Записки научных
семинаров ПОМИ
Том 367, 2009 г.

Х. Д. Икрамов, Х. Фассбендер

КВАДРАТИЧНО-НОРМАЛЬНЫЕ И КОНГРУЭНТНО-НОРМАЛЬНЫЕ МАТРИЦЫ

1. Введение

Матрица $A \in \mathbf{C}^{n\times n}$ называется нормальной, если

$$AA^* = A^*A,$$

и сопряженно-нормальной, если

$$AA^* = \overline{A^*A}.$$

Эти два матричных класса со столь похожими определениями действительно имеют много общего, как показывают списки их свойств, составленные соответственно в [1, 2] и [3]. Оба этих класса играют одинаково важные роли – один в теории унитарных подобий, а другой в теории унитарных конгруэнций.

Для настоящей статьи важен следующий аспект их сходства: всякая нормальная матрица A может быть приведена к диагональному виду посредством унитарного подобия; при этом диагональными элементами являются собственные значения этой матрицы. Всякая сопряженно-нормальная матрица A посредством подходящей унитарной конгруэнции может быть приведена к блочно-диагональной форме с диагональными блоками размеров 1×1 и 2×2 (см. [4, 5]). Из этой формы легко могут быть найдены так называемые псевдособственные значения матрицы A.

В этой статье нашей целью является проведение сравнительного анализа двух других матричных классов. Одним из этих классов являются унитарно квазидиагонализуемые матрицы, т.е. матрицы, которые посредством унитарных подобий могут быть приведены к блочно-диагональной форме с диагональными блоками размеров 1×1

Ключевые слова : квадратично-нормальные матрицы, сопряженно-нормальные матрицы, конгруэнтно-нормальные матрицы, унитарные подобия, унитарные конгруэнции, сингулярные числа.

Работа первого автора над этой статьей была поддержана грантом Немецкого научно-исследовательского общества (DFG).

и 2×2. Множество унитарно квазидиагонализуемых матриц порядка n будет обозначаться символом $\mathcal{Q}_n$. Всякая нормальная матрица унитарно квазидиагонализуема, а все остальные матрицы из $\mathcal{Q}_n$ можно описать как анормальные матрицы, структура которых наиболее близка к структуре нормальных матриц. Впрочем, в других отношениях матрицы из $\mathcal{Q}_n$ могут быть произвольно далеки от нормальности. Например, классу $\mathcal{Q}_n$ принадлежат (косые и ортогональные) проекторы и инволюции. Более общо, $\mathcal{Q}_n$ содержит все матрицы с квадратичными минимальными многочленами.

Второй клас – это класс конгруэнтно-нормальных матриц, введенных в [6].

Определение. *Матрица $A \in \mathbf{C}^{n\times n}$ называется конгруэнтно-нормальной, если матрица $B = A\overline{A}$ нормальна в обычном смысле.*

Таким образом, всякая конгруэнтно-нормальная матрица есть решение матричного уравнения

$$X\overline{X} = B \tag{1.1}$$

для некоторой нормальной матрицы B. Множество конгруэнтно-нормальных $n\times n$-матриц будем обозначать символом $\mathcal{C}_n$. Легко показать, что $\mathcal{C}_n$ содержит все сопряженно-нормальные $n\times n$-матрицы. С другой стороны, матрицы

$$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \quad \text{и} \quad \begin{bmatrix} 0 & 1 \\ 2 & 0 \end{bmatrix}$$

принадлежат классу $\mathcal{C}_2$, но не являются сопряженно-нормальными.

В разделе 3 мы покажем, что всякая матрица $A \in \mathcal{C}_n$ посредством подходящей унитарной конгруэнции может быть приведена к блочно-диагональной форме с диагональными блоками размеров 1×1 и 2×2. Отличие от случая сопряженно-нормальных матриц состоит в том, что блоки размера 2×2 не обязаны быть сопряженно-нормальными.

Следует отметить, что каноническая форма конгруэнтно-нормальных матриц относительно унитарных конгруэнций была найдена еще в [6]. Термин “каноническая форма”, употребленный в названии [6], не следует воспринимать слишком строго. В действительности, авторы [6] не показывают (и не утверждают), что нетривиальные элементы блочно-диагональной формы однозначно определяются классом унитарной конгруэнтности исходной матрицы. Точнее было бы говорить о “компактной” форме. Однако мы, вслед за [6], будем говорить о канонической форме в этом нестрогом смысле. Нашим вкладом в данной статье является не сама эта форма, а, скорее, подход к ее выводу, отличающийся от подхода, принятого в [6].

Мы рассматриваем уравнение (1.1) как конгруэнтный аналог уравнения

$$X^2 = B, \tag{1.2}$$

где B снова является нормальной матрицей.

Уравнение (1.2) описывает квадратные корни из нормальных матриц, или, в другой терминологии, квадратично-нормальные матрицы (см. [7]). Этот класс матриц был исследован в [8], где получен следующий результат.

Теорема 1. *Если $A \in \mathbf{C}^{n\times n}$ является квадратным корнем из нормальной матрицы, то A унитарно подобна матрице вида*

$$N = \begin{bmatrix} N_1 & N_2 \\ 0 & -N_1 \end{bmatrix} \oplus N_3, \tag{1.3}$$

где N_1, N_2 и N_3 – нормальные матрицы, причем N_1 и N_2 перестановочны.

Рассмотрим унитарную матрицу

$$Q = Q_1 \oplus Q_1 \oplus Q_3, \tag{1.4}$$

где Q_3 приводит N_3 к диагональному виду, а Q_1 приводит обе матрицы N_1 и N_2 к диагональной форме. (Одновременное приведение возможно, поскольку N_1 и N_2 перестановочны.) Унитарное подобие

$$N \to M = Q^* N Q$$

дает матрицу

$$M = \begin{bmatrix} M_1 & M_2 \\ 0 & -M_1 \end{bmatrix} \oplus M_3, \tag{1.5}$$

в которой блоки M_1, M_2 и M_3 суть диагональные матрицы. Простым симметричным переупорядочением строк и столбцов матрицы (1.5) первое прямое слагаемое можно превратить в прямую сумму 2×2-блоков вида

$$\begin{bmatrix} m_1 & m_2 \\ 0 & -m_1 \end{bmatrix}$$

и некоторой диагональной матрицы. Мы приходим к следующему выводу.

Теорема 2. *Всякая квадратично-нормальная матрица унитарно квазидиагонализуема.*

В разделе 2 мы приводим собственное доказательство теоремы 2. Наши рассуждения в разделе 3 построены по образцу именно этого доказательства. В заключительном разделе 4 даны некоторые комментарии относительно полученной блочно-диагональной формы.

2. Квадратично-нормальные матрицы

Доказательству теоремы 2 мы предпошлем два вспомогательных предложения.

Напомним, что нильпотентной называется матрица $A \in \mathbf{C}^{n\times n}$ такая, что $A^m = 0$ для некоторого натурального числа m. Наименьшее из таких m называется индексом нильпотентности.

Лемма 1. *Пусть $A \in \mathbf{C}^{n\times n}$ – нильпотентная матрица индекса два. Тогда A унитарно квазидиагонализуема.*

Доказательство. Пусть $q_1, \dots, q_k$ – произвольный ортонормированный базис ядра $\mathcal{N}$ матрицы A, а Q – произвольная унитарная $n \times n$-матрица, в которой $q_1, \dots, q_k$ являются первыми k столбцами. Тогда

$$\triangle = Q^* A Q = \begin{bmatrix} 0 & \triangle_{12} \\ 0 & 0 \end{bmatrix}. \tag{2.1}$$

Нижний правый блок нулевой по той причине, что $A^2 = 0$ и, следовательно,

$$Aq_i \in \mathcal{N} = \operatorname{span}\{q_1, \dots, q_k\}, \quad i = k+1, \dots, n.$$

Пусть

$$\triangle_{12} = U\Sigma V^*, \quad \Sigma = \operatorname{diag}(\sigma_1, \dots, \sigma_\ell), \quad \ell = \min\{k, n-k\}, \tag{2.2}$$

есть сингулярное разложение блока $\triangle_{12}$. Положим

$$P = U \oplus V. \tag{2.3}$$

Тогда

$$\Gamma = P^* \triangle P = \begin{bmatrix} 0 & \Sigma \\ 0 & 0 \end{bmatrix}. \tag{2.4}$$

Наконец, симметричным переупорядочением строк и столбцов матрицы Γ получаем блочно-диагональную матрицу с 2×2-блоками вида

$$\begin{bmatrix} 0 & \sigma_j \\ 0 & 0 \end{bmatrix}, \tag{2.5}$$

соответствующими ненулевым сингулярным числам блока $\triangle_{12}$. Итак, A может быть преобразована к желаемой блочно-диагональной форме. Это означает, что $A \in \mathcal{Q}_n$.

Следствие 1. *Пусть $A \in \mathbf{C}^{n\times n}$ – матрица вида*

$$A = \alpha I_n + N,$$

где N – нильпотентная матрица индекса два. Тогда $A \in \mathcal{Q}_n$.

Лемма 2. *Пусть $A \in \mathbf{C}^{n\times n}$ – диагонализуемая матрица, имеющая ровно два различных собственных значения. Тогда A унитарно квазидиагонализуема.*

Доказательство. Обозначим собственные значения матрицы A через λ_1 и λ_2; через k обозначим кратность числа λ_1. Пусть Q – унитарная матрица, трансформирующая A посредством подобия к форме Шура, в которой λ_1 занимает первые k диагональных позиций. Тогда

$$\triangle = Q^* A Q = \begin{bmatrix} \lambda_1 I_k & \triangle_{12} \\ 0 & \lambda_2 I_{n-k} \end{bmatrix}.$$

Как и в лемме 1, рассмотрим сингулярное разложение (2.2) блока $\triangle_{12}$ и введем унитарную матрицу (2.3). Тогда

$$\Gamma = P^* \triangle P = \begin{bmatrix} \lambda_1 I_k & \Sigma \\ 0 & \lambda_2 I_{n-k} \end{bmatrix}. \tag{2.6}$$

Симметричное переупорядочение строк и столбцов превращает Γ в блочно-диагональную матрицу с 2×2-блоками вида

$$\begin{bmatrix} \lambda_1 & \sigma_j \\ 0 & \lambda_2 \end{bmatrix}, \tag{2.7}$$

отвечающими ненулевым сингулярным числам блока $\triangle_{12}$. Полученная блочно-диагональная форма доказывает, что $A \in \mathcal{Q}_n$.

Леммы 1 и 2 и следствие 1 приводят к такому выводу.

Следствие 2. *Всякая матрица* $A \in \mathbf{C}^{n\times n}$ *с квадратичным минимальным многочленом унитарно квазидиагонализуема.*

Рассмотрим уравнение

$$f(X) = B, \tag{2.8}$$

несколько более общее, чем уравнение (1.2). Здесь B, как и прежде, – нормальная матрица, а f – некоторый квадратичный многочлен. Без потери общности, f можно считать унитальным многочленом:

$$f(z) = z^2 + az + b.$$

Теорема 3. *Всякое решение уравнения* (2.8) *является унитарно квазидиагонализуемой матрицей.*

Доказательство. Поскольку B есть многочлен от неизвестной матрицы X, обе матрицы должны коммутировать:

$$BX = XB. \tag{2.9}$$

Обозначим через $\lambda_1, \dots, \lambda_m$ все различные собственные значения матрицы B, а через $k_1, \dots, k_m$ ($k_1+\dots+k_m = n$) – их кратности. Найдем унитарную матрицу Q такую, что

$$C = Q^*BQ = \lambda_1 I_{k_1} \oplus \lambda_2 I_{k_2} \oplus \dots \oplus \lambda_m I_{k_m}. \tag{2.10}$$

Положим

$$Y = Q^*XQ.$$

Тогда

$$CY = YC \tag{2.11}$$

и

$$f(Y) = C. \tag{2.12}$$

Из соотношений (2.10) и (2.11) вытекает, что

$$Y = Y_1 \oplus Y_2 \oplus \dots \oplus Y_m. \tag{2.13}$$

Диагональный блок Y_i $(i = 1, \dots, m)$ имеет размер $k_i \times k_i$ и удовлетворяет уравнению

$$f(Y_i) = \lambda_i I_{k_i}.$$

Теперь мы опустим индексы и рассмотрим две возможные ситуации.

Случай 1: $a^2 = 4(b - \lambda)$. В этом случае Y есть решение уравнения

$$g(Y) = 0, \tag{2.14}$$

в котором многочлен

$$g(z) = f(z) - \lambda \tag{2.15}$$

имеет нулевой дискриминант и, следовательно, единственный корень:

$$g(z) = (z - \alpha)^2.$$

Отсюда вытекает, что Y – это матрица типа, описанного в следствии 1. Поэтому Y унитарно квазидиагонализуема.

Случай 2: $a^2 \neq 4(b - \lambda)$. Теперь многочлен (2.15) имеет два различных корня. Согласно (2.14), $g(z)$ является минимальным многочленом матрицы Y. В соответствии со следствием 2, Y унитарно квазидиагонализуема.

Мы приходим к выводу, что все блоки Y_i $(i = 1, \dots, m)$ в (2.13) унитарно квазидиагонализуемы. Следовательно, $Y \in \mathcal{Q}_n$, а потому $X \in \mathcal{Q}_n$.

Остается заметить, что теорема 2 соответствует специальному выбору многочлена f, а именно

$$f(z) = z^2.$$

Замечание. Априори не ясно, разрешимо ли уравнение (2.8) для заданных матрицы B и многочлена f. Приведенное выше доказательство заодно устанавливает разрешимость этого уравнения для всякой нормальной матрицы B и любого квадратичного многочлена f.

3. Конгруэнтно-нормальные матрицы

В этом разделе мы докажем следующее утверждение.

Теорема 4. *Всякая конгруэнтно-нормальная матрица A посредством подходящей унитарной конгруэнции может быть приведена к блочно-диагональной форме с диагональными блоками размеров 1×1 и 2×2.*

Как уже было отмечено во введении, каноническая форма конгруэнтно-нормальных матриц по отношению к унитарным конгруэнциям была впервые найдена в [6]. Наше доказательство теоремы 4 отличается от доказательства, приведенного в [6], тем, что оно пытается подчеркнуть и использовать сходство уравнений (1.1) и (1.2). Тем не менее, мы заимствуем несколько остроумных идей, содержащихся в [6].

Мы начнем с нескольких подготовительных предложений.

Лемма 3. *Пусть $A, B \in \mathbf{C}^{n\times n}$, причем B – нормальная матрица. Тогда из равенства*

$$A\overline{B} = BA \tag{3.1}$$

следует, что

$$A\overline{B^*} = B^*A,$$

т.е.

$$AB^T = B^*A. \tag{3.2}$$

Доказательство. Хорошо известно, что матрица B^*, сопряженная к нормальной матрице B, может быть представлена как многочлен от B:

$$B^* = f(B). \tag{3.3}$$

Менее известно, что для соотношения (3.3) может быть выбран многочлен f с вещественными коэффициентами (см. [9]). Пусть наш многочлен f именно таков. Тогда

$$B^*A = f(B)A = Af(\overline{B}) = A\overline{f(B)} = A\overline{B^*} = AB^T.$$

Замечание. Наша лемма 3 – это лемма 1 в [6], хотя доказательства различны.

Следствие 3. *Если $A \in \mathcal{C}_n$, то*

$$AA^*A^T = A^TA^*A. \tag{3.4}$$

Доказательство. Напомним, что включение $A \in \mathcal{C}_n$ означает, что матрица $A\overline{A}$ нормальна. Равенство (3.1) тривиально выполняется для $B = A\overline{A}$. Поскольку $B^T = A^*A^T$, из соотношения (3.2) вытекает (3.4).

Замечание. Как отмечено в [6], равенство (3.4) в действительности характеризует конгруэнтно-нормальные матрицы. В самом деле, предполагая, что (3.4) выполнено, положим $B = A\overline{A}$. Тогда

$$\begin{aligned} BB^* &= A\overline{A}A^TA^* = A\overline{(AA^*A^T)} = A\overline{(A^TA^*A)} = AA^*A^T\overline{A} \\ &= (AA^*A^T)\overline{A} = (A^TA^*A)\overline{A} = (A^TA^*)(A\overline{A}) = B^*B. \end{aligned}$$

Тем самым B нормальна и $A \in \mathcal{C}_n$.

Для заданной матрицы A определим матрицы

$$C = A^T\overline{A} \tag{3.5}$$

и

$$D = AA^*. \tag{3.6}$$

Лемма 4. *Если $A \in \mathcal{C}_n$, то*

$$CD = DC = BB^*. \tag{3.7}$$

Доказательство. Цепочка равенств, приведенная в предыдущем замечании, содержит соотношение

$$BB^* = AA^*A^T\overline{A}. \tag{3.8}$$

Его правая часть есть матрица DC. Беря сопряженное к равенству (3.8), получаем

$$BB^* = A^T\overline{A}AA^*,$$

т.е. матрицу CD.

Замечание. Для сопряженно-нормальной матрицы A (см. определение в разделе 1) матрицы C и D совпадают. Для произвольной матрицы $A \in \mathcal{C}_n$, согласно лемме 4, матрицы C и D могут различаться, однако должны коммутировать.

Мы будем также пользоваться соотношением

$$A\overline{C} = DA, \tag{3.9}$$

немедленно вытекающим из (3.5) и (3.6).

Прежде чем начать доказательство теоремы 4, напомним особенности спектра матриц вида $B = A\overline{A}$ (см. [10, с. 252–253]).

1. Спектр матрицы B симметричен относительно вещественной оси. Кроме того, собственные значения λ и $\overline{\lambda}$ имеют одинаковую кратность.
2. Отрицательные собственные значения матрицы B (если таковые имеются) обязательно имеют четную кратность.

Теперь мы приступим к доказательству теоремы 4.

Доказательство. Пусть $A \in \mathcal{C}_n$ и $B = A\overline{A}$. Тождество

$$A\overline{A}A = A\overline{A}A$$

можно переписать в виде

$$BA = A\overline{B}. \tag{3.10}$$

Это равенство выражает некоторое свойство коммутирования A и B, сохраняемое унитарными конгруэнциями матрицы A. Выберем конгруэнцию

$$A \to Y = Q^T A Q \tag{3.11}$$

так, чтобы превратить матрицу

$$\Lambda = Q^T A\overline{AQ} = (Q^T AQ)(Q^* \overline{AQ}) = Y\overline{Y} \tag{3.12}$$

в диагональную, притом имеющую следующие свойства:

(a) если λ – собственное значение матрицы B кратности $m > 1$, то m копий числа λ занимают последовательные позиции на главной диагонали матрицы Λ;

(b) если λ – комплексное собственное значение матрицы B, то на главной диагонали матрицы Λ копии числа $\overline{\lambda}$ (или копии λ) следуют сразу за копиями λ (или $\overline{\lambda}$).

Соотношение (3.10) переходит в равенство

$$\Lambda Y = Y\overline{\Lambda}, \tag{3.13}$$

означающее, что Y есть блочно-диагональная матрица:

$$Y = Y_1 \oplus Y_2 \oplus \cdots \oplus Y_k. \tag{3.14}$$

Каждый блок Y_i $(1 \leq i \leq k)$ соответствует или вещественному собственному значению матрицы B или паре ее комплексно сопряженных собственных значений. Чтобы доказать теорему 4, нужно показать, что каждая матрица Y_i посредством унитарной конгруэнции может быть приведена к блочно-диагональной форме с диагональными блоками размеров 1×1 и 2×2.

Будем различать четыре возможные случая.

1. Y_i соответствует собственному значению $\lambda = 0$.
2. Y_i соответствует положительному собственному значению матрицы B.
3. Y_i соответствует отрицательному собственному значению матрицы B.
4. Y_i соответствует паре комплексных собственных значений λ, $\overline{\lambda}$.

Мы дадим раздельный анализ этих четырех случаев. Чтобы упростить записи, опустим индекс i. Порядок матрицы Y будет обозначаться через t или (если он четен) через $2t$.

3.1. Случай 1: $\lambda = 0$.

В этом случае матрице Y соответствует нулевая матрица Λ. Согласно (3.12), имеем

$$Y\overline{Y} = 0. \tag{3.15}$$

Чтобы привести Y к блочно-диагональному виду, воспользуемся построением, аналогичным тому, что было применено в лемме 1. Пусть $p_1, \dots, p_k$ – произвольный ортонормированный базис ядра $\mathcal{N}$ матрицы Y (это означает, что векторы $\overline{p_1}, \dots, \overline{p_k}$ составляют базис ядра матрицы $\overline{Y}$). Пусть P – унитарная $t \times t$-матрица, в которой $p_1, \dots, p_k$ являются первыми k столбцами. Тогда

$$\triangle = P^T Y P = \begin{bmatrix} 0 & \triangle_{12} \\ 0 & 0 \end{bmatrix}.$$

Нижний правый блок нулевой вследствие (3.15). В самом деле, из (3.15) вытекает, что

$$Yp_i \in \overline{\mathcal{N}} = \operatorname{span}\{\overline{p_1}, \dots, \overline{p_k}\}, \quad i = k+1, \dots, t.$$

Пусть

$$\triangle_{12} = U\Sigma V^*, \quad \Sigma = \operatorname{diag}(\sigma_1, \dots, \sigma_\ell), \quad \ell = \min\{k, t-k\},$$

есть сингулярное разложение блока $\triangle_{12}$. Положим

$$R = \overline{U} \oplus V.$$

Тогда

$$\Gamma = R^T \triangle R = \begin{bmatrix} 0 & \Sigma \\ 0 & 0 \end{bmatrix}.$$

Теперь желаемая блочно-диагональная форма получается симметричным переупорядочением строк и столбцов матрицы Γ. Блоки размера 2×2 имеют вид (2.5) и соответствуют ненулевым сингулярным числам подматрицы $\triangle_{12}$.

3.2. Случай 2: $\lambda > 0$.

В этом случае основное уравнение имеет вид

$$Y\overline{Y} = \lambda I.$$

Мы можем избавиться от λ, вводя новую матрицу

$$\widetilde{Y} = \frac{1}{\sqrt{\lambda}} Y.$$

Чтобы не усложнять обозначения, опустим тильду и будем в дальнейшем работать с уравнением

$$Y\overline{Y} = I. \tag{3.16}$$

Уравнение (3.16) дает некоторую информацию о сингулярных числах матрицы Y. Действительно, $\overline{Y}$ имеет те же сингулярные числа, что и Y, а сингулярные числа матрицы Y^{-1} обратны сингулярным числам Y. Поскольку $\overline{Y} = Y^{-1}$, мы заключаем, что если $\gamma \neq 1$ является сингулярным числом для Y, то γ^{-1} также есть сингулярное число, причем оба числа имеют одну и ту же кратность.

Матрицы (3.5) и (3.6) преобразуются в

$$S = Y^T \overline{Y} = \overline{Y^* Y}$$

и

$$T = YY^*,$$

тогда как соотношение (3.7) переходит в

$$ST = I. \tag{3.17}$$

Выполним теперь унитарную конгруэнцию

$$Y \to Z = P^T Y P,$$

выбирая P так, чтобы сделать

$$\Gamma = P^T S \overline{P} = P^T Y^T \overline{Y P} = (P^T Y^T P)(P^* \overline{Y P}) = Z^T \overline{Z}$$

диагональной матрицей. Мы дополнительно предположим, что если γ – кратное сингулярное число матрицы Y, то все копии числа γ^2 занимают последовательные позиции на главной диагонали матрицы Γ и, если $\gamma > 1$, то копии числа γ^{-2} следуют сразу за копиями γ^2.

Вследствие (3.17), соотношение (3.9) принимает вид

$$Z\Gamma = \Gamma^{-1} Z.$$

Другими словами,

$$\gamma_j^2 z_{ij} = \frac{1}{\gamma_i^2} z_{ij} \quad \text{для всех} \quad i, j. \tag{3.18}$$

Равенства (3.18) приводят нас к следующим выводам.

1. Матрица Z яявляется блочно-диагональной:

$$Z = Z_1 \oplus Z_2 \oplus \cdots \oplus Z_\ell. \tag{3.19}$$

2. Если $\gamma = 1$ является сингулярным числом матрицы Y, то прямая сумма (3.19) содержит блок, отвечающий этому сингулярному числу. Порядок такого блока равен кратности числа $\gamma = 1$.
3. Каждый из остальных блоков в (3.19) соответствует некоторой паре γ_j, γ_j^{-1} взаимно обратных сингулярных чисел матрицы Y. Если s – кратность числа γ_j, то ассоциированный с ним блок Z_j имеет размер $2s \times 2s$. Будучи разбит на блоки размера $s \times s$, Z_j принимает вид

$$Z_j = \begin{bmatrix} 0 & G_j \\ H_j & 0 \end{bmatrix}. \tag{3.20}$$

Чтобы завершить анализ случая 2, остается показать, что каждый блок Z_j может быть приведен к желаемой форме. Мы сделаем это раздельно для $\gamma = 1$ и для пары γ, γ^{-1}, в которой $\gamma > 1$. Индекс j будет опущен (точно так же, как ранее был опущен индекс i).

3.2.1. Случай $\gamma = 1$.

Соответствующая матрица Z имеет единственное сингулярное число 1 и, следовательно, является унитарной матрицей. С другой стороны, из соотношения (3.16) вытекает, что

$$Z\overline{Z} = I. \tag{3.21}$$

Из $Z^{-1} = Z^*$ и $Z^{-1} = \overline{Z}$ мы делаем вывод, что матрица Z симметрична. Согласно теореме Такаги (см. [10, следствие 4.4.4]), симметричную матрицу Z посредством унитарной конгруэнции можно привести к диагональной форме, диагональными элементами которой являются сингулярные числа Z. В обсуждаемом случае такой диагональной формой оказывается единичная матрица.

3.2.2. Случай $\gamma > 1$.

Наша матрица Z имеет вид (3.20) (где индекс j следует опустить) и снова подчиняется соотношению (3.21). Отсюда следует, что

$$G\overline{H} = I. \tag{3.22}$$

Из равенства

$$S = \overline{Y^*Y}$$

вытекает, что

$$Z^*Z = \begin{bmatrix} \gamma^2 I & 0 \\ 0 & \gamma^{-2} I \end{bmatrix},$$

т.е.

$$H^*H = \gamma^2 I$$

и

$$G^*G = \frac{1}{\gamma^2} I. \tag{3.23}$$

Итак, H имеет единственное сингулярное число γ, тогда как G имеет единственное сингулярное число γ^{-1}. Сравнивая (3.22) и (3.23), заключаем, что

$$G^{-1} = \overline{H} = \gamma^2 G^*,$$

откуда

$$H = \gamma^2 G^T. \tag{3.24}$$

Пусть

$$G = \gamma^{-1} U V^*$$

есть сингулярное разложение матрицы G. Согласно (3.24), имеем

$$H = \gamma \overline{V} U^T.$$

Положим

$$R = \overline{U} \oplus V$$

и применим к Z конгруэнцию

$$Z \to W = R^T Z R.$$

В результате получается матрица

$$W = \begin{bmatrix} 0 & \gamma^{-1} I \\ \gamma I & 0 \end{bmatrix}. \tag{3.25}$$

Наконец, симметричное переупорядочение строк и столбцов матрицы (3.25) превращает ее в прямую сумму s блоков размера 2×2, имеющих форму

$$\begin{bmatrix} 0 & \gamma^{-1} \\ \gamma & 0 \end{bmatrix}. \tag{3.26}$$

Проведенный выше анализ можно суммировать в виде следующего предложения.

Теорема 5. *Всякое решение Y матричного уравнения* (3.16) *посредством подходящей унитарной конгруэнции можно привести к прямой сумме единичной матрицы (возможно, отсутствующей) и 2×2-блоков вида* (3.26). *Единичная матрица присутствует в такой сумме, если $\gamma = 1$ является сингулярным числом матрицы Y. Ее порядок равен кратности этого сингулярного числа. Каждый блок вида* (3.26) *соответствует паре γ, γ^{-1} взаимно обратных сингулярных чисел матрицы Y; при этом $\gamma \neq 1$.*

Замечание. Матрица $E \in \mathbf{C}^{n \times n}$ называется псевдоинволюцией, если $E\overline{E} = I$ (см. [11, с. 477]). Наша теорема 5 описывает каноническую форму псевдоинволюций относительно унитарных конгруэнций. С точностью до несущественных отличий, это теорема 1.5 из [12].

3.3. Случай 3: $\lambda < 0$.

Напомним, что отрицательное собственное значение матрицы B необходимо имеет четную кратность. С точностью до этого различия, анализ для $\lambda < 0$ очень схож с анализом, проведенным в предыдущем случае. Поэтому мы опустим некоторые детали, сосредоточившись, главным образом, на отличиях от случая $\lambda > 0$.

Вводя новую матрицу, можно свести основное уравнение рассматриваемого случая к виду

$$Y\overline{Y} = -I. \tag{3.27}$$

Подобно (3.16), это уравнение показывает, что если $\gamma \neq 1$ является сингулярным числом для Y, то γ^{-1} также есть сингулярное число, причем оба числа имеют одну и ту же кратность. Отсюда следует, что если $\gamma = 1$ является сингулярным числом, то оно имеет четную кратность.

Матрицы S и T определяются так же, как и выше, и выполнено соотношение (3.17). Ничто не меняется в описании унитарной конгруэнции

$$Y \to Z = P^T Y P,$$

которая снова приводит к равенствам (3.18), означающим, что Z имеет блочно-диагональную форму (3.19).

Предположим, что $\gamma = 1$ является сингулярным числом матрицы Y. Тогда один из блоков в правой части равенства (3.19) соответствует этому сингулярному числу. Опуская индекс, будем считать, что таким блоком является Z. Итак, Z имеет единственное сингулярное число 1 и, следовательно, является унитарной матрицей. С другой стороны, из соотношения (3.27) вытекает, что

$$Z\overline{Z} = -I. \tag{3.28}$$

Из $Z^{-1} = Z^*$ и $Z^{-1} = -\overline{Z}$ мы делаем вывод, что матрица Z кососимметрична. Всякая невырожденная кососимметричная матрица посредством унитарной конгруэнции может быть приведена к прямой сумме 2×2-блоков вида

$$\begin{bmatrix} 0 & \varsigma_j \\ -\varsigma_j & 0 \end{bmatrix}, \quad \varsigma_j \neq 0$$

(см. [10, раздел 4.4, задача 26]). Числа ς_j могут быть выбраны положительными и, в этом случае, они являются сингулярными числами исходной матрицы. Указанное приведение, будучи применено к нашей матрице Z, дает прямую сумму s копий блока

$$\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}. \tag{3.29}$$

Каждый из остальных блоков в (3.19) соответствует некоторой паре γ_j, γ_j^{-1} взаимно обратных сингулярных чисел матрицы Y и должен иметь форму (3.20). Снова опуская индекс j, из (3.28) выводим соотношение

$$G\overline{H} = -I, \tag{3.30}$$

отличающееся от (3.22). С другой стороны, равенство (3.23) остается верным. Вместе соотношения (3.23) и (3.30) означают, что

$$H = -\gamma^2 G^T \tag{3.31}$$

(ср. с (3.24)). Снова определяя унитарную матрицу R с помощью сингулярного разложения матрицы G и выполняя затем преобразование

$$Z \to W = R^T Z R,$$

получаем матрицу

$$W = \begin{bmatrix} 0 & \gamma^{-1} I \\ -\gamma I & 0 \end{bmatrix},$$

которая после симметричного переупорядочения строк и столбцов превращается в прямую сумму s копий блока

$$\begin{bmatrix} 0 & \gamma^{-1} \\ -\gamma & 0 \end{bmatrix}. \tag{3.32}$$

Приведенное обсуждение можно суммировать в виде следующего утверждения.

Теорема 6. *Всякое решение Y матричного уравнения* (3.27) *посредством подходящей унитарной конгруэнции можно привести к прямой сумме 2×2-блоков вида* (3.32). *Параметры γ_j суть сингулярные числа матрицы Y. Если $\gamma \neq 1$ есть сингулярное число кратности s, то с таким γ в прямой сумме ассоциированы в точности s блоков. Остальные блоки (если они имеются) соответствуют $\gamma = 1$ и имеют форму* (3.29).

Замечание. Матрица $E \in \mathbf{C}^{n \times n}$ называется косой псевдоинволюцией, если $E\overline{E} = -I$. Теорема 6 описывает каноническую форму косых псевдоинволюций относительно унитарных конгруэнций. С точностью до несущественных отличий, это теорема 12 из [13].

3.4. Случай 4: комплексная собственная пара $\lambda, \overline{\lambda}$.

Для определенности, будем считать, что число λ принадлежит верхней полуплоскости, т.е. $\operatorname{Im}\lambda > 0$. Вспомним, что λ и $\overline{\lambda}$ суть собственные значения исходной матрицы $B = A\overline{A}$. Предположим, что оба они имеют кратность t. Тогда Y есть матрица порядка $2t$, удовлетворяющая соотношению (3.13), в котором

$$\Lambda = \lambda I_t \oplus \overline{\lambda} I_t.$$

Разбивая матрицу Y на блоки размера $t \times t$, заключаем из (3.13), что Y – матрица вида

$$Y = \begin{bmatrix} 0 & K \\ L & 0 \end{bmatrix}. \tag{3.33}$$

Ввиду (3.12), имеем

$$K\overline{L} = \lambda I_t.$$

Отсюда следует, что

$$L = \overline{\lambda K}^{-1}. \tag{3.34}$$

Пусть

$$K = U\Sigma V^*, \quad \Sigma = \operatorname{diag}(\sigma_1, \dots, \sigma_t),$$

есть сингулярное разложение матрицы K. Тогда из (3.34) следует, что

$$L = \overline{\lambda V}\Sigma^{-1}U^T.$$

Положим

$$R = \overline{U} \oplus V$$

и выполним конгруэнцию

$$Y \to Z = R^T Y R,$$

приводящую к матрице

$$Z = \begin{bmatrix} 0 & \Sigma \\ \overline{\lambda}\Sigma^{-1} & 0 \end{bmatrix}.$$

Симметричным переупорядочением строк и столбцов матрица Z может быть преобразована в прямую сумму 2×2-блоков

$$\begin{bmatrix} 0 & \sigma_j \\ \overline{\lambda}/\sigma_j & 0 \end{bmatrix}, \quad j = 1, 2, \dots, t. \tag{3.35}$$

Этим завершается анализ случая 4 и тем самым все доказательство теоремы 4.

4. Заключительные замечания

Насколько канонична “каноническая” форма конгруэнтно-нормальных матриц, полученная нами в теореме 4?

Числа σ_j в разделе 3.1 суть сингулярные числа соответствующих блоков Y_i, т.е., в конечном счете, сингулярные числа исходной матрицы A. Нормирующий множитель $1/\sqrt{\lambda}$ в разделе 3.2 является квадратным корнем из положительного собственного значения матрицы $B = A\overline{A}$ (иначе называемым вещественным псевдособственным значением матрицы A). Точно так же, нормирующий множитель $1/\sqrt{|\lambda|}$ в разделе 3.3 есть квадратный корень из модуля отрицательного собственного значения для B. Нетривиальные элементы блочно-диагональных форм, полученных в этих разделах, суть сингулярные числа нормированных матриц и, следовательно, легко могут быть выражены через сингулярные числа матрицы A и собственные значения матрицы $A\overline{A}$. Таким образом, блочно-диагональные матрицы, построенные нами для первых трех случаев в теореме 4, действительно можно признать каноническими формами.

К сожалению, этого нельзя сказать о блочно-диагональной матрице в разделе 3.4: она уже не определяется только сингулярными числами матрицы A и собственными значениями матрицы B. Вот соответствующий пример (за который мы благодарим профессора Р. Хорна): матрицы

$$A_1 = \begin{bmatrix} 0 & 1 \\ 2i & 0 \end{bmatrix} \quad \text{и} \quad A_2 = \begin{bmatrix} 0 & 2 \\ i & 0 \end{bmatrix}$$

имеют одни и те же сингулярные числа 1 и 2; кроме того, $A_1\overline{A}_1 = A_2\overline{A}_2$. Однако A_1 и A_2 не могут приводиться к одинаковой блочно-диагональной форме, поскольку не являются унитарно конгруэнтными.

Последнее утверждение нуждается в разъяснении. Предположим поэтому, что A_1 и A_2 унитарно конгруэнтны, т.е. найдется унитарная матрица V такая, что

$$A_2 = V^T A_1 V. \tag{4.1}$$

Переписывая (4.1) в виде

$$A_1 V = \overline{V} A_2$$

и учитывая унитарность матрицы V, без труда покажем, что $v_{11} = v_{22} = 0$. Далее, следствием (4.1) является равенство

$$A_2\overline{A}_2 = V^T A_1 \overline{A}_1 \overline{V}.$$

Левая часть этого равенства – диагональная матрица

$$\mathrm{diag}\,(-2i, 2i),$$

тогда как для нашей матрицы V правая часть имеет вид

$$\mathrm{diag}\,(2i, -2i).$$

Это противоречие показывает, что A_1 и A_2 не могут быть унитарно конгруэнтны.

Попробуем разобраться, почему блочно-диагональная форма из раздела 3.4 не определяется только сингулярными числами матрицы A (помимо числа λ). Она, во всяком случае, определяется сингулярными числами блока K в (3.33), которые составляют половину сингулярных чисел матрицы Y. Проблема в том, что мы не можем заранее сказать, какая именно половина будет взята для нашей матрицы A. Так, матрицы A_1 и A_2 уже имеют форму (3.33); при этом блоком K для A_1 является число 1, а для A_2 – число 2.

Заметим, что блочно-диагональная форма конгруэнтно-нормальных матриц, найденная в теореме 4, с точностью до несущественных отличий, совпадает с канонической формой из теоремы 3 работы [6]. Однако нигде в [6] термины "сингулярное число" или "сингулярное разложение" не употребляются.

Как отмечено во введении, всякая сопряженно-нормальная матрица принадлежит классу $\mathcal{C}_n$; следовательно, теорема 4 применима к такой матрице. Однако, не все 2×2-блоки, перечисленные в разделе 3, могут присутствовать в блочно-диагональной форме сопряженно-нормальной матрицы A. Как очевидно из определений, сопряженно-нормальные матрицы и обычные нормальные матрицы имеют следующее общее свойство: для всякого i $(1 \leq i \leq n)$ 2-норма строки i равна 2-норме столбца i. Это свойство, будучи применено к нашей блочно-диагональной форме, исключает блоки типов (2.5) и (3.26), а также блоки типа (3.32) с $\gamma \neq 1$. Более того, для блоков типа (3.35) должно выполняться

$$\sigma_j = |\lambda|/\sigma_j,$$

т.е.

$$\sigma_j = |\lambda|^{\frac{1}{2}}.$$

Если учесть эти ограничения, то получится каноническая форма сопряженно-нормальных матриц, описанная в [4].

Литература

1. R. Grone, C. R. Johnson, E. M. Sa, H. Wolkowicz, *Normal matrices.* — Linear Algebra Appl. **87** (1987), 213–225.
2. L. Elsner, Kh. D. Ikramov, *Normal matrices: an update.* — Linear Algebra Appl. **285** (1998), 291–303.
3. H. Faßbender, Kh. D. Ikramov, *Conjugate-normal matrices: a survey.* — Linear Algebra Appl. **429** (2008), 1425–1441.
4. M. Vujičić, F. Herbut, G. Vujičić, *Canonical form for matrices under unitary congruence transformations. I. Conjugate-normal matrices.* — SIAM J. Appl. Math. **23** (1972), 225–238.
5. H. Faßbender, Kh. D. Ikramov, *Some observations on the Youla form and conjugate-normal matrices.* — Linear Algebra Appl. **422** (2007), 29–38.
6. F. Herbut, P. Loncke, M. Vujičić, *Canonical form for matrices under unitary congruence transformations. II. Congruence-normal matrices.* — SIAM J. Appl. Math. **26** (1974), 794–805.
7. F. Kittaneh, *On the structure of polynomially normal operators.* — Bull. Austral. Math. Soc. **30** (1984), 11–18.
8. H. Radjavi, P. Rosenthal, *On roots of normal operators.* – J. Math. Anal. Appl. **34** (1971), 653–664.
9. T.J. Laffey, *A normality criterion for an algebra of matrices.* — Linear Algebra Appl. **25** (1979), 169–174.
10. Р. Хорн, Ч. Джонсон, *Матричный анализ.* Мир, М., 1989.
11. R. A. Horn, C. R. Johnson, *Topics in Matrix Analysis.* Cambridge University Press, Cambridge, 1991.
12. R. A. Horn, D. I. Merino, *A real-coninvolutory analog of the polar decomposition.* — Linear Algebra Appl. **190** (1993), 209–227.
13. M. N. M. Abara, D. I. Merino, A. T. Paras. *Skew-coninvolutory matrices.* — Linear Algebra Appl. **426** (2007), 540–557.

Ikramov Kh. D., Fassbender H. Quadratically normal and congruence-normal matrices.

A matrix $A \in \mathbf{C}^{n\times n}$ is unitarily quasi-diagonalizable if A can be brought by a unitary similarity transformation to a block diagonal form with 1×1 and 2×2 diagonal blocks. In particular, the square roots of normal matrices, the so-called quadratically normal matrices, are unitarily quasi-diagonalizable.

A matrix $A \in \mathbf{C}^{n\times n}$ is congruence-normal if $B = A\overline{A}$ is a conventional normal matrix. We show that every congruence-normal matrix A can be brought by a unitary congruence transformation to a block diagonal form with 1×1 and 2×2 diagonal blocks. Our proof emphasizes and exploits the likeliness between the equations $X^2 = B$ and $X\overline{X} = B$ for a normal matrix B.

Московский государственный
университет
ГСП-1, Ленинские горы,
19991 Москва, Россия

E-mail: ikramov@cs.msu.su

Поступило 6 октября 2008 г.

Institute of Computational Mathematics,
TU Braunschweig,
Pockelstrasse 14, D-38023 Braunschweig, Germany

E-mail: h.fassbender@tu-bs.de