Rozdział 5

Obraz, rząd i jądro macierzy

5.1 Obraz i rząd macierzy

5.1.1 Rząd kolumnowy i rząd wierszowy

Niech $A \in \mathbf{K}^{m,n}$ będzie dana w postaci blokowej,

$$A = [\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n], \quad \vec{a}_j \in \mathbf{K}^m,\ 1 \le j \le n.$$

Obraz macierzy A definiujemy jako

$$\mathcal{R}(A) \;:=\; \{\, A * \vec{x} :\ \vec{x} \in \mathbf{K}^n \,\} \;=\; \mathrm{span}(\vec{a}_1, \vec{a}_2 \ldots, \vec{a}_n) \subseteq \mathbf{K}^m.$$

Dalej, *rząd kolumnowy* macierzy A definiujemy jako

$$\mathrm{rz}_k(A) \;:=\; \dim\left(\mathcal{R}(A)\right).$$

Oczywiście, $0 \le \mathrm{rz}_k(A) \le \min(m,n)$. Przedstawiając z kolei A jako wektory-wiersze (funkcjonały),

$$A = \begin{bmatrix} \hat{a}_1^T \\ \vdots \\ \hat{a}_m^T \end{bmatrix},$$

definiujemy *rząd wierszowy* macierzy A jako

$$\mathrm{rz}_w(A) \;=\; \dim\left(\mathcal{R}(A^T)\right) \;=\; \dim\left(\mathrm{span}(\hat{a}_1, \hat{a}_2, \ldots, \hat{a}_n)\right).$$

Podobnie jak dla rzędu kolumnowego, $0 \le \mathrm{rz}_w(A) \le \min(m,n)$.

5.1.2 Rząd macierzy

Mamy następujące ważne twierdzenie.

Twierdzenie 5.1 *Dla dowolnej macierzy* $A \in \mathbf{K}^{m,n}$

$$\mathrm{rz}_k(A) = \mathrm{rz}_w(A).$$

Dowód. Oznaczmy

$$k = \mathrm{rz}_k(A) \quad \text{oraz} \quad w = \mathrm{rz}_w(A).$$

Zauważmy najpierw, że permutacja kolumn macierzy nie zmienia ani jej rzędu kolumnowego (bo to tylko zmiana kolejności wektorów) ani jej rzędu wierszowego (bo to tylko przenumerowanie współrzędnych, identyczne dla każdego z wektorów). Podobnie rzędów nie zmienia permutacja wierszy.

Dokonajmy więc, dla uproszczenia, takiej permutacji kolumn, a potem wierszy, aby otrzymana macierz $\hat{A}$ była postaci

$$\hat{A} = \begin{bmatrix} A_I & A_{II} \end{bmatrix},$$

gdzie $A_I \in \mathbf{K}^{m,k}$, $A_{II} \in \mathbf{K}^{m,n-k}$, $\mathrm{rz}_k(A_I) = k$, oraz

$$A_I = \begin{bmatrix} A_1 \\ A_2 \end{bmatrix},$$

przy czym $A_1 \in \mathbf{K}^{w_1,k}$, $A_2 \in \mathbf{K}^{m-w_1,k}$, $w_1 := \mathrm{rz}_w(A_I) = \mathrm{rz}_w(A_1)$. Oczywiście

$$w_1 \leq w,$$

bo wiersze A_1 są "obciętymi" wierszami $\hat{A}$.

Ponieważ wektory-wiersze macierzy A_2 są liniowo zależne od wektorów-wierszy macierzy A_1 to istnieje macierz $B \in \mathbf{K}^{w_1,m-w_1}$ taka, że $A_2^T = A_1^T * B$ (gdzie kolejne kolumny B są współczynnikami odpowiednich kombinacji liniowych), czyli $A_2 = B^T * A_1$. Dla dowolnego $\vec{x} \in \mathbf{K}^k$ mamy więc

$$A_I * \vec{x} = \begin{bmatrix} A_1 * \vec{x} \\ A_2 * \vec{x} \end{bmatrix} = \begin{bmatrix} A_1 * \vec{x} \\ B^T * A_1 * \vec{x} \end{bmatrix}.$$

Stąd, $A_1 * \vec{x} = \vec{0}$ wtedy i tylko wtedy gdy $A_I * \vec{x} = \vec{0}$, a ponieważ kolumny macierzy A_I są liniowo niezależne, oznacza to także liniową niezależność kolumn macierzy A_1. A jeśli tak to ich liczba k nie może przekroczyć w_1, czyli wymiaru przestrzeni do której należą.

Otrzymaliśmy więc, że

$$\mathrm{rz}_k(A) = \mathrm{rz}_k(\hat{A}) = k \le w_1 \le w = \mathrm{rz}_w(\hat{A}) = \mathrm{rz}_w(A).$$

Przeprowadzając podobne rozumowanie dla macierzy A^T otrzymujemy $\mathrm{rz}_w(A) \le \mathrm{rz}_k(A)$, a stąd ostatecznie $\mathrm{rz}_w(A) = \mathrm{rz}_k(A)$, co należało pokazać.

Na podstawie twierdzenia 5.1 poprawna jest następująca definicja rzędu macierzy.

Definicja 5.1 Rzędem *macierzy A nazywamy liczbę*

$$\mathrm{rz}(A) := \mathrm{rz}_k(A) = \mathrm{rz}_w(A).$$

5.2 Przestrzeń zerowa (jądro) macierzy

Dla $A \in \mathbf{K}^{m,n}$ zbiór

$$\mathcal{N}(A) := \left\{ \vec{x} \in \mathbf{K}^n : \; A * \vec{x} = \vec{0} \right\}$$

nazywamy *jądrem* macierzy A.

Niech $k = \mathrm{rz}(A)$. Załóżmy, że kolumny macierzy A zostały tak przestawione, że otrzymana macierz $\hat{A}$ ma postać

$$\hat{A} = \left[\begin{array}{cc} A_I & A_{II} \end{array} \right],$$

gdzie $A_I \in \mathbf{K}^{m,k}$, $A_{II} \in \mathbf{K}^{m,n-k}$, oraz $\mathrm{rz}(A_I) = \mathrm{rz}(\hat{A})$ $(= \mathrm{rz}(A))$. Jeśli tak to kolumny macierzy A_{II} są liniowo zależne od kolumn macierzy A_I. W konsekwencji $A_{II} = A_I * B$ dla pewnej $B \in \mathbf{K}^{k,n-k}$. Załóżmy teraz, że $\vec{x} \in \mathcal{N}(\hat{A})$. Przedstawiając $\vec{x}$ w postaci

$$\vec{x} = \left[\begin{array}{c} \vec{x}_I \\ \vec{x}_{II} \end{array} \right],$$

$\vec{x}_I \in \mathbf{K}^k$, $\vec{x}_{II} \in \mathbf{K}^{n-k}$, mamy

$$\begin{aligned} \vec{0} &= \hat{A} * \vec{x} = \left[\begin{array}{cc} A_I & A_{II} \end{array} \right] \left[\begin{array}{c} \vec{x}_I \\ \vec{x}_{II} \end{array} \right] = A_I * \vec{x}_I + A_{II} * \vec{x}_{II} \\ &= A_I * \vec{x}_I + A_I * B * \vec{x}_{II} = A_I * (\vec{x}_I + B * \vec{x}_{II}). \end{aligned}$$

Ostatnie wyrażenie jest liniową kombinacją kolumn macierzy A_I, a ponieważ kolumny te są liniowo niezależne to kombinacja ta daje wektor zerowy tylko wtedy gdy $\vec{x}_I + B * \vec{x}_{II} = \vec{0}$, czyli $\vec{x}_I = -B * \vec{x}_{II}$. Stąd

$$\begin{aligned} \mathcal{N}(\hat{A}) &= \left\{ \begin{bmatrix} -B * \vec{x}_{II} \\ \vec{x}_{II} \end{bmatrix} : \vec{x}_{II} \in \mathbf{K}^{n-k} \right\} \\ &= \left\{ \begin{bmatrix} -B \\ I_{n-k} \end{bmatrix} * \vec{x}_{II} : \vec{x}_{II} \in \mathbf{K}^{n-k} \right\}. \end{aligned}$$

Przedstawiając B kolumnowo, $B = [\vec{b}_1, \ldots, \vec{b}_{n-k}]$, otrzymujemy ostatecznie

$$\mathcal{N}(\hat{A}) = \mathcal{R}\left(\begin{bmatrix} -B \\ I_{n-k} \end{bmatrix} \right) = \operatorname{span}\left(\begin{bmatrix} -\vec{b}_1 \\ \vec{e}_1 \end{bmatrix}, \ldots, \begin{bmatrix} -\vec{b}_{n-k} \\ \vec{e}_{n-k} \end{bmatrix} \right),$$

gdzie jak zwykle $\vec{e}_j \in \mathbf{K}^{n-k}$ jest j-tym wersorem. Ponieważ $\vec{e}_1, \ldots, \vec{e}_{n-k}$ są liniowo niezależne to liniowo niezależne są też wektory w powyższym "span". Stąd $\dim(\mathcal{N}(\hat{A})) = n - k = n - \operatorname{rz}(A)$. Wobec równości $\dim(\mathcal{N}(\hat{A})) = \dim(\mathcal{N}(A))$ (bo permutacja kolumn skutkuje jedynie przestawieniem współrzędnych w jądrze) dostajemy następujący wniosek.

Wniosek 5.1 *Dla dowolnej macierzy* $A \in \mathbf{K}^{m,n}$

$$\dim(\mathcal{N}(A)) + \dim(\mathcal{R}(A)) = n.$$

5.3 Rozkład względem obrazu i jądra

Zatrzymajmy się na chwilę na przypadku gdy $\mathbf{K} \subseteq \mathbf{C}$. Ponieważ wtedy

$$\overline{\left(\sum_{j=1}^{n} \vec{a}_j * x_j \right)} = \sum_{j=1}^{n} \overline{\vec{a}}_j * \overline{x}_j$$

(gdzie sprzężenie wektora oznacza sprzężenie "po współrzędnych") to wektory $(\vec{a}_1, \ldots, \vec{a}_n)$ oraz $(\overline{\vec{a}}_1, \ldots, \overline{\vec{a}}_n)$ są jednocześnie albo liniowo niezależne, albo liniowo zależne. Stąd $\operatorname{rz}(\overline{A}) = \operatorname{rz}(A)$ (gdzie znów sprzężenie macierzy oznacza sprzężenie "po współrzędnych"). W konsekwencji,

$$\operatorname{rz}(A^H) = \operatorname{rz}(\overline{A}^T) = \operatorname{rz}(A^T) = \operatorname{rz}(A).$$

Łatwo można też wywnioskować inną własność; mianowicie, jeśli

$$A = B * C,$$

$A \in \mathbf{K}^{m,n}$, $B \in \mathbf{K}^{m,k}$, $C \in \mathbf{K}^{k,n}$, to

$$\mathrm{rz}(A) \leq \min(\mathrm{rz}(B), \mathrm{rz}(C)).$$

Rzeczywiście, równość $A = B*C$ oznacza, że kolumny macierzy A są liniową kombinacją kolumn macierzy B, a stąd $\mathcal{R}(A) \subseteq \mathcal{R}(B)$ i w konsekwencji $\mathrm{rz}(A) \leq \mathrm{rz}(B)$. Biorąc z kolei transpozycję mamy $A^T = C^T * B^T$ i to samo rozumowanie daje $\mathcal{R}(A^T) \subseteq \mathcal{R}(C^T)$ oraz

$$\mathrm{rz}(A) = \mathrm{rz}(A^T) \leq \mathrm{rz}(C^T) = \mathrm{rz}(C).$$

Na koniec jeszcze jedno istotne twierdzenie.

Twierdzenie 5.2 *Niech* $\mathbf{K} \subseteq \mathbf{C}$ *i* $A \in \mathbf{K}^{m,n}$. *Wtedy*

$$\begin{array}{rcl} \mathbf{K}^m & = & \mathcal{R}(A) \oplus \mathcal{N}(A^H) \\ \mathbf{K}^n & = & \mathcal{R}(A^H) \oplus \mathcal{N}(A). \end{array}$$

Dowód. Wystarczy pokazać pierwszą z tych równości. W tym celu najpierw uzasadnimy, że suma jest sumą prostą. Rzeczywiście, jeśli $\vec{y} \in \mathcal{R}(A) \cap \mathcal{N}(A^H)$ to $A^H * \vec{y} = \vec{0}$ oraz istnieje $\vec{x} \in \mathbf{K}^n$ taki, że $A * \vec{x} = \vec{y}$. Stąd

$$\|\vec{y}\|_2^2 \;=\; \vec{y}^H * \vec{y} \;=\; (A * \vec{x})^H * \vec{y} \;=\; \vec{x}^H * (A^H * \vec{y}) \;=\; 0,$$

czyli $\vec{y} = \vec{0}$ i suma podprzestrzeni jest prosta.

Pozostaje pokazać, że wymiar sumy prostej wynosi m. Korzystając z wniosku 5.1 mamy bowiem

$$\begin{array}{rcl} \dim\big(\mathcal{R}(A) \oplus \mathcal{N}(A^H)\big) & = & \dim(\mathcal{R}(A)) + \dim\big(\mathcal{N}(A^H)\big) \\ & = & \dim(\mathcal{R}(A)) + \big[m - \dim\big(\mathcal{R}(A^H)\big)\big] \\ & = & \dim(\mathcal{R}(A)) + [m - \dim(\mathcal{R}(A))] \\ & = & m. \end{array}$$