

УДК: 519.85

Численное решение обратной задачи для уравнения гиперболической теплопроводности с малым параметром

Г. Д. Акиндинов^{1,2,a}, В. В. Матюхин^{1,b}, О. И. Криворотко^{1,3,c}

¹Московский физико-технический институт,
Россия, 141701, Московская обл., г. Долгопрудный, Институтский пер., 9

²Институт проблем передачи информации им. А. А. Харкевича,
Россия, 127051, г. Москва, Большой Каретный переулок, д. 19, стр. 1

³Новосибирский государственный университет,
Россия, 630090, Новосибирск, ул. Пирогова, 1

E-mail: ^a akindinov.gd@phystech.edu, ^b matyukhin@phystech.edu, ^c o.krivorotko@g.nsu.ru

Получено 19.02.2023.

Принято к публикации 23.02.2023.

В данной работе приведен алгоритм численного решения обратной начально-краевой задачи для гиперболического уравнения с малым параметром перед второй производной по времени, которая состоит в нахождении начального распределения по заданному конечному. Данный алгоритм позволяет для заданной наперед точности получить решение задачи (в допустимых пределах точности). Данный алгоритм позволяет избежать сложностей, аналогичных случаю с уравнением теплопроводности с обращенным временем. Предложенный алгоритм позволяет подобрать оптимальный размер конечно-разностной схемы путем обучения на относительно больших разбиениях сетки и малом числе итераций градиентного метода. Предложенный алгоритм позволяет получить оценку для константы Липшица градиента целевого функционала. Также представлен способ оптимального выбора малого параметра при второй производной для ускорения решения задачи. Данный подход может быть применен и в других задачах с похожей структурой, например в решении уравнений состояния плазмы, в социальных процессах или в различных биологических задачах. Новизна данной работы заключается в разработке оптимальной процедуры выбора размера шага путем применения экстраполяции Ричардсона и обучения на малых размерах сетки для решения задач оптимизации с неточным градиентом в обратных задачах.

Ключевые слова: обратные задачи, гиперболическая теплопроводность, неточный градиент, схема Ричардсона, регуляризация

Исследование выполнено за счет гранта Российского научного фонда (проект № 21-71-30005), <https://rscf.ru/project/21-71-30005/>.

UDC: 519.85

Numerical solving of an inverse problem of a hyperbolic heat equation with small parameter

G. D. Akindinov^{1,2,a}, V. V. Matyukhin^{1,b}, O. I. Krivorotko^{1,3,c}

¹Moscow Institute of Physics and Technology,
9 Institutskiy per., Dolgoprudny, Moscow region, 141701, Russia

²IITP RAS,

19/1 Bolshoy Karetny per., Moscow, 127051, Russia

³Novosibirsk State University,

1 Pirogova st., Novosibirsk, 630090, Russia

E-mail: ^a akindinov.gd@phystech.edu, ^b matyukhin@phystech.edu, ^c o.krivorotko@g.nsu.ru

Received 19.02.2023.

Accepted for publication 23.02.2023.

In this paper we describe an algorithm of numerical solving of an inverse problem on a hyperbolic heat equation with additional second time derivative with a small parameter. The problem in this case is finding an initial distribution with given final distribution. This algorithm allows finding a solution to the problem for any admissible given precision. Algorithm allows evading difficulties analogous to the case of heat equation with inverted time. Furthermore, it allows finding an optimal grid size by learning on a relatively big grid size and small amount of iterations of a gradient method and later extrapolates to the required grid size using Richardson's method. This algorithm allows finding an adequate estimate of Lipschitz constant for the gradient of the target functional. Finally, this algorithm may easily be applied to the problems with similar structure, for example in solving equations for plasma, social processes and various biological problems. The theoretical novelty of the paper consists in the developing of an optimal procedure of finding of the required grid size using Richardson extrapolations for optimization problems with inexact gradient in ill-posed problems.

Keywords: inverse and ill-posed problems, hyperbolic heat equation, inexact gradient, Richardson method, regularization

Citation: *Computer Research and Modeling*, 2023, vol. 15, no. 2, pp. 245–258 (Russian).

The research was supported by Russian Science Foundation (project No. 21-71-30005), <https://rscf.ru/en/project/21-71-30005/>.

Введение

В данной работе был исследован градиентный поход к решению задачи определения источника возмущений в уравнении гиперболической теплопроводности, которая рассматривается как модификация уравнения теплопроводности путем добавления второй производной по времени с малым параметром. Данное уравнение используется для описания быстротекущих процессов (состояние плазмы, социальные процессы и т. п.) [Четверушкин, 2004].

Основная задача (обратная задача) состоит в восстановлении начального распределения (температуры, концентрации, плотности населения и др.) по конечным наблюдениям волнового процесса. Для решения поставленной задачи в вариационной постановке были использованы градиентные методы с неточным градиентом целевого функционала. Для получения явной формулы для градиента используются методы вариационного исчисления и функционального анализа. После чего необходимо найти оптимальный способ решения дифференциальных уравнений в частных производных с учетом размера сетки в конечно-разностном методе, величины малого параметра и шага градиентного метода.

Постановка обратной задачи

Рассмотрим следующую начально-краевую задачу:

$$\begin{cases} \varepsilon u_{tt} + u_t = (u_x u)_x + u(1 - u), & t \in (0, T), x \in [0, 1], \\ u(0, x) = q(x), \quad \varepsilon u_t(0, x) = 0, & x \in [0, 1], \\ u_x(t, 0) = u_x(t, 1) = 0, & t \in (0, T). \end{cases} \quad (1)$$

Первое уравнение (1) называется уравнением гиперболической теплопроводности с дополнительным нелинейным членом в виде Колмогорова [Колмогоров, Петровский, Пискунов, 1937]. Начальное распределение $q(x)$ — ограниченная, интегрируемая по Лебегу с квадратом функция по $x \in [0, 1]$. В реальных приложениях (физика плазмы, социальные, эпидемиологические, экономические процессы) начальное распределение недоступно для измерений, а известна лишь функция вида

$$f(x) = u(T, x|q), \quad x \in [0, 1]. \quad (2)$$

Обратная задача (1)–(2) состоит в восстановлении начального распределения $q(x)$ по наблюдениям конечного распределения $f(x)$.

В работе [Четверушкин, 2004] показано, что добавление малого параметра $\varepsilon > 0$ улучшает устойчивость решения прямой задачи для параболического уравнения. В данной работе будет исследована устойчивость решения обратной задачи (1)–(2) в зависимости от параметра ε .

В данном случае подход с обращением времени и решением задачи с помощью разностных схем **некорректен** в силу возникновения эффектов, аналогичных тем, что возникают в обыкновенном уравнении теплопроводности с обращенным временем [Кабанихин, 2018]:

$$\begin{cases} u_t + u_{xx} = 0, & t \in (0, T), x \in [0, 1], \\ u(0, x) = f(x), & x \in [0, 1], \\ u_x(t, 0) = u_x(t, \pi) = 0, & t \in (0, T). \end{cases}$$

Решение последнего уравнения выражается в виде ряда Фурье по собственным функциям стационарного оператора и имеет вид

$$u(t, x) = \sum_{k=1}^{\infty} f_n \cos(kx) e^{k^2 t},$$

где f_n — коэффициенты разложения $f(x)$ в ряд по собственным функциям $\cos(kx)$. В таком случае квадрат нормы в $\mathbb{L}_2([0, \pi])$ решения уравнения будет иметь следующий вид:

$$\|u(t, x)\|_{\mathbb{L}_2([0, \pi])}^2 = \sum_{k=1}^{\infty} f_k^2 e^{2k^2 t}.$$

Получается, что если решение и существует, то коэффициенты разложения f_n должны быть $o(e^{-n^2 t})$, а значит, такое решение будет устойчивым лишь для очень узкого класса функций f . Даже в случае, когда ряд обрывается на некотором члене, малые изменения коэффициентов Фурье способны сильно повлиять на решение. В работе предложен градиентный подход к решению обратной задачи с учетом регуляризации ее решения.

Градиентный подход к решению обратной задачи

Будем искать решение обратной задачи в классе функций

$$\mathcal{Q} := \{q \in \mathbb{L}_2([0, 1]): 0 \leq q(x) < \infty, q'(0) = q'(1) = 0\}.$$

Идея состоит в построении градиентного метода в следующем виде:

$$q^{n+1}(x) = q^n(x) - h_n J'(q^n), \quad q^0(x) \in \mathcal{Q}.$$

В данном случае целевой функционал имеет вид

$$J(q) = \int_0^1 |f(x) - u(T, x|q)|^2 dx,$$

а производная функционала $J'(q)$ понимается в смысле производной Фреше:

$$J(q + \delta q) - J(q) = \langle J'(q), \delta q \rangle + o(\|\delta q\|).$$

Вывод градиента целевого функционала

Для получения градиента целевого функционала в явном виде будем рассматривать два допустимых начальных условия из класса \mathcal{Q} : q и $q + \delta q$; соответствующие им решения уравнения будем обозначать как $u(t, x|q)$ и $u(t, x|q) + \delta u$. Рассмотрим приращение целевого функционала:

$$J(q + \delta q) - J(q) = \int_0^1 2(u(T, x|q) - f(x))\delta u(T, x|q) dx + o(\|\delta u(T, x|q)\|).$$

Для того чтобы получить связь δu и δq , вычтем из уравнения на $(u + \delta u)$ уравнение на (u) :

$$\begin{cases} \varepsilon \delta u_{tt} + \delta u_t = \varepsilon \delta u_{xx} + u_{xx} \delta u + 2u_x \delta u_x + \delta u - 2u \delta u - \delta u^2 + \delta u \delta u_{xx} + \delta u_x^2, & t \in (0, T), x \in [0, 1], \\ \delta u(0, x) = \delta q, \quad \varepsilon \delta u_t(0, x) = 0, & x \in [0, 1], \\ \delta u_x(t, 0) = \delta u_x(t, 1) = 0, & t \in (0, T). \end{cases}$$

Будем рассматривать линейную часть уравнения по вариациям δu и ее производным вплоть до второго порядка. Запишем оператор $\mathcal{L}\delta u$:

$$\mathcal{L}\delta u = \varepsilon \delta u_{tt} + \delta u_t - \varepsilon \delta u_{xx} - u_{xx} \delta u - 2u_x \delta u_x - \delta u + 2u \delta u.$$

Рассмотрим скалярное произведение в пространстве $\mathbb{L}_2([0, T] \times [0, 1])$ оператора $\mathcal{L}\delta u$ с пробной функцией p :

$$\langle \mathcal{L}\delta u, p \rangle = \langle \delta u, \mathcal{L}^* p \rangle = 0.$$

С помощью интегрирования по частям можем получить сопряженный оператор \mathcal{L}^* и соответствующую систему для пробной функции с помощью интегрирования по частям:

$$\begin{aligned} \langle \delta u, \mathcal{L}^* p \rangle^{(t,x)} &= \langle \delta u, \varepsilon p_{tt} - p_t - u p_{xx} + p(2u - 1) \rangle^{(t,x)} + \\ &+ \langle \delta u(T, x), p(T, x) - \varepsilon p_t(T, x) \rangle^{(x)} + \langle \delta q, -p(0, x) + \varepsilon p_t(0, x) \rangle^{(x)} + \langle \delta u_t(T, x), \varepsilon p(T, x) \rangle^{(x)} + \\ &+ \langle u(t, 0) \delta u(t, 0), -p_x(t, 0) \rangle^{(t)} + \langle u(t, 1) \delta u(t, 1), p_x(t, 1) \rangle^{(t)}. \end{aligned}$$

Посмотрев на выражение для градиента с двух сторон, можем видеть, что

$$\langle J'(q), \delta q \rangle = J(q + \delta q) - J(q) + \langle \delta u, \mathcal{L}^* p \rangle^{(t,x)} + o(\|\delta q\|).$$

Подставив все, учитывая $J(q + \delta q) - J(q) = \langle \delta u(T, x), 2(u(T, x) - f(x)) \rangle^{(x)} + o(\|\delta q\|)$, получим сопряженную к исходной начально-краевую задачу

$$\begin{cases} \varepsilon p_{tt} - p_t - u p_{xx} + p(2u - 1) = 0, & t \in (0, T), x \in [0, 1], \\ \varepsilon p_t(T, x) - p(T, x) = 2(u(T, x) - f(x)), \quad \varepsilon p(T, x) = 0, & x \in [0, 1], \\ p_x(t, 0) = p_x(t, 1) = 0, & t \in (0, T), \end{cases}$$

и выражение для градиента целевого функционала примет вид:

$$J'(q) = -p(0, x) + \varepsilon p_t(0, x).$$

Преобразуем сопряженное уравнение путем инвертирования времени ($v(t, x) = p(T - t, x)$) для применения той же разностной схемы для численного решения, как в случае системы (1):

$$\begin{cases} \varepsilon \widetilde{v}_{tt} + \widetilde{v}_t = u(T - t, x) \widetilde{v}_{xx} + (1 - 2u(T - t, x)) \widetilde{v}, & t \in (0, T), x \in [0, 1], \\ \widetilde{v}(0, x) + \varepsilon \widetilde{v}_t(0, x) = 2(f(x) - u(T, x)), \quad \varepsilon \widetilde{v}(0, x) = 0, & x \in [0, 1], \\ \widetilde{v}_x(t, 0) = \widetilde{v}_x(t, 1) = 0, & t \in (0, T). \end{cases} \quad (3)$$

Тогда выражение для градиента целевого функционала запишется следующим образом:

$$J'(q) = -\{\widetilde{v}(T, x) + \varepsilon \widetilde{v}_t(T, x)\}.$$

Подходы к численному определению неточного градиента

Для вычисления градиента целевого функционала необходимо численно решить прямую (1) и сопряженную (3) задачи. Для численного решения указанных задач будем использовать конечно-разностный подход.

Явные разностные схемы

Дискретизируем область определения системы с шагами Δt по времени и Δx по пространству. Тогда значение функции в точке сетки будем обозначать как $U_j^i = u(t_i, x_j)$. Аппроксимируем частные производные следующим образом:

$$\begin{aligned} u_t &= \frac{U_j^{i+1} - U_j^{i-1}}{2\Delta t}, & u_{tt} &= \frac{U_j^{i+1} - 2U_j^i + U_j^{i-1}}{\Delta t^2}, \\ u_x &= \frac{U_{j+1}^i - U_{j-1}^i}{2\Delta x}, & u_{xx} &= \frac{U_{j+1}^i - 2U_j^i + U_{j-1}^i}{\Delta x^2}. \end{aligned}$$

Тогда исходное уравнение преобразуется к следующему виду:

$$U_j^{i+1} = a^T U + U^T A U,$$

где $U = [U_j^i, U_j^{i-1}, U_{j+1}^i, U_{j-1}^i]^T$, $a^T = [\frac{2\varepsilon+\Delta t^2}{2\varepsilon+\Delta t}, \frac{-2\varepsilon+\Delta t}{2\varepsilon+\Delta t}, 0, 0]$,

$$A = \begin{bmatrix} -4(\Delta x^2 + 2) & 0 & 2 & 2 \\ 0 & 0 & 0 & 0 \\ 2 & 0 & 1 & -1 \\ 2 & 0 & -1 & 1 \end{bmatrix} \cdot \frac{\Delta t^2}{2(2\varepsilon + \Delta t)\Delta x^2}.$$

Начально-краевые условия в случае $\varepsilon \neq 0$ будут иметь следующий вид:

$$U_j^0 = q_j, \quad U_j^1 = U_j^0, \quad U_0^i = U_2^i, \quad U_n^i = U_{n-2}^i.$$

Можно обратить внимание, что член $U^T A U$ может сильно повлиять на решение в силу своей квадратичности, поэтому имеет смысл уменьшить норму матрицы (в качестве нормы рассматривается спектральная норма), откуда можно будет получить оптимальный выбор размера сетки. Потребуем, чтобы норма этой матрицы была меньше единицы, но в реальных задачах для большей устойчивости используются числа значительно меньше 1:

$$\|A\|_2 < 1 \Rightarrow \frac{\sqrt{(\Delta x^2 + 2)^2 + (\Delta x^2 + 2) \sqrt{4 + \Delta x^2 + 1}}}{2} \cdot \frac{2\Delta t^2}{(2\varepsilon + \Delta t)\Delta x^2} < 1,$$

$$\sqrt{(\Delta x^2 + 2)^2 + (\Delta x^2 + 2) \sqrt{4 + \Delta x^2 + 1}} \leq 4,1,$$

$$\frac{4,1\Delta t^2}{(2\varepsilon + \Delta t)\Delta x^2} = c < 1.$$

«Хороший» размер сетки можно выбирать согласно соотношению выше.

Аналогично решается и сопряженная задача, для которой конечно-разностной метод выглядит следующим образом:

$$V_j^{i+1} = b^T V + V^T B V,$$

где $V = [V_j^i, V_j^{i-1}, V_{j+1}^i, V_{j-1}^i, U_j^{n-i}]^T$, $b^T = [\frac{2\varepsilon+\Delta t^2}{\Delta t^2}, \frac{2\varepsilon-\Delta t}{2\Delta t^2}, 0, 0, 0] \cdot \frac{2\Delta t^2}{2\varepsilon+\Delta t}$,

$$B = \begin{bmatrix} 0 & 0 & 0 & 0 & s \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 \\ s & 0 & 1 & 1 & 0 \end{bmatrix} \cdot \frac{\Delta t^2}{(2\varepsilon + \Delta t)\Delta x^2}, \quad s = -2(\Delta x^2 + 1).$$

Аналогично прямой задаче ограничим норму матрицы квадратичной формы:

$$\|B\|_2 = \sqrt{s^2 + 1} \cdot \frac{\Delta t^2}{(2\varepsilon + \Delta t)\Delta x^2} < \frac{4,2\Delta t^2}{(2\varepsilon + \Delta t)\Delta x^2} = c_1 < 1.$$

Оптимальный размер шага для прямой и сопряженной задачи вышел почти одинаковый, поэтому будем использовать в вычислениях размер сетки из последнего соотношения как оптимальный в обоих случаях.

Процедура, описанная выше, хорошо подошла для решения уравнений с различными начальными условиями, целевой функционал сходил к 0 (сначала линейный участок,

далее плато), квадрат нормы расхождения истинного начального условия и полученного по данной процедуре сходил к нулю таким же образом.

Минус такого подхода состоит в том, что зачастую явные схемы являются неустойчивыми, поэтому можно применить полученные результаты к неявным схемам, решение которых чуть сложнее за счет необходимости вычислять решение системы уравнений на каждом шаге по времени, однако большим плюсом для этого является их устойчивость.

Неявные разностные схемы

В данном разделе описывается вид, по которому производится замена производных, все выкладки аналогичны случаю явных схем.

Производные будем заменять следующим образом:

$$u_t = \frac{U_j^{i+1} - U_j^{i-1}}{2\Delta t}, \quad u_{tt} = \frac{U_j^{i+1} - 2U_j^i + U_j^{i-1}}{\Delta t^2},$$

$$u_x = \frac{U_{j+1}^{i+1} - U_{j-1}^{i+1}}{2\Delta x}, \quad u_{xx} = \frac{U_{j+1}^{i+1} - 2U_j^{i+1} + U_{j-1}^{i+1}}{\Delta x^2}.$$

В результате такой замены получаем связь между $(U_{j-1}^{i+1}, U_j^{i+1}, U_{j+1}^{i+1})$ в зависимости от U_j^i , после чего решаем на слое $i + 1$ систему уравнений для $j \in \{1, n-1\}$. Функция известна на слоях $i = \{0, 1\}$ из начальных условий, связь $U_0^{i+1} = U_2^{i+1}$ и $U_n^{i+1} = U_{n-2}^{i+1}$ следует из граничных условий.

В результате имеем большую устойчивость метода. Для него также будет справедлив аналог условия Куранта:

$$\frac{\Delta t^2}{(2\varepsilon + \Delta t)\Delta x^2} \lesssim \frac{1}{4}. \quad (4)$$

Выбор шага по x

Была установлена связь между Δt и Δx , но все же нужно разобраться с тем, какой шаг выбрать по одному из них. Проще будет разобраться с размером шага по Δx в силу того, что наш функционал качества как раз является интегралом по пространству, значит, таким образом нам будет проще разобраться с размером шага с помощью него. Мы хотим каким-то образом получить следующий вид функционала в смысле разложения по малому параметру — размеру сетки по x :

$$J(q^n) - J^* = J(q^n) = C(\Delta x)^p + O((\Delta x)^{p+1}).$$

Предлагается изначально взять наибольший допустимый размер сетки (разбиение каждой из областей по крайней мере на два отрезка), после чего, дробя шаг по x (например, делением на 2), получать меньшую ошибку. Данная идея как раз реализуется в **методе экстраполяции Ричардсона**, в котором исследуется связь следующих величин:

$$A(h) - A^* = Ch^p + O(h^{p+1}) \Rightarrow A_1(h) = \frac{s^p A_0\left(\frac{h}{s}\right) - A_0(h)}{s^p - 1} = A^* + O(h^{p+1}).$$

В соответствии с этим определяется рекуррентное соотношение:

$$A_{i+1}(h) = \frac{s^{p_i} A_i\left(\frac{h}{s}\right) - A_i(h)}{s^{p_i} - 1} = A^* + O(h^{p_{i+1}}).$$

Таким образом, формула Ричардсона позволяет нам получить достаточно мощный инструмент, позволяющий по небольшому числу точек (и относительно малому числу итераций градиентного

метода) оценить константу C и степень p при h , уже после чего можно задавать желаемую точность γ , по которой уже можно получить оптимальный в данном случае шаг по координате и после этого уже запускать градиентный метод на нужном числе итераций.

$$\Delta x \approx \left(\frac{C}{\gamma}\right)^{1/p}. \quad (5)$$

Шаг по времени будет определяться из аналога условия Куранта (4).

Анализ градиентного метода

Также интересно исследовать другое важное свойство — условие Липшица для градиента целевого функционала:

$$J'(q) = -(v(T, x) + \varepsilon v_t(T, x)).$$

Запишем вариацию градиента:

$$\delta J'(q) = -\delta v(T, x) - \varepsilon v_t(T, x),$$

причем $\delta v(t, x)$ мы можем получить как решение системы для (w) :

$$\begin{cases} \varepsilon w_{tt} + w_t = -g(T-t, x)(v_{xx} + 2v) + u(T-t, x)w_{xx} + (1 - 2u(T-t, x))w, & t \in (0, T), x \in [0, 1], \\ w(0, x) + \varepsilon w_t(0, x) = -2g(T, x), & x \in [0, 1], \\ w(0, x) = 0, & x \in [0, 1], \\ w_x(t, 0) = w_x(t, 1) = 0, & t \in (0, T), \end{cases}$$

где $g \equiv \delta u$, определяемая из системы

$$\begin{cases} \varepsilon g_{tt} + g_t = u g_{xx} + u_{xx} g + 2u_x g_x + g - 2ug, & t \in (0, T), x \in [0, 1], \\ g(0, x) = \delta q, & x \in [0, 1], \\ \varepsilon g_t(0, x) = 0, & x \in [0, 1], \\ g_x(t, 0) = g_x(t, 1) = 0, & t \in (0, T) \end{cases}$$

(системы для g и w получены как линейная часть вариации соответствующих уравнений).

Таким образом, решая эти две системы, мы можем получить гессиан нашего целевого функционала, и по мере работы градиентного метода оценивать константу Липшица градиента целевого функционала, и в соответствии с ней определять субоптимальный размер шага градиентного метода как $\max \frac{\|\delta J'(q^i)\|}{\|\delta q^i\|}$. Что касается вариации начального распределения, то его можно брать как $q^i - q^{i-1}$. Сделав несколько итераций по мере работы метода (либо же до момента достижения нужной точности), можно пользоваться уже полученной ранее константой Липшица. Данный функционал (градиент целевого функционала), очевидно, будет обладать данным свойством, так как является линейным и ограниченным.

$$J'(q) = \int_0^1 \int_0^T K(\tau, \xi, x, t) v(\tau, \xi) d\tau d\xi,$$

$$\delta u(t, x) = \int_0^1 \int_0^T G(\tau, \xi, t, x) u(\tau, \xi) d\tau d\xi.$$

Ядра этих функционалов содержат дельта-функцию Дирака (поскольку включают в себя начально-краевые условия), однако оба функционала будут ограниченными. А так как функционалы являются линейными и ограниченными, значит, у них конечная норма, а норма первого функционала и есть константа Липшица градиента нашего исходного целевого функционала.

Шаг градиентного метода на первой итерации выбирается произвольно, другими словами, можно считать, что два начальных распределения выбираются из каких-то эвристических соображений, после чего метод будет сам получать субоптимальный размер шага.

Условие Липшица для градиента — очень хорошее свойство, но в нашем случае это достаточно дорогое удовольствие, поэтому не следует его вычислять слишком много раз. Оптимальный вариант — на первых шагах запускать классический градиентный спуск, а потом уже использовать ускоренные методы, когда мы уже получили некоторое представление о константе Липшица целевого функционала. Ошибка вычисления градиента целевого функционала у нас полностью совпадает с неточностью решения системы для v .

Также хочется отметить, что было экспериментально проверено, что для первых итераций можно брать в качестве оценки константы Липшица норму гесса конечного распределения (в случае ее существования), после чего уже подстраиваться под новые данные.

Оптимизация малого параметра при второй производной

Для оптимизации малого параметра при второй производной можно проделать аналогичную процедуру путем вариации функционала по ε при фиксированном q . В таком случае $\delta_\varepsilon u \equiv r(t, x)$,

$$\begin{cases} \left\{ \varepsilon \frac{\partial^2}{\partial t^2} + \frac{\partial}{\partial t} - u_{xx} - u \frac{\partial^2}{\partial x^2} - 2u_x \frac{\partial}{\partial x} - 1 + 2u \right\} r = -\varepsilon u_{tt}, & t \in (0, T), x \in [0, 1], \\ r(0, x) = 0, & x \in [0, 1], \\ r_t(0, x) = 0, & x \in [0, 1], \\ r_x(t, 0) = r_x(t, 1) = 0, & t \in (0, T). \end{cases}$$

С рассмотрением такой задачи

$$R_\gamma(\varepsilon) \rightarrow \min, \quad \text{s. t.} \int_0^1 \int_0^T |u(t, x) - u_\varepsilon(t, x)|^2 dt dx = \gamma^2 \|u(t, x)\|^2,$$

где $R_\gamma(\varepsilon)$ — это некоторая характеристика (будь то вычислительная сложность, время выполнения, «точность вычисления»).

Экспериментально проверено, что зависимость (времени выполнения, числа итераций, точности) от ε является унимодальной по ε , поэтому целесообразно проводить отыскание оптимального ε с помощью бинарного поиска, изначально задав два произвольных значения ε (в наших экспериментах его целесообразно выбирать значительно меньше, чем характерный размер сетки по t), поэтому можно выбирать значения между $\varepsilon_1 = 0,001\Delta t$ и $\varepsilon_2 = 0,01\Delta t$. «Малость» параметра определяется относительно Δt исходя из метода размерностей и подтверждается экспериментами. При таком выборе целесообразно выбирать Δx и Δt из условия (4).

Применение методов на практике

Для проверки работоспособности алгоритма предлагается изначально задать точное начальное распределение q_{exact} , после чего вычислить (с желаемой точностью) функцию наблюдений для случая с $\varepsilon = 0$, после чего уже работать с f , полученной как $u(T, x|q_{exact})$. После этого предлагается вычислить функционал $J(f)$ по методу Ричардсона для того, чтобы определить оптимальный размер сетки по x . После этого мы уже выполняем градиентный метод. Процедура решения будет следующая.

а) Если задано q_{exact} :

0. Вычисляем $f(x) = U(T, x|q_{exact})$ и добиваемся требуемой точности с помощью метода экстраполяции Ричардсона. Забываем q_{exact} , инициализируемся другим $q = q_0$.

i.1. Получаем $u(t, x|q_i)$ с помощью метода конечных разностей.

i.2. Получаем $v(t, x|q_i)$ с помощью метода конечных разностей.

i.3. Вычисляем градиент $J'(q_i) = -(v(T, x) + \varepsilon v_i(T, x))$ и делаем шаг градиентного метода $q_{i+1} = q_i - hJ'(q_i)$.

б) Если задано $f(x) = u(T, x|q_{exact})$:

0. Инициализируемся некоторым q_0 из некоторых общих соображений.

i.1. Получаем $u(t, x|q_i)$ с помощью метода конечных разностей.

i.2. Получаем $v(t, x|q_i)$ с помощью метода конечных разностей.

i.3. Вычисляем градиент $J'(q_i) = -(v(T, x) + \varepsilon v_i(T, x))$ и делаем шаг градиентного метода $q_{i+1} = q_i - hJ'(q_i)$.

На практике получается, что $J(h) = Ch^p + o(h^p)$. Улучшить оценку для J можно путем применения метода экстраполяции Ричардсона. Условно: можно запустить градиентный метод на 20–40 итерациях с очень большим шагом по координате ($\Delta x = 0,1$ при $T = 0,01$), в соответствии с ним выбрать $\Delta t \approx \frac{\Delta x^2}{4}$, после чего уменьшить шаг вдвое, получить оценку J_2 и рассмотреть $J^* = J_1 + \frac{J_1 - J_2}{2^p - 1}$, которая будет на порядок точнее описывать действительность. Данную процедуру можно повторить несколько раз для получения представления о первом члене в вариационном ряде для $J(\Delta x) = C\Delta x^p + o(\Delta x^p)$, который мы исследуем на значение константы C и первую степень в вариационном ряду.

Для удобства прологарифмируем обе части и получим линейную зависимость, после чего уже можно будет определить значения коэффициентов $\ln(C)$ и p . Используя эти числа, можно получить нужный шаг для достижения требуемой точности по невязке функционала согласно (5). Для всех рассмотренных пределов выходила неплохая почти линейная зависимость, по которой с помощью метода наименьших квадратов удалось получить корректные константы C и p .

Данная зависимость была исследована для $T = 0,01$ для функций $q_{exact}(x) \in \{\sin^2(\pi x), \sin^2(\frac{\pi x}{2}), x\}$ и $f(x) = u(T, x|q_{exact}) \in \{\sin^2(\pi x), \sin^2(\frac{\pi x}{2}), x\}$. Хочется отметить, что не для всех функций оптимальный (в смысле желаемой точности) алгоритм работает похожим образом, причем иногда при выборе неоптимального шага метод выходил на чуть более плохую асимптоту, а на рис. 2 можно заметить немонотонный характер зависимости логарифма целевого функционала от числа итераций. Природу данного эффекта еще предстоит изучить. Кроме того, невозможно достичь любой желаемой точности при работе с данным алгоритмом из-за того, что $\Delta t \leq \frac{T}{3}$, а в соответствии с ним и $\Delta x \geq \sqrt{T}$. Поэтому для наибольшей возможной точности $\widehat{\gamma}$ можно ввести соотношение через p, C, T :

$$\widehat{\gamma} \sim CT^{p/2}.$$

С результатами экспериментов можно ознакомиться по ссылке <https://github.com/ujiffy/inverseproblems> (будут дополняться) и в разделе приложения.

Заключение

В данной работе была исследована связь между размером сетки по времени и пространству, получен способ определения оптимального размера сетки по пространству, предложен способ определения оптимального параметра при второй производной и получена адекватная оценка для константы Липшица целевого функционала. Все данные методы были проверены на синтетических примерах и показали неплохие результаты. Данный подход применим и для решения других задач со схожей структурой. Основное различие будет заключаться в вариации стационарного оператора и в вариации внешнего возмущения (а соответственно, и в изменении сопряженной задачи). Во всем остальном процедура решения задачи будет аналогичной.

Отдельно хочется поблагодарить Александра Владимировича Гасникова за неоценимый вклад в данную работу и безграничное число идей для исследования и развития.

Список литературы (References)

- Кабанихин С. И.* Обратные и некорректные задачи. — 4-е издание. — Новосибирск: Издательство СО РАН, 2018. — 512 с.
Kabanikhin S. I. Obratnye i nekorrektnye zadachi [Inverse and ill-posed problems]. — 4th edition. — Novosibirsk: Publishing House SB RAS, 2018. — 512 p. (in Russian).
- Колмогоров А. Н., Петровский И. Г., Пискунов Н. С.* Исследование уравнения диффузии, соединенной с возрастанием количества вещества, и его применение к одной биологической проблеме // Бюллетень Московского государственного университета. — М., 1937. — С. 1–26.
Kolmogorov A., Petrovskii A., Piscunov N. Issledovanie uravneniya diffuzii, soedinennoi s vozrastaniem kolichestva veshchestva, i ego primenenie k odnoi biologicheskoi probleme [A study of the equation of diffusion with increase in the quantity of matter, and its application to a biological problem] // Byul. Moskovskogo gos. univ. — Moscow, 1937. — Vol. 1, No. 6. — P. 1–26 (in Russian). Reprinted in: *Tikhomirov V. M.* (ed.) Selected Works of A. N. Kolmogorov. — Dordrecht: Kluwer, 1991. — Vol. 1. — P. 242–270. Also in: I. G. Petrowsky Selected Works. — Amsterdam: Gordon and Breach, 1996. — Part II. — P. 106–132.
- Моисеев Т. Е., Мышецкая Е. Е., Тишкин В. Ф.* О близости решений невозмущенных и гиперболизированных уравнений теплопроводности для разрывных начальных данных // Препринты ИПМ им. Келдыша. — 2017. — № 86. — 15 с.
Moiseev T. O blizosti reshenii nevozmushchennykh i giperbolizirovannykh uravnenii teploprovodnosti dlya razryvnykh nachal'nykh dannyykh [On the proximity of solutions of unperturbed and hyperbolized heat equations for discontinuous initial data] // Keldysh Institute Preprints. — 2017. — No. 86. — 15 p. (in Russian).
- Четвертушкин Б. Н.* Кинетические схемы и квазигазодинамическая система уравнений. — М.: Мир, 2004. — 332 с.
Chevtvrtushkin B. N. Kineticheskie skhemy i kvazigazodinamicheskaya sistema uravnenii [Kinetic schemes and quasi-gas-dynamic system of equations]. — Moscow: Mir, 2004. — 332 p. (in Russian).
- Gasnikov A., Kabanikhin S., Mohammed A., Shishlenin M.* Convex optimization in Hilbert space with applications to inverse problems // e-print, 2017. — URL: <https://arxiv.org/abs/1703.00267>
- Matyukhin V., Kabanikhin S., Shishlenin M., Novikov N., Vasin A., Gasnikov A.* Convex optimization with inexact gradients in Hilbert space and applications to elliptic inverse problems // Mathematical Optimization Theory and Operations Research: 20th International Conference, MOTOR 2021, Irkutsk, Russia, July 5–10, 2021. — Springer International Publishing. — Proceedings 20. — P. 159–175.
- Krivorotko O., Kabanikhin S. I., Zhang S., Kashtanova V.* Global and local optimization in identification of parabolic systems // Journal of Inverse and Ill-posed Problems. — 2020. — Vol. 28, No. 6. — P. 899–913.
- Polyak B. T.* Introduction to Optimization. — New York: Optimization Software, 1987.

Приложения

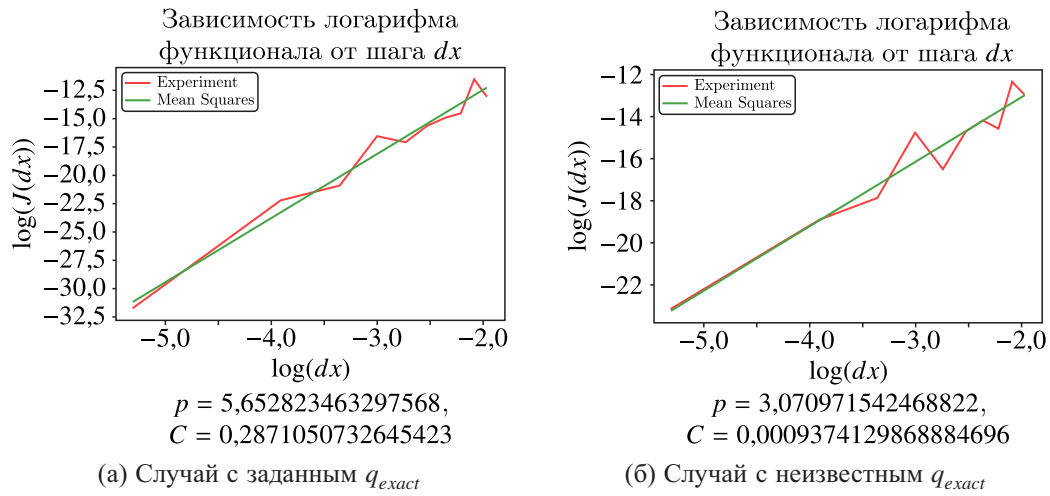


Рис. 1. Характер зависимости логарифма целевого функционала и соответствующие им C и p для $J(dx) = C dx^p + o(dx^p)$. Число итераций градиентного метода — 20, шаг градиентного метода — 0,1. $q_{exact} = \sin^2(\pi x)$ в первом случае и $f(x) = \sin^2(\pi x)$ во втором соответственно

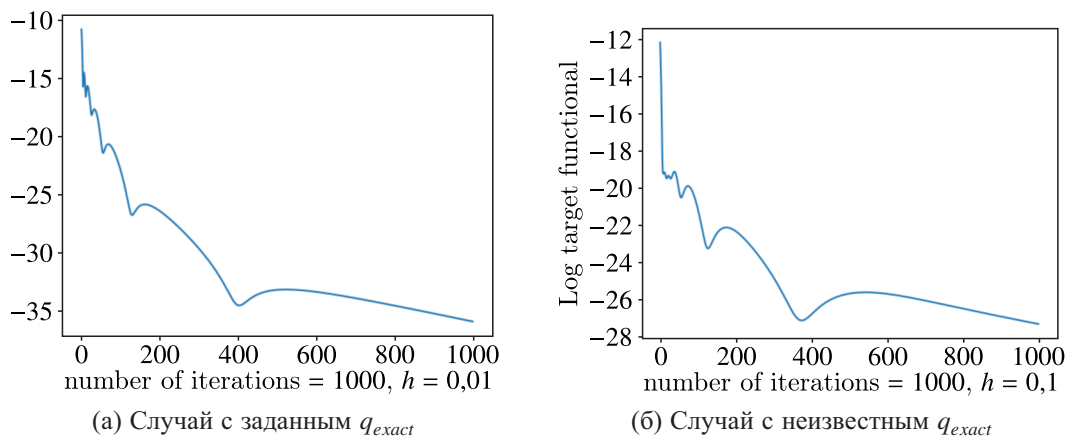


Рис. 2. Работа градиентного метода на оптимальных шагах, полученных в предыдущем эксперименте. $q_{exact} = \sin^2(\pi x)$ в первом случае и $f(x) = \sin^2(\pi x)$ во втором соответственно

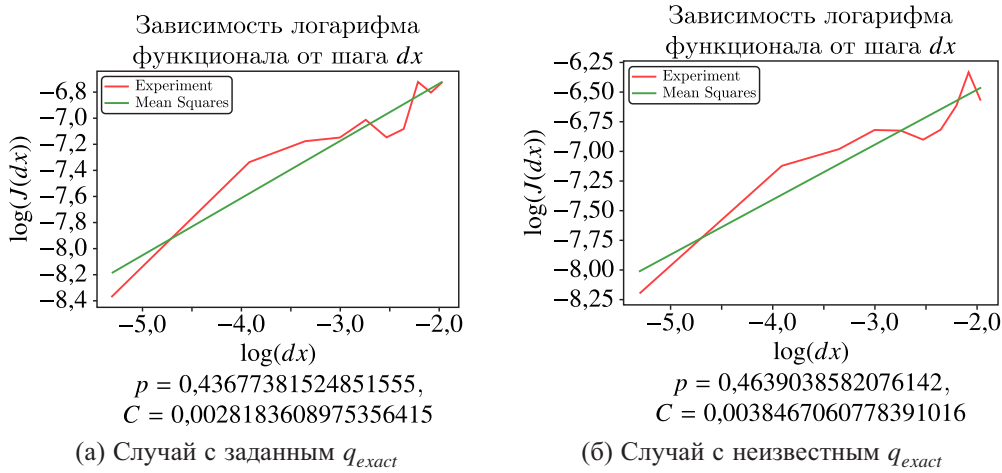


Рис. 3. Характер зависимости логарифма целевого функционала и соответствующие им C и p для $J(dx) = C dx^p + o(dx^p)$. Число итераций градиентного метода – 20, шаг градиентного метода – 0,1. $q_{exact} = \sin^2\left(\frac{\pi x}{2}\right)$ в первом случае и $f(x) = \sin^2\left(\frac{\pi x}{2}\right)$ во втором соответственно

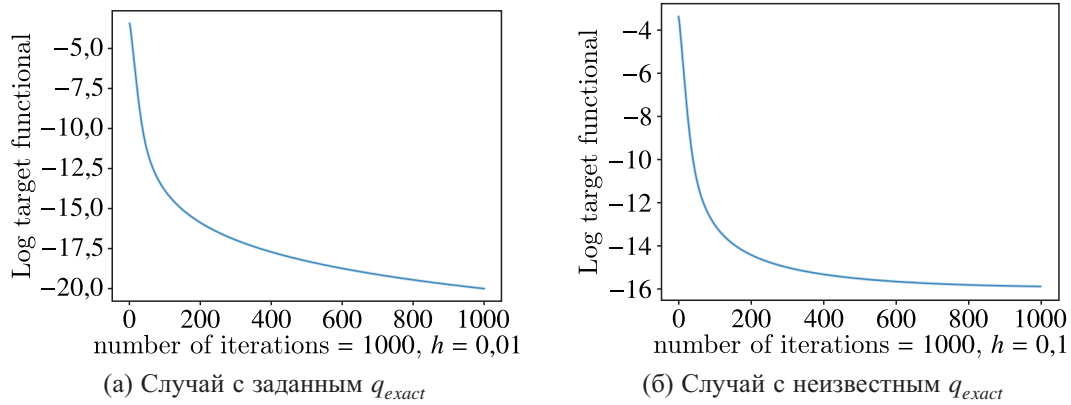


Рис. 4. Работа градиентного метода на оптимальных шагах, полученных в предыдущем эксперименте. $q_{exact} = \sin^2\left(\frac{\pi x}{2}\right)$ в первом случае и $f(x) = \sin^2\left(\frac{\pi x}{2}\right)$ во втором соответственно

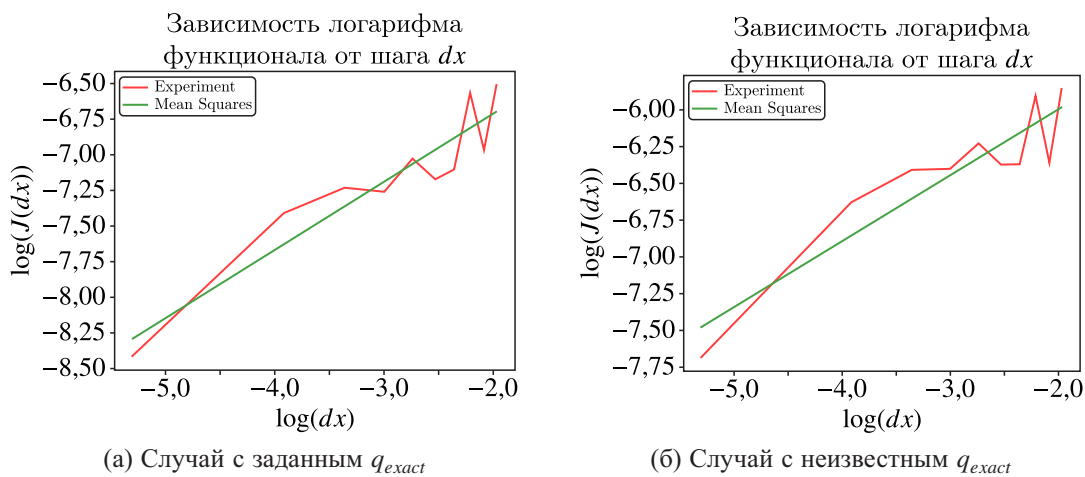


Рис. 5. Характер зависимости логарифма целевого функционала и соответствующие им C и p для $J(dx) = C dx^p + o(dx^p)$. Число итераций градиентного метода – 20, шаг градиентного метода – 0,1. $q_{exact} = x$ в первом случае и $f(x) = x$ во втором соответственно

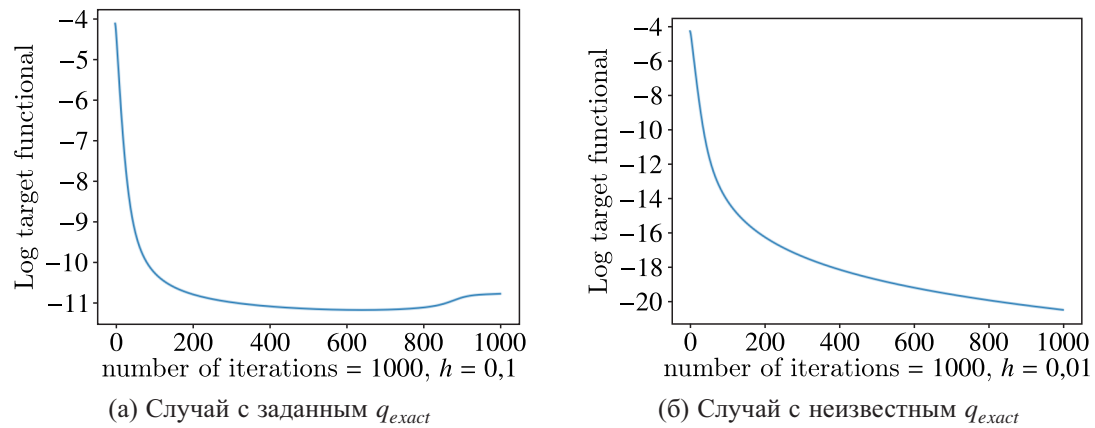


Рис. 6. Работа градиентного метода на оптимальных шагах, полученных в предыдущем эксперименте. $q_{exact} = x$ в первом случае и $f(x) = x$ во втором соответственно