Maciej Grzesiak
Instytut Matematyki Politechniki Poznańskiej

Macierze. Układy równań.

1. Macierze

Jeżeli każdej uporządkowanej parze liczb naturalnych (i, j), $1 \leqslant i \leqslant m$, $1 \leqslant j \leqslant n$ jest przyporządkowana dokładnie jedna liczba a_{ij}, to mówimy, że jest określona *macierz prostokątna* $A = [a_{ij}]$ typu $m \times n$. Macierz zapisujemy w postaci tablicy:

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \multicolumn{4}{c}{\dotfill} \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}.$$

Rzędy poziome tej tablicy nazywamy *wierszami*, a rzędy pionowe — *kolumnami*. Tak więc powyższa macierz ma m wierszy i n kolumn. Gdy $m = n$ macierz nazywamy *kwadratową*.
O elementach a_{ii} macierzy kwadratowej $\mathbf{A}$ mówimy, że tworzą *przekątną główną*.
Macierz, w której $a_{ij} = 0$ dla $i < j$ (odpowiednio: $a_{ij} = 0$ dla $i > j$) nazywamy *dolnotrójkątną* (odpowiednio: *górnotrójkątną*). Jeśli $a_{ij} = 0$ dla $i \neq j$, to macierz nazywamy *diagonalną*.
Jeśli w macierzy $\mathbf{A}$ zamienimy wiersze z kolumnami, to otrzymamy macierz, którą nazywamy macierzą *transponowaną* macierzy $\mathbf{A}$ i oznaczamy $\mathbf{A}^T$. Tak więc, jeśli $A = [a_{ij}]$ jest typu $m \times n$, to $\mathbf{A}^T = [a_{ji}]$ jest typu $n \times m$.

Przykład.

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 3 & 2 \\ 2 & 3 & 0 & 4 \\ 0 & 5 & -1 & 5 \end{bmatrix}, \qquad \mathbf{A}^T = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 3 & 5 \\ 3 & 0 & -1 \\ 2 & 4 & 5 \end{bmatrix}.$$

Określimy teraz działania na macierzach.
Sumę $\mathbf{A} + \mathbf{B}$ dwóch m-wierszowych i n-kolumnowych macierzy $\mathbf{A} = [a_{ij}]$ i $\mathbf{B} = [b_{ij}]$ otrzymujemy po dodaniu odpowiadających sobie wyrazów:

$$\mathbf{A} + \mathbf{B} = [a_{ij} + b_{ij}].$$

Działanie to jest łączne, przemienne, ma element neutralny $\mathbf{O}$ ($\mathbf{O}$ oznacza macierz, której wszystkie elementy są zerami):

$$\mathbf{A} + \mathbf{O} = \mathbf{O} + \mathbf{A} = \mathbf{A},$$

oraz dla każdej macierzy $\mathbf{A}$ istnieje element odwrotny względem dodawania.
Iloczyn $c\mathbf{A}$ macierzy $\mathbf{A}$ przez skalar $c \in \mathbb{R}$ określamy jako macierz $[ca_{ij}]$. Oczywiście

$$1 \cdot \mathbf{A} = \mathbf{A}, \; c(d\mathbf{A}) = (cd)\mathbf{A}, \; (c+d)\mathbf{A} = c\mathbf{A} + d\mathbf{A}, \; c(\mathbf{A} + \mathbf{B}) = c\mathbf{A} + c\mathbf{B}.$$

Ciekawszym działaniem jest mnożenie macierzy.

Definicja 1. *Niech* $\mathbf{A}(m \times n)$ *i* $\mathbf{B}(n \times p)$ *będą macierzami.* Iloczynem $\mathbf{AB}$ *nazywamy macierz* $\mathbf{C}(m \times p)$ *taką, że* $c_{ik} = \sum_{j=1}^{n} a_{ij} b_{jk}$.

Jak widać warunkiem istnienia iloczynu $\mathbf{AB}$ jest, by macierz $\mathbf{A}$ miała tyle kolumn, ile macierz $\mathbf{B}$ ma wierszy. W praktycznych rachunkach jest wygodny tzw. *schemat Falka*:

$$\begin{array}{c|c} & \mathbf{B} \\ \hline \mathbf{A} & \mathbf{AB} \end{array}\ .$$

Element c_{ij} macierzy $\mathbf{C} = \mathbf{AB}$ otrzymujemy na przecięciu linii wyznaczonych przez i-ty wiersz macierzy $\mathbf{A}$ i j-tą kolumnę macierzy $\mathbf{B}$.

Przykład.

Obliczymy iloczyn $\mathbf{AB}$ macierzy

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 3 & 2 \\ 2 & 3 & 0 & 4 \end{bmatrix} \qquad \text{i} \qquad \mathbf{B} = \begin{bmatrix} 2 & 1 \\ -1 & 2 \\ 3 & -2 \\ 0 & 1 \end{bmatrix}$$

Po zastosowaniu schematu Falka otrzymujemy:

$$\begin{array}{cccc|cc} & & & & 2 & 1 \\ & & & & -1 & 2 \\ & & & & 3 & -2 \\ & & & & 0 & 1 \\ \hline 1 & 0 & 3 & 2 & 11 & -3 \\ 2 & 3 & 0 & 4 & 1 & 12 \end{array}\ .$$

Obliczmy także iloczyn $\mathbf{BA}$:

$$\begin{array}{cc|cccc} & & 1 & 0 & 3 & 2 \\ & & 2 & 3 & 0 & 4 \\ \hline 2 & 1 & 4 & 3 & 6 & 8 \\ -1 & 2 & 3 & 6 & -3 & 6 \\ 3 & -2 & -1 & -6 & 9 & -2 \\ 0 & 1 & 2 & 3 & 0 & 4 \end{array}\ .$$

Przy mnożeniu macierzy specjalną rolę odgrywa tzw. *macierz jednostkowa*, dla której używamy oznaczenia $\mathbf{I}$ i określamy następujaco:

$$\mathbf{I} = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 \end{bmatrix}.$$

Wymienimy teraz własności iloczynu macierzy.

1) Mnożenie macierzy jest łączne, tj. $(\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC})$.

2) Dla macierzy $\mathbf{A}$ typu $(m \times n)$ i macierzy $\mathbf{I}$ stopnia n mamy $\mathbf{AI} = \mathbf{A}$.

3) Mnożenie macierzy jest rozdzielne względem dodawania, tj.

$$(\mathbf{A} + \mathbf{B}) \cdot \mathbf{C} = \mathbf{AC} + \mathbf{BC}, \qquad \mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC}.$$

4) $(a\mathbf{A}) \cdot \mathbf{B} = a(\mathbf{AB})$, $a(b(\mathbf{A})) = (ab)(\mathbf{A})$.

5) Iloczyn macierzy transponujemy według wzoru

$$(\mathbf{AB})^T = \mathbf{B}^T \mathbf{A}^T.$$

U w a g a. Mnożenie macierzy na ogół nie jest przemienne. Wystarczy spojrzeć na przykład wyżej: iloczyny $\mathbf{AB}$ i $\mathbf{BA}$ są innego typu. Nawet jeśli $\mathbf{AB}$ i $\mathbf{BA}$ są tego samego typu (co ma miejsce tylko wtedy, gdy macierze $\mathbf{A}$ i $\mathbf{B}$ są kwadratowe), to na ogół są różne.

Jeżeli $\mathbf{A}$ jest macierzą kwadratową, to można określić potęgę

$$\mathbf{A}^0 = \mathbf{I}, \qquad \mathbf{A}^n = \mathbf{A}^{n-1} \cdot \mathbf{A},\ n \in \mathbb{N}.$$

Następnie dla dowolnego wielomianu $f(x) = a^n x^n + a^{n-1}x^{n-1} + \cdots + a_1 x + a_0$ można obliczać jego wartość na macierzy $\mathbf{A}$ jako

$$f(\mathbf{A}) = a^n \mathbf{A}^n + a^{n-1}\mathbf{A}^{n-1} + \cdots + a_1\mathbf{A} + a_0\mathbf{I}.$$

Np. dla wielomianu $f(x) = 2x^2 - 3x + 4$ i macierzy $\mathbf{A} = \begin{bmatrix} 1 & 2 \\ -3 & 0 \end{bmatrix}$ mamy

$$f(\mathbf{A}) = 2\mathbf{A}^2 - 3\mathbf{A} + 4\mathbf{I} = 2\begin{bmatrix} -5 & 2 \\ -3 & -6 \end{bmatrix} - 3\begin{bmatrix} 1 & 2 \\ -3 & 0 \end{bmatrix} + 4\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} -9 & -2 \\ 3 & -8 \end{bmatrix}$$

2. Wyznacznik macierzy

Macierzy kwadratowej można przyporządkować pewną charakterystykę liczbową — *wyznacznik* . Są różne sposoby wprowadzenia tego pojęcia. Następująca definicja jest definicją *rekurencyjną* — jako punkt wyjścia określamy wyznacznik stopnia pierwszego i drugiego, a wyznacznik stopnia n definiujemy, używając wyznacznika stopnia $n - 1$.

Definicja 2. *1) Jeżeli* $\mathbf{A} = [a]$ *jest macierzą stopnia 1, to jej wyznacznikiem nazywamy liczbę* a *i piszemy*

$$\det \mathbf{A} = a \qquad \textit{lub} \qquad |\mathbf{A}| = a.$$

2) Jeżeli

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

jest macierzą stopnia 2, to jej wyznacznikiem nazywamy liczbę

$$\det \mathbf{A} = a_{11}a_{22} - a_{12}a_{21}.$$

Piszemy także

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}.$$

3) Jeżeli

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

jest macierzą stopnia n*, to*

$$\det \mathbf{A} = \sum_{k=1}^{n} a_{1k}A_{1k},$$

gdzie $A_{1k} = (-1)^{1+k}M_{1k}$*, a* M_{1k} *oznacza wyznacznik stopnia* $n-1$ *powstały przez skreślenie w macierzy* $\mathbf{A}$ *pierwszego wiersza i* k*-tej kolumny. Piszemy także:*

$$\det \mathbf{A} = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{12} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}.$$

Liczbę M_{1k} nazywamy *podwyznacznikiem* lub *minorem* macierzy $\mathbf{A}$, natomiast A_{1k} — to *dopełnienie algebraiczne* elementu a_{1k} macierzy $\mathbf{A}$. Analogicznie określa się M_{ik} oraz A_{ik}.
Równość nazywamy *rozwinięciem Laplace'a* wyznacznika macierzy $\mathbf{A}$ według pierwszego wiersza.
Nie ma żadnego powodu, aby pierwszy wiersz był uprzywilejowany. Mówi o tym następujące twierdzenie:

Twierdzenie 1. *Dla dowolnego $1 \leqslant i \leqslant n$:*

$$\det \mathbf{A} = \sum_{k=1}^{n} a_{ik} A_{ik}$$

(rozwinięcie Laplace'a według i-tego wiersza) oraz

$$\det \mathbf{A} = \sum_{i=1}^{n} a_{ik} A_{ik}$$

(rozwinięcie Laplace'a według k-tej kolumny).

Przykład.

Wyznacznik:

$$\begin{vmatrix} 1 & 2 & 3 \\ -4 & 5 & 6 \\ -2 & -1 & -1 \end{vmatrix}$$

można rozwinąć według pierwszego wiersza, otrzymując:

$$1 \cdot (-1)^2 \cdot \begin{vmatrix} 5 & 6 \\ -1 & -1 \end{vmatrix} + 2 \cdot (-1)^3 \cdot \begin{vmatrix} -4 & 6 \\ -2 & -1 \end{vmatrix} + 3 \cdot (-1)^4 \cdot \begin{vmatrix} -4 & 5 \\ -2 & -1 \end{vmatrix},$$

albo np. według drugiej kolumny:

$$2 \cdot (-1)^3 \cdot \begin{vmatrix} -4 & 6 \\ -2 & -1 \end{vmatrix} + 5 \cdot (-1)^4 \cdot \begin{vmatrix} 1 & 3 \\ -2 & -1 \end{vmatrix} + (-1) \cdot (-1)^5 \cdot \begin{vmatrix} 1 & 3 \\ -4 & 6 \end{vmatrix}.$$

Inny przykład:

$$\begin{vmatrix} 0 & 1 & -1 & 2 \\ 1 & 3 & 5 & 2 \\ 0 & 0 & 0 & 2 \\ -1 & -2 & 0 & 3 \end{vmatrix} = 2 \cdot (-1)^7 \begin{vmatrix} 0 & 1 & -1 \\ 1 & 3 & 5 \\ -1 & -2 & 0 \end{vmatrix}$$

(rozwinięcie Laplace'a według trzeciego wiersza).
Ostatni przykład pokazuje, że korzystnie jest rozwijać wyznacznik według wiersza (kolumny) z dużą liczbą zer.

3. Własności wyznaczników

Następujące własności, wynikające z definicji, pozwalają na zmniejszenie liczby rachunków potrzebnych do obliczenia wyznacznika.
1) Wyznacznik macierzy dolnotrójkątnej (górnotrójkątnej, diagonalnej) jest iloczynem elementów przekątnej głównej.
2) Wyznacznik macierzy $\mathbf{A}$ równy jest wyznacznikowi macierzy $\mathbf{A}^T$,

$$\det \mathbf{A} = \det \mathbf{A}^T.$$

3) Jeżeli macierz $\mathbf{A}$ ma wiersz (kolumnę) złożony z samych zer, to $\det \mathbf{A} = 0$.
4) Zamiana dwóch wierszy (kolumn) macierzy zmienia znak jej wyznacznika.
5) Jeżeli dwa wiersze (dwie kolumny) wyznacznika macierzy $\mathbf{A}$ są równe, to $\det \mathbf{A} = 0$.
6) Wspólny czynnik wszystkich elementów jednego wiersza (jednej kolumny) można wynieść przed znak wyznacznika.
7) Jeżeli dwa wiersze (dwie kolumny) wyznacznika macierzy $\mathbf{A}$ są proporcjonalne, to $\det \mathbf{A} = 0$.

8) Niech wszystkie elementy i-tego wiersza (j-tej kolumny) wyznacznika będą sumami dwóch składników. Wówczas wyznacznik jest równy sumie wyznaczników:

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \multicolumn{4}{c}{\dotfill} \\ a_{i1}+a_{i1}^* & a_{i2}+a_{i2}^* & \cdots & a_{in}+a_{in}^* \\ \multicolumn{4}{c}{\dotfill} \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} =$$

$$= \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \multicolumn{4}{c}{\dotfill} \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \multicolumn{4}{c}{\dotfill} \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \multicolumn{4}{c}{\dotfill} \\ a_{i1}^* & a_{i2}^* & \cdots & a_{in}^* \\ \multicolumn{4}{c}{\dotfill} \\ a_{n1} & a_{n2} & \cdots & a_{nn}. \end{vmatrix}$$

9) Jeżeli do elementów jednego wiersza (jednej kolumny) wyznacznika dodamy odpowiednie elementy innego wiersza (innej kolumny) pomnożone przez dowolną liczbę, to wartość wyznacznika nie ulegnie zmianie.

Przykład.

Obliczyć wyznacznik macierzy

$$\mathbf{A} = \begin{bmatrix} 6 & -5 & 8 & 4 \\ 9 & 7 & 5 & 2 \\ 7 & 5 & 3 & 7 \\ -4 & 8 & -8 & -3 \end{bmatrix}.$$

Skorzystamy z własności 9: od wiersza pierwszego odejmiemy podwojony drugi, od trzeciego — potrojony drugi, a do czwartego dodamy drugi:

$$\det \mathbf{A} = \begin{vmatrix} 6 & -5 & 8 & 4 \\ 9 & 7 & 5 & 2 \\ 7 & 5 & 3 & 7 \\ -4 & 8 & -8 & -3 \end{vmatrix} = \begin{vmatrix} -12 & -19 & -2 & 0 \\ 9 & 7 & 5 & 2 \\ -20 & -16 & -12 & 1 \\ 5 & 15 & -3 & -1 \end{vmatrix} =$$

(teraz do drugiego wiersza dodamy podwojony czwarty i do trzeciego dodamy czwarty; potem rozwiniemy według czwartej kolumny)

$$= \begin{vmatrix} -12 & -19 & -2 & 0 \\ 19 & 37 & -1 & 0 \\ -15 & -1 & -15 & 0 \\ 5 & 15 & -3 & -1 \end{vmatrix} = (-1)\cdot(-1)^8 \begin{vmatrix} -12 & -19 & -2 \\ 19 & 37 & -1 \\ -15 & -1 & -15 \end{vmatrix} =$$

(od pierwszej kolumny odejmiemy trzecią; od trzeciej — 15 razy drugą; rozwiniemy według drugiej)

$$= - \begin{vmatrix} -10 & -19 & 283 \\ 20 & 37 & -556 \\ 0 & -1 & 0 \end{vmatrix} = - \begin{vmatrix} -10 & 283 \\ 20 & -556 \end{vmatrix} = - \begin{vmatrix} -10 & 283 \\ 0 & 10 \end{vmatrix} = 100.$$

Odnotujmy jeszcze, że wyznacznik stopnia trzeciego można obliczać bezpośrednio (bez rozwijania). Mianowicie, po rozwinięciu według dowolnego wiersza i podstawieniu wartości wyznaczników stopnia drugiego otrzymamy wzór:

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = \begin{matrix} a_{11}a_{22}a_{33} & + & a_{12}a_{23}a_{31} & + & a_{13}a_{21}a_{32} & - \\ & - & a_{31}a_{22}a_{13} & - & a_{32}a_{23}a_{11} & - & a_{33}a_{21}a_{12} \end{matrix}$$

Sposób mnemotechniczny zapamiętania tego wzoru nazywa się *schematem Sarrusa.* Dopisujemy z prawej strony jeszcze raz pierwszą i drugą kolumnę, i następnie tworzymy iloczyny: w kierunku przekątnej głównej ze znakami plus, w kierunku drugiej przekątnej ze znakami minus.

4. Układy Cramera

Równanie postaci:

$$a_1x_1 + a_2x_2 + \ldots + a_nx_n = b$$

nazywamy *równaniem liniowym o n niewiadomych* $x_1, x_2, \ldots, x_n$. Ciąg n liczb $(s_1, s_2, \ldots, s_n)$ nazywa się *rozwiązaniem* tego równania, jeśli jest $a_1s_1 + a_2s_2 + \ldots + a_ns_n = b$.
Skończony zbiór równań liniowych o niewiadomych $x_1, x_2, \ldots, x_n$ nazywa się *układem równań liniowych.* Ciąg n liczb $(s_1, s_2, \ldots, s_n)$ nazywa się *rozwiązaniem układu*, jeśli jest rozwiązaniem każdego równania tego układu. Układ, który nie ma rozwiązań nazywamy *sprzecznym.*
Na razie zbadamy przypadek układu o tej samej liczbie równań i niewiadomych. Układ n równań o n niewiadomych

$$\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n & = & b_2 \\ \multicolumn{9}{c}{\dotfill} \\ a_{n1}x_1 & + & a_{n2}x_2 & + & \cdots & + & a_{nn}x_n & = & b_n \end{array} \tag{1}$$

nazywa się *układem Cramera* , jeśli

$$\det \mathbf{A} = \det[a_{ij}] \neq 0.$$

Macierz $\mathbf{A}$ nazywamy *macierzą układu* , a $\det \mathbf{A}$ — *wyznacznikiem układu* . Macierz $\mathbf{A}$ spełniającą warunek $\det \mathbf{A} \neq 0$ nazywamy *nieosobliwą.*

Twierdzenie 2. (Cramera) *Układ Cramera ma dokładnie jedno rozwiązanie. Jest one dane wzorem:*

$$x_k = \frac{\det \mathbf{A}_k}{\det \mathbf{A}} \qquad (k = 1, 2, \ldots, n) \tag{2}$$

gdzie macierz $\mathbf{A}_k$ *powstaje z macierzy* $\mathbf{A}$ *przez zastąpienie* k*-tej kolumny kolumną utworzoną z wyrazów* $b_1, b_2, \ldots, b_n$*.*

Przykład.
Rozwiązać układ:

$$\begin{array}{rcrcrcrcr} x_1 & + & 2x_2 & + & 3x_3 & + & 4x_4 & = & 5 \\ 2x_1 & + & x_2 & + & 2x_3 & + & 3x_4 & = & 1 \\ 3x_1 & + & 2x_2 & + & x_3 & + & 2x_4 & = & 1 \\ 4x_1 & + & 3x_2 & + & 2x_3 & + & 4x_1 & = & -5 \end{array} .$$

Obliczamy kolejno wyznaczniki:

$$\det \mathbf{A} = \begin{vmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 2 & 3 \\ 3 & 2 & 1 & 2 \\ 4 & 3 & 2 & 1 \end{vmatrix} = -20,$$

$$\det \mathbf{A}_1 = \begin{vmatrix} 5 & 2 & 3 & 4 \\ 1 & 1 & 2 & 3 \\ 1 & 2 & 1 & 2 \\ -5 & 3 & 2 & 1 \end{vmatrix} = 40, \qquad \det \mathbf{A}_2 = \begin{vmatrix} 1 & 5 & 3 & 4 \\ 2 & 1 & 2 & 3 \\ 3 & 1 & 1 & 2 \\ 4 & -5 & 2 & 1 \end{vmatrix} = -40,$$

$$\det \mathbf{A}_3 = \begin{vmatrix} 1 & 2 & 5 & 4 \\ 2 & 1 & 1 & 3 \\ 3 & 2 & 1 & 2 \\ 4 & 3 & -5 & 1 \end{vmatrix} = 60, \qquad \det \mathbf{A}_4 = \begin{vmatrix} 1 & 2 & 3 & 5 \\ 2 & 1 & 2 & 1 \\ 3 & 2 & 1 & 1 \\ 4 & 3 & 2 & -5 \end{vmatrix} = -60.$$

Zatem

$$x_1 = \frac{\det \mathbf{A}_1}{\det \mathbf{A}} = -2, x_2 = \frac{\det \mathbf{A}_2}{\det \mathbf{A}} = 2, x_3 = \frac{\det \mathbf{A}_3}{\det \mathbf{A}} = -3, x_4 = \frac{\det \mathbf{A}_4}{\det \mathbf{A}} = 3.$$

Układ (1) dla $b_1 = b_2 = \ldots = b_n = 0$ nazywamy układem *jednorodnym.* Jeżeli jest on układem Cramera, to ma tylko rozwiązanie zerowe (bo $\det \mathbf{A}_k = 0$ dla $k = 1, 2, \ldots, n$).

Wniosek 1. *Układ jednorodny ma rozwiązanie niezerowe wtedy i tylko wtedy, gdy* $\det \mathbf{A} = 0$.

Wówczas rozwiązań jest nieskończenie wiele. Sposób opisu zbioru rozwiązań podamy później.

5. Układy równań i ich macierze

Omówiliśmy już układy Cramera. Teraz zajmiemy się przypadkiem ogólniejszym. Przedmiotem badań będzie układ m równań liniowych o n niewiadomych $x_1, x_2, \ldots, x_n$:

$$\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n & = & b_2 \\ \multicolumn{9}{c}{\dotfill} \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n & = & b_m. \end{array} \tag{3}$$

Równania te można zapisać krócej:

$$\sum_{j=1}^{n} a_{ij}x_j = b_i \qquad (i = 1, \ldots, m).$$

Współczynniki układu są liczbami rzeczywistymi lub zespolonymi.
Jak poprzednio, ciąg n liczb $(s_1, s_2, \ldots, s_n)$ nazywamy rozwiązaniem układu, jeśli jest rozwiązaniem każdego równania tego układu. Układ, który nie ma rozwiązań nazywamy sprzecznym.
Omówimy dwie metody rozwiązywania takiego układu. Pierwszą będzie *metoda eliminacji Gaussa.*
Zaczniemy od przykładu. Symbolem R_i oznaczamy w nim i-te równanie ($i = 1, 2, 3, 4$), a np. zapis $(R_2 - 2R_1) \to R_2$ oznacza, że od drugiego równania odjęliśmy stronami równanie pierwsze pomnożone przez dwa i wynik zapisaliśmy jako nowe drugie równanie.

Przykład.
Rozwiążemy układ równań:

$$\begin{array}{ccrcrcrcrcr} R_1 & : & x_1 & + & x_2 & & & + & 3x_4 & = & 4 \\ R_2 & : & 2x_1 & + & x_2 & - & x_3 & + & x_4 & = & 1 \\ R_3 & : & 3x_1 & - & x_2 & - & x_3 & + & 2x_4 & = & -3 \\ R_4 & : & -x_1 & + & 2x_2 & + & 3x_3 & - & x_4 & = & 4 \end{array}$$

Użyjemy równania R_1 do eliminacji x_1 z równań R_2, R_3, R_4. Wykonujemy operacje: $(R_2 - 2R_1) \to R_2$, $(R_3 - 3R_1) \to R_3$, $(R_4 + R_1) \to R_4$. Otrzymujemy:

$$\begin{array}{ccrcrcrcrcr} R_1 & : & x_1 & + & x_2 & & & + & 3x_4 & = & 4 \\ R_2 & : & & - & x_2 & - & x_3 & - & 5x_4 & = & -7 \\ R_3 & : & & - & 4x_2 & - & x_3 & - & 7x_4 & = & -15 \\ R_4 & : & & & 3x_2 & + & 3x_3 & + & 2x_4 & = & 8 \end{array}$$

Następnie użyjemy R_2 do eliminacji x_2 z R_3 i R_4 : $(R_3 - 4R_2) \to R_3$, $(R_4 + 3R_2) \to R_4$. Otrzymujemy:

$$\begin{array}{ccrcrcrcrcr} R_1 & : & x_1 & + & x_2 & & & + & 3x_4 & = & 4 \\ R_2 & : & & - & x_2 & - & x_3 & - & 5x_4 & = & -7 \\ R_3 & : & & & & & 3x_3 & + & 13_4 & = & 13 \\ R_4 & : & & & & & & - & 13x_4 & = & -13 \end{array}$$

Układ jest teraz w postaci *trójkątnej* i można go łatwo rozwiązać przez *podstawienie wsteczne.* Ponieważ z R_4 mamy $x_4 = 1$, zatem z R_3 możemy obliczyć x_3:

$$x_3 = \frac{1}{3}(13 - 13x_4) = 0.$$

Dalej R_2 daje:
$$x_2 = -(-7 + 5x_4 + x_3) = 2,$$
a R_1 daje:
$$x_1 = 4 - 3x_4 - x_2 = -1.$$
Zatem rozwiązaniem układu jest czwórka liczb:
$$x_1 = -1, \qquad x_2 = 2, \qquad x_3 = 0, \qquad x_4 = 1.$$
Układ powyższy ma więc jednoznaczne rozwiązanie. Można je znaleźć także, stosując wzory Cramera (2). Wymagałoby to obliczenia pięciu wyznaczników stopnia czwartego. Metoda eliminacji ma widoczną przewagę — rachunki są krótsze. Więcej, metoda ta jest ogólniejsza. Można ją stosować wtedy, gdy $m \neq n$ (liczba równań jest różna od liczby niewiadomych) oraz wtedy, gdy wprawdzie $m = n$, ale wyznacznik układu jest równy zero.
Wprowadzimy pewną terminologię. Macierz
$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \multicolumn{4}{c}{\dotfill} \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$
zbudowaną ze współczynników przy niewiadomych układu (3) nazywamy *macierzą układu*, a macierz
$$\mathbf{B} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \multicolumn{5}{c}{\dotfill} \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{bmatrix}$$
poszerzoną o kolumnę wyrazów wolnych układu nazywamy *macierzą uzupełnioną*.
Dwa układy równań nazywamy *równoważnymi*, gdy mają taki sam zbiór rozwiązań. Jest oczywiste, że następujące typy operacji na układzie:
1) przestawienie dowolnych dwóch równań,
2) pomnożenie równania przez stałą $c \neq 0$, $c \in \mathbb{K}$,
3) dodanie wielokrotności jednego równania do innego równania
przekształcają układ w układ mu równoważny. Te przekształcenia nazywamy *elementarnymi operacjami na równaniach.*
Powyższym operacjom na równaniach odpowiadają *elementarne operacje na wierszach* macierzy układu:
1) przestawienie dowolnych dwóch wierszy,
2) pomnożenie wiersza przez stałą $c \neq 0$, $c \in \mathbb{K}$,
3) dodanie wielokrotności jednego wiersza do innego wiersza.
Niech $\mathbf{A}$ będzie macierzą prostokątną. Pierwszy niezerowy element wiersza nazywamy *elementem wiodącym (kierunkowym)* tego wiersza. Kolumny zawierające element wiodący nazywamy także wiodącymi. Jeżeli dla elementów wiodących kolejnych wierszy a_{ij}, a_{i+1k} jest spełniony warunek $j < k$, to macierz nazywamy *macierzą schodkową.*

Przykład.
Macierz
$$\mathbf{A} = \begin{bmatrix} 4 & 2 & 5 & 0 & 0 & 1 \\ 0 & 3 & 1 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 4 & 5 \\ 0 & 0 & 0 & 0 & 0 & 9 \end{bmatrix}$$
jest macierzą schodkową. Elementami wiodącymi wierszy są kolejno 4,3,4,9. Kolumny: pierwsza, druga, piąta i szósta są wiodące; trzecia i czwarta — niewiodące.

6. Metoda eliminacji Gaussa

Kolejne etapy rozwiązywania układu równań (3) metodą eliminacji Gaussa są następujące:
1) Tworzymy macierz uzupełnioną układu. Dla zaznaczenia specjalnej roli ostatniej kolumny oddzielamy ją kreską.
2) Jeśli $a_{11} \neq 0$, to dzielimy pierwszy wiersz przez a_{11} (wtedy wyraz wiodący wynosi 1) i posługując się tym wierszem, uzyskamy zera w pierwszej kolumnie — od wiersza drugiego odejmujemy wiersz pierwszy pomnożony przez a_{21} itd. Gdyby $a_{11} = 0$, a np. $a_{k1} \neq 0$, to przestawiamy najpierw wiersz pierwszy z k-tym — i dalej jak poprzednio.
3) Jeśli $a_{22} \neq 0$, to dzielimy drugi wiersz przez a_{22} i posługując się tym wierszem, uzyskamy zera w drugiej kolumnie poniżej a_{22}. Jeśli $a_{22} = 0$, a np. $a_{k2} \neq 0$, to przestawiamy wiersz drugi z k-tym — i dalej jak wyżej. Jeśli wszystkie $a_{k2} = 0$ dla $k = 2, 3, \ldots, m$, to przechodzimy do następnej kolumny.
4) Postępowanie kontynuujemy aż do n-tej kolumny.
Wynikiem tego postępowania jest macierz schodkowa, w której każdy element wiodący wynosi 1. Macierz ta pozwala na proste znalezienie rozwiązania. Są możliwe trzy przypadki:
1) Jeżeli w otrzymanej macierzy występuje wiersz $(0\,0\,\ldots\,0|1)$, to układ jest sprzeczny (taki wiersz odpowiada równaniu $0 = 1$).
2) Nie ma wierszy postaci $(0\,0\,\ldots\,0|1)$ i liczba niezerowych wierszy jest równa liczbie niewiadomych, tzn. macierz jest postaci:

$$\left[\begin{array}{ccccc|c} 1 & * & * & * & * & b_1 \\ 0 & 1 & * & * & * & b_2 \\ \multicolumn{6}{c}{\dotfill} \\ 0 & 0 & 0 & 0 & 1 & b_n \\ 0 & 0 & 0 & 0 & 0 & 0 \\ & & \cdots & & & \end{array}\right],$$

gdzie * oznacza jakiś element. Odpowiada to układowi:

$$\begin{array}{ccccccccc} x_1 & + & *x_2 & + & \ldots & + & *x_n & = & b_1 \\ & & x_2 & + & \ldots & + & *x_n & = & b_2 \\ & & & & \multicolumn{5}{c}{\dotfill} \\ & & & & x_{n-1} & + & *x_n & = & b_{n-1} \\ & & & & & & x_n & = & b_n \end{array}$$

Stąd $x_n = b_n$. Po podstawieniu do równania $(n-1)$-szego obliczamy x_{n-1} itd. Rozwiązanie jest tylko jedno.
3) W macierzy nie ma wierszy postaci $(0\,0\,\ldots\,0|1)$, ale występują kolumny niewiodące. Stosujemy, jak wyżej, podstawienie wsteczne, ale niewiadomym, które odpowiadają kolumnom niewiodącym nadajemy dowolne wartości (będą one parametrami rozwiązania).
Przykład.
Przypuśćmy, że po eliminacji uzyskaliśmy macierz:

$$\mathbf{A} = \left[\begin{array}{cccc|c} 1 & 2 & 3 & 0 & 5 \\ 0 & 1 & 2 & 2 & 1 \\ 0 & 0 & 0 & 1 & 2 \end{array}\right] \tag{4}$$

Niewiadoma x_3 jest niewiodąca. Zatem $x_4 = 2$, $x_3 = s$ (gdzie $s \in \mathbb{R}$), $x_2 = 1 - 2x_3 - 2x_4 = -3 - 2s$, $x_1 = 5 - 2x_2 - 3x_3 = 11 + s$.
Rozwiązań jest nieskończenie wiele. Są one postaci:

$$x_1 = 11 + s, \qquad x_2 = -3 - 2s, \qquad x_3 = s, \qquad x_4 = 2, \qquad (s \in \mathbb{R}).$$

7. Uwagi o eliminacji

Wróćmy do macierzy (4). Po wykonaniu na jej wierszach operacji: $w_2 - 2w_3$ i $w_1 - 2w_2$ uzyskamy zera *nad* elementami wiodącymi:

$$\left[\begin{array}{cccc|c} 1 & 2 & 3 & 0 & 5 \\ 0 & 1 & 2 & 2 & 1 \\ 0 & 0 & 0 & 1 & 2 \end{array}\right] \sim \left[\begin{array}{cccc|r} 1 & 2 & 3 & 0 & 5 \\ 0 & 1 & 2 & 0 & -3 \\ 0 & 0 & 0 & 1 & 2 \end{array}\right] \sim \left[\begin{array}{ccrc|r} 1 & 0 & -1 & 0 & 11 \\ 0 & 1 & 2 & 0 & -3 \\ 0 & 0 & 0 & 1 & 2 \end{array}\right].$$

Teraz możemy po prostu odczytać rozwiązanie (dla $x_3 = s$):

$$x_1 = 11 + s, \qquad x_2 = -3 - 2s, \qquad x_3 = s, \qquad x_4 = 2, \qquad (s \in \mathbb{R}).$$

Tę metodę "przedłużonej" eliminacji nazywa się *metodą Gaussa–Jordana*.
Na podstawie metody eliminacji opracowano metody rozwiązywania układów za pomocą komputerów. Algorytm nie jest, jak widać, skomplikowany. Problemem pozostaje efektywność. Często bardziej praktyczne okazują się metody iteracyjne, szczególnie wtedy, gdy dla danych współczynników szukamy rozwiązania z określoną dokładnością.

8. Macierz odwrotna

Definicja 3. *Niech* $\mathbf{A}$ *będzie macierzą kwadratową. Macierz* $\mathbf{A}^{-1}$ *spełniającą warunki*

$$\mathbf{A}\mathbf{A}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}. \tag{5}$$

nazywamy macierzą odwrotną *względem* $\mathbf{A}$.

Macierz odwrotna nie zawsze istnieje. Aby to uzasadnić skorzystamy z następującego twierdzenia:

Twierdzenie 3. (Cauchy'ego[1]) *Dla dowolnych macierzy kwadratowych* $\mathbf{A}$ *i* $\mathbf{B}$ *tego samego stopnia*

$$\det(\mathbf{AB}) = \det\mathbf{A} \cdot \det\mathbf{B}.$$

Po zastosowaniu twierdzenia Cauchy'ego do równości (5) otrzymujemy:

$$\det\mathbf{A} \cdot \det\mathbf{A}^{-1} = \det(\mathbf{A}\mathbf{A}^{-1}) = \det\mathbf{I} = 1.$$

Wnosimy stąd, że macierz mająca macierz odwrotną musi mieć wyznacznik różny od zera.

Definicja 4. *Macierz* $\mathbf{A}$ *nazywamy nieosobliwą gdy* $\det\mathbf{A} \neq 0$.

Z twierdzenia Cauchy'ego wynika także następujący wniosek: *wyznacznik macierzy odwrotnej jest równy odwrotności wyznacznika macierzy danej.*
Macierz odwrotna iloczynu dwóch macierzy nieosobliwych jest równa iloczynowi macierzy odwrotnych tych macierzy wziętych w odwrotnej kolejności:

$$(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$$

i analogicznie dla dowolnej skończonej liczby czynników.
Macierz odwrotną możemy obliczyć dwoma sposobami.
Sposób 1. Zastosować wzór:

$$\mathbf{A}^{-1} = \frac{1}{\det\mathbf{A}} \begin{bmatrix} A_{11} & A_{21} & \dots & A_{n1} \\ A_{12} & A_{22} & \dots & A_{n2} \\ & & \dots & \\ A_{1n} & A_{2n} & \dots & A_{nn} \end{bmatrix}. \tag{6}$$

[1] Louis Augustin Cauchy (1789–1857).

Wzór ten łatwo sprawdzić, obliczając $\mathbf{AA}^{-1}$, bo element c_{ij} tego iloczynu jest postaci:

$$c_{ij} = \frac{1}{\det \mathbf{A}} \sum_{k=1}^{n} a_{ik} A_{kj},$$

oraz z definicji wyznacznika:

$$\sum_{k=1}^{n} a_{ik} A_{kj} = \left\{ \begin{array}{ccc} 0 & \text{dla} & i \neq j \\ \det \mathbf{A} & \text{dla} & i = j \end{array} \right. .$$

Praktycznie wygląda to tak: obliczamy wyznacznik (musi być niezerowy), tworzymy macierz minorów $[M_{ij}]$, zmieniamy znaki odpowiednich elementów, tworząc macierz dopełnień algebraicznych $[A_{ij}]$, tę macierz transponujemy — wynikiem jest tzw. *macierz dołączona* $\mathbf{A}^D = [A_{ji}]$, wreszcie dzielimy ją przez wyznacznik.

Przykład.
Znaleźć macierz odwrotną do macierzy

$$\mathbf{A} = \left[\begin{array}{cc} 1 & 3 \\ 4 & 5 \end{array} \right].$$

Obliczamy $\det \mathbf{A} = -7$ i następnie

$$[M_{ij}] = \left[\begin{array}{cc} 5 & 4 \\ 3 & 1 \end{array} \right], \qquad [A_{ij}] = \left[\begin{array}{rr} 5 & -4 \\ -3 & 1 \end{array} \right], \qquad \mathbf{A}^D = \left[\begin{array}{rr} 5 & -3 \\ -4 & 1 \end{array} \right].$$

Zatem

$$\mathbf{A}^{-1} = -\frac{1}{7} \left[\begin{array}{rr} 5 & -3 \\ -4 & 1 \end{array} \right].$$

Sposób 2. Znalezienie macierzy odwrotnej do macierzy $\mathbf{A} = [a_{ij}]$ polega na obliczeniu macierzy $\mathbf{X} = [x_{ij}]$ takiej, że $\mathbf{AX} = \mathbf{I}$. Dla prostoty zapisu przeprowadzimy rachunki dla macierzy stopnia drugiego. Zatem

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Stąd

$$\begin{bmatrix} a_{11}x_{11} + a_{12}x_{21} & a_{11}x_{12} + a_{12}x_{22} \\ a_{21}x_{11} + a_{22}x_{21} & a_{21}x_{12} + a_{21}x_{22} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Porównując kolumny macierzy otrzymujemy dwa układy równań, każdy z dwiema niewiadomymi:

$$\left\{ \begin{array}{rcl} a_{11}x_{11} + a_{12}x_{21} & = & 1 \\ a_{21}x_{11} + a_{22}x_{21} & = & 0 \end{array} \right. \qquad \left\{ \begin{array}{rcl} a_{11}x_{12} + a_{12}x_{22} & = & 0 \\ a_{21}x_{12} + a_{21}x_{22} & = & 1 \end{array} \right.$$

Macierze uzupełnione tych układów to:

$$\begin{bmatrix} a_{11} & a_{12} & 1 \\ a_{21} & a_{22} & 0 \end{bmatrix} \qquad \begin{bmatrix} a_{11} & a_{12} & 0 \\ a_{21} & a_{22} & 1 \end{bmatrix}$$

Po wykonaniu eliminacji Gaussa – Jordana otrzymamy:

$$\begin{bmatrix} 1 & 0 & x_{11} \\ 0 & 1 & x_{21} \end{bmatrix} \qquad \begin{bmatrix} 1 & 0 & x_{12} \\ 0 & 1 & x_{22} \end{bmatrix}$$

W obu układach współczynniki przy niewiadomych są takie same (tworzą macierz $\mathbf{A}$), więc zamiast przeprowadzać eliminację na obu macierzach osobno, można ją przeprowadzić na macierzy

$$\begin{bmatrix} a_{11} & a_{12} & 1 & 0 \\ a_{21} & a_{22} & 0 & 1 \end{bmatrix},$$

której lewą część stanowi macierz $\mathbf{A}$, a prawą macierz $\mathbf{I}$. Jeśli dokonamy eliminacji Gaussa – Jordana, to uzyskamy

$$\begin{bmatrix} 1 & 0 & x_{11} & x_{12} \\ 0 & 1 & x_{21} & x_{22} \end{bmatrix}$$

czyli po prawej otrzymamy macierz $\mathbf{I}$, a po lewej macierz odwrotną $\mathbf{A}^{-1}$.
Przypomnijmy, że *elementarne operacje na wierszach macierzy* to:
1) przestawienie dowolnych dwóch wierszy,
2) pomnożenie wiersza przez stałą $c \neq 0$, $c \in \mathbb{K}$,
3) dodanie wielokrotności jednego wiersza do innego wiersza.
Przeprowadzone rozumowanie jest uzasadnieniem następującego twierdzenia:

Twierdzenie 4. *Jeżeli macierz* $\mathbf{I}$ *otrzymujemy przez operacje elementarne na wierszach z macierzy* $\mathbf{A}$*, to macierz* $\mathbf{A}^{-1}$ *powstaje z macierzy* $\mathbf{I}$ *w wyniku wykonania tych samych operacji elementarnych.*

Przykład.
Znajdziemy odwrotność macierzy z poprzedniego przykładu.
Zapisujemy macierze $\mathbf{A}$ i $\mathbf{I}$ obok siebie; kolejne etapy przekształcenia łączymy znakiem równoważności $\sim$:

$$\left[\begin{array}{cc|cc} 1 & 3 & 1 & 0 \\ 4 & 5 & 0 & 1 \end{array}\right] \sim \left[\begin{array}{cc|cc} 1 & 3 & 1 & 0 \\ 0 & -7 & -4 & 1 \end{array}\right] \sim \left[\begin{array}{cc|cc} 1 & 0 & -\frac{5}{7} & \frac{3}{7} \\ 0 & 1 & \frac{4}{7} & -\frac{1}{7} \end{array}\right].$$

Zadanie. Wyznaczyć macierze odwrotne do danych:

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 2 \\ 2 & 1 & -2 \\ 2 & -2 & 1 \end{bmatrix}, \ \mathbf{B} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}.$$

9. Równania macierzowe

Wygodnym sposobem zapisu układu równań:

$$\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n & = & b_2 \\ \multicolumn{9}{c}{\dotfill} \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n & = & b_m \end{array}$$

jest zapis macierzowy:

$$\mathbf{AX} = \mathbf{B}, \tag{7}$$

gdzie $\mathbf{A}$ jest macierzą układu, $\mathbf{X}$ — jednokolumnową macierzą niewiadomych, a $\mathbf{B}$ — jednokolumnową macierzą wyrazów wolnych.
Ogólniej, rozważmy równanie postaci (7), w którym $\mathbf{X}$ i $\mathbf{B}$ nie muszą być jednokolumnowe — o macierzach występujących w tym równaniu zakładamy tylko, że ich wymiary są takie, że równanie ma sens.
Jeżeli w równaniu (7) macierz $\mathbf{A}$ jest kwadratowa i nieosobliwa, to mnożąc to równanie z lewej strony przez $\mathbf{A}^{-1}$, otrzymamy:

$$\mathbf{A}^{-1}\mathbf{AX} = \mathbf{A}^{-1}\mathbf{B},$$

czyli

$$\mathbf{X} = \mathbf{A}^{-1}\mathbf{B}.$$

Analogicznie z równania:

$$\mathbf{XA} = \mathbf{B}$$

otrzymamy (mnożąc równanie z prawej strony przez $\mathbf{A}^{-1}$):

$$\mathbf{X} = \mathbf{BA}^{-1}.$$

Przykład.
Rozwiązać równanie $\mathbf{AX} + \mathbf{B} = \mathbf{C}$, gdzie:

$$\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix}, \ \mathbf{B} = \begin{bmatrix} 1 & -2 \\ 2 & -5 \end{bmatrix}, \ \mathbf{C} = \begin{bmatrix} -1 & 3 \\ -2 & 3 \end{bmatrix}.$$

Szukana macierz $\mathbf{X}$ musi być kwadratowa stopnia 2 oraz:

$$\mathbf{AX} = \mathbf{C} - \mathbf{B},$$

$$\mathbf{X} = \mathbf{A}^{-1}(\mathbf{C} - \mathbf{B}).$$

Obliczamy:

$$\mathbf{C} - \mathbf{B} = \begin{bmatrix} -2 & 5 \\ -4 & 8 \end{bmatrix}, \ \mathbf{A}^{-1} = \begin{bmatrix} 3 & -2 \\ -1 & 1 \end{bmatrix}.$$

Zatem:

$$\mathbf{X} = \begin{bmatrix} 2 & -1 \\ -2 & 3 \end{bmatrix}.$$

Układy równań macierzowych rozwiązujemy metodą podstawiania. Przy przekształceniach i podstawianiu należy pamiętać, że mnożenie macierzy nie jest przemienne.
Przykład 1. Wyznaczyć macierz $\mathbf{X}$ z równania

$$\begin{bmatrix} 2 & 1 \\ 3 & 2 \end{bmatrix} \mathbf{X} \begin{bmatrix} -3 & 2 \\ 5 & -3 \end{bmatrix} = \begin{bmatrix} -2 & 4 \\ 3 & -1 \end{bmatrix}.$$

Przykład 2. Rozwiązać układ:

$$\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \mathbf{X} + \begin{bmatrix} 3 & 1 \\ 1 & 1 \end{bmatrix} \mathbf{Y} = \begin{bmatrix} 3 & 5 \\ 1 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \mathbf{X} + \begin{bmatrix} 1 & 1 \\ 1 & 3 \end{bmatrix} \mathbf{Y} = \begin{bmatrix} 1 & 1 \\ 5 & 3 \end{bmatrix}.$$

10. Rząd macierzy

Definicja 5. Macierz $\mathbf{B}$ nazywamy *wierszowo równoważną* macierzy $\mathbf{A}$, jeżeli $\mathbf{B}$ można otrzymać z $\mathbf{A}$ przez zastosowanie skończonej liczby operacji elementarnych na wierszach.

Każdą macierz można za pomocą operacji elementarnych sprowadzić do postaci schodkowej.

Definicja 6. *Rząd macierzy* $\mathbf{A}$ jest równy liczbie niezerowych wierszy w postaci schodkowej tej macierzy.

Przykład. Obliczymy rząd macierzy:

$$\mathbf{A} = \begin{bmatrix} 1 & -1 & 1 & 3 \\ 2 & -5 & 3 & 10 \\ 3 & 3 & 1 & 1 \end{bmatrix}.$$

Aby przekształcić macierz do postaci schodkowej, wykonujemy operacje: $w_2 - 2w_1$ i $w_3 - 3w_1$, a następnie $w_3 + 2w_2$:

$$\begin{bmatrix} 1 & -1 & 1 & 3 \\ 2 & -5 & 3 & 10 \\ 3 & 3 & 1 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & -1 & 1 & 3 \\ 0 & -3 & 1 & 4 \\ 0 & 6 & -2 & -8 \end{bmatrix} \sim \begin{bmatrix} 1 & -1 & 1 & 3 \\ 0 & -3 & 1 & 4 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Są dwa wiersze niezerowe, więc $R(\mathbf{A}) = 2$.
Istnieje ważny związek rzędu z minorami.

Twierdzenie 5. *Jeżeli macierz zawiera minor stopnia r różny od zera, dla którego wszystkie zawierające go minory stopnia $r+1$ (minory obrzeżające) są równe zeru, to rząd tej macierzy jest równy r.*

A zatem rząd macierzy jest równy najwyższemu ze stopni różnych od zera minorów tej macierzy.
Obliczanie rzędu macierzy metodą obrzeżania należy prowadzić od stopni najniższych do najwyższych. Przykładowo, weźmy ponownie macierz

$$\mathbf{A} = \begin{bmatrix} 1 & -1 & 1 & 3 \\ 2 & -5 & 3 & 10 \\ 3 & 3 & 1 & 1 \end{bmatrix}.$$

Minor $|a_{11}| = 1$ jest niezerowy. Minor obrzeżający:

$$\begin{vmatrix} 1 & -1 \\ 2 & -5 \end{vmatrix} = -3$$

jest także niezerowy. Dla niego mamy dwa minory obrzeżające:

$$\begin{vmatrix} 1 & -1 & 1 \\ 2 & -5 & 3 \\ 3 & 3 & 1 \end{vmatrix} = 0 \quad , \quad \begin{vmatrix} 1 & -1 & 3 \\ 2 & -5 & 10 \\ 3 & 3 & 1 \end{vmatrix} = 0,$$

a więc $R(\mathbf{A}) = 2$.
Macierze $\mathbf{A}$ i $\mathbf{A}^T$ mają te same minory, więc mamy poniższy wniosek:

Wniosek 2. $R(\mathbf{A}) = R(\mathbf{A}^T)$.

Przy transponowaniu wiersze stają się kolumnami. Zatem przy obliczaniu rzędu metodą przekształcania macierzy do postaci schodkowej można wykonywać również operacje na kolumnach (co było niedopuszczalne w metodzie eliminacji).

11. Twierdzenie Kroneckera-Capellego

Twierdzenie 6. (Kroneckera–Capellego) *Układ*

$$\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n & = & b_2 \\ \multicolumn{9}{c}{\dotfill} \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n & = & b_m. \end{array} \tag{8}$$

ma rozwiązanie wtedy i tylko wtedy, gdy

$$R(\mathbf{A}) = R(\mathbf{B}).$$

D o w ó d. Oznaczmy:

$$\mathbf{v}_j = \begin{bmatrix} a_{1j} \\ a_{2j} \\ \cdots \\ a_{mj} \end{bmatrix}, \quad \mathbf{w} = \begin{bmatrix} b_1 \\ b_2 \\ .. \\ b_m \end{bmatrix}.$$

Wtedy układ (8) jest równoważny równaniu wektorowemu:

$$x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_n\mathbf{v}_n = \mathbf{w}. \tag{9}$$

Jeżeli układ ma rozwiązanie, to istnieją elementy $x_j \in \mathbb{K}$ spełniające układ (8), a więc i równanie wektorowe (9).
Z równania (9) wynika, że wektor $\mathbf{w}$ jest kombinacją liniową wektorów $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$.
Tym samym $\mathbf{w}$ należy do przestrzeni rozpiętej na wektorach $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$.

Zatem przestrzeń generowana przez zbiór $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ jest taka sama jak przestrzeń generowana przez zbiór $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n, \mathbf{w}\}$. To oznacza, że wymiar jest taki sam, czyli $R(\mathbf{A}) = R(\mathbf{B})$.
Odwrotnie, jeśli $R(\mathbf{A}) = R(\mathbf{B}) = r$, to wśród wektorów $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ jest r liniowo niezależnych — niech to będą $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r$.
Ale $R(\mathbf{B}) = r$, więc wśród wektorów $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n, \mathbf{w}$ jest też tylko r liniowo niezależnych. Muszą to być $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r$.
Pozostałe są od nich liniowo zależne. W szczególności $\mathbf{w}$ jest kombinacją liniową wektorów $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r$, a więc i wektorów $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$. Tym samym istnieją elementy $x_j \in \mathbb{K}$ spełniające równanie (9), a więc i układ (8). ■
Przykład. Stosując twierdzenie Kroneckera–Capellego sprawdzić, czy układ ma rozwiązanie:

$$\begin{cases} x + 2y + 3z + 3u = 6 \\ x + y + z = 1 \\ 3x - y + 2z + u = 2 \end{cases}$$

Ponieważ

$$\mathbf{A} \sim \begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & 2 & 3 & 3 \\ 3 & -1 & 2 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & 2 & 3 \\ 0 & -4 & -1 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 7 & 13 \end{bmatrix}$$

więc $R(\mathbf{A}) = 3$ oraz $R(\mathbf{A}) \leqslant R(\mathbf{B}) \leqslant 3$.
Zatem $R(\mathbf{A}) = R(\mathbf{B})$. Układ ma rozwiązanie.

12. Zastosowania do geometrii

.
Zadanie. Jaki warunek muszą spełniać punkty $M_1 = (x_1, y_1)$, $M_2 = (x_2, y_2)$, $M_3 = (x_3, y_3)$ leżące na jednej prostej?
Jeżeli istnieje prosta $Ax + By + C = 0$ na której leżą wszystkie te punkty, to układ

$$\begin{cases} Ax_1 + By_1 + C = 0 \\ Ax_2 + By_2 + C = 0 \\ Ax_3 + By_3 + C = 0 \end{cases}$$

z niewiadomymi A, B, C ma rozwiązanie niezerowe.
Jest tak wtedy, i tylko wtedy, gdy

$$\begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix} = 0.$$

Zadanie. Sprawdzić, czy punkty $(-2, 1), (1, -1), (7, -5)$ leżą na jednej prostej.
Sprawdzamy wyznacznik:

$$\begin{vmatrix} -2 & 1 & 1 \\ 1 & -1 & 1 \\ 7 & -5 & 1 \end{vmatrix} = 0.$$

Punkty leżą na jednej prostej.
Zadanie. Podać warunek na to, by proste $A_1x+B_1y+C_1 = 0$, $A_2x+B_2y+C_2 = 0$, $A_3x + B_3y + C_3 = 0$ przechodziły przez jeden punkt.
Jeżeli istnieje punkt (x, y) wspólny dla tych prostych, to układ

$$\begin{cases} A_1x + B_1y + C_1 = 0 \\ A_2x + B_2y + C_2 = 0 \\ A_3x + B_3y + C_3 = 0 \end{cases}$$

czyli

$$\begin{cases} A_1x + B_1y = -C_1 \\ A_2x + B_2y = -C_2 \\ A_3x + B_3y = -C_3 \end{cases}$$

z niewiadomymi x, y ma rozwiązanie.
Jest tak wtedy, i tylko wtedy, gdy rzędy macierzy

$$\begin{bmatrix} A_1 & B_1 \\ A_2 & B_2 \\ A_3 & B_3 \end{bmatrix}, \qquad \begin{bmatrix} A_1 & B_1 & -C_1 \\ A_2 & B_2 & -C_2 \\ A_3 & B_3 & -C_3 \end{bmatrix}$$

są równe. A zatem musi być

$$\begin{vmatrix} A_1 & B_1 & -C_1 \\ A_2 & B_2 & -C_2 \\ A_3 & B_3 & -C_3 \end{vmatrix} = 0.$$

Zadanie. Sprawdzić, czy proste $x+12y-4=0$, $x-2y+2=0$, $2x+3y+1=0$ mają punkt wspólny.
Sprawdzamy wyznacznik:

$$\begin{vmatrix} 1 & 12 & -4 \\ 1 & -2 & 2 \\ 2 & 3 & 1 \end{vmatrix} = 0.$$

Proste mają punkt wspólny.

13. Wartości i wektory własne macierzy

Definicja 7. Niech $\mathbf{A} = [a_{ij}]$ będzie macierzą kwadratową. Wektor $\mathbf{v}$, $\mathbf{v} \neq \mathbf{0}$, spełniający związek

$$\mathbf{A}\mathbf{v}^T = \lambda \mathbf{v}^T$$

dla pewnego $\lambda \in \mathbb{K}$ nazywamy *wektorem własnym*, a odpowiadającą liczbę λ — *wartością własną* macierzy $\mathbf{A}$.

Symbol $\mathbf{v}^T$ oznacza macierz jednokolumnową — transpozycję wektora $\mathbf{v}$.

Twierdzenie 7. (o istnieniu wektora własnego) *Każda macierz zespolona* $\mathbf{A} = [a_{ij}]$ *ma przynajmniej jeden wektor własny.*

D o w ó d. Jeżeli $\mathbf{v} = (x_1, x_2, \ldots, x_n)$ jest wektorem własnym, to:

$$\mathbf{A}\mathbf{v}^T = \lambda \mathbf{v}^T$$

Ta równość wektorowa jest równoważna układowi równań:

$$\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & \lambda x_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n & = & \lambda x_2 \\ \multicolumn{9}{c}{\dotfill} \\ a_{n1}x_1 & + & a_{n2}x_2 & + & \cdots & + & a_{nn}x_n & = & \lambda x_n \end{array}, \tag{10}$$

czyli

$$\begin{array}{cccccccc} (a_{11}-\lambda)x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = 0 \\ a_{21}x_1 & + & (a_{22}-\lambda)x_2 & + & \cdots & + & a_{2n}x_n & = 0 \\ \multicolumn{8}{c}{\dotfill} \\ a_{n1}x_1 & + & a_{n2}x_2 & + & \cdots & + & (a_{nn}-\lambda)x_n & = 0 \end{array}. \tag{11}$$

Układ taki ma rozwiązanie niezerowe wtedy i tylko wtedy, gdy liczba λ spełnia warunek:

$$\begin{vmatrix} a_{11}-\lambda & a_{12} & \cdots & a_{1n} \\ a_{12} & a_{22}-\lambda & \cdots & a_{2n} \\ \multicolumn{4}{c}{\dotfill} \\ a_{n1} & a_{n2} & \cdots & a_{nn}-\lambda \end{vmatrix} = 0. \tag{12}$$

Łatwo zauważyć, że lewa strona tej równości jest wielomianem zmiennej λ stopnia n. Na mocy zasadniczego twierdzenia algebry równanie (12) ma co najmniej jeden pierwiastek λ_0. Jeżeli $(x_1, x_2, \ldots, x_n)$ jest jakimkolwiek rozwiązaniem niezerowym układu otrzymanego z układu (11) po podstawieniu $\lambda = \lambda_0$, to wektor $(x_1, x_2, \ldots, x_n)$ jest własny. ■

Równanie (12) nazywamy *równaniem charakterystycznym* przekształcenia f, a wielomian tworzący lewą stronę tego równania —*wielomianem charakterystycznym* macierzy $\mathbf{A}$.

Równanie charakterystyczne można zapisać krótko:

$$\det(\mathbf{A} - \lambda\mathbf{I}) = 0$$

lub

$$|\mathbf{A} - \lambda\mathbf{I}| = 0.$$

Przykład. Obliczyć wartości i wektory własne macierzy:

$$\mathbf{A} = \begin{bmatrix} 1 & 3 & 0 \\ 3 & -2 & -1 \\ 0 & -1 & 1 \end{bmatrix}.$$

Piszemy równanie charakterystyczne:

$$\begin{vmatrix} 1-\lambda & 3 & 0 \\ 3 & -2-\lambda & -1 \\ 0 & -1 & 1-\lambda \end{vmatrix} = 0,$$

czyli

$$-\lambda^3 + 13\lambda - 12 = 0.$$

Pierwiastkami tego równania są -4, 1, 3. Macierz ma więc trzy wartości własne. Obliczymy teraz wektory własne.

Dla $\lambda = -4$ układ (11) przyjmuje postać:

$$\begin{array}{rcrcrcl} 5x & + & 3y & & & = & 0 \\ 3x & + & 2y & - & z & = & 0 \\ & & -y & + & 5z & = & 0 \end{array}.$$

Rozwiązaniem układu są liczby $x = -3k$, $y = 5k$, $z = k$ $(k \in \mathbb{R})$. Zatem wektory własne są postaci $k(-3, 5, 1)$ $(k \in \mathbb{R})$.

Analogicznie, dla $\lambda = 1$ układ (11) przyjmuje postać:

$$\begin{array}{rcrcrcl} & & 3y & & & = & 0 \\ 3x & - & 3y & - & z & = & 0 \\ & & -y & & & = & 0 \end{array}.$$

Znajdujemy: $x = k$, $y = 0$, $z = 3k$ $(k \in \mathbb{R})$. Zatem wektory własne są postaci $k(1, 0, 3)$, $k \in \mathbb{R}$.

Dla $\lambda = 3$ mamy:

$$\begin{array}{rcrcrcl} -2x & + & 3y & & & = & 0 \\ 3x & - & 5y & - & z & = & 0 \\ & & -y & - & 2z & = & 0 \end{array},$$

skąd $x = 3k$, $y = 2k$, $z = -k$ $(k \in \mathbb{R})$. Wektory własne: $k(3, 2, -1)$, $k \in \mathbb{R}$.

Aby wyjaśnić rolę wartości i wektorów własnych wprowadzimy najpierw nowe pojęcie.

Definicja 8. Macierze $\mathbf{A}$ i $\mathbf{B}$ nazywamy *podobnymi*, gdy istnieje taka macierz nieosobliwa $\mathbf{P}$, że $\mathbf{B} = \mathbf{P}^{-1}\mathbf{AP}$.

Twierdzenie 8. *Jeżeli macierze $\mathbf{A}$ i $\mathbf{B}$ są podobne, to mają ten sam wyznacznik, ten sam rząd, ten sam ślad, ten sam wielomian charakterystyczny i te same wartości własne.*

D o w ó d (częściowy). Niech $\mathbf{B} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$ dla pewnej macierzy nieosobliwej $\mathbf{P}$. Wtedy

$$\det \mathbf{B} = \det(\mathbf{P}^{-1}\mathbf{A}\mathbf{P}) = \det \mathbf{P}^{-1} \det \mathbf{A} \det \mathbf{P} = \frac{1}{\det \mathbf{P}} \det \mathbf{A} \det \mathbf{P} = \det \mathbf{A},$$

więc wyznacznik jest ten sam. Sprawdzimy jeszcze równość wielomianów charakterystycznych:

$$\begin{aligned} \det(\mathbf{P}^{-1}\mathbf{A}\mathbf{P} - \lambda\mathbf{I}) &= \det(\mathbf{P}^{-1}\mathbf{A}\mathbf{P} - \mathbf{P}^{-1}\lambda\mathbf{I}\mathbf{P}) = \\ &= \det(\mathbf{P}^{-1}(\mathbf{A} - \lambda\mathbf{I})\mathbf{P}) = \\ &= \det(\mathbf{P}^{-1}) \det(\mathbf{A} - \lambda\mathbf{I}) \det \mathbf{P} = \\ &= \det(\mathbf{A} - \lambda\mathbf{I}). \qquad \blacksquare \end{aligned}$$

To twierdzenie formułuje jedynie pewne warunki *konieczne*, by macierze były podobne.
Wektory własne macierzy $\mathbf{A}$ i $\mathbf{P}^{-1}\mathbf{A}\mathbf{P}$ są na ogół inne. Ale liczba liniowo niezależnych wektorów własnych jest taka sama dla obu macierzy.
Wiele rachunków macierzowych można uprościć znajdując macierz diagonalną podobną do danej macierzy, następnie wykonując (łatwo) rachunki na postaci diagonalnej i na końcu wracając do macierzy pierwotnej. Dokładniej, zauważmy, że

$$\begin{bmatrix} \lambda_1 & 0 & 0 & \dots & 0 \\ 0 & \lambda_2 & 0 & \dots & 0 \\ 0 & 0 & \lambda_3 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \lambda_n \end{bmatrix}^n = \begin{bmatrix} \lambda_1^n & 0 & 0 & \dots & 0 \\ 0 & \lambda_2^n & 0 & \dots & 0 \\ 0 & 0 & \lambda_3^n & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \lambda_n^n \end{bmatrix},$$

z czego dalej wynika, że jeśli $p(t)$ jest wielomianem, a $\mathbf{A}$ jest macierzą diagonalną, to

$$p(\mathbf{A}) = p\left(\begin{bmatrix} \lambda_1 & 0 & 0 & \dots & 0 \\ 0 & \lambda_2 & 0 & \dots & 0 \\ 0 & 0 & \lambda_3 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \lambda_n \end{bmatrix} \right) = \begin{bmatrix} p(\lambda_1) & 0 & 0 & \dots & 0 \\ 0 & p(\lambda_2) & 0 & \dots & 0 \\ 0 & 0 & p(\lambda_3) & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & p(\lambda_n) \end{bmatrix}.$$

Ponadto jeśli $\mathbf{A} = \mathbf{P}^{-1}\mathbf{B}\mathbf{P}$, gdzie $\det \mathbf{P} \neq 0$, to:

$$\mathbf{A}^n = \mathbf{P}^{-1}\mathbf{B}^n\mathbf{P} \quad \text{dla dowolnego } n \in \mathbb{N},$$

$$p(\mathbf{A}) = \mathbf{P}^{-1}p(\mathbf{B})\mathbf{P} \quad \text{dla dowolnego wielomianu } p(t).$$

Przykład. Dla macierzy

$$\mathbf{A} = \begin{bmatrix} 2 & 1 \\ 3 & 4 \end{bmatrix}, \quad \mathbf{P} = \begin{bmatrix} -1 & 1 \\ 1 & 3 \end{bmatrix}$$

mamy

$$\mathbf{P}^{-1} = \frac{1}{4} \begin{bmatrix} -3 & 1 \\ 1 & 1 \end{bmatrix}, \quad \mathbf{A} = \mathbf{P} \begin{bmatrix} 1 & 0 \\ 0 & 5 \end{bmatrix} \mathbf{P}^{-1}$$

Stąd np.

$$\mathbf{A}^{20} = \mathbf{P} \begin{bmatrix} 1 & 0 \\ 0 & 5^{20} \end{bmatrix} \mathbf{P}^{-1} = \frac{1}{4} \begin{bmatrix} 3 + 5^{20} & -1 + 5^{20} \\ -3 + 3 \cdot 5^{20} & 1 + 3 \cdot 5^{20} \end{bmatrix}$$

Przykład pokazuje jak można wykorzystać macierz diagonalną podobną do danej macierzy. Dlatego tak ważne jest następujące twierdzenie:

Twierdzenie 9. *Niech* $\mathbf{A}$ *będzie macierzą mającą* n *liniowo niezależnych wektorów własnych* $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$*. Niech* λ_i *będzie wartością własną odpowiadającą wektorowi własnemu* $\mathbf{v}_i$ *(dla* $i = 1, 2, \dots, n$*), tj.* $\mathbf{A}\mathbf{v}_i^T = \lambda_i \mathbf{v}_i^T$*. Utwórzmy macierz* $\mathbf{P}$*, której kolumnami są wektory* $\mathbf{v}_1^T, \mathbf{v}_2^T, ..., \mathbf{v}_n^T$ *:*

$$\mathbf{P} = [\ \mathbf{v}_1^T \quad \mathbf{v}_2^T \quad ... \quad \mathbf{v}_n^T\ ].$$

Wówczas:

$$\mathbf{P}^{-1}\mathbf{AP} = \begin{bmatrix} \lambda_1 & 0 & 0 & \ldots & 0 \\ 0 & \lambda_2 & 0 & \ldots & 0 \\ 0 & 0 & \lambda_3 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & \lambda_n \end{bmatrix}$$

czyli macierz $\mathbf{A}$ *jest podobna do macierzy diagonalnej.*

Jednak nie wszystkie macierze mają wystarczającą liczbę wektorów własnych (mogą ich wcale nie mieć).
Twierdzenie 7 zapewnia wprawdzie, że każda macierz ma zespoloną wartość własną, ale:

1. nie wynika z tego istnienie n niezależnych liniowo wektorów własnych;
2. w praktyce, gdy macierz jest rzeczywista szukamy tylko rzeczywistych wartości własnych, których może wcale nie być.

Przykład. Macierz

$$\begin{bmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{bmatrix}$$

ma wielomian charakterystyczny

$$c(\lambda) = \lambda^2 - 2\cos\varphi \cdot \lambda + 1$$

dla którego $\Delta = -4\sin^2\varphi$, więc pierwiastek rzeczywisty istnieje tylko gdy $\varphi = k\pi$.
Istnieje jednak klasa macierzy "porządna" pod tym względem.

Twierdzenie 10. *Niech* $\mathbf{A}$ *będzie rzeczywistą macierzą symetryczną (tj.* $\mathbf{A} = \mathbf{A}^T$*). Wtedy wszystkie pierwiastki równania charakterystycznego są rzeczywiste i istnieje* n *wartości własnych (licząc z krotnościami) oraz* n *liniowo niezależnych wektorów własnych* $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$*.*

Przykład. Macierz

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$$

ma wielomian charakterystyczny $c(\lambda) = (\lambda + 1)^2(\lambda - 2)$.
Wektory własne: dla $\lambda = -1$: $(1, -1, 0)^T, (1, 0, -1)^T$; dla $\lambda = 2$: $(1, 1, 1)^T$.
Przykład. Macierz

$$\begin{bmatrix} a & b \\ b & c \end{bmatrix}$$

ma wielomian charakterystyczny

$$c(\lambda) = \lambda^2 - (a + c)\lambda + (ac - b^2).$$

Zatem $\Delta = (a - c)^2 + 4b^2 \geqslant 0$. Gdy $a \neq c$ to $\Delta > 0$ i istnieją dwie wartości własne, a gdy $\Delta = 0$ to $a = c$ i $b = 0$, więc macierz jest diagonalna.

Wniosek 3. *Każda macierz symetryczna jest podobna do macierzy diagonalnej.*

Przykład. Dla wyżej zdefiniowanej macierzy $\mathbf{A}$ tworzymy macierz z jej wektorów własnych:

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

i znajdujemy jej odwrotność:

$$\mathbf{P}^{-1} = \frac{1}{3}\begin{bmatrix} 1 & -2 & 1 \\ 1 & 1 & -2 \\ 1 & 1 & 1 \end{bmatrix}.$$

Wtedy zachodzi równość

$$\mathbf{P}^{-1}\mathbf{AP} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{bmatrix}.$$

Uwaga. Można wykazać, że suma wszystkich wartości własnych macierzy jest równa sumie elementów przekątnej głównej tej macierzy.

Definicja 9. *Śladem* macierzy $\mathbf{A} = [a_{ij}]$ nazywamy sumę elementów jej przekątnej głównej.

Oznaczenie: $\operatorname{tr}\mathbf{A}$. Zatem z definicji:

$$\operatorname{tr}\mathbf{A} = \sum_{i=1}^{n} a_{ii},$$

oraz jak można wykazać $\operatorname{tr}\mathbf{A} = \sum_{i=1}^{n} \lambda_i$.

14. Macierze symetryczne

W tym rozdziale poszerzamy wiedzę o macierzach symetrycznych. [2]
Standardowy iloczyn skalarny w $\mathbb{R}^n$ określamy wzorem

$$\langle \mathbf{u}, \mathbf{v} \rangle = \sum_{i=1}^{n} u_i \cdot v_i,$$

gdzie u_i, v_i oznaczają współrzędne wektorów.
Wygodne będzie posługiwanie się zapisem macierzowym. Wektory $\mathbf{u}, \mathbf{v}$ będziemy traktować jak macierze wierszowe. Wtedy ich iloczyn skalarny określamy wzorem

$$\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u} \cdot \mathbf{v}^T$$

Twierdzenie 11. *Niech $\mathbf{A}$ będzie macierzą symetryczną stopnia n, $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$. Wtedy*

$$\langle \mathbf{A}\mathbf{u}^T, \mathbf{v}^T \rangle = \langle \mathbf{u}^T, \mathbf{A}\mathbf{v}^T \rangle.$$

D o w ó d.

$$\begin{aligned} \langle \mathbf{A}\mathbf{u}^T, \mathbf{v}^T \rangle &= \left(\mathbf{A}\mathbf{u}^T\right)^T \cdot \mathbf{v}^T = \left((\mathbf{u}^T)^T \mathbf{A}^T\right) \cdot \mathbf{v}^T = (\mathbf{u}^T)^T \left(\mathbf{A} \cdot \mathbf{v}^T\right) = \\ &= \langle \mathbf{u}^T, \mathbf{A}\mathbf{v}^T \rangle. \end{aligned}$$

Twierdzenie 12. *Dla dowolnej macierzy symetrycznej $\mathbf{A}$ wartości własne odpowiadające różnym wartościom własnym są ortogonalne.*

D o w ó d. Niech $\mathbf{A}\mathbf{u}^T = \lambda \mathbf{u}^T$, $\mathbf{A}\mathbf{v}^T = \mu \mathbf{u}^T$. Wtedy:

$$\begin{aligned} \lambda \langle \mathbf{u}^T, \mathbf{v}^T \rangle &= \langle \lambda \mathbf{u}^T, \mathbf{v}^T \rangle = \langle \mathbf{A}\mathbf{u}^T, \mathbf{v}^T \rangle = \langle \mathbf{u}^T, \mathbf{A}\mathbf{v}^T \rangle = \\ &= \langle \mathbf{u}^T, \mu \mathbf{v}^T \rangle = \mu \langle \mathbf{u}^T, \mathbf{v}^T \rangle, \end{aligned}$$

więc

$$(\lambda - \mu)\langle \mathbf{u}^T, \mathbf{v}^T \rangle = 0,$$

a stąd

$$\langle \mathbf{u}^T, \mathbf{v}^T \rangle = 0$$

Macierz $\mathbf{A}$ nazywamy *ortogonalnie diagonalizowalną* gdy można znaleźć macierz ortogonalną $\mathbf{P}$ taką, że $\mathbf{P}^{-1}\mathbf{AP}$ jest diagonalna.
Macierze symetryczne są ortogonalnie diagonalizowalne. Wynika to z twierdzeń, które podamy bez dowodów.

[2] Wymagana jest znajomość pojęć: iloczyn skalarny, norma (długość) wektora, macierz ortogonalna.

Twierdzenie 13. *Jeżeli* **A** *jest diagonalna i* λ *jest jej wartością własną, to krotność* λ *jako pierwiastka wielomianu charakterystycznego (krotność algebraiczna) równa się krotności geometrycznej (wymiarowi podprzestrzeni własnej* E_λ*).*

Twierdzenie 14. *Niech* $d_i = \dim(E_{\lambda_i})$ *dla* $i = 1, 2, \ldots, k$*. Następujące warunki są równoważne:*
a) **A** *jest diagonalizowalna,*
b) $\sum_{i=1}^{k} d_i = n$*.*

Twierdzenie 15. (spektralne) *Niech* **A** *będzie macierzą kwadratową stopnia* n*. Następujące warunki są równoważne:*
a) **A** *ma ortonormalny zbiór wektorów własnych,*
b) **A** *jest ortogonalnie diagonalizowalna,*
c) **A** *jest symetryczna.*
Jeśli spełniony jest którykolwiek (a więc wszystkie) z warunków (1)– (3), to algebraiczne krotności różnych wartości własnych są równe krotnościom geometrycznym i suma tych krotności wynosi n*.*

Przykład.
Znaleźć macierz ortogonalną $\mathbf{P}$ taką, że $\mathbf{P}^{-1}\mathbf{A}\mathbf{P}$ jest diagonalna, gdzie

$$\mathbf{A} = \begin{bmatrix} 1 & -2 & 0 \\ -2 & 0 & 2 \\ 0 & 2 & -1 \end{bmatrix}.$$

Mamy

$$c_{\mathbf{A}}(\lambda) = \det \begin{bmatrix} 1-\lambda & -2 & 0 \\ -2 & -\lambda & 2 \\ 0 & 2 & -1-\lambda \end{bmatrix} = \lambda^3 - 9\lambda = \lambda(\lambda-3)(\lambda+3).$$

Wartościami własnymi są $\lambda = 0, 3, -3$. Odpowiednie wektory własne:

$$\mathbf{v}_1 = (2,1,2),\ \mathbf{v}_2 = (-2,2,1),\ \mathbf{v}_3 = (1,2,-2)$$

są ortogonalne. Ponieważ $\|\mathbf{v}_1\| = \|\mathbf{v}_2\| = \|\mathbf{v}_3\| = 3$, więc $\frac{1}{3}\mathbf{v}_1,\ \frac{1}{3}\mathbf{v}_2,\ \frac{1}{3}\mathbf{v}_3$ są wektorami ortonormalnymi. Stąd macierz

$$\mathbf{P} = \frac{1}{3}\begin{bmatrix} 2 & -2 & 1 \\ 1 & 2 & 2 \\ 2 & 1 & -2 \end{bmatrix}$$

jest ortogonalna (czyli $\mathbf{P}^{-1} = \mathbf{P}^T$) oraz

$$\mathbf{P}^T\mathbf{A}\mathbf{P} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & -3 \end{bmatrix}$$

na podstawie algorytmu diagonalizacji.

15. Twierdzenie Cayleya–Hamiltona

Każda macierz spełnia swoje równanie charakterystyczne. Mamy bowiem następujące twierdzenie.

Twierdzenie 16. (Cayleya–Hamiltona) *Jeżeli* **A** *jest macierzą kwadratową, a* $c(\lambda)$ *jest jej wielomianem charakterystycznym, to* $c(\mathbf{A}) = \mathbf{O}$*.*

D o w ó d. Niech $\mathbf{A} = [a_{ij}]$ będzie macierzą stopnia n i niech

$$c(\lambda) = \det(\mathbf{A} - \lambda\mathbf{I}) = c_n\lambda^n + c_{n-1}\lambda^{n-1} + \cdots + c_1\lambda + c_0.$$

Dopełnienia algebraiczne macierzy $\mathbf{A}-\lambda\mathbf{I}$ są wielomianami zmiennej λ stopnia (co najwyżej) $n-1$. Zatem

$$[(\mathbf{A}-\lambda\mathbf{I})_{ij}]^T = \mathbf{D}_{n-1}\lambda^{n-1} + \cdots + \mathbf{D}_1\lambda + \mathbf{D}_0,$$

gdzie $\mathbf{D}_i$ są macierzami stopnia n.
Ponieważ

$$(\mathbf{A}-\lambda\mathbf{I})[(\mathbf{A}-\lambda\mathbf{I})_{ij}]^T = \det(\mathbf{A}-\lambda\mathbf{I})\cdot\mathbf{I} = c(\lambda)\mathbf{I},$$

więc

$$\begin{aligned}
c(\lambda)\mathbf{I} &= (\mathbf{A}-\lambda\mathbf{I})(\mathbf{D}_{n-1}\lambda^{n-1} + \cdots + \mathbf{D}_1\lambda + \mathbf{D}_0) = \\
&= -\mathbf{D}_{n-1}\lambda^{n-1} + (\mathbf{A}\mathbf{D}_{n-1} - \mathbf{D}_{n-2})\lambda^{n-1}\cdots + (\mathbf{A}\mathbf{D}_1 - \mathbf{D}_0)\lambda + \mathbf{A}\mathbf{D}_0,
\end{aligned}$$

ale z drugiej strony:

$$\begin{aligned}
c(\lambda)\mathbf{I} &= \begin{bmatrix} c(\lambda) & 0 & \dots & 0 \\ 0 & c(\lambda) & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & c(\lambda) \end{bmatrix} = \\
&= (c_n\mathbf{I})\lambda^n + (c_{n-1}\mathbf{I})\lambda^{n-1} + \cdots + (c_1\mathbf{I})\lambda^+ c_0\mathbf{I}.
\end{aligned}$$

Po porównaniu mamy:

$$\begin{aligned}
-\mathbf{D}_{n-1} &= c_n\mathbf{I} \\
\mathbf{A}\mathbf{D}_{n-1} - \mathbf{D}_{n-2} &= c_{n-1}\mathbf{I} \\
&\dots \\
\mathbf{A}\mathbf{D}_2 - \mathbf{D}_1 &= c_2\mathbf{I} \\
\mathbf{A}\mathbf{D}_1 - \mathbf{D}_0 &= c_1\mathbf{I} \\
\mathbf{A}\mathbf{D}_0 &= c_0\mathbf{I}.
\end{aligned}$$

Przez pomnożenie tych równości (z lewej strony) przez $\mathbf{A}^n, \dots, \mathbf{A}^2, \mathbf{A}, \mathbf{I}$ odpowiednio otrzymujemy:

$$\begin{aligned}
-\mathbf{A}^n\mathbf{D}_{n-1} &= c_n\mathbf{A}^n \\
\mathbf{A}^n\mathbf{D}_{n-1} - \mathbf{A}^{n-1}\mathbf{D}_{n-2} &= c_{n-1}\mathbf{A}^{n-1} \\
&\dots \\
\mathbf{A}^3\mathbf{D}_2 - \mathbf{A}^2\mathbf{D}_1 &= c_2\mathbf{A}^2 \\
\mathbf{A}^2\mathbf{D}_1 - \mathbf{A}\mathbf{D}_0 &= c_1\mathbf{A} \\
\mathbf{A}\mathbf{D}_0 &= c_0\mathbf{I},
\end{aligned}$$

a po dodaniu stronami uzyskujemy:

$$c_n\mathbf{A}^n + c_{n-1}\mathbf{A}^{n-1} + \cdots + c_1\mathbf{A}^1 + c_0\mathbf{I} = \mathbf{O}.\blacksquare$$

Twierdzenie Cayleya–Hamiltona ułatwia obliczanie wielomianów macierzowych, gdyż pozwala zredukować stopień wielomianu.

Wniosek 4. *Jeżeli $f(\lambda)$ jest dowolnym wielomianem o współczynnikach z ciała $\mathbb{K}$, $\mathbf{A}$ jest macierzą kwadratową stopnia n, to istnieje wielomian $r(\lambda)$ stopnia mniejszego od n, dla którego $f(\mathbf{A}) = r(\mathbf{A})$.*

D o w ó d. Z twierdzenia o dzieleniu wielomianów z resztą wynika, że istnieją takie wielomiany $g(\lambda)$ i $r(\lambda)$, że $f(\lambda) = c(\lambda)g(\lambda) + r(\lambda)$, przy czym $r(\lambda)$ jest stopnia mniejszego niż $n = \deg c(\lambda)$. Zatem

$$f(\mathbf{A}) = c(\mathbf{A})g(\mathbf{A}) + r(\mathbf{A}),$$

czyli $f(\mathbf{A}) = r(\mathbf{A})$, bo $c(\mathbf{A}) = \mathbf{O}$. $\blacksquare$

1. Sprawdzić twierdzenie Cayleya–Hamiltona dla macierzy:

$$\mathbf{A} = \begin{bmatrix} 2 & 1 \\ -1 & 3 \end{bmatrix}.$$

Tutaj $c(\lambda) = \lambda^2 - 5\lambda + 7$, $\mathbf{A} = \begin{bmatrix} 2 & 1 \\ -1 & 3 \end{bmatrix}$. Łatwo sprawdzić, że

$$\mathbf{A}^2 - 5\mathbf{A} + 7 = 0.$$

2. Obliczyć $\mathbf{A}^6 - 25\mathbf{A}^2 + 112\mathbf{A}$, posługując się twierdzeniem Cayleya–Hamiltona, gdy

$$\mathbf{A} = \begin{bmatrix} 0 & 0 & 2 \\ 2 & 1 & 0 \\ -1 & -1 & 3 \end{bmatrix}.$$

Tutaj $f(\lambda) = \lambda^6 - 25\lambda^2 + 112$, $c(\lambda) = -\lambda^3 + 4\lambda^2 - 5\lambda - 2$. Wielomian $f(\lambda)$ po podzieleniu przez $c(\lambda)$ daje iloraz $-\lambda^3 - 4\lambda^2 - 11\lambda - 22$ i resztę $-20\lambda - 44$. Zatem

$$\mathbf{A}^6 - 25\mathbf{A}^2 + 112\mathbf{A} = -20\mathbf{A} - 44\mathbf{I} = \begin{bmatrix} -44 & 0 & -40 \\ -40 & -64 & 0 \\ 20 & 20 & -104 \end{bmatrix}.$$

Z twierdzenia Cayleya–Hamiltona wynika, że każda macierz kwadratowa stopnia n spełnia pewne równanie stopnia n. Ale niektóre macierze spełniają również pewne równanie stopnia mniejszego niż n.

Definicja 10. Niech $\mathbf{A}$ będzie macierzą kwadratową. Wielomian $m(\lambda)$ taki, że $m(\mathbf{A}) = 0$, nazywamy *wielomianem minimalnym*, jeśli nie istnieje wielomian stopnia mniejszego niż $\deg m(t)$ mający tę samą własność.

Łatwo wykazać, że $c(\lambda)$ musi być podzielny przez $m(\lambda)$. Trudniej, że $m(\lambda)$ ma dokładnie takie same pierwiastki co $c(\lambda)$, co najwyżej z mniejszymi krotnościami. W szczególności, jeśli $c(\lambda)$ ma pierwiastki jednokrotne, to $m(\lambda) = c(\lambda)$.