Записки научных
семинаров ПОМИ
Том 268, 2000 г.

Л. Ю. Колотилина

ОПТИМАЛЬНО ОБУСЛОВЛЕННЫЕ БЛОЧНЫЕ 2×2 МАТРИЦЫ

§1. Введение

В этой заметке мы рассматриваем эрмитовы положительно определенные матрицы $A \in \mathbb{C}^{n\times n}$, $n \geqslant 2$, представленные в блочной 2×2 форме

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \tag{1.1}$$

где $A_{11} \in \mathbb{C}^{m\times m}$, $1 \leqslant m \leqslant n-1$, и $A_{21} = A_{12}^*$.

Через $\Delta(m, n-m)$ мы обозначаем группу всех невырожденных матриц из $\mathbb{C}^{n\times n}$, имеющих вид

$$D = \begin{bmatrix} D_1 & 0 \\ 0 & D_2 \end{bmatrix},$$

где $D_1 \in \mathbb{C}^{m\times m}$, $D_2 \in \mathbb{C}^{(n-m)\times(n-m)}$.

Будем говорить, что матрица A оптимально обусловлена относительно $\Delta(m, n-m)$, если

$$k(A) = \min_{D\in\Delta(m,n-m)} \{k(D^*AD)\}.$$

Здесь и далее $k(B)$ – это спектральное число обусловленности эрмитовой положительно определенной матрицы $B \in \mathbb{C}^{n\times n}$, т.е.

$$k(B) = \lambda_1(B)/\lambda_n(B),$$

где

$$\lambda_{\max}(B) = \lambda_1(B) \geqslant \lambda_2(B) \geqslant \ldots \geqslant \lambda_n(B) = \lambda_{\min}(B)$$

– это (вещественные и положительные) собственные значения матрицы B, упорядоченные в порядке невозрастания.

Как хорошо известно [1–4], если

$$A_{11} = cI_m, \quad A_{22} = cI_{n-m} \tag{1.2}$$

Работа выполнена при финансовой поддержке NWO (грант 047-008-007).

для некоторой положительной константы c (через I_k или просто I мы, как обычно, обозначаем единичную матрицу порядка k), то матрица A оптимально обусловлена относительно $\Delta(m, n-m)$.

С другой стороны, также известно, что оптимально обусловленные относительно $\Delta(m, n-m)$ матрицы не обязательно удовлетворяют условиям (1.2). Например, матрица

$$A = \begin{bmatrix} \alpha & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 1 & 2 \end{bmatrix},$$

имеющая собственные значения $\alpha, 1$ и 3, оптимально обусловлена относительно $\Delta(2,1)$ при $1 \leqslant \alpha \leqslant 3$ (см. [3]), хотя условия (1.2) для нее, очевидно, не выполняются.

Цель настоящей работы – описать все эрмитовы положительно определенные матрицы из $\mathbb{C}^{n\times n}$, оптимально обусловленные относительно $\Delta(m, n-m)$. Это описание содержится в теореме 2.2, см. §2. Доказательство этой теоремы основывается на новом доказательстве известного неравенства [7, теорема 2]

$$\lambda_{\max}(A_{11}S_1^{-1}(A)) \leqslant \frac{[k(A)+1]^2}{4k(A)}, \tag{1.3}$$

составляющего утверждение теоремы 2.1 данной работы.

На протяжении данной статьи используются следующие обозначения. Через

$$S_1(A) = A_{11} - A_{12}A_{22}^{-1}A_{21}, \quad S_2(A) = A_{22} - A_{21}A_{11}^{-1}A_{12}$$

мы обозначаем дополнения по Шуру в матрице A подматриц A_{22} и A_{11} соответственно, а константа

$$\begin{aligned} \xi(A) &= \lambda_{\max}(A_{11}S_1^{-1}(A)) \\ &= \lambda_{\min}^{-1}(I_m - A_{12}A_{22}^{-1}A_{21}A_{11}^{-1}) = \lambda_{\max}(A_{22}S_2^{-1}(A)) \end{aligned} \tag{1.4}$$

является количественной характеристикой рассматриваемого блочного 2×2 разбиения матрицы A. Наконец, для $B \in \mathbb{C}^{p\times q}$, $p, q \geqslant 1$, через $\|B\|$ мы обозначаем спектральную норму матрицы B, так что $\|B\|^2 = \lambda_{\max}(B^*B)$.

§2. Результаты

Сперва мы покажем, что неравенство (1.3), являющееся небольшой переформулировкой теоремы 2 из [7], в действительности легко следует из классических неравенств Канторовича и Коши–Шварца.

Теорема 2.1. *Пусть эрмитова положительно определенная матрица* $A \in \mathbb{C}^{n\times n}$ *имеет вид* (1.1) *и пусть константа* $\xi(A)$ *определена с помощью* (1.4). *Тогда*

$$\xi(A) \leqslant \frac{[\lambda_1(A)+\lambda_n(A)]^2}{4\lambda_1(A)\lambda_n(A)} = \frac{[k(A)+1]^2}{4k(A)}. \tag{2.1}$$

Доказательство. В силу неравенства Канторовича (см., например, [6, II. 4.3.1]) для любого ненулевого вектора $z \in \mathbb{C}^n$ мы имеем:

$$\frac{z^*Az\, z^*A^{-1}z}{(z^*z)^2} \leqslant \frac{[\lambda_1(A)+\lambda_n(A)]^2}{4\lambda_1(A)\lambda_n(A)} = \frac{[k(A)+1]^2}{4k(A)}. \tag{2.2}$$

Пусть x – собственный вектор матрицы $A_{11}^{1/2}S_1^{-1}(A)A_{11}^{1/2}$, ассоциированный с ее наибольшим собственным значением $\xi(A)$, и пусть $z = \left[\begin{smallmatrix} A_{11}^{1/2}x \\ 0 \end{smallmatrix}\right]$. Тогда с учетом представления

$$A^{-1} = \begin{bmatrix} S_1^{-1}(A) & * \\ * & * \end{bmatrix}$$

из неравенства (2.2) мы получаем:

$$\frac{x^*A_{11}^2x\, x^*x}{(x^*A_{11}x)^2}\, \frac{x^*A_{11}^{1/2}S_1^{-1}(A)A_{11}^{1/2}x}{x^*x} = \xi(A)\frac{x^*A_{11}^2x\, x^*x}{(x^*A_{11}x)^2} \leqslant \frac{[k(A)+1]^2}{4k(A)},$$

откуда и следует неравенство (2.1), если принять во внимание, что, в силу неравенства Коши–Шварца,

$$\frac{x^*A_{11}^2x\, x^*x}{(x^*A_{11}x)^2} \geqslant 1.$$

□

Поскольку, очевидно,

$$\xi^{-1}(A) = 1 - \|A_{11}^{-1/2}A_{12}A_{22}^{-1/2}\|^2, \tag{2.3}$$

неравенство (2.1) эквивалентно неравенству

$$\|A_{11}^{-1/2}A_{12}A_{22}^{-1/2}\| \leqslant \frac{\lambda_1(A)-\lambda_n(A)}{\lambda_1(A)+\lambda_n(A)} = \frac{k(A)-1}{k(A)+1}, \tag{2.4}$$

полученному Деммелем [2, теорема 1], и может также быть переписано в виде

$$k(A) \geqslant \frac{1+\|A_{11}^{-1/2}A_{12}A_{22}^{-1/2}\|}{1-\|A_{11}^{-1/2}A_{12}A_{22}^{-1/2}\|}. \tag{2.5}$$

Следует отметить, что величины $\xi(A)$, а также (например, в силу (2.3)) и $\|A_{11}^{-1/2}A_{12}A_{22}^{-1/2}\|$ обе инвариантны относительно преобразований подобия матрицы A с помощью блочно диагональных матриц из группы $\Delta(m, n-m)$. Действительно, обозначая

$$B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} = D^*AD,$$

где

$$D = \begin{bmatrix} D_1 & 0 \\ 0 & D_2 \end{bmatrix} = \Delta(m, n-m),$$

и используя соотношение

$$(B^{-1})_{11} = S_1^{-1}(B),$$

мы выводим

$$\xi(B) = \lambda_1(B_{11}S_1^{-1}(B)) = \lambda_1(D_1^*A_{11}D_1D_1^{-1}S_1^{-1}(A)D_1^{-*}) \\ = \lambda_1(A_{11}S_1^{-1}(A)) = \xi(A).$$

Принимая во внимание указанную выше инвариантность нормы $\|A_{11}^{-1/2}A_{12}A_{22}^{-1/2}\|$ и неравенство (2.5), мы немедленно приходим к следующему результату, являющемуся следствием теоремы 2.1.

Следствие 2.1. *Пусть эрмитова положительно определенная матрица* $A \in \mathbb{C}^{n\times n}$ *имеет вид* (1.1). *Тогда*

$$\min_{D\in\Delta(m,n-m)} k(D^*AD) \geqslant \frac{1+\|A_{11}^{-1/2}A_{12}A_{22}^{-1/2}\|}{1-\|A_{11}^{-1/2}A_{12}A_{22}^{-1/2}\|}. \tag{2.6}$$

В том случае, когда

$$A = \begin{bmatrix} cI_m & A_{12} \\ A_{12}^* & cI_{n-m} \end{bmatrix}, \quad c > 0, \tag{2.7}$$

мы, очевидно, имеем

$$\|A_{11}^{-1/2}A_{12}A_{22}^{-1/2}\| = c^{-1}\|A_{12}\|,$$

и выполнены соотношения

$$\lambda_1(A) = c + \|A_{12}\|, \quad \lambda_n(A) = c - \|A_{12}\|,$$

откуда следует, что

$$k(A) = \frac{c + \|A_{12}\|}{c - \|A_{12}\|} = \frac{1 + \|A_{11}^{-1/2}A_{12}A_{22}^{-1/2}\|}{1 - \|A_{11}^{-1/2}A_{12}A_{22}^{-1/2}\|}.$$

Полученное равенство в сочетании с неравенством (2.6) показывает, что, как уже отмечалось в §1, матрицы вида (2.7) оптимально обусловлены относительно $\Delta(m, n-m)$.

Если матрица A вида (1.1) блочно диагональна, т.е. $A_{12} = 0$, то, очевидно, она оптимально обусловлена относительно $\Delta(m, n-m)$ тогда и только тогда, когда $k(A) = 1$, т.е. когда $A = cI_n$, $c > 0$, – скалярная матрица.

Описание же всех отличных от скалярных эрмитовых положительно определенных матриц, оптимально обусловленных относительно $\Delta(m, n-m)$, дается следующей теоремой.

Теорема 2.2. *Эрмитова положительно определенная матрица*

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \in \mathbb{C}^{n\times n},$$

где $A_{11} \in \mathbb{C}^{m\times m}$, $A_{12} \neq 0$, *оптимально обусловлена относительно* $\Delta(m, n-m)$, *т.е.*

$$\xi(A) = \frac{[k(A)+1]^2}{4k(A)}, \tag{2.8}$$

в том и только том случае, когда справедливо любое из следующих двух равносильных утверждений:

(i) *существует вектор* $z = \begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{C}^n$, $x \in \mathbb{C}^m$, *такой что*

$$Az = \lambda_1(A)z$$

и

$$A\widetilde{z} = \lambda_n(A)\widetilde{z},$$

где $\widetilde{z} = \begin{bmatrix} x \\ -y \end{bmatrix}$;

(ii) *существует вектор* $z = \begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{C}^n$, $x \in \mathbb{C}^m$, *с ненулевыми блочными компонентами* x *и* y *такой, что*

$$A_z = \frac{1}{2}\begin{bmatrix} \lambda_1(A) + \lambda_n(A) & \lambda_1(A) - \lambda_n(A) \\ \lambda_1(A) - \lambda_n(A) & \lambda_1(A) + \lambda_n(A) \end{bmatrix},$$

где в соответствии с [5] *векторно аггрегированная матрица* A_z *определяется формулой*

$$A_z = \begin{bmatrix} \dfrac{x^*A_{11}x}{\|x\|^2} & \dfrac{x^*A_{12}y}{\|x\|\,\|y\|} \\ \dfrac{y^*A_{21}x}{\|x\|\,\|y\|} & \dfrac{y^*A_{22}y}{\|y\|^2} \end{bmatrix}. \tag{2.9}$$

Доказательство. Допустим сперва, что справедливо равенство (2.8), и в этом предположении покажем, что справедливо утверждение (i).

Пусть $x \in \mathbb{C}^m$ и $y \in \mathbb{C}^{n-m}$, $\|x\| = \|y\| = 1$, – это пара сингулярных векторов матрицы $A_{11}^{-1/2}A_{12}A_{22}^{-1/2}$, ассоциированная с ее наибольшим сингулярным значением

$$\sigma = \|A_{11}^{-1/2}A_{12}A_{22}^{-1/2}\|, \tag{2.10}$$

т.е.

$$x^*A_{11}^{-1/2}A_{12}A_{22}^{-1/2}y = \sigma \tag{2.11}$$

или, равносильно,

$$A_{11}^{-1/2}A_{12}A_{22}^{-1/2}y = \sigma x, \quad x^*A_{11}^{-1/2}A_{12}A_{22}^{-1/2} = \sigma y^*. \tag{2.12}$$

Тогда мы имеем

$$x^*A_{11}^{-1/2}A_{12}A_{22}^{-1}A_{21}A_{11}^{-1/2}x = \sigma^2, \tag{2.13}$$

$$y^*A_{22}^{-1/2}A_{21}A_{11}^{-1}A_{12}A_{22}^{-1/2}y = \sigma^2. \tag{2.14}$$

Используя (2.13), (2.10) и (2.3), мы выводим:

$$\begin{aligned} x^*A_{11}^{-1/2}S_1(A)A_{11}^{-1/2}x &= x^*(I_m - A_{11}^{-1/2}A_{12}A_{22}^{-1}A_{21}A_{11}^{-1/2})x \\ &= 1 - \sigma^2 = 1 - \|A_{11}^{-1/2}A_{12}A_{22}^{-1/2}\|^2 = \xi^{-1}(A), \end{aligned}$$

откуда следует, что x – это собственный вектор матрицы $A_{11}^{1/2}S_1^{-1}(A)A_{11}^{1/2}$, ассоциированный с ее наибольшим собственным значением $\xi(A)$. Заметим, что ввиду (2.3) равенство (2.8) равносильно соотношению

$$\sigma = \frac{\lambda_1(A) - \lambda_n(A)}{\lambda_1(A) + \lambda_n(A)}. \tag{2.15}$$

Далее, из равенства (2.8) и доказательства теоремы 2.1 вытекает, что

$$\frac{x^* A_{11}^2 x \, x^* x}{(x^* A_{11} x)^2} = 1.$$

Применяя неравенство Коши–Шварца, мы можем заключить, что векторы x и $A_{11}x$ должны быть коллинеарны, т.е., что x является собственным вектором первого диагонального блока A_{11}. Аналогичным образом можно убедиться в том, что из (2.8) и (2.14) с необходимостью следует, что y – собственный вектор второго диагонального блока A_{22}. Итак пусть

$$A_{11}x = \mu x, \quad A_{22}y = \nu y. \tag{2.16}$$

Используя (2.16) и (2.12), мы выводим:

$$\sigma \mu^{1/2} x = \sigma A_{11}^{1/2} x = A_{12} A_{22}^{-1/2} y = \nu^{-1/2} A_{12} y,$$

откуда следует, что

$$A_{12} y = \sigma \mu^{1/2} \nu^{1/2} x. \tag{2.17}$$

Аналогично выводится равенство

$$A_{21} x = \sigma \mu^{1/2} \nu^{1/2} y. \tag{2.18}$$

Теперь мы покажем, что

$$\mu = \nu = \frac{\lambda_1(A) + \lambda_n(A)}{2}. \tag{2.19}$$

Принимая во внимание нормировку $\|x\| = \|y\| = 1$ и используя соотношения (2.16), (2.17) и (2.18), мы получаем равенство

$$A_z = \begin{bmatrix} \mu & \sigma\mu^{1/2}\nu^{1/2} \\ \sigma\mu^{1/2}\nu^{1/2} & \nu \end{bmatrix}. \tag{2.20}$$

Два собственных значения $\lambda_\pm$ матрицы A_z определяются формулой

$$\lambda_\pm = \frac{\mu+\nu \pm [(\mu-\nu)^2 + 4\sigma^2\mu\nu]^{1/2}}{2}. \tag{2.21}$$

Как известно [5, следствие 2.1], имеют место неравенства

$$\lambda_n(A) \leqslant \lambda_- \leqslant \lambda_+ \leqslant \lambda_1(A).$$

Следовательно, с учетом (2.21) мы имеем:

$$\frac{\mu+\nu + [(\mu-\nu)^2 + 4\sigma^2\mu\nu]^{1/2}}{\mu+\nu - [(\mu-\nu)^2 + 4\sigma^2\mu\nu]^{1/2}} \leqslant \frac{\lambda_1(A)}{\lambda_n(A)}.$$

Это неравенство можно переписать в виде

$$\frac{\lambda_1(A) - \lambda_n(A)}{\lambda_1(A) + \lambda_n(A)}(\mu+\nu) \geqslant [(\mu-\nu)^2 + 4\sigma^2\mu\nu]^{1/2},$$

а также, ввиду соотношения (2.15), в виде

$$\sigma^2(\mu+\nu)^2 \geqslant (\mu-\nu)^2 + 4\sigma^2\mu\nu.$$

Последнее неравенство, очевидно, равносильно неравенству

$$\sigma^2(\mu-\nu)^2 \geqslant (\mu-\nu)^2.$$

Но поскольку (например, в силу (2.15)) $\sigma < 1$, из полученного неравенства с необходимостью следует, что

$$\mu = \nu.$$

Теперь, в силу (2.16)–(2.18) и (2.15), мы имеем:

$$\begin{aligned} A\begin{bmatrix} x \\ y \end{bmatrix} &= \mu(1+\sigma)\begin{bmatrix} x \\ y \end{bmatrix} = \mu\frac{2\lambda_1(A)}{\lambda_1(A)+\lambda_n(A)}\begin{bmatrix} x \\ y \end{bmatrix}, \\ A\begin{bmatrix} x \\ -y \end{bmatrix} &= \mu(1-\sigma)\begin{bmatrix} x \\ -y \end{bmatrix} = \mu\frac{2\lambda_n(A)}{\lambda_1(A)+\lambda_n(A)}\begin{bmatrix} x \\ -y \end{bmatrix}. \end{aligned} \tag{2.22}$$

Равенства (2.22) показывают, что как $\mu\frac{2\lambda_1(A)}{\lambda_1(A)+\lambda_n(A)}$, так и $\mu\frac{2\lambda_n(A)}{\lambda_1(A)+\lambda_n(A)}$ являются собственными значениями матрицы A. Следовательно, имеют место неравенства

$$\lambda_1(A) \geqslant \mu\frac{2\lambda_1(A)}{\lambda_1(A)+\lambda_n(A)}, \quad \mu\frac{2\lambda_n(A)}{\lambda_1(A)+\lambda_n(A)} \geqslant \lambda_n(A),$$

которые равносильны неравенствам

$$\frac{\lambda_1(A)+\lambda_n(A)}{2} \geqslant \mu \geqslant \frac{\lambda_1(A)+\lambda_n(A)}{2},$$

показывающим, что

$$\mu = \frac{\lambda_1(A)+\lambda_n(A)}{2},$$

что и завершает доказательство (2.19).

Заметим теперь, что из (2.19) и (2.22) немедленно следует справедливость утверждения (i). С другой стороны, используя соотношения (2.20), (2.19) и (2.15), мы получаем, что

$$A_z = \frac{\lambda_1(A)+\lambda_n(A)}{2}\begin{bmatrix}1 & \sigma\\ \sigma & 1\end{bmatrix} = \frac{1}{2}\begin{bmatrix}\lambda_1(A)+\lambda_n(A) & \lambda_1(A)-\lambda_n(A)\\ \lambda_1(A)-\lambda_n(A) & \lambda_1(A)+\lambda_n(A)\end{bmatrix}.$$

Тем самым (ii) также доказано.

Обратно, пусть справедливо утверждение (i). В этом случае мы имеем:

$$\begin{aligned} &A_{11}x + A_{12}y = \lambda_1(A)x, \quad A_{11}x - A_{12}y = \lambda_n(A)x,\\ &A_{21}x + A_{22}y = \lambda_1(A)y, \quad A_{21}x - A_{22}y = -\lambda_n(A)y.\end{aligned}$$

Как легко видеть, из этих равенств следуют соотношения

$$A_{11}x = \frac{\lambda_1(A)+\lambda_n(A)}{2}x, \quad A_{22}y = \frac{\lambda_1(A)+\lambda_n(A)}{2}y \tag{2.23}$$

и

$$A_{21}x = \frac{\lambda_1(A)-\lambda_n(A)}{2}y, \quad A_{12}y = \frac{\lambda_1(A)-\lambda_n(A)}{2}x. \tag{2.24}$$

Используя эти последние, нетрудно убедиться, что

$$\begin{aligned} S_1(A)A_{11}^{-1}x = (I - A_{12}A_{22}^{-1}A_{21}A_{11}^{-1})x &= \left[1 - \left(\frac{\lambda_1(A)-\lambda_n(A)}{\lambda_1(A)+\lambda_n(A)}\right)^2\right]x\\ &= \frac{4k(A)}{[k(A)+1]^2}x.\end{aligned}$$

Таким образом мы приходим к неравенству

$$\xi(A) = \lambda_{\max}(A_{11}S_1^{-1}(A)) \geqslant \frac{[k(A)+1]^2}{4k(A)},$$

которое в сочетании с противоположным неравенством, гарантируемым теоремой 2.1, и устанавливает равенство (2.8).

Наконец, мы покажем, что утверждения (i) и (ii) равносильны. Действительно, как уже было показано, из (i) следуют неравенства (2.23) и (2.24), из которых вытекают соотношения

$$\frac{x^* A_{11} x}{x^* x} = \frac{y^* A_{22} y}{y^* y} = \frac{\lambda_1(A) + \lambda_n(A)}{2} \tag{2.25}$$

и

$$\frac{x^* A_{12} y}{\|x\|\,\|y\|} = \frac{\lambda_1(A) - \lambda_n(A)}{2}\,\frac{\|x\|}{\|y\|},$$

$$\frac{y^* A_{21} x}{\|x\|\,\|y\|} = \frac{\lambda_1(A) - \lambda_n(A)}{2}\,\frac{\|y\|}{\|x\|}. \tag{2.26}$$

Поскольку матрица A эрмитова, ее собственные векторы $\begin{bmatrix} x \\ y \end{bmatrix}$ и $\begin{bmatrix} x \\ -y \end{bmatrix}$ необходимо ортогональны, что приводит нас к равенству

$$\|x\| = \|y\|. \tag{2.27}$$

Учитывая (2.27), соотношения (2.26) можно переписать в виде

$$\frac{x^* A_{12} y}{\|x\|\,\|y\|} = \frac{y^* A_{21} x}{\|x\|\,\|y\|} = \frac{\lambda_1(A) - \lambda_n(A)}{2}. \tag{2.28}$$

Теперь утверждение (ii) тривиально следует из (2.25) и (2.28).

Обратно, предположим, что выполнено условие (ii). В этом случае для вектора

$$\widehat{z} = \begin{bmatrix} \dfrac{x}{\|x\|} \\ \dfrac{y}{\|y\|} \end{bmatrix}$$

мы, очевидно, имеем:

$$A_{\widehat{z}} = A_z\,.$$

Поэтому, не теряя общности, мы можем предполагать, что (ii) справедливо при дополнительном предположении

$$\|x\| = \|y\| = 1. \tag{2.29}$$

С учетом (2.29) мы имеем:

$$\frac{z^* A z}{z^* z} = \frac{[1\ 1] A_z \begin{bmatrix} 1 \\ 1 \end{bmatrix}}{2},$$

а поскольку, ввиду (ii), имеет место равенство

$$A_z \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \lambda_1(A) \begin{bmatrix} 1 \\ 1 \end{bmatrix},$$

мы заключаем, что

$$\frac{z^* A z}{z^* z} = \lambda_1(A).$$

Значит z является собственным вектором матрицы A, отвечающим ее наибольшему собственному значению $\lambda_1(A)$.

Аналогичным образом устанавливаются равенства

$$\frac{\widetilde{z}^* A \widetilde{z}}{\widetilde{z}^* \widetilde{z}} = \frac{[1\ -1] A_z \begin{bmatrix} 1 \\ -1 \end{bmatrix}}{2} = \lambda_n(A),$$

показывающие, что $\widetilde{z}$ – собственный вектор матрицы A, отвечающий ее наименьшему собственному значению $\lambda_n(A)$. Это доказывает, что (i) следует из (ii), и завершает доказательство теоремы 2.2. □

Замечание 2.1. Тот факт, что из условия (i) следует оптимальная обусловленность матрицы A относительно $\Delta(m, n-m)$, может быть доказан в точности таким же образом, как это было сделано в работе [4] для блочных 2×2 матриц, отмасштабированных с помощью блочного метода Якоби. Действительно, поскольку

$$\widetilde{z} = Sz, \quad \text{где} \quad S = \begin{bmatrix} I_m & 0 \\ 0 & -I_{n-m} \end{bmatrix},$$

мы имеем равенство

$$SA^{-1}SAz = k(A)z,$$

из которого следует, что

$$k(A) \leqslant \lambda_{\max}(SA^{-1}SA).$$

С другой стороны, для произвольной матрицы $D \in \Delta(m, n-m)$, учитывая, что $\|S\| = 1$ и что матрицы D и S коммутируют, мы выводим:

$$\begin{aligned} \lambda_{\max}(SA^{-1}SA) \leqslant \|D^{-1}(SA^{-1}SA)D\| &= \|S(D^{-1}A^{-1}D^{-*})S(D^*AD)\| \\ &\leqslant \|(D^*AD)^{-1}\|\,\|D^*AD\| = k(D^*AD). \end{aligned}$$

Таким образом, для любой матрицы $D \in \Delta(m, n-m)$ справедливо неравенство

$$k(A) \leqslant k(D^*AD).$$

Теперь, основываясь на теореме 2.2, мы покажем, что отличная от скалярной эрмитова положительно определенная матрица A, которая оптимально обусловлена относительно $\Delta(m, n-m)$, может быть приведена с помощью унитарной матрицы вида

$$Q = \begin{bmatrix} Q_1 & 0 \\ 0 & Q_2 \end{bmatrix} \in \Delta(m, n-m)$$

к оптимально же обусловленной разложимой матрице

$$Q^*AQ = \begin{bmatrix} \frac{\lambda_1(A)+\lambda_n(A)}{2} & 0 & \dots & 0 & \frac{\lambda_1(A)-\lambda_n(A)}{2} & 0 & \dots & 0 \\ 0 & * & \dots & * & 0 & * & \dots & * \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & * & \dots & * & 0 & * & \dots & * \\ \frac{\lambda_1(A)-\lambda_n(A)}{2} & 0 & \dots & 0 & \frac{\lambda_1(A)+\lambda_n(A)}{2} & 0 & \dots & 0 \\ 0 & * & \dots & * & 0 & * & \dots & * \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & * & \dots & * & 0 & * & \dots & * \end{bmatrix}, \qquad (2.30)$$

имеющей 2×2 неразложимую компоненту

$$\frac{1}{2} \begin{bmatrix} \lambda_1(A) + \lambda_n(A) & \lambda_1(A) - \lambda_n(A) \\ \lambda_1(A) - \lambda_n(A) & \lambda_1(A) + \lambda_n(A) \end{bmatrix}.$$

Действительно, если $x \in \mathbb{C}^m$ и $y \in \mathbb{C}^{n-m}$, $\|x\| = \|y\| = 1$, – это соответственно левый и правый сингулярные векторы матрицы $A_{11}^{-1/2} A_{12} A_{22}^{-1/2}$, отвечающие ее наибольшему сингулярному значению

$$\sigma = \|A_{11}^{-1/2} A_{12} A_{22}^{-1/2}\|,$$

то, как было установлено в ходе доказательства теоремы 2.2 (см. (2.15)–(2.19)), имеют место следующие равенства:

$$\sigma = \frac{\lambda_1(A) - \lambda_n(A)}{\lambda_1(A) + \lambda_n(A)},$$

$$A_{11}x = \frac{\lambda_1(A) + \lambda_n(A)}{2} x, \quad A_{22}y = \frac{\lambda_1(A) + \lambda_n(A)}{2} y,$$

$$A_{12}y = \frac{\lambda_1(A) - \lambda_n(A)}{2}x, \quad A_{21}x = \frac{\lambda_1(A) - \lambda_n(A)}{2}y.$$

Выбирая унитарные матрицы $Q_1 \in \mathbb{C}^{m\times m}$ и $Q_2 \in \mathbb{C}^{(n-m)\times(n-m)}$ вида

$$Q_1 = [x, u_2, \dots, u_m], \quad Q_2 = [y, v_2, \dots, v_{n-m}],$$

мы легко убеждаемся, что

$$Q_1^* A_{11} Q_1 e_1^{(m)} = Q_1^* A_{11} x = \begin{bmatrix} \frac{\lambda_1(A)+\lambda_n(A)}{2} \\ 0 \\ \vdots \\ 0 \end{bmatrix},$$

$$Q_2^* A_{22} Q_2 e_1^{(n-m)} = Q_2^* A_{22} y = \begin{bmatrix} \frac{\lambda_1(A)+\lambda_n(A)}{2} \\ 0 \\ \vdots \\ 0 \end{bmatrix},$$

$$Q_1^* A_{12} Q_2 e_1^{(n-m)} = Q_1^* A_{12} y = \begin{bmatrix} \frac{\lambda_1(A)-\lambda_n(A)}{2} \\ 0 \\ \vdots \\ 0 \end{bmatrix},$$

$$e_1^{(m)*} Q_1^* A_{12} Q_2 = x^* A_{12} Q_2 = \left[\frac{\lambda_1(A) - \lambda_n(A)}{2}, 0, \dots, 0\right],$$

где $e_1^{(k)} = [1, 0, \dots, 0]^* \in \mathbb{C}^k$. Тем самым показано, что эрмитова матрица Q^*AQ имеет требуемый вид (2.30).

Обратно, если для некоторой блочно диагональной унитарной матрицы $Q \in \Delta(m, n-m)$ справедливо соотношение (2.30), то как Q^*AQ так и A обе оптимально обусловлены относительно $\Delta(m, n-m)$. Действительно, как легко видеть, векторы

$$z = \begin{bmatrix} e_1^{(m)} \\ e_1^{(n-m)} \end{bmatrix}, \quad \widetilde{z} = \begin{bmatrix} e_1^{(m)} \\ -e_1^{(n-m)} \end{bmatrix}$$

являются собственными для матрицы Q^*AQ и отвечают ее экстремальным собственным значениям $\lambda_1(Q^*AQ) = \lambda_1(A)$ и $\lambda_n(Q^*AQ) = \lambda_n(A)$. Следовательно, по теореме 2.2 матрица Q^*AQ оптимально обусловлена относительно $\Delta(m, n-m)$.

Таким образом, на основании теоремы 2.2 нами установлен следующий результат.

Следствие 2.2. *В условиях теоремы* 2.2 *матрица* A *оптимально обусловлена относительно* $\Delta(m, n-m)$ *тогда и только тогда, когда существует унитарная матрица* $Q \in \Delta(m, n-m)$, *такая что матрица* Q^*AQ *имеет вид* (2.30).

Литература

1. H. Bart, I. Gohberg, M. Kaashoek, and P. Van Dooren, *Factorizations of transfer functions.* SIAM J. Contr. Optimiz., **18** (1980), 675–696.
2. J. Demmel, *The condition number of equivalence transformations that block diagonalize matrix pencils.* In: *Matrix Pencils* (B. Kagström and A. Ruhe, eds.), *Lecture Notes in Mathematics*, **973**, Springer, 1982, pp. 2–16.
3. S. C. Eisenstat, J. W. Lewis, and M. H. Schultz, *Optimal block diagonal scaling of block 2-cyclic matrices.* Linear Algebra Appl., **44** (1982), 181–186.
4. L. Elsner, *A note on optimal block scaling of matrices.* Numer. Math., **44** (1984), 127–128.
5. L. Yu. Kolotilina, *Eigenvalue bounds and inequalities using vector aggregation of matrices.* Linear Algebra Appl., **271** (1998), 139–167.
6. M. Marcus and H. Minc, *A Survey of Matrix Theory and Matrix Inequalities.* Allyn and Balon, Boston, 1964.
7. Song-Gui Wang and Wai-Cheung Ip, *A matrix version of the Wielandt inequality and its applications to statistics.* Linear Algebra Appl., **296** (1999), 171–181.

Kolotilina L. Yu. Optimally conditioned block 2×2 matrices.

This paper provides a description of block 2×2 Hermitian positive-definite matrices that are optimally conditioned with respect to block diagonal similarity transformations.

Санкт-Петербургское отделение
Математического института
им. В. А. Стеклова

Поступило 10 апреля 2000 г.