15 Двунивови методи. Константа в усиленото неравенство на Коши-Буняковски-Шварц.

15.1 Константа в усиленото неравенство на Коши-Буняковски-Шварц.

Нека V_1 и V_2 са крайномерни векторни пространства, $W = V_1 \times V_2$ и нека

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

където блоковете на матрицата са съгласувани по размерност с V_1 и V_2. Предполагаме, че A е симетрична и положително полуопределена матрица с неособен блок A_{11} и следователно A_{11} е положително определена. Нека пространствата W_1 и W_2 са дефинирани, както следва:

$$W_1 = \left\{ \mathbf{v} = \begin{bmatrix} \mathbf{v_1} \\ 0 \end{bmatrix}, \mathbf{v_1} \in V_1 \right\}, W_2 = \left\{ \mathbf{v} = \begin{bmatrix} 0 \\ \mathbf{v_2} \end{bmatrix}, \mathbf{v_2} \in V_2 \right\}.$$

Нека γ е минималната положителна константа, удовлетворяваща за всеки ненулеви $\mathbf{w_i} \in W_i$, $i = 1, 2$, неравенството

$$|\mathbf{w_1}^T A \mathbf{w_2}| \leq \gamma \{\mathbf{w_1}^T A \mathbf{w_1} \;\; \mathbf{w_2}^T A \mathbf{w_2}\}^{1/2}. \tag{15.1}$$

Така дефинираната константа γ характеризира блочното разделяне на матрицата A и геометрично се интерпретира, като косинус на абстрактния ъгъл между W_1 и W_2 относно скаларното произведение породено от A. Ако A е положително определена, лесно може да се докаже, че $\gamma < 1$. Когато A е особена, се налага да се изследват ядрата на A и A_{ii}, $i = 1, 2$. В специализираната литература (15.1) е известно, като усиленото неравенство на Коши-Буняковски-Шварц (КБШ).

От блочното представяне на A следва, че (15.1) е еквивалентно на неравенството

$$|\mathbf{v_1}^T A_{12} \mathbf{v_2}| \leq \gamma \{\mathbf{v_1}^T A_{11} \mathbf{v_1} \;\; \mathbf{v_2}^T A_{22} \mathbf{v_2}\}^{1/2} \tag{15.2}$$

за всеки ненулеви $\mathbf{v_i} \in V_i$, $i = 1, 2$.

Лема 15.1 *Нека γ е дефинирана с помощта на (15.1) или (15.2). Тогава:*

(a) $\gamma \leq 1$;

(b) $\gamma = 1$, *ако съществува вектор* $\begin{bmatrix} \mathbf{v_1} \\ \mathbf{v_2} \end{bmatrix} \in \mathcal{N}(A)$, *за който* $\mathbf{v_2}$ *не принадлежи на ядрото* $\mathcal{N}(A_{22})$;

(c) $\gamma < 1$, *ако за всеки вектор* $\begin{bmatrix} \mathbf{v_1} \\ \mathbf{v_2} \end{bmatrix} \in \mathcal{N}(A)$, $\mathbf{v_2} \in \mathcal{N}(A_{22})$;

(d) *Ако е в сила условие (c),*

$$\gamma = \sup_{\mathbf{v_i} \in V_i \setminus \mathcal{N}(A_{ii}), \quad i=1,2} \frac{\mathbf{v_1}^T A_{12} \mathbf{v_2}}{\{\mathbf{v_1}^T A_{11} \mathbf{v_1} \ \mathbf{v_2}^T A_{22} \mathbf{v_2}\}^{1/2}}. \tag{15.3}$$

[Тъй като по предположение A_{11} е положително определена, то $\mathcal{N}(A_{11}) = \{\mathbf{0}\}$.]

Доказателство. Нека $\mathbf{w} = \begin{bmatrix} \mathbf{v_1} \\ \mathbf{v_2} \end{bmatrix}$, където $\mathbf{v_1}, \mathbf{v_2} \neq 0$. Тъй като A е симетрична и положително полуопределена, то $\mathbf{w}^T A \mathbf{w} \geq 0$ и следователно

$$\mathbf{v_1}^T A_{11} \mathbf{v_1} + \mathbf{v_2}^T A_{22} \mathbf{v_2} + 2\mathbf{v_1}^T A_{12} \mathbf{v_2} \geq 0. \tag{15.4}$$

Нека $\gamma(\mathbf{v_1}, \mathbf{v_2})$ е константата, съответстваща на векторите $\mathbf{v_1}, \mathbf{v_2}$. Очевидно $\gamma(\mathbf{v_1}, \mathbf{v_2}) = \gamma(\alpha \mathbf{v_1}, \beta \mathbf{v_2})$, за всеки $\alpha, \beta \neq 0$. Следователно винаги можем да предполагаме, че векторите са подходящо скалирани, при което стойността на γ остава непроменена.

Нека разгледаме най-напред случая, когато $\mathbf{v_2}$ е такова, че $\mathbf{v_2}^T A_{22} \mathbf{v_2} \neq 0$. Тъй като A_{11} е по условие положително определена, то можем да изберем $\mathbf{v_1}$ така, че $\mathbf{v_1}^T A_{11} \mathbf{v_1} = \mathbf{v_2}^T A_{22} \mathbf{v_2}$. Заместваме в (15.4) и получаваме

$$|\mathbf{v_1}^T A_{12} \mathbf{v_2}| \leq \mathbf{v_1}^T A_{11} \mathbf{v_1}$$

и следователно $\gamma(\mathbf{v_1}, \mathbf{v_2}) \leq 1$.

Ако $\mathbf{v_2}$ е такова, че $\mathbf{v_2}^T A_{22} \mathbf{v_2} = 0$, то нека $\mathbf{v_1} = \tau \hat{\mathbf{v}}_\mathbf{1}$. Тогава

$$\tau \hat{\mathbf{v}}_\mathbf{1}^T A_{11} \hat{\mathbf{v}}_\mathbf{1} + 2\hat{\mathbf{v}}_\mathbf{1}^T A_{12} \mathbf{v_2} \geq 0.$$

Тъй като това е в сила за всяко $\tau > 0$, то очевидно

$$\hat{\mathbf{v}}_{\mathbf{1}}^T A_{12} \mathbf{v_2} \geq 0.$$

Ако допуснем, че $A_{12}\mathbf{v_2} \neq 0$ и изберем $\hat{\mathbf{v}}_{\mathbf{1}} = -A_{12}\mathbf{v_2}$, ще получим

$$-||A_{12}\mathbf{v_2}|| \geq 0$$

и следователно

$$\mathbf{v_2}^T A_{22} \mathbf{v_2} = 0 \qquad ==> \qquad A_{12}\mathbf{v_2} = 0, \tag{15.5}$$

т.е. неравенството на КБШ е изпълнено за взяко γ. С това условие (a) е доказано.

Нека $A\mathbf{w} \neq 0$. Тогава в (15.4) неравенството е строго и следователно $\gamma < 1$.

Нека съществува вектор $\begin{bmatrix} \mathbf{v_1} \\ \mathbf{v_2} \end{bmatrix} \in \mathcal{N}(A)$, за който $\mathbf{v_2}$ не принадлежи на ядрото $\mathcal{N}(A_{22})$. Избираме $\mathbf{v_1}$ така, че $\mathbf{v_1}^T A_{11} \mathbf{v_1} = \mathbf{v_2}^T A_{22} \mathbf{v_2}$. Заместваме в (15.4) и получаваме

$$-\mathbf{v_1}^T A_{12} \mathbf{v_2} = \mathbf{v_1}^T A_{11} \mathbf{v_1}$$

и следователно $\gamma(\mathbf{v_1}, \mathbf{v_2}) = 1$. С това твърдение (b) е доказано.

Нека е в сила условието от твърдение (c), т.е. ако $A\mathbf{w} = 0$, то $A_{22}\mathbf{v_2} = \mathbf{0}$. Тогава от (15.5) и (15.4) следва, че $A_{11}\mathbf{v_1} = 0$. Тъй като A_{11} е положително определена, то $\mathbf{v_1} = 0$. Следователно този случай се изключва при изследване на константата в усиленото неравенство на КБШ, с което са доказани твърдения (c) и (d).

■

Лема 15.2 *Нека A е симетрична и положително полуопределена матрица с положително определен блок A_{11}. Нека освен това, ако $A\mathbf{w} = 0$ то $A_{22}\mathbf{v_2} = \mathbf{0}$, където $\mathbf{w} = \begin{bmatrix} \mathbf{v_1} \\ \mathbf{v_2} \end{bmatrix}$. Тогава:*

(a)

$$\gamma^2 = \sup_{\mathbf{v_2} \in V_2 \setminus \mathcal{N}(A_{22})} \frac{\mathbf{v_2}^T A_{21} A_{11}^{-1} A_{12} \mathbf{v_2}}{\mathbf{v_2}^T A_{22} \mathbf{v_2}}; \tag{15.6}$$

(*b*)

$$1-\gamma^2 \le \frac{\mathbf{v_2}^T S \mathbf{v_2}}{\mathbf{v_2}^T A_{22} \mathbf{v_2}} \le 1, \tag{15.7}$$

където неравенствата са точни.

Доказателство. От твърдение (d) на предходната лема получаваме

$$\begin{aligned}
\gamma &= \sup_{\mathbf{v_i} \in V_i \setminus \mathcal{N}(A_{ii}),\quad i=1,2} \frac{\mathbf{v_1}^T A_{12} \mathbf{v_2}}{\{\mathbf{v_1}^T A_{11} \mathbf{v_1}\ \mathbf{v_2}^T A_{22} \mathbf{v_2}\}^{1/2}} \\
&= \sup_{\mathbf{v_1} \neq \mathbf{0},\quad \mathbf{v_2} \in V_2 \setminus \mathcal{N}(A_{22})} \frac{\mathbf{v_1}^T A_{11}^{-1/2} A_{12} \mathbf{v_2}}{\{\mathbf{v_1}^T \mathbf{v_1}\ \mathbf{v_2}^T A_{22} \mathbf{v_2}\}^{1/2}},
\end{aligned}$$

където супремумът се достига за $\mathbf{v_1} = A_{11}^{-1/2} A_{12} \mathbf{v_2}$ и следователно

$$\gamma = \left\{ \sup_{\mathbf{v_2} \in V_2 \setminus \mathcal{N}(A_{22})} \frac{\mathbf{v_2}^T A_{21} A_{11}^{-1} A_{12} \mathbf{v_2}}{\mathbf{v_2}^T A_{22} \mathbf{v_2}} \right\}^{1/2}.$$

С това твърдение (a) е доказано. Твърдение (b) следва пряко от (a) и от дефиницията на допълнението на Шур $S = A_{22} - A_{21} A_{11}^{-1} A_{12}$. ■

15.2 Алгебрични двунивови методи.

От алгебрична гледна точка, съвременните многонивови методи за преобуславяне представляват рекурсивно обобщение на съответните двунивови методи. Алгебричните адитивен и мултипликативен двунивов преобусловител се дефинират, както следва:

$$C_A = \begin{bmatrix} A_{11} & \\ & A_{22} \end{bmatrix}, \tag{15.8}$$

$$C_M = \begin{bmatrix} A_{11} & \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} I & A_{11}^{-1} A_{12} \\ & I \end{bmatrix}, \tag{15.9}$$

т.е. адитивният преобусловител C_A представлява блочно-диагоналната част на A, а мултипликативният C_M се получава, като заместим с A_{22} допълнението на Шур

$$S = A_{22} - A_{21} A_{11}^{-1} A_{12}$$

в точната факторизация

$$A=\begin{bmatrix} A_{11} & \\ A_{21} & S \end{bmatrix}\begin{bmatrix} I & A_{11}^{-1}A_{12} \\ & I \end{bmatrix}.$$

Тук, с помощта на константата в усиленото неравенство на КБШ, ще оценим относителното число на обусловеност за така дефинираните алгебрични двунивови методи.

Теорема 15.1 *За всеки вектор* $\mathbf{w}=\begin{bmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \end{bmatrix}$ *са в сила оценките*

$$(1-\gamma)\left(\mathbf{v}_1^T A_{11}\mathbf{v}_1+\mathbf{v}_2^T A_{22}\mathbf{v}_2\right)\le \mathbf{w}^T A\mathbf{w}\le (1+\gamma)\left(\mathbf{v}_1^T A_{11}\mathbf{v}_1+\mathbf{v}_2^T A_{22}\mathbf{v}_2\right)$$

и следователно

$$(1-\gamma)\mathbf{w}^T C_A\mathbf{w}\le \mathbf{w}^T A\mathbf{w}\le (1+\gamma)\mathbf{w}_A^C\mathbf{w}. \tag{15.10}$$

Доказателство. Прилагаме последователно (15.2) и известното от училищния курс неравенство

$$2ab\le a^2+b^2.$$

Така получаваме оценката отдясно, т.е.

$$\begin{aligned} \mathbf{w}^T A\mathbf{w} &= \mathbf{v}_1^T A_{11}\mathbf{v}_1+\mathbf{v}_2^T A_{22}\mathbf{v}_2+\mathbf{v}_2^T A_{21}\mathbf{v}_1+\mathbf{v}_1^T A_{12}\mathbf{v}_2 \\ &\le \mathbf{v}_1^T A_{11}\mathbf{v}_1+\mathbf{v}_2^T A_{22}\mathbf{v}_2+2\gamma\sqrt{\mathbf{v}_1^T A_{11}\mathbf{v}_1}\sqrt{\mathbf{v}_2^T A_{22}\mathbf{v}_2} \\ &\le (1+\gamma)\left(\mathbf{v}_1^T A_{11}\mathbf{v}_1+\mathbf{v}_2^T A_{22}\mathbf{v}_2\right). \end{aligned}$$

По аналогичен начин от (15.2) и

$$-2ab\ge -a^2-b^2$$

следва

$$\begin{aligned} \mathbf{w}^T A\mathbf{w} &= \mathbf{v}_1^T A_{11}\mathbf{v}_1+\mathbf{v}_2^T A_{22}\mathbf{v}_2+\mathbf{v}_2^T A_{21}\mathbf{v}_1+\mathbf{v}_1^T A_{12}\mathbf{v}_2 \\ &\ge \mathbf{v}_1^T A_{11}\mathbf{v}_1+\mathbf{v}_2^T A_{22}\mathbf{v}_2-2\gamma\sqrt{\mathbf{v}_1^T A_{11}\mathbf{v}_1}\sqrt{\mathbf{v}_2^T A_{22}\mathbf{v}_2} \\ &\ge (1-\gamma)\left(\mathbf{v}_1^T A_{11}\mathbf{v}_1+\mathbf{v}_2^T A_{22}\mathbf{v}_2\right), \end{aligned}$$

с което теоремата е доказана. ■

Теорема 15.2 *За двунивовия мултипликативен преобуславител са в сила оценките*

$$(1-\gamma^2)\mathbf{w}^T C_M \mathbf{w} \le \mathbf{w}^T A \mathbf{w} \le \mathbf{w}^T C_M \mathbf{w} \tag{15.11}$$

за всеки вектор $\mathbf{w}$.

Доказателство. Твърдението следва непосредствено от (15.7) и равенствата

$$A = \begin{bmatrix} I & \\ A_{21}A_{11}^{-1} & I \end{bmatrix} \begin{bmatrix} A_{11} & \\ & S \end{bmatrix} \begin{bmatrix} I & A_{11}^{-1}A_{12} \\ & I \end{bmatrix}$$

и

$$C_M = \begin{bmatrix} I & \\ A_{21}A_{11}^{-1} & I \end{bmatrix} \begin{bmatrix} A_{11} & \\ & A_{22} \end{bmatrix} \begin{bmatrix} I & A_{11}^{-1}A_{12} \\ & I \end{bmatrix}.$$

Действително,

$$\begin{aligned} \mathbf{w}^T A \mathbf{w} &= \hat{\mathbf{w}}^T \begin{bmatrix} A_{11} & \\ & S \end{bmatrix} \hat{\mathbf{w}} \\ &= \hat{\mathbf{v}}_1^T A_{11} \hat{\mathbf{v}}_1 + \hat{\mathbf{v}}_2^T S \hat{\mathbf{v}}_2 \\ &\ge \hat{\mathbf{v}}_1^T A_{11} \hat{\mathbf{v}}_1 + (1-\gamma^2) \hat{\mathbf{v}}_2^T A_{22} \hat{\mathbf{v}}_2 \\ &\ge (1-\gamma^2)\hat{\mathbf{w}}^T \begin{bmatrix} A_{11} & \\ & A_{22} \end{bmatrix} \hat{\mathbf{w}} \\ &= (1-\gamma^2)\mathbf{w}^T C_M \mathbf{w}, \end{aligned}$$

където $\mathbf{w} = \begin{bmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \end{bmatrix}$ и $\hat{\mathbf{w}} = \begin{bmatrix} \hat{\mathbf{v}}_1 \\ \hat{\mathbf{v}}_2 \end{bmatrix} = \begin{bmatrix} I & A_{11}^{-1}A_{12} \\ & I \end{bmatrix} \mathbf{w}$. Дясното неравенство в (15.11) следва напълно аналогично, с което теоремата е доказана. ■

Следствие 15.1 *От (15.10) и (15.11) получаваме следните оценки за относителните числа на обусловеност на алгебричните двунивови методи:*

$$\kappa\left(C_A^{-1}A\right) \le \frac{1+\gamma}{1-\gamma} \tag{15.12}$$

$$\kappa\left(C_M^{-1}A\right) \le \frac{1}{1-\gamma^2} \tag{15.13}$$

Тук е важно да отбележим, че особено внимание от гледна точка на методите за преобуславяне заслужават такива блочни разделяния на матрицата A, за които γ е равномерно ограничена с константа, която не зависи от размерността на A.