WYZNACZNIK MACIERZY

Z dużym prawdopodobieństwem w szkole średniej uczyli się Państwo następującej metody rozwiązywania układów dwóch równań z dwiema niewiadomymi:

PRZYKŁAD

$$\begin{cases} 2x + 3y = 7 \\ -x + 2y = 0 \end{cases}$$

WYZNACZNIK GŁÓWNY $W = \begin{vmatrix} 2 & 3 \\ -1 & 2 \end{vmatrix} = 2\cdot 2 - 3(-1) = 4+3 = 7$

Warunkiem stosowalności metody jest $W \neq 0$

$W_x = \begin{vmatrix} 7 & 3 \\ 0 & 2 \end{vmatrix} = 7\cdot 2 - 0\cdot 3 = 14$ $\quad W_y = \begin{vmatrix} 2 & 7 \\ -1 & 0 \end{vmatrix} = 2\cdot 0 - (-1)\cdot 7 = 7$

Rozwiązanie

$x = \frac{W_x}{W} = \frac{14}{7} = 2 \qquad y = \frac{W_y}{W} = \frac{7}{7} = 1$ $\square$

Być może niektórzy uczyli się także rozszerzenia tej metody na przypadek układów trzech równań z trzema niewiadomymi:

PRZYKŁAD

$$\begin{cases} x + 2y - z = 2 \\ x - y + z = 2 \\ -x + y + z = 4 \end{cases}$$

WYZNACZNIK GŁÓWNY

$$W = \left|\begin{matrix} 1 & 2 & -1 \\ 1 & -1 & 1 \\ -1 & 1 & 1 \end{matrix}\right|\begin{matrix} 1 & 2 \\ 1 & -1 \\ -1 & 1 \end{matrix} = 1\cdot(-1)\cdot 1 + 2\cdot 1\cdot(-1) + (-1)\cdot 1\cdot 1 - (-1)(-1)(-1) +$$

$$-1\cdot 1\cdot 1 - 2\cdot 1\cdot 1 = -1 + (-2) + (-1) - (-1) - 1 - 2 =$$

$$= -1-2-1+1-1-2 = -6$$

$W_x = \begin{vmatrix} 2 & 2 & -1 \\ 2 & -1 & 1 \\ 4 & 1 & 1 \end{vmatrix} = -2+8-2-4-4-2 = -6$ $\quad W_y = \begin{vmatrix} 1 & 2 & -1 \\ 1 & 2 & 1 \\ -1 & 4 & 1 \end{vmatrix} = 2-2-4-2-2-4 = -12$

$W_z = \begin{vmatrix} 1 & 2 & 2 \\ 1 & -1 & 2 \\ -1 & 1 & 4 \end{vmatrix} = -4-4+2-2-8-2 = -18$

Rozwiązanie

$x = \frac{W_x}{W} = \frac{-6}{-6} = 1 \qquad y = \frac{W_y}{W} = \frac{-12}{-6} = 2 \qquad z = \frac{W_z}{W} = \frac{-18}{-6} = 3$ $\square$

Uogólnień na więcej niewiadomych raczej nikt nie uczy bo rachunki są bardziej wymagające, wzór na wyznacznik nie poddaje się łatwej „obróbce" obrazkowej. Pora zatem na rozwiązanie problemu w sposób ogólny, dla dowolnej liczby niewiadomych. Tym bardziej, że potrzebne w tym celu pojęcie wyznacznika macierzy kwadratowej ma znaczenie znacznie wykraczające poza konkretny problem rachunkowy jakim jest rozwiązywanie układów równań.

Funkcję $D: \underbrace{\mathbb{K}^n \times \dots \times \mathbb{K}^n}_{n} \longrightarrow \mathbb{K}$ o własnościach:

(1) D jest n-liniowa, tzn dla dowolnego $i \in \overline{1n}$ zachodzi

$$D(v_1, v_2, \dots, \alpha v_i + \beta v_i', v_{i+1}, \dots, v_n) = \alpha D(v_1 \dots v_i \dots v_n) + \beta D(v_1, \dots, v_i' \dots v_n)$$

(2) D jest antysymetryczne, tzn dla dowolnej pary i,j $(i \neq j)$

$$D(v_1, \dots, \underset{\substack{\uparrow \\ \text{miejsce } i}}{v_i}, \dots, \underset{\substack{\uparrow \\ \text{miejsce } j}}{v_j}, \dots, v_n) = -D(v_1, \dots, \underset{\substack{\uparrow \\ \text{miejsce } i}}{v_j}, \dots, \underset{\substack{\uparrow \\ \text{miejsce } j}}{v_i} \dots v_n)$$

(3) D jest unormowane, tzn $D(e_1, e_2, \dots, e_n) = 1$ gdzie e_i jest i-tym elementem bazy standardowej w $\mathbb{K}^n$,

Nazywamy wyznacznikiem. Wyznacznik stosujemy tu do układu n wektorów z $\mathbb{K}^n$, ale o wektorach tych myśleć będziemy jak o kolumnach macierzy i stosować wyznacznik do macierzy $n \times n$ czyli elementów $\mathbb{K}^n_n$.

W dalszym ciągu przekonamy się, że wyznacznik istnieje i że jest jedyny. W tym celu wprowadźmy wygodną notację:

Macierz A ma wyrazy macierzowe a^i_j ← numer wiersza, ← numer kolumny

tzn

$$A = \begin{bmatrix} a^1_1 & a^1_2 & a^1_3 & \cdots & a^1_n \\ a^2_1 & a^2_2 & a^2_3 & \cdots & a^2_n \\ \vdots & & & & \vdots \\ a^n_1 & a^n_2 & a^n_3 & \cdots & a^n_n \end{bmatrix}$$

Symbol a^i oznaczał będzie i-ty wiersz macierzy, tzn

$$a^i = [a^i_1 \ a^i_2 \ \dots \ a^i_n]$$

Symbol a_j to j-ta kolumna macierzy A tzn $a_j = \begin{bmatrix} a^1_j \\ a^2_j \\ \vdots \\ a^n_j \end{bmatrix}$

Mnożąc macierze przez siebie „mnożymy wiersze przez kolumny". Na przykład i-ty wiersz A przez j-tą kolumnę B

$$a^i b_j = \sum_k a^i_k b^k_j$$

i wynik wstawiamy w miejscu $*^i_j$ macierzy AB, tzn

$$AB = \begin{bmatrix} a^1 b_1 & a^1 b_2 & \cdots & a^1 b_n \\ a^2 b_1 & a^2 b_2 & \cdots & a^2 b_n \\ \vdots & & & \vdots \\ a^n b_1 & a^n b_2 & \cdots & a^n b_n \end{bmatrix}$$

samą macierz A możemy zapisać

$A = [a_1 \ a_2 \ \dots \ a_n]$ lub

$$A = \begin{bmatrix} a^1 \\ a^2 \\ \vdots \\ a^n \end{bmatrix}$$

Zauważmy najpierw, że z definicyjnych własności wyznacznika wynikają pewne dalsze własności:

STWIERDZENIE Jeśli D jest wyznacznikiem to (1) $D(0) = 0$,

(2) Jeśli $a_i = a_j$ dla pewnych i,j to $\det A = 0$,

(3) Oznaczamy $v = a_i + \lambda_1 a_1 + \dots + \lambda_{i-1} a_{i-1} + \lambda_{i+1} a_{i+1} + \dots + \lambda_n a_n$, wtedy

$$D(a_1, a_2, \dots, a_{i-1}, v, a_{i+1}, \dots, a_n) = D(a_1, \dots, a_i, \dots, a_n)$$

DOWÓD: Własność (1) wynika z wieloliniowości wyznacznika, własność (2) z antysymetrii, własność (3) z wieloliniowości i antysymetrii. Wszystkie trzy są oczywiste!

STWIERDZENIE: Jeśli D jest wyznacznikiem, prawdziwy jest wzór:

$$D(AB) = D(A) \sum_{\sigma \in S_n} \operatorname{sgn}\sigma \, b_1^{\sigma(1)} b_2^{\sigma(2)} \dots b_n^{\sigma(n)}$$

UWAGA Zanim przeprowadzimy dowód zauważmy, że jeśli powyższe stwierdzenie jest prawdziwe, to otrzymujemy wzór rachunkowy na wyznacznik, gdyż

$$D(A) = D(\mathbb{1} A) = \underbrace{D(\mathbb{1})}_{=1} \sum_{\sigma \in S_n} \operatorname{sgn}\sigma \, a_1^{\sigma(1)} a_2^{\sigma(2)} \dots a_n^{\sigma(n)}$$

D jest więc określona jednoznacznie. Istnienie wykażemy sprawdzając, że powyższy wzór czyni zadość warunkom definicyjnym. Warto więc zainwestować energię w dowodzenie stwierdzenia. Dowód ma naturę rachunkową.

DOWÓD:

$$D(AB) = D\left(\begin{bmatrix} a^1 b_1 & a^1 b_2 & \dots & a^1 b_n \\ a^2 b_1 & a^2 b_2 & \dots & a^2 b_n \\ \vdots & \vdots & & \vdots \\ a^n b_1 & a^n b_2 & \dots & a^n b_n \end{bmatrix}\right) = \sum_k b_1^k D\left(\begin{bmatrix} a^1 e_k & a^1 b_2 & \dots & a^1 b_n \\ a^2 e_k & a^2 b_2 & \dots & a^2 b_n \\ \vdots & \vdots & & \vdots \\ a^n e_k & a^n b_2 & \dots & a^n b_n \end{bmatrix}\right) =$$

$$a^i b_1 = a^i(b_1^1 e_1 + b_1^2 e_2 + \dots + b_1^n e_n) = b_1^1 a^i e_1 + b_1^2 a^i e_1 + \dots + b_1^n a^i e_n$$

$$= \sum_{k,l} b_1^k b_1^l D\left(\begin{bmatrix} a^1 e_k & a^1 e_l & \dots & a^1 b_n \\ a^2 e_k & a^2 e_l & \dots & a^2 b_n \\ \vdots & \vdots & & \vdots \\ a^n e_k & a^n e_l & \dots & a^n b_n \end{bmatrix}\right) =^{*}$$

$$\begin{array}{cc} \uparrow & \uparrow \\ a_k & a_l \end{array}$$

$$a^i e_k = [a_1^i \ a_2^i \ \dots \ a_n^i] \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} \begin{array}{c} \\ \\ \leftarrow k \\ \\ \\ \end{array} = a_k^i$$

$$= \sum_{k,l} b_1^k b_2^l D\left(\begin{bmatrix} & & a^1 b_3 & \dots & a^1 b_n \\ a_k & a_l & a^2 b_3 & \dots & a^2 b_n \\ & & \vdots & & \vdots \\ & & a^n b_3 & \dots & a^n b_n \end{bmatrix}\right) = \text{itd} = \sum_{k_1, k_2, \dots, k_n} b_1^{k_1} b_2^{k_2} \dots b_n^{k_n} D(a_{k_1}, a_{k_2}, \dots, a_{k_n}) =$$

jeśli w zbiorze $\{k_1, \dots, k_n\}$ jakiś indeks powtarza się, to $D(a_{k_1}, \dots, a_{k_n}) = 0$, zatem taki składnik nie daje wkładu do sumy. Pozostają zatem jedynie permutacje

$$= \sum_{\sigma \in S_n} b_1^{\sigma(1)} \dots b_n^{\sigma(n)} D(a_{\sigma(1)}, \dots, a_{\sigma(n)}) = \sum_{\sigma \in S_n} \operatorname{sgn}\sigma \, b_1^{\sigma(1)} \dots b_n^{\sigma(n)} D(A)$$

Wartość D na poprzestawianych wektorach różni się o wartości na uporządkowanych wektorach o znak permutacji, to znaczy

$$D(a_{\sigma(1)}, \dots, a_{\sigma(n)}) = \operatorname{sgn}\sigma \, D(a_1, \dots, a_n) = \operatorname{sgn}\sigma \, D(A) \quad \blacksquare$$

Pozostaje sprawdzić, że wzór $D(A) = \sum_{\sigma \in S_n} \operatorname{sgn}\sigma \, a_1^{\sigma(1)} a_2^{\sigma(2)} \dots a_n^{\sigma(n)}$ określa funkcję spełniającą warunki z definicji wyznacznika:

(1) wieloliniowość jest oczywista $D([a_1, a_2, \dots, \lambda a_i + \beta \tilde{a}_i, \dots a_n]) =$

$$= \sum_{\sigma \in S_n} \operatorname{sgn}\sigma \, a_1^{\sigma(1)} \dots (\lambda a_i^{\sigma(i)} + \mu \tilde{a}_i^{\sigma(i)}) \dots a_n^{\sigma(n)} = \lambda \sum_{\sigma \in S_n} \operatorname{sgn}\sigma \, a_1^{\sigma(1)} \dots a_i^{\sigma(i)} \dots a_n^{\sigma(n)} + \mu \sum_{\sigma \in S_n} \operatorname{sgn}\sigma \, a_1^{\sigma(1)} \dots \tilde{a}_i^{\sigma(i)} \dots a_n^{\sigma(n)} = \lambda D([a_1, \dots, a_i, \dots, a_n]) + \mu D([a_1, \dots, \tilde{a}_i, \dots a_n])$$

(3) $D([e_1, \dots, e_n]) = \sum_{\sigma \in S_n} \operatorname{sgn}\sigma \, e_1^{\sigma(1)} \dots e_n^{\sigma(n)} = \sum_{\sigma \in S_n} \operatorname{sgn}\sigma \, \delta_1^{\sigma(1)} \delta_2^{\sigma(2)} \dots \delta_n^{\sigma(n)} = 1$

(2) $D([a_1, \dots, \overset{i}{a_j}, \dots, \overset{j}{a_i} \dots a_n]) = \sum_{\sigma \in S_n} \operatorname{sgn}\sigma \, a_1^{\sigma(1)} \dots a_j^{\sigma(i)} \dots a_i^{\sigma(j)} \dots a_n^{\sigma(n)} =$

$\sigma' = \sigma \circ (i\, j)$ $\quad = \sum_{\sigma \in S_n} \underbrace{\operatorname{sgn}\sigma}_{(-1)\operatorname{sgn}\sigma'} a_1^{\sigma'(1)} \dots a_i^{\sigma'(i)} \dots a_j^{\sigma'(j)} \dots a_n^{\sigma'(n)} =$

$$= \sum_{\sigma'} \operatorname{sgn}\sigma' (-1) \, a_1^{\sigma'(1)} \dots a_n^{\sigma'(n)} = (-1) D([a_1, \dots, a_i, \dots, a_j \dots a_n])$$

sumować można równie dobrze po σ'

Podsumowując, udało nam się ustalić, że funkcja spełniająca warunki bycia wyznacznikiem istnieje i jest tylko jedna. Znamy także wzór na wyznacznik. Od tego momentu tę jedyną funkcję zapisywać będziemy det. Ze wzoru oraz własności definicyjnych wynikają następujące

WNIOSKI

(1) $\det(AB) = \det(A)\det(B)$ – wynika wprost z dowodu jednego z udowodnionych wcześniej stwierdzeń. Ten wzór jest czasem nazywany wzorem Cauchy'ego na wyznacznik iloczynu.

(2) $\det(A^T) = \det(A)$ – wynika z postaci wzoru rachunkowego. Istotnie:

$$\det(A) = \sum_{\sigma\in S_n} \operatorname{sgn}\sigma \, a_1^{\sigma(1)} a_2^{\sigma(2)} \dots a_n^{\sigma(n)} = \sum_{\sigma\in S_n} \operatorname{sgn}\sigma \, a^1_{\sigma^{-1}(1)} a^2_{\sigma^{-1}(2)} \dots a^n_{\sigma^{-1}(n)} =$$

W ramach każdego składnika porządkujemy czynniki ze względu na górny indeks a nie dolny. Korzystamy z przemienności mnożenia liczb.

$$= \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma^{-1}) a^1_{\sigma^{-1}(1)} \dots a^n_{\sigma^{-1}(n)} = \sum_{\rho\in S_n} \operatorname{sgn}\rho \, a^1_{\rho(1)} \dots a^n_{\rho(n)} =$$

$\operatorname{sgn}\sigma = \operatorname{sgn}\sigma^{-1}$ zmiana indeksu sumowania

$$= \sum_{\rho\in S_n} \operatorname{sgn}\rho \, (a^T)^{\rho(1)}_1 (a^T)^{\rho(2)}_2 \dots (a^T)^{\rho(n)}_n = \det(A^T)$$

(3) $\det(A)=0 \iff rk(A)<n$. Wiemy już, że jeśli $rk(A)<n$, tzn jedna z kolumn jest kombinacją liniową pozostałych, to $\det(A)=0$. Pozostaje wykazać, że jest odwrotnie, tzn dowodzimy

$$\det(A)=0 \implies rk(A)<n$$

a.c. załóżmy że $rk(A)=n$. Oznacza to że układ $(a_1,\dots,a_n)$ jest bazą w K_n. Istnieje więc nieosobliwa macierz zmiany bazy Q taka, że

$e_k = q^i_k a_i$, macierzowo: $\mathbb{1} = Q\cdot A$ zatem $1 = \det(QA) = \det Q \det A$.

Z tego wynika $\det A \neq 0$ (i $\det Q = 0$ oczywiście też).

(4) Nieznikający wyznacznik jest więc kryterium odwracalności macierzy:

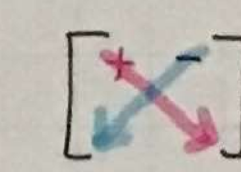
$$(\det A \neq 0) \iff (\exists A^{-1}: AA^{-1} = A^{-1}A = \mathbb{1}),\ \text{ponadto}$$

$$1 = \det\mathbb{1} = \det(A^{-1})\det(A) \implies \det(A^{-1}) = (\det A)^{-1}$$

Sprawdźmy teraz czy „nasz wyznacznik" to to samo co „szkolny wyznacznik" dla macierzy 2×2 i 3×3.

PRZYKŁAD: $n=2$

$$A = \begin{bmatrix} a^1_1 & a^1_2 \\ a^2_1 & a^2_2 \end{bmatrix} \quad S_2 = \{id, (1\ 2)\} \quad \det A = \operatorname{sgn} id\, a^1_1 a^2_2 + \operatorname{sgn}(1\ 2)\, a^2_1 a^1_2 = a^1_1 a^2_2 - a^2_1 a^1_2$$



$n=3$ $\begin{bmatrix} a^1_1 & a^1_2 & a^1_3 \\ a^2_1 & a^2_2 & a^2_3 \\ a^3_1 & a^3_2 & a^3_3 \end{bmatrix}$ $S_3 = \{id, (1\,2\,3), (1\,3\,2), (1\,2), (1\,3), (2\,3)\}$