Записки научных
семинаров ПОМИ
Том 395, 2011 г.

Х. Д. Икрамов

ПРОИЗВЕДЕНИЯ ОРТОПРОЕКТОРОВ И ТЕОРЕМА КРИММИНСА

§1. Введение

Пусть $M_n(\mathbf{C})$ – множество комплексных $n \times n$-матриц. В недавней публикации [1] цитируется следующий любопытный результат:

Теорема 1. *Матрица $A \in M_n(\mathbf{C})$ тогда и только тогда представима произведением*

$$A = PQ, \tag{1}$$

где P и Q – ортопроекторы, когда A удовлетворяет уравнению

$$A^2 = AA^*A. \tag{2}$$

Эта теорема приписывается в [1] Т. Кримминсу, но вместо ссылки на источник дается отсылка к статье [2]. Однако и в [2] упоминание об исходной публикации Кримминса отсутствует. Что касается доказательства теоремы 1, то в [2] говорится следующее:

"Необходимость тривиальна. Для доказательства достаточности определим P как ортопроектор на образ матрицы A, а Q – как ортопроектор на ортогональное дополнение к ядру $\ker A$. Тогда выполняется (1)".

Необходимость условия теоремы 1 действительно очевидна и проверяется подстановкой представления (1) в уравнение (2). Однако отнюдь не очевидна достаточность, и можно предположить, что разложение (1) матрицы A было в доказательстве Кримминса итогом некоторого не вполне тривиального (и, возможно, довольно протяженного) рассуждения. Целью данной статьи является обоснование этого разложения для конечномерного случая (теорема 1 верна и в бесконечномерном гильбертовом пространстве).

В §2 показано, что ненулевые собственные значения матрицы A, удовлетворяющей уравнению (2), вещественны и принадлежат полуинтервалу $(0, 1]$. В §3 исследуются корневые векторы этой матрицы.

Ключевые слова: ортопроектор, инвариантное подпространство, унитарно квазидиагонализуемые матрицы.

Выясняется, что ее собственные значения полупросты, т.е. корневые векторы высоты, большей единицы, отсутствуют. При этом, если A имеет собственное значение $\lambda = 1$, то соответствующее собственное подпространство является общим для A и A^*. Анализ, проведенный в §4, показывает, что матрица A унитарно квазидиагонализуема. Это означает, что $\mathbf{C}^n$ может быть разложено в ортогональную сумму одно- и двумерных подпространств, одновременно A-инвариантных и A^*-инвариантных. Исследование сужений A и A^* на общее двумерное инвариантное подпространство позволяет нам получить в §5 произведение ортопроекторов, указанных Кримминсом.

§2. Собственные значения

Пусть λ – ненулевое собственное значение матрицы A, удовлетворяющей уравнению (2), а x – собственный вектор, отвечающий λ:

$$Ax = \lambda x. \tag{3}$$

Собственные векторы всегда предполагаются нормированными:

$$||x|| = 1.$$

Символ $||\cdot||$ в данной статье обозначает 2-норму соответствующего вектора.

Из (2) и (3) выводим

$$A^2x = \lambda^2 x = AA^*Ax = \lambda AA^*x,$$

откуда

$$AA^*x = \lambda x, \tag{4}$$

т.е. x является собственным вектором и для AA^*, причем ему соответствует то же собственное значение λ. Поскольку матрица AA^* эрмитова и положительно полуопределена, отсюда выводим следующие заключения:

1. Все ненулевые собственные значения матрицы A положительны.

2. Собственные векторы, относящиеся к различным положительным собственным значениям, ортогональны.

Согласно (3) и (4), имеем

$$AA^*x = Ax, \tag{5}$$

или

$$A(A^*x - x) = 0. \tag{6}$$

Таким образом,

$$A^*x = x + y, \quad y \in \ker A. \tag{7}$$

Рассмотрим соотношение

$$A^{*2} = A^*AA^*, \tag{8}$$

получаемое взятием сопряжения от равенства (2). Используя (8) и (7), находим

$$A^{*2}x = A^*A(A^*x) = A^*A(x+y) = A^*Ax = \lambda A^*x,$$

откуда

$$A^*(A^*x - \lambda x) = 0, \tag{9}$$

или

$$A^*x = \lambda x + z, \quad z \in \ker A^*. \tag{10}$$

Поскольку $x \in \operatorname{im} A$, слагаемые суммы (10) ортогональны и

$$||A^*x||^2 = \lambda^2 + ||z||^2. \tag{11}$$

С другой стороны, из (4) выводим

$$(AA^*x, x) = ||A^*x||^2 = \lambda. \tag{12}$$

Подставляя в (11), имеем

$$\lambda^2 + ||z||^2 = \lambda \tag{13}$$

и, следовательно,

$$\lambda^2 \leqslant \lambda,$$

т.е.

$$\lambda \leqslant 1.$$

§3. Корневые и собственные векторы

Пусть λ – положительное собственное значение матрицы A, удовлетворяющей уравнению (2). Предположим, что имеется корневой вектор z высоты 2, относящийся к λ, т.е.

$$(A - \lambda I)z = x \neq 0 \tag{14}$$

и

$$(A - \lambda I)x = 0. \tag{15}$$

Тогда

$$Az = \lambda z + x,$$

$$A^2z = \lambda Az + \lambda x = \lambda(\lambda z + x) + \lambda x = \lambda^2 z + 2\lambda x,$$

$$AA^*Az = AA^*(\lambda z + x) = \lambda AA^*z + AA^*x = \lambda AA^*z + \lambda x.$$

В последнем переходе использовано равенство (4). Теперь из уравнения (2) следует, что

$$\lambda^2 z + 2\lambda x = \lambda AA^*z + \lambda x,$$

или

$$(AA^* - \lambda I)z = x. \tag{16}$$

Если положить

$$B = AA^* - \lambda I,$$

то, согласно (16) и (4),

$$Bz = x, \qquad Bx = 0,$$

т.е.

$$B^2 z = 0. \tag{17}$$

Однако для эрмитовой матрицы B соотношение (17) влечет за собой равенство $Bz = 0$, что противоречит исходному предположению (14). Это противоречие устанавливает еще одно свойство матрицы A, которое можно добавить к свойствам 1 и 2, указанным в предыдущем разделе:

3. Положительные собственные значения матрицы A полупросты, т.е. их геометрическая кратность (максимальное число линейно независимых ассоциированных собственных векторов) совпадает с алгебраической кратностью.

Пусть $\lambda = 1$ является собственным значением матрицы A, а x – соответствующий собственный вектор. Тогда из (13) следует, что $z = 0$, т.е. (см. (10))

$$A^*x = x.$$

Таким образом, собственное подпространство для $\lambda = 1$ общее для матриц A и A^*.

Напротив, если $\lambda < 1$, равенство (13) невозможно при $z = 0$. Это означает, что никакой собственный вектор матрицы A, относящийся к числу λ $(0 < \lambda < 1)$, не может быть одновременно собственным вектором для A^*.

Свойство полупростоты было выше доказано для положительных собственных значений матрицы A. В действительности оно выполняется и для $\lambda = 0$. В самом деле, предположим, что $A \neq I_n$ и

$$z \in \operatorname{im} A \cap \ker A,$$

т.е. найдется вектор y такой, что $z = Ay$. В то же время $Az = 0$. Тогда из уравнения (2) выводим, что

$$A^2 y = 0 = AA^*(Ay),$$

т.е. вектор $z = Ay$ принадлежит ядру матрицы AA^*, которое совпадает с $\ker A^*$. Так как подпространство $\ker A^*$ является ортогональным дополнением подпространства $\operatorname{im} A$, то одновременно имеем

$$z \in \operatorname{im} A \quad \text{и} \quad z \perp \operatorname{im} A,$$

откуда $z = 0$. Это и означает полупростоту собственного значения $\lambda = 0$. Заодно установлено еще одно полезное свойство матрицы A.

4. Область значений (образ) матрицы A есть линейная оболочка ее собственных векторов, относящихся к положительным собственным значениям.

§4. Унитарная квазидиагонализуемость

Обозначим через $\mathcal{L}_1$ собственное подпространство матрицы A для $\lambda = 1$. Как показано выше, $\mathcal{L}_1$ является собственным подпространством и для A^*.

Пусть $x_1, \ldots, x_k$ – максимальная линейно независимая система собственных векторов матрицы A, относящихся к собственным значениям $\lambda_1, \ldots, \lambda_k$, где $0 < \lambda_j < 1$ $(j = 1, 2, \ldots, k)$. Без ограничения общности эту систему можно считать ортонормированной. Действительно, векторы x_r и x_s, относящиеся к различным собственным значениям λ_r и λ_s, ортогональны вследствие свойства 2. Что касается векторов, соответствующих одному и тому же (кратному) собственному значению, то их всегда можно выбрать ортонормированными.

В соответствии с (7), каждый вектор $A^* x_j$ можно представить в виде

$$A^* x_j = x_j + y_j, \quad y_j \in \ker A, \quad 1 \leqslant j \leqslant k. \tag{18}$$

Покажем, что двумерное подпространство

$$\mathcal{M}_j = \operatorname{span}\{x_j, y_j\} \tag{19}$$

является общим инвариантным подпространством матриц A и A^*. Инвариантность $\mathcal{M}_j$ относительно A очевидна из определения векторов x_j и y_j:

$$Ax_j = \lambda_j x_j, \quad Ay_j = 0$$

(см. (18)). Вектор A^*x_j принадлежит $\mathcal{M}_j$ согласно равенству (18). Остается проверить вектор A^*y_j. Имеем

$$\begin{aligned} A^*y_j &= A^*(A^*x_j - x_j) = A^{*2}x_j - A^*x_j \\ &= A^*AA^*x_j - A^*x_j = (\lambda_j - 1)A^*x_j. \end{aligned} \tag{20}$$

В этих выкладках использованы равенства (8) и (4). Поскольку $A^*x_j \in \mathcal{M}_j$, то и $A^*y_j \in \mathcal{M}_j$, что доказывает A^*-инвариантность подпространства $\mathcal{M}_j$.

Теперь мы покажем, что подпространства $\mathcal{M}_j$ с различными индексами j попарно ортогональны и все такие подпространства ортогональны собственному подпространству $\mathcal{L}_1$. То, что вектор x_j ортогонален $\mathcal{L}_1$, следует из свойства 2. Пусть теперь x – произвольный вектор из $\mathcal{L}_1$, т.е. $Ax = x$. Используя (20), находим

$$\begin{aligned} (y_j, x) &= (y_j, Ax) = (A^*y_j, x) = (\lambda_j - 1)(A^*x_j, x) \\ &= (\lambda_j - 1)(x_j, Ax) = (\lambda_j - 1)(x_j, x) = 0. \end{aligned}$$

Итак,

$$y_j \perp \mathcal{L}_1,$$

откуда

$$\mathcal{M}_j \perp \mathcal{L}_1, \qquad j = 1, 2, \ldots, k.$$

Рассмотрим наряду с $\mathcal{M}_j$ подпространство $\mathcal{M}_l$, где $l \neq j$. Векторы x_j и x_l ортогональны в соответствии с выбором системы $x_1, \ldots, x_k$. Далее, имеем

$$(x_j, y_l) = (x_j, A^*x_l - x_l) = (x_j, A^*x_l) = (Ax_j, x_l) = \lambda_j(x_j, x_l) = 0.$$

Аналогично показываем, что $(x_l, y_j) = 0$. Наконец,

$$\begin{aligned} (y_j, y_l) &= (A^*x_j - x_j, y_l) = (A^*x_j, y_l) - (x_j, y_l) \\ &= (A^*x_j, y_l) = (A^*x_j, A^*x_l - x_l) = (A^*x_j, A^*x_l) - (A^*x_j, x_l) \\ &= (AA^*x_j, x_l) - (x_j, Ax_l) = \lambda_j(x_j, x_l) - \lambda_l(x_j, x_l) = 0. \end{aligned}$$

Итак, подпространства $\mathcal{M}_j$ и $\mathcal{M}_l$ ортогональны.

Пусть l – размерность подпространства $\mathcal{L}_1$. Подпространство

$$\mathcal{K} = \mathcal{L}_1 \oplus \mathcal{M}_1 \oplus \cdots \oplus \mathcal{M}_k \tag{21}$$

имеет размерность $l + 2k$ и является общим инвариантным подпространством матриц A и A^*. Как следует из его построения, $\mathcal{K}$ содержит все собственные векторы для A, относящиеся к положительным

собственным значениям. Тем самым

$$\operatorname{im} A \subset \mathcal{K}. \tag{22}$$

Однако $\mathcal{K}$ обладает тем же свойством и по отношению к A^*. Действительно, собственные векторы A^* для $\lambda = 1$ находятся в $\mathcal{L}_1$, и каждое подпространство $\mathcal{M}_j$ содержит собственный вектор для $\lambda_j > 0$. Итак,

$$\operatorname{im} A^* \subset \mathcal{K}. \tag{23}$$

Если

$$l + 2k = n,$$

то

$$\mathbf{C}^n = \mathcal{K} = \mathcal{L}_1 \oplus \mathcal{M}_1 \oplus \cdots \oplus \mathcal{M}_k. \tag{24}$$

Заменяя здесь $\mathcal{L}_1$ прямой суммой одномерных оболочек, натянутых на векторы $u_1, \ldots, u_l$ какого-либо ортонормированного базиса этого подпространства, получим разложение $\mathbf{C}^n$ в ортогональную сумму одно- и двумерных инвариантных подпространств, общих для A и A^*. Это и означает, что матрица A унитарно квазидиагонализуема.

Пусть теперь

$$m = n - (l + 2k) > 0. \tag{25}$$

Поскольку

$$\operatorname{rank} A = l + k, \quad \operatorname{def} A = n - (l + k),$$

то (25) эквивалентно неравенству

$$\operatorname{def} A = \dim \ker A > k.$$

Подпространство $\mathcal{K}^\perp$ имеет размерность m и, будучи ортогональным дополнением общего инвариантного подпространства матриц A и A^*, само инвариантно относительно обеих этих матриц. Из включений (22) и (23) следует, что

$$\mathcal{K}^\perp \subset \operatorname{im} A^\perp = \ker A^*$$

и

$$\mathcal{K}^\perp \subset \ker A,$$

т.е.

$$\mathcal{K}^\perp \subset \mathcal{N} = \ker A \cap \ker A^*. \tag{26}$$

Покажем, что включение в (26) является в действительности равенством. Пусть $z \in \mathcal{N}$. Тогда

$$\begin{aligned}(y_j, z) &= (A^* x_j - x_j, z) = (A^* x_j, z) - (x_j, z) \\ &= (x_j, Az) - (x_j, z) = -(x_j, z) = 0, \quad j = 1, 2, \ldots, k.\end{aligned}$$

Скалярное произведение (x_j, z) равно нулю, потому что $x_j \in \operatorname{im} A$, а $z \in \ker A^*$. Итак, подпространство $\mathcal{N}$ ортогонально каждому из векторов y_j и, следовательно, натянутому на них подпространству

$$\operatorname{span}\{y_1, \dots, y_k\}.$$

Поэтому размерность $\dim \mathcal{N}$ не превосходит числа

$$\operatorname{def} A - k = n - (l + k) - k = m.$$

С другой стороны, согласно (26),

$$\dim \mathcal{N} \geqslant \dim \mathcal{K}^{\perp} = m.$$

Отсюда следует, что

$$\mathcal{N} = \mathcal{K}^{\perp}.$$

Вместо (24) мы имеем теперь разложение

$$\mathbf{C}^n = \mathcal{L}_1 \oplus \mathcal{M}_1 \oplus \dots \oplus \mathcal{M}_k \oplus \mathcal{N}.$$

Прежним образом заменяем $\mathcal{L}_1$ ортогональной суммой одномерных подпространств, натянутых на собственные векторы для $\lambda = 1$, и то же самое (с заменой $\lambda = 1$ на $\lambda = 0$) проделываем для подпространства $\mathcal{N}$. Тем самым матрица A оказывается унитарно квазидиагонализуемой и в случае (25).

§5. Ортопроекторы Кримминса

Базис x_j, y_j подпространства $\mathcal{M}_j$ не является ортогональным, а вектор y_j даже не нормирован. Действительно,

$$\begin{aligned} (x_j, y_j) &= (x_j, A^* x_j - x_j) = (A x_j, x_j) - (x_j, x_j) \\ &= (\lambda_j - 1)(x_j, x_j) = \lambda_j - 1, \end{aligned} \tag{27}$$

$$\begin{aligned} (y_j, y_j) &= (A^* x_j - x_j, y_j) = (A^* x_j, y_j) - (x_j, y_j) \\ &= (A^* x_j, A^* x_j - x_j) - (\lambda_j - 1) \\ &= (A^* x_j, A^* x_j) - (A^* x_j, x_j) - (\lambda_j - 1) \\ &= (A A^* x_j, x_j) - (x_j, A x_j) - (\lambda_j - 1) \\ &= \lambda_j - \lambda_j - \lambda_j + 1 = 1 - \lambda_j. \end{aligned} \tag{28}$$

Заменим базис x_j, y_j на ортонормированный базис x_j, z_j. Найдем требуемый вектор z_j как вектор из $\mathcal{M}_j$, ортогональный к x_j. Поскольку x_j – собственный вектор A для $\lambda_j > 0$, то z_j должен принадлежать

ядру матрицы A^*. Записывая z_j в виде

$$z_j = \alpha x_j + \beta y_j \tag{29}$$

и полагая $v_j = A^* x_j$, имеем

$$\begin{aligned} A^* z_j &= \alpha v_j + \beta A^* y_j = \alpha v_j + \beta A^*(A^* x_j - x_j) \\ &= \alpha v_j + \beta A^{*2} x_j - \beta v_j = (\alpha - \beta) v_j + \beta A^* A A^* x_j \\ &= (\alpha - \beta) v_j + \beta A^* A (x_j + y_j) = (\alpha - \beta) v_j + \beta A^* A x_j \\ &= (\alpha - \beta) v_j + \lambda_j \beta A^* x_j = (\alpha - \beta + \lambda_j \beta) v_j. \end{aligned}$$

В этой цепочке равенств использованы соотношения (7) и (8).

Вектор $v_j = A^* x_j$ не может быть нулевым; в противном случае x_j есть общий собственный вектор для A и A^*, но он отвечает различным собственным значениям $\lambda_j > 0$ и 0. Поэтому из условия $A^* z_j = 0$ получаем соотношение между коэффициентами α и β в представлении (29):

$$\beta = \frac{\alpha}{1 - \lambda_j}. \tag{30}$$

Вектор (29) пока еще не нормирован. Будем считать α положительным числом. Используя равенства (27), (28) и (30), находим

$$\begin{aligned} (z_j, z_j) &= (\alpha x_j + \beta y_j, \alpha x_j + \beta y_j) = \alpha^2 + 2\alpha\beta (x_j, y_j) + \beta^2 (y_j, y_j) \\ &= \alpha^2 \left(1 - 2 + \frac{1}{1 - \lambda_j}\right) = \frac{\lambda_j}{1 - \lambda_j} \alpha^2. \end{aligned}$$

Итак, чтобы получить в $\mathcal{M}_j$ ортонормированный базис, вектор (29) следует заменить на

$$z_j = \frac{\sqrt{1 - \lambda_j}}{\sqrt{\lambda_j}} x_j + \frac{1}{\sqrt{\lambda_j}\sqrt{1 - \lambda_j}} y_j. \tag{31}$$

Если положить

$$\lambda_j = \cos^2 \gamma_j, \qquad 0 < \gamma_j < \frac{\pi}{2},$$

то вектор (31) можно записать как

$$z_j = \operatorname{tg} \gamma_j x_j + \frac{1}{\sin \gamma_j \cos \gamma_j} y_j. \tag{32}$$

Рассматривая A как оператор, действующий в $\mathcal{M}_j$, найдем матрицу этого оператора в построенном базисе x_j, z_j. Имеем

$$A x_j = \lambda_j x_j = \cos^2 \gamma_j x_j,$$

$$Az_j = \operatorname{tg}\gamma_j Ax_j + \frac{2}{\sin 2\gamma_j} Ay_j = \lambda_j \operatorname{tg}\gamma_j x_j = \sin\gamma_j \cos\gamma_j x_j.$$

Таким образом, искомая матрица A_j имеет вид

$$A_j = \begin{pmatrix} \cos^2\gamma_j & \sin\gamma_j\cos\gamma_j \\ 0 & 0 \end{pmatrix}.$$

Эта матрица может быть представлена произведением двух ортопроекторов

$$P_2 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

и

$$Q_j = \begin{pmatrix} \cos^2\gamma_j & \sin\gamma_j\cos\gamma_j \\ \cos\gamma_j\sin\gamma_j & \sin^2\gamma_j \end{pmatrix}.$$

Первый из них проектирует $\mathcal{M}_j$ на линейную оболочку вектора x_j, а второй – на линейную оболочку вектора r_j, имеющего в локальном базисе x_j, z_j координаты $\cos\gamma_j$ и $\sin\gamma_j$. Если переписать (32) как

$$y_j = -\sin\gamma_j\{\sin\gamma_j x_j - \cos\gamma_j z_j\},$$

то видно, что векторы r_j и y_j ортогональны. Напомним, что $y_j \in \ker A$.

Определим операторы P и Q, действующие во всем пространстве $\mathbf{C}^n$, формулами

$$P = I_l \oplus P_2 \oplus \cdots \oplus P_2 \oplus 0_m,$$

$$Q = I_l \oplus Q_1 \oplus \cdots \oplus Q_k \oplus 0_m.$$

Эти операторы суть ортопроекторы Кримминса, описанные во введении. Действительно, из проведенного в данной статье анализа следует, что P проектирует $\mathbf{C}^n$ на $\operatorname{im} A$, а Q есть ортопроектор на ортогональное дополнение к $\ker A$. (Оператор Q оставляет подпространство $\mathcal{L}_1$ на месте, а каждое подпространство $\mathcal{M}_j$ проектирует на линейную оболочку соответствующего вектора r_j; при этом r_j ортогонален подпространству $\ker A$.)

Итак,

$$A = PQ,$$

и теорема Кримминса доказана.

Замечание. Пусть H_1 и H_2 – эрмитовы положительно полуопределенные матрицы. Тогда собственные значения произведения

$$H = H_1 H_2 \tag{33}$$

вещественны и неотрицательны. Это утверждение получается предельным переходом из хорошо известного факта: если H_1 и H_2 положительно *определены*, то матрица (33) диагонализуема и имеет положительный спектр.

Если в (33) матрица H_2 по-прежнему положительно полуопределена, а H_1 – произвольная эрмитова матрица, то спектр матрицы H остается вещественным. Более точно спектральные свойства такой матрицы H описаны в [3]: ее ненулевые собственные значения полупросты, а собственному значению нуль могут отвечать жордановы клетки порядка не выше 2.

Таким образом, свойства 1 и 3, найденные в разделах 2 и 3, необходимы для того, чтобы решение A матричного уравнения (2) могло быть представлено произведением пары положительно полуопределенных матриц H_1 и H_2. При этом выясняется, что если H_1 и H_2 – ортопроекторы, то полупросты не только положительные собственные значения матрицы A, но и собственное значение нуль.

Литература

1. G. Corach, A. Maestripieri, *Products of orthogonal projections and polar decompositions.* — Linear Algebra Appl. **434** (2011), 1594–1609.
2. H. Radjavi, J. P. Williams, *Products of self-adjoint operators.* — Michigan Math. J., **16** (1969), 177–185.
3. Y. P. Hong, R. A. Horn, *The Jordan canonical form of a product of a Hermitian and a positive semidefinite matrix.* — Linear Algebra Appl. **147** (1991), 373–386.

Ikramov Kh. D. Products of orthoprojectors and a theorem of Crimmins. A proof of the following result, due to T. Crimmins, is proposed: A matrix $A \in M_n(\mathbf{C})$ can be represented as a product of orthoprojectors P and Q if and only if A satisfies the equation $A^2 = AA^*A$.

Московский государственный
университет ГСП-1, Ленинские горы,
119991 Москва, Россия

E-mail: ikramov@cs.msu.su

Поступило 25 мая 2011 г.