

Градиентные методы оптимизации для решения некоторых обратных задач математической физики

Плетнев Никита Вячеславович

Московский физико-технический институт
Факультет управления и прикладной математики
Кафедра интеллектуальных систем
Научный руководитель: д. ф.-м. н. А. В. Гасников

16 июня 2021 г.

Задача Коши для уравнения Гельмгольца

$$\begin{aligned}\Delta u + k^2 u &= 0, & (x, y) \in \Omega &= (0, 1) \times (0, 1), \\ u|_{x=0} &= f(y), & u_x|_{x=0} &= g(y), & u|_{y=0} &= u|_{y=1} = 0.\end{aligned}$$

Проблема и путь решения

Задача Коши для эллиптического уравнения поставлена некорректно.

Корректной она станет, если добавить условие: $u|_{x=1} = q(y)$.
Таким образом, задача сводится к обратной.

Обратная задача

$$\begin{aligned}\Delta u + k^2 u &= 0, & (x, y) \in \Omega &= (0, 1) \times (0, 1), \\ u_x|_{x=0} &= g(y), & u|_{y=0} &= u|_{y=1} = 0, & u|_{x=1} &= q(y).\end{aligned}$$

Найти $q(y)$ по известной $f(y) = u|_{x=0}$.

Проверка различных градиентных методов в данном подходе

Необходимо проверить эффективность используемого подхода к решению задачи Коши.

Поскольку для функционала и градиента возможно только приближённое вычисление, погрешность вычислений тоже влияет на качество.

Установление концепции шума

аддитивный шум	относительный шум
$\ \tilde{\nabla} J(q) - \nabla J(q)\ \leq \delta$	$\ \tilde{\nabla} J(q) - \nabla J(q)\ \leq \alpha \ \nabla J(q)\ $
погрешность накапливается, невязка растёт	сходимость замедляется, наклонная прямая в логарифмическом масштабе

Эксперименты позволяют определить, какая из концепций имеет место.

Оператор

Функция q — элемент гильбертова пространства $H = \{q \in L_2(0, 1) : q(0) = q(1) = 0\}$.

Определим оператор $A : H \rightarrow H$ следующим образом:
 $(Aq)(y) = u(0, y) \quad \forall y \in (0, 1)$, где u — решение соответствующей краевой задачи.

Операторное представление

Задача принимает вид $Aq = f$, или

$$J(q) = \frac{1}{2} \|Aq - f\|_H^2 \rightarrow \min_{q \in H}.$$

Для данной оптимизационной задачи применяются метод сопряжённых градиентов (Флетчера-Ривса) и ускоренный метод подобных треугольников.

$\nabla J(q) = A^*(Aq - f)$, где A^* — сопряжённый оператор.

Обозначение: $\Gamma[a(y), b(y)](x, y)$ — решение $v(x, y)$ задачи

$$\begin{cases} v_{xx} + v_{yy} + k^2 v = 0, & (x, y) \in (0, 1) \times (0, 1) \\ v|_{x=1} = a(y), & y \in [0, 1] \\ v_x|_{x=0} = b(y), & y \in [0, 1] \\ v|_{y=0} = v|_{y=1} = 0, & x \in [0, 1] \end{cases}$$

Gradient Precise Oracle

Вход: $q(y) \in H$; $f(y), g(y), k$ — параметры задачи

- 1 вычислить $u(x, y) = \Gamma[q(y), g(y)](x, y)$;
- 2 вычислить $\psi(x, y) = \Gamma[0, u(0, y) - f(y)](x, y)$;
- 3 вычислить $\nabla J(q)(y) = \psi_x(1, y)$.

Выход: $\nabla J(q)(y)$.

Similar Triangle Method (из статьи [Васин, Гасников, Спокойный])

Вход: $f : Q \rightarrow \mathbb{R}$, $x_{start} \in Q$, $L = 1$, N .

1 $\tilde{x}_0 = x_{start}$; $A_0 = \frac{1}{L}$; $\alpha_0 = \frac{1}{L}$; $z_0 = \tilde{x}_0 - \alpha_0 \tilde{\nabla} f(\tilde{x}_0)$; $x_0 = z_0$;

2 for $k = 1 \dots N$:

3 $\alpha_k = \frac{1}{2L} + \sqrt{\frac{1}{4L^2} + A_{k-1}}$;

4 $A_k = A_{k-1} + \alpha_k$;

5 $\tilde{x}_k = \frac{A_{k-1}x_{k-1} + \alpha_k z_{k-1}}{A_k}$;

6 $z_k = z_{k-1} - \alpha_k \tilde{\nabla} f(\tilde{x}_k)$;

7 $x_k = \frac{A_{k-1}x_{k-1} + \alpha_k z_k}{A_k}$;

Выход: x_N .

Теорема: $f(x_N) - f(x^*) \leq \frac{4LR^2}{N^2} + 3\tilde{R}\delta + \frac{N\delta^2}{2L}$.

Обозначение: $\Gamma_h[a_m, b_m]_{n,m}$ — решение $v_{n,m}$ задачи

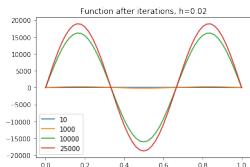
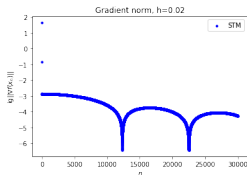
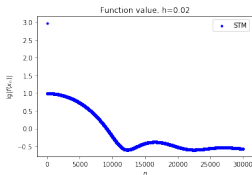
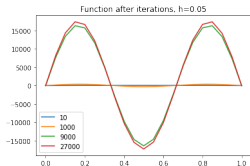
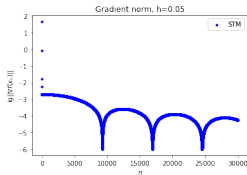
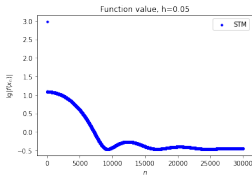
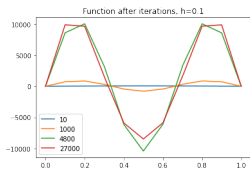
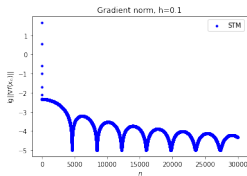
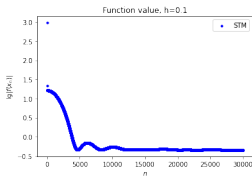
$$\begin{cases} \frac{v_{n+1,m} - 2v_{n,m} + v_{n-1,m}}{h^2} + \frac{v_{n,m+1} - 2v_{n,m} + v_{n,m-1}}{h^2} + k^2 v_{n,m} = 0, & 1 \leq n, m \leq N-1 \\ v_{N,m} = a_m, & 1 \leq m \leq N-1 \\ \frac{v_{1,m} - v_{0,m}}{h} = b_m, & 1 \leq m \leq N-1 \\ v_{n,0} = v_{n,N} = 0, & 0 \leq m \leq N \end{cases}$$

Gradient Approximate Oracle

Вход: $q(y) \in H$; $f_m = f(mh)$, $g_m = g(mh)$, k — параметры задачи, $h = \frac{1}{N}$ — шаг сетки

- 1 ВЫЧИСЛИТЬ $u_{n,m} = \Gamma_h[q_m, g_m]_{n,m}$;
- 2 ВЫЧИСЛИТЬ $\psi_{n,m} = \Gamma_h[0, u_{0,m} - f_m]_{n,m}$;
- 3 ВЫЧИСЛИТЬ $\tilde{\nabla} J(q)_m = \frac{\psi_{N,m} - \psi_{N-1,m}}{h}$.

Выход: $\tilde{\nabla} J(q)_m$.



Результаты экспериментов

h	time per iteration, s	function, min
0.1	0.34	0.4465
0.05	0.43	0.3448
0.02	1.74	0.2523

Оценка шума

Если используется n промежуточных точек, то $h = \frac{1}{n}$, шум аддитивный с $\delta = O\left(\frac{1}{n}\right)$.

Теорема о вычислительной эффективности (Плетнев, 2021)

Применение N итераций STM с приближённым вычислением градиента путём решения разностных задач требует $O(n^6 N)$ арифметических операций и обеспечивает невязку по функционалу $O\left(\frac{R^2}{N^2} + \frac{\tilde{R}}{n} + \frac{N}{n^2}\right)$

$\Gamma[a(y), b(y)](x, y)$ через ряды Фурье

$$a(y) = \sum_{n=1}^{\infty} \alpha_n \sin \pi n y, \quad b(y) = \sum_{n=1}^{\infty} \beta_n \sin \pi n y.$$

$\Gamma[a(y), b(y)](x, y) = \sum_{n=1}^{\infty} X_n(x) \sin \pi n y$, где $X_n(x)$ — решение задачи

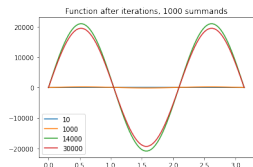
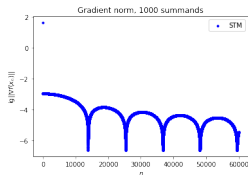
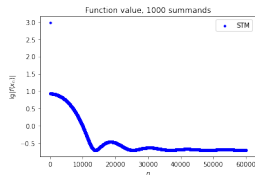
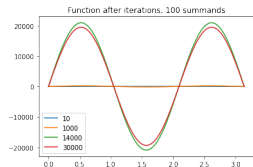
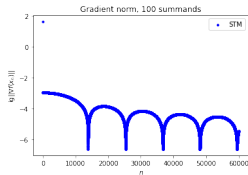
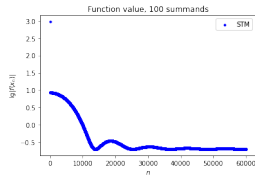
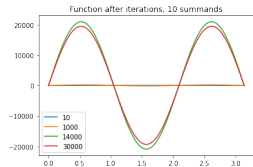
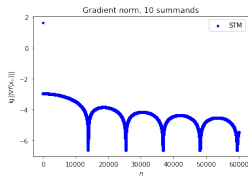
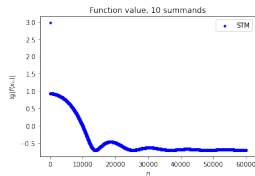
$$\begin{cases} X_n''(x) + (k^2 - \pi^2 n^2) X_n(x) = 0, & x \in (0, 1) \\ X_n(1) = \alpha_n \\ X_n'(0) = \beta_n \end{cases}$$

Градиент через ряды Фурье

$$q(y) = \sum_{n=1}^{\infty} q_n \sin \pi n y; \quad f(y) = \sum_{n=1}^{\infty} f_n \sin \pi n y, \quad g(y) = \sum_{n=1}^{\infty} g_n \sin \pi n y.$$

$$\nabla J(q)(y) = \sum_{n=1}^{\infty} \frac{1}{\operatorname{ch}^2 \gamma_n} \left(q_n - g_n \frac{\operatorname{sh} \gamma_n}{\gamma_n} - f_n \operatorname{ch} \gamma_n \right) \sin \pi n y.$$

Эксперименты с рядами Фурье



Результаты экспериментов

Fourier summands	time per iteration, s	function, min
10	0.87×10^{-3}	0.1984
100	0.91×10^{-3}	0.1993
1000	1.75×10^{-3}	0.1993

Оценка шума

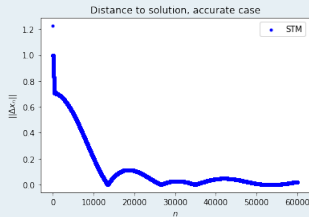
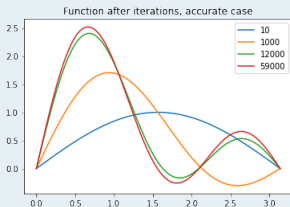
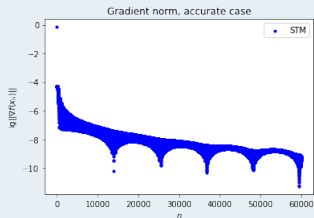
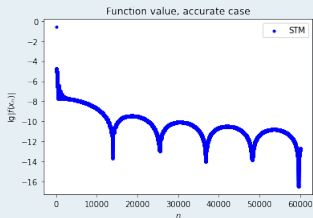
Если используется n слагаемых и p промежуточных точек при интегрировании, то шум аддитивный с $\delta = O\left(\frac{1}{n} + \frac{1}{p}\right)$.

Теорема о вычислительной эффективности (Плетнев, 2021)

Применение N итераций STM с использованием рядов Фурье для приближённого вычисления градиента требует $O((N+p)n)$ арифметических операций и обеспечивает невязку по функционалу $O\left(\frac{R^2}{N^2} + \frac{\tilde{R}}{n} + \frac{N}{n^2}\right)$.

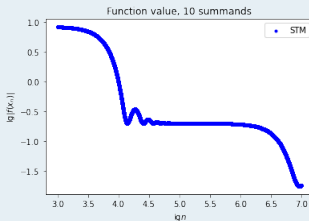
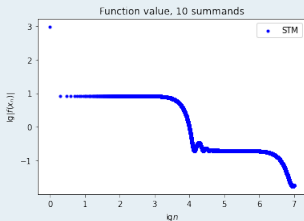
Модельный случай с точным решением

$$k = \pi, f(y) = \sin \pi y + \frac{\sin 3\pi y}{\operatorname{ch}\pi\sqrt{8}}, g(y) = \frac{\pi\sqrt{3}}{\operatorname{sh}\pi\sqrt{3}} \sin 2\pi y.$$
$$q^*(y) = \sin \pi y + \sin 2\pi y + \sin 3\pi y.$$



Определение концепции шума

10^7 итераций с рядами Фурье



Погрешность не накапливается, но стабилизируется. Отдельные локальные снижения.

Выводы

- Аддитивная концепция более адекватно описывает шум в задаче.
- Применение рядов Фурье ускоряет вычисления в тысячу раз и даёт меньшую невязку функционала.

Лемма (Плетнев, 2021)

Если верхняя оценка шума удовлетворяет неравенству

$$\delta \leq \min \left\{ \frac{f(x_0) - f(x^*)}{24R}, \frac{\sqrt[4]{L\mu} \sqrt{f(x_0) - f(x^*)}}{4\sqrt[4]{2}} \right\}, \text{ то } N = \sqrt{\frac{32L}{\mu}}$$

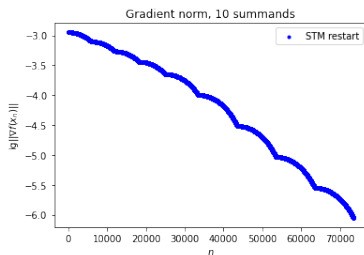
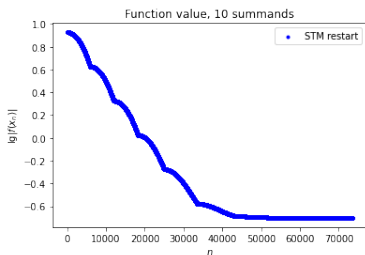
шагов метода подобных треугольников гарантируют уменьшение невязки функционала вдвое: $f(x_N) - f(x^*) \leq \frac{1}{2} (f(x_0) - f(x^*))$.

Теорема (Плетнев, 2021)

Если градиент выпуклой функции $f(x)$ вычисляется с аддитивной неточностью δ , и $\varepsilon \geq \max \left\{ 24R\delta, \frac{16\sqrt{2}\delta^2}{\sqrt{L\mu}} \right\}$, то достичь выполнения

условия $f(x) - f(x^*) < \varepsilon$ можно за не более, чем $\sqrt{\frac{32L}{\mu}} \log_2 \frac{f(x_0) - f(x^*)}{\varepsilon}$ итераций при использовании STM с рестартами при уменьшении функционала вдвое.

x_0 — начальная точка, x^* — точка минимума, $R = \|x_0 - x^*\|$, L — константа Липшица градиента, μ — (локальная) константа сильной выпуклости.



Эксперименты показывают, что применение рестартов не улучшает сходимость. Это подтверждает тот факт, что задача не является сильно выпуклой.

- 1 Алгоритм решения поставленной задачи оптимизации в гильбертовом пространстве с использованием STM и разложений в ряды Фурье, а также оценка его вычислительной эффективности;
- 2 Вывод об аддитивном характере неточности градиента в задаче;
- 3 Вычислительные эксперименты, подтверждающие выкладки;
- 4 Оценки сходимости STM с рестартами для оракула с аддитивным шумом.

Данные результаты планируются к публикации и обобщению на другие некорректно поставленные задачи.

- 1 Alexander Gasnikov, Sergey Kabanikhin, Ahmed Mohammed, Maxim Shishlenin: Convex optimization in Hilbert space with applications to inverse problems, <https://arxiv.org/abs/1703.00267>
- 2 Sergey Igorevich Kabanikhin, M. A. Shishlenin, D. B. Nurseitov, A. T. Nurseitova, and S.E.Kasenov: Comparative Analysis of Methods for Regularizing an InitialBoundary Value Problem for the Helmholtz Equation, <http://dx.doi.org/10.1155/2014/786326>
- 3 Ф. П. Васильев: Методы оптимизации, часть вторая: Оптимизация в функциональных пространствах. Регуляризация. Аппроксимация, ISBN 978-5-94057-708-9
- 4 Artem Vasin, Alexander Gasnikov, Vladimir Spokoiny: Stopping rules for accelerated gradient methods with additive noise in gradient, <https://arxiv.org/abs/2102.02921>