[f'_{v_i}(.,x_{\overline{u}},\overline{u})\chi_j\Big]\right| \leqslant \|\Phi\|_{(\widehat{\mathcal{Z}}^{\hat{m}})_*}\,D_v(u,\overline{u}) +$$

$$+ \left|\big\{\Psi_u - \Psi_{\overline{u}}\big\}\Big[f'_{v_i}(.,x_u,u)\chi_j\Big]\right| + \left|\Psi_{\overline{u}}\Big[\big\{f'_{v_i}(.,x_u,u) - f'_{v_i}(.,x_{\overline{u}},\overline{u})\big\}\chi_j\Big]\right| \leqslant$$

$$\leqslant \|\Phi\|_{(\widehat{\mathcal{Z}}^{\hat{m}})_*}\,D_v(u,\overline{u}) + \|\Psi_u - \Psi_{\overline{u}}\|_{(\mathcal{Z}^m)_*}\,\Big\|\big|f'_{v_i}(.,x_u,u)\big|\Big\|_{\mathcal{Z}} + \|\Psi_{\overline{u}}\|_{(\mathcal{Z}^m)_*}\,d_v(u,\overline{u}).$$

Пользуясь леммами 7, 3, 6, получаем

$$\partial J_{ij} \leqslant \|\Phi\|_{(\widehat{\mathcal{Z}}^{\hat m})_*} D_v(u,\overline{u}) + \|\eta_*\|_{\mathcal{Z}}\, \sigma_* \left[d_x(u,\overline{u}) + D_x(u,\overline{u})\right] + \Psi_*\, d_v(u,\overline{u}).$$

Остается воспользоваться леммой 10. Теорема 1 доказана.

§ 5. Пример: задача Коши–Дарбу

Рассмотрим управляемую задачу Коши–Дарбу:

$$\begin{cases} x''_{t_1t_2}(t) = f\Big(t, x(t), u(t)\Big), \quad t = (t_1,t_2) \in \Pi = [0,T_1]\times[0,T_2]; \\ x_{t_1}\Big|_{t_2=v(t_1)} = \pi(t_1), \quad t_1\in[0,T_1]; \quad x_{t_2}\Big|_{t_1=v^{-1}(t_2)} = \chi(t_2), \quad t_2\in[0,T_2]; \\ x(\overline{t}) = a. \end{cases} \tag{5.1}$$

Будем считать, что $t_2 = v(t_1)$ — строго монотонная, абсолютно непрерывная функция, определенная на отрезке $[0,T_1]$, и такая, что $v(0)=T_2$, $v(T_1)=0$; $t_1 = v^{-1}(t_2)$ — функция, обратная к функции $t_2 = v(t_1)$; $\overline{t}\in\Pi$, $\overline{t}_2 = v(\overline{t}_1)$, — фиксированная точка на кривой начальных данных; $a\in\mathbb{R}$ — заданное число; функция $f\in\mathbb{F}(\ell,s,m;\mathcal{X},\mathcal{Z})$ при $\ell=m=1$, $\mathcal{X}=L_q(\Pi)$, $\mathcal{Z}=L_p(\Pi)$, $1\leqslant p<q<+\infty$; $\pi\in L_p[0;T_1]$, $\chi\in L_p[0;T_2]$. Решение задачи (5.1) будем понимать в смысле п.в. и искать его среди функций из $L_q(\Pi)$, имеющих первые и смешанную производную класса $L_p(\Pi)$. Обращая дифференциальный оператор этой задачи, приводим ее к виду

$$x(t_1,t_2) = a + \int_{\overline{t}_1}^{t_1}\pi(\xi)\,d\xi + \int_{\overline{t}_2}^{t_2}\chi(\xi)\,d\xi + \int_{v^{-1}(t_2)}^{t_1} d\xi_1 \int_{v(\xi_1)}^{t_2} f(\xi,x,u)\,d\xi_2.$$

Полученное уравнение имеет вид (1.1), где $w\in L_q(\Pi)$, $A: L_p(\Pi)\to L_q(\Pi)$ — функция и оператор, определяемые формулами

$$w(t_1,t_2) = a + \int_{\overline{t}_1}^{t_1}\pi(\xi)\,d\xi + \int_{\overline{t}_2}^{t_2}\chi(\xi)\,d\xi, \quad A[z](t_1,t_2) = \int_{v^{-1}(t_2)}^{t_1} d\xi_1\int_{v(\xi_1)}^{t_2} z(\xi)\,d\xi_2.$$

Проверим, что ЛОО A удовлетворяет условиям $\mathbf{A}_1$), $\mathbf{A}_2$) при $B=A$. Поскольку оператор A положительный, то, принимая $B=A$, нам достаточно лишь обосновать выполнение условия $\mathbf{A}_1$). Для этого воспользуемся [10, теорема 2.1]. Напомним соответствующие понятия.

Пусть $\Sigma=\Sigma(\Pi)$ — σ-алгебра всех измеримых по Лебегу подмножеств множества Π, P_H — оператор умножения на характеристическую функцию χ_H множества $H\in\Sigma$. Тогда система $\mathcal{B}(A) = \Big\{H\in\Sigma:\quad P_HAP_H = P_HA\Big\}$ называется *системой вольтерровых множеств* ЛОО $A:\mathcal{Z}^m\to\mathcal{X}^\ell$. При этом для числа $\delta>0$ подсистема $\mathcal{T} = \Big\{\emptyset = H_0\subset H_1\subset\ldots\subset H_k=\Pi\Big\}\subset$ $\subset\mathcal{B}(A)$ именуется *вольтерровой δ-малой по мере цепочкой множеств* ЛОО A, если $\operatorname{mes}(h)<\delta$ для всех $h = H_i\backslash H_{i-1}$, $i=\overline{1,k}$. Согласно [10, теорема 2.1], нам достаточно убедиться, что $\forall\delta>0$ ЛОО A обладает δ-малой по мере цепочкой множеств. Опишем ее построение. Прежде всего, для произвольного $\sigma>0$ выберем разбиение $0=\xi_0<\xi_1<\ldots<\xi_\kappa=T_1$, $\xi_i-\xi_{i-1}<\sigma$, $i=\overline{1,\kappa}$, и обозначим $\eta_j = v(\xi_{\kappa-j})$, $j=\overline{0,\kappa}$. Определим множества $h[i;j] = [\xi_{i-1},\xi_i]\times[\eta_{j-1},\eta_j]$, $i,j=\overline{1,\kappa}$. Положим $k=\kappa^2$ и для каждого $\nu=\overline{1,k}$ определим числа $k_1=k_1(\nu)$, $k_2=k_2(\nu)$ следующим образом. Обозначим $\gamma_1=[\nu/\kappa]$, $\gamma_2=\nu-\gamma_1\cdot\kappa$. Если $\gamma_2>0$, то берем $k_1=\gamma_2$, $k_2=\gamma_1+1$. В противном случае принимаем $k_1=\kappa$, $k_2=\gamma_1$. После этого определим множества $h_\nu = h[k_1,k_2]$, $\nu=\overline{1,k}$. Фактически h_ν — это множества $h[i,j]$, линейно упорядоченные от левого нижнего в направлении движения по горизонтальным рядам. Положим $\Pi_+\equiv\Big\{t\in$ $\in\Pi:\ t_2\geqslant v(t_1)\Big\}$, $\Pi_-\equiv\Big\{t\in\Pi:\ t_2<v(t_1)\Big\}$. Организуем множества $H_0^+=\emptyset$, $\mathcal{H}_i^+ = \bigcup\limits_{\nu=1}^{i} h_\nu$, $H_i^+ = \Pi_+\cap\mathcal{H}_i^+$, $i=\overline{1,k}$. Нетрудно понять, что множества H_i^+, $i=\overline{0,k}$, являются вольтерровыми

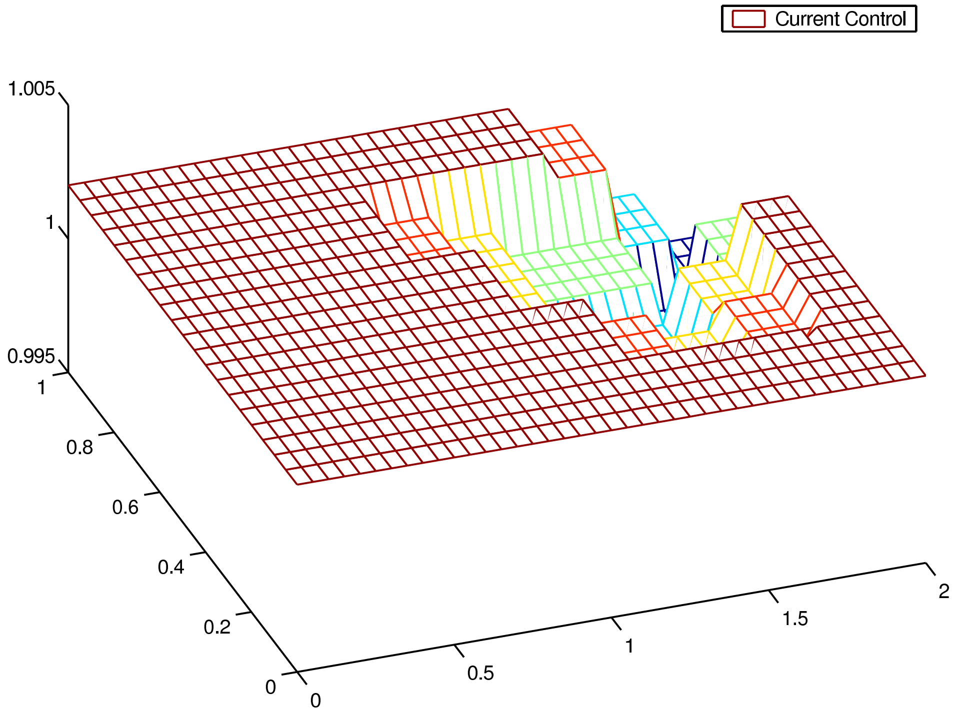

Рис. 1. Оптимальное управление (`fmincon`, $u_* = 7$, $\gamma = 6$)

множествами оператора A (в самом деле, значения оператора $A[z](t)$ при $t \in H_i^+$ зависят лишь от значений функции $z(\xi)$ при $\xi \in [0,t] \cap \Pi_+ \subset H_i^+$), содержатся в Π_+ и упорядочены по вложению. Аналогичным образом можно построить упорядоченную систему $\emptyset = H_0^- \subset H_1^- \subset \subset \ldots \subset H_k^- = \Pi_-$ вольтерровых множеств оператора A. Ясно, что каково бы ни было $\delta > 0$, всегда можно найти число $\sigma > 0$ так, что система

$$H_i = H_i^+, \quad i = \overline{0,k}, \quad H_i = \Pi_+ \cup H_{i-k}^-, \quad i = \overline{k+1, 2k}, \quad k = k(\sigma),$$

будет вольтерровой δ-малой по мере цепочкой оператора A. Тем самым условия $\mathbf{A}_1$), $\mathbf{A}_2$) выполняются. Условие $\mathbf{H}$) можно понимать как условие на функцию f. Будем считать его выполненным при $B = A$. В качестве функционала рассмотрим

$$J[u] = \int_\Pi F\big(t, x_u(t), u(t)\big)\, dt, \tag{5.2}$$

где x_u — решение задачи (5.1), отвечающее управлению u, а функция F принадлежит классу $\mathbb{F}(1, s, 1; \mathcal{X}, \mathcal{Z})$ при $\mathcal{X} = L_q(\Pi)$, $\mathcal{Z} = L_1(\Pi)$.

Воспользуемся результатами, сформулированными в § 2. В соответствии с формулой (2.3) можем записать

$$\nabla J\{w\} = \left\{ \frac{\partial}{\partial w_{11}} J\{w\}, \ldots, \frac{\partial}{\partial w_{s\kappa}} J\{w\} \right\}, \quad \frac{\partial}{\partial w_{ij}} J\{w\} = \int_{\Pi_j} G_i(t)\, dt,$$

$$G_i(.) = \psi_u(.) f'_{v_i}\Big(., x_u(.), u(.)\Big) + F'_{v_i}\Big(., x_u(.), u(.)\Big) \in L_1(\Pi),\ i = \overline{1,s},\ j = \overline{1,\kappa},$$

$\psi_u(.) \in L_{p'}(\Pi)$ $(p^{-1} + (p')^{-1} = 1)$ — единственное решение сопряженного уравнения (см. лемму 2; линейные непрерывные функционалы отождествляем с соответствующими функциями Рисса)

$$\psi(t) - A^*\Big[\Big(f'_y(., x_u, u)\Big)^* \psi\Big](t) = A^*\Big[\Big(F'_y(., x_u, u)\Big)^*\Big](t),$$

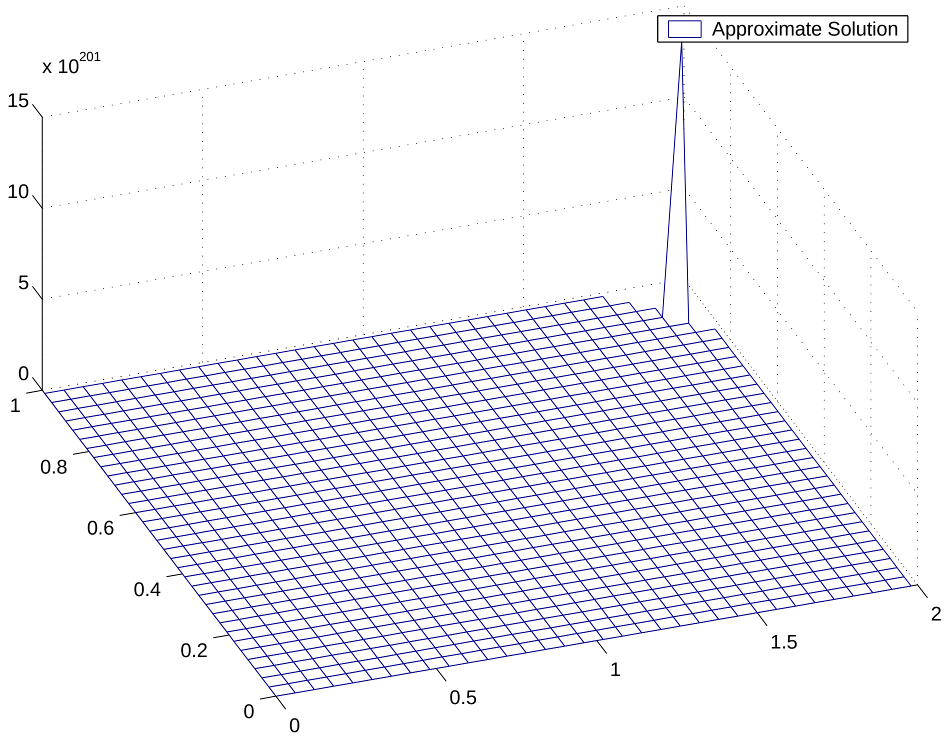

Рис. 2. Нарушение условия **H**)

в котором $A^* : L_{q'}(\Pi) \to L_{p'}(\Pi)$ — оператор, сопряженный к оператору A, то есть определяемый формулой

$$A^*[z](t) = \left\{ \int_{t_1}^{T^1} d\xi_1 \int_{t_2}^{T^2} z(\xi)\, d\xi_2,\ t \in \Pi_+; \quad \int_0^{t_1} d\xi_1 \int_0^{t_2} z(\xi)\, d\xi_2,\ t \in \Pi_- \right\}.$$

Понятно, что на самом деле функция ψ абсолютно непрерывна на каждом из множеств Π_-, Π_+ (с точностью до кривой начальных данных как множества нулевой меры). Поэтому сужение решения сопряженного уравнения на нижний треугольник Π_- можно понимать как решение задачи Гурса–Дарбу:

$$\begin{cases} \psi''_{t_1t_2}(t) = \Big(f'_y(t, x_u, u)\Big)^* \psi + \Big(F'_y(t, x_u, u)\Big)^*, \quad t \in \Pi_-; \\ \psi(t_1, 0) = 0, \quad t_1 \in [0, T_1]; \quad \psi(0, t_2) = 0, \quad t_2 \in [0, T_2). \end{cases} \tag{5.3}$$

Аналогичным образом сужение решения сопряженного уравнения на верхний треугольник Π_+ можно понимать как решение задачи

$$\begin{cases} \psi''_{t_1t_2}(t) = \Big(f'_y(t, x_u, u)\Big)^* \psi + \Big(F'_y(t, x_u, u)\Big)^*, \quad t \in \Pi_+; \\ \psi(t_1, T_2) = 0, \quad t_1 \in [0, T_1]; \quad \psi(T_1, t_2) = 0, \quad t_2 \in [0, T_2). \end{cases} \tag{5.4}$$

Кривая начальных данных может быть линией разрыва функции ψ. По теореме 1 градиент $\nabla J\{w\}$ удовлетворяет условию Липшица на множестве W.

§ 6. Результаты численных экспериментов

Рассмотрим следующий частный случай задачи (5.1):

$$T_1 = 2, \quad T_2 = 1, \quad f(t, x, u) = 2ux^3, \quad v(t_1) = 1 - 0.5\, t_1, \quad v^{-1}(t_2) = 2\,(1 - t_2),$$

Таблица 1. Использование программы `fmincon`

j	u_*	γ	`Iters`	`fEvals`	$\|\nabla J\|$	J
1	3	0	4	8	8.6609e-4	7.3885e-5
2	3	2	4	8	0.0020	9.2655e-5
3	3	3	5	10	0.0066	3.3650e-4
4	4	3	5	10	0.0066	3.3650e-4
5	5	4	8	16	6.5932e-5	6.9054e-5
6	6	5	10	20	2.9006e-4	6.9641e-5
7	7	6	15	30	0.0013	7.9313e-5
8	8	7	14	28	0.0024	1.0382e-4
9	9	8	24	48	0.0036	1.4954e-4
10	10	9	39	78	0.0062	3.1511e-4
11	1	0	4	8	0.0012	7.8413e-5
12	0.7	0	3	7	0.1703	0.1809

$$\pi(t_1) = (3 - 0.5\,t_1)^{-2}, \quad \chi(t_2) = (2 + t_2)^{-2}, \quad \overline{t} = (1, 0.5), \quad x(\overline{t}) = a = 0.4.$$

Зададим интегрант функционала (5.2) как $F(t,x,u) = \Big(x - \big(4 - t_1 - t_2\big)^{-1}\Big)^2 + (u-1)^2$. Возьмем $p = 1$, $q = 3$. Будем решать численно задачу оптимизации

$$J[u] = \int_\Pi F(t, x_u, u)\,dt \to \min_{\mathcal{D}_\infty}, \quad \mathcal{D}_\infty = \Big\{u \in L_\infty(\Pi) :\, u(t) \in [0; u_*]\Big\} \tag{6.1}$$

для различных случаев выбора верхней границы $u_* \in \{0.7, 2, 3, \ldots, 9\}$. Как и раньше, x_u — это решение задачи (5.1) (при указанном выборе параметров). Для $u_* \geqslant 1$ решение задачи оптимизации (6.1) очевидно: $\overline{u} = 1$, $\overline{x} = x_{\overline{u}} = \big(4 - t_1 - t_2\big)^{-1}$, $\overline{J} = J[\overline{u}] = 0$. Нас интересует здесь практическая работоспособность изучаемого подхода к ее численному решению.

Переходя к аппроксимирующей задаче, заменим множество $\mathcal{D}_\infty$ множеством конечномерных аппроксимаций $\mathcal{D} = \Big\{u \in L_\infty(\Pi) :\, u(t) = w_{ij} \in [0; u_*],\, t \in \Pi_{ij},\, i = \overline{1, \kappa_1},\, j = \overline{1, \kappa_2}\Big\}$, где $\kappa_1 = 10$, $\kappa_2 = 5$, $\Pi_{ij} = [(i-1)\,h; ih) \times [(j-1)\,h; jh)$, $h = 0.2$. Множество $\mathcal{D}$ отождествляется с множеством $W \subset \mathbb{R}^{50}$ всевозможных векторов $(w_{11}, \ldots, w_{\kappa_1 1}; \ldots; w_{1\kappa_2}, \ldots, w_{\kappa_1\kappa_2})$ с компонентами из $[0; u_*]$. Соответственно, функционал $J[u]$ преобразуется в функцию 50-ти переменных $J\{w\}$, $w \in W$. Вообще говоря, после этого можно использовать любую программу для численной минимизации функции $J\{w\}$ на множестве W. Воспользуемся, например, программой `fmincon.m` системы MatLab. Результаты выполнения программы для различных u_* и начальных приближений $u_0 \equiv \gamma$ при точности по аргументу `tolx=1e-5` и точности по функции `tolf=1e-4` приведены в таблице 1 (запись `e-5` означает умножение на 10^{-5}, `Iters` — количество итераций (вычислений градиента), `fEvals` — количество вычислений целевой функции).

Результаты использования программы `ch_cond_grad` (условного градиента) для минимизации функции $J\{w\}$ на множестве W при различных u_* и γ, точности по аргументу `tolx=1e-4` и точности по градиенту `tolg=0.01` приведены в таблице 2. При $u_* \geqslant 1$ оптимальное управление, полученное таким способом, в точности совпадает с теоретическим $\overline{u} \equiv 1$. При использовании программы `fmincon.m` ситуация чуть хуже, см. рисунок 1.

Отметим, что численное решение мажорантной задачи (соответствующей мажорантному уравнению) удается построить лишь при $u_* \leqslant 5.58 \equiv u_p$. При этом значение решения в точке $(T_1; T_2)$ оказывается равным $5.679 \cdot 10^{301}$. При $u_* > u_p$ численное решение мажорантной задачи разрушается в окрестности точки $(T_1; T_2)$, см. рисунок 2. Таким образом, выполнение условия **H**) гарантируется лишь при $u_* \leqslant u_p$. И в самом деле, для данных строк с седьмой по десятую таблиц 1, 2 решение управляемой системы разрушается уже на начальной итерации. В программе, реализующей целевую функцию АЗ, в случае когда обнаруживалось разрушение

Таблица 2. Использование программы `ch_cond_grad`

j	u_*	γ	Iters	fEvals	$\|\nabla J\|$	J	CurDelta
1	3	0	2	8	1.9872e-4	6.9135e-5	0.0018
2	3	2	2	9	1.9681e-4	6.9070e-5	0.0018
3	3	3	2	8	1.9872e-4	6.9135e-5	0.0018
4	4	3	2	8	1.9872e-4	6.9135e-5	0.0023
5	5	4	2	9	1.9676e-4	6.9078e-5	0.0030
6	6	5	2	9	2.0332e-4	6.9110e-5	0.0041
7	7	6	2	9	1.9687e-4	6.9091e-5	0.0041
8	8	7	2	10	1.9690e-4	6.9093e-5	0.0046
9	9	8	2	10	1.9711e-4	6.9101e-5	0.0051
10	10	9	2	10	1.9858e-4	6.9133e-5	0.0054
11	1	0	2	22	2.7066e-4	6.9354e-5	9.4750e-7
12	0.7	0	2	22	0.1703	0.1809	5.5645e-5

Таблица 3. Использование программы `fminunc`

j	u_*	γ	Iters	fEvals	$\|\nabla J\|$	J
1	3	0.001	53	67	2.5299e-4	6.9149e-5
2	3	2	4	11	0.0024	8.5600e-5
3	3	2.999	20	95	0.0014	7.4999e-5
4	4	3	4	11	0.0107	3.0618e-4
5	5	4	11	46	0.0027	8.0551e-5
6	6	5	13	51	0.0029	7.9128e-5
7	7	6	11	42	0.0037	8.3263e-5
8	8	7	16	68	0.0034	7.9483e-5
9	9	8	14	56	0.0224	4.4936e-4
10	10	9	14	54	0.0543	0.0024
11	1	0.001	4	14	2.1308e-5	7.1730e-5
12	0.7	0.001	7	17	0.0011	0.1809

численного решения управляемой системы (появлялись значения `Inf` или `NaN`), было предусмотрено формальное назначение некоторого достаточно большого числа в качестве значения функции. По идее, это должно было обеспечить своего рода защиту от сбоев в ходе одномерного поиска. В связи с этим интересны следующие результаты, полученные при $u_* > u_p$. При выборе начального приближения $\gamma > u_p$ при использовании программы `fmincon` для решения АЗ в некоторых случаях наблюдалось зацикливание, преждевременное прекращение вычислений и прочее. С другой стороны, метод условного градиента успешно работал и в такой ситуации. При выборе начального приближения $\gamma < u_p$ никаких сбоев при численном решении АЗ не наблюдалось. Вероятно, это можно объяснить тем, что поскольку мы двигаемся в направлении убывания целевой функции, то в некоторой окрестности текущего управления в ходе одномерного поиска удается найти точку–управление, которой отвечает приемлемое решение управляемой системы. В результате при хорошем начальном приближении метод оказывается работоспособным и в том случае, когда условие **H**) не выполнено.

Проводились также численные эксперименты с использованием известного *метода синус-параметризации*, позволяющего свести задачу условной оптимизации к задаче безусловной минимизации: $w = (d_1 + d_2)/2 + [(d_2 - d_1)/2] \sin\omega$, здесь синус от вектора $\omega \in \mathbb{R}^{50}$ понимается как вектор из синусов его компонент. Результаты численных экспериментов: для программы `fminunc` системы MatLab — см. таблицу 3 (строки 1–10: заданная точность вычислений

Таблица 4. Использование программы `ch_steepd`

j	u_*	γ	`Iters`	`fEvals`	$\|\nabla J\|$	J
1	3	0.001	2	18	0.0026	8.9841e-5
2	3	2	3	37	1.6397e-4	6.9033e-5
3	3	2.999	3	31	1.8737e-4	6.9059e-5
4	4	3	3	42	0.0014	7.2811e-5
5	5	4	3	39	0.0024	7.8222e-5
6	6	5	4	51	0.0028	7.8550e-5
7	7	6	8	134	1.9994e-4	6.9068e-5
8	8	7	11	196	8.3539e-4	6.9483e-5
9	9	8	11	202	5.3448e-5	6.8989e-5
10	10	9	8	143	5.1713e-4	6.9116e-5
11	1	0.001	2	22	8.0411e-7	6.9279e-5
12	0.7	0.001	2	17	3.8068e-4	0.1809

Таблица 5. Использование программы `ch_qnewton`

j	u_*	γ	`Iters`	`fEvals`	$\|\nabla J\|$	J
1	3	0.001	2	18	0.0026	8.9841e-5
2	3	2	4	37	1.3935e-4	6.9001e-5
3	3	2.999	3	35	0.0041	1.2240e-4
4	4	3	3	37	8.5659e-4	7.0439e-5
5	5	4	3	40	0.0033	8.6150e-5
6	6	5	5	72	5.5611e-4	6.9314e-5
7	7	6	10	157	0.0015	7.1549e-5
8	8	7	10	179	0.0039	8.2776e-5
9	9	8	13	233	2.0334e-4	6.8995e-5
10	10	9	11	201	0.0027	7.4114e-5
11	1	0.001	2	22	8.0411e-7	6.9279e-5
12	0.7	0.001	2	17	3.8068e-4	0.1809

`tolx=1e-3`, `tolf=2e-4`, остальные строки: `tolx=1e-4`, `tolf=2e-5`); для программы `ch_steepd` (метод наискорейшего спуска в авторской реализации) — см. таблицу 4 (здесь и дальше точность вычислений `tolx=1e-5`, `tolf=1e-4`, `tolg=5e-3`); для программы `ch_qnewton` (метод BFGS в авторской реализации) — см. таблицу 5.

Описанные выше численные эксперименты (и ряд других проведенных автором аналогичных экспериментов, связанных с полулинейными уравнениями колебаний струны, теплопроводности, системы Гурса–Дарбу и др.) позволяют сделать следующие *выводы* (см. также [14]).

1. Изучаемый подход достаточно прост и позволяет разделить исходную задачу на подзадачи (в частности, конечномерной оптимизации и решения НКЗ и сопряженной задачи), для каждой из которых можно использовать как собственные разработки, так и готовые программы, выбирая те, что наиболее эффективны или наиболее просты и доступны[8].

2. Метод синус-параметризации (см. выше), позволяющий преобразовать АЗ в задачу с существенно меньшим числом ограничений, оказывается вполне работоспособным.

[8]Задачи, аналогичные представленной в § 6, выдавались автором в качестве лабораторных заданий и заданий выпускных квалификационных работ студентам мех.-мат. ф-та Нижегородского гос. ун-та; лицензия MatLab № 855463.

3. Представленный подход позволяет получать хорошее приближение (по крайней мере, по значению функционала) к решению оптимизационной задачи даже при достаточно грубой сетке конечномерной аппроксимации управления.

4. При использовании для решения АЗ программ, реализующих квазиньютоновские методы, не обнаружился существенный выигрыш по времени по сравнению с методом условного градиента. Более того, оказалось, что метод условного градиента может быть более выигрышным и по количеству итераций и вычислений функции — сравните таблицу 2 и таблицы 1, 5. В данном случае это можно объяснить лишь тем, что границы множества допустимых значений управления, а также оптимальное управление являются постоянными (при большой размерности АЗ).

5. Исследуемый в статье метод хорошо работает и в том случае, когда начальное управление находится достаточно далеко от оптимального.

6. При хорошем начальном приближении метод оказывается работоспособным и в том случае, когда условие **H**) не выполнено.

СПИСОК ЛИТЕРАТУРЫ

1. Волин Ю.М., Островский Г.М. О методе последовательных приближений расчета оптимальных режимов некоторых систем с распределенными параметрами // Автоматика и телемеханика. 1965. Т. 26. № 7. С. 1197–1204.
2. Бутковский А.Г. Теория оптимального управления системами с распределенными параметрами. М.: Наука, 1965. 474 с.
3. Горнов А.Ю. Численные методы исследования задач оптимального управления в механических системах // Мехатроника, автоматизация, управление. 2010. № 8 (113). С. 2–7.
4. Teo K.L., Goh C.J., Wong K.H. A unified computational approach to optimal control problems. Pitman Monographs and Surveys in Pure and Applied Mathematics. Vol. 55. Harlow – New York: Longman Scientific & Technical, John Wiley & Sons, Inc., 1991. ix + 329 p.
5. Sadek I., Kucuk I. A robust technique for solving optimal control of coupled Burger's equations // IMA J. Math. Control Inf. 2011. Vol. 28. № 3. P. 239–250.
6. Варга Дж. Оптимальное управление дифференциальными и функциональными уравнениями. М.: Наука, 1977. 624 с.
7. Сумин В.И. Функционально-операторные вольтерровы уравнения в теории оптимального управления распределенными системами // ДАН СССР. 1989. Т. 305. № 5. С. 1056–1059.
8. Сумин В.И. Об обосновании градиентных методов для распределенных задач оптимального управления // Ж. вычисл. матем. и матем. физ. 1990. Т. 30. № 1. С. 3–21.
9. Афанасьев А.П., Дикусар В.В., Милютин А.А., Чуканов С.А. Необходимое условие в оптимальном управлении. М.: Наука, 1990. 320 с.
10. Чернов А.В. Об одном мажорантном признаке тотального сохранения глобальной разрешимости управляемого функционально–операторного уравнения // Изв. вузов. Математика. 2011. № 3. С. 95–107.
11. Чернов А.В. О мажорантно-минорантном признаке тотального сохранения глобальной разрешимости управляемого функционально-операторного уравнения // Изв. вузов. Математика. 2012. № 3. С. 62–73.
12. Чернов А.В. О достаточных условиях управляемости нелинейных распределенных систем // Ж. вычисл. матем. и матем. физ. 2012. Т. 52. № 8. С. 1400–1414.
13. Чернов А.В. Об управляемости нелинейных распределенных систем на множестве конечномерных аппроксимаций управления // Вестник Удмуртского университета. Математика. Механика. Компьютерные науки. 2013. Вып. 1. С. 83–98.
14. Чернов А.В. О гладких конечномерных аппроксимациях распределенных оптимизационных задач с помощью дискретизации управления // Ж. вычисл. матем. и матем. физ. 2013. Т. 53. № 12. С. 2029–2043.
15. Вайнберг М.М. Вариационный метод и метод монотонных операторов в теории нелинейных уравнений. М.: Наука, 1972. 416 с.

16. Сумин В.И., Чернов А.В. Операторы в пространствах измеримых функций: вольтерровость и квазинильпотентность // Дифференц. уравнения. 1998. Т. 34. № 10. С. 1402–1411.
17. Далецкий Ю.Л., Крейн М.Г. Устойчивость решений дифференциальных уравнений в банаховом пространстве. М.: Наука, 1970. 536 с.
18. Чернов А.В. О сходимости метода условного градиента в распределенных задачах оптимизации // Ж. вычисл. матем. и матем. физ. 2011. Т. 51. № 9. С. 1616–1629.

Поступила в редакцию 19.12.2013

Чернов Андрей Владимирович, к. ф.-м. н., доцент, Нижегородский государственный университет, 603950, Россия, г. Нижний Новгород, пр. Гагарина, 23; Нижегородский государственный технический университет, 603950, Россия, г. Нижний Новгород, ул. Минина, 24.
E-mail: chavnn@mail.ru

A. V. Chernov
On applicability of control parametrization technique to solving distributed optimization problems

Keywords: distributed parameter systems optimization, functional differentiation, piecewise constant approximation of control, control parametrization technique.

Mathematical Subject Classifications: 47J05, 47J35, 47N10, 49M25, 49M37

We study approximating finite-dimensional mathematical programming problems arising from piecewise constant discretization of the control (in the framework of control parametrization technique) in the course of optimization of distributed parameter systems of a rather wide class. We establish the Lipschitz continuity for gradients of approximating problems. We present their formulas involving analytical solutions of an original controlled system and their adjoint one, thereby giving the opportunity for algorithmic separation of the optimization problem itself and the problem of solving a controlled system. Application of the approach under study to numerical optimization of distributed systems is illustrated by example of the Cauchy–Darboux system controlled by an integral criterion. We present the results of numerical solving the corresponding approximation problem in MatLab with the help of the program `fmincon` and also an author-developed program based on the conditional gradient method. Moreover, the unconstrained minimization problem is investigated that arises from the constrained approximation problem with applying the sine parametrization method. We present the results of numerical solving this problem in MatLab with the help of the program `fminunc` and also two author-developed programs based on the steepest descent and BFGS methods, respectively. The results of all numerical experiments are analyzed in detail.

REFERENCES

1. Volin Ju.M., Ostrovskii G.M. A method of successive approximations for calculating optimal modes of some distributed-parameter systems, *Automation and Remote Control*, 1966, vol. 26, pp. 1188–1194.
2. Butkovskiy A.G. *Distributed control systems*, New York: American Elsevier Publishing Company, Inc., 1969, 446 p.
3. Gornov A.Yu. Numerical methods of investigation of optimal control problems in mechanic systems, *Mekhatronika, avtomatizatsiya, upravlenie*, 2010, no. 8 (113), pp. 2–7 (in Russian).
4. Teo K.L., Goh C.J., Wong K.H. *A unified computational approach to optimal control problems*, Pitman Monographs and Surveys in Pure and Applied Mathematics, vol. 55, Harlow – New York: Longman Scientific & Technical, John Wiley & Sons, Inc., 1991, ix + 329 p.
5. Sadek I., Kucuk I. A robust technique for solving optimal control of coupled Burger's equations, *IMA J. Math. Control Inf.*, 2011, vol. 28, no. 3, pp. 239–250.
6. Warga J. *Optimal control of differential and functional equations*, New York–London: Academic Press, Inc., 1972, xiii+531 p. Translated under the title *Optimal'noe upravlenie differentsial'nymi i funktsional'nymi uravneniyami*, Moscow: Nauka, 1977, 624 p.
7. Sumin V.I. Volterra functional-operator equations in the theory of optimal control of distributed systems, *Soviet Math. Dokl.*, 1989, vol. 39, no. 2, pp. 374-378.

8. Sumin V.I. The features of gradient methods for distributed optimal control problems, *USSR Comput. Math. Math. Phys.*, 1990, vol. 30, no. 1, pp. 1–15.

9. Afanas'ev A.P., Dikusar V.V., Milyutin A.A., Chukanov S.A. *Neobkhodimoe uslovie v optimal'nom upravlenii* (A necessary condition in optimal control), Moscow: Nauka, 1990, 320 p.

10. Chernov A.V. A majorant criterion for the total preservation of global solvability of controlled functional operator equation, *Russian Mathematics*, 2011, vol. 55, no. 3, pp. 85–95. DOI: 10.3103/S1066369X11030108.

11. Chernov A.V. A majorant-minorant criterion for the total preservation of global solvability of a functional operator equation, *Russian Mathematics*, 2012, vol. 56, no. 3, pp. 55–65. DOI: 10.3103/S1066369X12030085.

12. Chernov A.V. Sufficient conditions for the controllability of nonlinear distributed systems, *Comput. Math. Math. Phys.*, 2012, vol. 52, no. 8, pp. 1115–1127. DOI: 10.1134/S0965542512050053.

13. Chernov A.V. On controllability of nonlinear distributed systems on a set of discretized controls, *Vestn. Udmurt. Univ. Mat. Mekh. Komp'yut. Nauki*, 2013, no. 1, pp. 83–98 (in Russian).

14. Chernov A.V. Smooth finite-dimensional approximations of distributed optimization problems via control discretization, *Comput. Math. Math. Phys.*, 2013, vol. 53, no. 12, pp. 1839–1852. DOI: 10.1134/S096554251312004X.

15. Vainberg M.M. *Variational method and method of monotone operators in the theory of nonlinear equations*, New York–Toronto: John Wiley & Sons, 1973, xi+356 p. Original Russian text publishied in Vainberg M.M. *Variatsionnyi metod i metod monotonnykh operatorov v teorii nelineinykh uravnenii*, Moscow: Nauka, 1972, 416 p.

16. Sumin V.I., Chernov A.V. Operators in spaces of measurable functions: the Volterra property and quasinilpotency, *Differential Equations*, 1998, vol. 34, no. 10, pp. 1403–1411.

17. Daletsky Y., Krein M.G. *Stability of solutions of differential equations in Banach spaces*, Ann. Math. Soc. Transl., vol. 43, Providence, R.I.: American Mathematical Society, 1974, 386 p. Original Russian text publishied in Daletskii Yu.L., Krein M.G. *Ustoichivost' reshenii differentsial'nykh uravnenii v banakhovom prostranstve*, Moscow: Nauka, 1970, 536 p.

18. Chernov A.V. On the convergence of the conditional gradient method in distributed optimization problems, *Comput. Math. Math. Phys.*, 2011, vol. 51, no. 9, pp. 1510–1523. DOI: 10.1134/S0965542511090077.

Received 19.12.2013

Chernov Andrei Vladimirovich, Candidate of Physics and Mathematics, Associate Professor, Nizhni Novgorod State University, pr. Gagarina, 23, Nizhni Novgorod, 603950, Russia; Nizhni Novgorod State Technical University, ul. Minina, 24, Nizhni Novgorod, 603950, Russia.
E-mail: chavnn@mail.ru