Двойственный алгоритм для задач регуляризации с недифференцируемыми стабилизаторами

А. С. Величко
Институт автоматики и процессов управления ДВО РАН, Владивосток, Россия
e-mail: vandre@dvo.ru

Величко А.С. Двойственный алгоритм для задач регуляризации с недифференцируемыми стабилизаторами // Вычисл. технологии. 2014. Т. 19, № 2. С. 14–19.

Для класса некорректно поставленных задач рассматривается подход регуляризации с недифференцируемыми стабилизирующими функционалами с позиций теории и методов оптимизации. Предложено эквивалентное представление исходной негладкой задачи в виде квадратичной оптимизационной задачи с линейными ограничениями-неравенствами. Рассматривается также параллельный численный метод, основанный на двойственной постановке оптимизационной задачи и нелинейном методе Якоби. Представленный подход используется для решения некорректно поставленной задачи большой размерности для интегрального уравнения Фредгольма первого рода, которая, в частности, возникает в задаче восстановления гравитационного поля Земли в математической геофизике.

Ключевые слова: регуляризация, условная оптимизация, численные методы, параллельный алгоритм.

Velichko A.S. Dual algorithm for regularization problems with nondifferentiable stabilizing functionals // Comput. Technologies. 2014. Vol. 19, No. 2. P. 14–19.

The paper examines theoretical approaches and numerical methods for constrained quadratic optimization problems which arise in regularization problems with non-differentiable stabilizing functionals.

Key words: regularization problem, constrained optimization, numerical method, parallel algorithm.

Введение

В работах [1, 2] описан метод регуляризации для решения некорректно поставленной задачи для операторного уравнения $A\mathbf{u} = f$, основанный на использовании недифференцируемых стабилизирующих функционалов. Классическая тихоновская регуляризация не всегда позволяет с приемлемым качеством восстановить истинное решение $\mathbf{u}$, поскольку функционалы, содержащие норму пространства Соболева, могут обладать эффектом "заглаживания" решения. Для повышения точности восстановления истинных решений требуется измельчение вычислительной сетки, что влечёт необходимость решать задачи большой размерности. Малые значения параметра регуляризации вызывают неустойчивость численных методов решения возникающих задач оптимизации.

Для решения возникающих задач негладкой оптимизации в [3, 4] использовались субградиентные, проксимальные методы и методы решения систем неравенств с помощью проективных методов. Указанные численные методы часто подвержены медленной скорости сходимости, а для задач большой размерности эта проблема усиливается.

Анализ альтернативных возможностей построения эффективных численных методов, в том числе параллельных, является предметом данного исследования.

1. Методы регуляризации и задачи оптимизации

В качестве примеров использования подхода регуляризации для некорректно поставленной задачи в работах [5, 6] рассматривалось интегральное уравнения Фредгольма $\int\limits_a^b K(t,s)u(s)\,ds = y(t)$, для которого можно записать различные задачи негладкой оптимизации. В настоящей работе для регуляризации уравнения Фредгольма предлагается задача безусловной недифференцируемой оптимизации вида

$$\min_{\{u_j\}} \left\{ \sum_{i=1}^{m} h_m \left[\left| \sum_{j=1}^{n} h_n K(t_i,s_j)u_j - y_i \right| + \left(\sum_{j=1}^{n} h_n K(t_i,s_j)u_j - y_i \right)^2 \right] + \alpha h_n \sum_{j=1}^{n} u_j^2 \right\}, \quad (1)$$

где $h_m,\ h_n$ — шаги сетки, α — параметр регуляризации, здесь и далее для краткости обозначений $u_j = u(s_j),\ y_i = y(t_i)$.

Классическая тихоновская регуляризация предполагает решение квадратичной задачи безусловной оптимизации

$$\min_{\{u_j\}} \left\{ \sum_{i=1}^{m} h_m \left(\sum_{j=1}^{n} h_n K(t_i,s_j)u_j - y_i \right)^2 + \alpha h_n \sum_{j=1}^{n} u_j^2 \right\}. \quad (2)$$

Альтернативой использования специальных методов негладкой оптимизации для решения задач (1), (2) может быть их предварительное эквивалентное преобразование к задачам квадратичной условной оптимизации с линейными ограничениями-неравенствами, для решения которых существуют эффективные численные методы [7].

Для задачи (1) возможно её эквивалентное представление в виде оптимизационной задачи с квадратичной целевой функцией и $n+m$ неизвестными переменными

$$\min_{\{u_j,w_i\}} \left\{ h_m \sum_{i=1}^{m} w_i + h_m \sum_{i=1}^{m} w_i^2 + \alpha h_n \sum_{j=1}^{n} u_j^2 \right\}$$

и с $2m$ линейными ограничениями-неравенствами

$$w_i \geq \sum_{j=1}^{n} h_n K(t_i,s_j)u_j - y_i, \quad w_i \geq -\left(\sum_{j=1}^{n} h_n K(t_i,s_j)u_j - y_i \right).$$

Данный приём, позволяющий “избавиться” от негладкой оптимизируемой функции, не является универсальным, а обусловлен именно поиском минимума оптимизируемого выпуклого функционала. Возможна дополнительная замена переменных $z_j = \sqrt{\alpha}u_j$, позволяющая частично решить проблему неустойчивости по выбору малого параметра регуляризации численных методов решения возникающих задач оптимизации.

2. Двойственное представление задачи, численные методы и алгоритмы

Запишем последнюю задачу в виде

$$\min_x \left\{ \frac{1}{2}\mathbf{x}'Q\mathbf{x} + \mathbf{c}'\mathbf{x},\ A\mathbf{x} \le \mathbf{b} \right\},$$

где $\mathbf{x} = (\mathbf{u}, \mathbf{w})'$, $\mathbf{u}$ и $\mathbf{w}$ — векторы из элементов $\{u_j\}$ и $\{w_j\}$ соответственно, $Q = \begin{pmatrix} \alpha h_n E_n & \mathbf{0} \\ \mathbf{0} & h_n E_m \end{pmatrix}$ — диагональная матрица,

$$A = \begin{pmatrix} h_n K & -E_m \\ -h_n K & -E_m \end{pmatrix}, \quad \mathbf{c} = (\mathbf{0}_n, h_m \mathbf{e}_m), \quad \mathbf{b} = \begin{pmatrix} \mathbf{y} \\ -\mathbf{y} \end{pmatrix}.$$

Здесь E_k — единичная матрица размерности k, $\mathbf{0}$ — матрица из нулей соответствующей размерности, K — матрица из элементов $\{K(t_i, s_j)\}$, $\mathbf{0}_n$ — вектор-строка из нулей длины n, $\mathbf{e}_m$ — вектор-строка из единиц длины m, $\mathbf{y}$ — вектор из элементов $\{y_i\}$.

Следуя [8], можно записать двойственную постановку рассматриваемой задачи:

$$\Phi(\mathbf{p}^*) = \min_{\mathbf{p}} \left\{ \frac{1}{2}\mathbf{p}'D\mathbf{p} + \mathbf{d}'\mathbf{p},\ \mathbf{p} \ge 0 \right\}.$$

Здесь $D = AQ^{-1}A'$ — положительно определённая, симметричная матрица, $\mathbf{d} = \mathbf{b} + AQ^{-1}\mathbf{c}$ и для оптимальных решений прямой ($\mathbf{x}^*$) и двойственной ($\mathbf{p}^*$) постановки задачи выполняется соотношение $\mathbf{x}^* = -Q^{-1}(\mathbf{c} + A'\mathbf{p}^*)$. Размерность двойственного вектора $\mathbf{p}$ равна $2m$ — числу линейных ограничений-неравенств рассматриваемой квадратичной задачи условной оптимизации.

Идея параллельного алгоритма для решения двойственной задачи основана на применении последовательного нелинейного метода Якоби [8] и предлагаемой ниже модификации, позволяющей распараллеливание вычислений. Используемый для решения задач оптимизации нелинейный метод Якоби предполагает фиксацию значений всех, кроме одной, компонент вектора неизвестных. На каждом шаге алгоритма решается получающаяся одномерная задача оптимизации, обновляется значение компоненты вектора неизвестных на данной итерации алгоритма и производится переход к другой компоненте искомого вектора, при этом компоненты циклически перебираются до выполнения критерия останова алгоритма. В отличие от методов покоординатного спуска в рассматриваемом алгоритме не вычисляется градиент оптимизируемой функции в текущей точке и не осуществляется спуск в направлении антиградиента функции.

Шаг $s+1$ предлагаемого параллельного алгоритма состоит из трёх подшагов.

1. Пусть $\mathbf{p}^s = (p_1^s, \ldots, p_i^s, \ldots, p_k^s)'$ — приближённое решение, полученное на предыдущем шаге s. Формируется $k = 2m$ одномерных квадратичных подзадач ($2m$ — размерность двойственного вектора $\mathbf{p}$) вида $\min\left\{\frac{1}{2}\mathbf{p}^{(i)}D\mathbf{p}^{(i)\prime} + \mathbf{d}'\mathbf{p}^{(i)},\ p_i \ge 0\right\}$, где $\mathbf{p}^{(i)} = (p_1^s, \ldots, p_i, \ldots, p_k^s)'$, а индекс i принимает значения от 1 до k. Получаемые на данном подшаге одномерные задачи вида $\min\{av^2 + bv\}$ при условии $v \ge 0$ для $a > 0$ очевидно имеют решение $v^* = \max\{0, -b/2a\}$. Вычисление k векторов $\mathbf{p}^{(i)}$ осуществляется в виде параллельных процессов.

2. Определяется индекс i^* из условия

$$i^* = \arg\min_i \left\{ \frac{1}{2}\mathbf{p}^{(i)'} D\mathbf{p}^{(i)} + \mathbf{d}'\mathbf{p}^{(i)} \right\}.$$

Формируется новое приближенное решение $\mathbf{p}^{s+1} = (p_1^s, \ldots, p_{i^*}, \ldots, p_k^s)'$.

3. Если выполняется критерий останова $||\mathbf{p}^{s+1} - \mathbf{p}^s|| \leq \varepsilon$ для наперёд заданного малого числа $\varepsilon > 0$, то алгоритм завершает свою работу, возвращая приближённое решение $\mathbf{p}^{s+1}$, иначе — значение s увеличивается на единицу и алгоритм продолжает свое выполнение на следующем шаге.

Модификация исходного нелинейного алгоритма Якоби по сути состоит в дополнительной проверке условия на втором подшаге, что не составляет труда, поскольку это требует поиска минимального значения из уже вычисленных $2m$ значений оптимизируемой функции, однако эта проверка составляет непараллелизуемую часть алгоритма.

Параллельный алгоритм для многопроцессорной вычислительной архитектуры с распределённой памятью реализован на языке Matlab и Octave [9] с использованием библиотек функций интерфейса передачи сообщений MPI в реализации OpenMPI [10] и “OpenMPI Extension for Octave” [11]. Расчёты проводятся на многопроцессорном вычислительном комплексе Центра коллективного пользования ДВО РАН во Владивостоке “Дальневосточный вычислительный ресурс” [12].

В численных расчётах ядро $K(t,s)$ уравнения Фредгольма имеет вид $\dfrac{H}{(t-s)^2 + H^2}$, где $H = 10$, что приводит к числу обусловленности для матрицы с элементами $K(t_i, s_i)$ порядка 10^{16}.

Модельное решение

$$u(s) = \frac{10}{4\sqrt{2\pi}} \exp\left\{-\frac{s^2}{32}\right\} + \frac{10}{\sqrt{2\pi}} \exp\left\{-\frac{(s-4)^2}{2}\right\} -$$

сумма гауссианов на отрезке $[-10, 10]$. Значения правой части уравнения Фредгольма $y(t)$ брались не с использованием решения интегрального уравнения, а по аппроксимационной формуле для интеграла в узлах t_i, когда значения $u(s_j)$ берутся по истинной (наперёд заданной) функции $u(s)$ в узлах сетки s_j. Параметр регуляризации α принимался равным 10^{-6}.

На рис. 1, *а* показано восстановление модельного решения с помощью решения оптимизационной задачи (1) с её последующим эквивалентным преобразованием, предлагаемым в данной работе, на рис. 1, *б* — то же, для классической тихоновской регуляризации в результате решения задачи (2).

Величина “ускорения” R_n параллельного алгоритма определяется как отношение времени выполнения алгоритма на одном процессоре к времени выполнения алгоритма на n параллельных процессорах. Известная теоретическая оценка $R_n = \dfrac{n}{(n-1)a + 1}$ (закон Амдала), где a — доля непараллельного кода алгоритма. При “идеальном” распараллеливании процесса вычислений $a = 0$, и тогда $R_n = n$ [8]. Для рассматриваемого в настоящей работе параллельного алгоритма на рис. 2 показаны график фактического ускорения (speedup) (сплошная линия) и аппроксимация теоретической зависимости для R_n с оценкой для доли непараллельного кода алгоритма $a = 0.06$ (штриховая

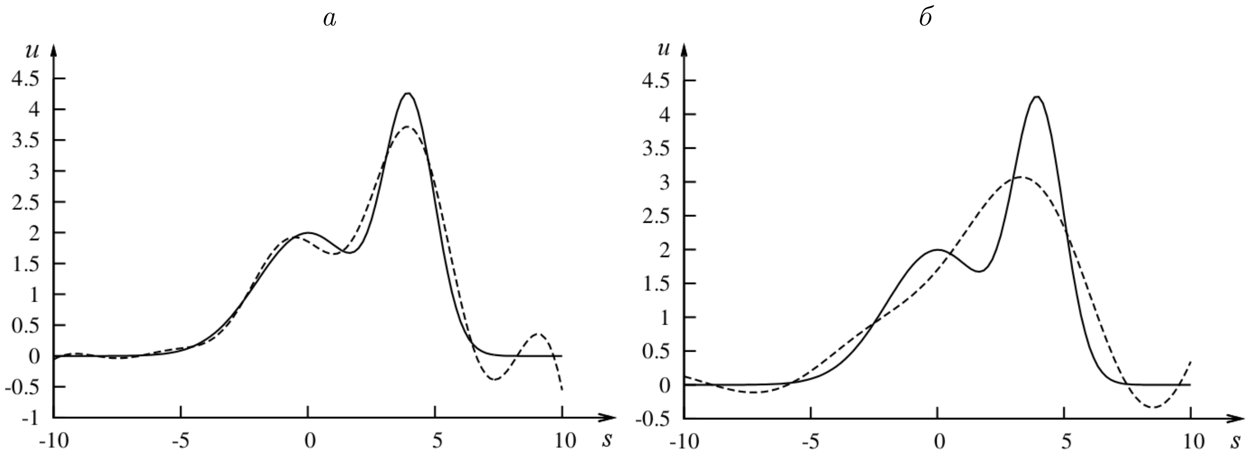

Рис. 1. Модельное (сплошная линия) и восстановленное (штриховая линия) решения для задач (1) (*а*) и (2) (*б*)

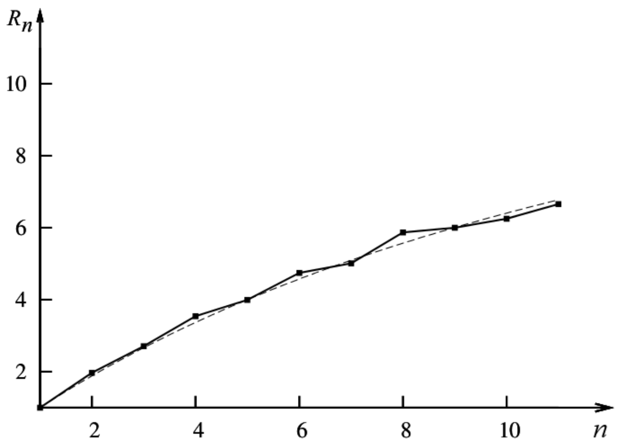

Рис. 2. График ускорения параллельного алгоритма

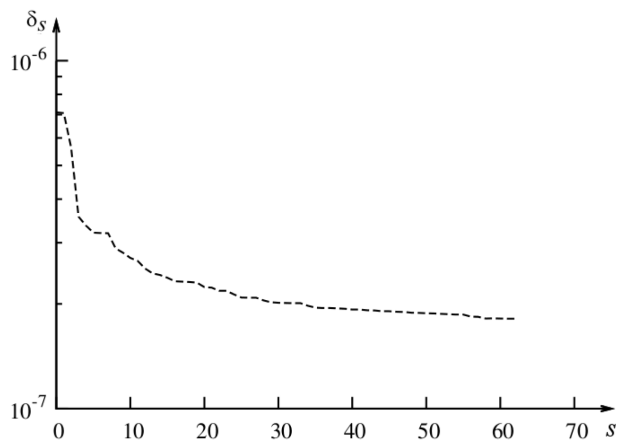

Рис. 3. Характер сходимости алгоритма

линия). На рис. 3 представлен график, демонстрирующий характер сходимости алгоритма: приведены значения $\delta_s = |\Phi(p^s) - \Phi(p^*)|$, т. е. абсолютные отклонения значений оптимизируемой функции, вычисленных для приближённых решений, получаемых на шагах алгоритма, от оптимального значения этой функции.

Для балансировки вычислительной нагрузки параллельный алгоритм может на втором подшаге алгоритма предусматривать решение не только одной, но и нескольких одномерных задач на одном физическом процессоре (ядре) многопроцессорной вычислительной системы.

Заключение

В работе на примере решения некорректно поставленной задачи для интегрального уравнения Фредгольма первого рода, возникающей в задаче восстановления гравитационного поля Земли в математической геофизике, с позиций теории и методов оптимизации предлагается решение квадратичной оптимизационной задачи с линейными ограничениями-неравенствами в двойственной постановке с помощью параллельного численного метода, родственного нелинейному методу Якоби.

Параллельный алгоритм демонстрирует хорошую степень масштабируемости, т. е. сохраняет характер поведения ускорения с ростом числа задействованных процессоров. Использование предлагаемой квадратичной задачи показывает хорошее восстановление двухпикового модельного решения в виде композиции гауссианов, что является обычным тестом качества применяемой методики в подобных задачах. Проведённые вычислительные эксперименты указывают на возможность эффективного использования предложенного подхода для решения других задач.

Список литературы

[1] Васин В.В. Устойчивая аппроксимация негладких решений некорректно поставленных задач // Докл. АН. 2005. Т. 402, № 5. С. 586–589.

[2] Васин В.В., Агеев А.Л. Некорректные задачи с априорной информацией. Екатеринбург: Наука, 1993. 262 с.

[3] Васин В.В., Ерёмин И.И. Операторы и итерационные процессы фейеровского типа. Теория и приложения. Москва, Ижевск: НИЦ РХД, 2005. 200 с.

[4] Васин В.В. Итерационные процессы фейеровского типа в некорректных задачах с априорной информацией // Изв. высших учебных заведений. Математика. 2009. № 2. С. 3–24.

[5] Vasin V.V., Korotkii M.A. Tikhonov regularization with nondifferentiable stabilizing functionals // J. of Inverse and Ill-posed Problems. 2007. Vol. 15, No. 8. P. 853–865.

[6] Васин В.В., Сережникова Т.И. Регулярный алгоритм аппроксимации негладких решений для интегральных уравнений Фредгольма первого рода // Вычисл. технологии. 2010. Т. 15, № 2. С. 15–23.

[7] Нестеров Ю.Е. Введение в выпуклую оптимизацию. М.: МЦНМО, 2010. 274 с.

[8] Bertsekas D.P, Tsitsiklis J.N. Parallel and Distributed Computation: Numerical Methods. Nashua: Athena Sci., 1997.

[9] GNU Octave. URL: http://www.octave.org.

[10] OpenMPI. URL: http://www.open-mpi.org.

[11] OpenMPI Extension for Octave. URL: http://octave.sourceforge.net/openmpi_ext/index.html.

[12] Центр коллективного пользования ДВО РАН “Дальневосточный вычислительный ресурс”. URL: http://www.cc.dvo.ru.

Поступила в редакцию 20 января 2014 г.