Отдельный оттиск

УСПЕХИ МАТЕМАТИЧЕСКИХ НАУК

ТОМ
XX
ВЫПУСК
6 (126)

1965

1965 г. ноябрь—декабрь *т. XX, вып. 6 (126)*

УСПЕХИ МАТЕМАТИЧЕСКИХ НАУК

КРИТЕРИЙ ПРЕДСТАВИМОСТИ МАТРИЦЫ В ВИДЕ ПРОИЗВЕДЕНИЯ НИЖНЕ-ТРЕУГОЛЬНОЙ И ВЕРХНЕ-ТРЕУГОЛЬНОЙ МАТРИЦ

В. М. Киржнер

Пусть A — матрица n-го порядка. Будем говорить, что матрица A T-факторизуема, если существует такая нижне-треугольная матрица B и такая верхне-треугольная матрица C, что $A=BC$.

Хорошо известно (см., например [1], стр. 35—39), что, если в матрице A все угловые миноры отличны от нуля, то она T-факторизуема. Наша цель заключается в установлении необходимых и достаточных условий T-факторизуемости.

Обозначим через $A_1^{(m)}$ прямоугольную матрицу, состоящую из первых m строк матрицы A, и через $A_2^{(m)}$ — прямоугольную матрицу, состоящую из первых m столбцов той же матрицы A. Квадратную матрицу m-го порядка, стоящую на пересечении $A_1^{(m)}$ и $A_2^{(m)}$, обозначим через $A_0^{(m)}$. Систему первых m строк и первых m столбцов матрицы A назовем ее окаймлением m-го порядка.

Теорема. *Для T-факторизуемости матрицы A необходимо и достаточно, чтобы выполнялись неравенства*

$$r(A_1^{(m)})+r(A_2^{(m)})-r(A_0^{(m)})\leqslant m \quad (1\leqslant m\leqslant n), \tag{1}$$

где $r(\cdot)$ — ранг матрицы.

Для доказательства теоремы мы приведем матрицу A к некоторому специальному виду. Рассмотрим элементарные преобразования матриц, состоящие в том, что из строк (столбцов) с бо́льшими номерами вычитаются предыдущие строки (столбцы), которые предварительно умножены на то или иное число. Этими преобразованиями, совместно с элементарными преобразованиями, состоящими в умножении строк (столбцов) на числа, можно произвольную матрицу A привести к такому виду, что все элементы будут равны 0 или 1 и в каждой строке (столбце) будет не более одной единицы. Будем называть такую матрицу ладейной, и будем обозначать через $\widetilde{A}$ какую-нибудь ладейную матрицу, эквивалентную A в смысле указанных преобразований.

Отметим, что ранг ладейной матрицы равен числу содержащихся в ней единиц.

Очевидно, существует такая невырожденная нижне-треугольная матрица P и такая невырожденная верхне-треугольная матрица Q, что $PAQ=\widetilde{A}$ и, следовательно, $A=P^{-1}\widetilde{A}Q^{-1}$. Так как P^{-1} — нижне-треугольная, а Q^{-1} — верхне-треугольная матрица, то для T-факторизуемости матрицы A необходимо и достаточно, чтобы была T-факторизуемой матрица $\widetilde{A}$.

Заметим, что неравенства (1) эквивалентны аналогичным неравенствам для матрицы $\widetilde{A}$, так как рассматриваемые элементарные преобразования не меняют рангов и, преобразуя матрицу A в матрицу $\widetilde{A}$, мы не переставляли строк и столбцов.

Но в силу ладейности матрицы $\widetilde{A}$ неравенства (1) означают, что количество единиц в окаймлении m-го порядка матрицы $\widetilde{A}$ будет не больше m.

Перейдем теперь непосредственно к доказательству теоремы.

Необходимость. Зафиксируем какое-нибудь значение m $(1 \leqslant m \leqslant n)$. Пусть $\widetilde{A}=BC$, где B — нижне-треугольная, а C — верхне-треугольная матрицы. Заметим, что в силу строения матриц B и C в образовании элементов окаймления m-го порядка матрицы $\widetilde{A}$ участвуют только первые m элементов каждой строки матрицы B и только первые m элементов каждого столбца матрицы C. В дальнейшем будем называть отрезки из первых m элементов строк или столбцов усеченными строками и столбцами. Назовем усеченную строку e_i матрицы B образующей, если в i-й строке окаймления m-го порядка матрицы $\widetilde{A}$ стоит 1. Это означает, что найдется такой усеченный столбец v^i матрицы C, что $(e_i, v^i)=1$, и при $i>m$ индекс столбца v^i не больше m. Докажем, что образующие усеченные строки матрицы B линейно независимы.

Пусть

$$\sum a_{i_k} e_{i_k}=0, \tag{2}$$

где суммирование распространяется на образующие усеченные строки, $i_1>i_2>\ldots$, $a_{i_k} \neq 0$. Обозначим через v такой усеченный столбец матрицы C, что: 1) $(e_{i_1}, v)=1$, 2) индекс столбца v не больше m, если $i_1>m$. Умножая равенство (2) на v скалярно, обнаружим, что существует такое $s>1$, что $(e_{i_s}, v)=1$. Таким образом, в столбце матрицы $\widetilde{A}$, отвечающем столбцу v, оказалось две единицы, что невозможно.

Поскольку усеченные строки m-мерны, то среди них линейно независимых строк не больше m. Тем самым образующих усеченных строк в матрице B не больше m. Следовательно, в окаймлении m-го порядка матрицы $\widetilde{A}$ не больше, чем m единиц.

Достаточность. Пусть a_{ii} — произвольный диагональный элемент матрицы $\widetilde{A}$. Назовем элементы a_{ik}, a_{ki} $(k>i)$ подчиненными элементу a_{ii}. Каждому элементу a_{ii} подчинено не более двух элементов, равных единице. Рассмотрим в матрице $\widetilde{A}$ окаймление m-го порядка и обозначим через $\nu_s^{(m)}$ число диагональных элементов в этом окаймлении, которым подчинено s элементов, равных 1. Неравенства (1) означают, что

$$0 \cdot \nu_0^{(m)}+1 \cdot \nu_1^{(m)}+2 \cdot \nu_2^{(m)} \leqslant m,$$

т. е., что

$$\nu_2^{(m)} \leqslant \nu_0^{(m)} \tag{3}$$

в силу очевидного соотношения $\nu_0^{(m)}+\nu_1^{(m)}+\nu_2^{(m)}=m$.

Обозначим первую строку матрицы $\widetilde{A}$, за исключением элемента a_{11}, — через $\widetilde{A}^{(1)}$, первый столбец матрицы $\widetilde{A}$, за исключением элемента a_{11}, — через $\widetilde{A}^{(2)}$, подматрицу матрицы $\widetilde{A}$, получающуюся вычеркиванием первой строки и первого столбца, — через $\widetilde{A}^{(3)}$. Аналогично определим $B^{(1)}$, $B^{(2)}$, $B^{(3)}$ и $C^{(1)}$, $C^{(2)}$, $C^{(3)}$, исходя из матриц B и C соответственно. Тогда равенство $\widetilde{A}=BC$ можно записать в виде серии равенств

$$a_{11}=b_{11} c_{11}, \quad \widetilde{A}^{(1)}=b_{11} C^{(1)}, \quad \widetilde{A}^{(2)}=c_{11} B^{(2)} \tag{4}$$

и равенства

$$\widetilde{A}^{(3)}-B^{(2)} C^{(1)}=B^{(3)} C^{(3)}. \tag{5}$$

В силу (3) $\nu_2^{(1)} \leqslant \nu_0^{(1)}$, а это обеспечивает разрешимость уравнений (4) относительно b_{11}, c_{11}, $B^{(2)}$, $C^{(1)}$.

Действительно, пусть $a_{11}=1$. Тогда из (4) видно, что матрицы $B^{(2)}$ и $C^{(1)}$ — нулевые. При этом для матрицы $R=\widetilde{A}^{(3)}-B^{(2)} C^{(1)}$, очевидно, выполняются неравенства (3). Пусть $a_{11}=0$, и только один элемент, подчиненный a_{11}, равен единице. Пусть, для определенности, этот элемент стоит в первой строке. Тогда из (4) видно, что матрицу $B^{(2)}$ можно выбрать произвольно. Положим $B^{(2)}=0$. Тогда для R также выполняются неравенства (3). Пусть $a_{11}=0$, а все элементы, подчиненные a_{11}, равны нулю. Тогда из (4) видно, что $B^{(2)}$ и $C^{(1)}$ можно выбрать произвольно. Рассмотрим ближайший к a_{11} диагональный элемент a_{ii} с двумя подчиненными единицами (если он есть; если его нет, то полагаем $B^{(2)}=0$; $C^{(1)}=0$). Пусть $a_{ik}=1$ и $a_{ki}=1$. Выберем тогда матрицу $B^{(2)}$ такой, чтобы у нее были равны 0 все элементы, кроме i-го элемента, а i-й элемент был равен 1; k-й элемент матрицы $C^{(1)}$ положим равным 1, а все остальные — нулю. Тогда у матрицы $B^{(2)}C^{(1)}$ будут равны 0 все элементы, кроме (i, k)-го, который будет равен 1. При этом у матрицы R элементу a_{ii} будет теперь подчинена только одна единица, так что для матрицы R неравенства (3) по-прежнему выполняются. Действуя по индукции, мы придем к матрице первого порядка, которая T-факторизуема тривиальным образом. Следовательно, исходная матрица T-факторизуема. Теорема доказана.

Упомянутое вначале достаточное условие T-факторизуемости непосредственно вытекает из этой теоремы, так как при его выполнении будет $r(A_1^{(m)})=r(A_2^{(m)})=r(A_3^{(m)})=m$. Но теперь легко обнаружить, что оно является и необходимым, если матрица A неособенная, так как в этом случае $r(A_1^{(m)})=r(A_2^{(m)})=m$, и, следовательно, в силу (1), $r(A_3^{(m)})\geqslant m$, т. е. $r(A_3^{(m)})=m$.

Отметим еще, что если ранг матрицы A равен ϱ и в ней отличны от нуля первые ϱ угловых миноров, то она T-факторизуема, так как в этом случае

$$r(A_1^{(m)})+r(A_2^{(m)})-r(A_3^{(m)})=\min(m, \varrho).$$

Этот факт известен ([1], стр. 37—38). Мы теперь можем его обратить в следующем смысле: если матрица A T-факторизуема и в ней отличен от нуля угловой минор ϱ-го порядка, то и все угловые миноры низших порядков отличны от нуля. Действительно, при указанных предположениях

$$r(A_1^{(m)})=r(A_2^{(m)})=m \quad (m<\varrho).$$

Автор выражает благодарность Ю. В. Гандилю за постановку задачи.

Поступило в редакцию 18 февраля 1965 г.

ЦИТИРОВАННАЯ ЛИТЕРАТУРА

[1] Ф. Р. Гантмахер, Теория матриц, М., Гостехиздат, 1953.