

МЕТОДЫ И ТЕХНОЛОГИИ РЕШЕНИЯ БОЛЬШИХ ЗАДАЧ

МОСКВА 2004

Российская академия наук Институт вычислительной математики

МЕТОДЫ И ТЕХНОЛОГИИ РЕШЕНИЯ БОЛЬШИХ ЗАДАЧ

Сборник научных трудов под редакцией В. И. Агошкова и Е. Е. Тыртышникова

Методы и технологии решения больших задач: Сборник научных трудов / Под ред. В. И. Агошкова и Е. Е. Тыртышникова. — М.: Институт вычислительной математики РАН, 2004. — 196 с. -ISBN 5-901854-10-1

В данном сборнике представлены работы, выполненные в рамках программы приоритетных фундаментальных исследований Отделения математических наук РАН "Вычислительные и информационные проблемы решения больших задач" по двум направлениям: "Методы решения задач вариационного усвоения данных наблюдений и управление сложными системами" и "Матричные методы в интегральных и дифференциальных уравнениях". Сборник рассчитан на широкий круг научных работников, специализирующихся в области численных методов и разработки современного программного обеспечения.

This book includes papers concluding results of the advanced basic research program Computational and Informational Problems in Solving Large-Scale Problems conducted in the Mathematical Department of the Russian Academy of Science within the following two lines: "Methods for Variational Data-Asimilation Problems and Control of Complex Systems" and "Matrix Methods in Integral and Differential Equations".

The book is intended for wide audience of scientists, specialising on numerical methods and development of contemporary software.

Печатается по решению Ученого совета Института вычислительной математики Российской академии наук

ISBN 5-901854-10-1 \bigcirc Институт вычислительной математики РАН, 2004.

Посвящается 25-летию Института вычислительной математики Российской академии наук

ПРЕДИСЛОВИЕ

В данном сборнике представлены работы, выполненные в рамках программы приоритетных фундаментальных исследований Отделения математических наук РАН "Вычислительные и информационные проблемы решения больших задач" по двум направлениям: "Методы решения задач вариационного усвоения данных наблюдений и управление сложными системами" и "Матричные методы в интегральных и дифференциальных уравнениях".

Разработка комплекса вычислительных алгоритмов решения многомерных нестационарных задач вариационного усвоения данных наблюдений (спутниковых, измерений с кораблей и др.) является одной из важнейших современных вычислительных и информационных проблем. Подходы к решению данных проблем базируются на ряде разделов современной математики: оптимальном управлении, теории некорректных задач, функциональном анализе, теории краевых задач математической физики, сопряженных уравнениях, современных вычислительных методах.

С другой стороны, решение систем дифференциальных и интегральных уравнений, возникающих при исследовании и численном решении в научно-технических и инженерных приложениях, таких как задачи вариационного усвоения данных наблюдений, часто приводит к большим матричным задачам с числом неизвестных и уравнений порядка сотен миллионов в случае разреженных матриц и порядка миллионов в случае плотных матриц. Такие задачи требуют особых вычислительных технологий и подходов.

Поэтому проведение исследований по отмеченным выше двум направлениям в свою очередь требует объединения усилий многих научных сотрудников ИВМ РАН с целью концентрации целенаправленного проведения научного поиска. Авторы работ настоящего сборника с благодарностью отмечают поддержку данных исследований Отделением математических наук РАН, а также грантами

РФФИ "Сопряженные уравнения и методы оптимального управления в обратных задачах и вычислительных процессах", "Параллельные многоуровневые технологии решения трехмерных задач на неструктурированных и/или сгущающихся сетках", "Теория и методы аппроксимации функций, операторов и матриц с использованием блоков малого ранга и иерархических базисов".

Кроме того, решение отмеченных выше задач невозможно без привлечения современной вычислительной техники. Поэтому в 2004 году в ИВМ РАН был внедрен собственный вычислительный кластер. Его появлению способствовала поддержка как со стороны упомянутых выше программы Отделения математических наук РАН, проектов и грантов РФФИ, а также многих других грантов, получаемых нашим институтом для проведения фундаментальных исследований в области вычислительной математики, математического моделирования и геофизической гидродинамики, математических методов в иммунологии, теории параллельных вычислений.

Весь комплекс всех данных исследований со всеми их взаимодополняющими идеями и подходами позволяет с оптимизмом смотреть на последующие 25 лет научной деятельности нашего института.

В. И. Агошков, Е. Е. Тыртышников

Часть І

Методы решения задач вариационного усвоения данных наблюдений и управления сложными системами

Об одной методике усвоения данных

В. М. Ипатова

Предлагается процедура усвоения данных, такая что функционал стоимости минимизируется на множестве управлений, при которых заданное начальное состояние является точкой минимума. Доказаны существование и единственность решения оптимизационной задачи. Вычислены собственные значения оператора управления.

Различные методики обработки данных наблюдений широко применяются в науке и практике как для уточнения наименее известных параметров моделей, так и для исправления самих наблюдаемых полей, обычно неполных и зашумлённых. В [1] предложен общий подход к решению задач идентификации и вариационного усвоения данных, основанный на использовании методов оптимального управления и аппарата сопряжённых уравнений в случае, когда за критерий оптимальности берётся минимум целевого функционала стоимости. Можно предположить, что в определённых ситуациях будут полезны и некоторые другие критерии оптимальности, поэтому представляет интерес изучение следующей задачи. Пусть точно известно начальное состояние системы u_0 , а распределённое внешнее силовое воздействие f требуется определить с помощью функционала S(u), характеризующего расхождение между модельным решением $u = u(u_0, f)$ и данными наблюдений. При поиске неизвестной функции f будем руководствоваться следующими соображениями:

1) потребуем, чтобы при данном фиксированном f функционал S(u) нельзя было уменьшить, выбирая вместо u_0 какое-либо другое начальное состояние. Иначе говоря, u_0 является предписанной заранее точкой минимума, и нам подходят только те функции f, для которых этот минимум имеет место;

2) среди всех f, удовлетворяющих условию 1, выберем ту функцию, на которой S(u) принимает своё наименьшее возможное значение.

Разумеется, функционал S(u) можно было бы уменьшить, если отказаться от первого требования и провести оптимизацию по всему пространству. Поэтому рассматриваемый здесь алгоритм характеризует ситуацию, в которой информация о начальном состоянии u_0 считается гораздо более достоверной, чем содержащаяся в S(u) информация о данных наблюдений. К тому же сам функционал S(u) обычно включает в себя регуляризующие добавки, которые никак не связаны с наблюдаемыми полями, но позволяют корректно поставить задачу и эффективно организовать вычисления. В этом случае становится очевидным, что наименьшее значение функционала ещё не означает наилучшего согласования с наблюдениями, и кажется вполне разумным требование о существовании предписанной точки минимума.

В настоящей работе исследуются вопросы существования и единственности решения задачи 1-2. Метод доказательства состоит в том, что непосредственно вычисляются собственные значения оператора, входящего в уравнение для управления. Аналогичный способ был использован мной в работе [2]. Перейдём к точной формулировке нашей задачи.

Пусть A – не зависящий от времени линейный замкнутый оператор, действующий в гильбертовом пространстве H с областью определения, плотной в H. Будем предполагать, что A – неограниченный самосопряженный положительно определенный оператор в H, имеющий вполне непрерывный обратный.

Обозначим: $[\cdot,\cdot]$ – скалярное произведение в $H; X^{\alpha}, \alpha \in \mathbf{R}$ – гильбертовы пространства со скалярным произведением $[u,v]_{\alpha}==[A^{\alpha}u,A^{\alpha}v]; Y^{\alpha}=L_2(0,T;X^{\alpha}),$ где $T<+\infty, Y=Y^0=L_2(0,T;H);$ $(\cdot,\cdot),$ $\|\cdot\|$ – скалярное произведение и норма в Y; $(\cdot,\cdot)_{\alpha},$ $\|\cdot\|_{\alpha}$ – скалярное произведение и норма в $Y^{\alpha};$ $W=\{u\in Y^{\alpha+1/2},\frac{du}{dt}\in Y^{\alpha-1/2}\};$ W^* – сопряженное с W пространство при отождествлении Y со своим сопряженным.

Рассмотрим эволюционную задачу

$$\frac{du}{dt} + Au = f, \quad t \in (0, T); \quad u(0) = u_0,$$

ее разрешающий оператор $G: X^{\alpha} \times Y^{\alpha-1/2} \to W$ определим равенством $u = G(u_0; f) \quad \forall u_0 \in X^{\alpha} \quad \forall f \in Y^{\alpha - 1/2}$. Пусть $Q_0 : X^{\alpha} \to W$ и $Q:Y^{\alpha-1/2} \to W$ есть сужения G на множества $X^{\alpha} \times \{0\}$ и $\{0\} imes Y^{lpha-1/2}$ соответственно. Сопряженные операторы $Q_0^*:W^* o$ X^{lpha} и $Q^*:W^* o Y^{1/2-lpha}$ определим сооотношениями

$$(Q_0 y, g) = [y, Q_0^* g]_{\alpha} \quad \forall y \in X^{\alpha} \quad \forall g \in W^*,$$

$$(Qf, g) = (f, Q^* g) \quad \forall f \in Y^{\alpha - 1/2} \quad \forall g \in W^*.$$

Пусть на W задан функционал

$$S(u) = \frac{\gamma}{2} \|u_t + Au\|_{\alpha - 1/2}^2 + \frac{\sigma}{2} \|u\|_{\alpha + 1/2}^2 + \frac{1}{2} \|u - \hat{u}\|_{\beta}^2,$$

где $\gamma > 0, \ \sigma \geqslant 0, \ \beta \leqslant \alpha + 1/2$. При заданных $u_0 \in X^{\alpha}$ и $\hat{u} \in Y^{\beta}$ введём множество всех функций $g \in Y^{\alpha-1/2}$, для которых $S(G(u_0;g)) \leqslant S(G(v_0;g))$ $\forall v_0 \in X^\alpha$, и рассмотрим задачу об отыскании функции f такой, что:

- (i) $f \in F$;
- (ii) $S(G(u_0; f)) \leq S(G(u_0; g)) \quad \forall g \in F.$

Пусть $\{w_n\}$ – полная ортонормированная система в H, составленная из собственных функций $A: Aw_n = \lambda_n w_n, \ \lambda_n > 0,$ $\lambda_n \to +\infty$ при $n \to \infty$. Справедлива

Теорема 1. Если $\gamma > 0$ и $\sigma > 0$, $\beta \leqslant \alpha + 1/2$ или $\sigma \geqslant 0$, $\beta = \alpha + 1/2$, то задача (i)-(ii) имеет единственное решение при всех $u_0 \in X^{\alpha}$, $\hat{u} \in Y^{\beta}$.

Доказательство. Покажем, что множество F не пусто. Функция $f \in F$, если первая вариация S(u) равна нулю на $u = G(u_0; f)$ и всех $v \in Q_0 X^{\alpha}$.

Возьмём $v = v_n = Q_0 w_n = e^{-\lambda_n t} w_n$, тогда

$$\delta S(u)v_n = \sigma(u, \lambda_n^{2\alpha+1}v_n) + (u - \hat{u}, \lambda_n^{2\beta}v_n) = 0.$$

Обозначим через $\psi_n(t) = e^{-\lambda_n t} - e^{\lambda_n (t-2T)}, \ c_n = \int_0^T \psi_n^2(t) dt = \frac{1}{2\lambda_n} (1-t)$

 $e^{-4\lambda_n T}) - 2Te^{-2\lambda_n T}, \ m_n = \frac{1}{2}(\sigma \lambda_n^{2\alpha} + \lambda^{2\beta-1}).$ Из уравнения $-q_t + Aq = v_n, \ q(T) = 0,$ находим $q = q_n = \frac{1}{2\lambda_n} \psi_n(t) w_n.$ Используя равенство $(u,v_n)=(f,q_n)+\frac{1}{2\lambda_n}(1-e^{-2\lambda_n T})[u_0,w_n]$, имеем

$$\delta S(u)v_n = m_n \int_0^T \psi_n(t)[f,w_n]dt + m_n(1-e^{-2\lambda_n T})[u_0,w_n] - \lambda_n^{2\beta}(\hat{u},v_n) = 0,$$

то есть

$$\int_0^T \psi_n(t)[f, w_n] dt = \frac{\lambda_n^{2\beta}}{m_n} (\hat{u}, v_n) - (1 - e^{-2\lambda_n T})[u_0, w_n].$$

Следовательно, минимальное по норме $L_2(0,T)$ значение $[f,w_n]$ равно

$$\hat{f}_n(t)dt = \frac{1}{c_n} \left(\frac{\lambda_n^{2\beta}}{m_n} \int_0^T e^{-2\lambda_n t} [\hat{u}, w_n] dt - (1 - e^{-2\lambda_n T}) [u_0, w_n] \right) \psi_n(t).$$

Нетрудно проверить, что функция $\hat{f} = \sum_{n=1}^{\infty} \hat{f}_n(t) w_n$ принадлежит $Y^{\alpha-1/2}$ и по построению $\hat{f} \in F$. Теперь осталось заметить, что линейное многообразие F слабо замкнуто в $Y^{\alpha-1/2}$, функционал $J(f) = S(G(u_0; f)), f \in F$, является строго выпуклым и слабо полунепрерывным снизу в $Y^{\alpha-1/2}$, а любая его минимизирующая последовательность ограничена, из чего следует, что J(f) имеет на F точку минимума и эта точка минимума единственна.

Обозначим через $B=\sigma A^{2\alpha+1}+A^{2\beta},\ D=\gamma A^{2\alpha-1}+Q^*BQ.$ Условия (i)-(ii) можно записать в виде

(i)
$$Q_0^*BQf = Q_0^*(A^{2\beta}\hat{u} - BQ_0u_0),$$

(ii)
$$Df - Q^*A^{2\beta}\hat{u} = Q^*BQ_0r_0, \quad r_0 \in X^{\alpha}.$$

Исключая из этой системы f, получаем для r_0 уравнение $Mr_0=\eta$, где $M=Q_0^*BQD^{-1}Q^*BQ_0$, $\eta=Q_0^*(A^{2\beta}\hat{u}-BQ_0u_0)-Q_0^*BQD^{-1}Q^*A^{2\beta}\hat{u}$. **Теорема 2.** Оператор M имеет собственные функции w_n , соответствующие собственным значениям

$$\mu_n = \frac{1}{2} (\sigma + \lambda_n^{2\beta - 2\alpha - 1}) (1 - e^{-2\lambda_n T}) - \frac{\lambda_n (\sigma + \lambda_n^{2\beta - 2\alpha - 1})}{a_n + \lambda_n + (a_n - \lambda_n) e^{-2a_n T}} (1 - e^{-2\lambda_n T}),$$

где
$$a_n = \lambda_n \sqrt{1 + (\sigma + \lambda_n^{2\beta - 2\alpha - 1})/\gamma}$$
 .

Если выполнены условия теоремы 1, то решением задачи (i)-(ii) является

$$f = D^{-1}Q^*A^{2\beta}\hat{u} + \frac{1}{\gamma} \sum_{n=1}^{\infty} \frac{f_n}{\mu_n} \Big(e^{-a_n t} e^{a_n (t-2T)} \Big) w_n,$$

где

$$f_n = \frac{\lambda_n^2}{\rho_n} (\sigma + \lambda_n^{2\beta - 2\alpha - 1}) \left(\frac{\lambda_n^{2\beta - 2\alpha}}{\rho_n} \int_0^T h_n(t) [\hat{u}, w_n] dt - \frac{1}{2} (\sigma + \lambda_n^{2\beta - 2\alpha - 1}) (1 - e^{-2\lambda_n T}) [u_0, w_n] \right),$$

$$h_n(t) = (a_n + \lambda_n)e^{-a_n t} + (a_n - \lambda_n)e^{a_n(t-2T)}, \ \rho_n = h_n(0).$$

Доказательство. Пусть w – собственная функция $A,\ Aw = \lambda w,$ тогда

$$Q^*BQ_0w=rac{b}{2\lambda}\psi(t)w,$$
 где $b=\sigma\lambda^{2lpha+1}+\lambda^{2eta},$ $\psi(t)=e^{-\lambda t}-e^{\lambda(t-2T)}.$

Для $\xi = D^{-1}Q^*BQ_0w$ имеем $\xi = y(t)w$

$$\gamma \lambda^{2\alpha-1} \xi + bQ^*Q\xi = \frac{b}{2\lambda} \psi(t)w,$$

из чего находим $y(T)=0,\ \gamma\lambda^{2\alpha-1}(-y^{'}+\lambda y)w+bQ\xi=be^{-\lambda t}w.$ Следовательно,

$$-y^{'}(0) + \lambda y(0) = b\lambda^{1-2\alpha}/\gamma, \qquad \gamma \lambda^{2\alpha-1}(-y^{''} + \lambda^{2}y) + by = 0,$$

то есть $y^{''}=a^2y$, где $a=\sqrt{\lambda^2+b\lambda^{1-2lpha}/\gamma}$. Таким образом,

$$y(t) = \frac{b(e^{-at} - e^{a(t-2T)})}{\gamma \lambda^{2\alpha - 1}(\lambda + a + (a - \lambda)e^{-2aT})}.$$

Далее рассмотрим уравнение $-q_t + Aq = bQy(t)w$, q(T) = 0. Его решением будет q = z(t)w, где

$$-z'' + \lambda^2 z = by(t),$$
 $z(T) = 0,$ $-z'(0) + \lambda z(0) = 0,$

то есть

$$z(t) = \frac{1}{2}\lambda^{2\alpha}(\sigma + \lambda^{2\beta - 2\alpha - 1})(e^{-\lambda t} - e^{\lambda(t - 2T)}) - \frac{\lambda^{2\alpha + 1}(\sigma + \lambda^{2\beta - 2\alpha - 1})(e^{-at} - e^{a(t - 2T)})}{\lambda + a + (a - \lambda)e^{-2aT}}.$$

Заметив, что $Mw=\lambda^{-2\alpha}z(0)w$, получаем собственные значения M. Для того чтобы построить разложение $\eta=\sum_{n=1}^\infty \eta_n w_n$, вычислим

$$[\eta, w] = [\eta, \lambda^{-2\alpha} w]_{\alpha} = \lambda^{2\beta - 2\alpha} (\hat{u}, Q_0 w - Q y(t) w) - b \lambda^{-2\alpha} (Q_0 u_0, Q_0 w) = \frac{\lambda^{2\beta - 2\alpha}}{h(0)} \int_0^T h(t) [\hat{u}, w] dt - \frac{1}{2} (\sigma + \lambda^{2\beta - 2\alpha - 1}) (1 - e^{-2\lambda T}) [u_0, w],$$

где $h(t)=(\lambda+a)e^{-at}+(a-\lambda)e^{a(t-2T)}$. Далее воспользуемся равенствами $r_0=\sum_{n=1}^\infty \eta_n w_n/\mu_n$ и $D^{-1}Q^*BQ_0w=y(t)w$ для вычисления $f=D^{-1}(Q^*A^{2\beta}\hat{u}+Q^*BQ_0r_0)$.

В рассмотренном нами функционале S(u) функция \hat{u} представляет собой некоторые данные наблюдений, в то время как первые два члена введены из чисто технических соображений и служат для регуляризации задачи. Интересно отметить, что при использовании для этой цели эквивалентных норм могут возникать ситуации, когда исследуемая задача не имеет решения. Введём функционал

$$S(u) = \frac{\gamma}{2} \|u_t\|_{\alpha - 1/2}^2 + \frac{\sigma}{2} \|u\|_{\alpha + 1/2}^2 + \frac{1}{2} \|u - \hat{u}\|_{\beta}^2 \tag{1}$$

и для него вновь будем рассматривать задачу (i)-(ii).

Теорема 3. Если $\gamma > 0$ и $\sigma > 0$, $\beta < \alpha + 1/2$, $\sigma \neq \gamma$, или $\sigma \geqslant 0$, $\beta = \alpha + 1/2$, $\sigma \neq \gamma - 1$, то задача (i)-(ii) для функционала S(u), заданного в (1), имеет единственное решение при всех $u_0 \in X^{\alpha}$, $\hat{u} \in Y^{\beta}$.

Доказательство. Достаточно показать, что множество F не пусто. Функция $f \in F$, если для $u = G(u_0; f)$ и всех $v \in Q_0 X^{\alpha}$ имеет место равенство

$$\delta S(u)v = \gamma(u_t, A^{2\alpha - 1}v_t) + \sigma(Au, A^{2\alpha}v) + (u - \hat{u}, A^{2\beta}v) = 0.$$

Возьмём $v = v_n = Q_0 w_n = e^{-\lambda_n t} w_n$. Используя соотношения $u_t = f - Au$,

$$(u, v_n) = (f, q_n) + \frac{1}{2\lambda_n} (1 - e^{-2\lambda_n T}) [u_0, w_n], \ q_n = \frac{1}{2\lambda_n} (e^{-\lambda_n t} - e^{\lambda_n (t - 2T)}) w_n,$$

находим

$$\delta S(u)v_n = \gamma(f, -\lambda_n^{2\alpha}v_n) + \left((\gamma + \sigma)\lambda_n^{2\alpha+1} + \lambda_n^{2\beta}\right)(u, v_n) - \lambda_n^{2\beta}(\hat{u}, v_n)$$

$$\frac{1}{2} \int_0^T \psi_n(t)[f, w_n]dt + \frac{1}{2} m_n (1 - e^{-2\lambda_n T})[u_0, w_n] - \lambda_n^{2\beta}(\hat{u}, v_n) = 0,$$

где
$$\psi_n(t) = p_n e^{-\lambda_n t} - m_n e^{\lambda_n (t-2T)}, p_n = (\sigma - \gamma) \lambda_n^{2\alpha} + \lambda^{2\beta - 1}, m_n = (\sigma + \gamma) \lambda_n^{2\alpha} + \lambda^{2\beta - 1}.$$

Минимальное по норме $L_2(0,T)$ значение $[f,w_n]$ равно

$$\hat{f}_n(t)dt = rac{1}{c_n} \Biggl(2\lambda_n^{2eta} \int_0^T e^{-\lambda_n t} [\hat{u}, w_n] dt - m_n (1 - e^{-2\lambda_n T}) [u_0, w_n] \Biggr) \psi_n(t),$$

где

$$c_n = \int_0^T \psi_n^2(t) dt = \frac{p_n^2}{2\lambda_n} (1 - e^{-2\lambda_n T}) - 2p_n m_n T e^{-2\lambda_n T} + \frac{m_n^2}{2\lambda_n} e^{-2\lambda_n T} (1 - e^{-2\lambda_n T}) .$$

В условиях нашей теоремы функция $\hat{f} = \sum_{n=1}^{\infty} \hat{f}_n(t) w_n$ принадлежит $Y^{\alpha-1/2}$ и по построению $\hat{f} \in F$.

Из доказательства теоремы 3 следует, что при $\sigma = \gamma$, $\beta < \alpha + 1/2$ или $\sigma = \gamma - 1$, $\beta = \alpha + 1/2$, ряд для функции \hat{f} может расходиться в $Y^{\alpha - 1/2}$, то есть множество F может быть пустым для заданных $u_0 \in X^{\alpha}$ и $\hat{u} \in Y^{\beta}$. Таким образом, применённый в (1) способ регуляризации гарантирует только ограниченность минимизируюцих последовательностей, но не гарантирует существование хотя бы одного элемента f, такого что на u_0 реализуется минимум функционала $I(v_0) = S(G(v_0;f)), \quad v_0 \in X^{\alpha}$.

Рассмотрим один простой пример неразрешимости задачи (i)-(ii). Пусть

$$S(u) = \frac{1}{2} ||u_t||^2 + \frac{1}{2} ||Au||^2$$

что соответствует значениям $\gamma=1,\ \sigma=0,\ \alpha=1/2,\ \beta=1,\ \hat{u}=0.$ Если $u_0=w$, где $Aw=\lambda w$, то искомым решением является $f=2\lambda e^{\lambda t}w$. Действительно, для $u=G(w;f)=e^{\lambda t}w$ и всех $v\in Q_0X^{1/2}$ имеем

$$\delta S(u)v = (u_t, v_t) + (Au, Av) = (-u_t + Au, Av) = 0.$$

Пусть функция \tilde{f} такова, что $\tilde{f}+f\in F$, тогда для $\tilde{u}=Q\tilde{f}$ и всех $v\in Q_0X^{1/2}$ также верно равенство $(\tilde{u}_t,v_t)+(A\tilde{u},Av)=0$. Поэтому

$$\delta S(u)\tilde{u} = (u_t, \tilde{u}_t) + (Au, A\tilde{u}) = (U_t, \tilde{u}_t) + (AU, A\tilde{u}) = 0,$$

где $U=u+e^{2\lambda T}Q_0w=(e^{\lambda t}+e^{\lambda(2T-t)})w$. Так как $U_t(T)=0$ и $-U_{tt}+A^2U=0$, то $\delta S(u)\tilde{u}=(-U_{tt},\tilde{u})+(A^2U,\tilde{u})=0$.

Для произвольного начального состояния u_0 решением задачи (i)-(ii) будет

$$f = 2\sum_{n=1}^{\infty} \lambda_n[u_0, w_n]e^{\lambda_n t}w_n.$$

Однако ясно, что этот ряд сходится в Y не при всех $u_0 \in X^{1/2}$.

В заключение отмечу, что результаты этой работы могут быть использованы на практике для усвоения данных наблюдений в линейных моделях и в моделях с малой нелинейностью.

Автор благодарит В.И. Агошкова за ценные консультации и обсуждения.

Работа выполнена в рамках проекта "Методы решения задач вариационного усвоения данных наблюдений и управления сложными системами" по программе N3 фундаментальных исследований Отделения математических наук РАН "Вычислительные и информационные проблемы решения больших задач" и при частичной поддержке РФФИ (грант 04-01-00615).

Литература

- 1. Агошков В.И. Методы оптимального управления и сопряжённых уравнений в задачах математической физики. М.: ИВМ РАН, 2003.
- 2. Ипатова В.М. Сопряженные уравнения в задачах нечувствительного управления. В кн.: Вычислительная математика и математическое моделирование. М: ИВМ РАН, Т. 2, с. 10-16.

Задача восстановления начального условия в локально одномерной модели вертикального теплообмена

В. П. ШУТЯЕВ, Е. И. ПАРМУЗИН

Рассматривается задача вариационного усвоения данных наблюдений с целью восстановления начального условия в локальноодномерной модели вертикального теплообмена в океане, описываемой нестационарным уравнением теплопроводности с нелинейным коэффициентом диффузии. Исследована разрешимость задачи об усвоении данных, разработаны и обоснованы алгоритмы ее численного решения. Приводятся результаты численных экспериментов.

1. Введение

Настоящая работа связана с разработкой методов решения задач вариационного усвоения данных наблюдений в моделях динамики океана. Здесь рассматривается локально-одномерная модель вертикального теплообмена в океане (Венцель, Залесный, 1996), базирующаяся исследована разрешимость задачи об усвоении данных, разработаны и обоснованы алгоритмы ее численного решения. В заключение приводятся результаты численных расчетов с использованием разработанных алгоритмов.

2. Локально-одномерная модель вертикального теплообмена в океане

Математическая модель локально-одномерного вертикального теплообмена в океане может быть описана с помощью задачи:

$$\begin{cases}
T_t + a_1 w T_z + a_2 (w T)_z = (\mu T_z)_z - \lambda T + f, \\
\mu T_z = R(T - T_{surf}), \quad z = 0 \\
\mu T_z = 0, \quad z = H \\
T|_{t=0} = u,
\end{cases}$$
(1)

где $t \in (0,\bar{t}), \ z \in (0,H), \ a_1,a_2 = const \geqslant 0, a_1+a_2=1, w=w(t,z), \lambda=\lambda(t,z)\geqslant 0, R=R(t)\geqslant 0, T_{surf}=T_{surf}(t)$ – ограниченные функции, $u\in L_2(0,H), \ f\in L_2((0,\bar{t})\times(0,H)), \ \bar{t}, H<\infty.$

Здесь T — температура, w — вертикальная скорость, μ — коэффициент турбулентного теплообмена, z=0 — невозмущенная поверхность океана, H — глубина океана, z — вертикальная координата, направленная вниз.

Коэффициент теплообмена μ (Schmidt, Mysak, 1996; Zalesny et al, 1998) задается по формуле :

$$\mu = \mu_{min} exp \left\{ \frac{1}{2} \left[1 - \tanh(\gamma T_z) \right] \ln \frac{\mu_{max}}{\mu_{min}} \right\},$$

где $\gamma, \mu_{min}, \mu_{max} = const > 0$. Под "упрощенной" динамикой понимается следующее. Вертикальная скорость w может состоять из суммы трех компонент: линейной функции по z, описывающей, например, влияние длинных поверхностных волн:

$$w = W_0 \frac{H - z}{H} \sin \omega t,$$

компоненты, параметризующей эффект конвективного вовлечения:

$$w = W_c \max(0, T_z),$$

и составляющей, описывающей влияние крупномасштабной геострофической циркуляции согласно формулам:

$$\begin{cases} u = -\frac{1}{f\rho_0} p_y, \\ v = \frac{1}{f\rho_0} p_x, \\ p_z = g\rho, \\ u_x + v_y + w_z = 0, \\ w = 0 \quad \text{при} \quad z = 0, z = H. \end{cases}$$

Составляющая w может быть выражена из последних формул в явном виде (Венцель, Залесный, 1996).

В дальнейшем, считая функцию w заданной, мы рассмотрим задачу о восстановлении начального условия u.

3. Операторная формулировка задачи

Пусть

$$Y = L_2(0, \bar{t}; W_2^1(0, H)), \quad Y^* = L_2(0, \bar{t}; W_2^{-1}(0, H)),$$

$$W = \{ \varphi \in Y : \ \varphi_t \in Y^* \}.$$

Рассмотрим интегральное тождество

$$(T_t, v) + a_1(wT_z, v) - a_2(wT, v_z) + a_2 \int_0^{\bar{t}} (wTv) \Big|_{z=0}^{z=H} dt =$$

$$= -(\mu T_z, v_z) - \int_0^{\bar{t}} R(T|_{z=0} - T_{surf}) v|_{z=0} dt - (\lambda T, v) + (f, v),$$

 $\forall v \in Y$, где $(\cdot, \cdot) = (\cdot, \cdot)_{L_2(0, \bar{t}; (0,H))}$.

Определим операторы $A, F: Y \to Y^*$ по формулам:

$$(AT, v) = a_1(wT_z, v) - a_2(wT, v_z) + a_2 \int_0^{\bar{t}} (wTv) \Big|_{z=0}^{z=H} dt$$

$$+ \int_{0}^{t} RT|_{z=0}v|_{z=0}dt + (\lambda T, v), \quad (F(T), v) = (\mu T_{z}, v_{z}), \quad T, v \in Y.$$

Тогда задача (1) из раздела 2 запишется в виде

$$\begin{cases}
T_t + AT + F(T) = \bar{f}, & t \in (0, \bar{t}) \\
T|_{t=0} = u,
\end{cases}$$
(1)

где $\bar{f} \in Y^*$ определяется по формуле:

$$(\bar{f}, v) = (f, v) + \int_{0}^{\bar{t}} RT_{surf}v|_{z=0}dt, \ v \in Y.$$

4. Свойства операторов и разрешимость

Лемма 4.1. Оператор $A:Y\to Y^*$ – линейный ограниченный с областью определения D(A)=Y. При $a_1=a_2$ оператор A симметричен и справедливо представление

$$(AT,T) = a_2 \int_0^{\bar{t}} (wT^2) \Big|_{z=0}^{z=H} dt + \int_0^{\bar{t}} RT^2|_{z=0} dt + (\lambda T, T), \quad T \in Y.$$

Доказательство. В силу вложения $W_2^1(0,H)$ в C(0,H) имеем

$$|T/_{z=0}| \le c_1 ||T||_{W_2^1(0,H)}, \quad c_1 = const > 0.$$

Тогда

$$\begin{split} |(AT,v)| &\leqslant a_1 \|w\|_{L_{\infty}} \|T_z\| \|v\| + a_2 \|w\|_{L_{\infty}} \|T\| \|v_z\| + \left(2a_2 \|w\|_{L_{\infty}} + \|R\|_{L_{\infty}}\right) c_1^2 \|T\|_Y \|v\|_Y + \|\lambda\|_{L_{\infty}} \|T\| \|v\| &\leqslant [(a_1+a_2) \|w\|_{L_{\infty}} + (2a_2 \|w\|_{L_{\infty}} + \|R\|_{L_{\infty}} + \|\lambda\|_{L_{\infty}}) c_1^2 \|T\|_Y \|v\|_Y, \end{split}$$

где $\|\cdot\|=(\cdot,\cdot)^{1/2}$. Значит, $A:Y\to Y^*$ – линейный ограниченный оператор с областью определения D(A)=Y.

Для (AT, T) имеем:

$$(AT,T) = a_1(wT_z,T) - a_2(wT,T_z) + a_2 \int_0^{\bar{t}} (wT^2) \Big|_{z=0}^{z=H} dt + \int_0^{\bar{t}} RT^2 \Big|_{z=0} dt + \lambda(T,T), \quad T \in Y.$$

Отсюда при $a_1 = a_2$ получаем

$$(AT,T) = a_2 \int_0^t (wT^2) \Big|_{z=0}^{z=H} dt + \int_0^t RT^2 \Big|_{z=0} dt + \lambda(T,T), \quad T \in Y$$

Нетрудно видеть, что при $a_1 = a_2$

$$(AT, v) = (T, Av), t, v \in Y,$$

т.е. оператор A – симметричен. Лемма доказана.

Следствие. Если $w|_{z=H}\geqslant 0,\ R\geqslant a_2w|_{z=0},\ \lambda\neq 0,$ то оператор A положителен:

$$(AT,T) > 0, \ \forall T \in Y, T \neq 0.$$

Лемма 4.2. One pamop $F:Y \to Y^*$ ограничен.

Доказательство. Для (F(T), v) имеем:

$$|(F(T), v)| \le \mu_{max} ||T||_Y ||v||_Y, T, v \in Y,$$

откуда следует утверждение леммы.

Лемма 4.3. В любой точке $T \in Y$ оператор F имеет производную Γ ато F'(T), определяемую по формуле

$$(F'(T)u, v) = (\mu u_z, v_z) + (\mu' T_z u_z, v_z), \quad T, u, v \in Y,$$

 $e \partial e$

$$\mu' = -\mu q_1 \gamma (1 - \tanh^2(\gamma T_z)), \quad q_1 = \frac{1}{2} \ln \frac{\mu_{max}}{\mu_{min}}.$$

Onepamop F'(T) ограничен из Y в Y^* и

$$||F'(T)u||_{Y^*} \leqslant k||u||_Y$$

 $e \partial e$

$$k = \mu_{max}(1 + q_1q_2), \quad q_2 = 2(e^{2x_0} - e^{-2x_0})^{-1},$$

а x_0 – положительный корень уравнения $\tanh x = 1/(2x)$.

Доказательство. Существование производной F'(T) доказывается на основе определения производной Гато, следуя [7]. Рассмотрим функцию

$$g(x) = x(1 - \tanh^2 x), \quad x \in (-\infty, \infty).$$

Поскольку

$$g'(x) = 4(e^x + e^{-x})^{-3}(e^x + e^{-x} - 2x(e^x - e^{-x})),$$

то g(x) ограничена и достигает своего максимального значения при $x=x_0$, где x_0 – положительное решение уравнения $\tanh x=1/(2x)$. При этом $1/2\leqslant x_0\leqslant 1$ и

$$g(x_0) = x_0(1 - \tanh^2 x_0) = \frac{4x_0}{(e^{x_0} + e^{-x_0})^2} = \frac{2}{\tanh x_0(e^{x_0} + e^{-x_0})^2}$$

$$= \frac{2}{e^{2x_0} + e^{-2x_0}} = q_2 \leqslant \frac{2}{e + e^{-1}} < 1.$$

Тогда

$$|\mu'T_z| = |\mu q_1(1 - \tanh^2 \gamma T_z)\gamma T_z| \leqslant \mu_{\max} q_1 q_2$$

И

$$|(F'(t)u, v)| \leq \mu_{\max} ||u_z|| ||v_z|| + \mu_{\max} q_1 q_2 ||u_z|| ||v_z|| \leq \mu_{\max} (1 + q_1 q_2) ||u||_Y ||v||_Y.$$

Отсюда следует второе утверждение леммы.

Лемма 4.4. Пусть $w|_{z=H} \geqslant 0$, $R \geqslant a_2 w|_{z=0}$. При $a_1 = a_2$, $\lambda \geqslant \lambda_0 > 0$, $q_1 q_2 < 1$ оператор A + F сильно монотонный:

$$(Au + F(u) - Av - F(v), u - v) \ge m||u - v||_Y^2, \ \forall u, v \in Y,$$

 $e \partial e \quad m = \min(\lambda_0, \mu_{min}(1 - q_1 q_2)).$

Доказательство. Для F(u) имеет место представление [7]:

$$F(u) - F(v) = \int_{0}^{1} F'(v + \tau(u - v)) d\tau(u - v).$$

Отсюда и из леммы 4.3 получаем:

$$(F(u) - F(v), u - v) = \left(\int_{0}^{1} F'(v + \tau(u - v)) d\tau(u - v), u - v \right) =$$

$$(\mu(u-v)_z,(u-v)_z)-\Big(\mu q_1\int\limits_0^1\Big(1-\tanh^2\gamma(v+ au(u-v))\Big)\Big) imes$$

$$\gamma(v+\tau(u-v))d\tau(u-v)_z, (u-v)_z) \geqslant \mu_{\min}(1-q_1q_2)\|(u-v)_z\|^2.$$

Для оператора A имеем при $a_1 = a_2$:

$$(Au - Av, u - v) = (A(u - v), u - v) = \int_{0}^{\bar{t}} R(u - v)^{2} \Big|_{z=0} dt$$
$$+ (\lambda(u - v), u - v) \ge \lambda_{0} ||u - v||^{2}.$$

Тогда

$$(A + F(u) - Av - F(v), u - v) \ge m \|u - v\|_{Y}^{2}$$

где $m = \min(\lambda_0, \mu_{\min}(1 - q_1q_2))$. Лемма доказана.

Замечание. Утверждение леммы 4.4 о сильной монотонности оператора A+F остается справедливым, если условие $\lambda \geqslant \lambda_0>0$ заменить на условие $R\neq 0,\,w|_{z=0}=0.$

Из лемм 4.1-4.4 с использованием результатов [4] вытекает следующая

Теорема 4.1.[12] Пусть $w|_{z=H} \ge 0$, $R \ge a_2 w|_{z=0}$, $a_1 = a_2$, $\lambda = \lambda_0 > 0$, $q_1 q_2 < 1$, $\bar{f} \in L_2((0,\bar{t}) \times (0,H))$, $u \in L_2(0,H)$. Тогда существует единственное решение $T \in W$ задачи (1) из раздела 3.

5. Задача об усвоении данных

Рассмотрим задачу (1) из раздела 2 с неизвестной функцией $u \in L_2(0,H)$ (т.е. неизвестным управлением u) в начальном условии. Задача об усвоении данных может быть сформулирована следующим образом: найти функцию $u \in L_2(0,H)$ такую, что:

$$\begin{cases}
T_t + AT + F(T) = \bar{f}, & t \in (0, \bar{t}) \\
T|_{t=0} = u, \\
S(u) = \inf_{v} S(v),
\end{cases} \tag{1}$$

где

$$S(u) = \frac{\alpha}{2} \|u - \hat{T}_0\|_{L_2(0,H)}^2 + \frac{1}{2} \int_0^{\bar{t}} \|T - \hat{T}\|_{L_2(0,H)}^2 dt, \tag{2}$$

а \hat{T}, \hat{T}_0 – заданные функции, $\alpha = const \geqslant 0$.

Таким образом, задача об усвоении данных состоит в отыскании функций u и T таких, что они удовлетворяют задаче (1), причем на множестве решений уравнения (1) функционал (2) принимает наименьшее значение. Функции \widehat{T} , \widehat{T}_0 определяются, как правило, априорными данными наблюдений [1, 2, 8, 10, 14].

Теорема 5.1.В условиях теоремы 4.1 при $\alpha > 0$ существует $u \in L_2(0,H)$ – решение задачи (1).

Доказательство. Пусть u_n – минимизирующая последовательность для функционала S(u), т.е. $S(u_h) \to \inf_{v \in L_2(0,H)} S(v), \ n \to \infty$.

Поскольку $S(u) \geqslant \frac{\alpha}{2} \|u - \widehat{T}_0\|_{L_2(0,H)}^2$, $\forall u \in L_2(0,H)$, то при $\alpha > 0$ по-

следовательность u_n ограничена: $||u_n||_{L_2(0,H)} \leqslant const.$ Следовательно, у нее существует слабо сходящаяся подпоследовательность (мы ее также обозначим через u_n). Поскольку гильбертово пространство $L_2(0,H)$ слабо замкнуто, то существует элемент u из $L_2(0,H)$ такой, что $u_n \to u$ (слабо в $L_2(0,H)$), т.е. $(u_n,\theta)_{L_2(0,H)} \to (u,\theta)_{L_2(0,H)}$, $\forall \theta \in L_2(0,H)$.

Пусть T_n и T – решения исходной задачи (1) при $T_n\big|_{t=0}=u_n$ и $T\big|_{t=0}=u$ соответственно. Для разности $T-T_n$ справедливо

$$((T - T_n)_t, v) + (A(T - T_n), v) + (F(T) - F(T_n), v) = 0, \ v \in Y.$$

Пусть $v \in W_T = \{ \varphi \in W : \varphi \big|_{t=\bar{t}} = 0 \}$. Тогда интегрированием по частям получаем для $v \in W_T$

$$-(T - T_n, v_t) + ((T - T_n), A^*v) + (F(T) - F(T_n), v) = (u - u_n, v).$$

Последнюю задачу запишем в виде

$$(T - T_n, -v_t + A^*v + (F'(\xi))^*v) = (u - u_n, v), v \in W_T,$$

где $(F'(\xi))^*: Y \to Y^*$ – оператор, сопряженный к $F'(\xi)$. Рассмотрим, $p \in Y^*$ и введем вспомогательную сопряженную задачу:

$$\begin{cases} -v_t + A^*v + (F'(\xi))^*v = p, \\ v|_{t=\bar{t}} = 0. \end{cases}$$

Эта задача при $p \in Y^*$ имеет единственное решение $v \in W_T$ [6]. Тогда $\forall p \in Y^*$ имеет место равенство

$$(T - T_n, p) = (u - u_n, v),$$

где v — решение сопряженной задачи. Поскольку $(u-u_n,v)\to 0$, $n\to\infty$, то $(T-T_n,p)\to 0$, $n\to\infty$. Это означает, что $T_n\to T$ слабо в $L_2((0,\bar t)\times(0,H))$. Функционал $J(\cdot)=\|\cdot\|^2$ слабо полунепрерывен снизу в слабой топологии [6], т.е.

$$\lim\inf S(u_n)\geqslant S(u),$$

и, следовательно,

$$\inf_{v \in L_2(0,H)} S(v) \geqslant S(u).$$

Отсюда

$$\inf_{v \in L_2(0,H)} S(v) = S(u),$$

т.е. и доставляет наименьшее значение функционалу S(v). Теоре-

Если $u \in L_2(0, H)$ – решение задачи (1), то оно удовлетворяет системе оптимальности (необходимое условие оптимальности):

$$\begin{cases}
T_t + AT + F(T) = \bar{f}, & t \in (0, \bar{t}) \\
T|_{t=0} = u,
\end{cases}$$
(3)

$$\begin{cases}
T_t + AT + F(T) = \bar{f}, & t \in (0, \bar{t}) \\
T|_{t=0} = u, \\
-T_t^* + A^*T^* + (F'(T))^*T^* = (\hat{T} - T), & t \in (0, \bar{t}) \\
T^*|_{t=\bar{t}} = 0, \\
\alpha(u - \hat{T}_0) - T^*|_{t=0} = 0
\end{cases} \tag{4}$$

$$\alpha(u - \hat{T}_0) - T^*|_{t=0} = 0 \tag{5}$$

с неизвестными T, T^*, u .

Для гладких T, T^* эта система принимает вид:

$$\begin{cases}
T_t + a_1 w T_z + a_2 (w T)_z = (\mu T_z)_z - \lambda T + f, \\
\mu T_z = R (T - T_{surf}), \quad z = 0 \\
\mu T_z = 0, \quad z = H \\
T|_{t=0} = u,
\end{cases}$$
(6)

$$\begin{cases}
T_{t} + a_{1}wT_{z} + a_{2}(wT)_{z} = (\mu T_{z})_{z} - \lambda T + f, \\
\mu T_{z} = R(T - T_{surf}), \quad z = 0 \\
\mu T_{z} = 0, \quad z = H \\
T|_{t=0} = u,
\end{cases}$$

$$\begin{cases}
-T_{t}^{*} - a_{1}(wT^{*})_{z} - a_{2}wT_{z}^{*} = \\
((\mu + \mu'T_{z})T_{z}^{*})_{z} - \lambda T^{*} + (\hat{T} - T), \\
wT^{*} + (\mu + \mu'T_{z})T_{z}^{*} = RT^{*}, \quad z = 0 \\
wT^{*} + \mu T_{z}^{*} = 0, \quad z = H \\
T|_{t=T} = 0,
\end{cases}$$
(6)

$$\alpha(u - \hat{T}_0) - T^*|_{t=0} = 0. (8)$$

6. Метод последовательных приближений

Рассмотрим итерационный процесс для решения системы (3)-(5) из раздела 5:

$$\left\{ \begin{array}{l} T_t^{(n+1)} + A T^{(n+1)} + B T^{(n+1)} + F(T^{(n)}) = B T^{(n)} + \bar{f}, \\ T^{(n+1)}\big|_{t=0} = u^{(n+1)}, \end{array} \right. \tag{1}$$

$$-T_{t}^{*(n+1)} + A^{*}T^{*(n+1)} + B^{*}T^{*(n+1)} + (F'(T^{(n)}))^{*}T^{*(n)} =$$

$$\begin{cases} B^{*}T^{*(n)} + (\hat{T} - T^{(n+1)}), \\ T^{*(n+1)}|_{t=\bar{t}} = 0, \end{cases}$$
(2)

$$\alpha(u^{(n+1)} - \hat{T}_0) - T^{*(n+1)}|_{t=0} = 0, \tag{3}$$

где

$$t \in (0, \bar{t}), \ B: Y \to Y^*, \ (BT, v) = \frac{1}{\beta}(T_z, v_z),$$

$$\beta = 2[\mu_{min}(1 - q_1q_2) + \mu_{max}(1 + q_1q_2)]^{-1}.$$

Из результатов работы [13] вытекает

Теорема 6.1. В условиях теоремы 4.1 метод (1)–(3) $cxo\partial umcs$.

7. Итерационные методы для решения линейной за-ДАЧИ

На каждом шаге метода последовательных приближений (1)-(3) из раздела 6 мы имеем линейную задачу об усвоении данных вида

$$\begin{cases}
T_t + AT + BT = \tilde{f}, & t \in (0, \bar{t}) \\
T|_{t=0} = u,
\end{cases}$$
(1)

$$\begin{cases}
T_t + AT + BT = \tilde{f}, & t \in (0, \bar{t}) \\
T|_{t=0} = u, \\
-T_t^* + A^*T^* + B^*T^* = (\tilde{T} - T), & t \in (0, \bar{t}) \\
T^*|_{t=\bar{t}} = 0,
\end{cases} \tag{2}$$

$$\alpha(u - \hat{T}_0) - T^*|_{t=0} = 0, (3)$$

где \tilde{f}, \tilde{T} – заданные функции. Для решения этой системы можно воспользоваться классом итерационных алгоритмов, разработанных в [13], вида

$$\begin{cases}
T_t^k + AT^k + BT^k = \tilde{f}, & t \in (0, \bar{t}) \\
T^k|_{t=0} = u^k,
\end{cases}$$
(4)

$$\begin{cases}
T_t^k + AT^k + BT^k = \tilde{f}, & t \in (0, \bar{t}) \\
T^k\big|_{t=0} = u^k, \\
-T_t^{*k} + A^*T^{*k} + B^*T^{*k} = (\tilde{T} - T^k), & t \in (0, \bar{t}) \\
T^{*k}\big|_{t=\bar{t}} = 0,
\end{cases} \tag{5}$$

$$u^{k+1} = u^k - \alpha_{k+1} C_k (\alpha(u^k - \hat{T}_0) - T^{*k}/_{t=0}) + \beta_{k+1} D_k (u^k - u^{k-1}),$$
(6)

где $C_k, D_k: L_2(0,H) o L_2(0,H)$ – некоторые операторы, α_k, β_k – итерационные параметры. Выбор итерационных параметров обсуждается в [13].

Метод сопряженных градиентов для линейной задачи имеет вид:

$$\begin{cases}
T_t^k + AT^k + BT^k = \tilde{f}, & t \in (0, \bar{t}) \\
T^k|_{t=0} = u^k,
\end{cases}$$
(7)

$$\begin{cases}
T_t^k + AT^k + BT^k = \tilde{f}, & t \in (0, \bar{t}) \\
T^k|_{t=0} = u^k, \\
-T_t^{*k} + A^*T^{*k} + B^*T^{*k} = (\tilde{T} - T^k), & t \in (0, \bar{t}) \\
T^{*k}|_{t=\bar{t}} = 0,
\end{cases} \tag{8}$$

$$u^{k+1} = u^k - \alpha_{k+1}(\alpha(u^k - \hat{T}_0) - T^{*k}/_{t=0}) + \beta_{k+1}(u^k - u^{k-1}), \quad (9)$$

где $\alpha_{k+1} = 1/p_{k+1}, \ \beta_{k+1} = e_k/p_{k+1}$ и

$$e_k = \begin{cases} 0, & k = 0 \\ p_k \|\xi^k\|_{L_2}^2 / \|\xi^{k-1}\|_{L_2}^2, & k > 0, \end{cases}$$
$$p_{k+1} = \alpha + \|\eta^k\|^2 / \|\xi^k\|_{L_2}^2 - e_k, & k = 0, 1, \dots$$

Здесь
$$\xi^k=\alpha u^k-T^{*k}\big|_{t=0},$$
 а η^k – решение задачи
$$\frac{d\eta^k}{dt}+(A+B)\eta^k=0,\;t\in(0,\bar t);\quad\eta^k(0)=\xi^k.$$

Из общих результатов работы [13] вытекает **Теорема 7.1.** В условиях теоремы 4.1 при $\alpha > 0$ итерационный процесс (7)–(9) сходится.

В отличие от метода простой итерации и чебышевских методов вида (4)–(6), метод сопряженных градиентов не требует знания границ спектра оператора управления [13].

8. Численное решение задачи об усвоении данных

В настоящем параграфе мы приведем некоторые результаты численных экспериментов на основе сформулированного выше вычислительного алгоритма.

Для решения системы (6)–(8) из раздела 5 воспользуемся итерационными методами, сформулироваными в разделах 6 и 7. Один из таких методов (после аппроксимации задачи по временной переменной на основе явной схемы) формулируется следующим образом:

$$\frac{T_i^{k+1(j)} - T_i^{k(j)}}{\tau} + a_1 w_i \frac{T_{i+1}^{k(j)} - T_{i-1}^{k(j)}}{2h} + a_2 \frac{w_{i+1} T_{i+1}^{k(j)} - w_{i-1} T_{i-1}^{k(j)}}{2h} =$$

$$\mu_{i+1/2} \frac{T_{i+1}^{k(j)} - T_i^{k(j)}}{h^2} - \mu_{i-1/2} \frac{T_i^{k(j)} - T_{i-1}^{k(j)}}{h^2} - \lambda T_i^{k(j)}, \qquad (1)$$

$$\mu_0 \frac{T_1^{k(j)} - T_{i-1}^{k(j)}}{2h} = R(T_0^{k(j)} - T_{surf}), \qquad T_{N+1}^{k(j)} = T_{N-1}^{k(j)},$$

$$T_{N+1}^{0(j)} = u^j$$

$$-\frac{T_{i}^{*k+1(j)} - T_{i}^{*k(j)}}{\tau} - a_{1} \frac{w_{i+1} T_{i+1}^{*k(j)} - w_{i-1} T_{i-1}^{*k(j)}}{2h} - a_{2} w_{i} \frac{T_{i+1}^{*k(j)} - T_{i-1}^{*k(j)}}{2h} = \tilde{\mu}_{i+1/2} \frac{T_{i+1}^{*k(j)} - T_{i}^{*k(j)}}{h^{2}} - \tilde{\mu}_{i-1/2} \frac{T_{i}^{*k(j)} - T_{i-1}^{*k(j)}}{h^{2}} - \lambda T_{i}^{*k(j)} + (\hat{T}_{i} - T_{i}^{k(j)}),$$

$$(2)$$

$$w_{0} T_{0}^{*k(j)} + \left(\mu_{0} + \mu_{0}' \frac{T_{1}^{k(j)} - T_{-1}^{k(j)}}{2h}\right) \frac{T_{1}^{*k(j)} - T_{-1}^{*k(j)}}{2h} = R T_{0}^{*k(j)},$$

$$w_{N} T_{N}^{*k(j)} + \mu_{N} \frac{T_{N+1}^{*k(j)} - T_{N-1}^{*k(j)}}{2h} = 0, \qquad T^{*M(j)} = 0,$$

$$u^{j+1} = u^{j} + \alpha_{j+1} (T^{*0(j)} - \alpha(u^{j} - \hat{T}_{0})) + \beta_{j+1} (u^{j} - u^{j-1}),$$

$$(3)$$

где

$$\begin{split} \mu_{i+1/2} &= \mu_{\min} \exp \left\{ \frac{1}{2} \Big[1 - \tanh \left(\gamma_{i+1/2} \frac{T_{i+1}^{k(j)} - T_i^{k(j)}}{h} \right) \ln \frac{\mu_{\max}}{\mu_{\min}} \right\}, \\ \mu_{i-1/2} &= \mu_{\min} \exp \left\{ \frac{1}{2} \Big[1 - \tanh \left(\gamma_{i-1/2} \frac{T_i^{k(j)} - T_{i-1}^{k(j)}}{h} \right) \ln \frac{\mu_{\max}}{\mu_{\min}} \right\}, \\ \tilde{\mu}_{i+1/2} &= \mu_{i+1/2} + \mu'_{i+1/2} \frac{T_{i+1}^{k(j)} - T_i^{k(j)}}{h}, \\ \tilde{\mu}_{i-1/2} &= \mu_{i-1/2} + \mu'_{i-1/2} \frac{T_i^{k(j)} - T_{i-1}^{k(j)}}{h}, \end{split}$$

 $\alpha_{j+1},\ \beta_{j+1}$ — итерационные параметры, $T^{k(j)},\ T^{*k(j)},\ u^j$ — итерационные последовательности, τ — шаг сетки по времени, $k=\overline{0,M-1},\ \tau=\mathcal{T}/M,\ h$ — шаг сетки по пространству $i=\overline{0,N},\ h=1/N.$ Аппроксимация рассматриваемой схемы — $O(\tau+h^2)$ на гладких решениях.

Приведем результаты численных экспериментов. В рассматриваемых примерах шаги по пространству выбирались, одинаковыми и равными h=0.02, при этом $\tau=1.33\times 10^{-5}$. Критерием остановки итерационного процесса являлось условие

$$\frac{\|u^{j+1} - u^j\|}{\|u^j\|} < \varepsilon,$$

где u^j – решение системы (1)–(3) после j итераций, $\varepsilon = h^2 = 4 \cdot 10^{-4}$, а $\|\cdot\|$ – разностный аналог нормы в $L_2(0, H)$.

В численных экспериментах исследовалась сходимость итерационных методов при следующих входных параметрах и функциях:

$$a_i = 0.5, \ i = 1, 2, \ R = 1, \ w = 1, \ \lambda = 1, \ \gamma = 1,$$

 $\mu_{min} = 1, \ \mu_{max} = 2, \ H = 1, \ \bar{t} = 1, \ \beta_{i+1} = 0.$ (4)

Решение T прямой задачи (1) из раздела 2 было найдено численно при $u=x^2(1-x)^2$, и оно рассматривалось в качестве $\widehat{T}=\widehat{T}(x,t)$, где $t\in(0,1),x\in(0,1)$. Затем предполагалось, что u неизвестно и для его отыскания из (6)-(8) раздел 5 мы воспользовались алгоритмом (1)-(3) с начальным приближением $u^0=0$, полагая согласно [13] $\alpha_{j+1}=\tau_{onm}=\frac{1}{\alpha+\overline{t}/2}.$

Таблица 1

Параметр α	1	0.1	0.01	0.001	0.0001
количество итераций n	6	22	109	205	226
Погрешность \mathcal{R}	0.864	0.555	0.298	0.069	0.023

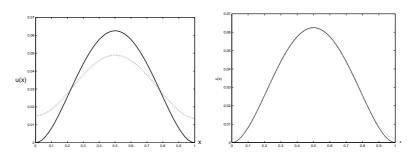
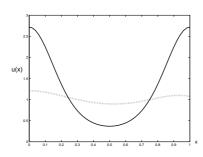


Рис. 1.

Введение регуляризатора с параметром α , как известно [2, 9], приводит к решению u задачи (6)–(8) из раздела 5, которое отличается от решения \bar{u} , полученного из (6)–(8) раздела 5 при $\alpha=0$, на величину $O(\alpha)$. Обозначим через $\mathcal R$ относительную погрешность $\mathcal R=\|u^j-\bar{u}\|/\|\bar{u}\|$, где u^j – приближенное решение, найденное с помощью рассматриваемого алгоритма. В таблице 1 приведены зависимость числа итераций, необходимых для отыскания приближенного решения с требуемой точностью ε и значения погрешности $\mathcal R$ при различных α . Из таблицы видно, что с уменьшением параметра регуляризации α величина погрешности $\mathcal R$ уменьшается, однако при этом растет число итераций в алгоритме. На рисунке 1 приведены результаты двух численных эспериментов (указанных в таблице 1 с параметрами $\alpha=10^{-3}$ и $\alpha=10^{-4}$ соответственно.

В следующем эксперименте в качестве начального условия для построения $\widehat{T}(x,t)$ выбиралась функция $u(x)=e^{\cos 2\pi x}$. При этом для удовлетворения граничному условию функция T_{surf} выбиралась равной $T_{surf}=e$. Значения остальных параметров и функций приведены в (4). Результаты численных экспериментов для вы-



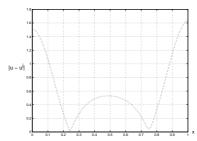
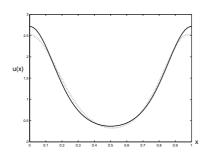


Рис. 2.



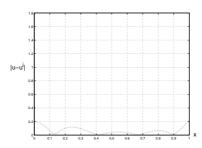


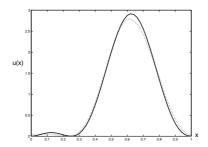
Рис. 3.

бранной функции начального условия приведены на рисунках 2 и 3.

Одной из важнейших характеристик задач данного типа является значение минимизируемого функционала (2) из раздела 5. В таблице 2 представлены значения функционала S(u) при различных значениях α после окончания процесса инициализации. Результаты численных экспериментов для других начальных условия приведены на рисунках 4 и 5. Значение параметров и функций

Таблица 2

Параметр α	1	0.25	0.1	0.01	0.0001
S(u)	0.177	0.115	0.069	0.017	0.009



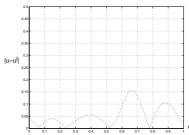


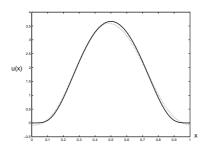
Рис. 4.

для этих экспериментов приведены в (4).

Так, в эксперименте, представленном на рисунке 4, в качестве начального условия для построения $\widehat{T}(x,t)$ выбиралась функция $u(x)=(1-\sin 2\pi x)(1-\cos 2\pi x)$. В следующем эсперименте в качестве начального условия выбиралась функция $u(x)=100e^{-1/(x(1-x))}$. Точное значение начального условия и восстановленная функция при $\alpha=10^{-4}$ показаны на рисунке 5. Данные графики также подтверждают результаты теоретических исследований об уменьшении погрешности при уменьшении параметра регуляризации.

Одной из особенностей данной задачи является наличие нелинейной части в прямом и сопряженном уравнении. В следующем эксперименте исследуется зависимость сходимости предложенного итерационного метода от величины нелинейности. Для этого проводились численные эксперименты с различным соотношением $\kappa = \mu_{\rm max}/\mu_{\rm min}$. Решение задачи (1) из раздела 5 с начальным условием $u(x) = x^2(1-x^2)$ рассматривалось в качестве $\widehat{T}(x,t)$. Остальные параметры (кроме значений $\mu_{\rm max}$ и $\mu_{\rm min}$) задавались согласно (4). Результаты численных экспериментов (количество итераций и ошибка \mathcal{R}) для некоторых значений κ при $\alpha = 10^{-3}$ приведены в таблице 3. На рисунке 6 приведены графики зависимости числа итераций, необходимых для отыскания приближенного решения

с требуемой точностью ε , от параметра регуляризации α для различных значений μ_{\max} и μ_{min} . Из таблицы 3 и рисунка 6 видно, что при $0.1 < \alpha < 1$ влияние нелинейности минимально. Однако при уменьшении параметра регуляризации влияние нелинейной



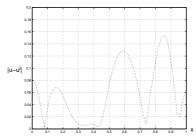


Рис. 5.

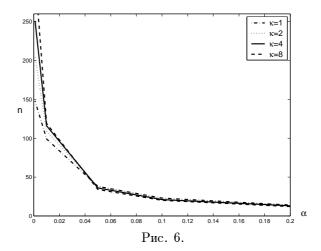
Таблица 3

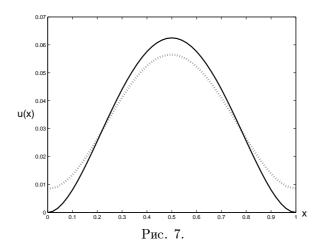
	$\kappa = 1$	$\kappa = 2$	$\kappa = 4$	$\kappa = 8$
количество итераций n	180	205	251	312
Погрешность \mathcal{R}	0.045	0.069	0.093	0.128

составляющей замедляет сходимость итерационого процесса.

Для иллюстрации сходимости метода на рисунке 7 приведены результаты численного эксперимента для функции $u(x)=x^2(1-x)^2$ при $\alpha=10^{-3}$ и $\kappa=8$. Остальные параметры и функции заданы в (4). Сформулируем основные выводы, к которым приводят результаты численных экспериментов. Метод последовательных приближений (1)–(3) раздел 6, использованный для решения нелинейной системы (6)–(8) раздел 5, сходится достаточно быстро и может рассматриваться как эффективное средство для решения задач об усвоении данных. Сходимость всего итерационного процесса определяется величиной параметра регуляризации α . Если $0.05 \leqslant \alpha \leqslant 1$, то количество итераций невелико, — в этом преимущество введения регуляризатора. Недостатком же является тот факт, что при введении регуляризатора в решение задачи вносится погрешность. С уменьшением α погрешность уменьшается, но при этом общее число итераций метода резко возрастает.

Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований (проект 03-01-00779).





Литература

- [1] Агошков В.И. Методы оптимального управления и сопряженных уравнений в задачах математической физики. М.: ИВМ РАН, 2003.
- [2] Agoshkov V.I., Marchuk G.I. On solvability and numerical solution of data assimilation problems. – Russ. J. Numer. Anal. Math. Modelling, 1993, vol. 8, No. 1, pp. 1-16.
- [3] Le Dimet F.-X., Shutyaev V.P. On data assimilation for quasilinear parabolic problems. Russ. J. Numer. Anal. Math. Modelling (2001), 16 (3), 247–259.
- [4] Гаевский Х., Грегер К., Захариас К. Нелинейные операторные уравнения и операторные дифференциальные уравнения.
 М.: Мир, 1978.
- [5] Kazantsev Chr., Moshonkin S.N., Zalesny V.B. Mathematical modelling of the global ocean dynamics: solvability, numerical algorithm, calculations. In: Variability and Predictability of Atmospheric and Oceanic Circulations. – Moscow: MGU, 1999.
- [6] Lions J.-L. Optimal control by systems described by partial differential equations.— Moscow: "Mir", 1972.
- [7] Marchuk G.I., Agoshkov V.I., Shutyaev V.P. Adjoint Equations and Perturbation Algorithms in Nonlinear Problems. New York: CRC Press, 1996.
- [8] Marchuk G.I., Zalesny V.B. A numerical technique for geophysical data assimilation problem using Pontryagin's principle and splitting-up method. Russian J. Num. Anal. Math. Mod., 1993, v. 8, No. 4, p. 311-326.
- [9] Marchuk G.I., Shutyaev V.P. Iteration methods for solving a data assimilation problem.-Russ. J. Numer. Anal. Math. Modelling, 1994, Vol. 9, No.3, pp.265-279.
- [10] Marchuk G., Shutyaev V., Zalesny V. Approaches to the solution of data assimilation problems. In: Optimal Control and Partial Differential Equations. Amsterdam: IOS Press (2001), 489-497.

- [11] Schmidt G.A., Mysak L.A. The stability of a zonally averaged thermohaline circulation model. *Tellus*, 1996, v.48 A, pp.158–178.
- [12] Шутяев В.П. Задача об усвоении данных для локальноодномерной модели вертикального теплообмена. В сб.: "Отчет ИВМ РАН о научной и научно-организационной деятельности в 2001 году. Тезисы научных докладов на отчетной сессии". – М.: ИВМ РАН, 2001, с.69–70.
- [13] Шутяев В.П. Операторы управления и итерационные алгоритмы в задачах вариационного усвоения данных. М.: Наука, 2001.
- [14] Венцель М., Залесный В.Б. Усвоение данных в одномерной модели конвекции-диффузии тепла в океане // Изв.АН. Физика атмосферы и океана.1996.Т.32, N5. С.613-629.

Часть II

Матричные методы в интегральных и дифференциальных уравнениях

О параллельном многосеточном решении консервативных неструктурированных конечноэлементных систем

Ю. В. Василевский, В. Н. Чугунов

Рассматривается многосеточный метод для неконформного конечно-элементного пространства, степени свободы которого ассоциированы с ребрами треугольников (гранями тетраэдров). Описываются последовательный и параллельный варианты алгоритма и обсуждаются тонкости реализации на многопроцессорных вычислительных системах. Приведены результаты численных экспериментов на разных кластерах и дана их оценка.

1. Введение

Одним из подходов численного решения дифференциальных уравнений в частных производных является метод смешанных конечных элементов [2]. Он обладает рядом ценных преимуществ: локально консервативен, явно вычисляет потоки. Легко параллелизуемый процесс алгебраической конденсации [3] смешанной конечноэлементной системы приводит ее к системе более простого вида, эквивалентной некоторой неконформной конечноэлементной системе. В случае самосопряженности исходного оператора эффективным способом решения полученной системы является метод сопряженных градиентов с многосеточным переобуславливателем, который характеризуются быстрой скоростью сходимости и невысокой стоимостью итерации. Применение многосеточных переобуславливателей на адаптивных локально сгущающихся сетках значительно сокращает арифметическую стоимость решения этих задач. Параллельный многосеточный метод дает возможность решать большие дискретные системы на вычислительных параллельных машинах быстрее последовательного.

В данной статье мы рассмотрим многосеточный метод для неконформного конечноэлементного пространства, степени свободы которого ассоциированы с ребрами треугольников (гранями тетраэдров). Решение этой задачи — главная составляющая эффективного решения смешанного метода конечных элементов, поскольку алгебраическая конденсация может быть выполнена поэлементно и, следовательно, легко параллелизуема. Параллелизация реберно-ориентированного многосеточного метода на локально сгущающихся сетках — сложная техническая задача [5, 6]. Она включает в себя равномерное распределение данных по процессорам даже в случае неравномерной сетки, учет топологии сетки, обусловленный тем фактом, что степени свободы соответствуют серединам ребер треугольников (граней тетраэдров). Последнее обстоятельство ведет как к росту заполненности галеркинских проекций матрицы на грубые сетки, так и является основной причиной падения выигрыша при увеличении числа процессоров в силу увеличения количества глобальных обменов между всеми процессорами, как будет показано в статье. В этом главная сложность предлагаемого нами метода.

Тестирование предлагаемого алгоритма дает различные результаты на разных кластерах. Это обусловлено особенностями параллельных машин и описываемого параллельного алгоритма. В работе мы укажем на основные причины этих различий.

2. Постановка и дискретизация задачи

Рассмотрим следующую задачу:

$$-div(\mathcal{A} \ grad \ u) + u = f \text{ в } \Omega,$$

$$u = 0 \text{ на } \mathcal{G}_0,$$

$$\frac{\partial u}{\partial \nu_{\mathcal{G}_1}} = 0 \text{ на } \mathcal{G}_1,$$

$$(1)$$

где Ω — ограниченная прямоугольная (полигональная) область в $\mathcal{R}^2(\mathcal{R}^3)$ с границей $\partial\Omega=\bar{\mathcal{G}}_0\cap\mathcal{G}_1,\ \mathcal{G}_0\cap\mathcal{G}_1=\varnothing,\ \mathcal{A}$ — симметричный положительно определенный тензор диффузии и $f\in L_2(\Omega)$ — правая часть. Пусть

$$\epsilon_0, \ \epsilon_1, ..., \epsilon_M = \epsilon$$
 (2)

иерархическая последовательность симплексных (треугольных в \mathcal{R}^2 и тетраэдральных в \mathcal{R}^3) сеток. Каждая сетка ϵ_i , $i = 1, \ldots, M$,

получается локальным разбиением ϵ_{i-1} . Разбиение осуществляется бисекцией определенных (задаваемых пользователем) симплексов из ϵ_{i-1} [1, 7, 8].

Для решения задачи (1) будем использовать неконформный метод конечных элементов. Для этого введем пространство функций $N_k = \{v \subset L_2(\Omega), v|_E \in P_1(E), \forall E \in \epsilon_k\}$, причем v непрерывны в серединах внутренних ребер, а E – стандартный симплекс ϵ_k . Конечноэлементная формулировка задачи (1) принимает вид: найти функцию $\psi^h \in N_M$ такую, что

$$a_h(\psi^h, \varphi^h) = (P^h f, \varphi^h), \quad \varphi^h \in N_M,$$
 (3)

где

$$a_h(\psi^h, \varphi^h) = \sum_{E \in \epsilon_M} (\mathcal{A} \nabla \psi^h, \nabla \varphi^h)_E + (\psi^h, \varphi^h)_E,$$

и $P^h - L_2$ -проектор в $P_0(E)$. Алгебраическая форма (3) — линейная система алгебраических уравнений

$$A\lambda = \tilde{f}.\tag{4}$$

Сделаем несколько важных замечаний. Во-первых, степени свободы (4) ассоцированы с центрами сторон треугольников (граней тетраэдров). Порядок системы равен числу ребер (граней) сетки ϵ_M . Во-вторых, матрица системы A — симметричная положительно определенная и очень сильно разреженная.

3. Многосеточный метод

Для эффективного решения задачи (4) будем использовать многосеточный метод, основанный на иерархической последовательности сеток. А точнее, задачу (4) будем решать с помощью переобусловленного метода сопряженных градиентов с многосеточным методом в качестве переобуславливателя, который опишем, как в работе [4]. Определим оператор продолжения (оператор перехода на более мелкую сетку) $(I_k v)q: N_{k-1} \to N_k$ как

$$I_{k} = \begin{cases} 0, & q \in G_{0}, \\ v(q), & q \in G_{1}, \\ v(q), & q \notin \partial E \text{ if } E \in \epsilon_{k-1}, \\ \frac{1}{2} \{ v|_{E_{1}}(q) + v|_{E_{2}}(q) \}, & q \in \partial E_{1} \cap \partial E_{2} \text{ if } E_{1}, E_{2} \in \epsilon_{k-1}, \end{cases}$$
(5)

где $v \in N_{k-1}$, q — центр масс ребра (грани в трехмерном случае) сетки ϵ_k . Определим оператор проектирования $P_k: N_k \to N_{k-1}$, $k=1,\ldots,M$, как сопряженный к оператору продолжения:

$$(P_k v, w)_{k-1} = (v, I_k w)_k \quad \forall w \in N_{k-1}, v \in N_k, \quad k = 1, \dots, M,$$

где

$$(v,w)_k := \sum_q h_k(q)^d v(q) w(q) \quad v, w \in N_k, \ k = 0, \dots, M,$$

 $h_k(q)$ обозначает диаметр ребра (грани) с центром q, d — размерность пространства.

С помощью операторов продолжения и проектирования сформируем галеркинские проекции матрицы A на грубые сетки:

$$A_M = A, \ A_k = P_{k+1}A_{k+1}I_{k+1}, \ k = M - 1, ..., 0.$$
 (6)

Также введем оператор R_k в виде

$$R_k = \frac{1}{2} (diag A_k)^{-1}, \ k = M, ..., 0.$$
 (7)

Теперь действие многосеточного метода на $g \in N_M$ определим как результат V-цикла с одним пред- и одним постсглаживанием

- 1) Set $u_M = g$;
- 2) For k = M, M 1, ..., 1: $x_k = R_k u_k, u_{k-1} = P_k (u_k A_k x_k)$;

3)
$$w_0 = M_0^{-1} u_0;$$
 (8)

- 4) For k = 1, 2, ..., M: $y_k = x_k + I_k w_{k-1}, w_k = y_k + R_k(u_k A_k y_k)$;
- 5) Set $B_M g = w_M$.

Анализ сходимости (8) говорит о робастности данного метода [4]. Однако скорость работы параллельного аналога целиком зависит от эффективности распараллеливания. Основная трудность

заключается в том, что степени свободы ассоциированы с ребрами, а не с вершинами. Предлагаемый нами алгоритм справляется с данной трудностью.

При описании параллельного алгоритма примем следующие соглашения. Будем описывать двумерный алгоритм, так как трехмерный случай полностью ему идентичен. Пусть $mes_2(\epsilon)$ — площадь области ϵ . Будем считать, что все ребра всех сеток имеют сквозную нумерацию. Сначала нумеруются ребра самой грубой сетки произвольным образом. В процессе построения сетки ϵ_i из ϵ_{i-1} при помощи процедуры бисекции каждое новое ребро получает постепенно увеличивающийся номер. Такая нумерация ребер позволяет для каждого ребра r сетки ϵ_i определить множество \mathcal{K}_r как минимальное подмножество ребер $\{1,...,r-1\}$, благодаря которому можно полностью определить правую часть (5) для вычисления $I_i v$ для ребра r. Согласно этому определению множество \mathcal{K}_r может содержать ребра разных уровней.

4. Параллельный алгоритм

Прежде чем описывать алгоритм, сформулируем признаки, которым будет удовлетворять параллельный аналог многосеточного метода.

4.1. Принципы параллелизации.

- Π 1. Параллельный алгоритм эквивалентен последовательному (8).
- П2. На параллельных компьютерах с умеренным числом процессоров он дает приемлемое ускорение как для равномерных, так и для локально сгущающихся сеток.
- ПЗ. Алгоритм идентичен для двух- и трехмерных сеток.
- П4. Время разбиения области на подобласти в соответствии с числом процессоров пропорционально числу ребер мелкой сетки.
- П5. Число внутренних ребер больше, чем число граничных.
- Пб. Число неизвестных (ребер) равномерно распределено по процессорам.

- П7. Рабочая память алгоритма пропорциональна числу неизвестных.
- **4.2. Разбиение сеточной области.** Параллельный алгоритм решения задачи начинается с разбиения самой мелкой сетки ϵ на множество непересекающихся подобластей-полос (слоев в трехмерном случае):

$$\epsilon^1, \ \epsilon^2, ..., \epsilon^L,$$

где L — число процессоров. Каждая подобласть будет ассоциироваться со своим процессором. Для достижения этой цели используем алгоритм инерциальной бисекции [9].

Такое разбиение соответствует принципам П4-П6 и позволяет легко определить параллельные аналоги операторов продолжения и проектирования.

- 4.3. Квазииерархические последовательности сеток в подобластях. Операторы продолжения и проектирования являются основными в многосеточном методе (8). Они участвуют как в вычислении галеркинских проекций матрицы А на грубые сетки (6), так и при действии самого переобуславливателя. Для определения параллельных аналогов этих операторов каждой подсетке $\epsilon^i, i = 1, ..., L$, поставим в соответствие такие последовательности сеток $\epsilon_k^i,\ k=M-1,...,0,$ чтобы ϵ_k^i состояли из множеств \mathcal{K}_r для любого ребра $r \in \epsilon^i_{k+1} \setminus \epsilon^i_k, \; k=M-1,...,0.$ Подсетки ϵ^i_k строятся следующим образом. Подсетку ϵ_M^i берем совпадающую с ϵ^i , а ϵ^i_k , k=M-1,...,0, получаются из ϵ^i_{k+1} исключением ребер, принадлежащих $\epsilon_{k+1} \setminus \epsilon_k$, и добавлением множеств \mathcal{K}_r для каждого исключенного ребра. Подсетки грубых уровней k < M могут перекрываться из-за возможных расширений. Следовательно, получаем для каждой подобласти квазииерархическую последовательность сеток.
- **4.4. Типы ребер.** Для номеров k < M могут перекрываться не только соседние подсетки ϵ_k^i и $\epsilon_k^{i\pm 1}$, но и подсетки с номерами, отличающимися больше чем на единицу $mes_2(\epsilon_k^i\cap\epsilon_k^j)>0,\ |i-j|>1.$ Этот факт заставляет нас ввести классификацию ребер. Для фиксированного уровня k и фиксированной подсетки ϵ_k^i для ребер, общих с ϵ_k^j , введем следующие типы.

- Если |i-j|=1, ребро назовем локально граничным (тип 1), которое может быть или левым i-j=1, или правым i-j=-1.
- Если |i-j| > 1, то ребро определим как глобальное (тип 2).

Ребра, глобальные для пары подобластей, будем считать глобальными для каждой подсетки фиксированного уровня. Остальные ребра будем называть *внутренними* (тип 0).

Процесс определения типов ребер для каждой подсетки следующий. Сначала классифицируем ребра для подсеток самого мелкого уровня k = M по описанному выше правилу. Далее переходим к сеткам более грубого уровня. При этом классификация ребер производится вместе с генерацией подсетки ϵ_k^i из ϵ_{k+1}^i и только для тех ребер, которые добавляются к сетке ϵ_k^i в процессе получения ϵ_{k+1}^i . Рассмотрим множества \mathcal{K}_r для $r \in (\epsilon_{k+1} \setminus \epsilon_k) \cap \epsilon_{k+1}^i$. Каждое ребро из $r \in (\epsilon_{k+1} \setminus \epsilon_k) \cap \epsilon_{k+1}^i$ передает свой тип ребрам из \mathcal{K}_r . Для некоторых ребер множества \mathcal{K}_r пересекаются и возможны конфликты при определении типов. Для ребра $r_0 \in \mathcal{K}_{r_1} \cup \mathcal{K}_{r_2}$ в случае конфликтной ситуации установим следующее правило: а) если тип ребра r_1 отличается от типа r_2 , то r_0 получает максимальный из типов r_1 и r_2 ; b) если r_1 и r_2 — левое и правое граничные ребра, то r_0 определим как глобальное. После разрешения конфликтных ситуаций синхронизируем глобальные ребра для каждой подобласти. В результате получаем классификацию ребер для каждой подобласти и каждого сеточного уровня. Последовательность сеток ϵ_k^i , k = 1, ..., M, для фиксированного номера i будем хранить на процессоре, отвечающем подобласти с номером i.

4.5. Параллельные аналоги операторов продолжения и проектирования. Построенные квазииерархические последовательности сеток $\epsilon_k^i,\ k=1,...,M,$ для каждой подобласти дают возможность определить параллельные аналоги операторов продолжения и проектирования. Для каждого $r\in\epsilon_k^i$ оператор продолжения зависит от значений на ребрах, входящих в \mathcal{K}_r , которые содержатся на процессоре i в силу определения ϵ_{k-1}^i . Это позволяет ввести оператор продолжения для каждого процессора независимо от значений на других процессорах. Действие оператора проектирования P_k^i заключается в суммировании влияний ребра

r на степени свободы, ассоциированные с \mathcal{K}_r . Это включает в себя обмен между процессорами. Алгоритм разбиения всей области на подобласти может производить подсетки такие, что для некоторых ребер r элементы множества \mathcal{K}_r принадлежат сразу нескольким подобластям. Поэтому после вычисления значений \mathcal{K}_r для каждого процессора необходимо взять соответствующие данные с других процессоров.

4.6. Генерация матриц уровней. Матрицы уровней A_k занимают большую часть памяти, поэтому они распределены по процессорам. Матрица самого мелкого уровня представлена как сумма матриц $A_M^i, i=1,\ldots,L$, которые получены ассемблированием локальных матриц по ячейкам сетки, соответствующей процессору i. Матрицы других уровней $A_k, k=M-1,\ldots,1$, являются галеркинскими проекциями (6) матриц самой мелкой сетки на ϵ_k . Проекции матрицы A_k на сетки ϵ_k^i вычисляются на основании A_{k+1} и локальных операторов продолжения и проектирования P_{k+1}^i, I_{k+1}^i : $A_k^i = P_{k+1}^i A_{k+1}^i I_{k+1}^i, k=M-1,\ldots,1$. При этом перемножение осуществляется локально на каждом процессоре.

Как было отмечено выше, подсетки ϵ_k^i при увеличении i могут как расширяться, так и содержать для некоторых r множество \mathcal{K}_r , состоящее из ребер разных уровней. Это ведет к тому, что матрицы A_k при уменьшении k, то есть при переходе к более грубым сеткам, будут постепенно терять свойство разреженности, становясь более плотными.

4.7. Умножение на матрицу уровня. Умножение матрицы A_k^i на вектор выполняется по-разному для разных типов ребер. Для внутренних ребер выполняется матрично-векторное умножение без обмена данными между процессорами. Для локально граничных степеней свободы используется обмен информации между соседними процессорами, для глобальных ребер выполняется процедура обмена между всеми процессорами.

Постепенная заполняемость матриц уровней и наличие обменов делают операцию умножения матрицы A_k^i на вектор достаточно трудоемкой. Расширение подсеток ϵ_k^i при увеличении i ведет к возрастанию числа глобальных ребер, что приводит к увеличению времени, требуемого на глобальные обмены между процессорами. Это является основной сложностью описанного здесь

Таблица 1

Сетка	Число	Число процессоров				
	треугольников	1	2	4	8	16
Равномерная	131072	21.88	17.85	9.79	5.36	4.16
Равномерная	261144	=	-	22.56	12.42	7.90
Сгущение						
к точке	113808	22.56	18.97	10.80	7.26	5.87
Сгущение						
к точке	193288	-	-	29.07	18.44	13.23

Таблица 2

Сетка	Число	Число процессоров				
	тетраэдров	1	2	4	8	16
Равномерная	65536	27.95	22.03	12.59	9.03	8.34
Равномерная	131072	-	-	32.51	21.18	17.46
Сгущение						
к точке	56176	37.66	30.38	23.18	18.59	19.96
Сгущение						
к точке	88944	-	-	50.82	41.32	40.02

Таблица 3

Сетка	Число	Число процессоров			3	
	треугольников	1	2	4	8	12
Равномерная	131072	4.94	3.05	2.01	1.28	1.04
Равномерная	261144	-	-	4.99	2.92	2.20
Сгущение						
к точке	113808	4.59	2.78	1.78	1.25	1.03
Сгущение						
к точке	193288	-	-	3.81	2.38	1.91

параллельного аналога многосеточного метода. Таким образом, параллельный эффект зависит от числа глобальных ребер: чем оно меньше, тем матрично-векторное умножение будет более быстрым.

5. Численные эксперименты и их анализ

В этом разделе приведены результаты численных экспериментов и их анализ. Эксперименты проводились на кластерах МГУ и ИВМ РАН (табл. 1 и 2).

Эксперименты показывают, что использование параллельного аналога многосеточного метода дает выигрыш сразу по двум направлениям. Во-первых, при умеренном увеличении числа процессоров мы получаем уменьшение времени решения задачи, а вовторых, параллельные версии позволяют решать линейные системы большего порядка по сравнению с последовательными. Это возможно благодаря распределению всех данных по процессорам. Однако эффективность параллелизации рассматриваемого метода наблюдается лишь при малом числе процессоров. Увеличение числа процессоров влечет за собой рост количества глобальных ребер, что в свою очередь делает операцию умножения на матрицы уровней более трудоемкой за счет увеличения числа глобальных обменов. Это приводит к замедлению работы программы и, как следствие, к уменьшению выигрыша. На локально сгущающихся сетках мы получаем более медленное ускорение, чем на равномерных, которое кроме увеличения числа глобальных обменов обусловлено более сильной заполненностью галеркинских проекций матрицы A на более грубые сетки. Число глобальных ребер в трехмерном случае оказывается больше, чем в двумерном, — имен-

Таблица 4

Сетка	Число	Число процессоров				3
	тетраэдров	1	2	4	8	12
равномерная	65536	5.95	3.57	2.53	1.75	1.76
равномерная	131072	-	-	6.23	4.12	4.13
сгущение						
к точке	56176	6.76	4.81	3.65	3.56	4.56
сгущение						
к точке	88944	-	-	5.94	5.47	6.73

но этим объясняется больший выигрыш в двумерном случае при увеличении числа процессоров. Если сравнить работу двух кластеров, на которых производились численные эксперименты, то в ИВМ РАН параллельная машина обладает более мощными процессорами. Это видно из численных экспериментов. Поэтому на первый план выходят обмены, которые уменьшают выигрыш при увеличении числа процессоров про сравнению с кластером МГУ.

Литература

- [1] Bänsch E. Local mesh refinement in 2 and 3 dimensions // IM-PACT of Computing in Science and Engrg. Vol. 3. P.181–191. 1991.
- [2] Raviart P., Thomas J. A mixed finite element method for 2-nd order elliptic problems // Mathematical Aspects of Finite Element Methods / Lecture. Eds. I.Galligani and E.Magenes.
- [3] Chen Z. Equivalence between multigrid algorithms for nonconforming and mixed methods for second-order elliptic problems // East-West J.Numer.Math. 1996. Vol. 4. P. 1–33.
- [4] Chen Z. On the convergence of Galerkin-multigrid methods for nonconforming finite elements // East-West J.Numer.Math. 1999. Vol. 7. P. 79–104.
- [5] Chugunov V., Vassilevski Yu. Parallel multilevel data structures for a non-conforming finite element problem on unstructured meshes // Russ.J.Numer.Anal.Math.Modelling. 2003. Vol.18, No. 1. P. 1-11.
- [6] Chugunov V., Svyatski D., Tyrtyshnikov E. Vassilevski Yu. Parallel iterative multilevel solution of mixed finite element systems for scalar equations // To appear in Concurrency and Computation: Practice and Experience
- [7] Rivara M. Mesh refinement processes based on the generalized bisection of simplexes.// SIAM J.Numer.Anal. 1984. Vol. 21. P. 604–613.

- [8] Rivara M. Selective refinement/derefinement algorithms for sequences of nested triangulations // Int.J.Numer.Meth.Engrg. 1989. Vol. 28. P. 2889–2906.
- [9] Williams R. Performance of dynamic load balancing algorithms for unstructured mesh calculations // Concurrency. 1991. Vol. 3. P. 457.

Факторизация сеточного Лапласиана на последовательностях сеток: экспериментальные оценки вычислительной работы

Ю. В. Василевский, А. В. Прокопенко

В работе рассмотрены вычислительные аспекты прямого решения больших систем с разреженными матрицами в рамках нескольких общедоступных современных реализаций метода исключения. Анализ осуществлялся в рамках последовательностей матриц, соответствующих общей структуре разреженности, определяемой сеткой и методом дискретизации. Экспериментальным путем получены асимптотические формулы числа ненулевых элементов в верхнем и нижнем треугольных множителях и объема арифметических вычислений, а также сравнительные времена факторизации и последующего решения треугольных систем, характеристики разреженности треугольных множителей и запросов на компьютерную память. На основании анализа сделано несколько практических выводов по использованию методов исключения.

1. Введение

Бурное развитие итерационных методов решения линейных систем с разреженными матрицами обусловлено тем обстоятельством, что эффективные прямые методы применимы лишь к узкому классу матриц, а методы факторизации произвольных разреженных матриц на нижний и верхний треугольные множители A = LU налагают серьезные ограничения на порядок матриц. Эти ограничения выражаются как в большом времени факторизации, так и в больших объемах используемой машинной памяти. С другой стороны, методы факторизации обладают несомненным достоинством, а именно, минимальными входными данными в терминах исключительно ненулевых элементов матрицы системы. Это позволяет использовать методы как технологии "черного ящика", что очень удобно в прикладных расчетах. За последнее

десятилетие было предложено несколько современных реализаций [3, 4, 5, 7] методов факторизации разреженных матриц, которые являются общедоступными (через интернет). Целью настоящей работы является экспериментальное сравнение этих реализаций для: а) выбора наилучшей реализации для конкретного приложения и б) априорной оценки границы применимости выбранного метода.

Сравнения, проведенные ранее [7, 8], осуществлялись для конкретных матриц, и из них нельзя сделать вывод о поведении метода при увеличении порядка матрицы, а также о границах применимости метода. Мы рассмотрим несколько последовательностей наиболее популярных разреженных матриц с одинаковой структурой разреженности, которая порождается фиксированной дискретизацией оператора Лапласа и набором измельчающихся сеток. Сравнение и анализ методов будут проводиться в рамках каждой последовательности матриц различных порядков, что позволит делать экстраполяционные оценки границ применимости методов. Несмотря на то, что рассматриваемые матрицы являются симметричными и положительно определенными, исследуемые методы не использовали эту информацию, работая с матрицами как с разреженными матрицами общего вида. Другим принципиальным моментом является использование параметров по умолчанию, т.е. таких значений управляющих факторизацией параметров, которые предлагаются для матриц общего вида. Объектом настоящего исследования является стандартный алгоритм ("черный ящик") факторизации разреженных матриц, поэтому мы не рассматриваем возможности некоторого улучшения производительности того или иного алгоритма с учетом дополнительной информации о матрице. Будем тестировать следующие программные (однопроцессорные) реализации LU-разложения:

- MUMPS (version 4.3.1) MUlti-frontal Massively Parallel Solver [3];
- WSMP (version 4.01.31) Watson Sparse Matrix Package [7];
- UMF (PACK, version 4.1) Unsymmetric Multi-Frontal method [4];
- SuperLU(version 3.0) SuperLU package [5].

Первые три алгоритма используют многофронтальный метод факторизации, последний применяет разреженную факторизацию по суперузлам, причем все алгоритмы заранее переупорядочивают матрицы с целью уменьшения заполнения в треугольных множителях [2]. В целях более полного сравнения мы приводим также данные, полученные с помощью реализации других методов разреженной факторизации:

- ТМ алгоритм минимальной степени Марковица [10] (реализация А.Н. Беспалова);
- ND алгоритм вложенных сечений [6] (реализация С.А. Горейнова).

Главная стратегия метода ТМ — локальная минимизация заполнения на каждом шаге исключения за счет выбора главного элемента в той строке и столбце, которые обеспечивают внесение наименьшего числа ненулевых элементов в треугольные множители Lи U. Полученное число ненулевых элементов в L и U, NZ_{UU} , может считаться разумной оценкой минимально возможного количества ненулевых элементов в треугольных множителях данной матрицы. Метод вложенных сечений систематически разбивает граф матрицы при помощи разделителей, которые должны находиться автоматически. Для матриц с определенной структурой разреженности такие разделители можно указать явно, и поэтому удается априорно оценить как NZ_{LU} , так и объем арифметических затрат на факторизацию [1, 9]. К сожалению, эффективность автоматического выбора разделителей зависит от структуры разреженности матрицы, поэтому эффективность метода (хотя бы с точки зрения минимизации NZ_{LU}) наблюдалась нами лишь для матриц, связанных с равномерными сетками.

Расчеты проводились на персональном компьютере с процессором Pentium 4 (2.5GH) и оперативной памятью 750 MB, на операционной системе Linux (SuSE 8.1). Исходные Fortran77- и С-коды компилировались на компиляторах проекта Gnu, Fortran90-коды компилировались на компиляторе фирмы Intel со стандартными оптимизационными опциями.

Статья построена следующим образом. В первом разделе описываются последовательности разреженных матриц. В разделах

2 и 3 представлены экспериментально определенные асимптотические зависимости количества ненулевых элементов в множителях L и U и объем арифметической работы при выполнении факторизации. В разделах 4 и 5 исследуемые схемы сравниваются с точки зрения компактности треугольных множителей и быстроты факторизации. В разделе 6 мы попытались очертить современный предел порядков популярных разреженных матриц, для которых применим последовательный метод факторизации.

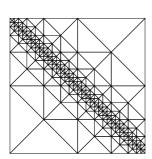
2. Типы и последовательности сеток

Процедуру факторизации будем исследовать для последовательностей матриц возрастающей размерности, соответствующих разным структурам разреженности. Выбор структуры разреженности основан на практическом опыте: рассматриваются системы, наиболее часто встречающиеся в приложениях. Как правило, системы с разреженными матрицами возникают из различных дискретизаций краевых задач. Шаблон дискретизации, определяющий структуру разреженности, зависит от оператора краевой задачи, метода дискретизации и сетки. Поскольку сеточный оператор Лапласа предоставляет наиболее репрезентативный шаблон для многих краевых задач, мы ограничились только сеточным лапласианом. Таким образом, в нашем исследовании структура разреженности будет определяться только сеткой и методом дискретизации, а последовательность матриц возрастающей размерности будет задаваться последовательностью измельчающихся сеток. С помощью изменения метода дискретизации мы варьируем локальный шаблон и локальную геометрию структуры разреженности, а с помощью задания сетки мы определяем взаимосвязи локальных шаблонов и глобальную структуру разреженности.

Множество исследуемых матриц можно разбить на группы согласно следующим формальным признакам: двумерная или трехмерная сетка, прямоугольная или симплициальная (т.е. состоящая из треугольников или тетраэдров), а также сеточный шаблон, порожденный дискретизацией определенного типа. Мы рассмотрим аппроксимации конечно-разностные на прямоугольных сетках и конечно-элементные на симплициальных сетках, причем во втором случае рассмотрим аппроксимации кусочно-линейными базисными функциями как конформного типа (функции Куранта),

так и неконформного типа (обеспечивающие непрерывность решения только в серединах симплициальных граней). Последний тип имеет важное прикладное значение, поскольку разреженность получаемой матрицы точно такая же, к какой сводится смешанная конечно-элементная консервативная дискретизация эллиптических операторов путем поэлементной алгебраической конденсации.

Набор прямоугольных двумерных сеток с квадратными ячейками, в которых количество ячеек по обоим направлениям одинаково, будем называть последовательностью регулярных двумерных изотропных сеток (тип 1). Набор прямоугольных трехмерных сеток с кубическими ячейками, в которых количество ячеек по всем трем направлениям одинаково, будем называть последовательностью регулярных трехмерных изотропных сеток (тип 2). Если же количество ячеек растет только по одному направлению, а по другим двум количества ячеек фиксированы (в нашем случае они равны пяти), получим последовательность регулярных трехмерных анизотропных сеток (тип 3). Набор триангуляций единичного квадрата, в котором каждая более мелкая сетка строится из предыдущей путем бисекции (измельчения) тех треугольников, которые пересекаются с главной диагональю квадрата, назовем последовательностью двумерных неструктурированных сеток (тип 4), (рис.1,слева). Аналогично, набор тетраэдризаций полиэдра с вершинами (0,0,0), (1,0,0), (0,1,0), (0,0,1), (0,0,-1), в котором каждая более мелкая сетка строится из предыдущей путем бисекции (измельчения) тех тетраэдров, которые пересекаются с плоскостью z = 0, назовем последовательностью трехмерных неструктурированных сеток (тип 5), (рис.1,справа). Отметим, что каждая последовательность сеток получается последовательным измельчением: для типа 1 равномерным уменьшением шага сетки в два раза в обоих направлениях; для типа 2 равномерным уменьшением шага сетки в обоих направлениях в рамках последовательности (1/20,1/30, 1/40, 1/50); для типа 3 равномерным уменьшением шага сетки в четыре раза в одном направлении; для типа 4 локальным уменьшением размера треугольников, пересекающихся с главной диагональю, в четыре раза; для типа 5 локальным уменьшением размера тетраэдров, пересекающихся с плоскостью z=0, в два раза.



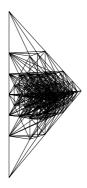


Рис. 1. Неструктурированные сетки

При выборе последовательностей сеток мы руководствовались следующими соображениями: регулярные изотропные сетки обеспечивают самые распространенные сеточные шаблоны простейших конечно-разностных аппроксимаций; регулярные анизотропные сетки порождают матрицы, обладающие фиксированной шириной ленты после соответствующего переупорядочения; неконформные аппроксимации на неструктурированных сетках (типы 4, 5) порождают матрицы, в которых 5 (2D) или 7 (3D) ненулевых элементов в строке, т.е. их разреженность точно такая же, как у конечно-разностных аппроксимаций на регулярных сетках, однако графы матриц радикально отличаются; конформные аппроксимации на неструктурированных сетках порождают матрицы с примерно 7 (2D) или 15 (3D) ненулевыми элементами в строках, причем их граф совпадает с сеткой. Помимо разного заполнения строк, конформная аппроксимация отличается от неконформной разными носителями степеней свободы (узлы и грани ячеек соответственно). Для удобства обозначений набор матриц, возникающих на разных типах сеток, будем различать по типам, причем конформные аппроксимации порождают последовательность матриц типа 4с и 5с, а неконформные аппроксимации порождают последовательность матриц типа 4n и 5n. В табл.1 приведены основные сведения о матрицах A_k из последовательностей различных типов: их порядок N, указанное в скобках число ненулевых элементов NZ, а также дополнительная геометрическая информация в виде количества шагов сетки вдоль координатных осей n, или количества уровней локального измельчения симплексов $Lv=n_1$ - n_2 , где n_1 соответствует первой сетке последовательности, а n_2 — последней.

Таблица 1

тип $\setminus A_k$	A_1	A_2	A_3	A_4
$1, n^2 =$	10000(49600)	40000 (199200)	160000 (798400)	640000 (3196800)
$2, n^3 =$	8000 (53600)	27000 (183600)	64000 (438400)	125000 (860000)
3, 25n =	10000(61950)	40000 (247950)	160000 (991950)	640000 (3967950)
4c, Lv = 18-30	3563 (24855)	14311 (100075)	57315 (401087)	$229343 \ (1605267)$
4n, Lv = 18-30	10646 (53150)	42882 (214314)	171886 (859318)	687962 (3439682)
5c, $Lv = 15-24$	1783 (23875)	7039 (98011)	27983 (397835)	111599 (1603883)
5n, Lv = 12-21	4144 (27760)	18044 (123404)	75832 (524440)	311636 (2167940)

3. АСИМПТОТИКА КОЛИЧЕСТВА НЕНУЛЕВЫХ ЭЛЕМЕНТОВ В МНОЖИТЕЛЯХ

В этом разделе мы исследуем асимптотику степепени разреженности получаемых треугольных множителей (в терминах суммарного числа ненулевых элементов NZ_{III}), которая является важной характеристикой компактности данных, необходимых для последующего решения системы (конкретные значения NZ_{LU} представлены в разделе 4). Отношения $\ln(NZ_{LU})/\ln f(N)$ указаны в таблицах для наибольшей матрицы в последовательности каждого типа, где f(N) — примерная функция асимптотической зависимости $NZ_{LU}\simeq f(N),\ N\to\infty$. Число в скобках указывает коридор отклонения $\ln(NZ_{LU})/\ln f(N)$ в пределах каждой последовательности, причем его знак определяет тенденцию этого отклонения. Так, например, данные первого столбца и второй строки табл.2 показывают, что для матрицы типа 4с порядка N=229343с NZ = 1605267 ненулевыми элементами и метода наименьшей степени ТМ $\ln(NZ_{LU})/(\ln C + \ln N + \ln \ln N) = 1.001$, причем при уменьшении порядка матриц в пределах последовательности 4с это отношение растет, но не более чем на 2.7 %. Это означает, что при дальнейшем увеличении N оотношение будет уменьшаться, но в пределах малой ошибки, поэтому $f(N) \leq N \ln N$. Отметим, что выстраивание асимптотик по четырем точкам является весьма сильным допущением, которое может быть правомерным только в случае малых ошибок в пределах последовательности каждого типа и большого разброса величины N.

Таблица 2

	Метод					
тип	TM	UMF	WSMP	MUMPS		
1	C = 1.8	C = 4.88	C = 4.95	C = 3.8		
	1.008 (-0.006)	1.014 (-0.024)	1.014 (-0.009)	1.015(-0.005)		
4c	C = 1	C = 1.9	C = 2.5	C = 2.4		
	1.001 (+0.027)	1.001 (+0.027)	1.002 (+0.028)	1.001 (+0.027)		
4n	C = 1	C = 1	C = 1	C = 1		
	$0.986 \ (+0.025)$	1.015 (+0.047)	1.035 (+0.065)	1.033 (+0.061)		

3.1. Двумерные сетки. В табл.2 представлены данные (вида $\ln(NZ_{LU})/(\ln C + \ln N + \ln \ln N)$) по двумерным сеткам, которые с указанной точностью удовлетворяют асимптотической формуле

$$NZ_{LU} \simeq C \cdot N \cdot \ln N$$

для матриц с плоскими графами [2, 1, 6, 9]. Для равномерных сеток C>1, причем минимальное значение C обеспечено алгоритмом ТМ. Отметим, что для метода ND C=3.62 для последовательностей матриц типа 1, т.е. алгоритм вложенных сечений обеспечивает следующее за алгоритмом ТМ минимальное заполнение. Для последовательностей на двумерных сгущающихся сетках C=1, причем тенденция отклонения указывает на то, что асимптотически $NZ_{LU} < N \ln N$ при $N \to \infty$. Это замечание кажется естественным, если учесть, что на последовательности локально измельчающихся сеток все большая часть графа матрицы становится "одномерной", а для одномерных графов $NZ_{LU} \simeq CN$.

3.2. ТРЕХМЕРНЫЕ СЕТКИ. В табл.3 приведены данные (вида $\ln(NZ_{LU})/\ln N$) по трехмерным сеткам, которые с указанной точностью удовлетворяют асимптотической формуле

$$NZ_{LU} \simeq N^{\alpha}$$
. (1)

Отметим, что $1.5\leqslant \alpha\leqslant 1.59$ для последовательности матриц типа 2. Подчеркнем, что на этой же последовательности $\alpha=1.54$ для метода вложенных сечений, что есть опять же второй результат после алгоритма минимальной степени. Интересно, что ни один из методов даже не приблизился к оценке [9]

$$NZ_{LU} \simeq N^{4/3}$$
.

Таблица 3

	Метод					
тип	TM	UMF	WSMP	MUMPS		
2	1.5 (+0.0)	1.587 (+0.003)	1.57 (+0.026)	1.55 (+0.03)		
5c	1.453 (+0.019)	1.508 (+0.054)	1.479 (+0.113)	1.478 (+0.104)		
5n	1.357 (+0.005)	$1.423 \ (+0.023)$	1.377 (+0.076)	1.395 (+0.07)		

Таблица 4

			Метод		
тип	TM	UMF	SuperLU	MUMPS	WSMP
3	C = 16.5	C = 31.8	C = 50	C = 52.8	C = 62.7
	0.997 (-0.005)	0.997 (-0.002)	$0.999(\pm 0.001)$	0.997(-0.004)	0.997 (+0.003)

При измельчении сеток типа 5 граф матрицы приближается к плоскому графу, поэтому α меньше, чем степень 1.5, наблюдаемая у строго трехмерного графа. Характер отклонений в рамках последовательностей показывает, что оценки (2) являются несколько завышенными для указанных степеней α .

3.3. Матрицы, сводимые к ленточным.

Трехмерные сетки с анизотропным измельчением (тип 3) позволяют переупорядочить матрицы так, чтобы они стали ленточными с полушириной ленты 25. Таким образом, после переупорядочения даже ленточная факторизация обеспечила бы $NZ_{LU}=50N$. Интересно отметить, что эффективные алгоритмы MUMPS и WSMP не смогли уложиться в эту тривиальную оценку, хотя асимптотически данные (вида $\ln(NZ_{LU})/(\ln C + \ln N)$) табл.4 подтверждают, что

$$NZ_{LU} \simeq C \cdot N$$
.

4. Асимптотика арифметической работы

Некоторые программные комплексы позволяют подсчитать количество $Flops_{LU}$ арифметических операций с плавающей точкой, выполненных в процессе факторизации. В табл.5 приведены значения отношений $\ln(Flops_{LU})/(\ln C + \ln N)$, которые с указанной

точностью удовлетворяют асимптотической формуле

$$Flops_{LU} \simeq C \cdot N^{\alpha}$$

для факторизации матриц разных типов, связанных с прямоугольными сетками. Отметим, что $1.5\leqslant\alpha\leqslant1.6$ для последовательности матриц типа 1 (теоретическая оценка [1] $\alpha=1.5$); $2.2\leqslant\alpha\leqslant2.3$ для последовательности матриц типа 2 (теоретическая оценка [9] $\alpha=2$); $\alpha=1$ для последовательности матриц типа 3 (теоретическая оценка $\alpha=1$); на неструктурированных сетках показатель α не превышает 1.2 для двумерных сеток и 2 для трехмерных. Уменьшение показателей по сравнению с равномерными сетками обусловлено тем фактом, что графы матриц типа 4 становятся все более одномерными (плоскими для типа 5) при $N\to\infty$.

5. Что компактнее?

Сравнительные характеристики заполняемости множителей Lи U представлены в табл.6 для наибольших (и некоторых других) матриц в каждой последовательности. Отметим, что для метода вложенных сечений $NZ_{LU}=7.35\cdot 10^6$ для типа $2,\,k=2,\,$ а для типа $1 NZ_{LU} = 8.7 \cdot 10^6$ при k = 3 и $40 \cdot 10^6$ при k = 4, соответственно. Метод наименьшей степени (ТМ) дает наиболее разреженные матрицы, однако его низкая арифметическая эффективность не позволяет использовать его для практической факторизации разреженных матриц порядка 10^5 . Для последовательностей на трехмерных сетках наиболее компактными оказываются технологии WSMP и MUMPS, причем на равномерной сетке (тип 2) и для конформных аппроксимаций (тип 5c) предпочтительнее выглядит MUMPS, в то время как для неконформных аппроксимаций (тип 5n) — WSMP, хотя число ненулевых элементов в строке в последнем случае такое же, как у матриц типа 2. Для двумерных сеток преимущество MUMPS наблюдается только на равномерной сетке, а на сгущающихся сетках (как и на сетках типа 3) более компактными оказываются множители UMF. Отметим, что значения NZ_{LU} у методов MUMPS и ND очень близки на равномерных сетках.

6. Что выстрее?

Времена факторизации и последующего решения (в скобках) представлены в табл.7.

Таблица 5

тип	1	2	3
WSMP	C = 30	C = 1	C = 2906
	$1.503\ (\pm0.014)$	2.211 (+0.014)	$1 \ (\pm 0.005)$
UMF	C = 5	C = 1	C = 672
	$1.58 \ (\pm 0.015)$	2.265 (-0.015)	$1 \ (\pm 0.05)$

Таблица 6

тип, k	TM	MUMPS	WSMP	UMF	SuperLU
2,k = 2	4.45	8.15	10.6	11.2	26
2,k = 4		82	100	123	
5c, k = 3	2.9	4.7	5.2	6.4	19
5c, k = 4		28.8	29.3	41	117
5n,k = 3	4.2	7.7	6.8	9.6	40
5n,k=4		45.9	36.5	66	
1, k = 3	3.9	8.9	11.8	11	17.6
1, k = 4		42	54.7	54	86
4c, k = 4	3.8	6.9	7.3	5.5	8.7
4n,k=4	7.4	15.8	16.3	11.7	28.9
3, k = 4	10	32	38	19.3	31.5

Таблица 7

тип, к	MUMPS	WSMP	UMF	SuperLU
2, k = 2	34.6 (0.09)	42.4 (0.12)	62.9 (0.26)	75.5 (0.2)
2, k = 4	1107 (1.57)	1241	2448	
5c, k = 4	80.3 (0.44)	75.7 (0.42)	215(1.29)	1169(2.7)
5n, k = 4	169.6(0.8)	76.5 (1.23)	435(2.1)	
1, k = 4	82 (0.97)	127 (2.0)	136 (2.2)	112 (1.6)
4c, k = 4	14.9 (0.6)	11.2 (0.7)	5.6 (0.6)	4.2 (0.18)
4n, k = 4	43.3 (1.23)	34.5 (1.4)	17.5 (1.3)	11 (0.54)
3, k = 4	39.5 (0.86)	37.8 (1.5)	154.5(1.0)	5.7 (0.41)

Поскольку факторизация наиболее трудоемкая процедура, проанализируем время факторизации более подробно. Во-первых, на равномерных сетках (типы 1 и 2) MUMPS-факторизация является самой быстрой. Во-вторых, для матриц типов 5с и 5п наиболее эффективным является метод WSMP, а для матриц типов 4c, 4n и 3 самой быстрой оказалась факторизация SuperLU, за которой следует UMF (за исключением типа 3, где UMF-факторизация оказалась совершенно не эффективной). Отметим быстроту SuperLUфакторизации матриц, у которых разделители графа содержат ограниченное число вершин (типы 4c, 4n), несмотря на то, что $NZ_{LU}(SuperLU) > 1.5NZ_{LU}(UMF)$. Другим интересным наблюдением является несогласованность времени факторизации, времени решения факторизованной системы и величины NZ_{LU} . Действительно, WSMP-факторизация для наибольших матриц типов 5с и 5п требует примерно 76 с, а решения систем с матрицами L и U занимают 0.42 и 1.23 с, в то время как NZ_{LU} равны 29.3и 36.5 миллиона соответственно. Таким образом, близость числа ненулевых элементов в треугольных множителях не означает близость времен решения соответствующих систем даже в рамках одной реализации. С другой стороны, для наибольшей матрицы типа 5n время MUMPS-факторизации более чем в два раза превышает время WSMP-факторизации, $NZ_{LU}(MUMPS)$ больше, чем $NZ_{LU}(WSMP)$, на 30%, а время MUMPS-решения в 1.5 раза меньше времени WSMP-решения. Это означает, что при сравнении различных реализаций разные критерии могут порождать разные выводы.

7. Где практический предел порядка матрицы?

Наряду с объемом арифметической работы серьезным ограничением на использование разреженной факторизации являются размеры машинной памяти. В табл.8 приведены размеры используемой памяти (в МВ) для наибольших матриц разных типов и некоторых наиболее удобных, на наш взгляд, пакетов. Для матриц типа 2 (семиточечный шаблон на равномерной 3D сетке) сетки $60 \times 60 \times 60$ являются, по-видимому, максимальными для современных компьютеров. Отметим, что параллельная версия WSMP [7] обеспечила факторизацию матрицы типа 2 на сетке $100 \times 100 \times 100$ на 32 процессорах. Предел для пятиточечного ша-

Таблица 8

тип	MUMPS	WSMP	UMF
2	926	993	1286
5c	319	310	377
5n	510	410	615
1	485	647	517
4c	115	112	79
4n	265	330	210
3	463	540	271

Таблица 9

сетка	1	2	3	4
$N = 10^3 \times$	8	27	64	125
MUMPS	18	97	333	926
UMF	21	138	405	1286

блона на равномерной 2D сетке задается, по-видимому, на сетке порядка 1000×1000 для факторизации на одном процессоре, поскольку сетка 800×800 требовала 500-700 MB оперативной памяти. Отметим, что при факторизации матриц типа 1 максимальная используемая память растет почти пропорционально порядку матриц: при увеличении порядка в 4 раза память возрастает в 4.5-5 раз. Динамика запросов на память (в MB) для матриц типа 2 представлена в табл.9.

Заключение

В работе рассмотрены некоторые практические аспекты прямого решения больших систем с разреженными матрицами. Было исследовано несколько общедоступных современных реализаций (MUMPS, WSMP, UMFPACK, SuperLU) метода исключения в приложении к разреженным матрицам. Анализ осуществлялся в рамках последовательностей матриц, соответствующих общей структуре разреженности, определяемой сеткой и методом дискретизации. Экспериментальным путем получены асимптотические формулы числа ненулевых элементов в верхнем и нижнем треугольных множителях, а также объема арифметических вычислений для некоторых методов. Получены сравнительные време-

на факторизации и последующего решения треугольных систем, а также характеристики разреженности треугольных множителей и запросов на компьютерную память. Принципиальным моментом являлся выбор стандартных параметров при вызове процедур, при котором не передавалась никакая дополнительная информация о разреженной матрице (симметричность, структура разреженности и т.д.). Анализ полученных данных позволил сделать несколько практических выводов:

- Не существует метода, являющегося наилучшим для последовательностей матриц всех типов.
- На эффективность метода влияет не только количество ненулевых элементов в строке, но и структура графа матрицы (сетка и дискретизация).
- Время факторизации, время решения систем с треугольными матрицами, и число ненулевых элементов в множителях не всегда согласованы. Поэтому выбирать оптимальный метод нужно исходя из какого-то одного критерия.
- Современные технологии позволяют решать системы с разреженными матрицами порядка сотен тысяч на однопроцессорном компьютере с 500-700 MB оперативной памяти.
- Представленные данные могут быть справочным материалом при априорной оценке целесообразности использования прямых методов для решения систем с разреженными матрицами

Работа финансировалась за счет гранта РФФИ 04-07-90336-в и госпрограммы РАН "Вычислительные и информационные проблемы решения больших задач".

Литература

- [1] Джордж А., Лю Дж. Численное решение больших разреженных систем уравнений. М.: Мир, 1984.
- [2] Писсанецки С. Технология разреженных матриц. М.: Мир, 1988.

- [3] Amestoy P., Duff I., L'Excellent J., Li X. Analysis and comparison of two general sparse solvers for distributed memory computers. // ACM Transactions on Mathematical Software, 2001, Vol.27, p.388-421.
- [4] Davis T., Duff I. An unsymmetric-pattern multifrontal method for sparse LU factorization. // SIMAX, 1997, Vol.18, p.140-158.
- [5] Demmel J., Gilbert J., Li X. An asynchronous parallel supernodal algorithm for sparse Gaussian elimination. //SIAM J.Matr.Anal.Appl., 1999, Vol.20, p.915-952.
- [6] George A., Liu J. An automatic nested dissection for irregular finite element problems. //SIAM J.Numer.Anal., 1978, Vol.15, p.1053-1069.
- [7] Gupta A. Improved symbolic and numerical factorization for unsymmetric sparse matrices. // SIAM J.Matr.Anal.Appl., 2002, Vol.24, p.529-552.
- [8] Gupta A. Recent advances in direct methods for solving unsymmetric sparse systems of linear equations. // ACM Transactions on Mathematical Software, 2002, Vol.28, p.301-324.
- [9] Lipton R., Rose D., Tarjan R. Generalized nested dissection. // SIAM J.Numer.Anal., 1979, Vol.16, p.346-358.
- [10] Markowitz H. The elimination form of the inverse and its application to linear programming. // Management Sci., 1957, Vol.3, p.255-269.

Параллельное моделирование особенностей кровотока в окрестности кава-фильтра с захваченным тромбом

Ю. В. Василевский, С. А. Капранов

В работе рассмотрены основные аспекты численного моделирования течения крови в вене с имплантированным кава-фильтром с захваченным тромбом. Геометрические особенности модели требуют большого числа расчетных узлов, поэтому решение возникающих систем возможно только благодаря современным эффективным численным методам. В статье представлен алгоритмический и технологический базис для параллельной компьютерной реализации модели течения, включающий как методы дискретизации, так и методы решения возникающих систем.

1. Введение

Целью работы является краткое описание первого этапа совместного сотрудничества специалистов по эндоваскулярной хирургии и специалистов по компьютерному моделированию в рамках поиска решения конкретной прикладной задачи. Задача описывается в медицинских терминах, формулируются предположения, в рамках которых выбирается математическая модель течения жидкости, обсуждаются аспекты компьютерного моделирования как с точки зрения методов дискретизации, так и с точки зрения решения возникающих сеточных уравнений. Наконец, обсуждаются результаты первых численных экспериментов и рассматриваются некоторые перспективы применения более адекватных математических и геометрических моделей.

2. МЕДИЦИНСКАЯ ПОСТАНОВКА ЗАДАЧИ

Распространенность острых тромбозов глубоких вен нижних конечностей и таза очень велика. Ежегодно сотни тысяч пациентов с этой патологией в развитых странах госпитализируются и получают лечение. Наиболее грозным осложнением заболевания

является отрыв венозного тромба от места его фиксации к стенке сосуда и затем миграция с током крови по нижней полой вене с последующим попаданием в сердце и легочную артерию (т.н. тромбоэмболия). Это осложнение сопровождается высокой частотой летальных исходов. Для предотвращения легочной эмболии наиболее эффективным способом профилактики является имплантация кава-фильтров (интравенозных фильтров) в нижнюю полую вену. В настоящее время использование кава-фильтров находит все большее признание в России и за рубежом [1, 3]. Кава-фильтры представляют собой конусообразную конструкцию из тонких проволочек, закрепленных на стенках вены свободными концами, а на вершине — собранных вместе и образующих своеобразную сетку, фильтрующую кровоток и захватывающую мигрирующий тромб с целью недопущения его в легочную артерию. К настоящему времени разработано более 100 различных моделей кава-фильтров, отличающихся по своей технической конструкции, эффективности и количеству осложнений [8, 4].

Создание и оценка качества интравенозных фильтров является весьма сложной комплексной проблемой, при решении которой отыскивается компромисс между качеством улавливания тромбов вновь создаваемой конструкцией и побочным влиянием фильтра, как инородного тела, на окружающие ткани и органы. Противоречие состоит в том, что чем лучше решается задача тромбоулавливания, тем более фильтр влияет как на кровоток, так и на окружающее его пространство [2]. Разрешение этого противоречия связано с комплексом проблем гидродинамики вязких жидкостей, знанием анатомических особенностей того участка венозного русла, где предполагается имплантация фильтра, и наличием качественно новых и оригинальных идей, реализуемых в новой конструкции. Существующие сегодня способы решения задач являются полуэмпирическими и требуют большого количества времени и финансовых затрат на реализацию экспериментов in vitro и in vivo.

Использование численных методов при разработке и анализе качества новых устройств является современным способом решения задач, базирующимся на использовании мощных современных алгоритмов и вычислительных средств. Подход базируется на создании формальных описаний:

- вновь создаваемой и/или исследуемой конструкции;
- тромбов как инородных тел разной геометрической формы и размеров;
- венозного кровотока.

Формализация вышеперечисленных аспектов возможна при определенных допущениях, адекватность которых в значительной мере определяет обоснованность выбранной математической модели, а следовательно, и результата численного моделирования. Адекватность математической модели формирует основу для создания механизма оценивания качества новых устройств без проведения экспериментальных работ, что позволит быстро и достоверно:

- оценить целесообразность реализации новых конструкций;
- рассчитать качество работы новых устройств;
- проводить сравнительный анализ результатов для разных модификаций элементов конструкций.

Целью расчетов на первом этапе моделирования является выявление общей картины установившегося течения, его стационарность или квазипериодичность, а также вычисление характеристик установившегося потока крови (скорости, давления) в заданных точках.

3. ГЕОМЕТРИЧЕСКАЯ МОДЕЛЬ

Рассматривается течение крови в венозном русле на промежутке в 12 см. Сосуд (область $\Omega \in \mathbb{R}^3$) представляет собой основную вену эллиптического сечения $(1.6 \times 2.4 \,\mathrm{cm})$ с возможными ответвлениями - малыми венами эллиптического сечения $(0.533 \times 0.8 \,\mathrm{cm})$. В работе рассматривается простейший случай основной вены без ответвлений. Поле скорости на концах всех вен стационарно и плоскопараллельно, причем осевая компонента вектора скорости имеет профиль Пуазейля с максимальным значением $8.8 \,\mathrm{cm/c}$ на сечениях втока. Кава-фильтр представляет собой систему $12 \,\mathrm{пр}$ ямых проволок толщины $0.5 \,\mathrm{mm}$, имеющих общую концевую точку и упирающихся другими концами в стенки основной вены. Длина

фильтра (вдоль оси основной вены) равна 2 см. Рассматриваются шарообразные тромбы различных диаметров с центром на оси основной вены, находящиеся внутри фильтра и касающиеся двух ножек фильтра (рис.1).

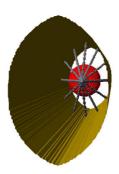


Рис. 1.

4. Гидродинамическая модель

При моделировании течения крови были сделаны следующие предположения. Венозная кровь считается несжимаемой ньютоновской жидкостью с постоянной плотностью $\rho=1.5$ г/куб.см и постоянной динамической вязкостью $\mu=0.0055$ Па·с. При этом кинематическая вязкость равна $\nu=\mu/\rho=0.036$ кв.см/с. Стенки вен, кава-фильтр и тромб неподвижны, причем кровь на их поверхностях также неподвижна (условие прилипания). Система уравнений, описывающих течение жидкости при таких предположениях, называется системой трехмерных уравнений Навье-Стокса. Обозначим через $\mathbf{u}(x,t)$ скорость кровотока с компонентами (u_1,u_2,u_3) , а через $p(x,t)=P(x,t)/\rho$ кинематическое давление жидкости. Значения скорости \mathbf{g} являются заданными стационарными величинами на границе области $\partial\Omega$. Уравнения Навье-Стокса выписываются следующим образом:

$$\frac{\partial \mathbf{u}}{\partial t} - \nu \Delta \mathbf{u} + (\mathbf{u} \cdot \nabla) \mathbf{u} + \nabla p = f \mathbf{B} \Omega; \tag{1}$$

$$\mathbf{u} = \mathbf{g}$$
 на $\partial \Omega$,

$$div \mathbf{u} = 0 \mathbf{B} \Omega. \tag{2}$$

Важным параметром подобия течений является число Рейнольдса $Re=\frac{\bar{U}\bar{D}}{\nu}$, где \bar{U} и \bar{D} — характерные скорости и размер препятствия. В данной модели $\bar{U}\simeq 5$ см, $\bar{D}\simeq 1$ см, поэтому $Re\simeq 130$. При подобном числе Рейнольдса даже незначительная асимметрия потока может привести к квазипериодичности установившегося течения, поэтому выяснение временного характера установившегося течения возможно только в рамках нестационарных уравнений.

5. Ограничения на сетку, выбор аппроксимации и решатели сеточных систем

Специфичность и сложность геометрии обтекаемого препятствия накладывают серьезные ограничения на размер сетки в окрестности препятствия. Даже в случае схемы аппроксимации второго порядка точности локальный размер сетки не может превышать 0.2 мм, в противном случае сетка просто не разрешит элементы конструкции кава-фильтра. Другой проблемой является сложность геометрии препятствия в виде кава-фильтра с захваченным тромбом, а также большие размеры области наибольшего влияния на кровоток: $3 \times 2.4 \times 1.6$ см. Построение блочноструктурированной локально сгущающейся сетки исключается самой постановкой проблемы, поскольку необходимо тестирование кава-фильтров различных конструкций с разными видами тромбов в кровеносных сосудах различных форм, а генерация одной блочно-структурированной сетки — очень трудоемкая операция, которая должна выполняться заново для каждой новой геометрии. Построение неструктурированной локально сгущающейся сетки (если это выполнимо) приведет к миллионам ячеек сетки для аппроксимации только давления. Известно, что аппроксимация скоростей потребует значительно больше степеней свободы, что породит неструктурированные сеточные уравнения с десятками миллионов неизвестных. Практическая разрешимость подобных уравнений неочевидна. Кроме того, существует проблема надежной генерации неструктурированных сеток с ячейками правильной формы для различных геометрий. Учитывая вышесказанное, нами была выбрана альтернативная стратегия: использование равномерных прямоугольных сеток в зоне наибольшего влияния, плавно разгрубляющихся в остальной части расчетной области, а также сеточных шаблонов, обеспечивающих второй порядок аппроксимации границы области и эффективное решение сеточных систем. При этом расчетная сетка содержит от 1.5 до 2 миллионов узлов для аппроксимации как давления, так и скорости. Поскольку количество неизвестных в сеточных уравнениях весьма велико, необходимо использовать параллельные компьютеры и эффективные численные методы, использующие прямоугольную структуру сетки.

6. Аппроксимация по времени

Для аппроксимации по времени используем проекционный алгоритм, хорошо себя зарекомендовавший при относительно больших числах Рейнольдса.

ullet Шаг 1: Предварительная скорость $\hat{f u}^{k+1}$ вычисляется из уравнения

$$\frac{\hat{\mathbf{u}}^{k+1} - \mathbf{u}^k}{\Delta t} - \nu \Delta \frac{\hat{\mathbf{u}}^{k+1} + \mathbf{u}^k}{2} + \left(\left(\frac{3}{2} \mathbf{u}^k - \frac{1}{2} \mathbf{u}^{k-1} \right) \cdot \nabla \right) \frac{\hat{\mathbf{u}}^{k+1} + \mathbf{u}^k}{2}$$

$$= \mathbf{f}^{k+1/2} - \nabla p^k \, \mathbf{B} \, \Omega, \tag{3}$$

$$\hat{\mathbf{u}}^{k+1} = \mathbf{g} \text{ Ha } \partial\Omega. \tag{4}$$

• Шаг 2: Проекция предварительной скорости в пространство бездивергентных вектор-функций происходит за счет решения уравнения

$$-\operatorname{div}\nabla\delta p = -\frac{1}{\Delta t}\operatorname{div}\hat{\mathbf{u}}^{k+1},\tag{5}$$

$$\mathbf{u}^{k+1} = \hat{\mathbf{u}}^{k+1} - \Delta t \nabla \delta p, \quad p^{k+1} = p^k + \delta p.$$
 (6)

Отметим два важных свойства этой схемы. Во-первых, уравнение (3) представляет собой схему Кранка-Николсон с полунеявной аппроксимацией конвективной части со вторым порядком по времени. Во-вторых, в уравнении для поправки для давления нет краевого условия. Оно присутствует через операторы div и ∇ , функционально определяемые через пространства для скоростей и давления. На дискретном уровне матрица уравнения (5) получается призведением матриц сеточной дивергенции и градиента, а

не аппроксимацией уравнения (5). В рамках данной стратегии необходимо аппроксимировать только (3) как краевую задачу для $\hat{\mathbf{u}}^{k+1}$ с правой частью $\mathbf{f}^{k+1/2} - \nabla p^k$ плюс члены, зависящие от \mathbf{u}^k :

$$-\frac{\nu}{2}\Delta\hat{\mathbf{u}}^{k+1} + \left(\left(\frac{3}{2}\mathbf{u}^k - \frac{1}{2}\mathbf{u}^{k-1}\right)\cdot\nabla\right)\frac{\hat{\mathbf{u}}^{k+1}}{2} + \frac{\hat{\mathbf{u}}^{k+1}}{\Delta t} = rhs^{k,k+1/2} \text{ в }\Omega,$$

$$\hat{\mathbf{u}}^{k+1} = \mathbf{g} \text{ на }\partial\Omega.$$
(7)

7. Аппроксимация по пространству

Уравнения (7) для компонент скорости u_i , i=1,2,3, аппроксимируются на прямоугольных разнесенных сетках Ω_i^h независимо друг от друга. Рассмотрим уравнение (7) для некоторой компоненты i, которое есть уравнение конвекции-реакции-диффузии. Множество сеточных узлов Ω_i^h разбивается на внутренние (принадлежащие Ω) и внешние узлы. Внутреннему узлу приписывается тип 1, если в его разностный шаблон входят только внутренние узлы, или тип 2 – в противном случае (рис.2; здесь и далее, для простоты изображен двумерный случай).

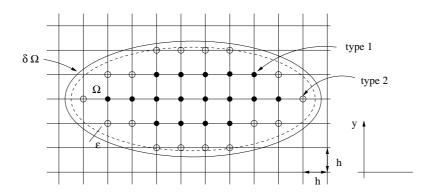


Рис. 2.

Оператор Лапласа аппроксимируется стандартным 7-точечным шаблоном

$$-\Delta u \approx \frac{(u_r - u_0)/h_r - (u_0 - u_l)/h_l}{(h_r + h_l)/2} + \frac{(u_u - u_0)/h_u - (u_0 - u_d)/h_d}{(h_u + h_d)/2}$$

$$+ \frac{(u_f - u_0)/h_f - (u_0 - u_b)/h_b}{(h_f + h_b)/2}$$

со вторым порядком точности в узлах типа 1, если шаги сетки меняются плавно, а оператор конвекции — линейной комбинацией центральной

$$(\mathbf{b}\nabla)u \approx b_1 \frac{u_r - u_l}{h_r + h_l} + b_2 \frac{u_u - u_d}{h_u + h_d} + b_3 \frac{u_f - u_b}{h_f + h_b},$$

и односторонней (вверх по потоку) конечных разностей

$$(\mathbf{b}\nabla)u \approx \max(b_1, 0)\frac{u_0 - u_l}{h_l} + \min(b_1, 0)\frac{u_r - u_0}{h_r} + \max(b_2, 0)\frac{u_0 - u_d}{h_d} + \min(b_2, 0)\frac{u_u - u_0}{h_u} + \max(b_3, 0)\frac{u_0 - u_b}{h_b} + \min(b_3, 0)\frac{u_f - u_0}{h_f},$$

причем односторонние разности используются только вдали от обтекаемого препятствия, где решение исходного уравнения - достаточно гладкая функция. Отметим, что соотношение шага расчетной сетки (~ 0.02) в районе обтекаемого препятствия и значения числа Рейнольдса позволяют обойтись лишь центральными разностями для обеспечения устойчивости аппроксимации. В тех зонах, где шаг сетки по какому-нибудь направлению достаточно большой $(\sim 0.3-0.5)$, односторонние разности не портят аппроксимацию в силу высокой гладкости решения. Для узлов типа 2 сеточная функция определена не во всех точках шаблона, поэтому предлагается заменить внешнюю точку шаблона на ее проекцию на границу (рис.3), где известно граничное значение. Для уравнения Пуассона второй порядок аппроксимации может быть показан и в этом случае. На практике вводятся дополнительные узлы типа 3 как координатные проекции внешних узлов на границу $\partial\Omega$, которые вносят вклад в шаблоны узлов типа 2, но в которых уравнение (7) не аппроксимируется. Из соображений устойчивости тем узлам типа 2, которые находятся в узкой полосе (по отношению к локальному шагу сетки) вдоль границы, также приписывается тип 3. Разнесенные сетки Ω_i^h , i=1,2,3, строятся с помощью сдвижки узлов некоторой прямоугольной сетки Ω^h_0 на полшага вдоль соответствующей координатной оси. На сетке Ω^h_0 определяется сеточная функция давления, а также множество значений сеточной

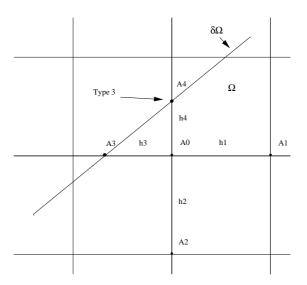


Рис. 3.

дивергенции. Оператор дивергенции аппроксимируется центральными разностями со вторым порядком точности на сетках с плавно меняющимся шагом, если все участвующие в шаблоне узлы являются узлами типа 1 или 2 сеток Ω_i^h , i=1,2,3. В противном случае используется трехточечный (по направлению) шаблон аппроксимации первой частной производной, также обеспечивающий второй порядок точности. В некоторых узлах сетки Ω_0^h не все компоненты дивергенции могут быть аппроксимированы конечными разностями, поскольку узлы, входящие в шаблон сеточной дивергенции, могут оказаться внешними для какой-нибудь компонентной сетки Ω_i^h . В этом случае недостающая частная производная заменяется приближенным диффиренцированием заданного граничного значения. В силу предполагаемой гладкости граничного условия подобная замена исказит аппроксимацию оператора дивергенции очень незначительно. Оператор градиента для давления аппроксимируется центральными разностями на сетке Ω_0^h в предположении, что давление ассоциировано с узлами сетки Ω^h_0 . Прозведение $div_0\nabla_0$ определенных таким образом сеточных операторов дивергенции div_0 и градиента ∇_0 не может быть симметричным оператором в случае произвольных областей. Дискретная схема проекционного алгоритма не рассматривает аппроксимацию уравнения (5) с некоторыми краевыми условиями, а является прямым аналогом проекционного алгоритма в сеточном случае. Отметим, что для разнесенных сеток выполняется так называемое LBB условие:

$$\sup_{\mathbf{u}\perp Ker\Delta} \frac{(p, div\,\mathbf{u})}{\|\mathbf{u}\|_{H^1}} \geqslant C_0 \|p\|_{L_2}, \quad \forall p \perp Ker\nabla,$$

с константой C_0 , не зависящей от числа степеней свободы [9].

Обозначив конечно-разностные дискретизации оператора Лапласа и конвекции через Δ_i и ($\mathbf{b}\nabla_i$), выпишем дискретный аналог проекционного алгоритма.

• Шаг 1: Предсказать сеточные компоненты скорости \hat{u}_i^{k+1} за счет решения уравнения импульса

$$\frac{\hat{u}_{i}^{k+1} - u_{i}^{k}}{\Delta t} - \nu \Delta \frac{\hat{u}_{i}^{k+1} + u_{i}^{k}}{2} + \left(\prod_{i} \left(\frac{3}{2} \mathbf{u}^{k} - \frac{1}{2} \mathbf{u}^{k-1} \right) \cdot \nabla_{i} \right) \frac{\hat{u}_{i}^{k+1} + u_{i}^{k}}{2} \\
= f_{i}^{k+1/2} - (\nabla_{0} p^{k})_{i} \mathbf{B} \Omega, \tag{8}$$

$$\hat{u}_i^{k+1}=g_i$$
 на $\partial\Omega,\quad i=1,2,3.$

• Шаг 2: Найти поправку для давления и спроектировать предсказанную скорость

$$-div_0\nabla_0\delta p = -\frac{1}{\Delta t}div_0\,\hat{\mathbf{u}}^{k+1},\,\,\mathbf{u}^{k+1} = \hat{\mathbf{u}}^{k+1} - \Delta t\nabla_0\delta p,\,\,p^{k+1} = p^k + \delta p.$$
(9)

8. Параллельные методы решения линейных систем

Дискретный проекционный алгоритм (8)-(9) требует решения четырех линейных на каждом шаге по времени. Первые три системы (для каждой компоненты скорости) порождаются уравнением импульса и являются сингулярно возмущенными уравнениями конвекции-реакции-диффузии с малым коэффициентом при операторе диффузии (ν) и относительно большим коэффициентом при единичном операторе ($1/\Delta t$). Кроме того, эти системы являются независимыми друг от друга и могут разрешаться на разных

процессорах. Четвертая система является аналогом некоторой сеточной аппроксимации уравнения Пуассона в расчетной области и не главными краевыми условиями. Для решения систем с разными свойствами применяются методы, наиболее полно использующие эти свойства, а также тот факт, что аппроксимация осуществлялась на прямоугольных сетках.

8.1. Параллельный метод Шварца с минимальными перекрытиями между подобластями. Для решения сингулярно возмущенного уравнения конвекции-реакции-диффузии применяется параллельный двухуровневый итерационный метод Шварца [6]. На внешнем уровне итераций расчетная область разбивается на параллельные слабо налегающие подобласти-слои Ω_m , ортогональные доминирующему направлению течения в канале (сосуде). Классический метод Шварца альтернирования по подобластям в направлении вниз по потоку определяет внешние итерации. Скорости сходимости подобных итераций очень высока для уравнения конвекции-диффузии, если задачи в подобластях решаются точно, и на границе между подобластями нет противотока или больших зон застоя. В случае наличия сингулярно возмущенного единичного оператора даже эти ограничения существенно ослаблены [5]. Однако классический одноуровневый метод Шварца обладает двумя существенными недостатками: вопервых, он очень плохо параллелизуется и, во-вторых, он требует точного решения сеточных задач в подобластях, что в трехмерном случае является весьма трудоемкой операцией. Поэтому в двухуровневом методе Шварца точное решение задач в подобластях заменяется на параллельное итерационное решение. Каждая подобласть-слой Ω_m разбивается на слабо налегающие подобластибруски $\Omega_{m,l}$ вдоль одной из координатных осей, ортогональных доминирующему направлению течения в канале (рис.4). Внутренние итерации представляют собой несколько итераций метода Шварца с альтернированием между подобластями-брусками с четными и нечетными индексами l. Параллелизм алгоритма заключается в несвязности подобластей $\Omega_{m,l}$ с (не)четными индексами l между собой: каждую подобласть-брусок может обрабатывать свой процессор. Дополнительным преимуществом разбиения второго уровня оказывается ленточная структура матрицы сеточного оператора в $\Omega_{m,l}$, для которой можно применять эффективные методы факторизации и оценки арифметической работы. Интересно отметить, что скорость сходимости итераций не зависит от числа подобластей ни на внешнем, ни на внутреннем уровне, поэтому арифметическая сложность одной внешней итерации пропорциональна числу неизвестных при фиксированной сеточной толщине подобластей-брусков. Общая эффективность двухуровневого метода определяется малыми перекрытиями (в один шаг сетки) между подобластями и высокой скоростью сходимости как внешних, так и внутренних итераций. На практике число внутренних итераций не превышает 2, а число внешних итераций не превышает 3-5. Отметим, что хорошие свойства алгоритма обусловлены сингулярной возмущенностью исходного оператора [6].



8.2. МЕТОД ФИКТИВНЫХ КОМПОНЕНТ ДЛЯ НАХОЖДЕНИЯ ПОПРАВКИ К ДАВЛЕНИЮ. Несмотря на то, что матрица системы (9) вырождена, система является совместной в силу алгебраического происхождения матрицы $div_0\nabla_0$. Пусть B – невырожденный переобуславливатель для $div_0\nabla_0$ в подпространстве векторов, ортогональных ядру $div_0\nabla_0$. Тогда система $-div_0\nabla_0 x = y$ может быть преобразована к системе

$$Ax = b$$

с матрицей $A=-B^{-1}div_0\nabla_0$ и правой частью $b=B^{-1}y$. Для решения последней системы можно применить любой итерационный процесс на подпространствах Крылова, требующий умноже-

ния вектора на матрицу A, т.е. последовательного умножения вектора на $div_0\nabla_0$ и решения системы с B. В данном приложении мы применяем обобщенный метод сопряженных невязок:

- 1. Найти $r_0 = b Ax_0$. Положить $p_0 = r_0$.
- 2. Для $j=0,1,2,\ldots$, выполнять, пока не сработает критерий остановки:

3.
$$\alpha_i = \frac{(r_j, Ap_j)}{(Ap_j, Ap_j)}.$$

- $4. x_{j+1} = x_j + \alpha_j p_j.$
- $5. r_{j+1} = r_j \alpha_j A p_j.$

6. Найти
$$\beta_{ij} = -\frac{(Ar_{j+1}, Ap_i)}{(Ap_i, Ap_i)}$$
, for $i = 0, 1, \dots, j$.

7.
$$p_{j+1} = r_{j+1} + \sum_{i=0}^{j} \beta_{ij} p_i$$
.

8. Конец цикла.

Переобуславливатель метода фиктивных компонент [10] формируется следующим образом. Пусть n_1 – количество узлов сетки Ω_0^h , обеспечивающих степени свободы для давления, и они пронумерованы первыми, а n_2 – количество оставшихся узлов. Определим переобуславливатель $B^{-1} = TC^{-1}T^T$, где $T = [I \ O]$, $I \in \mathbb{R}^{n_1 \times n_1}$, $T \in \mathbb{R}^{n_1 \times (n_1 + n_2)}$, $C \in \mathbb{R}^{(n_1 + n_2) \times (n_1 + n_2)}$. Матрица C – легко обратимая матрица тензорного вида:

$$C = (M_1 \otimes M_2 \otimes M_3)^{-1} (A_1 \otimes M_2 \otimes M_3 + M_1 \otimes A_2 \otimes M_3 + M_1 \otimes M_2 \otimes A_3),$$

где $A_i,\ M_i$ — трехдиагональная и диагональная матрицы соответственно, получаемые из аппроксимации второй и нулевой производных на одномерных следах сетки Ω_0^h . Можно показать, что матрица C эквивалентна по спектру конечно-разностному оператору Лапласа с однородными неймановскими краевыми условиями [10]. Учитывая свойства матрицы $-div_0\nabla_0$, можно также ожидать хорошей обусловленности матрицы A. Линейные системы с матрицей C могут эффективно решаться с помощью параллельного решателя [11]. Дополнительным преимуществом использования

обобщенного метода сопряженных невязок является возможность уменьшения числа итераций не только за счет высокой скорости сходимости, но и за счет выбора лучшего начального приближения, если накапливать крыловские векторы p_i и Ap_i . Фактически мы проектируем $x_0 = 0$ на накопленное крыловское пространство $\{p_i\}_{i=1}^k$ [12]:

$$x_0 := \sum_{i=1}^k \frac{(b, Ap_i)}{(Ap_i, Ap_i)} p_i,$$

то есть обращаем матрицу системы, спроектированной на крыловское пространство. Процедура подобного проектирования стоит k скалярных произведений и не требует никаких матричновекторных произведений. Когда течение установилось, данный подход позволяет уменьшить количество итераций в несколькораз.

8.3. Особенности параллелизации алгоритмов. Характерной особенностью используемых параллельных вычислений является их кластеризация. Множество всех доступных процессоров разбивается на три равные группы — кластеры для компонент скорости. Каждый кластер решает свою сеточную задачу (8) независимо от других, причем на каждом кластере применяется параллельная реализация двухуровневого метода Шварца. Кроме этого, из общего множества процессоров выделяется подмножество, на котором можно использовать параллельный метод фиктивных компонент (9). В идеальном случае это подмножество совпадает с объединением кластеров компонент скорости, поскольку решение задачи (9) следует за решением задач (8). Однако используемая версия быстрого параллельного сепарабельного решателя накладывает сильные ограничения на количество процессоров (2^n) , на которых может решаться задача (9). Поэтому кластер для давления имеет максимально возможное число процессоров вида 2^n . Кластеры для компонент скорости и кластер для давления обмениваются данными между решением дискретных задач (8) и (9).

Двухуровневый метод Шварца параллелизуется следующим образом. Пусть в кластере для компоненты скорости NPr_i процессоров. Процессор с индексом l обрабатывает данные из подобластей $\bigcup_{m} (\Omega_{m,2l-1} \cup \Omega_{m,2l}), \ l=1,\ldots,NPr_i$, т.е. обрабатывает слой

расчетной области, параллельный доминирующему направлению потока. При этом он обменивается данными только с процессорами с индексами l-1 и l+1 (рис.4). Данные в кластере для давления также разбиты на слои вдоль по потоку, но толщина слоя определяется числом процессоров в кластере для давления. Кроме того, каждый процессор обменивается не только с соседями [11], потому использование быстрого сепарабельного решателя требует больших коммуникационных затрат по сравнению с двухуровневым методом Шварца.

9. ПРИМЕР ПАРАЛЛЕЛЬНОГО РАСЧЕТА

Поскольку размеры кава-фильтра накладывают сильные ограничения на шаг сетки снизу, особенности геометрии не позволяют сравнить результаты на нескольких последовательно измельчающихся сетках. Кроме того, минимально необходимая для расчета сетка может быть обработана не меньше чем на 36 процессорах в силу ограничений на память. Поэтому не представляется возможным сравнительный анализ поведения рассматриваемого параллельного метода для рассматриваемой конфигурации препятствия с точки зрения как арифметической, так и параллельной масштабируемости. Подобный анализ был приведен для более простого цилиндрического препятствия в работе [7].

В настоящей работе мы приведем результаты расчета обтекания кава-фильтра с захваченным тромбом диаметра $1.2~{\rm cm}$, выполненные на сетке $160\times120\times80$, в которой около $1.5~{\rm muллионa}$ расчетных узлов для каждой из компонент скорости и давления. Таким образом, на каждом шаге по времени обрабатывается около 6 миллионов неизвестных. В той части вены, где расположен кава-фильтр с тромбом, шаг сетки по каждому направлению равен $0.2~{\rm km}$, что достаточно для геометрического разрешения ножек фильтра. Минимальное количество временных шагов, необходимых для отыскания установившегося решения, оказывается равным $200~{\rm (при}~ucxодных~данных)}$. Уменьшение количества шагов не представляется возможным в силу ограничения на временной шаг из-за численной неустойчивости применяемого алгоритма. Таким образом, один расчет требует решения $800~{\rm nuhe\"{u}}$ ных систем с $1.5~{\rm mun}$ лиона неизвестных.

Численный расчет проводился на 36-процессорном кластере

Таблица 1

n	1	4	15	20	108	121	147	150	$\begin{array}{c} 200 \\ 221 \end{array}$
K	0	171	0	125	0	115	0	35	221
it	60	47	30	12	22	6	17	6	4

BiAthlon (1.6 GH). Критерий остановки итераций – евклидова норма невязки – не превышает 0.01 как для уравнения (8), так и для уравнения (9). При этом количество итераций двухуровневого метода Шварца на каждом шаге по времени равнялось 3 или 4, а количество итераций метода фиктивных компонент зависело от номера временного шага и накопленной информации в пространстве крыловских векторов: чем богаче пространство, тем ближе начальное приближение к решению. Для выбора начального приближения в методе фиктивных компонент использовалось крыловское пространство размерностью не более 250 векторов, причем при переполнении заданного объема крыловское пространство обнулялось и создавалось заново. В табл.1 приведено количество итераций метода фиктивных компонент *it* в зависимости от шага по времени n и размера крыловского пространства K, используемого при построении начального проближения. Если накопленное крыловское пространство пусто, то нулевое начальное приближение становится все более близким к текущей поправке для давления, при стремлении течения к стационарному, поэтому на шагах $n=1,\ 15,\ 108,\ 147$ мы наблюдаем последовательное уменьшение числа итераций. Если накопленное крыловское пространство не пусто, эффективность проекции начального приближения зависит от того, насколько хорошо приближает пространство искомую поправку. На начальном этапе, когда решение не установилось ни в коей мере, даже богатое (больше 100 векторов) пространство уменьшает число итераций всего на 30 %, в то время как проекции для более гладких поправок к давлению на последующих шагах уменьшают количество итераций в 3-4 раза, по отношению к нулевому начальному приближению. Более того, в окрестности установившегося решения достаточно иметь всего несколько десятков векторов в крыловском пространстве, чтобы иметь такое же радикальное уменьшение. Общее время решения задачи равнялось 3500 с, из которых 2890 с были потрачены на решение 200 линейных систем (9), и 220 с на решение 200 линейных систем (8) для каждой компоненты скорости. Столь значительная разница во времени обусловлена тем, что среднее количество итераций при решении (9) и (8) равно 11 и 3.5 соответственно.

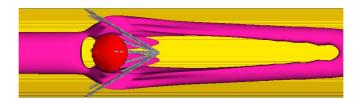


Рис. 5.

На рис.5 приведена зона, где осевая компонента скорости больше 7 см/с. Отметим локальное влияние ножек кава-фильтра на течение крови. Изолинии осевой компоненты скорости и давления в плоскости z=0.8 приведены на рис.7. Наконец, векторное поле скоростей представлено на рис.6, где сечения проходят через центр масс тромба.

Заключение

В работе рассмотрены основные аспекты численного моделирования течения крови в вене с имплантированным кава-фильтром с захваченным тромбом. Сформулированы допущения, в рамках которых рассматриваемая расчетная модель может считаться адекватной. Особое внимание уделено вопросам параллельной компьютерной реализации модели течения, включающим как методы дискретизации, так и методы решения возникающих систем. Геометрические особенности модели требуют большого числа

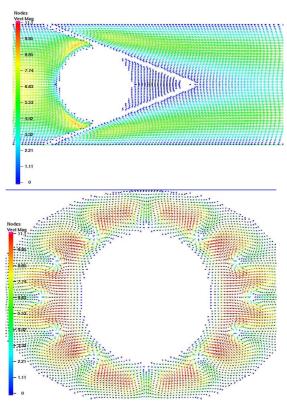


Рис. 6.

расчетных узлов, поэтому решение возникающих систем возможно только благодаря современным эффективным численным методам. В дальнейшем мы планируем применить модель для расчета системы вен и интравенозных фильтров и тромбов разных форм и размеров. Рассматривается также обобщение модели на геометрии области, взятые из индивидуальных особенностей пациента. В перспективе возможен переход к более сложным моделям реологии крови, а также учет взаимодействия упругого имплантанта (стента) и течения крови.

Работа финансировалась за счет гранта РФФИ 04-07-90336-в, госпрограммы РАН "Вычислительные и информационные проблемы решения больших задач", а также благодаря спонсорству компании Комед. Авторы признательны C.C.Ломкову и A.H.Балану за постановку задачи, полезные дискусии и поддержку.

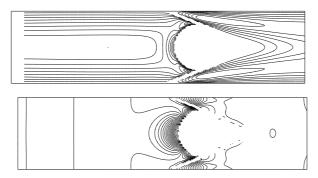


Рис. 7.

Литература

- [1] Прокубовский В.И., Капранов С.А., Эндоваскулярная профилактика тромбоэмболии легочной артерии. // Флебология /Под ред. В.С.Савельева. М.: Медицина, 2001, с.359-375.
- [2] Прокубовский В.И., Капранов С.А., Москаленко Е.П., Анатомические и гемодинамические изменения нижней полой вены при профилактике тромбоэмболии легочной артерии.// Ангиология и сосудистая хирургия, Vol.9, N.2, 2003, с. 51-60.

- [3] Савельев В.С., Прокубовский В.И., Капранов С.А., Эндоваскулярная хирургия в профилактике тромбоэмболии легочной артерии и лечении острых венозных тромбозов. // Хирургия, N.2, 2003, c. 6-11.
- [4] Asch M.R., Initial experience in humans with a new retrievable inferior vena cava filter. // Radiology, Vol.225, 2002, p.835-844.
- [5] Garbey M., Kuznetsov Yu., Vassilevski Yu., Parallel Schwarz Method for Convection-Diffusion Constituent of the Navier-Stokes Solver. // Proceedings of ParCFD-98 conference, Taiwan, 1998.
- [6] Garbey M., Kuznetsov Yu., Vassilevski Yu., Parallel Schwarz method for a convection-diffusion problem. // SIAM J.Sci.Comp., Vol.22, 2000, p.891-916.
- [7] Garbey M., Vassilevski Yu., A parallel solver for unsteady incompressible 3D Navier-Stokes equations. // Parallel Computing, Vol.27, 2001, p.363-389.
- [8] Ichikura K., Yamada N., Oota M. et al. Clinical experience retrievable vena cava filters for prevention of pulmonary thromboembolism. // J. Cardiol., Vol.40, 2002, p.267-273.
- [9] Kobel'kov G. and V. Valedinskiy V., On the inequality $||p||_{L_2} \leq c||\nabla p||_{W_2^{-1}}$ and its finite dimensional image. // Sov.J.Numer.Anal.Math.Modelling, Vol.1, 1986, p.189-200.
- [10] Marchuk G., Kuznetsov Yu., Matsokin A., Fictitious Domain and Domain Decomposition Methods. // Sov.J.Numer.Anal.Math.Modelling, Vol.1, 1986, p.3-35,.
- [11] Rossi T., Toivanen J., A Parallel Fast Direct Solver for the Discrete Solver for Block Tridiagonal Systems with Separable Matrices of Arbitrary Dimension. // SIAM J.Sci. Comp., Vol.20, 1999, p.1778-1793.
- [12] Tromeur-Dervout D., Résolution des Equations de Navier-Stokes en Formulation Vitesse Tourbillon sur Systèmes multiprocesseurs à Mémoire Distribuée, // Thesis, Univ. Paris VI, 1993.

Параллельная реализация метода решения объемного интегрального уравнения электродинамики на основе многоуровневых теплицевых матриц

Д. В. Савостьянов

Рассмотрено распространение электромагнитной волны в неоднородной среде, содержащей идеально проводящую плоскость. В модели локально неоднородной среды задача сведена к объемному интегральному уравнению. На равномерных декартовых сетках с использованием базисных функций специального вида, методом Галеркина получена матрица, обладающая трехуровневой блочной структурой вида ТТТ + ТНТ. С учетом структуры полученной матрицы предложен параллельный алгоритм решения сформулированной задачи, использование которого позволило существенно повысить точность вычисления значений полей вблизи неоднородности. Показано, что применение параллельной версии алгоритма может быть использовано для решения обратной задачи, т.е. исследования структуры неоднородности.

1. Постановка задачи

Рассмотрим распространение электромагнитной волны в полупространстве, ограниченном плоскостью, которая представляет собой идеальный проводник. Будем считать, что зависимость полей от времени гармоническая $e^{i\omega t}$. Электрические параметры среды будем считать переменными, но ограничимся моделью ложально неоднородной среды, согласно которой среда представляет из себя однородное пространство в котором располагается конечных размеров неоднородность \mathcal{V} [1, 2]. Размеры неоднородности предполагаются столь большими, чтобы их увеличение не оказывало существенного влияния на поведение полей в области, где производятся измерения. Таким образом, тот факт, что неоднородное пространство сведено к неоднородности конечных размеров, не отражается на значениях измеряемых величин.

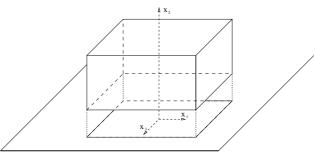


Рис. 1.

Мы ограничим рассмотрение случаем, когда область \mathcal{V} по форме является параллелепипедом. Введем прямоугольную систему координат так, чтобы идеальный проводник содержался в плоскости $x_2 = 0$, неоднородность располагалась в области $x_2 \geqslant 0$, а ось x_2 проходила через ее центр. При таком представлении неоднородности следует выбирать ее размеры так, чтобы значения электрического и магнитного полей на боковых и верхней стенке параллелепипеда достаточно сильно затухали в рамках выбранной точности. Геометрия задачи представлена на рисунке 1.

Будем считать, что сама неоднородность $\mathcal V$ представляет собой набор подобластей, в каждой из которых проводимость и диэлектрическая проницаемость среды постоянна (но различна в разных подобластях). Магнитная проницаемость внутри неоднородности совпадает с магнитной проницаемостью внешней среды. Внутри каждой подобласти справедлива система уравнений Максвелла, имеющая в гармоническом случае вид

$$rot \mathbf{H} = -i\omega \varepsilon \mathbf{E} + \mathbf{j}^{0},
rot \mathbf{E} = i\omega \mu \mathbf{H},$$
(1)

где комплекснозначная ε определена равенством

$$\varepsilon = \varepsilon_d - i\sigma/\omega$$
.

Источником электромагнитного поля является магнитный диполь, направленный вдоль оси x_2 и расположенный вне неоднородности в точке (0,h,0). Магнитное поле измеряется в нескольких точках в плоскости $x_2=h$. Типичная частота источника $100~\mathrm{M}\Gamma\mathrm{n}$, сопротивление внешней среды $10^3~\mathrm{Om}\cdot\mathrm{m}$.

2. Интегральное уравнение

Определим nepsuvhoe none $\mathbf{E}^0, \mathbf{H}^0$ равенством

$$rot \mathbf{H}^{0} = -i\omega\varepsilon_{0}\mathbf{E}^{0} + \mathbf{j}^{0},
rot \mathbf{E}^{0} = i\omega\mu\mathbf{H}^{0},$$
(2)

где ε_0 — диэлектрическая проницаемость внешней среды. Представим полное поле как сумму первичного и *аномального* поля

$$\mathbf{H} = \mathbf{H}^0 + \mathbf{H}^s, \mathbf{E} = \mathbf{E}^0 + \mathbf{E}^s.$$
 (3)

Для аномального поля \mathbf{E}^{s} , \mathbf{H}^{s} из (1), (2) и (3) имеем

$$\operatorname{rot} \mathbf{H}^{s} = -\mathrm{i}\omega\varepsilon_{0}\mathbf{E}^{s} + \mathrm{i}\omega(\varepsilon_{0} - \varepsilon)\mathbf{E},$$

$$\operatorname{rot} \mathbf{E}^{s} = \mathrm{i}\omega\mu\mathbf{H}^{s}.$$
(4)

Поскольку div $\mathbf{H}^s = 0$, то

$$\mathbf{H}^s = \operatorname{rot} \mathbf{A},$$

что вместе с предыдущим уравнением дает

$$\mathbf{E}^s = \mathrm{i}\omega\mu\mathbf{A} + \mathrm{grad}\,\Phi.$$

А и Φ называются векторным и скалярным потенциалом. Теперь первое уравнение из (4) переписывается в виде

$$\operatorname{rot}\operatorname{rot}\mathbf{A}=k_{0}^{2}\mathbf{A}-\mathrm{i}\omega\varepsilon_{0}\operatorname{grad}\Phi+\mathrm{i}\omega\varepsilon_{0}\left(1-\frac{\varepsilon}{\varepsilon_{0}}\right)\mathbf{E},$$

где $k_0 = \sqrt{\varepsilon_0 \mu \omega^2}$.

В декартовых координатах rot rot = $\operatorname{grad}\operatorname{div}-\Delta$. Учитывая это, получаем

$$\operatorname{grad}\left(\operatorname{div}\mathbf{A}+\mathrm{i}\omega\varepsilon_{0}\Phi\right)=\left(\Delta+k_{0}^{2}\right)\mathbf{A}+\mathrm{i}\omega\varepsilon_{0}\left(1-\frac{\varepsilon}{\varepsilon_{0}}\right)\mathbf{E}.$$

Учитывая условие калибровки (Лоренца)

$$\Phi = -\frac{1}{\mathrm{i}\omega\varepsilon_0}\operatorname{div}\mathbf{A},$$

имеем

$$\mathbf{E}^{s} = \frac{1}{-\mathrm{i}\omega\varepsilon_{0}} \left(k_{0}^{2} + \mathrm{grad}\,\mathrm{div}\right)\mathbf{A},$$

после чего векторный потенциал ${\bf A}$ должен удовлетворять векторному уравнению Лапласа

$$(\Delta + k_0^2) \mathbf{A} = i\omega \varepsilon_0 \mathbf{J}, \tag{5}$$

где ${f J}$ определяется равенством

$$\mathbf{J} = \gamma \mathbf{E}, \qquad \gamma = \frac{\varepsilon}{\varepsilon_0} - 1.$$

Отсюда заключаем

$$\mathbf{A}(x) = -\mathrm{i}\omega\varepsilon_0 \int_{\mathcal{V}} G(x, y) \mathbf{J}(y) \, dy,$$

где G(x,y) — тензорная функция Грина. Для аномальных полей таким образом справедливы формулы

$$\mathbf{E}^{s}(x) = \left(k_0^2 + \operatorname{grad}\operatorname{div}\right) \int_{\mathcal{V}} G(x, y) \mathbf{J}(y) \, dy, \tag{6}$$

$$\mathbf{H}^{s}(x) = -i\omega\varepsilon_{0} \operatorname{rot} \int_{\mathcal{V}} G(x, y) \mathbf{J}(y) \, dy. \tag{7}$$

Для определения тока $\mathbf{J} = \gamma(\mathbf{E}^0 + \mathbf{E}^s)$ перепишем уравнение (6) в виде объемного интегрального уравнения (см. [3])

$$\gamma^{-1}\mathbf{J}(x) - \left(k_0^2 + \operatorname{grad}\operatorname{div}\right) \int_{\mathcal{V}} G(x, y)\mathbf{J}(y) \, dy = \mathbf{E}^0, \qquad x \in \mathcal{V}. \tag{8}$$

 Φ ункция Грина для однородной среды без идеально проводящей плоскости имеет вид

$$G(x,y) = \begin{bmatrix} g(x,y) & & \\ & g(x,y) & \\ & & g(x,y) \end{bmatrix}, \qquad g(x,y) = \frac{e^{ik_0 ||x-y||}}{4\pi ||x-y||}. \quad (9)$$

В случае, когда в плоскости $x_2 = 0$ размещен идеальный проводник, функция Грина принимает вид

$$G(x,y) = \begin{bmatrix} g_1(x,y) & & & \\ & g_2(x,y) & & \\ & & g_3(x,y) \end{bmatrix},$$
(10)

$$g_1(x,y) = g_3(x,y) = \frac{e^{\mathrm{i}k_0\|x-y\|}}{4\pi\|x-y\|} - \frac{e^{\mathrm{i}k_0\|x-y^*\|}}{4\pi\|x-y^*\|},$$
$$g_2(x,y) = \frac{e^{\mathrm{i}k_0\|x-y\|}}{4\pi\|x-y\|} + \frac{e^{\mathrm{i}k_0\|x-y^*\|}}{4\pi\|x-y^*\|},$$

где точка y^* получается "отражением" точки y относительно идеально проводящей плоскости

$$y^* = (y_1, y_2, y_3)^* = (y_1, -y_2, y_3).$$

3. Метод дискретизации

Для дискретизации интегрального уравнения (8) применяется метод Галеркина. В качестве базисных функций используются конечные элементы, кусочно-линейные по одному из направлений и кусочно-постоянные по двум другим (см. [3]).

Введем на области $\mathcal{V} = [a_1, b_1] \times [a_2, b_2] \times [a_3, b_3]$ сетку, построенную декартовым произведением трех равномерных одномерных сеток:

$$\begin{aligned}
 x_1^{j_1} &= a_1 + (j_1 - 1)h_1, & h_1 &= (b_1 - a_1)/n_1, \\
 x_2^{j_2} &= a_2 + (j_2 - 1)h_2, & h_2 &= (b_2 - a_2)/n_2, \\
 x_3^{j_3} &= a_3 + (j_3 - 1)h_3, & h_3 &= (b_3 - a_3)/n_3.
 \end{aligned} \tag{11}$$

Неизвестную функцию ${\bf J}$ будем приближенно искать в виде следующей комбинации

$$\mathbf{J}(x_{1}, x_{2}, x_{3}) = \sum_{j_{1}=1}^{n_{1}-1} \sum_{j_{2}=1}^{n_{2}} \sum_{j_{3}=1}^{n_{3}} u_{j_{1}j_{2}j_{3}}^{1} \mathbf{F}_{j_{1}j_{2}j_{3}}^{1}(x_{1}, x_{2}, x_{3}) +$$

$$\sum_{j_{1}=1}^{n_{1}} \sum_{j_{2}=1}^{n_{2}-1} \sum_{j_{3}=1}^{n_{3}} u_{j_{1}j_{2}j_{3}}^{2} \mathbf{F}_{j_{1}j_{2}j_{3}}^{2}(x_{1}, x_{2}, x_{3}) +$$

$$\sum_{j_{1}=1}^{n_{1}} \sum_{j_{2}=1}^{n_{2}} \sum_{j_{3}=1}^{n_{3}-1} u_{j_{1}j_{2}j_{3}}^{3} \mathbf{F}_{j_{1}j_{2}j_{3}}^{3}(x_{1}, x_{2}, x_{3})$$

$$(12)$$

$$\mathbf{F}_{j_1j_2j_3}^1 = \left[egin{array}{c} \Phi_{j_1}^1 \Psi_{j_2}^2 \Psi_{j_3}^3 \ 0 \ 0 \end{array}
ight], \mathbf{F}_{j_1j_2j_3}^2 = \left[egin{array}{c} 0 \ \Psi_{j_1}^1 \Phi_{j_2}^2 \Psi_{j_3}^3 \ 0 \end{array}
ight], \mathbf{F}_{j_1j_2j_3}^3 = \left[egin{array}{c} 0 \ 0 \ \Psi_{j_1}^1 \Psi_{j_2}^2 \Phi_{j_2}^3 \end{array}
ight],$$

где $\Phi_{j_k}^k \ (j_k = 1, \dots, n_k - 1)$ — кусочно-линейные функции

$$\Phi_{j_k}^k(x_k) = \begin{cases} 1 - |x_k - x_k^{j_k+1}|/h_k, & |x_k - x_k^{j_k+1}| \leqslant h_k, \\ 0, & \text{при прочих } x_k, \end{cases}$$

а $\Psi_k^{j_k}$ $(j_k=1,\ldots,n_k)$ — кусочно-постоянные функции

$$\Psi_{j_k}^k(x_k) = \begin{cases} 1, & x_k^{j_k} \leqslant x_k \leqslant x_k^{j_k+1}, \\ 0, & \text{при прочих } x_k. \end{cases} \quad (k = 1, 2, 3)$$

Неизвестные коэффициенты из разложения (12) определяются, согласно методу Галеркина, из следующей системы линейных уравнений

$$\int_{\mathcal{V}} \left(\gamma^{-1} \mathbf{J}(x), \mathbf{F}_{i_1 i_2 i_3}^k(x) \right) dx - \int_{\mathcal{V}} \left((k_0^2 + \operatorname{grad div}) \int_{\mathcal{V}} G(x, y) \mathbf{J}(y) dy, \mathbf{F}_{i_1 i_2 i_3}^k(x) \right) dx = \int_{\mathcal{V}} \left(\mathbf{E}^0(x), \mathbf{F}_{i_1 i_2 i_3}^k(x) \right) dx. \tag{13}$$

Полученная линейная система имеет вид

$$\left(\begin{bmatrix} D_1 & & & \\ & D_2 & & \\ & & D_3 \end{bmatrix} + \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix} \right) \begin{bmatrix} u^1 \\ u^2 \\ u^3 \end{bmatrix} = \begin{bmatrix} f^1 \\ f^2 \\ f^3 \end{bmatrix}. (14)$$

Блочно-диагональная часть матрицы соответствует вне-интегральному слагаемому из (13) и определяется равенством

$$\left(D_k\right)_{i_1i_2i_3}^{j_1j_2j_3} = \int_{\mathcal{V}} \gamma^{-1} \left(\mathbf{F}_{i_1i_2i_3}^k(x), \mathbf{F}_{j_1j_2j_3}^k(x)\right) dx. \tag{15}$$

Плотная часть матрицы, отвечающая интегральному слагаемому из (13), определяется формулой

$$\left(A_{kl}\right)_{i_{1}i_{2}i_{3}}^{j_{1}j_{2}j_{3}} = -\int_{\mathcal{V}} \left(\left(k_{0}^{2} + \operatorname{grad div}\right) \int_{\mathcal{V}} G(x, y) \mathbf{F}_{j_{1}j_{2}j_{3}}^{l}(y) dy\right),
\mathbf{F}_{j_{1}j_{2}j_{3}}^{k}(x)\right) dx.$$
(16)

4. Специфика полученной матрицы

Именно действия с плотной частью матрицы требуют основных затрат для вычисления элементов и проведения операции умножения матрицы на вектор. Поэтому вопрос о ее эффективном хранении и умножении представляет наибольший практический интерес. Интегральная часть уравнения (13), представленная матрицей (16), обладает определенной спецификой, учитывая которую, мы можем значительно снизить затраты на вычисление, хранение и умножение на эту матрицу.

Рассмотрим сначала случай, когда сверхпроводник в плоскости $x_2=0$ отсутствует. Тогда функция Грина имеет вид (9). Отметим, что в этом случае G(x,y)=G(|x-y|), то есть функция Грина зависит только от расстояния между точками x и y. Поскольку сетки (11), на которых введены базисные функции, равномерны по всем трем направлениям, аналогичное свойство распространяется и на матрицу задачи, а именно

$$\left(A_{kl}
ight)_{i_1i_2i_3}^{j_1j_2j_3} = A_{kl}(i_1-j_1,i_2-j_2,i_3-j_3).$$

В случае, когда сверхпроводник в плоскости $x_2 = 0$ присутствует, функция Грина имеет более сложный вид (10). Функция Грина в этом случае представима в виде суммы функции, зависящей от разности координат двух точек и функции, зависящей от разности и суммы координат двух точек. Для матрицы задачи это означает, что

$$\left(A_{kl}\right)_{\substack{j_1j_2j_3\\j_1j_2j_2}}^{j_1j_2j_3} = A_{kl}^{(1)}(i_1-j_1,i_2-j_2,i_3-j_3) + A_{kl}^{(2)}(i_1-j_1,i_2+j_2,i_3-j_3). \tag{17}$$

Матрица, элементы которой зависят только от разности номера строки и столбца, называется mennuqesoй. Матрица, элементы которой зависят от суммы индекса строки и столбца, называется sankenesoй. Итак, в нашем случае матрица задачи представляется в виде суммы двух блочных 3×3 матриц, каждый блок которых обладает следующими свойствами:

- имеет трехуровневую структуру;
- на верхнем уровне в обоих случаях блоки являются блочнотеплицевыми

- на втором уровне для матрицы $A^{(1)}$ блоки блочно-теплицевы, для матрицы $A^{(2)}$ блочно-ганкелевы;
- на нижнем уровне в обоих случаях блоки являются теплицевыми матрицами.

Матрицы, обладающие подобными структурами, рассматривались в [3, 4, 5, 6].

Как известно, для хранения теплицевой матрицы размера n необходимо O(n) ячеек памяти, а умножение на нее можно произвести за $O(n\log(n))$ арифметических действий [7]. То же самое справедливо и для ганкелевой матрицы, так как перестановкой строк и столбцов ее можно привести к теплицевой. Таким образом, матрицу нашей задачи можно хранить в $O(n_1n_2n_3)$ ячеек памяти и умножать на вектор за $O(n_1n_2n_3\log(n_1n_2n_3))$ арифметических операций, пользуясь процедурами быстрого умножения Фурье (см. [8]).

Имея оптимизированную процедуру умножения на матрицу, мы можем найти решение линейной системы (14) с помощью какогонибудь итерационного метода (в нашей работе применяется метод минимальных невязок). По найденному решению \mathbf{J} строятся поля в точках наблюдения, согласно формулам (6) и (7).

5. Параллельный алгоритм

Предварительные расчеты, проведенные по указанному алгоритму, показывают, что точность вычисления величин полей заметно ухудшается в случае, когда плоскость, содержащая источник и точки наблюдения, приближается к неоднородности. В этом случае для поддержания определенной точности ответа, приходится увеличивать количество элементов сеток (11), что приводит к резкому росту памяти, необходимого для хранения матриц и векторов задачи и арифметических операций, необходимых для решения линейной системы.

Рассмотрим в качестве модельного примера неоднородность \mathcal{V} как параллелепипед $[-a,a] \times [H,H+a/2] \times [-a,a]$. Сетку для дискретизации выберем размером $n \times n/2 \times n$. В таблице 1 приведена зависимость количества памяти, необходимой для хранения векторов и матриц, использованных при расчете, от размера выбранной сетки. Данные приведены в предположении, что память, необ-

Таблица 1

n	16	32	64	128	256
$N = n^3/2$	2^{11}	2^{14}	2^{17}	2^{20}	2^{23}
$\mathtt{mem}(N)$	$32~\mathrm{Kb}$	$256~\mathrm{Kb}$	$2~\mathrm{Mb}$	$16~\mathrm{Mb}$	128 Mb
$\mathtt{mem}(A^{(1)})$	$2.3~\mathrm{Mb}$	$18.4~\mathrm{Mb}$	$144~\mathrm{Mb}$	1.1 Gb	$9.2~\mathrm{Gb}$

ходимая на хранение одного элемента составляет mem(1)=16b. Матрица $A^{(1)}$ занимает порядка $9\cdot 8\cdot N$ ячеек памяти.

На персональных компьютерах, доступных в текущее время, количество оперативной памяти не превышает 2 Гб. Таким образом, провести интересующие нас расчеты на персональной станции возможно только при числе точек сетки n=16,32,64. Для расчетов с большими значениями числа точек n необходимо привлекать многопроцессорные станции, обладающие большими ресурсами оперативной памяти и большей вычислительной мощностью. В этой работе мы изучаем метод реализации описанного алгоритма на многопроцессорных платформах с распределенной памятью, а именно на кластерных станциях, которые представляют собой вычислительную систему на основе обычных рабочих станций, соединенных высокоскоростной сетью. Такой принцип построения вычислительных систем является весьма недорогим, но кластерные станции обладают рядом особенностей:

- скорость межпроцессорных соединений невысока, много медленнее скорости операция внутри одного вычислительного узла.
- время передачи массива данных складывается из фиксированного времени инициализации и времени пересылки, которое пропорционально длине массива. На практике время инициализации может быть весьма велико.
- скорость обмена данными между соседними узлами выше, чем скорость обмена данными между случайной парой узлов.

Исходя из этих особенностей вычислительной платформы, при разработке параллельного кода стоит по возможности придерживаться следующих принципов:

- уменьшить число обменов данными между вычислительными узлами, возможно даже за счет увеличения вычислительной сложности на каждом узле.
- производить передачу данных крупными блоками
- производить обмен данными между соседними вычислительными узлами, пользуясь процедурами глобального обмена данными только по необходимости
- организовывать алгоритм так, чтобы минимизировать "простой" вычислительного узла, связанный с ожиданием им прибытия необходимых данных с других узлов. Этого можно достигнуть следующими способами:
 - организовать алгоритм так, чтобы на каждом узле всегда хранился необходимый запас данных для проведения расчета. при этом каждый вычислительный узел сам отслеживает количество оставшихся у него данных для расчета, своевременно информирует корневой узел о необходимости прислать дополнительные данные и получает новые данные раньше, чем закончит работу с текущими
 - организовать алгоритм так, чтобы операции пересылки данных происходили одновременно с операциями, требующими существенны расчетов на вычислительном узле.
 Например, обмениваться частями распределенного вектора в то же время, когда происходит умножение этого вектора на локальные части матриц.

Для параллельной реализации описанного метода необходимо распределить по **пр** процессорам следующие объекты:

- векторы размера $3 \cdot n_1 n_2 n_3$, предназначенные для хранения
 - правой части уравнения (14)
 - вектора неизвестных
 - вспомогательных векторов итерационного метода
- составные части матрицы A (см. (17))

- $-~A^{(1)}$ состоит из 3×3 трехуровневых трижды теплицевых
- $-A^{(2)}$ состоит из 3×3 трехуровневых теплиц-ганкельтеплицевых

В силу указанной специфики количество памяти, необходимой для хранения каждой из матриц, пропорционально их размеру и составляет

$$3 \cdot 3 \cdot (2n_1 - 1)(2n_2 - 1)(2n_3 - 1) \sim 72n_1n_2n_3$$

для каждой.

Векторы распределяются по процессорам путем "разрезания" множества неизвестных на ${\bf np}$ частей вдоль оси x_3 (рис. 2). Конкретнее, определим числа

$$k_3 = \lfloor n_3/\mathtt{np} \rfloor$$
, $r_3 = n_3 - k_3 n_3$

и разобьем сетку вдоль оси x_3 на **пр** частей, в каждую из которых войдет $n_3^{\{p\}}$ точек:

$$n_3^{\{0\}}=k_3+r_3, \qquad n_3^{\{p\}}=k_3, \qquad$$
 для прочих $p=1,\dots,$ np $-1.$

Таким образом, "корневой" процессор с индексом 0 берет на себя большую из частей сетки x_3 , если ее невозможно распределить равномерно. На каждом из процессоров хранится, с учетом трехуровневой структуры $3n_1n_2n_3^{\{p\}}$ элементов вектора.

Определим теперь схему размещения на процессорах элементов матрицы. Поскольку матрица обладает трехуровневой структурой, а разрезание множества неизвестных происходит только вдоль оси x_3 , можно рассматривать только разрезание верхнего уровня. На верхнем уровне, отвечающем сетке вдоль x_3 , все 3×3 блока матриц $A^{(1)}$ и $A^{(2)}$ являются теплицевыми, поэтому можно поставить вопрос об эффективном распределении по процессорам элементов теплицевой матрицы, не учитывая тот факт, что ее элементы сами по себе являются двухуровневыми матрицами специальной структуры. Поскольку для теплицевой матрицы элементы зависят только от разности индексов строки и столбца $a_{ij}=a_{i-j}$, реально мы распределяем по процессорам вектор длины 2n-1, однако для эффективного умножения на распределенную матрицу

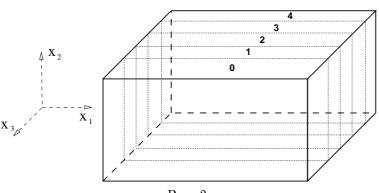


Рис. 2.

структура распределения этого вектора должна быть согласована с распределением элементов вектора неизвестных и правой части, описанном выше. Проиллюстрируем это распределение картинкой, см. рис. 3.

Корневой процессор размещает у себя один сегмент матрицы, все прочие по два. Индексы $i_3^{\{0\}}=i\!-\!j$ элементов a_s , размещенных на корневом процессоре, следующие

$$i_3^{\{0\}} = -n_3^{\{0\}} + 1, \dots, n_3^{\{0\}} - 1.$$

Прочие процессоры размещают у себя по два сегмента элементов, отвечающие индексам

$$egin{aligned} i_3^{\{p\}} &= (-p-1)k_3 - r_3 + 1, \cdots, (-p+1)k_3 - 1, \ i_3^{\{p\}} &= (\mathsf{np} - p - 1)k_3 + 1, \cdots (\mathsf{np} - p + 1)k_3 + r_3 - 1 \end{aligned}$$

Алгоритм умножения на распределенную матрицу $y := A \cdot x$. Каждому процессору присвоен уникальный номер p, лежащий в пределах от 0 до $\mathbf{np}-1$. Алгоритм представлен для процессора с номером p, хотя выполняется, конечно же, на всех процессорах одновременно.

- ${\bf do}\ \ {\tt ip}=0,\ {\tt np}-1\ \ //{\tt pacчeт}$ локальной части вектора y для процессора ${\tt ip}$
 - 1. [irecv] предоставить локальный вектор \tilde{x} для приема локальной части вектора x с соседнего процессора;

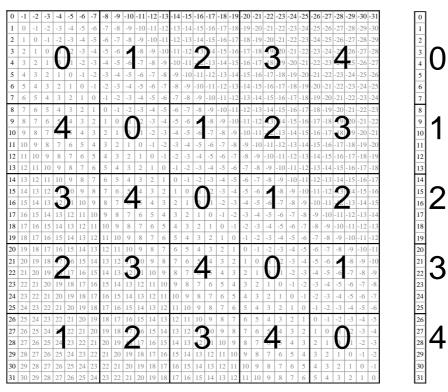


Рис. 3.

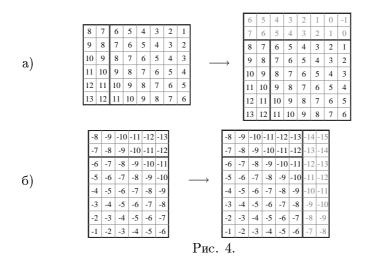
- 2. [isend] $x \to \tilde{x}$ выслать локальную часть вектора x процессору с индексом $p+1 \mod np$ ("вниз");
- 3. [A*x] $\tilde{y} := A \cdot x$ произвести умножение подходящей локальной части матрицы на локальную часть вектора x(подробнее ниже);
- 4. [allreduce] $y^{\{\mathrm{ip}\}}:=\sup_{\mathrm{all}} \tilde{y}$ просуммировать локально вычисленные вектора \tilde{y} на процессор $\mathrm{ip};$
- 5. [wait isend] дождаться завершения операции 2, если необходимо;
- 6. [wait irecv] дождаться завершения операции 1, если необходимо.

enddo

Опишем способ быстрого умножения на локальные части матриц, использованный в пункте 3 описанного алгоритма. Вопервых, отметим, что блоки матрицы, хранимые на каждом из процессоров, как и вся матрица, обладают свойством $a_{ij}=a_{i-j}$, следовательно, являются теплицевыми матрицами (для процессора p=0) или могут быть расширены до теплицевой матрицы (для прочих процессоров). Для умножения теплицевой матрицы на вектор существует эффективный алгоритм на основе быстрого умножения Фурье, позволяющий произвести умножение за $O(n \log n)$ действий для матрицы порядка n.

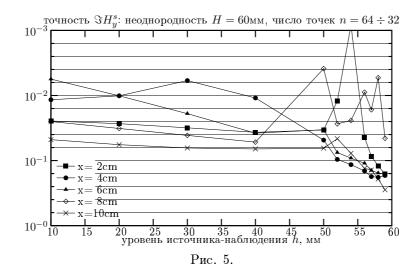
Теперь результат умножения на прямоугольные матрицы можно получить путем умножения на расширенные (квадратные теплицевы) матрицы следующим образом (см. рис. 4):

- а) для блоков в нижнем треугольнике матрицы:
 - 1. умножить расширенную матрицу на исходный вектор,
 - 2. отбросить первые r_3 компонент полученного вектора;
- б) для блоков в верхнем треугольнике матрицы:
 - 1. дополнить исходный вектор снизу r_3 нулевыми компонентами,
 - умножить расширенную матрицу на расширенный вектор.



Кроме умножения на полученную после дополнения теплицеву матрицу, нам потребуется алгоритм умножения на ее квадратную подматрицу меньшего размера (отмечена на рисунке 4). Он строится на основе алгоритма умножения на теплицеву матрицу следующим способом:

- а) для блоков в нижнем треугольнике матрицы:
 - 1. расширить исходный вектор сверху r_3 нулевыми компонентами,
 - 2. умножить расширенную матрицу на расширенный вектор,
 - 3. отбросить первые r_3 компонент полученного вектора;
- б) для блоков в верхнем треугольнике матрицы:
 - 1. расширить исходный вектор снизу r_3 нулевыми компонентами,
 - умножить расширенную матрицу на расширенный вектор,
 - 3. отбросить первые r_3 компонент полученного вектора;
- 0) для блоков на главной диагонали (корневого процессора):



- 1. расширить исходный вектор снизу r_3 нулевыми компонентами,
- 2. умножить исходную матрицу на расширенный вектор,
- 3. отбросить последние r_3 компонент полученного вектора.

Таким образом, мы описали процедуры умножения для всех процессоров на все хранимые ими части исходной матрицы. Тем самым, пункт 3 из алгоритма умножения на распределенную матрицу, а вместе с ним и сам алгоритм полностью описан.

Построение параллельного алгоритма для умножения на блочнодиагональную часть матрицы (15) не представляет труда. Таким образом, умножение на матрицу задачи полностью параллелизовано. На основе параллельного алгоритма умножения на матрицу производится решение задачи (14) с помощью какого-либо итерационного метода. В нашей работе мы использовали метод минимальных невязок [9].

6. Численные эксперименты

6.1. Повышение точности измерения полей. Суть эксперимента в следующем. Измеряются значения магнитных полей в 5 точках наблюдения, расположенных на одном уровне с источником. Высота этого уровня изменяется в наших экспериментах в

Таблица 2

геометрия	
размер неоднородности	(0.8, 0.2, 0.8)
уровень неоднородности	H = 0.06
положение источника	(0, h, 0)
тип источника	у-диполь
частота источника	100 МГц

электрические параметры							
пар	аметр	среда	неоднор.				
сопроз	гивление	10^{3}	50				
магн.	прониц.		0				
диэл.	прониц.		1				

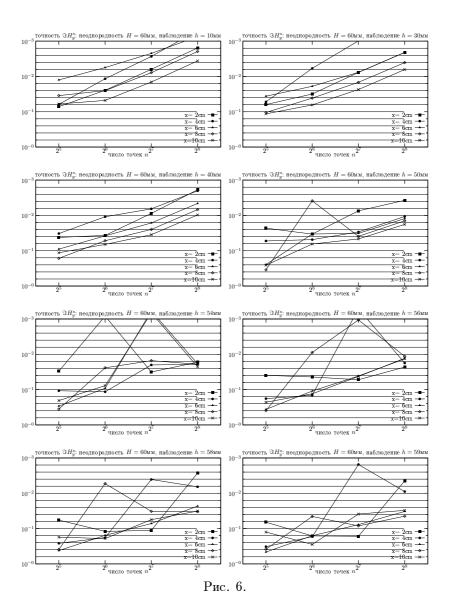
точки наблюдения								
1 2 3 4 5								
(0.02, h, 0.0)	(0.04, h, 0.0)	(0.06, h, 0.0)	(0.08, h, 0.0)	(0.10, h, 0.0)				

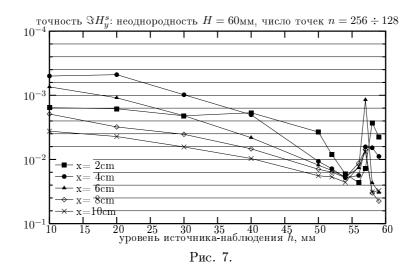
пределах от $h=10\,\mathrm{mm}$ до $h=59\,\mathrm{mm}$. В измеренных полях рассматриваемой величиной является мнимая часть проекции поля, направленной вдоль оси диполя-источника, то есть вдоль оси x_2 (эту ось мы в дальнейшем будем обозначать также y). Нас интересует точность измерения этой компоненты, причем так как первичное поле вычисляется, в принципе, достаточно просто и надежно, мы будет исследовать точность измерения аномального поля $\Im H_y^s$. В качестве меры точности принимается критерий внутренней $\operatorname{cxodu-mocm} u$, выраженный через два измерения одной и той же величины, произведенные при двух значениях числа точек дискретизации n, отличающихся вдвое.

$$arepsilon_{ ext{im}}(n) = rac{\left| \Im H^s_y(n) - \Im H^s_y(n/2)
ight|}{\left(\left| \Im H^s_y(n)
ight| + \left| \Im H^s_y(n/2)
ight|
ight)/2}$$

Численные параметры эксперимента приведены в таблице 2.

Из таблицы 1 следует, что в том случае, когда расчет по обсуждаемому методу происходит на стандартной персональной вычислительной платформе с количеством памяти до $2\,\Gamma$ б, количество элементов сетки ограничено величиной n=64. На картинке приведены результаты следующего эксперимента: мы фиксируем две сетки с количеством узлов n=32 и n=64 и, приближая плоскость, содержащую источник и точки наблюдения, к неоднородности, измеряем точность вычисления интересующих нас компонент полей. Рассматривая рисунок 5, можно заметить, что при приближении к неоднородности на расстояние ближе $10\,\mathrm{мм}$, то есть при $h\geqslant 50\,\mathrm{мм}$, происходит резкое ухудшение точности измерения ком-





понент полей практически для всех точек наблюдения. Для наших практических вычислений необходимо поддерживать точность вычисления компонент полей хотя бы на уровне 10%. Для того, чтобы получить возможность производить расчеты с такой точностью в областях, более близких к неоднородности, нам приходится увеличивать количество точек сетки n, что вызывает необходимость использовать многопроцессорные вычислительные системы.

Увеличение числа точек сетки приводит к повышению точности расчета, что проиллюстрировано экспериментами (рис. 6).

Возможность увеличить количество узлов сетки, конечно же, не отменяет ухудшения точности расчета при приближении плоскости измерения к неоднородности, но заметно смягчает ее, делая возможным проведения измерений с требуемой точностью на расстояниях порядка 1мм от неоднородности (рис. 7).

Число итераций, необходимых для обращения матрицы задачи, заметно увеличивается при приближении плоскости источниковприемников к неоднородности. При этом этот рост смягчается на больших размерах сетки, что проиллюстрировано таблицей 3. В ней приведено число итераций, необходимых для решения задачи при разных значениях числа узлов n. Параметр остановки итерационного метода выбран равным $\varepsilon = 0.05$, поскольку предварительные эксперименты показывают, что при этом погрешность

Таблица 3

h, мм											
n = 64	2	2	2	2	4	6	6	7	8	8	8
n = 64 $n = 256$	2	2	2	2	2	2	2	2	3	5	7

итерационного решения уже не сказывается существенно на значениях вычисляемых компонент поля.

6.2. Улучшение разрешающей спосовности привора. Рассмотрим теперь случай, когда неоднородность по существу является таковой, то есть содержит области с различными элек-

является таковой, то есть содержит области с различными электрическими свойствами. Для наглядности остановимся на следующем случае: неоднородность представляет собой (как и в прошлом эксперименте) параллелепипед

$$V = [-0.4, 0.4] \times [0.06, 0.26] \times [-0.4, 0.4]$$

с небольшим "вкраплением" $\tilde{\mathcal{V}}$, в качестве которого рассмотрим один из двух вариантов:

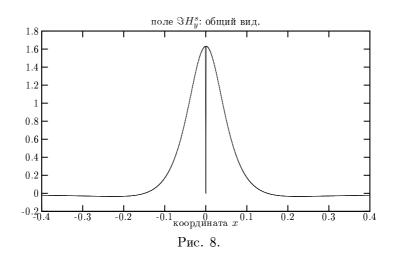
$$\tilde{\mathcal{V}}_1 = [0.29677 : 0.32258] \times [0.06000 : 0.06645] \times [-0.01290 : 0.01290]$$

$$\tilde{\mathcal{V}}_2 = [0.19355 : 0.21935] \times [0.06000 : 0.06645] \times [-0.01290 : 0.01290]$$

Геометрически эти вкрапления возникают следующим образом: вся неоднородность разбивается на $31\times31\times31$ макроячеек, и электрические параметры для одной из них устанавливаются отличными от всех других. Вкрапление $\tilde{\mathcal{V}}_1$ занимает место макроячейки с индексом (28,1,16), а вкрапление $\tilde{\mathcal{V}}_2$ — позицию (24,1,16). Соотношения электрических параметров представлены в таблице 4.

Таблица 4

электрические параметры							
	параметр	внеш. среда	неоднор.	вкрапление			
сопрот	ивление, Ом·м	10^{3}	50	2			
магн.	проницаемость	0	0	0			
диэл.	проницаемость	1	1	1			



Цель проводимых измерений состоит в том, чтобы с помощью величин магнитных полей, измеренных в удачно для этого выбранной плоскости наблюдения, "почувствовать" наличие вкрапления и максимально точно определить его месторасположение. Пример, рассматриваемый нами, является в определенном смысле модельным, но позволяет качественно понять проблему, для решения которой нам необходимо увеличивать точность дискретизации.

Итак, при зафиксированной геометрии, проведем измерения величин интересующих нас компонент магнитных полей (как и прежде, это мнимая часть компоненты аномального поля $\Im H_y^s$) в плоскости измерения $x_2=h$ на отрезке x=[-0.4,0.4], y=h, z=0. Общий вид аномального поля представлен на картинке 8.

Измеряемая компонента поля весьма сильно изменяется на рассматриваемом отрезке, поэтому для рассмотрения поведения полей вблизи вкрапления мы будем приводить графики с подходящим выбором рассматриваемого диапазона значений поля. Рассмотрим графики на стр. 162-163. На них изображено поведение исследуемой компоненты на рассматриваемом отрезке, пунктиром отмечена область, которая соответствует вкраплению. Слева представлены графики для вкрапления $\tilde{\mathcal{V}}_1$, справа для вкрапления $\tilde{\mathcal{V}}_2$.

По первой серии графиков, сделанной при количестве узлов сетки n=64 с использованием обычной однопроцессорной рабочей

станции, можно сделать следующие наблюдения.

- На существенном расстоянии плоскости измерения от неоднородности ($h=40\,\mathrm{mm}$, то есть в $20\,\mathrm{mm}$ от неоднородности) обнаружить по измеренным полям наличие вкрапления практически невозможно.
- При приближении плоскости измерения к неоднородности график измеряемой компоненты от координаты x начинает претерпевать изменения в районе, соответствующем положению вкрапления. Эти изменения слегка заметны при h = 50 мм (плоскость измерения в 10 мм от неоднородности), и становятся все более заметны при дальнейшем приближении к неоднородности.
- Уже при $h=52\,\mathrm{mm}$ и особенно при $h=54\,\mathrm{mm}$ мы можем достаточно хорошо определить положение вкрапления по поведению графика измеряемого поля для вкрапления $\tilde{\mathcal{V}}_1$. Однако положение вкрапления $\tilde{\mathcal{V}}_2$ по соответствующим ему графикам, продолжает вызывать определенные сомнения.
- Данные, полученные в плоскости измерения, размещенной на уровне $h=56\,\mathrm{mm}$, позволяют еще более точно определить расположение вкрапления $\tilde{\mathcal{V}}_1$. Однако эти измерения содержат уже достаточно сильные погрешности, что выражается в существенных их биениях, "шумах", которые могут быть восприняты как наличие других вкраплений и тем иным способом внести ошибки в процедуру локации неоднородности. Возникающие биения делают невозможным построить сколь-либо точные предположения относительно положения вкрапления $\tilde{\mathcal{V}}_2$.
- Дальнейшее приближение плоскости измерения к неоднородности ($h=58\,\mathrm{mm}$) приводит к тому, что результаты измерений вследствие возникающих погрешностей становятся неприменимыми для анализа.

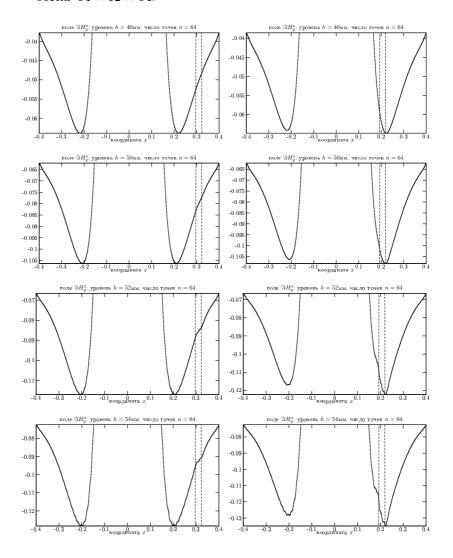
Сделанные наблюдения приводят к выводу о необходимости увеличивать точность измерений с целью подавить биения и добиться более аккуратных измерений полей вблизи неоднородно-

сти. Для достижения этой цели приходится увеличивать количество узлов сетки дискретизации, что уже возможно реализовать только на параллельной платформе.

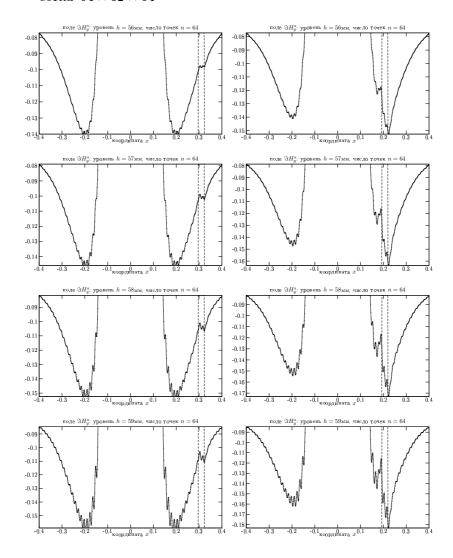
Вторую серию экспериментов проведем с количеством узлов n=128. Результаты представлены на стр. 164–165. Видим, что определить местоположения включения $\tilde{\mathcal{V}}_1$ теперь стало существенно проще, особенно при выборе плоскости наблюдения на уровне $h=57-58\,\mathrm{mm}$ (2-3 мм от неоднородности). Определение местоположения включения $\tilde{\mathcal{V}}_2$ по-прежнему, затруднено ввиду больших "шумов", но с определенной точностью все же возможно (на уровне $h=56-57\,\mathrm{mm}$ от плоскости). Таким образом, использование для дискретизации сетки размером n=128, эксперименты с которой можно проводить с использованием 16 процессоров на кластерной станции (что довольно немного для такого типа машин), уже существенно улучшает чувствительность измерений.

Третью серию экспериментов проведем с количеством узлов n=256. Рассматривая результаты этой серии экспериментов (стр. 166-167), можно заметить, что положение как неоднородности $\tilde{\mathcal{V}}_1$, так и $\tilde{\mathcal{V}}_2$ хорошо просматривается по графикам измеренных компонент полей при положении плоскости источников на уровне $h=57-59\,\mathrm{mm}$. Это означает, что точность решения прямой задачи, то есть определения величин полей при заданной геометрии неоднородности, при таком значении параметров уже достаточно высока для того, чтобы на основе алгоритма, реализующего прямую задачу, строить алгоритм для решения обратной задачи, т.е. определения геометрии и электрических параметров среды внутри неоднородности.

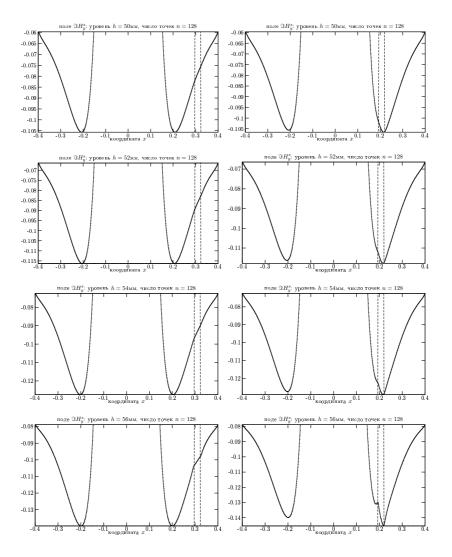
Результаты численных экспериментов: сетка $64 \times 32 \times 64$.



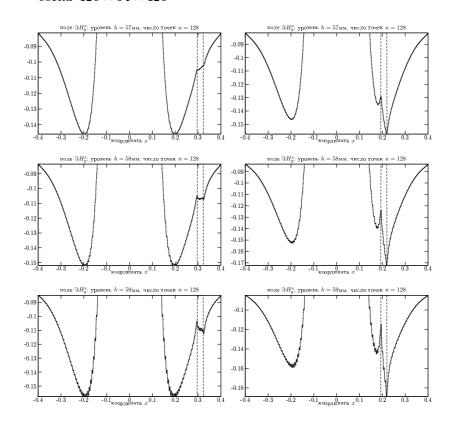
Результаты численных экспериментов: сетка $64 \times 32 \times 64$



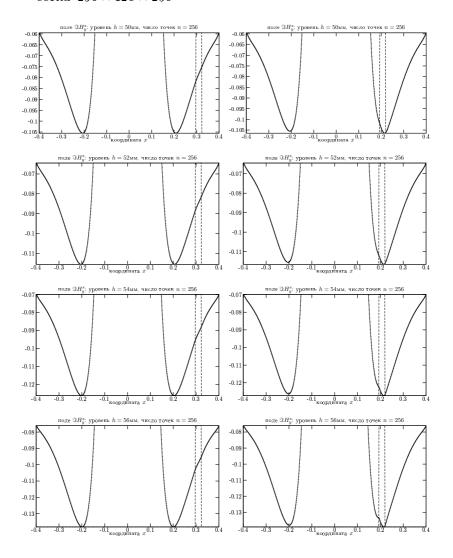
Результаты численных экспериментов: сетка $128 \times 64 \times 128$



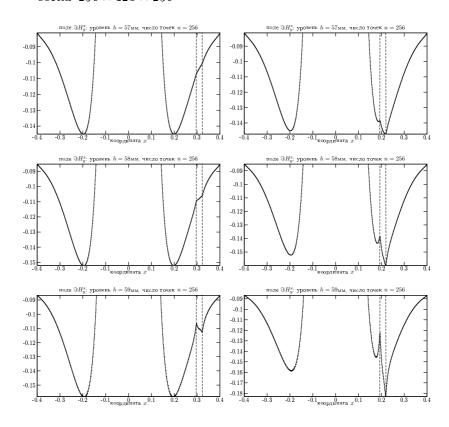
Результаты численных экспериментов: сетка $128 \times 64 \times 128$



Результаты численных экспериментов: сетка $256 \times 128 \times 256$



Результаты численных экспериментов: сетка $256 \times 128 \times 256$



6.3. Повышение скорости решения.

Ускорение существующего алгоритма, как таковое, не являлось целью нашей работы. Проблема состояла не в том, что существующий алгоритм работал медленно, а в том, что с его помощью мы не могли получить решение задачи в интересующих нас областях пространства с требуемой точностью. Потребность в использовании многопроцессорных вычислительных систем возникла не из-за нехватки времени, а из-за нехватки оперативной памяти, необходимой для того, чтобы хранить матрицы больших плотных систем, возникающих при дискретизации интегрального уравнения (8) на больших сетках. Тем не менее, мы приведем характерные времена исполнения параллельного алгоритма на используемых вычислительных платформах приводится — не как характеристику ускорения, а просто для сравнения.

Тестирование произведено на следующих системах:

inm Кластер ИВМ РАН

Узлы HP rx2600 на основе процессоров Intel Itanium2, 1.3 GHz, сеть Myrinet 2000, 2x250 Mb/s.

ncs Кластер факультета математики Университета земли Саар, Германия.

Узлы на основе процессоров Intel Xeon, $2.2~\mathrm{GHz}$, сеть на основе двух параллельных Fast Ethernet коммутаторов, $2\mathrm{x}100~\mathrm{Mb/s}$.

mvs Кластер MBC-1000М в Межведомственном Суперкомпьютерном Центре.

Узлы на основе процессоров Alpha 21264A, 667 MHz, сеть Myrinet 2 Gbit/s.

В таблице 5 представлены времена исполнения алгоритма на различных вычислительных платформах. В первой колонке указано название используемой платформы и число процессоров, на которых производился расчет.

Выражаем благодарность наших зарубежным коллегам С. Рязанову и И. Ибрагимову за предоставленную вычислительную платформу для проведения экспериментов.

 $\begin{tabular}{ll} {\bf Таблица} \begin{tabular}{ll} {\bf 5} \\ B peмя \ pa fomu \ nporpammu \ na \ cem ke \ n=32, \ cen. \end{tabular}$

h, mm	10	20	30	40	50	52	54	56	57	58	59
iter	3	3	4	5	7	7	7	8	8	8	8
2*inm	4.33	4.32	5.17	6.05	7.78	7.77	7.89	8.69	8.76	8.75	8.75
4*inm	3.66	3.68	4.44	5.16	6.81	6.77	6.79	7.61	7.77	7.75	7.74
8*inm	3.52	3.52	4.29	5.06	6.64	6.63	6.66	7.43	7.42	7.39	7.43
2*ncs	9.11	9.11	11.1	13.3	17.6	17.7	17.7	20.0	19.9	19.8	20.1
4*ncs	4.20	4.23	5.11	5.96	7.97	7.99	8.22	9.16	9.15	9.15	9.21
8*ncs	3.61	3.81	4.71	5.63	7.49	7.30	7.53	8.43	8.50	8.55	8.43
16*ncs	5.17	5.17	6.47	7.41	9.52	9.72	9.61	10.7	10.6	10.8	10.7
2*mvs	12.1	12.0	15.1	16.8	24.7	25.0	24.4	26.5	27.8	27.6	26.6
4*mvs	12.3	11.4	14.5	17.4	22.8	23.6	23.0	25.7	24.1	24.6	24.5
8*mvs	11.3	11.1	13.7	16.1	21.0	21.5	21.7	24.2	23.3	23.0	23.7
16*mvs	10.5	10.4	12.9	15.4	19.7	20.3	20.1	23.0	23.2	23.4	22.7

Время работы программы на сетке $n=64,\ ce\kappa.$

h, mm	10	20	30	40	50	52	54	56	57	58	59
iter	2	2	2	2	4	6	6	7	8	8	8
2*inm	50.5	53.4	53.4	53.7	80.8	109	109	123	137	137	137
4*inm	44.3	44.4	44.4	44.2	69.7	95.7	95.6	105	116	116	117
8*inm	20.7	20.8	20.8	20.8	33.2	45.6	45.5	52.0	58.2	58.2	58.2
2*ncs	78.5	78.8	79.5	78.8	128	180	179	206	232	232	232
4*ncs	72.0	71.8	72.5	71.9	119	166	166	190	214	215	214
8*ncs	62.0	61.9	61.9	62.1	102	144	144	164	185	185	184
16*ncs	36.2	36.8	36.1	36.1	58.1	79.7	80.2	91.6	102	102	102
2*mvs	103	103	102	107	159	213	217	248	274	295	308
4*mvs	83.8	85.1	86.9	84.6	134	188	182	216	199	199	198
8*mvs	66.2	69.6	68.4	68.7	109	151	155	168	187	186	185
16*mvs	63.3	64.0	63.8	62.9	107	149	149	171	193	193	193

Литература

- [1] А. Б. Самохин. Исследование задач дифракции электромагнитных волн в локально-неоднородных средах. $\mathcal{H}BMuM\Phi$, 30(1):107-121, 1990.
- [2] А. Б. Самохин. Интегральные уравнения и итерационные методы в электромагнитном рассеянии, *Радио и связь*, М., 1998.
- [3] Yu. G. Smirnov, A. A. Tsupack. Volume singular integral equations for solving diffraction problem of electromagnetic waves in mocrowave oven. *Proc. of European Symp. on Numer. Meth. in Electromagnetics* (March 6-8, 2002, Tolouse, France), 172-176.
- [4] V. I. Ivakhnenko, A. V. Kukk, E. E. Tyrtyshnikov. Application of 3D volume integral equations to solution of electromagnetic wave scattering problems. Research Report EM-RR 22/95. Elegant Mathematics, March 8, 1995.
- [5] V. I. Ivahnenko, E. E. Tyrtyshnikov, Block-Toeplitz-Structure-based solution strategies for CEM problems. 11th Annual Review of Progress in Applied Comp. Electromagnetics, Conf. Proceedings, Monterey, CA, March 20-25, 1995, 181-188.
- [6] Ю.А.Еремин, В.И.Ивахеннко. Строгие и приближенные модели царапины на основе метода интегральных уравнений. Дифф. уравнения, т. 37, №10, 1386-1394, 2001.
- [7] Воеводин В. В., Тыртышников Е. Е., Вычислительные процессы с теплицевыми матрицами. М.: Наука, 1987.
- [8] A.P.M. Zwamborn, Van der Berg. The three-dimensional weak form of the conjugate gradient FFT method for solving scattering problems. *IEEE Trans. Microwave Theory Tech.*, 1992, MTT-40, 9:1757-1765.
- [9] Y.Saad, M.H.Schultz, GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems. SIAM F. Scientific and Stat. Comp. 7:856-869, 1986.

Применение нелинейных методов аппроксимаций для быстрого решения задачи распространения звука в мелком море

И. В. Оселедец, Д. В. Савостьянов, С. Л. Ставцев

Проведено численное решение задачи распространения звука в мелком море с помощью недавно разработанных методов нелинейных приближений больших матриц: мозаично-скелетонного метода, тензорных и вейвлет-аппроксимаций. Использование быстрых методов позволило решать задачу на гораздо более мелких сетках, чем это было возможно ранее.

1. Введение

Многие пространственные задачи дифракции акустических и электромагнитных волн сводятся к решению системы интегральных уравнений. Особый интерес представляет решение задачи распространения звука в мелком море. Звуковые волны — единственные из волн известной физической природы способны распространяться под водой на расстояние многих сотен километров. Этой уникальной способностью и объясняется интерес, проявляемый учеными и инженерами самых разных специальностей.

Распространение звуковых волн под водой — крайне сложный процесс, трудность описания которого обусловлена многообразием явлений и свойств, присущих разным районам Мирового океана. В первую очередь это относится к шельфовым зонам океана, где особенности распространения звуковых волн связаны с их взаимодействием с морским дном. Вместе с тем, океанский шельф наиболее важен для жизнедеятельности человека, а следовательно и наиболее интересен для исследователей. К настоящему времени усилиями разных групп получено большое количество экспериментальных и теоретических результатов [3, 5, 12, 13], позволяю-

щих выделить акустику мелкого моря в самостоятельный раздел акустики океана.

Понятие "мелкое море" существует в подводной акустике давно. Причиной этого стали хорошо известные особенности распространения звука в водной среде, связанные главным образом с взаимодействием с морским дном. Величины характерных параметров волновода, определяемого как "мелкое море", могут изменяться в достаточно широких пределах. Это зависит не только от геометрических размеров волновода, но и от частоты звука. Исследование "мелкого моря" проводится чаще всего с помощью модовых методов. Действительно, при акустическом мониторинге шельфовых зон оказывается невозможным выделение (как по пространству, так и по времени) отдельных лучей, так как акустические сигналы, распространяющиеся по соседним лучам, перекрываются. Однако использование модовых методов возможно только для случая, когда звуковые волны распространяются на расстоянии десятков и сотен километров. При этом рассматривается изменение энергетических спектров на большом расстояние. В соответствии с этим влияние дна производится с некоторым усреднением и производится усреднение рассчитываемого поля давления.

Но возникают задачи, когда требуется рассчитать акустическое поле как вблизи, так и вдали от звуковых источников. В этом случае решить задачу с необходимой точностью, используя только "энергетические" усреднения, не удается. Поэтому можно воспользоваться другими (численными) методами для решения граничных задач акустики.

Сейчас существует большое количество численных методов, направленных на решение граничных задач математической физики. Значительное место среди них занимают сеточные методы [1, 2, 7]. Они дают численные результаты, хорошо согласующие с экспериментом, обладают большой гибкостью и адаптивностью к решаемым задачам. Однако для решения акустических проблем они не являются универсальным средством. Поскольку в ряде задач возникает необходимость расчета поля от источника, то требуется подстраивать сетку под источники, а также значительно измельчать сетку вблизи особых точек, что существенно усложняет расчеты и требует значительных компьютерных ресурсов. Вследствие

того, что исходное акустическое поле содержит особенности, при решении задач сеточными методами наблюдается слабая сходимость численных расчетов. При решении задач сеточными методами возникают трудности, если область распространения звука является неограниченной. В этом случае сетку приходится обрывать и считать поле на больших расстояниях сферическим или цилиндрическим (в зависимости от симметрии задачи), и возникает необходимость сшивать решения на границе обрыва. Кроме этого, как показывают численные исследования, размеры области, в которой необходимо проводить численные расчеты, велики, и поэтому решение задачи требует значительных компьютерных ресурсов.

Другим направлением в численных исследованиях акустических полей, которые стали развиваться с середины 50-х годов XX века, являются численные методы, использующие интегральные уравнения [4, 6, 10, 11]. Получение и решение интегральных уравнений основано чаще всего на использовании теории потенциала и теории обобщенных функций (в частности, теории потенциала простого и двойного слоев). Использование интегральных методов позволяет избежать многих недостатков сеточных методов. Вследствие нелокальности интегрального оператора появляется возможность сразу учитывать потенциалы источника и снизить размерность задачи. Однако это же свойство (нелокальности) ухудшает структуру матриц, получаемых после дискретизации оператора, — матрица становится плотной. Это приводит к необходимости развития новых быстрых методов матричного анализа.

При решении задач дифракции звука на акустически жестких телах в ряде случаев [6, 9, 14, 15, 16] используют гиперсингулярные интегральные уравнения (или интегральные уравнения с сильной особенностью), для решения которых разрабатываются особые методы [6, 8, 9]. Однако те методы, которые описываются в данных работах, позволяют провести быстрые расчеты только для канонических поверхностей (сферы, эллипсоида), в которых необходимые параметры и коэффициенты можно быстро рассчитать. Недостатком этих методов является также то, что они применимы только при небольших значениях $k\,r$, где k — волновое число, r — характерный размер тела, на котором происходит дифракция звуковой

волны.

В отличие от вышеуказанных методов, использующих интегральные преобразования, метод дискретных особенностей (МДО) [14, 15], в частности метод замкнутых дискретных вихревых рамок (МЗДВР) [15, 17], является достаточно универсальным для широкого класса поверхностей (достаточно лишь гладкости). С другой стороны, этот метод позволяет с определенной степенью точности рассчитать акустическое поле как на больших, так и на малых расстояниях от источника. Однако, как и всякий метод, МДО имеет определенные ограничения в использовании. Как описано в [14, 15], при решении гиперсингулярного интегрального уравнения искомая функция аппроксимируется кусочно-постояннными функциями. Поэтому при высоких частотах (более 100 Гц), то есть при больших значениях k, для получения необходимого решения требуется построить мелкую сетку. Таким образом, после дискретизации интегрального уравнения (или системы интегральных уравнений) возникает необходимость решения систем с матрицами больших размерностей.

Далеко не всегда быстрые методы, разработанные в вычислительной линейной алгебре, успешно применяются для решения прикладных задач. Это может происходить по различным причинам. Например, новые технологии могут оказаться применимыми только для узкого класса задач или для задач, в которых нет острой необходимости для быстрых методов. В частности, создано большое количество быстрых методов для решения систем, появляющихся при дискретизации одномерных уравнений. В таких задачах размер матриц редко превышает 1000, и нет особой надобности использовать специальные методы решения систем линейных уравнений с такими матрицами. Размер же матриц, возникающих при численном решении двумерных уравнений (в частности, рассматриваемых в данной работе уравнений акустики), достигает десятков-сотен тысяч или даже миллиона. Поэтому использование быстрых методов жизненно необходимо.

Мы используем два метода. Первый использует перестановку неизвестных, вычисляемую на основе информации о геометрическом распложении точек сетки, с которой эти неизвестные ассоциированы. В результате после перестановки в матрице можно выделить достаточно большие блоки, к которым можно с хорошей

точностью найти приближение матрицами малого ранга с помощью метода неполной крестовой аппроксимации [22, 23, 24].

Второй основан на представлении матрицы в виде суммы тензорных произведений с помощью алгоритма неполной крестовой аппроксимации [22, 23, 24] (см. адаптированную версию в [28]). Далее, следуя [28, 26], тензорные факторы сжимаются с помощью вейвлет-преобразований, что приводит к суперлинейному сжатию и почти линейной сложности умножения матрицы на вектор.

2. Основные уравнения задачи акустики

Процесс распространения звука в воде, в том числе и в "мелком море", довольно сложен. Поэтому для его изучения построено множество моделей. Рассмотрим самую простую модель, когда скорость распространения звука в воде постоянна и равна c. Пусть слой воды занимает область D (рис. 1). Границы этой области — поверхность воды σ_1 , которая является плоскостью, и поверхность σ_2 — абсолютно отражающее звук морское дно.

Найдем P(M), $M \in D$, удовлетворяющее уравнению Гельмгольца с волновым числом k:

$$\Delta P(M) + k^2 P(M) = 0; \ M \in D \tag{1}$$

и следующим граничным условиям:

$$P(M) = f_1(M), M \in \sigma_1; \tag{2}$$

$$\frac{\partial P(M)}{\partial n_M} = f_2(M), M \in \sigma_2, \tag{3}$$

где \vec{n}_{M} – вектор внешней нормали к поверхности σ_{2} .

Будем искать P(M), которые удовлетворяют условию излучения на бесконечности:

$$\left(\frac{\vec{r}_M}{r_M}, \operatorname{grad} P(M)\right) - ikP(M) = O\left(\frac{1}{r_{MN}}\right) \tag{4}$$

при $r_{MN} \to \infty$, где N – фиксированная точка области $D,\ r_M=|\vec{r}_M|=|\vec{xi}+y\vec{j}+z\vec{k}|=\sqrt{x^2+y^2+z^2}.$

Введем систему координат Oxyz таким образом, чтобы плоскость Oxy совпадала с плоскостью σ_1 (см. рис. 1). Пусть в точке

 $M_Q(0,0,z_Q)\in D$ расположен неподвижный тональный с частотой ν источник звука, создающий звуковое давление p(M) по формуле

$$p_Q(M) = \frac{Q}{4\pi} \frac{e^{ikr_{MM_Q}}}{r_{MM_Q}},\tag{5}$$

где Q — мощность источника, $k = \frac{2\pi\nu}{c}$. Тогда в граничном условии (2)

$$f_1(M) = -p_Q(M), M \in \sigma_1, \tag{6}$$

а в граничном условии (3)

$$f_2(M) = -\frac{\partial p_Q(M)}{\partial n_M}, M \in \sigma_2.$$
 (7)

Для выполнения граничного условия Дирихле на плоскости расположим симметрично плоскости z=0 поверхность σ_2^* . То есть если поверхность σ_2 задается функцией h(x,y), определенной на всей действительной плоскости, то поверхность σ_2^* задается противоположенной функцией -h(x,y). Обозначим через D' область, ограниченную поверхностями σ_2 и σ_2^* . Симметрично источнику расположим сток той же интенсивности, то есть

$$p_Q^*(M) = -\frac{Q}{4\pi} \frac{e^{ikr_{MM_Q}^*}}{r_{MM_Q}^*},\tag{8}$$

где $r_{MM_Q}^* = \sqrt{(x_M - x_{M_Q})^2 + (y_M - y_{M_Q})^2 + (z_M + z_{M_Q})^2}$.

Будем искать давление $P_1(M) = P(M) - p_Q^*(M)$. На поверхности σ_2^* зададим граничное условие

$$\frac{\partial P_1(M)}{\partial n_M} = f_2(M), M \in \sigma_2^*. \tag{9}$$

Так как в D' содержатся источник и сток, давления которых вычисляются по формулам (5) и (8) соответственно, то в граничных условиях на σ_2 и σ_2^* функция в правой части принимает вид

$$f_2^*(M) = -\frac{\partial}{\partial n_M} (p_Q(M) + p_Q^*(M)), M \in \sigma_2, M \in \sigma_2^*.$$
 (10)

Таким образом, задача сводится к нахождению функции $P_1(M)$, удовлетворяющей граничному условию

$$\frac{\partial P_1(M)}{\partial n_M} = f_2(M), M \in \sigma_2, M \in \sigma_2^*. \tag{11}$$

Решение задачи Неймана будем искать через потенциал двойного слоя, распределенный поверхностям σ_2 и σ_2^* :

$$P_{1}(M) = \int_{\sigma_{2}} \frac{\partial F(M, N)}{\partial n_{N}} g(N) d\sigma_{2,N} +$$

$$+ \int_{\sigma_{2}^{*}} \frac{\partial F(M, N)}{\partial n_{N}} g^{*}(N) d\sigma_{2,N}^{*}, \quad M \notin \sigma_{2} \cup \sigma_{2}^{*},$$

$$(12)$$

где

$$F(M,N) = \frac{1}{4\pi} \frac{e^{ikr_{MN}}}{r_{MN}}.$$
 (13)

В силу симметрии задачи $g^*(x,y,-h(x,y)) = -g(x,y,h(x,y)),$ поэтому для $P_1(M)$ выполняется (11), если g(M) удовлетворяет интегральному уравнению на поверхности σ_2 :

$$\int_{\sigma_2} \frac{\partial}{\partial n_M} \frac{\partial G(M, N)}{\partial n_N} g(N) d\sigma_{2,N} =
= -\frac{\partial}{\partial n_M} (p_Q(M) - p_Q^*(M)), M \in \sigma_2,$$
(14)

где $G(M, N) = F(M, N) - F(M, N^*).$

3. Большие матрицы в задачах акустики.

3.1. МЕТОД ЗАМКНУТЫХ ДИСКРЕТНЫХ ВИХРЕВЫХ РАМОК.

Как было показано в предыдущем параграфе, решение задачи (1)—(3) сводится к решению гиперсингулярного интегрального уравнения (14) для случая, когда поверхность σ_1 является плоскостью z=0. Очевидно, что для произвольной гладкой поверхности σ_2 уравнение (14) необходимо решать численно. Воспользуемся МЗДВР [15, 17]. Для численного решения зададим поверхность σ параметрически $\sigma=\sigma(\xi,\zeta)$, где ξ и ζ — параметры, заданные на единичном квадрате $\Pi=\{0\leqslant \xi<1,0\leqslant \zeta<1\}$. В соответствии с МЗДВР [15, 17] разобьем поверхность равномерно по параметрам на $s_1\cdot s_2$ частей. При этом s_1+1 — число координатных линий $\xi_i=$ const, $i=1,2,\ldots,s_1$, а s_2+1 — число линий $\zeta_i=$ const, $i=1,2,\ldots,s_2$. Точи пересечения координатных линий обозначим через $M_{ij}^0\left(x_{ij}^0,y_{ij}^0,z_{ij}^0\right)$ Таким образом поверхность σ состоит из $s_1\cdot s_2$ криволинейных ячеек, которые имеют вершины

 $M_{i-1\,j-1}^0, M_{i\,j-1}^0 M_{i\,j}^0, M_{i-1\,j}^0$. Расчетную точку M_{ij} поместим на пересечении отрезков, соединяющих середины противоположных боковых сторон соответствующей ячейки.

Положим, что величины $g_{ij} = g(M_{ij})$ постоянны для каждой ячейки, и уравнение (14) записывается в расчетных точках M_{ij} :

$$\sum_{i=1}^{s_1} \sum_{j=1}^{s_2} g_{ij} \Phi_{ij}(M_{lm}) = -f(M_{lm}), l = 1, \dots, s_1, m = 1, \dots, s_2,$$
 (15)

где f(M) — правая часть уравнения (14), а $\Phi_{ij}(M)$ вычисляется по формуле [17], являющейся следствием формулы Стокса

$$\Phi_{ij}(M) = -\int_{L_{\sigma_{ij}}} \vec{n}_M \cdot \left(\overrightarrow{dl} \times \nabla_N G(M, N) \right) + \\
+ k^2 \int_{\sigma_{ij}} G(M, N) \vec{n}_M \cdot \vec{n}_N d\sigma. \tag{16}$$

3десь $L_{\sigma_{ij}}$ – граница ячейки σ_{ij} .

Если мы вычислим функцию g(N) для следующих параметров задачи: $\nu=4\Gamma$ ц, $z_Q=50$ м, $c=1450\frac{\rm M}{\rm C}$, то учитывая осесимметричность задачи (то есть $g(M)=g\left(\sqrt{x_M^2+y_M^2}\right)=g(r_M)$) получим графики зависимости действительной (а) и мнимой (b) частей $g(r_M)$, изображенные на рис. 2. Из рисунка видно, что искомая функция является осциллирующей функцией. Для аппроксимации осциллирующей функции кусочно-постоянными функциями требуется около 10 ячеек на период. То есть при данных параметрах задачи для получения g(M) на расстояниях 750 м необходимо решить систему порядка 400 уравнений. Однако в случае увеличения частоты до $\nu=30\Gamma$ ц возникает необходимость в решении $3\cdot 10^4$ уравнений. В акустике моря требуется исследовать поля с еще более высокими частотами, вплоть до 1000-2000 Γ ц.

3.2. Вычисление элементов матриц. Как вычислять $\Phi_{ij}(M_{lm})$? Как видно из выражения (16), $\Phi_{ij}(M_{lm})$ представляется в виде суммы контурного интеграла и интеграла по площадке, которые не вычисляются аналитически, поэтому их нужно вычислять численно. Контурный интеграл будем вычислять по формуле прямоугольников. Интегралы по площадке имеют вид (так

как площадка плоская, то $\vec{n}_N \vec{n}_M = \text{const}$):

$$\int_{\sigma_{ij}} \frac{e^{ikr_{MN}}}{r_{MN}} d\sigma = \int_{\sigma_{ij}} \frac{\cos kr_{MN}}{r_{MN}} d\sigma + \int_{\sigma_{ij}} \frac{\sin kr_{MN}}{r_{MN}} d\sigma.$$

Функция $\sin kr_{MN}/r_{MN}$ является гладкой. Для вычисления интеграла от нее заменим интегрирование по площадке σ_{ij} на интегрирование по двум треугольникам с вершинами в точках $(M^0_{i-1\,j-1},M^0_{i-j},M^0_{ij})$ и $(M^0_{i-1\,j},M^0_{ij},M^0_{ij-1})$. Для вычисления интегралов по треугольнику воспользуемся квадратурными формулами типа Гаусса порядка p. В наших вычислениях p бралось равным p, и число точек для каждого треугольника равнялось p.

Интеграл от $\cos kr_{MN}/r_{MN}$ запишем в виде

$$\int_{\sigma_{ij}} \frac{\cos kr_{MN}}{r_{MN}} d\sigma = \int_{\sigma_{ij}} \frac{(\cos kr_{MN} - 1)}{r_{MN}} d\sigma + \int_{\sigma_{ij}} \frac{1}{r_{MN}} d\sigma,$$

где $(\cos kr_{MN}-1)/r_{MN}$ — гладкая функция, интеграл от которой вычислим по квадратурной формуле типа Гаусса, а интеграл от $1/r_{MN}$ по треугольнику вычисляется аналитически.

Выделение главной части (1/r) и использование квадратурных формул высокого порядка позволяет вычислять элементы матрицы с достаточной точностью. Это необходимо по следующим причинам:

• Матрица A системы оказывается плохо обусловленной, поэтому ошибка в элементах матрицы приводит к ошибке в решении

$$\triangle x \sim \operatorname{cond}(A)\varepsilon$$

и при больших n нет внутренней сходимости, решение "разбалтывается".

• Для вычисления интегралов типа

$$\int_{\sigma_{ij}} \frac{\cos k r_{MN}}{r_{MN}} d\sigma$$

по простейшим квадратурным формулам (таким как формула прямоугольников) для достижения необходимой точности

нужно брать много точек (порядка 100^2 вблизи особенности). Это очень сильно замедляет вычисление элементов матрицы.

4. НЕЛИНЕЙНЫЕ АППРОКСИМАЦИИ МАТРИЦ

4.1. Мозаично-скелетонный метод. Первым этапом мозаично-скелетонного метода является построение дерева кластеров. Кластер — это набор точек, объединенных геометрической близостью. Практически дерево кластеров можно получить следующим образом: поместить в его корне все множество точек сетки и на каждом этапе строить дочерние кластеры, разрезая уже существующие на две части с помощью разумного метода разделения множества точек. Поскольку в рассматриваемой задаче точки сетки расположены равномерно, разумно на каждом этапе деления разрезать кластеры строго пополам (см рис. 1).

По дереву кластеров мозаичное разбиение матрицы. Для этого в матрице производится перестановка строк и столбцов (другими словами, перенумерация неизвестных) с тем, чтобы элементы, соответствующие одному кластеру, составляли в векторе неизвестных неразрывный подвектор. В результате элементы матрицы, отвечающие взаимодействию двух кластеров точек, образуют после перестановки неразрывный блок. При построении мозаичного биения мы начинаем с полной матрицы и последовательно дробим ее на блоки в согласии с дроблением неизвестных на кластеры на каждом шаге построения дерева кластеров. При этом если полученный блок отвечает хорошо отделенным в пространстве кластерам (обычно такие блоки лежат вдали от главной диагонали), то его дробление прекращается, иначе он дробится дальше. Пример мозаично-скелетонного разбиения представлен на рисунке 2.

Поскольку элементы каждого блока мозаичного разбиения отвечают взаимодействию точкек, принадлежащих достаточно удаленным кластерам, а зависимость элементов матрицы от расстояния между взаимодействующими точками — гладкая функция, то каждый блок можно приблизить матрицей малого ранга, найденной с помощью метода неполной крестовой аппроксимации [23, 24]. При заданной точности аппроксимации ε каждый блок B прибли-

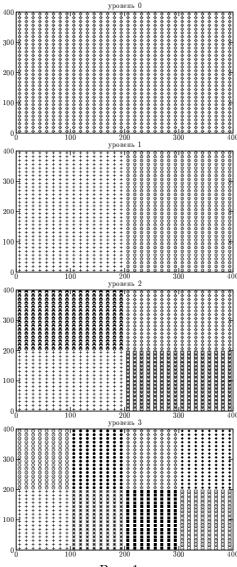


Рис. 1.

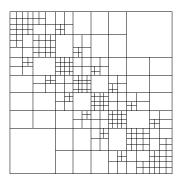


Рис. 2.

жается матрицей ранга k согласно критерию

$$\|B - \tilde{B}\|_F \leqslant \text{ерs} \|\tilde{B}\|_F$$
, где $\tilde{B} = \sum_{p=1}^k c_p r_p^T$. (17)

При этом для хранения блока \tilde{B} размера $l \times m$ нам нужно уже не ml, а только k(m+l) чисел. По аналогии с этой формулой, мозаичным рангом mr называют величину

$$mr = \frac{\text{mem}(A)}{2n} \tag{18}$$

где mem(A) — суммарный объем памяти, необходимый аппроксимаций всех блоков матрицы, n — размер матрицы.

Верхние оценки на мозаичный ранг представлены в работе [25]. Поскольку в качестве константы в них входят константы Липшица от функций, определяющих элементы матрицы, стоит ожидать, что мозаичный ранг будет расти при увеличении волнового числа k.

4.2. Тензорные аппроксимации и вейвлет-преобразования. Как было показано, при дискретизации уравнения (14) получаются большие системы уравнений с размером матриц $p^2 \times p^2 \sim 10^5 \times 10^5$. Эти матрицы плотные и не имеют какой-либо специальной структуры. Обычные методы решения линейных систем (такие как метод Гаусса) требуют $O(n^3)$ операций, что очень много. Однако матрицу такого размера невозможно даже сохранить

в оперативной памяти обычного компьютера, поэтому необходимо применение специальных быстрых методов аппроксимации матриц и решения линейных систем. Рассматриваемые матрицы являются двухуровневыми, и поэтому для них можно строить аппроксимации вида

$$A \approx A_r = \sum_{k=1}^r U_k \otimes V_k \tag{19}$$

 $([u_{kl}] \otimes V$ — блочная матрица вида $[u_{kl}V]$). Такие аппроксимации называются тензорными аппроксимациями, а r — тензорным рангом матрицы A_r . Минимальное число r, для которого существует матрица A_r тензорного ранга r такая, что $||A-A_r|| \leqslant \varepsilon ||A||$, будем называть тензорным ε -рангом. Такие аппроксимации эффективны, только если $r \ll n$. Тогда для хранения матрицы A в виде (19) нам нужно только 2rn чисел, что намного меньше, чем n^2 . Действительно ли $r \ll n$? Для класса матриц, порождённых асимптотически гладкими функциями, верхние оценки для r были получены в [21, 20, 19]. Однако рассматриваемые ядра не являются асимптотически гладкими, поэтому численное построение тензорных аппроксимаций представляет большой интерес. В частности, существенную роль имеет то, как тензорный ε -ранг зависит от волнового числа k.

Для построения тензорных аппроксимаций будем использовать метод неполной крестовой аппроксимации. (см. подробное описание алгоритма в [28]). Важно отметить, что в этом алгоритме используется только O(nr) элементов матрицы A (т.е. вычисляются не все элементы, что важно, т.к. вычисление элементов матрицы в нашем случае довольно трудоемкая задача).

Решение системы уравнений будем проводить с помощью итерационного метода GMRES. На каждом шаге требуется умножать матрицу на вектор. Используя представление (19), можно умножать матрицу на вектор за $O(n^{3/2})$ операций, что уже гораздо меньше, чем $O(n^2)$ операций при стандартном правиле умножения матрицы на вектор, но все еще довольно много. Для уменьшения вычислительной сложности применим к каждому фактору U_k и V_k вейвлет-преобразование [18]. Применяем вейвлет-преобразование к каждому тензорному фактору:

$$P_k = WU_kW^T, Q_k = WV_kW^T, \qquad 1 \leqslant k \leqslant r.$$

Здесь W — матрица вейвлет-преобразования порядка m; P_k и Q_k — псевдоразреженные матрицы. Выбирая подходящий порог $\tau=\tau(\varepsilon,P_k,Q_k)$ и полагая равными нулю все элементы в P_k и Q_k , которые по модулю меньше τ , мы приближаем A_r матрицей

$$A_{r, au} = (W^{-T} \otimes W^{-T})D(W \otimes W) pprox A_r, D = \sum_{k=1}^r P_k^ au \otimes Q_k^ au,$$

которая приближает A_r с необходимой точностью ε :

$$||A_r - A_{r,\tau}||_F \leqslant \varepsilon ||A_r||_F.$$

Если ν — число ненулевых элементов во всех P_k и Q_k , то матрицу A можно умножить на вектор за $O(\sqrt{n}\nu)$ операций. В качестве вейвлет-преобразований можно использовать преобразования Добеши.

Матрицы A, получающиеся при дискретизации нашего гиперсингулярного интегрального уравнения, оказываются плохо обусловленными, и поэтому итерационный метод требует много итераций. Необходимо использовать предобуславливание. В данной статье в качестве предобуславливателя будем использовать двухуровневый циркулянтный предобуславливатель. К сожалению, эксперименты показывают, что с увеличением частоты эффективность такого предобуславливателя падает.

Нужно отметить, что в рассматриваемом случае матрицы являются комплексными. Возможно несколько путей построения аппроксимаций для комплексных матриц. В нашей статье мы отдельно приближаем действительную и мнимую части матрицы A, т.е. сначала строим для них тензорные аппроксимации, а потом представляем факторы в вейвлетовском базисе.

5. Численные эксперименты

5.1. ВРЕМЯ СЧЁТА. Приведем времена аппроксимаций матрицы и решения линейных систем для TW-метода (Tensors+Wavelets), мозаично-скелететонного метода и стандартного метода (полное вычисление всех элементов матрицы и решение системы методом Гаусса-Зейделя). В качестве поверхности σ_2 была взята плоскость, параллельная плоскости Oxy и находящаяся на расстоянии 100 м. Расчет производился для следующих параметров задачи:

Таблица 1

Сетка	20×20	40×40	80×80	150×150
TW	1 c.	4 c.	16.3 c.	54 c.
Мозаично-скелетонный	20c.	2.5 м.	13.3 м.	1 ч.
Стандартный	7 м.	2 часа	> 1 дня	_

Таблица 2

n	20×20	40×40	80×80	150×150
TW	0.3 c.	1.1 c.	4.7 c.	17 c.
Мозаично скелетонный	7 c.	50 c.	4.3 м.	3.8 ч.
Стандартный	6 c.	3 м. 7 с.	∼1ч.	=

 $z_Q = 50$ м, c = 1450м/с, k = 0.04. Матрицы аппроксимировались с точностью 10^{-4} . Результаты представлены в таблицах 1 и 2.

5.2. Зависимость мозаичного ранга от волнового числа. Экспериментально исследуем зависимость мозаичного ранга от волнового числа k (табл. 3). Матрицы аппроксимировались с точностью $\varepsilon=10^{-4}$. В таблице представлены величины mr_{re} и mr_{im} — мозаичные ранги действительной и мнимой частей соответственно. Вычисления проводились на сетке 250×250 .

Из таблицы 3 видно, что с увеличением частоты мозаичный ранг увеличивается.

5.3. Зависимость тензорного ранга от волнового чи-

СЛА. Экспериментально исследуем зависимость тензорного ранга от волнового числа k (таблица 4). Матрицы аппроксимировались с точностью $\varepsilon=10^{-4},\,r_{re}$ и r_{im} - тензорные ранги действительной и мнимой частей соответственно. Вычисления проводились на сетке $250\times250.$

Таблица 3

k	0.01	0.02	0.04	0.08	0.1
mr_{re}	627	740	950	1430	1660
mr_{im}	322	506	775	1320	1590

Таблица 4

k	0.01	0.02	0.04	0.08
r_{re}	6	7	12	26
r_{im}	6	10	16	28

Таблица 5

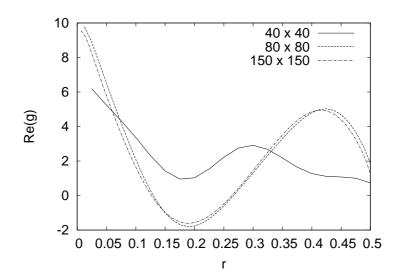
k	0.01	0.02	0.04	0.08
Число итераций	15	49	69	Нет сходимости

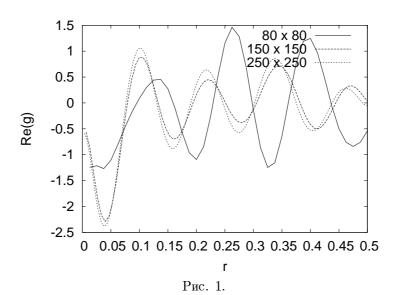
Видно, что с увеличением частоты ранг увеличивается, но не слишком значительно. Гораздо хуже становится обусловленность матрицы, что будет проиллюстрированно в следующем пункте.

- **5.4.** Число итераций в зависимости от волнового числа. Экспериментально исследуем зависимость числа итераций от волнового числа k при использовании предобуславливателя и без него (табл. 5 и 6). Матрицы аппроксимировались с точностью $\varepsilon=10^{-4}$. Вычисления проводились на сетке 150×150 .
- **5.5.** Внутренняя сходимость. Проиллюстрируем внутреннюю сходимость выбранного метода дискретизации для k=0.02,0.04. Воспользуемся тем, что задача радиально-симметрична, и будем откладывать по осям Re(g) и r (рис. 1).
- 6. Выводы и дальнейшая работа Мы показали, что используя методы нелинейной аппроксимации больших матриц, можно очень быстро решать системы линейных уравнений, возникающих при дискретизации уравнений акустики (в данном случае задачи распространения звука в мелком море) на сетках гораздо более мелких, чем это было возможно до этого. Для эффек-

Таблица 6

k	0.01	0.02	0.04	0.08
Число итераций	46	194	246	Нет сходимости





тивного вычисления элементов образующихся матриц используется выделение главной части, а сами интегралы вычисляются по квадратурным формулам типа Гаусса высокого порядка для треугольника.

Для разных частот показана внутренняя сходимость приближенного решения и экспериментально обоснована необходимость использования мелких сеток для решения задачи с приемлемой точностью. Однако было выявлено некоторое количество трудностей и проблем, которые требуют дальнейшего исследования. В частности:

- Получающиеся матрицы очень плохо обусловлены, причём обусловленность тем хуже, чем выше частота. Поэтому необходимо использовать предобуславливание. Предварительные эксперименты с разными предобуславливателями не дали хороших результатов. Возможно использование в качестве предобуславливателей приближенных обратных, построенных с помощью метода Ньютона [27].
- При увеличении частоты необходимо вычислять интегралы от сильно осциллирующих функций, что является довольно сложной задачей. Возможно, нужно разрабатывать специальные квадратурные формулы.

Нуждается в исследовании вопрос о возможности использования неравномерных сеток, в частности метода коллокации по чебышевским сеткам, предложенного для другого интегрального уравнения [27], где была экспериментально показана исключительная эффективность такого метода.

Эффективность используемых методов ухудшается с увеличением частоты. Поэтому для высоких частот необходимо использовать не кусочно-постоянные, а какие-то другие элементы. В частности, имеет смысл искать решение в виде

$$g = (p_Q(M) - p_Q^*(M))\widetilde{g}.$$

Выяснилось, что функция \tilde{g} оказывается не столь осциллирующей, поэтому возможно использование гораздо более грубых сеток (однако сохраняется необходимость вычисления интегралов от осциллирующих функций; для не очень больших размеров сеток их можно вычислять по адаптивным квадратурным формулам).

С практической точки зрения гораздо бо́льший интерес представляет решение задачи не с плоской, а с какой-то другой гладкой поверхностью дна, (например, при наличии на дне горы). Возможно ли применение быстрых методов для такого случая — это открытый вопрос, однако предварительные численные эксперименты для похожих задач позволяют на это надеяться. Работа по улучшению наших результатов продолжается.

Литература

- [1] Alfredson R.J. A Note on the Use of the Finite Difference Method for Predicting Steady State Sound Fields // Acoustica. Vol. 28. 1973. P. 296-301.
- [2] Arlett P.L., Bahrani, Zienkiewicz O.C. Application of FEM to Solution of Helmholtz's Equation // Proc. IEEE. 1968. P. 115-118.
- [3] Belov A.I., Serebryany A.N., Zhuravlev V.A. Observation of internal wave effects on acoustic and oceanographic field study in the Middle Atlantic Bight // Proceed. of the International conference on shallow water acoustics. Beijing, 1997.
- [4] Burton A.J., Miller G.F. The application of integral equation methods to the numerical solution of some exterior boundary value problems // Proc. R. London. Ser 1971. P. 201-210.
- [5] Caille G.W., Dahl P.H., Zijun Gan et al. Overview of the joint China-U.S. "Yellow Sea '96" experiment // Proceed. of the International conference on shallow water acoustics. Bejing, 1997.
- [6] Chien C.C., Rajiyah H., Atluri S.N. An effective method for solving the hypersingular integral equations in 3-D acoustics // J. of the Acoust. Society of Am. 1990. Vol 88. №2.
- [7] Gladwell G.M.L. A Finite Element Method for Acoustics. Proceedings of the 5th International Congress on Acoustics. L33. 1965. P. 1-4.

- [8] Kaya A.C., Erdogan F. On the Solution of Integral equations with Strongly Singular Kernels // Q. Appl. Math. XLV. 1987. P. 105-122.
- [9] Krishnasamy G., Schmerr L.W., Rudolphi T.J., Rizzo F.J. Hyper Singular Boundary Integral Equations: Some Applications in Acoustics and Elastic Wave Scattering // ASME J. App. Mech. 1989. P. 227-237.
- [10] Seybert A.F., Soenarko B., Rizzo F.J., Shippy J.F. A Special Integral Equation Formulation for Acoustic Radiation and Scattering for Axi-symmetric Bodies and Boundary Conditions // J. Acoust. Soc.Am. 1986. Vol 80. P. 1241-1247.
- [11] Seybert A.F., Rengarajan The Use of CHIEF to Obtain Unique Solu-tions for Acoustic Radiation Using Boundary Integral Equations // J. Acoust. Soc. Am. 1987. Vol. 81. P. 1299-1306.
- [12] Андреев М.Ю., Кацнельсон Б.Г., Кулапин Л.Г., Петников В.Г. Влияние гидродинамической изменчивости на фазу звукового поля в мелком море // Акуст. журнал. 1996. Т.42, №4. С.459-464.
- [13] Бондарь Л.Ф., Борисов С.В, Гриценко А.В. Результаты исследований флуктуаций интенсивности и фазы акустических сигналов на стационарных трассах и шельфовой зоне Японского моря // Акуст. журнал. 1994. Т. 40, №4, С.561-570
- [14] Вайникко Г.М., Лифанов И.К., Полтавский Л.Н. Численные методы в гиперсингулярных интегральных уравнениях и их приложения. М.: ТОО Янус, 2001. 510 с.
- [15] Довгий С.А., Лифанов И.К. Методы решения интегральных уравнений. Киев: Наукова думка, 2002. 344 с.
- [16] Колтон Д., Кресс Р. Методы интегральных уравнений в теории рассеивания. М.: Мир, 1987. 311с.
- [17] Лифанов И.К. Метод сингулярных интегральных уравнений и численный эксперимент. М.: ТОО Янус, 1995 520с.

- [18] I. Daubechies. Ortonormal bases of compactly supported wavelets. // Communications of Pure and Applied MAth. 41(7):909-996, 1988.
- [19] Hackbush W., Khoromskij B.N. Tyrtyshnikov E.E. Hierarchical Kronecker tensor-product approximations: Preprint №. 35. Leipzig: Max-Planck Institut f. Math. in Naturwissenschaften. 2003.
- [20] Tyrtyshnikov E.E. Kronecker-product approximations for some function-related matrices // Linear Algebra Appl. 2004. Vol. 379. P. 423–437.
- [21] Тыртышников Е.Е. Тензорные аппроксимации матриц, порождённых асимптотически гладкими функциями // Матем. сб. 2003. Т. 194, б. С. 147–160.
- [22] Горейнов С. А., Замарашкин Н. Л., Тыртышников Е. Е. Псевдоскелетные аппроксимации матриц // Доклады РАН. 1995. Т. 343 2 Р. 151–152.
- [23] Tyrtyshnikov E. E. Mosaic-skeleton approximations // Calcolo. 1996. Vol. 33 1-2. P. 47-57.
- [24] Tyrtyshnikov E. E. Incomplete cross approximation in the mosaic-skeleton method // Computing 2000. Vol.64 4. P. 367–380.
- [25] Горейнов С. А. Мозаично-скелетонные аппроксимации// Матричные методы и вычисления. (Сборник научных трудов под редакцией Е. Е. Тыртышникова). 1999. ИВМ РАН.
- [26] Ford J. M., Oseledets I. V., Tyrtyshnikov E. E. Matrix approximations and solvers using tensor products and non-standard wavelet transforms related to irregular grids // Rus. J. Numer. Anal. and Math. Modelling. 2004. Vol. 19. 2. P. 185-204.
- [27] Оселедец И.В., Тыртышников Е.Е. Приближённое обращение матриц при численном решении гиперсингулярного уравнения // ЖВМиМФ. 2005. т. 45 2 с. 315-326
- [28] Ford J. M., Tyrtyshnikov E. E. Combining Kronecker product approximation with discrete wavelet transforms to solve dense,

function-related systems // SIAM J. Sci. Comp. 2003. Vol. 25. 3 P. 961–981.

СОДЕРЖАНИЕ

Предисловие
Часть І. Методы решения задач вариационного усвоения данных наблюдений и управления сложными системами
Агошков В. И., Залесный В. Б., Минюк Ф. П., Русаков А. С. Восстановление внешних источников в задачах теплообмена
в океане 7 Ипатова В. М. Об одной методике усвоения данных наблюдений 61
Шутяев В. П., Пармузин Е. И. Задача восстановления начального условия в локально-одномерной модели вертикального теплообмена
Часть II. Матричные методы в интегральных и диффе-
ренциальных уравнениях
Василевский Ю. В., Чугунов В. Н. О параллельном многосеточном решении консервативных неструктурированных конечно- элементных систем
Лапласиана на последовательностях сеток: экспериментальны оценки вычислительной работы
Василевский Ю. В., Капранов С. А. Параллельное моделирование особенностей кровотока в окрестностях кава-фильтра с захваченным тромбом
Савостьянов Д. В. Паралельная реализация метода решения объемного интегрального уравнения электродинамики на
основе многоуровневых теплицевых матриц
Оселедец И. В., Савостьянов Д. В., Ставцев С. Л. Применение нелинейных методов аппроксимаций для быстрого решения задачи распространения звука в мелком море

Научное издание

МЕТОДЫ И ТЕХНОЛОГИИ РЕШЕНИЯ БОЛЬШИХ ЗАДАЧ

 $\begin{picture}(2000) \put(0,0){\line(0,0){100}} \put(0,0){\line(0,0){1$

Институт вычислительной математики Российской академии наук 119991 Москва, ул. Губкина, д. 8.

Редактор Лебедева В.В.

Лицензия ИД №03991 от 12.02.2004.

Оригинал-макет изготовлен в ИВМ РАН
Подписано в печать 15.12.2004 г.

Формат 60 × 90 № Печать офсетная. Бумага офсетная №1
Печ. л. 12.252.25. Тираж 200 экз. Заказ
Отпечатано согласно предоставленному оригинал-макету
в ФГУП "Производственно-издательский комбинат ВИНИТИ".
140010 г. Люберцы Московской обл., Октябрьский пр-т, 403.

Тел. 554-21-86