Записки научных
семинаров ПОМИ
Том 248, 1998 г.

А. Ю. Еремин, Л. Ю. Колотилина, А. А. Никишин

ПЕРЕОБУСЛАВЛИВАНИЕ С ПОМОЩЬЮ ФАКТОРИЗОВАННЫХ РАЗРЕЖЕННЫХ ПРИБЛИЖЕННЫХ ОБРАТНЫХ III. ИТЕРАЦИОННОЕ ПОСТРОЕНИЕ ПЕРЕОБУСЛАВЛИВАТЕЛЕЙ

1. Введение

В работе рассматриваются переобуславливания с помощью факторизованных разреженных обратных (так называемые FSAI-переобуславливания) для систем линейных алгебраических уравнений с симметричными, положительно определенными (SPD) матрицами коэффициентов. FSAI-переобуславливания были предложены и теоретически исследованы в [18]. Применение FSAI-переобуславливаний к решению трехмерных конечно-элементных задач на массивно-параллельных компьютерах рассматривалось в [19]. Альтернативный подход к построению разреженных приближенных обратных в факторизованном виде, основанный на алгоритме построения двух наборов A-бисопряженных векторов, был предложен в [3, 4]. По сравнению с переобуславливаниями с помощью нефакторизованных разреженных приближенных обратных (т.н. SAI-переобуславливаниями), которые были введены гораздо раньше [2] и достаточно интенсивно изучались в последнее десятилетие (см., например, [17, 10, 13, 5, 6, 9]), FSAI-переобуславливания обладают тем преимуществом, что они сохраняют симметрию и положительную определенность исходной матрицы.

В этой работе предполагается, что нижний треугольный шаблон разреженности S для FSAI-переобуславливателя SPD-матрицы A фиксирован заранее, и проблема оптимального выбора S не рассматривается. Этим наши рассмотрения отлича-

Работа выполнена при частичной финансовой поддержке Организации научных исследований Нидерландов (the Netherlands Organisation for Scientific Research (NWO)), грант 047 003 017; ИНТАСа (INTAS), грант INTAS-93-377 EXT; ИНТАСа и РФФИ, грант INTAS-RFBR 95-0098.

ются от подходов, использованных в работах [3, 4] и [13, 5, 6, 9], где шаблон разреженности выбирается в процессе вычисления переобуславливателя. В данной работе главным образом рассматриваются два вопроса. Во-первых, в §2 мы приводим дополнительные свидетельства необходимости использования апостериорного масштабирования по Якоби при построении FSAI-переобуславливателей. С этой целью мы проводим теоретическое сравнение двух типов FSAI-переобуславливателей (различающихся диагональной масштабирующей матрицей), первый из которых соответствует безусловной минимизации фробениусовой нормы соответствующей матричной невязки по всем нижним треугольным матрицам-переобуславливателям, имеющим фиксированный шаблон разреженности S, тогда как второй тип соответствует минимизации того же самого функционала, но при дополнительном ограничении, состоящим в том, что все диагональные элементы переобусловленной матрицы должны равняться единице.

Результаты этого сравнения показывают, что безусловная минимизация фробениусовой нормы приводит к переобусловленным матрицам с меньшими собственными значениями, в результате чего спектральное число обусловленности может быть значительно больше, чем при использовании апостериорного масштабирования по Якоби. Это выявляет потенциальную опасность построения FSAI-переобуславливателей на основе безусловной минимизации фробениусовой нормы.

Второй вопрос, рассматриваемый в данной работе, – это возможность итерационного построения FSAI-переобуславливателей. Необходимость рассмотрения их итерационного построения вызвана двумя причинами. Во-первых, при таком подходе появляется возможность снизить затраты на построение переобуславливателя, которые могут быть весьма велики, особенно, при использовании последовательных компьютеров. Во-вторых, в некоторых случаях (например, когда факторизованные разреженные приближенные обратные используются для аппроксимации обратных к дополнениям по Шуру при построении некоторых блочных факторизаций) нецелесообразно в явном виде формировать матрицу, приближенная обратная к которой нам нужна, например, в силу соображений экономии памяти. В подобных случаях прямые методы решения оказываются попросту неприменимыми.

Наш главный вывод относительно возможности итерационного построения FSAI-переобуславливателей, которая исследуется теоретически в §3 и численно в §4, состоит в том, что при построении FSAI-переобуславливателей высокая точность не требуется, и можно использовать простой критерий остановки, основанный на относительном изменении диагональных элементов строящегося переобуславливателя.

§2. Оптимальные и квазиоптимальные FSAI-переобуславливания

В этом параграфе мы напоминаем (см. [18, 19]) два типа FSAI-переобуславливаний для симметричной, положительно определенной матрицы A и проводим их сравнение. Переобуславливания первого типа вводятся по чисто теоретическим соображениям. Мы называем их оптимальными, поскольку они, по определению, минимизируют фробениусову норму соответствующей матричной невязки по множеству всех нижних треугольных матриц с фиксированным шаблоном разреженности. Мы также рассматриваем другой тип FSAI-переобуславливаний, которые в действительности обладают тем же свойством минимизации, но при дополнительном ограничении, состоящим в том, что диагональные элементы результирующей переобусловленной матрицы все должны быть равны единице. Таким образом переобуславливания второго типа являются квазиоптимальными с точки зрения минимизации соответствующей фробениусовой нормы. С другой стороны, оказывается, что с точки зрения оптимизации распределения спектра переобусловленных матриц квазиоптимальные переобуславливания предпочтительнее, чем оптимальные.

Пусть A – $n \times n$ SPD-матрица, и пусть S, где

$$\{(i,j):\ i<j\} \subseteq S \subseteq \{(i,j):\ i \neq j\}\,, \tag{2.1}$$

– некоторый фиксированный нижний треугольный шаблон разреженности. Пусть, далее, $A = L_A\, L_A^T$ есть разложение Холецкого матрицы A. Определим нижнюю треугольную матрицу $G^{(0)} = (g_{ij}^{(0)})_{i,j=1}^n$ как минимизатор значения функционала $\|I - HL_A\|_F^2 = \mathrm{tr}\,[(I-HL_A)(I-HL_A)^T]$ по всем матрицам $H = (h_{ij})_{i,j=1}^n$ таким, что $h_{ij} = 0$ при $(i,j) \in S$, т.е. по всем матрицам H с шаблоном разреженности S. Таким образом, матрица $G^{(0)}$ может рассматриваться как оптимальная разреженная приближенная обратная с шаблоном разреженности S к сомножителю Холецкого L_A, и, следовательно, естественно использовать $G^{(0)}$ в качестве переобуславливателя для A, когда переобусловленная матрица имеет вид $G^{(0)}AG^{(0)^T}$.

Как нетрудно убедиться [18], матрица $G^{(0)}$ определяется уравнениями

$$\begin{aligned} &g_{ij}^{(0)} = 0, && (i,j) \in S; \\ &(G^{(0)}A)_{ij} = 0, && i \neq j \,\&\, (i,j) \notin S; \\ &(G^{(0)}A)_{ii} = \ell_{ii}, && i = 1, \dots, n, \end{aligned} \tag{2.2}$$

где $L_A = (\ell_{ij})_{i,j=1}^n$. Поскольку все главные подматрицы SPD-матрицы A невырождены, система (2.2) однозначно разрешима, а значит $G^{(0)}$ определена единственным образом.

Мы также рассмотрим другую нижнюю треугольную матрицу-переобуславливатель $G = (g_{ij})_{i,j=1}^n$ с шаблоном разреженности S, которая определяется следующим образом. Сначала строится вспомогательная матрица $\widetilde{G} = (\widetilde{g}_{ij})_{i,j=1}^n$, определенная соотношениями

$$\begin{aligned} &\widetilde{g}_{ij} = 1, && i = j; \\ &\widetilde{g}_{ij} = 0, && (i,j) \in S; \\ &(\widetilde{G}A)_{ij} = 0, && (i,j) \notin S \,\&\, i \neq j. \end{aligned} \tag{2.3}$$

Затем вычисляется диагональная матрица $D = \mathrm{diag}\,(d_1, \dots, d_n)$, заданная равенством

$$D^2 = \mathrm{diag}\left(\widetilde{G}\,A\,\widetilde{G}^T\right). \tag{2.4}$$

Наконец, матрица G определяется по формуле

$$G = D^{-1}\,\widetilde{G}. \tag{2.5}$$

Как легко видеть, диагональные элементы соответствующей переобусловленной матрицы GAG^T все равны единице:

$$(GAG^T)_{ii} = 1, \quad i = 1, \dots, n, \tag{2.6}$$

т.е. матрица GAG^T является отмасштабированной по Якоби.

Насколько нам известно, матрицы определяемые соотношениями (2.3)–(2.5) впервые появились в работе [8], где они были введены другим (рекуррентным) способом в контексте аппроксимации ленточных SPD-матриц. В [1] (см. также [14–16]) эти матрицы (описанные несколько по-другому) возникают как минимизаторы значения функционала Капорина

$$\beta(\,HAH^T\,) = \frac{\operatorname{tr} HAH^T}{n} \left(\det HAH^T\,\right)^{-1/n} \tag{2.7}$$

по множеству всех нижних треугольных матриц с фиксированным шаблоном разреженности. Наконец, в [18] те же самые матрицы были предложены в качестве вычислимых на практике заместителей оптимальных разреженных приближенных обратных, определяемых соотношениями (2.2).

Наш первый результат показывает, что переобуславливатель G, определенный соотношениями (2.3)–(2.5), являющийся оптимальным относительно минимизации функционала β, в действительности обладает более общим свойством оптимальности, которое может быть сформулировано в терминах диагональных элементов матрицы G.

Лемма 2.1. *Пусть A есть SPD-матрица, и пусть S – фиксированный нижний треугольный шаблон разреженности. Если матрица G определена соотношениями (2.3)–(2.5), а матрица $H = (h_{ij})_{i,j=1}^n$ удовлетворяет условиям*

$$\forall\,(i,j) \in S \quad h_{ij} = 0 \tag{2.8}$$

и

$$\left(\,HAH^T\,\right)_{ii} \leqslant 1, \quad i = 1, \dots, n, \tag{2.9}$$

то

$$g_{ii} \geqslant h_{ii}, \quad i = 1, \dots, n. \tag{2.10}$$

Доказательство. Пусть индекс i, $1 \leqslant i \leqslant n$, фиксирован. В силу (2.6) и (2.9) мы имеем

$$0 \leqslant \left[(G-H)A(G-H)^T\,\right]_{ii} \leqslant 2\left[1-(GAH^T)_{ii}\,\right],$$

откуда следует, что

$$\left(GAH^T\right)_{ii} \leqslant 1. \tag{2.11}$$

Принимая во внимание, что по (2.3)–(2.5) и (2.8)

$$\left(GAH^T\right)_{ii} = \sum_{j=1}^{n} (GA)_{ij}\, h_{ij} = (GA)_{ii}\, h_{ii},$$

и, аналогично,

$$1 = \left(GAG^T\right)_{ii} = (GA)_{ii}\, g_{ii}, \tag{2.12}$$

из (2.11) мы получаем, что

$$(GA)_{ii}\, h_{ii} \leqslant (GA)_{ii}\, g_{ii},$$

откуда и следует (2.10). $\square$

Основываясь на этом простом результате, легко вывести некоторые важные следствия и, в частности, квазиоптимальность матриц, определяемых соотношениями (2.3)–(2.5) и рассматриваемых в качестве разреженных приближенных обратных к сомножителю Холецкого L_A.

Теорема 2.1. *Пусть для данной $n{\times}n$ SPD-матрицы A и некоторого фиксированного нижнего треугольного шаблона разреженности S матрица $G = (g_{ij})_{i,j=1}^n$ определяется соотношениями (2.3)–(2.5), и пусть матрица $H = (h_{ij})_{i,j=1}^n$ с шаблоном разреженности S удовлетворяет условиям*

$$\left(HAH^T \right)_{ii} = 1, \quad i = 1, \ldots, n. \tag{2.13}$$

Тогда

$$\|I - GL_A\|_F \leqslant \|I - HL_A\|_F.$$

Доказательство. Ясно, что нам достаточно убедиться в справедливости неравенств

$$\begin{aligned} &\left[(I - GL_A)(I - GL_A)^T\right]_{ii} \\ &\leqslant \left[(I - HL_A)(I - HL_A)^T\right]_{ii}, \quad i = 1, \ldots, n. \end{aligned} \tag{2.14}$$

Поскольку ввиду (2.6) и (2.13) мы имеем

$$\left[(I - GL_A)(I - GL_A)^T\right]_{ii} = 2\,(1 - g_{ii}\,\ell_{ii})\,, \quad i = 1, \ldots, n, \tag{2.15}$$

$$\left[(I - HL_A)(I - HL_A)^T\right]_{ii} = 2\,(1 - h_{ii}\,\ell_{ii})\,, \quad i = 1, \ldots, n, \tag{2.16}$$

то соотношения (2.14) тривиально следуют из неравенств (2.10), установленных в лемме 2.1. $\square$

Теорема 2.1 утверждает, что матрица G, определенная с помощью (2.3)–(2.5), минимизирует значение функционала $\|I-HL_A\|_F$ по всем матрицам H с тем же шаблоном разреженности S, удовлетворяющим дополнительному условию $\operatorname{diag}(HAH^T) = I_n$. Таким образом G оказывается квазиоптимальной разреженной приближенной обратной к L_A.

Основываясь на лемме 2.1, мы можем также очень легко передоказать приведенное выше свойство оптимальности матрицы G, определенной с помощью (2.3)–(2.5), относительно функционала (2.7).

Теорема 2.2. *Пусть $n \times n$ SPD-матрица A и нижний треугольный шаблон разреженности S фиксированы. Тогда матрица G, удовлетворяющая соотношениям (2.3)–(2.5), минимизирует значение функционала (2.7) по всем невырожденным матрицам H с тем же шаблоном разреженности S.*

Доказательство. Пусть

$$\Gamma = \operatorname{diag}(\gamma_1, \dots, \gamma_n), \quad \text{где} \quad \gamma_i = \left[(HAH^T)_{ii}\right]^{1/2}, \quad i = 1, \dots, n.$$

Тогда

$$\beta(HAH^T) = \frac{\Sigma\gamma_i^2}{n}\left(\det HAH^T\right)^{-1/n}$$

$$\geqslant (\Pi\gamma_i^2)^{1/n}\left(\det HAH^T\right)^{-1/n} = \det\left(\overline{H}A\overline{H}^{\,T}\right)^{-1/n},$$

где мы полагаем $\overline{H} = \Gamma^{-1} H$. Поскольку $(\overline{H}A\overline{H})_{ii} = 1$, $i = 1, \dots, n$, а $\overline{H} = (\overline{h}_{ij})_{i,j=1}^n$ – матрица с тем же шаблоном разреженности S, то по лемме 2.1 мы имеем

$$\overline{h}_{ii} = h_{ii}/\gamma_i \leqslant g_{ii}, \quad i = 1, \dots, n,$$

откуда следует, что

$$\det\left(\overline{H}A\overline{H}^{\,T}\right)^{-1/n} \geqslant \left(\det GAG^T\right)^{-1/n} = \beta\left(GAG^T\right).$$

$\square$.

Теперь мы установим явное соотношение между G и $G^{(0)}$.

Лемма 2.2. *Пусть $A = L_A\, L_A^T$, где $L_A = (\ell_{ij})_{i,j=1}^n$ – сомножитель Холецкого, есть некоторая $n \times n$ SPD-матрица, S – нижний треугольный шаблон разреженности, а матрицы $G^{(0)}$ и G определены соотношениями* (2.2) *и* (2.3)–(2.5), *соответственно. Тогда*

$$G^{(0)} = \Delta G, \tag{2.17}$$

где

$$\Delta = \mathrm{diag}\,(\delta_1, \dots, \delta_n), \quad \delta_i = \ell_{ii}\, g_{ii} \leqslant 1, \quad i = 1, \dots, n. \tag{2.18}$$

Доказательство. Заметим сперва, что в силу (2.3) мы имеем

$$0 < \left(\widetilde{G} A \widetilde{G}^T\right)_{ii} = \sum_{j=1}^n (\widetilde{G} A)_{ij}\, \widetilde{g}_{ij} = (\widetilde{G} A)_{ii}, \quad i = 1, \dots, n, \tag{2.19}$$

и значит, ввиду (2.3)–(2.5), матрица G удовлетворяет следующим соотношениям:

$$\begin{aligned}
&(GA)_{ii} = d_i^{-1}\ (\widetilde{G} A)_{ii} > 0, && i = 1, \dots, n;\\
&(GA)_{ij} = d_i^{-1}\ (\widetilde{G} A)_{ij} = 0, && i \neq j \,\&\, (i, j) \notin S;\\
&g_{ij} = d_i^{-1}\, \widetilde{g}_{ij} = 0, && (i, j) \in S.
\end{aligned}$$

Сравнивая эти соотношения с соотношениями (2.2) и используя однозначную разрешимость (2.2), мы приходим к заключению, что существует диагональная матрица $\Delta = \mathrm{diag}\,(\delta_1, \dots, \delta_n)$ с положительными диагональными элементами такая, что выполнено (2.17). Очевидно, что $\delta_i = g_{ii}^{(0)}/g_{ii}$, $i = 1, \dots, n$, а тогда по (2.17) и (2.6) мы имеем

$$\left(G^{(0)} A G^{(0)T}\right)_{ii} = \left(g_{ii}^{(0)}/g_{ii}\right)^2 (GAG^T)_{ii} = \left(g_{ii}^{(0)}/g_{ii}\right)^2, \quad i = 1, \dots, n. \tag{2.20}$$

С другой стороны, в силу (2.2)

$$\left(G^{(0)} A G^{(0)T}\right)_{ii} = \sum_{j=1}^n (G^{(0)} A)_{ij}\, g_{ij}^{(0)} = \ell_{ii}\, g_{ii}^{(0)}, \quad i = 1, \dots, n. \tag{2.21}$$

Сравнивая (2.20) с (2.21), мы приходим к равенствам

$$g_{ii}^{(0)} = g_{ii}^2\, \ell_{ii}, \quad i = 1, \dots, n, \tag{2.22}$$

из которых следует, что

$$\delta_i = g_{ii}^{(0)} / g_{ii} = g_{ii}\,\ell_{ii}, \quad i = 1, \ldots, n.$$

Таким образом для завершения доказательства леммы 2.2 нам остается лишь убедиться, что $g_{ii}\,\ell_{ii} \leqslant 1$, $i = 1, \ldots, n$. Действительно, для любого i, $1 \leqslant i \leqslant n$, мы имеем

$$1 = (GAG^T)_{ii} = \left[GL_A(GL_A)^T\right]_{ii} = (GL_A)_{ii}^2 + \sum_{j \neq i}(GL_A)_{ij}^2 \geqslant (g_{ii}\,\ell_{ii})^2.$$

□

Основываясь на лемме 2.2, мы сперва покажем, что вне зависимости от A и S

$$1 \leqslant \frac{\|I - GL_A\|_F}{\|I - G^{(0)}\,L_A\|_F} \leqslant \sqrt{2},$$

и, кроме того, $\|I - GL_A\|_F$ приближается к $\|I - G^{(0)}\,L_A\|_F$, когда $G^{(0)}$ приближается к L_A^{-1}. Значит, хотя матрица G, определенная соотношениями (2.3)–(2.5), и является худшей, чем матрица $G^{(0)}$, определенная с помощью 2.2, аппроксимацией обратной к L_A, значения фробениусовых норм $\|I - GL_A\|_F$ и $\|I - G^{(0)}\,L_A\|_F$ не могут различаться на коэффициент больший, чем $\sqrt{2}$, а разность между этими нормами убывает, когда оптимальная приближенная обратная $G^{(0)}$ приближается к L_A^{-1}. Заметим также, что в силу (2.22) $g_{ii}^{(0)}\,\ell_{ii} = 1$ тогда и только тогда, когда $g_{ii}\,\ell_{ii} = 1$.

Теорема 2.3. *Пусть для данной $n{\times}n$ SPD-матриц A и фиксированного нижнего треугольного шаблона разреженности S разреженные приближенные обратные $G^{(0)}$ и G к матрице L_A определяются с помощью* (2.2) *и* (2.3)–(2.5)*, соответственно. Тогда*

$$\|I - G^{(0)}L_A\|_F \leqslant \|I - G\,L_A\|_F$$
$$\leqslant \left[\frac{2}{1 + \min\limits_{1 \leqslant i \leqslant n}\{g_{ii}^{(0)}\,\ell_{ii}\}^{1/2}}\right]^{1/2} \|I - G^{(0)}\,L_A\|_F. \qquad (2.23)$$

Доказательство. Левое неравенство в (2.23) следует из оптимальности матрицы $G^{(0)}$ относительно $\|I - G^{(0)}\,L_A\|_F$. Правое неравенство в (2.23) доказывается следующим образом:

$$\left[(I - G^{(0)}\,L_A)(I - G^{(0)}\,L_A)^T\right]_{ii} = 1 - 2g_{ii}^{(0)}\,\ell_{ii} + \left(G^{(0)}AG^{(0)^T}\right)_{ii}$$

$$
\begin{aligned}
&\overset{(2.21)}{=} 1 - g_{ii}^{(0)}\,\ell_{ii} \overset{(2.22)}{=} 1 - g_{ii}^{2}\,\ell_{ii}^{2} = (1 - g_{ii}\,\ell_{ii})(1 + g_{ii}\,\ell_{ii}) \\
&\geqslant \left(1 + \min_{1\leqslant i\leqslant n}\{g_{ii}\,\ell_{ii}\}\right)(1 - g_{ii}\,\ell_{ii}) \\
&\overset{(2.22),(2.15)}{=} \left(1 + \min_{1\leqslant i\leqslant n}\{g_{ii}^{(0)}\,\ell_{ii}\}^{1/2}\right) \frac{[(I - GL_A)(I - GL_A)^T]_{ii}}{2}, \\
&\qquad i = 1, \ldots, n.
\end{aligned}
$$

□

Результат следующей теоремы, касающейся собственных значений матриц GAG^T и $G^{(0)}AG^{(0)^T}$, также устанавливается с помощью леммы 2.2.

Теорема 2.4. *В предположениях теоремы 2.3 справедливы неравенства*

$$
\begin{aligned}
&\lambda_i\left(G^{(0)}\,A\,G^{(0)^T}\right) \leqslant \lambda_i(GAG^T) \\
&\leqslant \left[\min_{1\leqslant i\leqslant n}\{g_{ii}^{(0)}\,\ell_{ii}\}\right]^{-1} \lambda_i\left(G^{(0)}\,A\,G^{(0)^T}\right), \quad i = 1, \ldots, n. \qquad (2.24)
\end{aligned}
$$

Здесь предполагается, что собственные значения обеих матриц занумерованы в одном и том же монотонном порядке, скажем, по невозрастанию.

Доказательство. Используя (2.17) и (2.18), мы доказываем левое неравенство в (2.24) следующим образом:

$$
\begin{aligned}
\lambda_i(GAG^T) &= \lambda_i\left(\Delta^{-1}\,G^{(0)}\,A\,G^{(0)^T}\,\Delta^{-1}\right) \\
&\geqslant \lambda_i\left(G^{(0)}\,A\,G^{(0)^T}\right)\lambda_{\min}(\Delta^{-2}) \\
&= \lambda_i\left(G^{(0)}\,A\,G^{(0)^T}\right)\min_{1\leqslant i\leqslant n}\{\delta_i^{-2}\} \\
&\geqslant \lambda_i\left(G^{(0)}\,A\,G^{(0)^T}\right), \quad i = 1, \ldots, n.
\end{aligned}
$$

Остальные неравенства устанавливаются аналогично:

$$\begin{aligned}\lambda_i(GAG^T) &\leqslant \lambda_i\left(G^{(0)}\,A\,G^{(0)^T}\right)\lambda_{\max}(\Delta^{-2})\\ &= \lambda_i\left(G^{(0)}\,A\,G^{(0)^T}\right)\max_{1\leqslant i\leqslant n}\{\delta_i^{-2}\}\\ &= \lambda_i\left(G^{(0)}\,A\,G^{(0)^T}\right)\max_{1\leqslant i\leqslant n}\left\{(g_{ii}\,\ell_{ii})^{-2}\right\}\\ &= \lambda_i\left(G^{(0)}\,A\,G^{(0)^T}\right)\max_{1\leqslant i\leqslant n}\left\{(g_{ii}^{(0)}\,\ell_{ii})^{-1}\right\},\quad i=1,\ldots,n.\end{aligned}$$

□

Как показывает теорема 2.4, ни одно из собственных значений матрицы GAG^T не может быть меньше, чем соответствующее собственное значение матрицы $G^{(0)}AG^{(0)^T}$. В частности, мы можем ожидать, что $\lambda_{\min}(GAG^T) > \lambda_{\min}(G^{(0)}AG^{(0)^T})$, что демонстрирует предпочтительность использования G в качестве переобуславливателя для A, если известно, что $\lambda_{\max}(GAG^T)$ ограничено сверху разумной константой. Например, если A является H-матрицей, то $\lambda_{\max}(GAG^T) < 2$ [18, Теорема 4.2]. С другой стороны, из теоремы 2.4 следует, что соответствующие собственные значения матриц GAG^T и $G^{(0)}AG^{(0)^T}$ приближаются друг к другу, когда диагональные элементы $g_{ii}^{(0)}$ (или g_{ii}) приближаются к ℓ_{ii}^{-1}, $i = 1,\ldots,n$. Таким образом, чем ближе к единичной матрице I_n любая из переобусловленных матриц GAG^T или $G^{(0)}AG^{(0)^T}$, тем меньше их разность.

В качестве следствия из теоремы 2.4 можно получить соотношения между минимальными собственными значениями матриц $G^{(0)}AG^{(0)^T}$ и GAG^T.

Следствие 2.1. *В предположениях теоремы* 2.3 *справедливы неравенства*

$$\lambda_{\min}\left(G^{(0)}AG^{(0)^T}\right) \leqslant \lambda_{\min}(GAG^T) \leqslant \lambda_{\min}^{1/2}\left(G^{(0)}AG^{(0)^T}\right). \qquad (2.25)$$

Доказательство. Если принять во внимание левое неравенство в (2.24), то нам требуется установить лишь неравенство

$$\lambda_{\min}\left(G^{(0)}AG^{(0)^T}\right) \geqslant \lambda_{\min}^2(GAG^T),$$

которое следует из правого неравенства в (2.24), равенства (2.22) и соотношения

$$\lambda_{\min}(GAG^T) \leqslant \min_{1\leqslant i\leqslant n} \left\{g_{ii}^2\, \ell_{ii}^2\right\}.$$

Последнее соотношение немедленно вытекает из известного неравенства (см., например, [12, с. 191])

$$\sigma_{\min}(B) \leqslant \min_{1\leqslant i\leqslant n} \mid \lambda_i(B) \mid, \tag{2.26}$$

где $\sigma_{\min}(B)$ есть наименьшее сингулярное значение B, примененного к матрице $B = GL_A$. $\square$

Двусторонние оценки младших собственных значений $G^{(0)}AG^{(0)^T}$ и GAG^T в терминах $g_{ii}^{(0)}\,\ell_{ii} = (G^{(0)}AG^{(0)^T})_{ii}$ приведены в следующей теореме.

Теорема 2.5. *В предыдущих обозначениях справедливы неравенства*

$$\begin{gathered}\min_{1\leqslant i\leqslant n} \left\{g_{ii}^{(0)}\,\ell_{ii}\right\} \left(\frac{n-1}{n}\right)^{n-1} \prod_{i=1}^{n} g_{ii}^{(0)}\,\ell_{ii} \\ \leqslant \lambda_{\min}\left(G^{(0)}AG^{(0)^T}\right) \leqslant \left[\min_{1\leqslant i\leqslant n} \left\{g_{ii}^{(0)}\,\ell_{ii}\right\}\right]^2,\end{gathered} \tag{2.27}$$

$$\left(\frac{n-1}{n}\right)^{n-1} \prod_{i=1}^{n} g_{ii}^{(0)}\,\ell_{ii} \leqslant \lambda_{\min}(GAG^T) \leqslant \min_{1\leqslant i\leqslant n} \left\{g_{ii}^{(0)}\,\ell_{ii}\right\}. \tag{2.28}$$

Доказательство. Верхние оценки в (2.27) и (2.28) следуют из неравенства (2.26) и соотношения $g_{ii}^2\,\ell_{ii} = g_{ii}^{(0)}$ (см. (2.22)). Обе нижние оценки в (2.27) и (2.28) легко выводятся из оценки [11, Теорема 1]

$$\sigma_{\min}(A) \geqslant \left(\frac{n-1}{n}\right)^{\frac{n-1}{2}} \mid \det A \mid \frac{r_{\min}(A)}{\prod\limits_{i=1}^{n} r_i(A)}, \tag{2.29}$$

где $A \in \mathbb{C}^{n\times n}$, $r_i(A) = (\sum\limits_{i=1}^{n} |a_{ij}|^2)^{1/2}$, $r_{\min}(A) = \min\limits_{1\leqslant i\leqslant n} r_i(A)$, если применить ее к матрицам $G^{(0)}L_A$ и GL_A и использовать соотношения $r_i^2(G^{(0)}\,L_A) = \left(G^{(0)}\,A\,G^{(0)^T}\right)_{ii} \overset{(2.21)}{=} g_{ii}^{(0)}\,\ell_{ii}$, $r_i^2(GL_A) = (GAG^T)_{ii} \overset{(2.6)}{=} 1$,

$$\det(G^{(0)}\,L_A) = \prod_{i=1}^{n} g_{ii}^{(0)}\,\ell_{ii}, \;\; \det(GL_A) = \prod_{i=1}^{n} g_{ii}\,\ell_{ii} \overset{(2.22)}{=} \left[\,\prod_{i=1}^{n} g_{ii}^{(0)}\,\ell_{ii}\right]^{1/2}.$$

□

Замечание 2.1 Как было показано в [11], нижняя оценка (2.29) в своем общем виде легко выводится (за счет применения подходящего масштабирования) из оценки

$$\sigma_{\min}(A) \geqslant \left(\frac{n-1}{n}\right)^{\frac{n-1}{2}} |\det A| \quad \left(\geqslant e^{-1/2}\,|\det A|\right), \qquad (2.30)$$

справедливой для произвольной матрицы $A \in \mathbb{C}^{n\times n}$ такой, что $(AA^*)_{ii} = r_i^2(A) = 1$, $i = 1, \ldots, n$. Заметим попутно, что доказательство неравенства (2.30), представленное в [11], является излишне усложненным, а сама оценка почти тривиальна. Действительно, для любой матрицы $A \in \mathbb{C}^{n\times n}$ мы имеем

$$\sigma_{\min}^2(A) = \lambda_n(AA^*) = \frac{\det AA^*}{\prod\limits_{i=1}^{n-1} \lambda_i(AA^*)} \geqslant \frac{|\det A|^2}{\left[\frac{\sum_{i=1}^{n-1} \lambda_i(AA^*)}{n-1}\right]}$$

$$\geqslant |\det A|^2 \left(\frac{n-1}{\operatorname{tr} AA^*}\right)^{n-1},$$

откуда немедленно следует (2.30), если только справедливо неравенство $\operatorname{tr} AA^* \leqslant n$.

Заметим, что интервал (2.27) для $\lambda_{\min}(G^{(0)}\,A\,G^{(0)^T})$ получается из интервала (2.28) для $\lambda_{\min}(GAG^T)$ умножением последнего на число $\min\limits_{1\leqslant i\leqslant n} g_{ii}^{(0)}\,\ell_{ii}$, которое строго меньше единицы за исключением случая $G = G^{(0)} = L_A^{-1}$.

Наше сравнение двух переобуславливателей $G^{(0)}$ и G мы завершим рассмотрением простого примера, показывающего, что соотношение

$$\lambda_{\min}(GAG^T) = O\left(\lambda_{\min}^{1/2}\left(G^{(0)}\,A\,G^{(0)^T}\right)\right)$$

действительно может реализоваться в случае, когда сомножитель Холецкого L_A матрицы A имеет большой внедиагональный элемент. Ввиду (2.25) последнее соотношение отвечает случаю, когда отношение $\lambda_{\min}(GAG^T)/\lambda_{\min}\left(G^{(0)}\,A\,G^{(0)^T}\right)$ имеет наибольший возможный порядок. Заметим, что поскольку G и $G^{(0)}$ различаются лишь на диагональный множитель, нам достаточно рассмотреть случай диагональных переобуславливающих матриц $G = D$ и $G^{(0)} = D^{(0)}$.

Пример 2.1. Рассмотрим 2×2 параметрическую матрицу

$$A = \begin{bmatrix} 1 & \alpha \\ \alpha & 1+\alpha^2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \alpha & 1 \end{bmatrix} \begin{bmatrix} 1 & \alpha \\ 0 & 1 \end{bmatrix}, \quad \alpha > 0.$$

Применяя к A симметричное масштабирование по Якоби, мы приходим к матрице

$$B = \begin{bmatrix} 1 & \frac{\alpha}{\sqrt{1+\alpha^2}} \\ \frac{\alpha}{\sqrt{1+\alpha^2}} & 1 \end{bmatrix},$$

имеющей собственные значения

$$\lambda_1(B) = 1 + \frac{\alpha}{\sqrt{1+\alpha^2}}, \quad \lambda_2(B) = 1 - \frac{\alpha}{\sqrt{1+\alpha^2}}.$$

Поскольку, очевидно,

$$\frac{\alpha}{1+\alpha} \leqslant \frac{\alpha}{\sqrt{1+\alpha^2}} \leqslant 1,$$

мы имеем

$$2 - \frac{1}{1+\alpha} \leqslant \lambda_1(B) \leqslant 2,$$

откуда следует, что

$$\lambda_1(B) = O(1), \quad \alpha \to +\infty.$$

Следовательно,

$$\lambda_2(B) = \frac{\det B}{\lambda_1(B)} = \frac{1}{(1+\alpha^2)\lambda_1(B)} = O(\alpha^{-2}), \quad \alpha \to +\infty.$$

С другой стороны, исходя из минимизации фробениусовой нормы $\|DL_A - I\|_F$, мы получаем матрицу

$$B^{(0)} = D^{(0)}\,A\,D^{(0)} = \begin{bmatrix} 1 & \frac{\alpha}{1+\alpha^2} \\ \frac{\alpha}{1+\alpha^2} & \frac{1}{1+\alpha^2} \end{bmatrix}.$$

Совершенно ясно, что

$$1 \leqslant \lambda_1(B^{(0)}) \leqslant \operatorname{tr} B^{(0)} = 1 + \frac{1}{1+\alpha^2} = O(1), \quad \alpha \to +\infty.$$

Но тогда

$$\lambda_2(B^{(0)}) = \frac{\det B^{(0)}}{\lambda_1(B^{(0)})} = \frac{1}{(1+\alpha^2)^2 \lambda_1(B^{(0)})} = O(\alpha^{-4}), \quad \alpha \to +\infty.$$

Таким образом, при $\alpha \to +\infty$ мы имеем

$$\lambda_{\min}^2(B)/\lambda_{\min}(B^{(0)} \to 1,$$

тогда как

$$\lambda_{\max}(B) = O(1), \quad \lambda_{\max}(B^{(0)}) = O(1).$$

Следовательно, для данного примера, в котором наличие в L_A большого внедиагонального элемента α является причиной появления малого диагонального элемента $(B^{(0)})_{22} = \frac{1}{1+\alpha^2} = O(\alpha^{-2})$, с точки зрения минимизации спектрального числа обусловленности масштабирование по Якоби действительно оказывается более предпочтительным, чем минимизация $\|DL_A - I\|_F$, причем

$$(\operatorname{cond} B)^2 / \operatorname{cond} B^{(0)} \to 1, \quad \alpha \to +\infty.$$

§3. Приближенные FSAI-переобуславливатели

Для того, чтобы построить FSAI-переобуславливатель (2.3)–(2.5) для некоторой SPD-матрицы A, необходимо решить линейные уравнения (2.3), которые, как нетрудно понять, распадаются на n независимых линейных систем (называемых локальными), каждая из которых определяет ненулевые элементы соответствующей строки матрицы $\widetilde{G}$. Как было упомянуто в §1, в некоторых случаях использование прямых методов при решении локальных систем может оказаться либо вообще невозможным, либо же слишком дорогостоящим. По этой причине в данном параграфе мы рассматриваем возможность построения приближенных FSAI-переобуславливателей (2.3)–(2.5) с помощью итерационных методов.

Самое общее теоретическое соображение, позволяющее заменить явный SPD-переобуславливатель $K \approx A^{-1}$ для SPD-матрицы

A на SPD-аппроксимацию $\widetilde{K}$ к K, состоит в следующем. Пусть $\widetilde{K} = K + E$, и предположим, что $\|E\|_2$ мала. Тогда

$$\widetilde{K}A = KA + EA,$$

и значит возмущение матрицы KA, индуцируемое возмущением K, есть EA, а для спектральной нормы EA мы имеем

$$\|EA\|_2 \leqslant \lambda_{\max}(A)\,\|E\|_2.$$

Следовательно, если $\lambda_{\max}(A)$ не превосходит существенно 1, то норма матрицы $\widetilde{K}A - KA$ будет почти столь же мала, сколь и норма матрицы $E = \widetilde{K} - K$.

Это простое рассуждение приводит нас к заключению, что чем меньше норма исходной матрицы A, тем более безопасно использовать вместо данного явного переобуславливателя для A некоторую его аппроксимацию.

В случае симметричных переобуславливаний вида $A \to GAG^T$ мы имеем следующий результат.

Теорема 3.1. *Пусть A есть $n \times n$ SPD-матрица, а $n \times n$ матрицы G и H обратимы, причем*

$$\|I - HG^{-1}\|_2 \leqslant \varepsilon < 1. \tag{3.1}$$

Тогда

$$\begin{aligned}(1-\varepsilon)^2\lambda_i(GAG^T) &\leqslant \lambda_i(HAH^T)\\ &\leqslant (1+\varepsilon)^2\lambda_i(GAG^T),\ i = 1,\dots,n,\end{aligned} \tag{3.2}$$

где собственные значения обеих матриц GAG^T и HAH^T упорядочены по невозрастанию.

Доказательство. Обозначим $F = I - HG^{-1}$. Поскольку по предположению $\|F\|_2 \leqslant \varepsilon < 1$, мы имеем

$$\|GH^{-1}\|_2 = \|(I-F)^{-1}\|_2 \leqslant (1 - \|F\|_2)^{-1} \leqslant (1-\varepsilon)^{-1}. \tag{3.3}$$

Отсюда, используя стандартные неравенства для сингулярных значений (см., например, [12, теорема 3.3.16(d)]) и (3.3), мы выводим

$$\begin{aligned}\lambda_i(GAG^T) &= \sigma_i^2(GL_A) = \sigma_i^2(GH^{-1}HL_A)\\ &\leqslant \sigma_1^2(GH^{-1})\sigma_i^2(HL_A) = \left\|GH^{-1}\right\|_2^2\lambda_i(HAH^T)\\ &\leqslant (1-\varepsilon)^{-2}\lambda_i(HAH^T),\quad i = 1,\dots,n,\end{aligned}$$

где $A = L_A\, L_A^T$, как и в §2, а через σ_i мы обозначаем сингулярные значения матриц, упорядоченные по невозрастанию. Таким образом левое неравенство в (3.2) установлено. Аналогично, мы имеем

$$\lambda_i(HAH^T) = \sigma_i^2(HG^{-1}\, GL_A) \leqslant \lambda_i(GAG^T)\big\|HG^{-1}\big\|_2^2, \quad i = 1, \dots, n,$$

и для завершения доказательства нам остается лишь воспользоваться неравенством

$$\|HG^{-1}\|_2 = \|I - F\|_2 \leqslant 1 + \varepsilon,$$

которое следует из (3.1). $\square$

Теперь, после представления общих соображений, мы возвратимся к FSAI-переобуславливателям. Поскольку наш подход к построению FSAI-переобуславливателей базируется на минимизации фробениусовой нормы, представляется интересным изучить поведение целевого функционала при замене точного FSAI-переобуславливателя G на некоторое приближение H к G, шаблон разреженности которого совпадает с шаблоном разреженности G.

Пусть некоторая SPD матрица $A = L_A\, L_A^T$ и некоторый нижний треугольный шаблон разреженности S, удовлетворяющий (2.1), фиксированы, а матрицы $\widetilde{G}$ и G определены в соответствии с (2.3)–(2.5). Рассмотрим произвольную матрицу $\widetilde{H} = (\widetilde{h}_{ij})_{i,j=1}^n$, удовлетворяющую соотношениям

$$\begin{aligned} &\widetilde{h}_{ii} = 1, \qquad && i = 1, \dots, n, \\ &\widetilde{h}_{ij} = 0, \qquad && (i,j) \in S, \end{aligned} \tag{3.4}$$

и определим матрицу H по формуле

$$H = \widetilde{D}^{\,-1}\, \widetilde{H}, \tag{3.5}$$

где $\widetilde{D} = \operatorname{diag}(\, \widetilde{d}_1, \dots, \widetilde{d}_n\,)$ и

$$\tilde{d}_i^{\,2} = (\, \widetilde{H}\, A\, \widetilde{H}^T\,)_{ii}, \quad i = 1, \dots, n. \tag{3.6}$$

Тогда для матрицы $H = (h_{ij})_{i,j=1}^n$, очевидно, выполнены равенства

$$\begin{aligned} &h_{ij} = 0, \qquad && (i,j) \in S, \\ &(HAH^T)_{ii} = 1, \qquad && i = 1, \dots, n. \end{aligned} \tag{3.7}$$

Следовательно, по лемме 2.1 мы имеем

$$h_{ii} \leqslant g_{ii}, \quad i = 1, \dots, n, \tag{3.8}$$

а теорема 2.1 гарантирует, что

$$\|I - HL_A\|_F \geqslant \|I - GL_A\|_F\,. \tag{3.9}$$

В действительности, в рассматриваемом случае для $\|I–HL_A\|_F^2$ и для разности $\|I - HL_A\|_F^2 - \|I - GL_A\|_F^2$ можно вывести явные выражения.

Лемма 3.1. *В предыдущих обозначениях имеют место соотношения*

$$\|I - HL_A\|_F^2 = 2\sum_{i=1}^{n}(1 - h_{ii}\,\ell_{ii}) \tag{3.10}$$

и

$$\|I - HL_A\|_F^2 - \|I - GL_A\|_F^2 = \sum_{i=1}^{n}(EAE^T)_{ii}\, g_{ii}\,\ell_{ii}, \tag{3.11}$$

где мы используем обозначения $L_A = (\ell_{ij})_{i,j=1}^n$ *и* $E = G - H$.

Доказательство. Равенство (3.10) тривиально следует из соотношений

$$\left[(I - HL_A)(I - HL_A)^T\right]_{ii} = 2\,(1 - h_{ii}\,\ell_{ii})\,, \quad i = 1, \dots, n.$$

С другой стороны, мы имеем

$$\begin{aligned} 2\,(1 - h_{ii}\,\ell_{ii}) &= 2\,(1 - g_{ii}\,\ell_{ii}) + 2\,(g_{ii} - h_{ii})\,\ell_{ii} \\ &= 2\,(1 - g_{ii}\,\ell_{ii}) + 2\left(1 - \frac{h_{ii}}{g_{ii}}\right) g_{ii}\,\ell_{ii}, \quad i = 1, \dots, n, \end{aligned}$$

а поскольку

$$\left[(I - GL_A)(I - GL_A)^T\right]_{ii} = 2(1 - g_{ii}\,\ell_{ii}), \quad i = 1, \dots, n, \tag{3.12}$$

то мы приходим к заключению, что

$$\|I - HL_A\|_F^2 = \|I - GL_A\|_F^2 + 2\sum_{i=1}^{n}\left(1 - \frac{h_{ii}}{g_{ii}}\right) g_{ii}\,\ell_{ii}. \tag{3.13}$$

Таким образом для завершения доказательства нам необходимо лишь показать, что

$$(EAE^T)_{ii} = 2\left(1 - \frac{h_{ii}}{g_{ii}}\right), \quad i = 1, \ldots, n. \tag{3.14}$$

Действительно,

$$\begin{aligned}(EAE^T)_{ii} &= \left[(G-H)A(G-H)^T\right]_{ii} \\ &= (GAG^T)_{ii} + (HAH^T)_{ii} - 2(GAH^T)_{ii} \\ &= 2\Big[1 - \sum_{j=1}^{n}(GA)_{ij}\, h_{ij}\Big] = 2\,[1 - (GA)_{ii}\, h_{ii}] \\ &\overset{(2.12)}{=} 2\left[1 - (GAG^T)_{ii}\frac{h_{ii}}{g_{ii}}\right] = 2\left(1 - \frac{h_{ii}}{g_{ii}}\right), \\ &\qquad i = 1, \ldots, n.\end{aligned} \tag{3.15}$$

Лемма 3.1 доказана. □

Здесь уместно сделать три замечания. Во-первых, для фиксированного FSAI-переобуславливателя G разность $\|I - HL_A\|_F^2 - \|I - GL_A\|_F^2$ зависисит только от диагональных элементов H (см. (3.13)). Во-вторых, соотношение (3.11) дает представление $\|I - HL_A\|_F^2$ в виде суммы двух слагаемых, первое из которых зависит только от шаблона разреженности S, тогда как второе учитывает неточность в вычислении G. В-третьих, чем ближе $g_{ii}\,\ell_{ii}$ к нулю [т.е. (см. (3.12)) чем хуже качество i-ой строки точного FSAI-переобуславливателя G], тем с меньшим коэффициентом входит $(EAE^T)_{ii}$ [т.е. квадрат A-нормы i-ой строки матрицы ошибки $E = G - H$] в (3.11). Тем самым с точки зрения минимизации $\|I - HL_A\|_F^2$ наиболее существенным является найти качественные аппроксимации к тем строкам G, для которых соответствующие произведения $g_{ii}\,\ell_{ii}$ близки к 1, тогда как строки, для которых $g_{ii}\,\ell_{ii}$ малы, не требуют аппроксимации с высокой точностью.

Используя лемму 3.1, мы можем легко вывести нижнюю оценку в терминах величин, связанных с минимизацией фробениусовой нормы, для спектральной нормы матрицы $I - HG^{-1}$, участвующей в формулировке теоремы 3.1.

Следствие 3.1. *В предыдущих предположениях справедлива оценка*

$$\left\| I - HG^{-1} \right\|_2 \geqslant \Big[\frac{\|I - HL_A\|_F^2 - \|I - GL_A\|_F^2}{n} \Big]^{1/2}.$$

Доказательство. Принимая во внимание, что в силу (2.18) $\ell_{ii}\, g_{ii} \leqslant 1$, $i = 1, \ldots, n$, мы имеем:

$$\begin{aligned} \sum_{i=1}^{n} (EAE^T)_{ii}\, g_{ii}\, \ell_{ii} &\leqslant \operatorname{tr} EAE^T = \operatorname{tr} L_A^T\, E^T\, EL_A \\ &= \operatorname{tr} L_A^T\, G^T \left(G^{-T}\, E^T\, EG^{-1} \right) GL_A \\ &\leqslant \lambda_{\max} \left(G^{-T}\, E^T\, EG^{-1} \right) \operatorname{tr} L_A^T\, G^T\, GL_A \\ &= \left\| EG^{-1} \right\|_2^2 \operatorname{tr} GAG^T = n \left\| I - HG^{-1} \right\|_2^2, \end{aligned}$$

и для завершения доказательства остается лишь воспользоваться равенством (3.11). □

Заметим, что при $S = \{(i, j) : i < j\}$ FSAI-переобуславливатель G совпадает с обратной матрицей L_A^{-1}, априведенная выше нижняя оценка сводится к тривиальной оценке

$$\|I - HL_A\|_2 \geqslant \frac{\|I - HL_A\|_F}{\sqrt{2}}.$$

В действительности было бы очень полезно иметь нетривиальную верхнюю оценку для спектральной нормы $\left\| I - HG^{-1} \right\|_2$ в терминах величин, связанных с целевым функционалом. К сожалению, нам удалось вывести только почти тривиальную оценку

$$\begin{aligned} \left\| I - HG^{-1} \right\|_2^2 &= \left\| EG^{-1} \right\|_2^2 \leqslant \|EL\|_2^2 \left\| L^{-1}\, G^{-1} \right\|_2^2 \\ &\leqslant \frac{\operatorname{tr} EAE^T}{\lambda_{\min}(GAG^T)} \overset{(3.15)}{=} 2 \frac{\sum_{i=1}^{n} \left(1 - \frac{h_{ii}}{g_{ii}} \right)}{\lambda_{\min}(GAG^T)}. \end{aligned} \tag{3.16}$$

Эта оценка, в частности, показывает, что если

$$\frac{g_{ii} - h_{ii}}{g_{ii}} < \frac{\lambda_{\min}(GAG^T)}{2n}, \quad i = 1, \ldots, n, \tag{3.17}$$

то заведомо

$$\left\| I - HG^{-1} \right\|_2^2 < 1,$$

а значит теорема 3.1 оказывается применимой.

Теперь мы выведем двусторонние оценки для величин $(EAE^T)_{ii}$, участвующих в (3.11). Для этого мы воспользуемся следующим результатом.

Лемма 3.2. *В предыдущих обозначениях имеют место соотношения*

$$h_{ii}^{-2} - g_{ii}^{-2} = (\,\widetilde{E}\,A\,\widetilde{E}^T\,)_{ii}, \quad i = 1, \dots, n, \tag{3.18}$$

где $\widetilde{E} = \widetilde{G} - \widetilde{H}$ *– промежуточная матрица ошибок.*

Доказательство. Мы имеем

$$\begin{aligned}(\,\widetilde{E}\,A\,\widetilde{E}^T\,)_{ii} &= \Big[(\,\widetilde{G} - \widetilde{H}\,)A(\,\widetilde{G} - \widetilde{H}\,)^T\Big]_{ii} \\ &= (\,\widetilde{G}\,A\,\widetilde{G}^T\,)_{ii} + (\,\widetilde{H}\,A\,\widetilde{H}^T\,)_{ii} - 2\left(\widetilde{G}\,A\,\widetilde{G}^T\right)_{ii}, \\ &\qquad i = 1, \dots, n.\end{aligned} \tag{3.19}$$

Но поскольку в силу (2.3) и (3.4)

$$(\,\widetilde{G}\,A\,\widetilde{H}^T\,)_{ii} = \sum_{j=1}^{n} (\,\widetilde{G}\,A)_{ij}\,\tilde{h}_{ij} = (\,\widetilde{G}\,A)_{ii}, i = 1, \dots, n,$$

и, аналогично,

$$(\,\widetilde{G}\,A\,\widetilde{G}^T\,)_{ii} = (\,\widetilde{G}\,A)_{ii}, \quad i = 1, \dots, n,$$

из (3.19) следует, что

$$(\,\widetilde{E}\,A\,\widetilde{E}^T\,)_{ii} = (\,\widetilde{H}\,A\,\widetilde{H}^T\,)_{ii} - (\,\widetilde{G}\,A\,\widetilde{G}^T\,)_{ii}, \quad i = 1, \dots, n,$$

откуда (3.18) получается с помощью (2.3)–(2.5) и (3.4)–(3.6). □

Замечание 3.1. Равенство (3.18) показывает, в частности, что убывание $(\,\widetilde{E}\,A\,\widetilde{E}^T\,)_{ii}$, т.е. квадрата A-нормы i-ой строки матрицы $\widetilde{E}$, эквивалентно возрастанию соответствующего диагонального элемента h_{ii} приближенного FSAI-переобуславливателя H и (ввиду (3.10) убыванию $\|I - HL_A\|_F$.

Лемма 3.3. *В предыдущих обозначениях справедливы оценки*

$$h_{ii}^2(\,\widetilde{E}\,A\,\widetilde{E}^T\,)_{ii} \leqslant (EAE^T)_{ii} \leqslant 2h_{ii}^2(\,\widetilde{E}\,A\,\widetilde{E}^T\,)_{ii}, \quad i = 1, \dots, n; \tag{3.20}$$

$$(EAE^T)_{ii} \leqslant g_{ii}^2 (\, \widetilde{E}\, A\, \widetilde{E}^T\,)_{ii}, \quad i = 1, \ldots, n. \tag{3.21}$$

Доказательство. Чтобы доказать правое неравенство в (3.20), мы воспользуемся соотношениями (3.14) и леммой 3.2. Таким образом мы получаем

$$\begin{aligned}(EAE^T)_{ii} &= 2\Big(1 - \frac{h_{ii}}{g_{ii}}\Big) \leqslant 2\Big(1 - \frac{h_{ii}^2}{g_{ii}^2}\Big)\\ &= 2h_{ii}^2 (\, \widetilde{E}\, A\, \widetilde{E}^T\,)_{ii}, \quad i = 1, \ldots, n.\end{aligned}$$

Левое неравенство в (3.20) также выводится с использованием (3.14) и леммы 3.2:

$$\begin{aligned}(EAE^T)_{ii} &= 2\Big(1 - \frac{h_{ii}}{g_{ii}}\Big) = 1 + \Big(1 - \frac{h_{ii}}{g_{ii}}\Big)^2 - \frac{h_{ii}^2}{g_{ii}^2}\\ &\geqslant 1 - \frac{h_{ii}^2}{g_{ii}^2} = h_{ii}^2 (\, \widetilde{E}\, A\, \widetilde{E}^T\,)_{ii}, \quad i = 1, \ldots, n.\end{aligned}$$

Наконец, неравенство (3.21) получается с помощью леммы 3.2 и неравенств (3.8) следующим образом:

$$\begin{aligned}(EAE^T)_{ii} &= 2\Big(1 - \frac{h_{ii}}{g_{ii}}\Big) = 2h_{ii}\Big(\frac{1}{h_{ii}} - \frac{1}{g_{ii}}\Big) = h_{ii} \frac{(\, \widetilde{E}\, A\, \widetilde{E}^T\,)_{ii}}{\big(h_{ii}^{-1} + g_{ii}^{-1}\big)\,/2}\\ &\leqslant h_{ii} (\, \widetilde{E}\, A\, \widetilde{E}^T\,)_{ii}\, h_{ii}^{1/2}\, g_{ii}^{1/2} \leqslant g_{ii}^2 (\, \widetilde{E}\, A\, \widetilde{E}^T\,)_{ii}, \quad i = 1, \ldots, n.\end{aligned}$$

□

Теперь вместо пары матриц $\widetilde{H}$ и H мы рассмотрим две последовательности матриц $\{\, \widetilde{H}^{(k)}\}_{k \geqslant 1}$ и $\{H^{(k)}\}_{k \geqslant 1}$, где для каждого k матрицы $\widetilde{H}^{(k)}$ и $H^{(k)}$ удовлетворяют соответствующим аналогам соотношений (3.4)–(3.6). Мы предполагаем, что матрицы $\widetilde{H}^{(k)}$, $k \geqslant 1$, получены как последовательные приближения к матрице $\widetilde{G}$, определенной в (2.3), с помощью (переобусловленного) метода сопряженных градиентов. Более точно это можно описать следующим образом. Как уже отмечалось, основные уравнения системы (2.3) распадаются на независимые “локальные” системы уравнений вида

$$\widetilde{g}_i\, A_i = b_i, \quad i \geqslant 2, \quad n_i \geqslant 1, \tag{3.22}$$

где

$$\begin{aligned} &A_i = A[J_i, J_i], \quad \widetilde{g}_i = \widetilde{G}[\{i\}, J_i], \quad b_i = -A[\{i\}, J_i], \\ &J_i = \{j : j \neq i \,\&\, (i,j) \notin S\}, \quad |\, J_i \,| = n_i, \end{aligned} \tag{3.23}$$

а через $B[I, J]$ обозначена подматрица матрицы B, отвечающая строчным индексам из I и столбцовым индексам из J.

Мы предполагаем, что строчные векторы $\tilde{h}_i^{(k)} = \widetilde{H}^{(k)}[\{i\}, J_i]$, $i \geqslant 2$, $n_i \geqslant 1$, $k \geqslant 1$, которые вполне определяют матрицу $\widetilde{H}^{(k)}$, являются приближенными решениями соответствующей локальной системы (3.22), полученными после k итераций PCG-метода [7]. Тогда, очевидно, мы имеем:

$$(\widetilde{E}^{(k)} A \, \widetilde{E}^{(k)^T})_{ii} = \begin{cases} 0, & n_i = 0, \\ \tilde{e}_i^{(k)} A_i \, \tilde{e}_i^{(k)^T}, & n_i \geqslant 1, \end{cases} \qquad i = 1, \ldots, n, \tag{3.24}$$

где $\widetilde{E}^{(k)} = \widetilde{G} - \widetilde{H}^{(k)}$ и $\tilde{e}_i^{(k)} = \tilde{g}_i - \tilde{h}_i^{(k)}$, $i = 1, \ldots, n$, $k \geqslant 0$. Применяя известные оценки для PCG-метода [7], мы получаем оценку

$$\begin{gathered} (\widetilde{E}^{(k)} A \, \widetilde{E}^{(k)^T})_{ii} \leqslant 4\Big(\frac{\sqrt{\varkappa_i} - 1}{\sqrt{\varkappa_i} + 1}\Big)^{2k} (\widetilde{E}^{(0)} A \, \widetilde{E}^{(0)^T})_{ii}, \\ n_i \geqslant 1, \quad k \geqslant 1, \end{gathered} \tag{3.25}$$

где $\varkappa_i = \lambda_{\max}(M_i^{-1} A_i) / \lambda_{\min}(M_i^{-1} A_i)$ – это спектральное число обусловленности i-ой переобусловленной матрицы $M_i^{-1} A_i$ и $A_i = M_i - N_i$.

Принимая во внимание леммы 3.1 и 3.3 и оценку (3.25), мы приходим к следующей верхней оценке для $\big\| I - H^{(k)} L_A \big\|_F^2$ как функции от k.

Теорема 3.2. *Пусть для фиксированной $n \times n$ SPD-матрицы A и фиксированного шаблона разреженности S вида (2.1) матрица $G = (g_{ij})_{i,j=1}^n$ есть точный FSAI-переобуславливатель, определенный соотношениями (2.3)–(2.5), и пусть последовательности $\{H^{(k)}\}$ и $\{\widetilde{H}^{(k)}\}$ $n \times n$ матриц с тем же шаблоном разреженности S, удовлетворяющих условиям*

$$\begin{gathered} (\widetilde{H}^{(k)})_{ii} = 1, \quad H^{(k)} = \widetilde{D}^{(k)^{-1}} \widetilde{H}^{(k)}, \quad (H^{(k)} A \, H^{(k)^T})_{ii} = 1, \\ i = 1, \ldots, n, \quad k \geqslant 1, \end{gathered}$$

построены в результате применения PCG метода к локальным системам (3.22)–(3.23). Тогда

$$\begin{gathered}\|I - GL_A\|_F^2 \leqslant \left\| I - H^{(k)} L_A \right\|_F^2 \leqslant \|I - GL_A\|_F^2 \\ +4 \sum_{i: n_i \geqslant 1} g_{ii}^2 \Big(\frac{\sqrt{\varkappa_i} - 1}{\sqrt{\varkappa_i} + 1} \Big)^{2k} (\widetilde{E}^{(0)} A \, \widetilde{E}^{(0)T})_{ii} \, g_{ii} \, \ell_{ii}, \quad k \geqslant 1,\end{gathered} \tag{3.26}$$

где для $1 \leqslant i \leqslant n$ таких, что $n_i \geqslant 1$, $\varkappa_i = \lambda_{\max}(M_i^{-1} A_i) / \lambda_{\min}(M_i^{-1} A_i)$, $A_i = M_i - N_i$, а $\widetilde{E}^{(0)} = \widetilde{G} - \widetilde{H}^{(0)}$ – это начальная промежуточная матрица ошибок.

Замечание 3.2. Ясно, что при построении приближенного FSAI-переобуславливателя H с помощью PCG-метода для различных локальных систем можно выполнить различное число PCG-итераций. В этом случае в правой части (3.26) следует заменить k на соответствующее k_i.

Наш теоретический анализ итерационного построения приближенных FSAI-переобуславливателей мы завершим рассмотрение функционала Капорина $\beta(H^{(k)} A \, H^{(k)^T}) = \beta(H^{(k)^T} H^{(k)} A)$ как функции от k. Значение этого функционала для теории переобуславливания определяется следующей теоремой.

Теорема 3.3 [14]. *Пусть A есть SPD-матрица и $M \approx A^{-1}$ – некоторый SPD-переобуславливатель для A. Тогда для k-го приближения PCG-метода y_k к решению системы $Ay = f$ справедлива оценка*

$$\|r_k\|_M \leqslant \Big[\beta(MA)^{n/k} - 1 \Big]^{k/2} \|r_0\|_M, \quad 1 \leqslant k \leqslant n - 1,$$

где $r_k = f - Ay_k$.

Таким образом с помощью β (если только $\beta < 2$ достаточно близко к 1) можно оценить скорость сходимости PCG-метода в терминах соответствующих норм невязок. Следовательно, ухудшение качества переобуславливания, связанное с переходом от точного FSAI-переобуславливателя G к приближенному FSAI-переобуславливателю H, может быть оценено в терминах отношения

$$
\begin{aligned}
\frac{\beta(HAH^T)}{\beta(GAG^T)} = \Bigl(\prod_{i=1}^{n} \frac{g_{ii}}{h_{ii}}\Bigr)^{2/n} &= \Bigl[\prod_{i=1}^{n} \Bigl(1 + g_{ii}^2(\,\widetilde{E}\,A\,\widetilde{E}^T\,)\Bigr)_{ii}\Bigr]^{1/n} \\
&\leqslant 1 + \max_{i:n_i \geqslant 1} \Bigl\{ g_{ii}^2(\,\widetilde{E}\,A\,\widetilde{E}^T)_{ii} \Bigr\},
\end{aligned}
\tag{3.27}
$$

где мы использовали лемму 3.2. Заметим, что по теореме 2.2 $\beta(HAH^T)/\beta(GAG^T) \geqslant 1$.

Из (3.27) и (3.25) непосредственно вытекает следующая оценка для отношения $\beta(H^{(k)}\,A\,{H^{(k)}}^T)/\beta(GAG^T)$, где матрицы $H^{(k)}$ являются PCG-аппроксимациями к G, имеющими тот же шаблон разреженности.

Теорема 3.4. *В предположениях теоремы* 3.2 *справедливы неравенства*

$$
\begin{aligned}
1 &\leqslant \frac{\beta(H^{(k)}\,A\,{H^{(k)}}^T)}{\beta(GAG^T)} \\
&\leqslant 1 + 4 \max_{i:n_i \geqslant 1} \Bigl\{ g_{ii}^2 \Bigl(\frac{\sqrt{\varkappa_i} - 1}{\sqrt{\varkappa_i} + 1}\Bigr)^{2k} (\,\widetilde{E}^{(0)}\,A\,{\widetilde{E}^{(0)}}^T)_{ii} \Bigr\}, \ k \geqslant 1.
\end{aligned}
$$

§4. Численные результаты

В этом параграфе представлены результаты численных экспериментов по использованию приближенных FSAI-переобуславливателей вместо точных при решении линейных систем уравнений с симметричными, положительно определенными матрицами коэффициентов, возникающих из конечно-элементных аппроксимаций задач линейной теории упругости.

Для наших экспериментов были выбраны две задачи линейной теории упругости для ортотропных материалов, которые ранее рассматривались в [20]. Первая из них соответствует кубической области с нулевыми перемещениями на трех попарно-смежных гранях в качестве краевых условий. Вторая задача – это стандартная задача вычислительной механики о швейлере. Три тестовые системы линейных уравнений вида $Ax = b$ с симметричными, положительно определенными матрицами коэффициентов*

*Тестовые линейные системы можно получить от авторов в ASCII-формате через ftp.

были получены применением иерархической p-версии конформного метода конечных элементов с $p = 3$ для кубической области и с $p = 1, 4$ для задачи о швейлере к соответствующим вариационным постановкам (подробнее см. в [20]) и симметричным масштабированием по Якоби результируюших матриц коэффициентов. Первая линейная система, отвечающая кубической области, достаточно хорошо обусловлена (cond$A = 1.03 \cdot 10^7$), а две другие системы являются весьма плохообусловленными (cond$A > 2.7 \cdot 10^9$).

Для каждой из тестовых систем были построены (с помощью метода Холецкого) точные FSAI-переобуславливатели, соответствующие двум шаблонам разреженности, называемым "исходным" и "расширенным." "Исходный" шаблон разреженности совпадает с шаблоном разреженности нижней треугольной части рассматриваемой матрицы коэффициентов, а "расширенный" шаблон – это шаблон с одним уровнем заполнения.

Приближенные FSAI-переобуславливатели были построены с помощью (несимметрично) переобусловленного блочного метода сопряженных градиентов (PBCG-метода) [21]. Выбор блочного варианта CG-метода связан с тем, что для рассматриваемых тестовых задач естественно возникают локальные системы с одними и теми же матрицами коэффициентов и несколькими ($\geqslant 3$) правыми частями. Далее, как следует из теоремы 2 в [21], PBCG-метод обладает свойством минимизации A-норм векторов ошибки, что гарантирует монотонную сходимость диагональных элементов последовательно строящихся приближенных FSAI-переобуславливателей (см. §3). Наконец, следует отметить, что при решении локальных систем было использовано их FSAI-переобуславливание с шаблонами разреженности, дублирующими шаблоны нижних треугольных частей "локальных" матриц коэффициентов.

Основной практический вывод из приведенных в §3 рассмотрений состоит в том, что последовательности $\{h_{ii}^{(k)}\}$ монотонно сходятся к соответствующим элементам g_{ii}, причем значения диагональных элементов $h_{ii}^{(k)}$, определяющие значения $\|I - H^{(k)} L_A\|_F$ и $\beta(H^{(k)} A\, H^{(k)^T})$, могут быть приняты в качестве меры качества соответствующих строк переобуславливателя $H^{(k)}$.

Основываясь на таком понимании, для остановки локальных итераций мы использовали следующий критерий:

$$\max_{1 \leqslant i \leqslant n} \frac{h_{ii}^{(k+1)} - h_{ii}^{(k)}}{h_{ii}^{(k)}} \leqslant \mathrm{tol} \quad (\mathrm{tol} > 0). \tag{4.1}$$

Заметим, что на каждой итерации для i-ой локальной системы, определяющей i-ую строку матрицы $\widetilde{H}^{(k)}$, необходимо вычислять величину

$$h_{ii}^{(k)^{-1}} = (\, \widetilde{H}^{(k)} A \, \widetilde{H}^{(k)^T})_{ii},$$

тогда как внедиагональные элементы окончательной переобуславливающей матрицы H вычисляются один раз после окончания всех итерационных процессов решения локальных систем.

Численные результаты для трех тестовых линейных систем, для переобуславливания каждой из которых использовалось несколько точных и приближенных FSAI-переобуславливателей, представлены в таблице 1. В таблице приняты следующие обозначения:

N – порядок линейной системы;

NZ – число ненулевых элементов в матрице коэффициентов;

tol – пороговый параметр, используемый в локальном критерии остановки (4.1), а tol $= 0$ означает, что переобуславливатель вычислен с помощью метода Холецкого;

IT_{loc} – число локальных итераций, необходимых для удовлетворения критерия (4.1) при соответствующем значении tol;

IT – число глобальных FSAI-CG-итераций, требующихся для удовлетворения критерия остановки глобальных итераций

$$\log_{10}(\|r_k\|_2/\|r_0\|_2) < -10,$$

где r_k – это невязка на k-ой итерации PCG-метода;

$\lambda_{\max}(\lambda_{\min})$ – наибольшее (наименьшее) собственное значение трехдиагональной матрицы, составленной из коэффициентов PCG-метода;

cond $= \lambda_{\max}/\lambda_{\min}$

Анализируя данные, приведенные в таблице 1, можно сделать два основных вывода. Во-первых, локальный критерий остановки

(4.1) работает вполне удовлетворительно, и число необходимых глобальных итераций монотонно убывает при уменьшении значения tol. Интересно отметить, что для итерационно построенных FSAI-переобуславливателей число глобальных итераций, необходимых для удовлетворения глобального критерия остановки, может быть даже меньшим, чем для соответствующего "точного" переобуславливателя (см. данные для тестовых систем 2 и 3), если только значение tol, используемое в локальном критерии остановки (4.1) достаточно мало. Объяснением этого явления мы пока не располагаем. Во-вторых, если матрица коэффициентов достаточно хорошо обусловлена и/или точный FSAI-переобуславливатель имеет относительно невысокое качество, то разумным представляется выбор $\mathrm{tol} = 10^{-1}$ или $\mathrm{tol} = 10^{-2}$, тогда как для плохо обусловленных систем и относительно качественных FSAI-переобуславливателей значение tol следует уменьшить до $10^{-3} - 10^{-4}$.

Таблица 1.

Тестовая система 1: N = 5040, NZ = 688108

S	tol	$\mathrm{IT}_{\mathrm{loc}}$	IT	$\lambda_{\min}$	$\lambda_{\max}$	COND
Исход.	0.0	–	218	8.9145e-03	3.9629	444.55
Исход.	1.0-e1	1	218	8.9145e-03	3.9629	444.55
Расшир.	0.0	–	73	3.4061e-02	2.0882	61.307
Расшир.	1.0e-1	3	82	2.7588e-02	2.2096	80.095
Расшир.	1.0e-2	4	76	3.1572e-02	2.1416	67.830
Расшир.	1.0e-3	8	73	3.3242e-02	2.0910	62.901
Расшир.	1.0e-4	12	73	3.3905e-02	2.0876	61.571

Тестовая система 2: N = 3088, NZ = 733872

S	tol	$\mathrm{IT}_{\mathrm{loc}}$	IT	$\lambda_{\min}$	$\lambda_{\max}$	COND
Исход.	0.0	–	2855	3.8945e-07	4.0550	1.0412e+07
Исход.	1.0-e2	1	2823	3.8945e-07	4.0550	1.0412e+07
Расшир.	0.0	–	418	3.1321e-06	2.0657	6.5952e+05
Расшир.	1.0e-1	9	440	3.0968e-06	2.2011	7.1075e+05
Расшир.	1.0e-2	11	425	3.1796e-06	2.0703	6.5112e+05
Расшир.	1.0e-3	21	417	3.1602e-06	2.0638	6.5306e+05
Расшир.	1.0e-4	28	411	3.1403e-06	2.0635	6.5712e+05

Тестовая система 3: N = 6183, NZ = 790755

S	tol	$\mathrm{IT}_{\mathrm{loc}}$	IT	$\lambda_{\min}$	$\lambda_{\max}$	COND
Исход.	0.0	–	> 3000	5.8359e-07	4.0447	6.9307e+06
Расшир.	0.0	–	1065	2.1491e-06	2.8593	1.3304e+06
Расшир.	1.0e-1	9	1385	1,3032e-06	3.1800	2.4401e+06
Расшир.	1.0e-2	33	1172	1.8281e-06	2.9511	1.6143e+06
Расшир.	1.0e-3	43	1050	2.1332e-06	2.8609	1.3411e+06
Расшир.	1.0e-4	60	1025	2.0805e-06	2.8604	1.3748e+06

§5. Выводы и заключительные замечания

В работе представлены новые результаты теоретического и экспериментального исследования переобуславливаний с помощью факторизованных разреженных приближенных обратных. Объяснена важность включения масштабирования по Якоби в конструкцию FSAI-переобуславливателей. Также показано, что точные FSAI-переобуславливатели можно заменить на приближенные практически без потери качества переобуславливания, что достигается благодаря использованию простого критерия остановки локальных итераций, основанного на монотонной сходимости диагональных элементов строящейся последовательности приближенных матриц-переобуславливателей. Важно подчеркнуть, что приближенные FSAI-переобуславливатели не требуется вычислять с большой точностью, и, как свидетельствуют имеющиеся экспериментальные данные, значение $\mathrm{tol} = 10^{-3}$ в (4.1) является приемлемым на практике.

Другой вывод, вытекающий из представленных экспериментальных данных, состоит в том, что (по крайней мере для рассматриваемого класса задач) разумный шаблон разреженности FSAI-переобуславливателя может быть предписан априори, и одним из возможных вариантов выбора является шаблон, называемый в §4 расширенным.

Наконец, следует отметить, что в данной статье мы сознательно избегали обсуждения вопросов, связанных со сравнением точных и приближенных FSAI-переобуславливателей (в предположении, что точные переобуславливатели действительно могут быть построены) с точки зрения их эффективности. Это связано с тем, что без фиксирования достаточно узкого класса задач и вычислительной платформы такое сравнение было бы спекулятивным. В соответствии с этим, основной целью данного исследования было продемонстрировать принципиальную возможность замены точных FSAI-переобуславливателей на приближенные без значительного ухудшения результирующей скорости сходимости.

Литература

1. И. Е. Капорин, *Альтернативный подход к оценке числа итераций метода сопряженных градиентов*. — В кн.: Численные методы и программное обеспечение. М., Отдел вычислительной математики АН СССР, 1990, 55–72.

2. M. W. Benson, P. O. Frederickson, *Iterative solution of large sparse linear systems arising in certain multidimensional approximation problems.* — Utilitas Math. **22** (1982), 127–140.
3. M. Benzi, C. D. Meyer, M. Tůma, *A sparse approximate inverse preconditioner for the conjugate gradient method.* — SIAM J. Sci. Comput. **17** (1996).
4. M. Benzi, M. Tůma, *A sparse approximate inverse preconditioner for nonsymmetric linear systems.* Tech. Rep. No. 653, Oct. 1995, Inst. Computer Science, Academy of Sciences of the Czech Republic, 1995.
5. E. Chow, Y. Saad, *Approximate inverse preconditioners for general sparse matrices.* Res. Rep. UMSI 94/101, Univ. of Minnesota, Supercomputer Inst., Minneapolis, Minnesota, USA, 1994.
6. E. Chow, Y. Saad, *Approximate inverse techniques for block-partitioned matrices.* Res. Rep. UMSI 95/13, Univ. of Minnesota, Supercomputer Inst., Minneapolis, Minnesota, USA, 1995.
7. P. Concus, G. H. Golub, D. P. O'Leary, *A generalized conjugate gradient method for the numerical solution of elliptic partial differential equations.* Sparse Matrix Computations, под ред. J. R. Bunch & D. J. Rose, New York, Academic Press, 1976, 309–332.
8. P. Dewilde, E. F. Deprettere, *Approximate inversion of positive matrices with applications to modelling.* Modelling, Robustness, and Sensitivity Reduction in Control Systems, под ред. R. F. Curtain, NATO ASI Series, **F 34** (1987), 211–238.
9. N. I. M. Gould, J. A. Scott, *On approximate-inverse preconditioners.* Tech. Rep. RAL 95-026, Rutherford Appleton Lab., Chilton, England, 1995.
10. M. Grote, H. Simon, *Parallel preconditioning and approximate inverses on the Connection Machine.* Proceedings of the Sixth SIAM Conference on Parallel Processing for Scientific Computing под ред. R. Sincovec и др., Philadelphia, SIAM, 1993, 519–523.
11. Y. P. Hong, C. T. Pan, *A lower bound for the smallest singular value.* — Linear Algebra Appl. **172** (1992), 27–32.
12. R. Horn, C. Johnson, *Topics in Matrix Analysis.* New York, Cambridge Univ. Press., 1991.
13. T. Huckle, M. Grote, *A new approach to parallel preconditioning with sparse approximate inverses.* Manuscript SCCM 94-03, Scietific Computing and Computational Mathematics Program, Stanford Univ., Stanford, CA, USA, 1994.
14. I. E. Kaporin, *Explicitly preconditioned conjugate gradient method for the solution of unsymmetric linear systems.* — Int. J. Comput. Math. **40** (1992), 169–187.
15. I. E. Kaporin, *New convergence results and preconditioning strategies for the conjugate gradient method.* — Numer. Linear Algebra Appl. **1** (1994), 179–210.
16. I. E. Kaporin, *The trace/determinant condition number: several properties and application to polynomial preconditionings.* — Numer. Linear Algebra Appl.
17. L. Yu. Kolotilina, A. Yu. Yeremin, *On a family of two-level preconditionings of the incomplete block factorization type.* — Sov. J. Numer. Anal. Math. Model. **1** (1986), 293–320.
18. L. Yu. Kolotilina, A. Yu. Yeremin, *Factorized sparse approximate inverse preconditionings. I: Theory.* — SIAM J. Matrix Anal. Appl. **14** (1993), 45–58.
19. L. Yu. Kolotilina, 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.
20. L. Yu. Kolotilina, A. Yu. Yeremin, *Block SSOR preconditionings for high order FE systems, III: Incomplete BSSOR preconditionings.* — Numer. Linear Algebra Appl..
21. D. P. O'Leary, *The block conjugate gradient algorithm and related methods.* — Linear Algebra Appl. **29** (1980), 293–322.

Yeremin A. Yu., Kolotilina L. Yu., and Nikishin A. A. Factorized sparse approximate inverse preconditionings III. Iterative construction of preconditioners.

This paper presents new results of the theoretical study of factorized sparse approximate inverse (FSAI) preconditioners. In particular, the effect of the a posteriori Jacobi scaling and the possibility of constructing such preconditioners iteratively are analyzed. A simple stopping criterion for the termination of local iterations for constructing approximate FSAI preconditioners using the PCG method is proposed. The results of numirical experimernts with 3D finite-element problems from linear elasticity are presented.

С.-Петербургское отделение
Математического института
им.В. А. Стеклова РАН

Поступило 3 ноября 1997 г.

Научно-исследовательский
вычислительный центр
Московского государственного университета

Вычислительный центр РАН
г. Москва