А. Джордж, Х. Д. Икрамов, Блочное LU-разложение устойчиво для матриц, обратных к матрицам с преобладающей блочной диагональю, *Зап. научн. сем. ПОМИ*, 2003, том 296, 15–26

Записки научных
семинаров ПОМИ
Том 296, 2003 г.

А. Джордж, Х. Д. Икрамов

БЛОЧНОЕ LU-РАЗЛОЖЕНИЕ УСТОЙЧИВО ДЛЯ МАТРИЦ, ОБРАТНЫХ К МАТРИЦАМ С ПРЕОБЛАДАЮЩЕЙ БЛОЧНОЙ ДИАГОНАЛЬЮ

1. Введение

Пусть $M_n(\mathbf{C})$ — множество комплексных $n \times n$ матриц. Говорят, что матрица $B \in M_n(\mathbf{C})$ имеет диагональное преобладание (по строкам), если

$$\sigma_i |b_{ii}| = \sum_{\substack{j=1 \\ j \neq i}}^{n} |b_{ij}|, \quad i = 1, \ldots, n, \tag{1}$$

где $0 \leqslant \sigma_i \leqslant 1$, $i = 1, \ldots, n$. Для краткости, мы будем говорить о dd-матрицах. Число

$$\sigma = \max_{1 \leqslant i \leqslant n} \sigma_i \tag{2}$$

будем называть *коэффициентом преобладания* матрицы B. Если $\sigma < 1$, то имеем *строгое* диагональное преобладание; в противном случае, говорим о *слабом* преобладании. Аналогичным образом определяется диагональное преобладание по столбцам.

Хорошо известны свойства dd-матриц в отношении гауссова исключения. Для удобства читателя мы суммируем их в приводимой ниже теореме. Для определенности, всюду в статье говорится о диагональном преобладании по строкам. Аналогичные факты могут быть доказаны для преобладания по столбцам.

Теорема 1. *Пусть $B \in M_n(\mathbf{C})$ — dd-матрица со строчными коэффициентами преобладания σ_i (см. (1)). При $\sigma = 1$ дополнительно предполагаем, что матрица B невырожденна. Тогда*

1. *Гауссово исключение применимо к B при любом способе выбора ведущих элементов на главной диагонали.*

Эта работа поддержана грантом OGP0008111 Канадского Совета по Исследованиям в области Естественных и Технических Наук.

2. *Пусть* $B^{(k)} = (b_{ij}^{(k)})$ *— матрица, полученная из* B *в результате* $k-1$ *шагов гауссова исключения. Тогда активная подматрица матрицы* $B^{(k)}$ *(иными словами, дополнение Шура ведущей главной подматрицы порядка* $k-1$ *в матрице* B*) также является dd-матрицей. Более того, для каждого значения* i *строчный коэффициент преобладания* σ_i' *для активной подматрицы не превосходит соответствующего коэффициента* σ_i *для матрицы* B *(предполагается, что строкам активной подматрицы приписаны те же индексы, что и в исходной матрице* B*).*
3. *Коэффициент роста для матрицы* B*, определяемый формулой*

$$\rho_n(B) = \frac{\max_{i,j,k} |b_{ij}^{(k)}|}{\max_{i,j} |b_{ij}|}, \tag{3}$$

ограничен числом $1+\sigma$*:*

$$\rho_n(B) \leqslant 1 + \sigma. \tag{4}$$

Ранее (см. [1]) авторы показали, что матрицы, обратные к dd-матрицам, имеют в отношении гауссова исключения по существу те же свойства, что перечислены в теореме 1. Опять-таки для удобства читателя, мы объединим эти свойства в Теореме 2.

Теорема 2. *Пусть* $A \in M_n(\mathbf{C})$ *— невырожденная матрица, причем* $B = A^{-1}$ *есть dd-матрица, удовлетворяющая соотношениям* (1). *Тогда*

1. *Для каждого значения* i $(i = 1, \dots, n)$ *наибольший модуль в* i*-м столбце матрицы* A *имеет диагональный элемент* a_{ii}*. Более точно, выполняются неравенства*

$$|a_{ji}| \leqslant \sigma_j |a_{ii}|, \qquad j \neq i. \tag{5}$$

2. *Свойство быть обратной к dd-матрице наследуется активными подматрицами (дополнениями Шура), получаемыми в ходе гауссова исключения. Более того, строчные коэффициенты преобладания для матрицы, обратной к дополнению Шура, не превосходят соответствующих коэффициентов* σ_i *для матрицы* B.
3. *Коэффициент роста для* A *ограничен числом* $1+\sigma$*:*

$$\rho_n(A) \leqslant 1 + \sigma. \tag{6}$$

Предположим теперь, что для $n \times n$-матриц фиксирован некоторый тип блочного разбиения, так что $B \in M_n(\mathbf{C})$ может быть представлена как блочная матрица

$$B = \begin{pmatrix} B_{11} & B_{12} \cdots & B_{1m} \\ B_{21} & B_{22} \cdots & B_{2m} \\ \cdots & \cdots \cdots \cdots & \cdots \\ B_{m1} & B_{m2} \cdots & B_{mm} \end{pmatrix} \tag{7}$$

с квадратными блоками B_{ii} $(i = 1, \ldots, m)$. Предположим также, что фиксирована некоторая матричная норма $\|\cdot\|$. Следуя [2, раздел 12.3.1], можно считать $\|\cdot\|$ произвольной подчиненной нормой. Однако для наших целей достаточно, чтобы выбранная норма была мультипликативной и подчинялась условию $\|I\| = 1$ для единичной матрицы I любого порядка. Например, норма

$$\|B\| = \max\{\|B\|_1, \|B\|_\infty\}$$

обладает обоими указанными свойствами, но не подчинена никакой векторной норме.

Говорят, что блочная матрица (7) имеет блочное диагональное преобладание (по строкам), если все ее диагональные блоки $B_{11}, \ldots, B_{mm}$ невырожденны и

$$\|B_{ii}^{-1}\| \sum_{\substack{j=1 \\ j \neq i}}^{m} \|B_{ij}\| = \sigma_i, \qquad i = 1, \ldots, m, \tag{8}$$

где $0 \leqslant \sigma_i \leqslant 1$, $i = 1, \ldots, m$. Для краткости, будем говорить о bdd-матрицах. Число

$$\sigma = \max_{1 \leqslant i \leqslant m} \sigma_i \tag{9}$$

называется *коэффициентом блочного преобладания* матрицы B. Заметим, что определение (12.16) в [2] соответствует соотношениям (8), в которых все коэффициенты σ_i равны 1 (слабое блочное диагональное преобладание; wbdd-матрицы). Аналогичным образом определяется блочное диагональное преобладание по столбцам.

В [2] доказано (теоремы 12.5 и 12.6; см. также [3, 4]), что свойства bdd-матриц в отношении блочного LU-разложения аналогичны свойствам скалярных dd-матриц, перечисленным в теореме 1. Мы снова соберем эти свойства в следующей теореме:

Теорема 3. *Предположим, что матрица $B \in M_n(\mathbf{C})$, представленная в блочной форме (7), является bdd-матрицей. Если B — wbdd-матрица, то дополнительно предположим, что B невырожденна. Тогда*

1. *B допускает блочное LU-разложение, в котором L и U имеют то же разбиение на блоки, что и матрица (7).*
2. *Пусть $B^{(k)} = (B_{ij}^{(k)})$ — матрица, полученная из B в результате $k-1$ шагов блочного исключения. Тогда активная подматрица матрицы $B^{(k)}$, т.е. блочная матрица*

$$\begin{pmatrix} B_{kk}^{(k)} \cdots B_{km}^{(k)} \\ \cdots\ \cdots\ \cdots \\ B_{mk}^{(k)} \cdots B_{mm}^{(k)} \end{pmatrix}, \tag{10}$$

   также является bdd-матрицей. Более того, для $i = k, \dots, m$ коэффициент блочного преобладания σ_i' строки i матрицы (10) не превосходит соответствующего коэффициента σ_i для матрицы B.
3. *Коэффициент роста в LU-разложении, определяемый формулой*

$$\rho_m(B) = \frac{\max_{i,j,k} \|B_{ij}^{(k)}\|}{\max_{i,j} \|B_{ij}\|}, \tag{11}$$

   ограничен числом $1+\sigma$:

$$\rho_m(B) \leqslant 1 + \sigma. \tag{12}$$

Поскольку доказательство теоремы 3, приведенное в [2], покрывает лишь случай $\sigma = 1$, мы даем собственное доказательство для $\sigma < 1$ в разделе 2.

В отзыве рецензента нашей статьи [1] содержался вопрос: можно ли распространить результаты теоремы 2 на случай блочного диагонального преобладания? Ответ на этот вопрос положительный, и основной целью данной статьи является доказательство следующей теоремы:

Теорема 4. *Пусть $A \in M_n(\mathbf{C})$ — невырожденная матрица, причем $B = A^{-1}$ есть bdd-матрица, удовлетворяющая соотношениям (8) при заданном разбиении на блоки типа разбиения (7) и заданной матричной норме. Тогда*

1. *Для каждого значения i ($i = 1, \dots, m$) матрица A_{ii} невырожденна и выполняются неравенства*
$$\|A_{ji}A_{ii}^{-1}\| \leqslant \sigma_j, \qquad j \neq i. \tag{13}$$
2. *Свойство быть обратной к bdd-матрице наследуется активными подматрицами (дополнениями Шура), получаемыми в ходе блочного LU-разложения. Более того, строчные коэффициенты блочного преобладания для матрицы, обратной к дополнению Шура, не превосходят соответствующих коэффициентов σ_i для матрицы B.*
3. *Коэффициент роста для A (определяемый, как в (11)) ограничен числом $1 + \sigma$:*
$$\rho_m(A) \leqslant 1 + \sigma. \tag{14}$$

Теорема 4 доказана в разделе 3. Заметим, что из (13) вытекает неравенство
$$\|A_{ji}\| \leqslant \sigma_j\|A_{ii}\|, \qquad j \neq i.$$
В самом деле,
$$\|A_{ji}\| = \|A_{ji}A_{ii}^{-1}A_{ii}\| \leqslant \|A_{ji}A_{ii}^{-1}\|\|A_{ii}\| \leqslant \sigma_j\|A_{ii}\|.$$

2. Доказательство теоремы 3

Начнем с наблюдения, которое пригодится в последующих рассуждениях, а именно: равенство (8) влечет за собой соотношение
$$\sum_{\substack{j=1\\ j\neq i}}^{m} \|B_{ij}\| \leqslant \sigma_i\|B_{ii}\|, \qquad i = 1, \dots, m. \tag{15}$$
Действительно,
$$\sum_{\substack{j=1\\ j\neq i}}^{m} \|B_{ij}\| = \sum_{\substack{j=1\\ j\neq i}}^{m} \|B_{ii}B_{ii}^{-1}B_{ij}\|$$
$$\leqslant \|B_{ii}\|\|B_{ii}^{-1}\| \sum_{\substack{j=1\\ j\neq i}}^{m} \|B_{ij}\| = \sigma_i\|B_{ii}\|.$$

Будет достаточно рассмотреть первый шаг блочного LU-разложения. Более того, поскольку блочные строки в B обрабатываются независимо друг от друга, достаточно показать, что вторая блочная строка сохраняет блочное преобладание при исключении блока B_{21}. Отсюда будет вытекать справедливость второго утверждения теоремы.

Прежде всего, нужно показать, что новый диагональный блок $B_{22}^{(2)}$ невырожден. Имеем

$$B_{22}^{(2)} = B_{22} - B_{21}B_{11}^{-1}B_{12} = B_{22}(I - \Delta), \qquad (16)$$

где $\Delta = B_{22}^{-1}B_{21}B_{11}^{-1}B_{12}$ и

$$\|\Delta\| \leqslant (\|B_{22}^{-1}\|\|B_{21}\|)(\|B_{11}^{-1}\|\|B_{12}\|) \leqslant \sigma^2 < 1.$$

Это означает, что матрица в соотношении (16) невырожденна. Кроме того,

$$\|(I-\Delta)^{-1}\| \leqslant \frac{1}{1-\|\Delta\|}, \qquad \|(B_{22}^{(2)})^{-1}\| \leqslant \frac{\|B_{22}^{-1}\|}{1-\|\Delta\|}. \qquad (17)$$

Вместо проверки неравенства

$$\|(B_{22}^{(2)})^{-1}\| \sum_{j=3}^{m} \|B_{2j}^{(2)}\| \leqslant \sigma_2$$

мы проверим сейчас более сильное неравенство (см. (17))

$$\|B_{22}^{-1}\| \left[\sum_{j=3}^{m} \|B_{2j}\| + \|B_{21}\|\|B_{11}^{-1}\| \sum_{j=3}^{m} \|B_{1j}\| \right] \leqslant$$
$$\sigma_2 - \|B_{22}^{-1}\|\|B_{21}\|\|B_{11}^{-1}\|\|B_{12}\|. \qquad (18)$$

В самом деле,

$$\|B_{22}^{-1}\| \left[\sum_{j=3}^{m} \|B_{2j}\| + \|B_{21}\|\|B_{11}^{-1}\| \sum_{j=3}^{m} \|B_{1j}\| \right] +$$
$$\|B_{22}^{-1}\|\|B_{21}\|\|B_{11}^{-1}\|\|B_{12}\| =$$
$$\|B_{22}^{-1}\| \left[\sum_{j=3}^{m} \|B_{2j}\| + \|B_{21}\|\|B_{11}^{-1}\| \sum_{j=2}^{m} \|B_{1j}\| \right] =$$

$$\|B_{22}^{-1}\|\left[\sum_{j=3}^{m}\|B_{2j}\|+\|B_{21}\|\sigma_1\right]\leqslant\|B_{22}^{-1}\|\sum_{\substack{j=1\\ j\neq 2}}^{m}\|B_{2j}\|=\sigma_2.$$

В этих выкладках использовалось равенство (8) при $i=1$ и $i=2$. Полученное неравенство есть лишь иная форма неравенства (18).

То обстоятельство, что блочное диагональное преобладание сохраняется при исключении внедиагональных блоков, устанавливает также первое утверждение теоремы. Для доказательства третьего утверждения заметим, что

$$\sum_{j=2}^{m}\|B_{2j}^{(2)}\|\leqslant\sum_{j=1}^{m}\|B_{2j}\|. \tag{19}$$

Это своего рода "закон невозрастания массы" для блочных строк bdd-матрицы. Доказывается он точно так же, как в [2, теорема 12.6], если заменить блочные столбцы блочными строками. Теперь мы выводим из (19) оценку

$$\|B_{22}^{(2)}\|\leqslant\sum_{j=1}^{m}\|B_{2j}\|=\|B_{22}\|+\sum_{\substack{j=1\\ j\neq 2}}^{m}\|B_{2j}\|\leqslant(1+\sigma_2)\|B_{22}\|.$$

Здесь было использовано соотношение (15). Аналогичным образом получаем

$$\|B_{ii}^{(k)}\|\leqslant(1+\sigma_i)\|B_{ii}\|,\qquad k=2,\dots,m,\quad i=k,\dots,m.$$

Это дает оценку (12). Теорема доказана.

Замечание. Пусть

$$B=LU \tag{20}$$

есть блочное разложение матрицы B, существование которого было только что установлено. Из доказательства видно, что U является bdd-матрицей со строчными коэффициентами блочного преобладания, не превосходящими соответственно $\sigma_1,\dots,\sigma_m$. Это верно независимо от того, в какой из матриц, L или U, диагональные блоки суть единичные матрицы. Соответственно, если B имеет блочное диагональное преобладание по столбцам, этот тип преобладания сохраняется матрицей L в (20).

Замечание. Анализ доказательства теоремы 3 показывает, что, в действительности, блочное диагональное преобладание сохраняется *любой* последовательностью исключений внедиагональных блоков, а не только последовательностью гауссова типа. Например, исключив блок B_{21}, мы могли бы затем исключить в полученной матрице блок (1,2) и в результате снова имели бы bdd-матрицу.

3. Доказательство теоремы 4

Начнем со следующего очевидного замечания. Любое симметричное переупорядочение строк и столбцов bdd-матрицы B, согласованное с принятым блочным разбиением и не меняющее порядок скалярных строк внутри каждой блочной строки, сохраняет свойство блочного диагонального преобладания. Аналогичное замечание справедливо в отношении матрицы A, обратной к bdd-матрице $B = A^{-1}$; т.е. при описанных выше переупорядочениях свойство быть обратной к bdd-матрице сохраняется.

Из сделанного только что замечания следует, что неравенства (13) достаточно доказать, скажем, для $j = m$, т.е. для блоков в последнем блочном столбце матрицы A. Для доказательства воспользуемся блочным LU-разложением (20) bdd-матрицы $B = A^{-1}$. Для определенности, предположим, что диагональные блоки в L суть единичные матрицы. Будем рассматривать все матрицы, участвующие в равенстве

$$UA = L^{-1},$$

как блочные и приравняем одноименные блоки последнего блочного столбца. Двигаясь по этому столбцу снизу вверх и учитывая, что последний блочный столбец у матрицы L^{-1} тот же, что у единичной матрицы I_n, последовательно получаем

$$U_{mm}A_{mm} = I,$$

$$U_{m-1,m-1}A_{m-1,m} + U_{m-1,m}A_{mm} = 0,$$

$$U_{m-2,m-2}A_{m-2,m} + U_{m-2,m-1}A_{m-1,m} + U_{m-2,m}A_{mm} = 0,$$

$$\dots \qquad \dots \qquad \dots$$

Первое из этих соотношений означает, что матрица A_{mm} невырожденна. Перепишем второе соотношение в виде

$$A_{m-1,m}A_{mm}^{-1} = -U_{m-1,m-1}^{-1}U_{m-1,m}.$$

Отсюда

$$\|A_{m-1,m}A_{mm}^{-1}\| \leqslant \|U_{m-1,m-1}^{-1}\|\|U_{m-1,m}\| \leqslant \sigma_{m-1}.$$

Предположим, что уже доказаны неравенства

$$\|A_{pm}A_{mm}^{-1}\| \leqslant \sigma_p, \qquad p = l+1, \dots, m-1. \tag{21}$$

Мы покажем, что

$$\|A_{lm}A_{mm}^{-1}\| \leqslant \sigma_l.$$

Имеем

$$U_{ll}A_{lm} + \sum_{p=l+1}^{m-1} U_{lp}A_{pm} + U_{lm}A_{mm} = 0,$$

откуда

$$\|A_{lm}A_{mm}^{-1}\| = \| - U_{ll}^{-1}\left(\sum_{p=l+1}^{m-1} U_{lp}A_{pm}A_{mm}^{-1} + U_{lm}\right)\| \leqslant$$

$$\|U_{ll}^{-1}\|\left(\sum_{p=l+1}^{m-1} \|U_{lp}\|\|A_{pm}A_{mm}^{-1}\| + \|U_{lm}\|\right) \leqslant$$

$$\|U_{ll}^{-1}\|\left(\sum_{p=l+1}^{m-1} \sigma_p\|U_{lp}\| + \|U_{lm}\|\right) \leqslant$$

$$\|U_{ll}^{-1}\| \sum_{p=l+1}^{m} \|U_{lp}\| \leqslant \sigma_l.$$

Здесь были использованы неравенства (21). Таким образом, первое утверждение теоремы доказано.

Для доказательства второго утверждения заметим, что активная подматрица матрицы $A^{(k)}$, полученной из A в результате $k-1$ шагов блочного разложения, иными словами, матрица

$$\begin{pmatrix} A_{kk}^{(k)} \cdots A_{km}^{(k)} \\ \cdots\cdots\ \cdots \\ A_{mk}^{(k)} \cdots A_{mm}^{(k)} \end{pmatrix} \tag{22}$$

является обратной к подматрице

$$\begin{pmatrix} B_{kk} & \cdots & B_{km} \\ \cdots & \cdots & \cdots \\ B_{mk} & \cdots & B_{mm} \end{pmatrix} \tag{23}$$

матрицы B. Поскольку эта подматрица есть bdd-матрица со строчными коэффициентами блочного преобладания, не превосходящими соответственно $\sigma_k, \ldots, \sigma_m$, мы получаем отсюда требуемое утверждение.

Для доказательства последнего утверждения введем новое разбиение матрицы B на блоки, а именно

$$B = \begin{pmatrix} \hat{B}_{11} & \hat{B}_{12} \\ \hat{B}_{21} & \hat{B}_{22} \end{pmatrix}, \tag{24}$$

где $\hat{B}_{22}$ есть подматрица (23). Определим блочную матрицу

$$C = \begin{pmatrix} \hat{B}_{11} - \hat{B}_{12}\hat{B}_{22}^{-1}\hat{B}_{21} & \hat{B}_{12} \\ 0 & \hat{B}_{22} \end{pmatrix}.$$

Ее можно получить, умножая (24) справа на матрицу

$$F = \begin{pmatrix} I_1 & 0 \\ -\hat{B}_{22}^{-1}\hat{B}_{21} & I_2 \end{pmatrix}.$$

В свою очередь, F можно получить, применяя к bdd-матрице

$$\begin{pmatrix} I_1 & 0 \\ \hat{B}_{21} & \hat{B}_{22} \end{pmatrix}$$

нечто вроде блочного исключения по Гауссу–Жордану: вначале в подматрице (23) исключаются внедиагональные блоки, а затем каждая блочная строка умножается слева на соответствующую матрицу B_{ii}^{-1}. Согласно замечанию, сделанному в конце раздела 2, блочное диагональное преобладание сохраняется операциями первого типа. Аналогичное свойство операций второго типа очевидно. Из сказанного следует, что если блок (2,1) матрицы F представить в виде

$$\begin{pmatrix} F_{k1} & \cdots & F_{k,k-1} \\ \cdots & \cdots & \cdots \\ F_{m1} & \cdots & F_{m,k-1} \end{pmatrix}, \tag{25}$$

то

$$\sum_{l=1}^{k-1} \|F_{il}\| \leqslant \sigma_i, \qquad i = k, \dots, m.$$

Далее, равенство

$$C = BF$$

влечет за собой

$$C^{-1} = F^{-1}A. \tag{26}$$

Заметим, что матрица $D = C^{-1}$ имеет вид

$$D = \begin{pmatrix} D_{11} & D_{12} \\ 0 & D_{22} \end{pmatrix},$$

где блоком D_{22} является матрица (22). Матрица $G = F^{-1}$ получается из F изменением знаков у поддиагональных элементов. Следовательно, если принять для блока (2,1) матрицы G такое же разбиение, как в (25), то

$$G_{ij} = -F_{ij}, \qquad i = k, \dots, m, \quad j = 1, \dots, k-1.$$

Теперь из соотношения (26) вытекают равенства

$$A_{ij}^{(k)} = -\sum_{l=1}^{k-1} F_{il}A_{lj} + A_{ij}.$$

Поэтому для $i, j = k, \dots, m$ имеем

$$\|A_{ij}^{(k)}\| \leqslant \max_{s,t} \|A_{st}\| \left(1 + \sum_{l=1}^{k-1} \|F_{il}\| \right) \leqslant (1 + \sigma_i) \max_{s,t} \|A_{st}\|.$$

Этим доказана оценка (14).

Литература

1. A. George, Kh. D. Ikramov, *Gaussian elimination is stable for the inverse of a diagonally dominant matrix.* — Math. Comp. (принята к печати).
2. N. J. Higham, *Accuracy and Stability of Numerical Algorithms.* — SIAM, Philadelphia, 1996.
3. J. W. Demmel, N. J. Higham, R. S. Schreiber, *Stability of block LU factorization.* — Numerical Linear Algebra with Applications, **2** (1995), 173–190.
4. Х. Д. Икрамов, *О блочном аналоге свойства диагонального преобладания.* — Вестник Моск. ун-та. Серия 15. Вычисл. мат. кибернет., 1983, No. 4, 52–55.

George A., Ikramov K. D. Block LU factorization is stable for block matrices whose inverses are block diagonally dominant.

Let $A \in M_n(\mathbf{C})$ and let its inverse $B = A^{-1}$ be represented as an $m \times m$ block matrix that is block diagonally dominant either by rows or by columns w.r.t. a certain matrix norm. We show that A possesses a block LU factorization w.r.t. the partitioning defined by B, and the growth factor for A in this factorization is bounded above by $1 + \sigma$, where $\sigma = \max_{1 \leqslant i \leqslant m} \sigma_i$ and the σ_i, $0 \leqslant \sigma_i \leqslant 1$, are the row (column) block dominance factors of B. Further, the off-diagonal blocks of A (and of its block Schur complements) satisfy the relations

$$\|A_{ji} A_{ii}^{-1}\| \leqslant \sigma_j, \qquad j \neq i.$$

School of Computer Science,
University of Waterloo,
Waterloo, Ontario, Canada
Московский
государственный университет

Поступило 17 марта 2003 г.