Rozdział 2

Macierze liczbowe

2.1 Podstawowe definicje

Macierzą (nad ciałem $\mathbf{K}$) nazywamy tablicę prostokątną

$$A = \begin{bmatrix} a_{1,1} & a_{1,2} & \ldots & a_{1,n} \\ a_{2,1} & a_{2,2} & \ldots & a_{2,n} \\ \vdots & \vdots & & \vdots \\ a_{m,1} & a_{m,2} & \ldots & a_{m,n} \end{bmatrix},$$

gdzie $a_{i,j} \in \mathbf{K}$, $1 \le i \le m$, $1 \le j \le n$. Będziemy mówić, że A jest macierzą *formatu* $m \times n$, tzn. macierzą o m *wierszach* i n *kolumnach.* Zbiór wszystkich takich macierzy oznaczamy przez $\mathbf{K}^{m,n}$.

2.1.1 Macierze szczególnych formatów

- $n \times n$ Macierze *kwadratowe* $\mathbf{K}^{n,n}$.

- $m \times 1$ Macierze jednokolumnowe nazywane *wektorami.*
  Zbiór wektorów oznaczamy przez $\mathbf{K}^{m,1} = \mathbf{K}^m$,

$$\mathbf{K}^m \ni A = (a_{i,1}) = \vec{a} = \hat{a} = (a_i)_{i=1}^m = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{bmatrix}.$$

- $1 \times n$ Macierze jednowierszowe nazywane *funkcjonałami.* Zbiór funkcjonałów oznaczamy przez $\mathbf{K}^{1,n} = \mathbf{K}^{nT}$ (albo $\mathbf{K}^{nH}$),
$$\mathbf{K}^{\mathbf{n}T} \ni A = (a_{1,j}) = \vec{a}^T = \hat{a}^T = (a_j)_{j=1}^n = [a_1, \ldots, a_n]\,.$$
- 1×1 Macierze jednoelementowe, utożsamiane z $\mathbf{K}$, tzn. $\mathbf{K}^{1,1} = \mathbf{K}$.

2.1.2 Podział blokowy

Często wygodnie jest przedstawić macierz w postaci blokowej, która w ogólności wygląda następująco:

$$A = \begin{bmatrix} A_{1,1} & \ldots & A_{1,t} \\ \vdots & & \vdots \\ A_{s,1} & \ldots & A_{s,t} \end{bmatrix} \in \mathbf{K}^{m,n}, \tag{2.1}$$

gdzie $A_{p,q} \in \mathbf{K}^{m_p,n_q}$, $1 \le p \le s$, $1 \le q \le t$, $\sum_{p=1}^s m_p = m$, $\sum_{q=1}^t n_q = n$.

Na postać blokową można patrzyć jak na macierz, której elementami są macierze. Z drugiej strony, macierz liczbową można interpretować jako macierz w postaci blokowej z blokami formatu 1×1.

Ważne szczególne przypadki to podział kolumnowy macierzy,

$$A = [\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n]\,, \qquad \text{gdzie} \quad \vec{a}_j = \begin{bmatrix} a_{1,j} \\ a_{2,j} \\ \vdots \\ a_{m,j} \end{bmatrix}, \quad 1 \le j \le n,$$

oraz podział wierszowy macierzy,

$$A = \begin{bmatrix} \hat{a}_1^T \\ \hat{a}_2^T \\ \vdots \\ \hat{a}_m^T \end{bmatrix}, \qquad \text{gdzie} \quad \hat{a}_i^T = [a_{i,1}, a_{i,2}, \ldots, a_{i,n}]\,, \quad 1 \le i \le m.$$

2.2 Działania na macierzach

2.2.1 Podstawowe działania

Możemy na macierzach wykonywać różne działania. Podstawowe z nich to:

$u \in \mathbf{K}, A \in \mathbf{K}^{m,n} \implies B = u * A \in \mathbf{K}^{m,n}$, $b_{i,j} = u * a_{i,j}$
(mnożenie macierzy przez liczbę)

$A, B \in \mathbf{K}^{m,n} \implies C = A + B \in \mathbf{K}^{m,n}$, $c_{i,j} = a_{i,j} + b_{i,j}$
(dodawanie macierzy)

$A \in \mathbf{K}^{m,n} \implies B = A^T \in \mathbf{K}^{n,m}$, $b_{j,i} = a_{i,j}$ (transpozycja macierzy)

$A \in \mathbf{C}^{m,n} \implies B = A^H \in \mathbf{K}^{n,m}$, $b_{j,i} = \overline{a}_{i,j}$ (sprzężenie hermitowskie)

$A \in \mathbf{C}^{m,n} \implies B = |A| \in \mathbf{C}^{m,n}$, $b_{i,j} = |a_{i,j}|$ (moduł macierzy)

W szczególności, mamy też dla $u, v \in \mathbf{K} \subset \mathbf{C}$, $A, B \in \mathbf{C}^{m,n}$,

$$\begin{aligned} &(u * A \pm v * B)^H = \overline{u} * A^H \pm \overline{v} * B^H, \\ &\left(A^T\right)^T = A = \left(A^H\right)^H. \end{aligned}$$

Zauważmy, że macierze formatu $m \times n$ z działaniem dodawania są grupą przemienną, przy czym elementem neutralnym jest macierz zerowa (gdzie $a_{i,j} = 0 \ \forall i, j$), a przeciwną do $(a_{i,j})$ jest macierz $(-a_{i,j})$.

Jeśli macierze dane są w postaci blokowej (2.1) to:

$B = u * A \implies B_{p,q} = u * A_{p,q}$

$C = A + B \implies C_{p,q} = A_{p,q} + B_{p,q}$

$B = A^T \implies B_{p,q} = A^T_{q,p}$

$B = A^H \implies B_{p,q} = A^H_{q,p}$

2.2.2 Mnożenie macierzy

Jeśli $A \in \mathbf{K}^{m,l}$ i $B \in \mathbf{K}^{l,n}$ to

$$C = A * B \in \mathbf{K}^{m,n},$$

gdzie

$$c_{i,j} = \sum_{k=1}^{l} a_{i,k} * b_{k,j}, \qquad 1 \le i \le m, 1 \le j \le n.$$

Zauważmy, że mnożenie $A * B$ jest wykonalne wtedy i tylko wtedy gdy liczba kolumn macierzy A jest równa liczbie wierszy macierzy B. Jeśli A jest w postaci wierszowej, a B kolumnowej,

$$A = \begin{bmatrix} \hat{a}_1^T \\ \vdots \\ \hat{a}_m^T \end{bmatrix}, \qquad B = \left[\vec{b}_1, \ldots, \vec{b}_l\right],$$

to $c_{i,j} = \hat{a}_i^T * \vec{b}_j \; \forall i, j$.

Podstawowe własności mnożenia macierzy są następujące. (Zakładamy, że macierze są odpowiednich formatów tak, że działania są wykonalne.)

$$(A + B) * C = A * C + B * C$$

$$C * (A + B) = C * A + C * B$$
(rozdzielność mnożenia względem dodawania)

$$u * (A * B) = (u * A) * B = A * (u * B) = (A * B) * u \quad (u \in \mathbf{K})$$

$$(A * B) * C = A * (B * C) \quad \text{(łączność mnożenia)}$$

Dowody tych własności polegają na zwykłym sprawdzeniu. Dlatego, dla przykładu, pokażemy tu jedynie łączność. Niech macierze A, B, C będą odpowiednio formatów $m \times k$, $k \times l$, $l \times n$. (Zauważmy, że tylko wtedy odpowiednie mnożenia są wykonalne.) Mamy

$$\begin{aligned} ((A * B) * C)_{i,j} &= \sum_{s=1}^{l} (A * B)_{i,s} c_{s,j} = \sum_{s=1}^{l} \left(\sum_{t=1}^{k} a_{i,t} b_{t,s} \right) c_{s,j} \\ &= \sum_{t=1}^{k} a_{i,t} \sum_{s=1}^{l} b_{t,s} c_{s,j} = \sum_{t=1}^{k} a_{i,t} (B * C)_{t,j} \\ &= (A * (B * C))_{i,j} . \end{aligned}$$

Mamy też

$$(A * B)^T = B^T * A^T, \qquad (A * B)^H = B^H * A^H.$$

Rzeczywiście,

$$\left((A * B)^H\right)_{i,j} = \overline{(A * B)_{j,i}} = \overline{\sum_{k=1}^{l} a_{j,k} b_{k,i}}$$

$$
\begin{aligned}
&= \sum_{k=1}^{l} \overline{a}_{j,k}\overline{b}_{k,i} = \sum_{k=1}^{l} \overline{b}_{k,i}\overline{a}_{j,k} \\
&= \sum_{k=1}^{l} \left(B^H\right)_{i,k} \left(A^H\right)_{k,j} = \left(B^H * A^H\right)_{i,j}.
\end{aligned}
$$

2.2.3 Mnożenie macierzy w postaci blokowej

Jeśli macierze są podane w postaci blokowej to można je mnożyć 'blok-po-bloku' (tak jak w przypadku bloków 1×1) o ile formaty odpowiednich bloków są zgodne. Dokładniej, jeśli $A = (A_{i,k})$, $B = (B_{k,j})$, $1 \le i \le m$, $1 \le k \le l$, $1 \le j \le n$, oraz dla wszystkich i, j, k liczba kolumn bloku $A_{i,k}$ macierzy A jest równa liczbie wierszy bloku $B_{k,j}$ macierzy B to iloczyn

$$C = A * B = (C_{i,j}),$$

$1 \le i \le m$, $1 \le j \le n$, gdzie

$$C_{i,j} = \sum_{k=1}^{l} A_{i,k} * B_{k,n}.$$

Pokażemy to na przykładzie. Niech

$$A = \begin{bmatrix} A_{1,1} & A_{1,2} & A_{1,3} & A_{1,4} \\ A_{2,1} & A_{2,2} & A_{2,3} & A_{2,4} \\ A_{3,1} & A_{3,2} & A_{3,3} & A_{3,4} \end{bmatrix}, \qquad B = \begin{bmatrix} B_{1,1} & B_{1,2} \\ B_{2,1} & B_{2,2} \\ B_{3,1} & B_{3,2} \\ B_{4,1} & B_{4,2} \end{bmatrix}.$$

Wtedy

$$C = \begin{bmatrix} C_{1,1} & C_{1,2} \\ C_{2,1} & C_{2,2} \\ C_{3,1} & C_{3,2} \end{bmatrix},$$

gdzie

$$
\begin{aligned}
C_{1,1} &= A_{1,1} * B_{1,1} + A_{1,2} * B_{2,1} + A_{1,3} * B_{3,1} + A_{1,4} * B_{4,1}, \\
C_{1,2} &= A_{1,1} * B_{1,2} + A_{1,2} * B_{2,2} + A_{1,3} * B_{3,2} + A_{1,4} * B_{4,2}, \\
C_{2,1} &= A_{2,1} * B_{1,1} + A_{2,2} * B_{2,1} + A_{2,3} * B_{3,1} + A_{2,4} * B_{4,1}, \\
C_{2,2} &= A_{2,1} * B_{1,2} + A_{2,2} * B_{2,2} + A_{2,3} * B_{3,2} + A_{2,4} * B_{4,2}, \\
C_{3,1} &= A_{3,1} * B_{1,1} + A_{3,2} * B_{2,1} + A_{3,3} * B_{3,1} + A_{3,4} * B_{4,1}, \\
C_{3,2} &= A_{3,1} * B_{1,2} + A_{3,2} * B_{2,2} + A_{3,3} * B_{3,2} + A_{3,4} * B_{4,2},
\end{aligned}
$$

o ile formaty bloków $A_{i,k}$ i $B_{k,j}$ są zgodnie.

Bardzo ważnym przypadkiem szczególnym mnożenia blokowego jest

$$\begin{aligned} A * B &= A * \left[\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_l\right] \\ &= \left[A * \vec{b}_1, A * \vec{b}_2, \ldots, A * \vec{b}_l\right]. \end{aligned}$$

Zwróćmy jeszcze uwagę na fakt, że jeśli $\vec{a} \in \mathbf{K}^m$ oraz $\vec{b} \in \mathbf{K}^n$ to

$$C = \vec{a} * \vec{b}^T \in \mathbf{K}^{m,n}$$

jest macierzą formatu $m \times n$, nazywaną *iloczynem wewnętrznym* wektorów. Jeśli natomiast wektory są tych samych formatów, $\vec{a}, \vec{b} \in \mathbf{K}^n$, to

$$c = \vec{a}^T * \vec{b} = \vec{b}^T * \vec{a} \in \mathbf{K}$$

jest liczbą, nazywaną *iloczynem zewnętrznym.* W przypadku $\vec{a}, \vec{b} \in \mathbf{C}^n$ definiujemy również *iloczyn skalarny* wektorów jako liczbę zespoloną

$$g = \vec{b}^H * \vec{a} \in \mathbf{C}.$$

2.3 Dalsze oznaczenia

2.3.1 Macierze trójkątne i jednostkowe

Wyróżnimy następujące podzbiory macierzy formatu $m \times n$ (niekoniecznie kwadratowych):

$$\begin{aligned} \mathrm{TRIU}^{m,n} &= \{\, A \in \mathbf{K}^{m,n} \,:\, \forall i > j \ a_{i,j} = 0 \,\}, \\ \mathrm{TRIL}^{m,n} &= \{\, A \in \mathbf{K}^{m,n} \,:\, \forall i < j \ a_{i,j} = 0 \,\}, \\ \mathrm{DIAG}^{m,n} &= \{\, A \in \mathbf{K}^{m,n} \,:\, \forall i \neq j \ a_{i,j} = 0 \,\}. \end{aligned}$$

Są to odpowiednio macierze *trójkątne górne*, *trójkątne dolne* i *diagonalne*. Zauważmy, że każdy z tych podzbiorów macierzy stanowi grupę ze względu na działanie dodawania macierzy (są to *podgrupy* $\{\mathbf{K}^{m,n}, +\}$), oraz

$$\mathrm{DIAG}^{m,n} = \mathrm{TRIU}^{m,n} \cap \mathrm{TRIL}^{m,n}.$$

Ponieważ macierze diagonalne $D \in \mathrm{DIAG}^{m,n}$ mają elementy niezerowe jedynie na głównej diagonali, powiedzmy d_i, $1 \le i \le \min(m,n)$, będziemy pisać

$$D = \mathrm{diag}\left(d_1, d_2, \ldots, d_{\min(m,n)}\right).$$

Szczególnie ważnymi macierzami diagonalnymi są (kwadratowe) macierze *jednostkowe*

$$I_n = \mathrm{diag}_n(\underbrace{1, 1, \ldots, 1}_{n}) \in \mathbf{K}^{n,n}.$$

Jeśli $A \in \mathbf{K}^{m,n}$ to

$$I_m * A \;=\; A \;=\; A * I_n,$$

co oznacza, że I_m i I_n są elementami neutralnymi mnożenia (odpowiednio lewostronnym i prawostronnym).

2.3.2 Układ równań jako równanie macierzowe

Rozpatrzmy następujący układ równań:

$$\left\{ \begin{array}{ccccccccc} a_{1,1}x_1 & + & a_{1,2}x_2 & + & \cdots & + & a_{1,n}x_n & = & b_1 \\ a_{2,1}x_1 & + & a_{2,2}x_2 & + & \cdots & + & a_{2,n}x_n & = & b_2 \\ \vdots & & \vdots & & & & \vdots & & \vdots \\ a_{m,1}x_1 & + & a_{m,2}x_2 & + & \cdots & + & a_{m,n}x_n & = & b_m \end{array} \right. . \tag{2.2}$$

Mówimy, że jest to układ m równań *liniowych* z n niewiadomymi. Liczby $a_{i,j} \in \mathbf{K}$ nazywamy i *współczynnikami* układu, b_i wyrazami *wolnymi*, a x_j to *niewiadome*.

Oznaczmy

$$A = (a_{i,j}) \in \mathbf{K}^{m,n}, \quad \vec{b} = (b_i) \in \mathbf{K}^m, \quad \vec{x} = (x_j) \in \mathbf{K}^n.$$

Wtedy układ (2.2) możemy równoważnie zapisać po prostu jako równanie macierzowe

$$A * \vec{x} = \vec{b}.$$

2.4 Macierze nieosobliwe

2.4.1 Grupa macierzy nieosobliwych

W zbiorze $\mathbf{K}^{n,n}$ mnożenie macierzy jest działaniem wewnętrznym. Ponadto, jak wcześniej zauważyliśmy, mnożenie jest łączne, a macierz jednostkowa

$I_n = \mathrm{diag}(1, \ldots, 1) \in \mathbf{K}^{n,n}$ jest elementem neutralnym mnożenia,

$$\forall A \in \mathbf{K}^{n,n} \quad A * I_n = A = I_n * A.$$

(Przypomnijmy, że element neutralny, jeśli istnieje, jest tylko jeden.) Naturalnym staje się teraz pytanie, czy istnieją elementy odwrotne. Niestety, nie zawsze. Na przykład, łatwo sprawdzić, że (niezerowa) macierz

$$\begin{bmatrix} 1 & -2 \\ -2 & 4 \end{bmatrix}$$

nie ma odwrotności (zarówno lewostronnej jak i prawostronnej). Z drugiej strony, wiele macierzy niezerowych mają odwrotności. Na przykład, macierze

$$A = \begin{bmatrix} 1 & 0 \\ -1 & 2 \end{bmatrix} \quad \text{oraz} \quad B = \begin{bmatrix} 1 & 0 \\ 1/2 & 1/2 \end{bmatrix}$$

stanowią parę macierzy do siebie wzajemnie odwrotnych, $A*B = I_2 = B*A$, tak, że możemy napisać $B = A^{-1}$ i $A = B^{-1}$. (Przypomnijmy, że element odwrotny, jeśli istnieje, jest wyznaczony jednoznacznie.)

Definicja 2.1 *Macierz kwadratową $A \in \mathbf{K}^{n,n}$ dla której istnieje macierz odwrotna $A^{-1} \in \mathbf{K}^{n,n}$ nazywamy* odwracalną *albo* nieosobliwą. *Macierz, która nie posiada macierzy odwrotnej nazywamy* osobliwą.

Zwróćmy uwagę na fakt, że pojęcie macierzy (nie)osobliwej przysługuje jedynie macierzy kwadratowej.

Iloczyn macierzy nieosobliwych jest macierzą nieosobliwą. Rzeczywiście, jeśli $A, B \in \mathbf{K}^{n,n}$ to sprawdzamy bezpośrednio, że odwrotnością $C = A * B$ jest macierz

$$C^{-1} = B^{-1} * A^{-1}.$$

Stąd wniosek, że

ZBIÓR MACIERZY NIEOSOBLIWYCH FORMATU $n \times n$ Z DZIAŁANIEM MNOŻENIA MACIERZY JEST GRUPĄ (NIEPRZEMIENNĄ).

2.4.2 Warunek nieosobliwości macierzy

Twierdzenie 2.1 *Aby macierz $A \in \mathbf{K}^{n,n}$ była nieosobliwa potrzeba i wystarcza, aby dla każdego $\vec{b} \in \mathbf{K}^n$ układ równań $A * \vec{x} = \vec{b}$ miał jednoznaczne rozwiązanie $\vec{x} \in \mathbf{K}^n$.*

Dowód. (Konieczność.) Jeśli A jest nieosobliwa to łatwo sprawdzić, że $\vec{x} = A^{-1} * \vec{b}$ jest rozwiązaniem. Z drugiej strony, jeśli $\vec{x}$ jest rozwiązaniem, $A * \vec{x} = \vec{b}$, to $A^{-1} * (A * \vec{x}) = A^{-1} * \vec{b}$, czyli $\vec{x} = A^{-1} * \vec{b}$ jest rozwiązaniem jednoznacznym.

(Dostateczność.) Układy równań $A*\vec{b}_j = \vec{e}_j$, gdzie $\vec{e}_j$ jest j-tym *wersorem*,

$$\vec{e}_j = [0, \ldots, 0, 1, 0, \ldots, 0]^T,$$

(gdzie jedynka stoi na j-tym miejscu) mają jednoznaczne rozwiązania $\vec{b}_j$, $1 \le j \le n$. Biorąc $B = [\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_n]$ mamy

$$A * B = [A * \vec{b}_1, \ldots, A * \vec{b}_n] = [\vec{e}_1, \ldots, \vec{e}_n] = I_n.$$

Pozostaje jeszcze pokazać, że $B*A = I_n$. Rzeczywiście, mamy $(A*B)*A = A$, czyli $A * (B * A) = A$. Rozwiązaniem równania $A * Z = A$ jest $Z = I_n$, a ponieważ z założenia rozwiązanie to jest jednoznaczne to $B * A = I_n$. Stąd $B = A^{-1}$, co kończy dowód.

Jednym z ważnych wniosków z tego twierdzenie jest następujący.

Wniosek 2.1 *Macierz trójkątna (górna lub dolna) $T \in \mathbf{K}^{n,n}$ jest nieosobliwa wtedy i tylko wtedy gdy wszystkie elementy na głównej diagonali są niezerowe.*

Rzeczywiście, wystarczy sprawdzić jednoznaczną rozwiązywalność odpowiedniego układu równań. Dodajmy, że macierz odwrotna do trójkątnej dolnej (górnej), jeśli istnieje, jest też trójkątna dolna (górna).

2.4.3 Permutacje

Niech $p = [p(1), p(2), \ldots, p(n)] \in \mathrm{Perm}(n)$ będzie *permutacją* n elementową. Odpowiadającą tej permutacji macierz $P = (p_{i,j}) \in \mathbf{K}^{n,n}$ zdefiniowaną jako

$$p_{i,j} = \begin{cases} 1 & \text{gdy } j = p(i), \\ 0 & \text{gdy } j \neq p(i), \end{cases}$$

nazywamy *macierzą permutacji.* Na przykład, jeśli $p = [3, 1, 4, 2] \in \mathrm{Perm}(4)$ to

$$P = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{bmatrix}.$$

Równoważnie, macierz kwadratowa P jest macierzą permutacji wtedy i tylko wtedy gdy w każdym wierszu i w każdej kolumnie występuje dokładnie jedna jedynka, a pozostałe elementy są zerami.

Łatwo sprawdzić, że permutacje n-elementowe $\mathrm{Perm}(n)$ tworzą grupę ze względu na ich złożenie,

$$(q \circ p)(i) = q(p(i)) \qquad 1 \le i \le n.$$

Elementem neutralnym jest permutacja identycznościowa $\mathrm{id}(i) = i \ \forall i$, a elementem odwrotnym do p jest permutacja odwrotna p' zdefiniowana równością $p'(p(i)) = i \ \forall i$.

Podobnie, macierze permutacji tworzą grupę ze względu na mnożenie macierzy, przy czym

$$P(q \circ p) = P(p) * P(q).$$

Rzeczywiście, $(P(q \circ p))_{i,j} = 1$ w.t.w. gdy $q(p(i)) = j$. Z drugiej strony, $(P(p) * P(q))_{i,j} = 1$ w.t.w gdy $(P(q))_{p(i),j} = 1$, czyli znów $q(p(i)) = j$.

Podobnie pokazujemy, że

$$P(p') = (P(p))^{-1} = (P(p))^T.$$

Zauważmy jeszcze, że jeśli $P = P(p)$, $p \in \mathrm{Perm}(n)$, to

$$P * \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} x_{p(1)} \\ \vdots \\ x_{p(n)} \end{bmatrix},$$

czyli mnożenie wektora z lewej strony przez macierz permutacji skutkuje zamianą kolejności współrzędnych. Podobnie,

$$P * \begin{bmatrix} \hat{a}_1^T \\ \vdots \\ \hat{a}_n^T \end{bmatrix} = \begin{bmatrix} \hat{a}_{p(1)}^T \\ \vdots \\ \hat{a}_{p(n)}^T \end{bmatrix}$$

powoduje przestawienie wierszy macierzy zgodnie z p. Ponieważ

$$A * P = \left((A * P)^T\right)^T = \left(P^T * A^T\right)^T,$$

dochodzimy do wniosku, że

$A * P$ permutuje kolumny A zgodnie z p',

$A * P^T$ permutuje kolumny A zgodnie z p.