Записки научных
семинаров ПОМИ
Том 284, 2002 г.

А. Ю. Еремин, А. А. Никишин

ПЕРЕОБУСЛАВЛИВАНИЕ ЛИНЕЙНЫХ СИСТЕМ С НЕСИММЕТРИЧНЫМИ МАТРИЦАМИ С ПОМОЩЬЮ ФАКТОРИЗОВАННЫХ РАЗРЕЖЕННЫХ ПРИБЛИЖЕННЫХ ОБРАТНЫХ

1. Введение

Во многих приложениях возникает необходимость итерационного решения систем линейных алгебраических уравнений с несимметричными, возможно, незнакоопределенными матрицами большой размерности. Эффективное решение таких систем на параллельных компьютерах требует построения (качественных) переобуславливателей, обладающих достаточным ресурсом параллелизма. Многие популярные переобуславливатели обладают слабым ресурсом параллельности как на стадии построения, так и на стадии применения в итерационном алгоритме. В последние годы все большее внимание привлекают к себе методы переобуславливания, обладающие "естественным" параллелизмом, особенно методы явного переобуславливания. В контексте параллельных вычислений явные методы имеют неоспоримое преимущество перед методами неявного переобуславливания. Если реализация последних требует на каждой итерации решения линейных систем с матрицей переобуславливателя, то для реализации явных методов необходимо лишь умножать переобуславливатель на вектор, что, с точки зрения параллельных вычислений, является гораздо более эффективной операцией. Наиболее широко известными и изученными среди явных переобуславливателей, применяемых для решения линейных систем с несимметричными незнакоопределенными матрицами, являются разреженные приближенные обратные типа SAI и AINV (см., например, [1, 2, 3, 15]).

В случае линейных систем с симметричной положительно

Работа выполнена при финансовой поддержке NWO, Нидерланды (грант 047-008-007).

определенной матрицей коэффициентов хорошо зарекомендовали себя переобуславливания с помощью факторизованных разреженных приближенных обратных типа FSAI [7, 11]. Суть этого подхода состоит в вычислении разреженной аппроксимации G к матрице, являющейся обратной для сомножителя Холецкого L_A исходной матрицы A, на которой норма Фробениуса $\|I - GL_A\|_F$ достигает минимума по множеству невырожденных треугольных матриц предписанной структуры разреженности. Хотя факторизованные приближенные обратные уже использовались в качестве существенной компоненты при построении различных переобуславливателей для несимметричных матриц (см., например, [12]), авторам не известно ни одной работы, в которой были бы даны подробное описание и функциональная интерпретация этого подхода для произвольной несимметричной невырожденной матрицы. В данной работе мы представим такое описание для несимметричных матриц, допускающих треугольное разложение. Мы также покажем на численных экспериментах, что пред и постфильтрация FSAI-переобуславливателей позволяет повысить их эффективность также и в несимметричном случае.

Работа построена следующим образом. В разделе 2 описан процесс построения сомножителей FSAI-переобуславливателя в случае невырожденной несимметричной матрицы. В разделе 3 дается функциональная интерпретация FSAI-переобуславливания. В разделе 4 описаны выбор структуры разреженности FSAI-переобуславливателя и процедура постфильтрации полученного переобуславливателя. Предфильтрация структуры ненулевых элементов, осуществляемая с помощью агрегирования, описана в разделе 5. Численные результаты для тестовых задач различного происхождения и заключительные замечания приведены в разделе 6.

2. Построение несимметричных факторизованных разреженных приближенных обратных

Пусть $A = (A_{ij}) \in \mathbb{R}^{n\times n}$ – произвольная невырожденная, возможно незнакоопределенная и несимметричная, матрица, допускающая треугольное разложение

$$A = LU.$$

Для некоторой выбранной треугольной структуры разреженности S_L, такой что

$$\{(i,j) : i < j\} \subseteq S_L \subseteq \{(i,j) : 1 \leqslant i \neq j \leqslant n\},$$

построим нижнюю треугольную матрицу $\bar{G}_L = (\{\bar{G}_L\}_{ij})$ и верхнюю треугольную матрицу $\bar{G}_U = (\{\bar{G}_U\}_{ij})$ по следующим формулам [10]:

$$\begin{cases} \{\bar{G}_L A\}_{ij} = \delta_{ij}, & (i,j) \notin S_L, \\ \{\bar{G}_L\}_{ij} = 0, & (i,j) \in S_L, \end{cases} \quad \text{и} \quad \begin{cases} \{A\bar{G}_U\}_{ij} = \delta_{ij}, & (j,i) \notin S_L, \\ \{\bar{G}_U\}_{ij} = 0, & (j,i) \in S_L. \end{cases} \tag{2.1}$$

Легко видеть, что ненулевые элементы i-той строки матрицы $\bar{G}_L$ могут быть найдены, как решение системы линейных уравнений с матрицей $A^{(i)} \equiv A[\xi^{(i)}]$, где $A[\xi^{(i)}]$ обозначает главную подматрицу исходной матрицы A, состоящей из элементов A_{kj} с индексами k и j такими, что $k,\ j \in \xi^{(i)} \equiv \{l : (i,l) \notin S_L\}$. Аналогичное утверждение верно и для ненулевых элементов j-того столбца матрицы $\bar{G}_U$. Очевидно, что элементы матриц $\bar{G}_L$ и $\bar{G}_U$ будут единственным образом определены, если все подматрицы $A^{(i)}$, $i = 1, \ldots, n$, будут невырожденными, более того, в этом случае диагональные элементы матриц $\bar{G}_L$ и $\bar{G}_U$ совпадают.

Пусть диагональные элементы $\{\bar{G}_L\}_{ii}$ и $\{\bar{G}_U\}_{ii}$, $i = 1, \ldots, n$, матриц $\bar{G}_L$ и $\bar{G}_U$ не равны нулю. Определим нижний треугольный сомножитель G_L FSAI-переобуславливателя по формуле

$$G_L = \operatorname{diag}(|\{\bar{G}_L\}_{11}|^{-1/2}, \ldots, |\{\bar{G}_L\}_{nn}|^{-1/2})\, \bar{G}_L.$$

Аналогично, верхний треугольный сомножитель G_U определяется соотношением

$$G_U = \bar{G}_U \operatorname{diag}(\operatorname{sign}(\{\bar{G}_U\}_{11})|\{\bar{G}_U\}_{11}|^{-1/2}, \ldots, \\ \operatorname{sign}(\{\bar{G}_U\}_{nn})|\{\bar{G}_U\}_{nn}|^{-1/2}),$$

где $\operatorname{sign}(a)$ означает знак числа a. Легко показать, что в этом случае переобусловленная матрица $G_L A G_U$ будет отмасштабирована по Якоби, то есть

$$\{G_L A G_U\}_{ii} = 1, \quad i = 1, \ldots, n.$$

Замечание. Если все локальные матрицы $A^{(k)}$, $k = 1, \ldots, n$, невырождены, то i-ый диагональный элемент матрицы $\bar{G}_L$ (матрицы

$\bar{G}_U$) обращается в ноль только в том случае, когда последний диагональный элемент матрицы $(A^{(i)})^{-1}$ равен нулю. Чтобы избежать подобной ситуации, на практике можно использовать различные виды диагональной коррекции локальной матрицы (например, использовать диагональный сдвиг $A^{(i)} + \tau I$, где τ – малый параметр). В этом случае матрицы $\bar{G}_L$ и $\bar{G}_U$ не будут уже в точности удовлетворять системам уравнений (2.1), но будут гарантированно невырожденными.

3. Функциональная интерпретация факторизованных приближенных обратных для несимметричной матрицы

Теорема 3.1. *Пусть $A \in \mathbb{R}^{n\times n}$ – произвольная вещественная матрица, для которой существует точное треугольное разложение $A = LU$. И пусть заданы не содержащие диагональных элементов структуры разреженности S_X и S_Y матриц X и $Y \in \mathbb{R}^{n\times n}$, такие что*

$$|\xi^{(i)}| = |\eta^{(i)}|, \quad i = 1, \dots, n,$$

где $|\xi^{(i)}|(|\eta^{(i)}|)$ – число элементов в множестве $\xi^{(i)}(\eta^{(i)})$ индексов ненулевых элементов i-ой строки X (i-ого столбца Y). Тогда нахождение экстремума билинейного функционала

$$F(X,Y) = tr((I - XL)(I - UY)) \tag{3.2}$$

с ограничениями в виде равенств

$$X|_{S_X} = 0 \quad и \quad Y|_{S_Y} = 0 \tag{3.3}$$

эквивалентно решению следующих систем линейных уравнений:

$$\begin{cases} \{XA\}_{ij} = U_{ij}, & (j,i) \notin S_Y, \\ X_{ij} = 0, & (i,j) \in S_X, \end{cases} \quad и \quad \begin{cases} \{AY\}_{ij} = L_{ij}, & (j,i) \notin S_X, \\ Y_{ij} = 0, & (i,j) \in S_Y. \end{cases} \tag{3.4}$$

Замечание. Если все подматрицы $A^{(i)} = A[\xi^{(i)}, \eta^{(i)}]$, $i = 1, \dots, n$, невырождены, то решение упомянутых выше задач существует и единственно.

Доказательство. Легко показать, что

$$\left.\frac{\partial F}{\partial X}\right|_{\bar{S}_X} = 0 \quad \Leftrightarrow \quad \forall\ (j,i) \notin S_X \quad \{AY\}_{ij} = L_{ij}.$$

Здесь и далее $\bar{S}_X$ обозначает дополнение к множеству S_X.

Действительно, $F(X,Y) = \mathrm{tr}\,((I - XL)(I - UY)) = \mathrm{tr}\,(I - XL - UY + XAY)$ и

$$\frac{\partial F}{\partial X} = \frac{\partial\ \mathrm{tr}\,(I - XL - UY + XAY)}{\partial X} = -L^T + Y^T A^T.$$

Выписывая эти равенства поэлементно, получаем

$$\forall\ (j,i) \notin S_X \quad \left\{\frac{\partial F}{\partial X}\right\}_{ji} = -L_{ij} + \{YA\}_{ij}.$$

Аналогично,

$$\left.\frac{\partial F}{\partial Y}\right|_{\bar{S}_Y} = 0 \quad \Leftrightarrow \quad \forall\ (j,i) \notin S_Y \ \ \{AX\}_{ij} = U_{ij},$$

так как

$$\frac{\partial F}{\partial Y} = -U^T + A^T X^T,$$

и, следовательно,

$$\forall\ (j,i)\ \notin S_Y \quad \left\{\frac{\partial F}{\partial Y}\right\}_{ji} = -U_{ij} + \{XA\}_{ij}.$$

Добавляя ограничения (3.3), окончательно получаем полные системы линейных уравнений (3.4) для нахождения элементов матриц X и Y соответственно. □

Следствие 3.1. *Пусть структуры разреженности S_X и S_Y удовлетворяют следующим дополнительным условиям:*

$$(i,j) \in S_X \quad \Leftrightarrow \quad (j,i) \in S_Y \tag{3.5}$$

и

$$\{(i,j) : i < j\} \subseteq S_X \subseteq \{(i,j) : 1 \leqslant i \neq j \leqslant n\}. \tag{3.6}$$

Тогда матрицы $\bar{G}_L$ и $\bar{G}_U$, определяемые для невырожденной вещественной матрицы A, совпадают с точностью до диагонального масштабирования с решением экстремальной задачи для билинейного функционала (3.2) с ограничениями (3.3).

Доказательство. Для доказательства следствия 3.1 достаточно заметить, что, налагая дополнительно условия (3.5), мы получим

$$\begin{cases} \{XA\}_{ij} = U_{ij}, & (i,j) \notin S_X, \\ X_{ij} = 0, & (i,j) \in S_X, \end{cases} \quad \text{и} \quad \begin{cases} \{AY\}_{ij} = L_{ij}, & (j,i) \notin S_X, \\ Y_{ij} = 0, & (j,i) \in S_X. \end{cases}$$

□

Следствие 3.2. *Пусть A – вещественная симметричная положительно определенная матрица. Потребуем дополнительно к условиям теоремы 3.1, чтобы $X \equiv Y^T$. Тогда функционал (3.2) может быть представлен в виде неотрицательного квадратичного функционала от X:*

$$\mathcal{F}(X) \equiv F(X, X^T) = ||I - XL||_F^2.$$

Более того, решение задачи минимизации

$$\min \mathcal{F}(X) \quad \text{с ограничением } X|_{S_X} = 0$$

эквивалентно нахождению неотмасштабированного сомножителя FSAI-переобуславливателя [10] симметричной положительно определенной матрицы.

Доказательство.

$$\mathcal{F}(X) = \mathrm{tr}\,((I - XL)(I - L^T X^T)) = \|I - XL\|_F^2.$$

С другой стороны, так как $\mathcal{F}(X) = \mathrm{tr}\,(I - XL - L^T X^T + XAX^T)$, получаем

$$\left.\frac{\partial \mathcal{F}}{\partial X}\right|_{\bar{S}_X} = 0 \quad \Leftrightarrow \quad \{XA\}_{ij} = L^T_{ij}, \quad \forall\,(i,j) \notin S_X.$$

□

Введем новые переменные:

$$\left\{ \begin{array}{l} \Omega_+ = \frac{1}{2}(XL + Y^T U^T), \\ \Omega_- = \frac{1}{2}(XL - Y^T U^T), \end{array} \right. \quad \text{или} \quad \left\{ \begin{array}{l} X = (\Omega_+ + \Omega_-)L^{-1}, \\ Y = U^{-1}(\Omega_+ - \Omega_-)^T. \end{array} \right. \tag{3.7}$$

Легко показать, что после замены переменных (3.7) билинейный функционал (3.2) принимает канонический гиперболический вид:

$$F(\Omega_+, \Omega_-) = \|I - \Omega_+\|_F^2 - \|\Omega_-\|_F^2.$$

4. Выбор структуры разреженности и постфильтрация

Как видно из формул (2.1), сомножители FSAI-переобуславливателя полностью определяются структурой разреженности S_L. В случае симметричной положительно определенной матрицы известно довольно много стратегий выбора структуры разреженности S_L [4, 5, 11, 18, 20]. Наиболее простой из них является выбор в качестве S_L структуры разреженности нижнего треугольника исходной матрицы A. Аналогично, в случае произвольной матрицы простейшим выбором структуры разреженности будет структура разреженности нижнего треугольника симметризованной матрицы $A+A^T$. Однако, как и в случае симметричных положительно определенных матриц, можно ожидать, что такой выбор структуры разреженности будет приводить, как правило, к построению FSAI-переобуславливателя заведомо низкого качества. Для того чтобы улучшить качество вычисляемого FSAI-переобуславливателя, нам необходимо увеличить число ненулевых элементов в сомножителях G_L и G_U. Увеличить число ненулевых элементов в сомножителях можно, например, выбрав в качестве структуры разреженности S_L структуру разреженности нижнего треугольника матрицы $(A + A^T)^p$, где p – некоторая положительная степень. Чем большую степень p мы будем использовать, тем точнее мы будем приближать факторы точного треугольного разложения, и тем выше будет качество вычисляемого FSAI-переобуславливателя. К сожалению, в большинстве случаев результирующий FSAI-переобуславливатель становится слишком плотным, и быстрый рост затрат на умножение переобусловленной матрицы на вектор не компенсируется уменьшением числа итераций основной итерационной схемы.

Число вычисляемых ненулевых элементов сомножителей FSAI-переобуславливателя при таком выборе структуры разреженности S_L можно сократить, если использовать структуру разреженности нижнего треугольника матрицы $(\tilde{A}+\tilde{A}^T)^p$, где матрица $\tilde{A}$ получена путем отбрасывания небольших по модулю элементов исходной матрицы A [4, 5]. Однако такой подход невозможно эффективно применить в том случае, когда большая часть элементов исходной матрицы варьируется в небольших пределах. Другим недостатком этого подхода является тот факт, что свойства исходной матрицы могут достаточно сильно видоизменяться после отбрасывания даже сравнительно маленьких по модулю элементов, и в этом случае качество FSAI-переобуславливателя такой структуры разреженности будет плохо предсказуемым.

Другим хорошо известным подходом к уменьшению затрат на итерационную фазу является постфильтрация построенного переобуславливателя (в нашем случае сомножителей переобуславливателя) по некоторому фиксированному порогу ε. Такой подход прекрасно себя зарекомендовал при использовании FSAI-переобуславливания в случае симметричной положительно определенной матрицы A. Детальное описание постфильтрации FSAI-переобуславливателя в симметричном положительно определенном случае может быть найдено в работе [19]. Мы предлагаем использовать данный подход и для произвольной матрицы A. Однако, при использовании итерационных схем типа метода обобщенных минимальных невязок GMRES, порог постфильтрации ε следует выбирать меньшим, поскольку затраты на собственно итерационную схему растут квадратично с ростом числа итераций. Заметим, что при использовании низких порогов постфильтрации перемасштабирование отфильтрованных сомножителей $\tilde{G}_L$ и $\tilde{G}_U$, описанное в работе [19], можно не проводить, так как диагональные элементы переобусловленной матрицы $\tilde{G}_L A \tilde{G}_U$ мало отличаются от единицы.

5. Предфильтрация структуры ненулевых элементов

В случае систем с симметричной положительно определенной матрицей эффективным способом минимизации тотальных арифметических затрат при итерационном решении линейных систем, переобусловленных с помощью факторизованных разреженных обратных, является предварительная фильтрация структуры ненулевых элементов (расширение структуры разреженности S_L) с помощью векторной агрегации (подробнее см. [14]). В этом разделе мы обобщим метод, описанный в работе [14], на случай произвольной разреженной матрицы.

Пусть $A = (A_{ij})_{i,j=1}^m \in \mathbb{R}^{n\times n}$, $1 \leqslant m \leqslant n$ – разреженная блочная матрица с квадратными диагональными блоками $A_{ii} \in \mathbb{R}^{n_i\times n_i}$, $i = 1,\dots,m$, такими, что $n = \sum\limits_{i=1}^m n_i$, и пусть S_A – блочная структура разреженности A, т.е.

$$S_A \subseteq \{(i,j) : 1 \leqslant i \neq j \leqslant m\},$$

и блок A_{ij} является структурно нулевым тогда и только тогда, когда $(i,j) \in S_A$. Пусть $c = (c_i)_{i=1}^m \in \mathbb{R}^n$ – ненулевой вектор разбитый на блочные компоненты в соответствии с блочным биением матрицы A, то есть, $c_i \in \mathbb{R}^{n_i}, i = 1,\dots,m$. Будем говорить, что вектор c не вырождается для данного блочного биения, если ни одна из его блочных компонент c_i не равна нулю. Легко видеть, что *"единичный"* вектор $e = (1, 1, \dots, 1)^T$ будет невырождающимся для любого блочного биения.

Пусть задан вектор c, невырождающийся для блочного биения $\mathcal{P} = \{n_1, n_2, \dots, n_m\}$, $n = \sum\limits_{i=1}^{m} n_i$, тогда с блочной матрицей A мы будем ассоциировать так называемую (векторно) агрегированную матрицу $A_c = (a^c_{ij})^m_{i,j=1} \in \mathbb{R}^{m\times m}$, элементы которой определены следующим образом:

$$a^c_{ij} = \frac{c_i^T A_{ij} c_j}{||c_i||\,||c_j||}, \quad 1 \le i, j \le m. \tag{5.8}$$

Вектор c будем называть вектором агрегации для матрицы A и блочного биения $\mathcal{P}$. Спектральные свойства агрегированной матрицы в случае симметричной матрицы A подробно описаны в работе [9].

Алгоритм предфильтрации для произвольной матрицы выглядит следующим образом:

Алгоритм предфильтрации

Шаг 1. Зафиксируем блочное биение $\mathcal{P}$ и вектор c, невырождающийся для этого блочного биения.

Шаг 2. Для блочного биения $\mathcal{P}$ и вектора c построим разреженную агрегированную матрицу $A_c = (a^c_{ij})^m_{i,j=1}$, элементы которой определяются по формуле (5.8).

Шаг 3. Вычислим сомножители FSAI-переобуславливателя G^c_L и G^c_U для агрегированной матрицы A_c по некоторой предписанной структуре разреженности S_c.

Шаг 4. Определим структуру разреженности S_δ, $S_\delta \supseteq S_c$, в соответствии с предписанным порогом предфильтрации δ по следующей формуле:

$$S_\delta = \{(i,j) : i \ne j,\ |(G^c_L)_{ij}| < \delta \text{ и } |(G^c_U)_{ji}| < \delta, 1 \leqslant i, j \leqslant m\}.$$

Шаг 5. Структура разреженности S_δ выбирается в качестве блочной структуры разреженности сомножителей FSAI-переобуславливателя для матрицы A.

Замечание 1. Здесь мы предполагаем, что агрегированная матрица обладает всеми свойствами, необходимыми для того, чтобы FSAI-переобуславливатель для нее существовал и был корректно определен. Очевидно, что даже если все главные подматрицы исходной матрицы A невырождены, в общем случае главные подматрицы агрегированной матрицы A_c не обязаны быть невырожденными.

Замечание 2. Начальная структура разреженности S_c выбирается по структуре разреженности нижнего треугольника матрицы $(A_c + A_c^T)^p$, $p \geqslant 1$.

6. Результаты численных экспериментов

В этом разделе представлены результаты численных экспериментов с системами линейных алгебраических уравнений $Ax = b$ для пяти тестовых несимметричных невырожденных матриц нерегулярной структуры разреженности, для которых в качестве переобуславливателя рассматривались факторизованные разреженные обратные с различными параметрами p, δ и ε. Первая тестовая матрица (UTM5940) возникает в процессе моделирования ТОКАМАКа и взята из коллекции матриц Matrix Market Collection [13]. Вторая тестовая матрица (RAEFSKY4) возникает при моделировании механических процессов прогибания и может быть найдена в коллекции матриц Davis Collection [6]. Для этих двух задач мы использовали искусственную правую часть $b = Ae$, где $e = (1, \ldots, 1)^T$. Остальные три модельные задачи (Tern, Osprey, PernisWest) возникают при моделировании процессов фильтрации в нефтяных резервуарах и предоставлены авторам компанией Шелл.

В качестве основной итерационной схемы для всех тестовых задач был выбран метод обобщенных минимальных невязок GMRES без рестартов [17] с нулевым начальным приближением и критерием остановки вида

$$\|r_k\|/\|r_0\| < 10^{-8}, \qquad (6.9)$$

где r_k обозначает невязку k-ой итерации метода GMRES. Блочное разбиение $\mathcal{P}$ для матриц Tern, Osprey и PernisWest соответствовало числу степеней свободы на узел; для матриц UTM5940 и RAEFSKY4 блочное разбиение было построено при помощи алгоритма аналогичного алгоритму автоматического определения блочной структуры разреженной матрицы, описанному в работе

[16]. Все матрицы, за исключением матриц Tern и RAEFSKY4, были предварительно отмасштабированы с помощью блочного метода Якоби и симметрично переупорядочены по обобщенному методу вложенных сечений (General Nested Disection) [8]. Вычисление треугольных сомножителей блочно диагональной части матрицы RAEFSKY4 было неустойчивым. Результаты для исходной матрицы Tern были существенно лучше, чем для переупорядоченной по методу GND матрицы, поэтому матрица Tern была лишь предварительно отмасштабирована с помощью блочного метода Якоби. В качестве вектора агрегации c во всех случаях был использован *"единичный"* вектор $e = (1, 1, \ldots, 1)^T$.

В таблице 1 представлены структурные характеристики тестовых матриц. Используются следующие обозначения:

N – размер линейной системы;
NZA – число ненулевых элементов исходной матрицы A;
M – число блоков в блочном разбиении $\mathcal{P}$;
MXDOF – максимальная размерность блоков блочного разбиения $\mathcal{P}$.

Таблица 1. Структурные характеристики тестовых линейных систем.

Матрица	N	NZA	M	MXDOF
UTM5940	5940	83842	2419	3
RAEFSKY4	19779	1328611	6743	3
Tern	18349	326535	6084	4
Osprey	35102	655182	11697	4
PernisWest	60248	1221300	200745	4

Таблицы 2–6 содержат результаты численных экспериментов с переобусловленными тестовыми матрицами и используют следующие обозначения:

p – степень матрицы $(A + A^T)^p$, или $(A_c + A_c^T)^p$, $p \geqslant 1$, используемой для определения структуры разреженности переобуславливателя;
δ – значение порога предфильтрации, $\delta \geqslant 0$;
ε – значение порога постфильтрации, $\varepsilon \geqslant 0$;
Dens – относительная плотность FSAI-переобуславли-

	вателя, т.е. $Dens = (NZG_L + NZG_U)/NZA$;
Constr. Costs	– арифметические затраты на построение FSAI-переобуславливателя в терминах числа умножений исходной матрицы A на вектор;
Iters	– число итераций переобусловленного метода GMRES, необходимых для удовлетворения критерия остановки;
MVM Costs	– арифметические затраты итерационной фазы на умножение переобусловленной матрицы на векторы в терминах числа умножений исходной матрицы A на вектор;
Iter. Costs	– суммарные арифметические затраты на итерационную фазу решения в терминах числа умножений исходной матрицы A на вектор
Total Costs	– суммарные арифметические затраты, т.е. Total Costs = Constr. Costs + Iter. Costs.

Таблица 2. Результаты численных экспериментов для матрицы RAEFSKY4.

p	δ	ε	$Dens$	$Constr.$ $Costs$	$Iters$	MVM $Costs$	$Iter.$ $Costs$	$Total$ $Costs$
1	–	–	1.05	79.28	61	124.79	190.48	269.76
2	–	–	4.30	1650.73	12	63.61	74.36	1725.09
1	–	$1e-8$	1.01	79.28	44	88.34	125.43	204.71
1	–	$1e-7$	1.00	79.28	53	106.15	157.23	236.51
2	–	$1e-8$	3.40	1650.73	17	74.85	86.29	1737.02
2	–	$1e-7$	3.36	1650.73	21	91.58	106.53	1757.26
1	$1e-5$	$1e-8$	1.00	80.92	46	92.17	132.16	213.08
1	$1e-4$	$1e-8$	1.00	79.48	55	109.88	164.39	243.87
1	$1e-3$	$1e-8$	0.98	77.15	93	184.34	327.37	404.52
2	$1e-3$	$1e-8$	3.07	1361.43	17	69.26	80.70	1442.13
2	$1e-2$	$1e-8$	2.11	619.85	18	55.97	67.67	687.52
2	$5e-2$	$1e-8$	0.98	162.29	17	33.69	45.13	207.42

Следует заметить, что, как указано в работе [12], критерий остановки (6.9) является слишком грубым для тестовой линейной системы с матрицей RAEFSKY4, а именно, в этом случае точность получаемого решения $\tilde{x}$ является неприемлемой (во всех вариантах, приведенных в таблице 2, мы получили

$$\frac{||\tilde{x} - x^*||}{||x^*||} \simeq 10^2,$$

где x^* – точное решение рассматриваемой системы). В силу этого, следуя работе [12], мы приводим в таблице 7 результаты численных экспериментов, полученных для более жесткого критерия остановки:

$$||r_k||/||r_0|| < 10^{-13}, \tag{6.10}$$

гарантирующего более приемлемую точность решения

$$\frac{||\tilde{x} - x^*||}{||x^*||} \simeq 10^{-2}.$$

Таблица 3. Результаты численных экспериментов для матрицы UTM5940.

p	δ	ε	*Dens*	*Constr. Costs*	*Iters*	*MVM Costs*	*Iter. Costs*	*Total Costs*
1	–	–	1.79	137.55	769	2146.66	37224.59	37362.14
2	–	–	7.93	1122.60	433	3865.26	15056.48	16179.08
3	–	–	17.64	4385.64	287	5350.20	10315.05	14700.69
4	–	–	30.72	12126.27	211	6692.77	9418.54	21544.81
3	–	$1e{-}3$	6.43	4385.64	289	2148.45	7169.59	11555.23
3	–	$5e{-}3$	5.31	4385.64	300	1892.25	7297.22	11682.86
4	–	$1e{-}3$	9.27	12126.27	213	2187.17	4941.04	17067.31
4	–	$5e{-}3$	7.58	12126.27	220	1887.69	4820.51	16946.78
3	$1e{-}4$	$1e{-}3$	6.14	3670.54	293	2092.93	7252.12	10922.66
3	$1e{-}3$	$1e{-}3$	5.57	2988.89	305	2002.48	7587.41	10576.30
3	$5e{-}3$	$1e{-}3$	4.63	2257.93	338	1902.67	8746.46	11004.39
4	$1e{-}3$	$1e{-}3$	7.30	7217.10	242	2007.92	5545.75	12762.85
4	$5e{-}3$	$1e{-}3$	5.88	5402.16	287	1973.93	6926.09	12328.25
4	$1e{-}2$	$1e{-}3$	5.14	4510.81	318	1951.38	8016.84	12527.65

Замечание. В таблицах 2–7 все затраты на построение агрегированной матрицы A_c и факторов FSAI-переобуславливателя для матрицы A_c включены в **Constr. Costs**.

Анализируя данные численных экспериментов, представленные в таблицах 2–7, можно сделать следующие выводы.

1. Как и в случае линейных систем с симметричной положительно определенной матрицей коэффициентов, в несимметричном случае, начиная с некоторой степени p, рост арифметических

Таблица 4. Результаты численных экспериментов для матрицы Tern.

p	δ	ε	*Dens*	*Constr. Costs*	*Iters*	*MVM Costs*	*Iter. Costs*	*Total Costs*
1	–	–	1.30	36.77	321	737.72	8548.20	8584.97
2	–	–	5.02	148.25	188	1131.14	3850.05	3998.30
3	–	–	11.77	776.88	98	1251.13	2024.69	2801.57
4	–	–	21.94	2958.54	65	1490.78	1861.31	4819.85
3	–	$1e-5$	5.69	776.88	102	682.42	1508.61	2285.49
3	–	$1e-4$	4.78	776.88	102	589.79	1415.98	2192.86
4	–	$1e-3$	5.33	2958.54	67	423.76	797.32	3755.86
4	–	$1e-2$	2.66	2958.54	76	278.04	748.21	3706.75
3	$1e-5$	$1e-4$	4.07	470.28	104	527.02	1384.71	1854.99
3	$1e-4$	$1e-4$	3.08	258.79	112	456.90	1445.39	1704.18
3	$1e-3$	$1e-4$	1.92	118.55	140	408.84	1932.52	2051.07
4	$1e-5$	$1e-2$	2.13	785.75	89	278.89	913.27	1699.02
4	$1e-4$	$1e-2$	1.68	376.86	102	273.43	1099.62	1476.48
4	$1e-3$	$1e-2$	1.20	186.50	135	296.39	1714.97	1901.47

Таблица 5. Результаты численных экспериментов для матрицы Osprey.

p	δ	ε	*Dens*	*Constr. Costs*	*Iters*	*MVM Costs*	*Iter. Costs*	*Total Costs*
1	–	–	1.29	36.76	> 800			> 47000
2	–	–	5.22	152.10	> 800			> 51000
3	–	–	12.50	700.92	175	2361.82	4633.28	5334.20
4	–	–	23.48	2800.95	118	2888.12	3957.55	6758.50
3	–	$5e-4$	2.49	700.92	178	620.74	2958.13	3659.05
3	–	$1e-3$	2.24	700.92	179	580.75	2943.72	3644.64
4	–	$5e-4$	3.68	2800.95	121	566.43	1667.56	4468.51
4	–	$1e-3$	3.29	2800.95	124	532.31	1686.60	4487.55
3	$1e-7$	$1e-3$	2.22	599.19	179	576.09	2939.06	3538.25
3	$1e-6$	$1e-3$	2.18	526.37	185	587.41	3108.42	3634.79
3	$1e-5$	$1e-3$	2.05	392.55	231	704.57	4608.47	5001.02
4	$1e-7$	$5e-4$	3.48	1896.24	125	559.83	1732.46	3628.70
4	$1e-6$	$5e-4$	3.24	1362.97	129	547.00	1793.37	3156.34
4	$1e-5$	$5e-4$	2.81	801.86	181	690.12	3105.80	3907.66

затрат на построение FSAI-переобуславливателя, структура разреженности которого совпадает со структурой разрежен-

Таблица 6. Результаты численных экспериментов для матрицы PernisWest.

p	δ	ε	$Dens$	$Constr.$ $Costs$	$Iters$	MVM $Costs$	$Iter.$ $Costs$	$Total$ $Costs$
1	–	–	1.26	36.48	179	405.06	2581.07	2617.55
2	–	–	5.63	172.59	123	815.15	1865.23	2037.82
3	–	–	14.48	925.29	89	1377.51	1951.35	2876.64
2	–	$5e-4$	2.03	172.59	125	378.20	1457.01	1629.60
2	–	$1e-3$	1.86	172.59	125	357.25	1436.06	1608.65
3	–	$5e-4$	3.72	925.29	97	458.11	1120.87	2046.16
3	–	$1e-3$	3.30	925.29	99	425.93	1114.70	2039.99
2	$1e-4$	$1e-3$	1.79	147.59	125	349.06	1427.87	1575.46
2	$5e-4$	$1e-3$	1.66	125.28	130	346.16	1510.22	1635.50
2	$1e-3$	$1e-3$	1.56	110.48	132	338.16	1537.24	1647.72
3	$1e-4$	$1e-3$	2.66	445.37	105	383.84	1154.41	1599.78
3	$5e-4$	$1e-3$	2.11	255.40	117	364.03	1313.16	1568.56
3	$1e-3$	$1e-3$	1.85	189.52	125	356.24	1435.05	1624.57

Таблица 7. Результаты численных экспериментов для матрицы RAEFSKY4 с критерием остановки (6.10).

p	δ	ε	$Dens$	$Constr.$ $Costs$	$Iters$	MVM $Costs$	$Iter.$ $Costs$	$Total$ $Costs$
1	–	–	1.05	79.28	> 800			> 10000
2	–	–	4.30	1650.73	229	1213.88	2036.98	3687.71
2	–	$1e-8$	3.40	1650.73	229	1008.29	1831.39	3482.12
2	–	$1e-7$	3.36	1650.73	227	989.92	1798.44	3449.17
2	–	$1e-6$	3.32	1650.73	249	1075.38	2045.01	3695.74
2	$1e-3$	$1e-8$	3.07	1361.43	232	945.17	1788.75	3150.18
2	$1e-3$	$1e-7$	3.07	1361.43	233	947.35	1798.07	3159.50
2	$1e-2$	$1e-8$	2.11	619.85	260	808.50	1863.01	2482.86
2	$1e-2$	$1e-7$	2.11	619.85	260	807.33	1861.84	2481.69

ности нижнего треугольника матрицы $(A + A^T)^p$, не компенсируется уменьшением затрат на итерационную фазу решения.

2. Постфильтрация FSAI-переобуславливателя, как и в симметричном случае, является эффективным способом снижения арифметических затрат на итерационную фазу решения, поскольку при значительном сокращении числа ненулевых элементов переобуславливателя число итераций метода GMRES

растет несущественно, если порог постфильтрации выбран удачно.

3. Величина оптимального порога постфильтрации в несимметричном случае на один-два порядка ниже, чем в случае симметричной положительно определенной матрицы коэффициентов. Данный факт можно объяснить тем, что затраты на собственно итерационную схему метода GMRES растут квадратично с ростом числа итерации, а не линейно, как в методе сопряженных градиентов. Поэтому, чтобы не допускать слишком заметного увеличения числа итераций (приемлемого в симметричном случае), приходится применять более низкие значения порога постфильтрации.
4. Как и в симметричном случае, при использовании FSAI-переобуславливателей для решения линейных систем с несимметричными матрицами предфильтрация блочной структуры ненулевых элементов с помощью векторной агрегации является универсальным инструментом, позволяющим минимизировать суммарные арифметические затраты.

Таблица 8. Сравнительные результаты численных экспериментов для метода IBBLU-GMRES.

Матрица	Параметры метода IBBLU	Dens	Constr. Costs	Iters	Iter. Costs	Total Costs
$UTM5940$	$(210, 3, 2e-3, 1e-4)$	4.70	295.7	43	399.6	695.3
$RAEFSKY4$	$(560, 2, 6e-1, 0.0)$	1.43	80.5	17	51.4	131.9
$Tern$	$(815, 3, 1e-2, 2e-4)$	1.92	74.4	64	439.2	513.6
$Osprey$	$(874, 3, 1e-2, 7e-5)$	3.02	311.4	44	297.5	608.9
$PernisWest$	$(871, 1, 1e-2, 1e-3)$	1.54	120.5	74	480.1	600.6

В заключение в таблице 8 мы приводим результаты численных экспериментов для тех же тестовых линейных систем при их переобуславливании с помощью метода неполного блочного треугольного разложения IBBLU, описанного в работе [12]. Результаты по первым двум модельным задачам UTM5940 и RAEFSKY4 взяты из работы [12], для остальных трех задач результаты были получены авторами. При выборе конкретного переобуславливателя мы использовали параметризацию, принятую в этой же работе.

Заметим, что, несмотря на то, что использование пост и предфильтрации приводит к значительному повышению эффективности FSAI-переобуславливания, как видно из таблицы 8, арифметические затраты при этом все же существенно превышают затраты на решение этих же линейных систем при использовании в качестве переобуславливателей неполных треугольных разложений. Поэтому FSAI-переобуславливания следует рассматривать в первую очередь в контексте параллельных вычислений. С другой стороны, поскольку целью данной работы было исследование методов повышения эффективности FSAI-переобуславливания, получаемых формальным обобщением методов, используемых для симметричных положительно определенных матриц, мы сознательно не касались здесь проблемы выбора дополнительных параметров, специфичных лишь для несимметричного случая (например, выбора существенно различных структур разреженности для нижнего и верхнего факторов FSAI-переобуславливателя). Изучение влияния таких специфических параметров на качество FSAI-переобуславливания в случае несимметричной незнакоопределенной матрицы коэффициентов, а также возможность дальнейшей минимизации последовательных вычислительных затрат представляется нам наиболее интересным направлением последующих исследований.

Литература

1. M. Benzi, C.D.Meyer, and M. Tůma, *A sparse approximate inverse preconditioner for the conjugate gradient method.* SIAM J. Sci. Comput., **17** (1996), 1135–1149.
2. M. Benzi and M. Tůma, *A sparse approximate inverse preconditioner for nonsymmetric linear systems.* SIAM J. Sci. Comput., **19** (1998), 968–994.
3. M. Benzi and M. Tůma, *A comparative study of sparse approximate inverse preconditioners.* Appl. Numer. Math., **30** (1999), 305–340.
4. E. Chow, *A priori sparsity patterns for parallel sparse approximate inverse preconditioners.* SIAM J. Sci. Comput., **21** (2000), 1804-1822.
5. E. Chow, *Parallel implementation and practical use of sparse approximate inverses with a priori sparsity patterns.* Int. J. High Perf. Comput. Appl., **15** (2001), 56-74.
6. T. Davis, *Sparse matrix collection.* http://www.cise.ufl.edu/~davis/sparse, University of Florida, June 1997.
7. M. R. Field, *An efficient parallel preconditioner for the conjugate gradient algorithm.* Tech. Rep. HDL-TR-97-175, O'Reilly Institute, Hitachi Dublin Laboratory (1997).
8. Alan George and Joseph W. Liu., *Computer Solution of Large Sparse Positive Definite Systems.* Prentice-Hall, Englewood Cliffs, NewJersey (1981).
9. L. Yu. Kolotilina, *Eigenvalue bounds and inequalities using vector aggregation of matrices.* Linear Algebra Appl., **271** (1998), 139–167.

10. L. Yu. Kolotilina and A. Yu. Yeremin, *Factorized sparse approximate inverse preconditionings. I: Theory.* SIAM J. Matrix Anal. Appl., **14** (1993), 45–58.
11. L. Yu. Kolotilina and A. Yu. Yeremin, *Factorized sparse approximate inverse preconditionings. II: Solution of 3D FE systems on massively parallel computers.* Int. J. High Speed Computing, **7** (1995), 191-215.
12. L. Yu Kolotilina, A. A. Nikishin and A. Yu. Yeremin, *An incomplete LU-factorization algorithm based on block bordering.* Numer. Linear Algebra Appl., **7** (2000), 543–567.
13. *Matrix Market sparse matrix collection,.* National Institute of Standards and Technology, http://www.nist.gov/MatrixMarket, 1999.
14. A. A. Nikishin and A. Yu. Yeremin, *Prefiltration technique via aggregation for constructing low density high quality factorized sparse approximate inverse preconditionings.* Submitted to Numer. Linear Algebra Appl. (2001).
15. Y. Saad, *Iterative Methods for Sparse Linear Systems.* PWS Publ. Co., Boston (1996).
16. Y. Saad, *Finding exact and approximate block structures for ILU preconditioning.* Report UMSI-2001-93, Minnesota Supercomputer Institute, University of Minnesota, Minneapolis, MN (2001).
17. Y. Saad and Martin H. Schultz, *GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems.* SIAM J. Sci. and Stat. Comput., **7** (1986), 856–869.
18. W.-P. Tang, *Toward an effective sparse approximate inverse preconditioner.* SIAM J. Matrix Anal. Appl., **20** (1998), 970–986.
19. A. Yu. Yeremin, L. Yu. Kolotilina, and A. A. Nikishin, *Factorized sparse approximate inverse preconditionings. III: Iterative construction of preconditioners.* Zap. Nauchn. Semin. POMI, **248** (1998), 17–48.
20. A. Yu. Yeremin, L. Yu. Kolotilina, and A. A. Nikishin, *Factorized sparse approximate inverse preconditionings. IV: Simple approaches to rising efficiency.* Numer. Linear Algebra Appl., **6** (1999), 515–531.

Yeremin A. Yu., Nikishin A. A. Factorized-sparse-approximate-inverse preconditionings of linear systems with unsymmetric matrices.

The paper considers the application of factorized-sparse-approximate-inverse (FSAI) preconditionings to linear algebraic systems with nonsingular unsymmetric coefficient matrices. Special attention is paid to methods for optimizing the sparsity pattern of such preconditioners. It is numerically demonstrated that the efficiency of FSAI preconditionings can substantially be improved by applying the so-called post-filtering and prefiltering in order to sparsify preconditioning matrices.

Научно-исследовательский
Вычислительный Центр МГУ,
E-mail:badger@vvv.srcc.msu.su
Вычислительный Центр РАН
им. А. А. Дородницына,
E-mail:nikishin@ccas.ru

Поступило 2 марта 2002 г.