Macierze i układy równań liniowych

Ostatnia aktualizacja: 27.10.2021 r.

Wykład z geometrii z algebrą liniową zaczynamy omówieniem teorii układów równań liniowych o współczynnikach w ciele. Czym jest ciało dowiemy się na kolejnym wykładzie. W tym momencie chcemy raczej uporządkować i usystematyzować język dotyczący samych układów i ich rozwiązań. Ważnym elementem tego uporządkowania pojęć będzie wprowadzenie tak zwanych macierzy.

Zbiór liczb rzeczywistych oznaczamy przez $\mathbb{R}$.

Definicja 1. *Niech $X = \{x_1, \ldots, x_n\}$ będzie zbiorem skończonym. **Równanie liniowe** na **zbiorze zmiennych** X o **współczynnikach** w zbiorze K to wyrażenie postaci*

$$a_1x_1 + a_2x_2 + \ldots + a_nx_n = b, \text{ gdzie } a_1, a_2, \ldots, a_n, b \in K.$$

***Rozwiązanie** powyższego równania to ciąg $(s_1, s_2, \ldots, s_n)$ elementów należących do zbioru K taki, że*

$$a_1 \cdot s_1 + a_2 \cdot s_2 + \ldots + a_n \cdot s_n = b.$$

Przykłady

- Dla $X = \{x\}$ oraz $K = \mathbb{R}$, równaniem liniowym na zbiorze zmiennych X i zbiorze współczynników K jest $\sqrt{2}x = \pi$. Natomiast definicji powyższej nie spełniają równania $2x + 1 = 3$ lub $2x = x$.
- Dla $X = \{x_1, x_2, x_3\}$ oraz $K = \mathbb{Q}$, równaniem liniowym na zbiorze zmiennych X i zbiorze współczynników K jest: $\frac{1}{2}x_1 + x_2 + 0 \cdot x_3 = 0$, ale nie jest nim równanie $\sqrt{2}x_1 = 2$.

Aby powyższa definicja miała sens, przyjmujemy na razie, że K jest równe $\mathbb{R}$ lub jego niepustemu podzbiorowi. Na kolejnym wykładzie omawiać będziemy pojęcie ciała oraz pojęcie równania liniowego o współczynnikach w ciele, dla których przeniesiemy bez właściwie żadnych poważnych modyfikacji wyniki uzyskane dla równań liniowych o współczynnikach w $\mathbb{R}$.

Definicja 2. ***Układ** m **równań liniowych** na zbiorze $X = \{x_1, \ldots, x_n\}$ o współczynnikach rzeczywistych to ciąg m równań liniowych o współczynnikach rzeczywistych postaci:*

$$\begin{cases} a_{11}x_1 + \ a_{12}x_2 + \ldots + \ a_{1n}x_n = b_1 \\ a_{21}x_1 + \ a_{22}x_2 + \ldots + \ a_{2n}x_n = b_2 \\ a_{31}x_1 + \ a_{32}x_2 + \ldots + \ a_{3n}x_n = b_3 \\ \qquad\qquad \cdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m. \end{cases} \tag{1}$$

Rozwiązanie powyższego układu równań liniowych to ciąg elementów $(s_1, \ldots, s_n)$ ze zbioru $\mathbb{R}$, który jest rozwiązaniem każdego z m równań liniowych tego układu.

Ważne jest zwrócenie uwagi na duży stopień dowolności, jaki daje ta definicja. Nic nie stoi na przeszkodzie by układ składał się z pojedynczego równania, albo zawierał kilka identycznych równań. Co więcej, dołożenie do układu choćby najbardziej trywialnego równania (np. $0 = 0$) daje nam inny układ równań! Oto kilka przykładów:

- $\begin{cases} 2x_1 + 3x_2 + x_3 & = 1 \\ \quad -\frac{1}{2}x_2 + x_3 & = 0 \end{cases}$
- $$\begin{cases} x_1 + \ x_2 & = 1 \\ x_1 + \ x_2 & = 1 \\ x_1 + \ x_2 & = 1 \\ 0x_1 + 0x_2 & = 3 \end{cases}$$
- $$\begin{cases} 0x_1 + 0x_2 + 0x_3 & = 0 \\ 0x_1 + 0x_2 + \ x_3 & = 0 \\ x_1 + \ x_2 + \ x_3 & = 0 \end{cases}$$

Przedstawimy teraz dwa proste fakty dotyczące zbiorów rozwiązań układów równań o n zmiennych o współczynnikach rzeczywistych.

Obserwacja 1. *Jeśli* $(s_1, s_2, \ldots, s_n)$ *oraz* $(s'_1, s'_2, \ldots, s'_n)$ *są rozwiązaniami układu równań* (1)*, to ciąg*

$$(s_1 - s'_1, s_2 - s'_2, \ldots, s_n - s'_n)$$

jest rozwiązaniem układu równań:

$$\begin{cases} a_{11}x_1 + \ a_{12}x_2 + \ldots + \ a_{1n}x_n = 0 \\ a_{21}x_1 + \ a_{22}x_2 + \ldots + \ a_{2n}x_n = 0 \\ a_{31}x_1 + \ a_{32}x_2 + \ldots + \ a_{3n}x_n = 0 \\ \qquad\qquad \ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = 0. \end{cases} \tag{2}$$

Dowód. Ciągi $(s_1, s_2, \ldots, s_n)$ oraz $(s'_1, s'_2, \ldots, s'_n)$ sa rozwiązaniami każdego z równań

$$a_{i1}x_1 + a_{i2}x_2 + \ldots + a_{in}x_n = b_i,$$

gdzie $i = 1, 2, \ldots, m$. Pisząc wprost mamy:

$$a_{i1}s_1 + a_{i2}s_2 + \ldots + a_{in}s_n = b_i$$
$$a_{i1}s'_1 + a_{i2}s'_2 + \ldots + a_{in}s'_n = b_i,$$

czyli po odjęciu stronami widzimy, że $(s_1 - s'_1, s_2 - s'_2, \ldots, s_n - s'_n)$ spełnia każde z m równań układu (2):

$$a_{i1}(s_1 - s'_1) + a_{i2}(s_2 - s'_2) + \ldots + a_{in}(s_n - s'_n) = 0,$$

co oznacza, że jest to rozwiązanie całego układu (2). □

Druga obserwacja sprowadza problem znajdowania zbioru rozwiązań układu równań (1) do dwóch kroków:

- znalezienia jakiegokolwiek rozwiązania układu (1)
- znalezienia wszystkich rozwiązań układu (2).

Obserwacja 2. *Załóżmy, że* $(s_1, \ldots, s_n)$ *jest rozwiązaniem układu równań* (1)*. Wówczas każde rozwiązanie układu* (1) *jest postaci:*

$$(s_1 + u_1, \ldots, s_n + u_n) \qquad (\diamondsuit),$$

gdzie $(u_1, \ldots, u_n)$ *jest rozwiązaniem układu* (2)*.*

Dowód. Weźmy dowolne rozwiązanie $(s_1, \ldots, s_n)$ układu (1) oraz dowolne rozwiązanie $(u_1, \ldots, u_n)$ układu (2). Wówczas ciąg

$$(s_1 + u_1, \ldots, s_n + u_n)$$

jest rozwiązaniem układu (1), bo spełnia dowolne z jego równań:

$$a_{i1}(s_1 + u_1) + \ldots + a_{in}(s_n + u_n) = (a_{i1}s_1 + \ldots + a_{in}s_n) + (a_{i1}u_1 + \ldots + a_{in}u_n) = b_i + 0 = b_i.$$

Pozostaje wykazać, że dowolne rozwiązanie $(s'_1, \ldots, s'_n)$ układu (1) mozna przedstawić w postaci $(\diamondsuit)$, dla pewnego rozwiązania $(u_1, \ldots, s_n)$ układu (2). Istotnie, zgodnie z poprzednią Obserwacją wystarczy wziąć:

$$u_1 = s'_1 - s_1, \quad \ldots \quad, u_n = s'_n - s_n$$

i dostajemy:

$$(s'_1, \ldots, s'_n) = (s_1 + (s'_1 - s_1), \ldots, s_n + (s'_n - s_n)).$$

□

Dla zilustrowania powyższych faktów rozważmy układ równań liniowych o współczynnikach w $\mathbb{R}$ postaci:

$$\left\{\begin{array}{ccccccc} x & + & 2y & + & 3z & = & 1 \\ 4x & + & 5y & + & 6z & = & 1 \\ 7x & + & 8y & + & 9z & = & 1 \end{array}\right. . \qquad (\dagger)$$

Zgodnie z procedurą opisaną wyżej do opisania rozwiązań tego układu potrzebujemy dowolne jego rozwiązanie, na przykład:

$$(s_1, s_2, s_3) = (-1, 1, 0)$$

oraz wszystkie rozwiązania układu postaci:

$$\left\{\begin{array}{ccccccc} x & + & 2y & + & 3z & = & 0 \\ 4x & + & 5y & + & 6z & = & 0 \\ 7x & + & 8y & + & 9z & = & 0 \end{array}\right. . \qquad (\dagger\dagger)$$

Można sprawdzić, że rozwiązań tych jest nieskończenie wiele oraz, że są one postaci:

$$(z, -2z, z), \text{ gdzie } z \in \mathbb{R}.$$

Geometryczną interpretację tego faktu znajdą Państwo w dalszej części wykładu. W związku z tym każde rozwiązanie układu ($\dagger$) ma postać:

$$(-1 + z, 1 - 2z, z), \text{ gdzie } z \in \mathbb{R}.$$

Definicja 3. *Układ równań liniowych postaci* (1) *nazywamy:*

- ***jednorodnym****, jeśli* $b_i = 0$*, dla każdego* $i = 1, 2, \ldots, m$*,*
- ***niesprzecznym****, jeśli zbiór rozwiązań układu* (1) *jest niepusty,*
- ***sprzecznym****, jeśli zbiór rozwiązań układu* (1) *jest pusty.*

Układ (2) nazywać będziemy czasem układem jednorodnym odpowiadającym układowi (1).

Definicja 4. *Układy równań liniowych określone na tym samym zbiorze* n *zmiennych* X *o zbiorze współczynników* K *nazwiemy* ***równoważnymi****, jeśli mają one te same zbiory rozwiązań (traktowane jako podzbiory zbioru wszystkich ciągów* n*-elementowych o wyrazach z* K*).*

Przykład. Następujące układy równań liniowych zmiennych x_1, x_2 nad $\mathbb{R}$ są równoważne:

$$\begin{cases} x_1 + x_2 = 0 \\ x_1 + x_2 = 0 \\ x_1 - x_2 = 2 \end{cases}, \qquad \begin{cases} x_1 + x_2 = 0 \\ x_1 - x_2 = 2 \end{cases}, \qquad \begin{cases} x_1 = 1 \\ x_2 = -1 \end{cases}$$

Naszym celem jest opis rozwiązań układów równań liniowych o współczynnikach rzeczywistych. Metoda polega na zastępowaniu układu innym – równoważnym, i z jakiegoś powodu prostszym do rozwiązania.

Definicja 5. *Dane sa równania liniowe* U, U' *nad zbiorem zmiennych* $x_1, \ldots, x_n$ *o współczynnikach w* $\mathbb{R}$*:*

$$U: \quad a_1x_1 + a_2x_2 + \ldots + a_nx_n = b, \qquad U': \quad a_1'x_1 + a_2'x_2 + \ldots + a_n'x_n = b'.$$

Dla każdego $\lambda \in \mathbb{R}$ *przez* λU *określamy równanie:*

$$\lambda \cdot a_1x_1 + \lambda \cdot a_2x_2 + \ldots + \lambda \cdot a_nx_n = \lambda \cdot b.$$

Przez $U + U'$ *rozumieć będziemy równanie:*

$$(a_1 + a_1')x_1 + (a_2 + a_2')x_2 + \ldots + (a_n + a_n')x_n = b + b'.$$

Obserwacja 3. *Jeśli* $(s_1, \ldots, s_n)$ *jest rozwiązaniem równań liniowych* U_1 *oraz* U_2*, gdzie* $s_1, \ldots, s_n \in \mathbb{R}$*, to jest też rozwiązaniem każdego równania postaci:*

$$\lambda_1 U_1 + \lambda_2 U_2, \text{ gdzie } \lambda_1, \lambda_2 \in \mathbb{R}.$$

Dowód. Niech $U_1 : a_1x_1 + \ldots + a_nx_n = b$ oraz $U_2 : a'_1x_1 + \ldots + a'_nx_n = b'$. Mamy wówczas

$$(\lambda_1 a_1 + \lambda_2 a'_1)s_1 + \ldots + (\lambda_1 a_n + \lambda_2 a'_n)s_n = \lambda_1 b + \lambda_2 b'.$$

W szczególności $(s_1, \ldots, s_n)$ jest rozwiązaniem równania $\lambda_1 U_1 + \lambda_2 U_2$. □

Podkreślmy, że nie pokazaliśmy równoważności układu $U_1 \wedge U_2$ z układem $\lambda_1 U_1 + \lambda_2 U_2$. Pokazaliśmy jedynie zawieranie się zbiorów rozwiązań układu $U_1 \wedge U_2$ w zbiorze rozwiązań układu $\lambda_1 U_1 + \lambda_2 U_2$. Dla zilustrowania tego odnotujmy, że układ równań liniowych zmiennych x_1, x_2 nad $\mathbb{R}$ złożony z równań $x_1 = 1$ oraz $x_2 = 1$ ma jedno rozwiązanie $(1, 1)$, podczas, gdy układ: $x_1 + x_2 = 2$ ma ich nieskończenie wiele (nad $\mathbb{R}$). Wskażemy teraz takie operacje na układach równań, które zachowują równoważność.

Twierdzenie 1. *Następujące operacje przeprowadzają układ równań liniowych U nad $\mathbb{R}$ w układ równoważny:*

(1) Dodanie do równania innego równania pomnożonego przez liczbę.

(2) Zamiana dwóch równań miejscami.

(3) Pomnożenie równania przez liczbę różną od zera.

Dowód. Dowód jest oczywisty dla operacji (2) i (3). Niech układ U' powstaje z U przez dodanie do i-tego równania U_i równania U_j przemnożonego przez $a \in \mathbb{R}$. Wtedy U powstaje z U' przez dodanie do i-tego równania $U_i + aU_j$ równania U_j przemnożonego przez $-a \in \mathbb{R}$. Na mocy Obserwacji 3: jeśli $(s_1, \ldots, s_n)$ jest rozwiązaniem równań U_i oraz U_j, to także jest rozwiązaniem $U_i + aU_j$. A zatem jeśli $(s_1, \ldots, s_n)$ spełnia U, to także U'. Z drugiej strony, jeśli $(s_1, \ldots, s_n)$ jest rozwiązaniem $U_i + aU_j$ oraz U_j, to jest też rozwiązaniem $U_i = (U_i + aU_j) - aU_j$. A zatem jeśli $(s_1, \ldots, s_n)$ spełnia U', to spełnia U. □

Definicja 6. *Operacje (1)-(3) nazywać będziemy* ***operacjami elementarnymi na układzie*** *U.*

A zatem operacje elementarne pozwalają przeprowadzić układ do innej, równoważnej postaci. Do jakich postaci chcemy dochodzić w ramach takiego „upraszczania” przy pomocy operacji elementarnych?

Definicja 7. *Niech U oraz U' będą równoważnymi układami równań liniowych o n zmiennych (niewiadomych). Przypuśćmy, że układ U'* ***można przepisać*** *w postaci:*

$$\begin{cases} x_{j_1} & = c_{11}x_1 + c_{12}x_2 + \ldots + c_{1n}x_n + d_1 \\ \ldots & \\ x_{j_k} & = c_{k1}x_1 + c_{k2}x_2 + \ldots + c_{kn}x_n + d_k \end{cases}$$

przy czym $j_1 < \ldots < j_k$ oraz zmienne $x_{j_1}, \ldots, x_{j_k}$ nie występują po prawej stronie (to znaczy: stoją przy nich współczynniki zerowe, czyli $c_{ij} = 0$, dla $j = j_1, \ldots, j_k$).

Mówimy wówczas, że:

- *układ U'* ***jest rozwiązaniem ogólnym*** *(zadaje rozwiązanie ogólne) układu U,*
- *w rozwiązaniu tym $x_{j_1}, \ldots, x_{j_k}$ nazywamy* ***zmiennymi zależnymi****,*
- *pozostałe x_i nazywamy* ***zmiennymi niezależnymi****, albo* ***parametrami****.*

Znowu potrzebna jest chwila uwagi by przeczytać powyższą definicję. Zdefiniowaliśmy właśnie owe „proste równoważne postaci układów równań”, z których łatwo jest odczytywać rozwiązania.

Przykład 1. *Rozważmy układ równań liniowych:*

$$\begin{cases} x_1 + x_2 + x_3 + x_4 = 1 \\ x_1 - x_2 - x_3 - x_4 = 0 \\ x_1 - x_2 - x_3 - x_4 = 0 \end{cases} .$$

To nie przypadek, że dwa razy wpisane jest to samo równanie. Przyzwyczaimy się do tego po wprowadzeniu pojęcia macierzy i stanie się to naturalne. Układ ten ma rozwiązanie ogólne. Jest nim na przykład:

$$\begin{cases} x_1 & = \frac{1}{2} \\ x_2 & = -x_3 - x_4 + \frac{1}{2} \end{cases}$$

A zatem jak przedstawić zbiór wszystkich rozwiązań? Jest to zbiór ciągów postaci

$$\{(\frac{1}{2}, -x_3 - x_4 + \frac{1}{2}, x_3, x_4), \quad gdzie\ x_3, x_4 \in \mathbb{R}\}.$$

Na marginesie istnieje także inny sposób zapisania rozwiązania ogólnego:

$$\begin{cases} x_1 &= \frac{1}{2} \\ x_4 &= -x_2 - x_3 + \frac{1}{2} \end{cases}$$

Ma ono zatem postać $\{(\frac{1}{2}, x_2, x_3, -x_2 - x_3 + \frac{1}{2}), \quad gdzie\ x_2, x_3 \in \mathbb{R}\}$.

Mamy zatem ilustrację definicji. W pierwszym rozwiązaniu ogólnym zmiennymi zależnymi są x_1, x_2 (w zasadzie x_1 jest tutaj akurat stałą), parametrami zaś x_3, x_4. W drugim rozwiązaniu ogólnym jest nieco inaczej. Parametrami są x_2, x_3. A zatem formalnie rzecz biorąc jeden układ równań posiadać może kilka ogólnych postaci rozwiązań, co nie wydaje się do końca wygodne. Intuicja podpowiada jednak, że druga z powyższych konfiguracji nie jest „elegancka", ponieważ zmienna zależna ma większy indeks niż zmienna niezależna. W dalszej części poradzimy sobie z tą drobną delikatnością.

Przejdziemy teraz do systematycznego opisu rozwiązywania układów równań liniowych o współczynnikach rzeczywistych. W tym celu wprowadzamy fundamentalne dla całego wykładu pojęcie macierzy.

Definicja 8. ***Macierzą rozmiaru*** *$m \times n$ (inaczej: macierzą o m wierszach i n kolumnach) o* ***wyrazach*** *ze zbioru X nazywamy tablicę:*

$$D = \begin{bmatrix} d_{11} & d_{12} & \dots & d_{1n} \\ d_{21} & d_{22} & \dots & d_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ d_{m1} & d_{m2} & \dots & d_{mn} \end{bmatrix}$$

gdzie $d_{ij} \in X$ dla $1 \leqslant i \leqslant m$ oraz $1 \leqslant j \leqslant n$.

Rzędy poziome macierzy D nazywamy ***wierszami****, wiersze pionowe zaś nazywamy* ***kolumnami****. Wyrazy $d_{i1}, d_{i2}, \dots, d_{in}$ tworzą i-ty wiersz, a wyrazy $d_{1j}, d_{2j}, \dots, d_{mj}$ tworzą j-tą kolumnę macierzy D. Będziemy też pisać $D = [d_{ij}]$. Zbiór wszystkich macierzy $m \times n$ o wyrazach ze zbioru X oznaczamy jako $M_{m\times n}(X)$.*

O macierzy można myśleć też jako o funkcji $\{1, \dots, m\} \times \{1, \dots, n\} \to X$ przypisującej każdej parze (i, j) (dla odpowiednich indeksów) element $d_{ij} \in X$, który nazywać będziemy wyrazem i, j macierzy D.

Definicja 9. *Każdemu układowi równań liniowych postaci:*

$$U : \begin{cases} a_{11}x_1 + \ a_{12}x_2 + \ldots + \ a_{1n}x_n &= b_1 \\ a_{21}x_1 + \ a_{22}x_2 + \ldots + \ a_{2n}x_n &= b_2 \\ \dots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n &= b_m. \end{cases}$$

możemy przypisać macierz $m \times (n + 1)$ postaci:

$$\left[\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{array}\right].$$

Nazywamy ją ***macierzą (albo macierzą rozszerzoną) układu*** *U. Macierz powstałą przez pominięcie ostatniej kolumny w macierzy układu U będziemy nazywać* ***macierzą współczynników układu*** *U.*

Rozważmy kilka przykładów macierzy i odpowiadających im układów równań liniowych:

układ	macierz rozszerzona	macierz współczynników
$\begin{cases} 2x_1 + 3x_2 + x_3 &= 1 \\ \quad - \frac{1}{2}x_2 + x_3 &= 0 \end{cases}$	$\left[\begin{array}{ccc\|c} 2 & 3 & 1 & 1 \\ 0 & -\frac{1}{2} & 1 & 0 \end{array}\right]$	$\begin{bmatrix} 2 & 3 & 1 \\ 0 & -\frac{1}{2} & 1 \end{bmatrix}$
$\begin{cases} x_1 + \ x_2 &= 1 \\ x_1 + \ x_2 &= 1 \\ x_1 + \ x_2 &= 1 \\ 0x_1 + 0x_2 &= 3 \end{cases}$	$\left[\begin{array}{cc\|c} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \\ 0 & 0 & 3 \end{array}\right]$	$\begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 0 & 0 \end{bmatrix}$
$\begin{cases} 0x_1 + 0x_2 + 0x_3 &= 0 \\ 0x_1 + 0x_2 + \ x_3 &= 0 \\ \ x_1 + \ x_2 + \ x_3 &= 0 \end{cases}$	$\left[\begin{array}{ccc\|c} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 \end{array}\right]$	$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix}$

Przyjrzyjmy się teraz macierzom układów będących w postaci zredukowanej.

układ	macierz rozszerzona
$\begin{cases} x_1 & & &= 2 \\ & x_2 & &= 1 \\ & & x_3 &= 4 \end{cases}$	$\left[\begin{array}{ccc\|c} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 4 \end{array}\right]$
$\begin{cases} x_1 & &= \frac{1}{2} \\ & x_2 + x_3 + x_4 &= \frac{1}{2} \end{cases}$	$\left[\begin{array}{cccc\|c} 1 & 0 & 0 & 0 & \frac{1}{2} \\ 0 & 1 & 1 & 1 & \frac{1}{2} \end{array}\right]$
$\begin{cases} x_1 \quad + x_3 \quad - 2x_5 &= 0 \\ \quad x_2 \quad\quad + \ x_5 &= 0 \\ \quad\quad\quad x_4 + \ x_5 &= 0 \end{cases}$	$\left[\begin{array}{ccccc\|c} 1 & 0 & 1 & 0 & -2 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 \end{array}\right]$

W powyższych przykładach pokreślono kolorem współczynniki odpowiadające zmiennym zależnym. Zauważmy, że w każdym kolejnym wierszu współczynniki te znajdują się w kolumnach o większym indeksie niż wyraz tego samego typu w poprzednim wierszu. Konfigurację tę zidentyfikujemy za chwilę w osobnej definicji. W tym momencie przeformułujemy na język macierzy operacje elementarne na równaniach.

Definicja 10. *Niech $A \in M_{m\times n}(\mathbb{R})$. Następujące transformacje macierzy A nazywamy* ***operacjami elementarnymi na wierszach****:*

(1) Dodanie do wiersza innego wiersza przemnożonego przez liczbę rzeczywistą.

(2) Zamiana dwóch wierszy miejscami.

(3) Pomnożenie wiersza przez liczbę różną od zera.

Analogicznie definiuje się ***operacje elementarne na kolumnach*** *macierzy A.*

Widzimy zatem, że operacjom elementarnym na układzie U odpowiadają operacje elementarne na wierszach macierzy tego układu. Proces znajdowania rozwiązania ogólnego układu U będzie polegał na kolejnym upraszczaniu macierzy tego układu, przy zastosowaniu operacji elementarnych na wierszach.

Przykłady operacji elementarnych na wierszach macierzy i notacja, którą będziemy stosować:

$$\begin{bmatrix} 1 & 2 & -1 & -1 & 0 \\ 2 & -1 & -2 & 1 & 1 \\ -1 & 0 & 1 & 4 & 1 \end{bmatrix} \xrightarrow{w_2 - 2\cdot w_1} \begin{bmatrix} 1 & 2 & -1 & -1 & 0 \\ 0 & -5 & 0 & 3 & 1 \\ -1 & 0 & 1 & 4 & 1 \end{bmatrix}$$
$$\xrightarrow{w_3 + w_1} \begin{bmatrix} 1 & 2 & -1 & -1 & 0 \\ 0 & -5 & 0 & 3 & 1 \\ 0 & 2 & 0 & 3 & 1 \end{bmatrix}$$
$$\xrightarrow{w_2 \leftrightarrow w_3} \begin{bmatrix} 1 & 2 & -1 & -1 & 0 \\ 0 & 2 & 0 & 3 & 1 \\ 0 & -5 & 0 & 3 & 1 \end{bmatrix}$$
$$\xrightarrow{2\cdot w_2} \begin{bmatrix} 1 & 2 & -1 & -1 & 0 \\ 0 & 4 & 0 & 6 & 2 \\ 0 & -5 & 0 & 3 & 1 \end{bmatrix}$$

Definicja 11. *Mówimy, że macierz $A = [a_{ij}]$ jest w **postaci schodkowej**, jeśli A spełnia następujące warunki:*

- *każdy wiersz zerowy (tzn. złożony z samych zer) znajduje się poniżej każdego wiersza niezerowego (czyli jeśli i-ty wiersz tej macierzy jest niezerowy, to j-ty wiersz jest zerowy tylko gdy $j > i$),*
- *dla każdego $i > 1$ pierwszy licząc od lewej niezerowy wyraz w i-tym wierszu znajduje się w kolumnie stojącej na prawo od pierwszego niezerowego wyrazu $(i-1)$-szego wiersza*[1]*.*

Rozważmy, za skryptem, kilka prostych przykładów ilustrujących nową definicję: dwie pierwsze macierze są w postaci schodkowej, dwie kolejne nie. Czy Czytelnik widzi owe „schodki”?

$$\begin{bmatrix} 0 & 4 & 7 & 2 \\ 0 & 0 & 1 & 12 \\ 0 & 0 & 0 & 5 \end{bmatrix}, \quad \begin{bmatrix} 2 & 5 & 1 & 1 \\ 0 & 3 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 & 3 & 4 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 5 \end{bmatrix}, \quad \begin{bmatrix} 2 & 8 & 3 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 12 \end{bmatrix}.$$

Dlaczego interesują nas macierze w postaci schodkowej? Przyjrzyjmy się następującemu przykładowi. Weźmy układ równań z dokładnie jednym rozwiązaniem, na przykład o macierzy:

$$\begin{bmatrix} 1 & 2 & 3 & 4 & 1 \\ 0 & 5 & 6 & 7 & 1 \\ 0 & 0 & 8 & 9 & 1 \\ 0 & 0 & 0 & 10 & 1 \end{bmatrix}$$

Oczywiście ma ona postać schodkową a po przetłumaczeniu jej z powrotem na język równań dostalibyśmy układ:

$$\begin{cases} x_1 + 2x_2 + 3x_3 + 4x_4 = 1 \\ \phantom{x_1 + {}} 5x_2 + 6x_3 + 7x_4 = 1 \\ \phantom{x_1 + 2x_2 + {}} 8x_3 + 9x_4 = 1 \\ \phantom{x_1 + 2x_2 + 3x_3 + {}} 10x_4 = 1 \end{cases},$$

który rozwiązywalibyśmy dalej szkolną metodą „przez podstawienie”. W wielu praktycznych zastosowaniach uzyskanie postaci schodkowej macierzy (układu, lub nie) będzie dla nas wystarczające do uzyskania opisu zbioru rozwiązań układu równań..

Definicja 12. *Mówimy, że macierz jest w **zredukowanej postaci schodkowej**, jeśli jest w postaci schodkowej oraz w każdym niezerowym wierszu pierwszy niezerowy wyraz wynosi 1 i jest jedynym niezerowym wyrazem w swojej kolumnie.*

Przykłady:

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 4 \end{array}\right], \quad \left[\begin{array}{cccc|c} 1 & 0 & 0 & 0 & \frac{1}{2} \\ 0 & 1 & 1 & 1 & \frac{1}{2} \end{array}\right], \quad \left[\begin{array}{ccccc|c} 1 & 0 & 1 & 0 & -2 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 \end{array}\right], \quad \left[\begin{array}{cc|c} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{array}\right].$$

Jeśli układ ma macierz w postaci schodkowej zredukowanej, to są takie kolumny gdzie powstaje „schodek”, powiedzmy $j_1 < j_2 < \ldots < j_k$ i zmienna z kolumny j_i raz jeden występuje w całym układzie właśnie w i-tym równaniu. Co więcej stoi przy niej współczynnik 1. A zatem jest to macierz, z której łatwo uzyskać

[1] Dla miłośników precyzji: dla każdego $i > 1$ jeśli $a_{ik} \neq 0$, to dla każdego $j < i$ istnieje $k' < k$ takie, że $a_{jk'} \neq 0$.

rozwiązanie ogólne układu równań, o którym mówiliśmy wcześniej. Chyba, że... jedna z tych kolumn to ostatnia kolumna. Wtedy sytuacja jest nieco inna. Czy Czytelnik widzi jaka? Pokażemy teraz kluczową obserwację: do uzyskania postaci schodkowej i schodkowej zredukowanej wystarczy stosować na wierszach macierzy odpowiednie operacje elementarne.

Twierdzenie 2. *Każdą macierz $A \in M_{n\times n}(\mathbb{R})$ można:*

(i) za pomocą operacji elementarnych typu (1) i (2) sprowadzić do postaci schodkowej,

(ii) za pomocą operacji elementarnych typu (1), (2) i (3) sprowadzić do postaci schodkowej zredukowanej.

Dowód. Pokażemy jedynie (i). Dowód (ii) będzie ćwiczeniem. Stosujemy zasadę indukcji matematycznej względem liczby wierszy macierzy. Dla $m = 1$ twierdzenie jest oczywiste. Załóżmy, że twierdzenie jest udowodnione dla macierzy o co najwyżej $m-1$ wierszach. Niech $A \in M_{m\times n}(\mathbb{R})$. Jeśli A jest macierzą zerową (to znaczy każdy jej wyraz jest zerem), to oczywiście jest schodkowa. Załóżmy więc, że A nie jest macierzą zerową. Niech s będzie numerem pierwszej niezerowej kolumny macierzy A. Wybierzmy takie r, że $a_{rs} \neq 0$.

$$A = \begin{bmatrix} 0 & \dots & 0 & a_{1s} & * & * & * \\ 0 & \dots & 0 & a_{2s} & * & * & * \\ \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & 0 & {\color{red}a_{rs}} & * & * & * \\ \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & 0 & a_{ms} & * & * & * \end{bmatrix}$$

Zamieniamy miejscami wiersze: pierwszy i r-ty (operacja typu (2)). Za pomocą operacji (1) zerujemy[2] wszystkie, poza pierwszym, wyrazy w s-tej kolumnie (od i-tego wiersza odejmujemy pierwszy przemnożony przez $\frac{a_{is}}{a_{rs}}$). Otrzymujemy w ten sposób macierz $A' = [a'_{ij}]$, w której kolumny o numerach $1, \dots, s-1$ są zerowe oraz $a'_{1s} \neq 0$ i $a'_{is} = 0$, dla $i > 1$.

$$A' = \begin{bmatrix} 0 & \dots & 0 & a_{rs} & * & * & * \\ 0 & \dots & 0 & {\color{red}a_{2s} - \frac{a_{2s}}{a_{rs}} \cdot a_{rs}} & * & * & * \\ \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & 0 & {\color{red}a_{1s} - \frac{a_{1s}}{a_{rs}} \cdot a_{rs}} & * & * & * \\ \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & 0 & {\color{red}a_{ms} - \frac{a_{ms}}{a_{rs}} \cdot a_{rs}} & * & * & * \end{bmatrix} = \begin{bmatrix} 0 & \dots & 0 & a_{rs} & * & * & * \\ 0 & \dots & 0 & 0 & * & * & * \\ \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & 0 & 0 & * & * & * \\ \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & 0 & 0 & * & * & * \end{bmatrix}$$

Niech $A'' \in M_{(m-1)\times n}(\mathbb{R})$ będzie macierzą otrzymaną z A' przez usunięcie pierwszego wiersza. Stosujemy do A'' założenie indukcyjne. Otrzymujemy macierz, która wraz z pierwszym wierszem macierzy A' tworzy macierz A''' w postaci schodkowej. Oczywiście A''' jest uzyskana z A przez ciąg operacji typu (1) i (2). □

Wniosek 1. *Każdy niesprzeczny układ równań liniowych ma rozwiązanie ogólne, przy czym:*

- *aby znaleźć rozwiązanie układu U wystarczy macierz tego układu sprowadzić do zredukowanej postaci schodkowej elementarnymi operacjami na wierszach,*
- *jeśli otrzymana macierz nie zawiera wiersza postaci $0 \dots 0\, 1$, to można z niej odczytać rozwiązanie ogólne układu U,*
- *jeśli otrzymana macierz zawiera wiersz postaci $0 \dots 0\, 1$, to układ U jest sprzeczny.*

Dowód. Sprowadzamy macierz układu U do zredukowanej postaci schodkowej. Jeśli otrzymana macierz ma wiersz $0 \dots 0\, 1$, to odpowiadający jej układ równań zawiera równanie $0x_1 + 0x_2 + \ldots + 0x_n = 1$, więc układ ten (i równoważny z nim układ U) jest sprzeczny. Jeśli otrzymana macierz w postaci schodkowej zredukowanej nie zawiera wiersza $0 \dots 0\, 1$, to odpowiadający jej układ równań ma (po pominięciu

[2]Krok ten nazywany jest często **eliminacją** i pochodzi od niego nazwa algorytmu opisującego uzyskiwanie postaci schodkowej (a także zredukowanej), jak również wielu równoważnych mu rozkładów macierzy.

ewentualnych równań postaci $0x_1 + 0x_2 + \ldots + 0x_n = 0$) postać

$$\begin{cases} x_{j_1} + a_{1(j_1+1)}x_{1(j_1+1)} + \ldots + a_{1n}x_n & = b_1 \\ x_{j_2} + a_{2(j_2+1)}x_{1(j_2+1)} + \ldots + a_{2n}x_n & = b_2 \\ \vdots \\ x_{j_k} + a_{k(j_k+1)}x_{1(j_k+1)} + \ldots + a_{kn}x_n & = b_k \end{cases}$$

przy czym $j_1 < j_2 < \ldots < j_k$ oraz dla każdego $1 \leqslant s \leqslant k$ mamy $a_{ij_s} = 0$, dla wszystkich i. Stąd układ jest równoważny układowi:

$$\begin{cases} x_{j_1} = -a_{1(j_1+1)}x_{1(j_1+1)} - \ldots - a_{1n}x_n + b_1 \\ x_{j_2} = -a_{2(j_2+1)}x_{1(j_2+1)} - \ldots - a_{2n}x_n + b_2 \\ \vdots \\ x_{j_k} = -a_{k(j_k+1)}x_{1(j_k+1)} - \ldots - a_{kn}x_n + b_k \end{cases}$$

spełniającego założenia o rozwiązaniu ogólnym. A więc jest to rozwiązanie ogólne układu U. □

Zauważmy jeszcze, że jeśli na pewnym etapie procesu sprowadzania macierzy układu U do zredukowanej postaci schodkowej otrzymamy macierz posiadającą wiersz postaci $0 \ldots 0\, b$, gdzie $b \neq 0$, to odpowiadający jej układ (a więc i układ U) jest sprzeczny i dalsze operacje prowadzące do otrzymania wiersza $0 \ldots 0\, 1$ są niepotrzebne.

Metodę rozwiązywania układów równań liniowych opisaną w powyższym wniosku, a w zasadzie w dowodzie (i), nazywamy – po ewentualnym sformalizowaniu kolejności wykonywania operacji elementarnych na wierszach – **algorytmem eliminacji Gaussa** rozwiązywania układów równań. Będzie on przećwiczony na ćwiczeniach. O szczegółach formalnych i analizę dotyczącą zastosowań praktycznych warto skierować się na przykład do notatek do wykładu prof. Przemysława Kiciaka z Matematyki obliczeniowej.

* * *

Zakończyliśmy notatki do pierwszego wykładu. Poniżej znajdą Państwo uzupełnienie, dodatek, trivię, a także notatkę historyczną. W kolejnych notatkach do wykładów również będę zamieszczał takie materiały i mają one charakter nieobowiązkowy. Pozwolą one Państwu być może lepiej skorzystać z wykładanego materiału i zrozumieć go w nieco szerszym kontekście. Polecam je szczególnie osobom myślącym o wyborze potoku gwiazdkowego w drugim semestrze.

W pierwszej części – uzupełnieniu, poruszam wątki bezpośrednio związane z treścią wykładu i dające być może nową perspektywę na omawiane pojęcia. Gdybyśmy dysponowali czterema godzinami wykładu w tygodniu, materiał ten włączony byłby do zajęć (po odpowiednim poszerzeniu). Elementy materiału przedstawianego w tej części będą się pojawiały na ćwiczeniach oraz w dalszej części wykładu.

W drugiej części – dodatku, poruszam zagadnienia bardziej zaawansowane, odnoszące się do ważnych motywów pojawiających się na wykładzie. We wprowadzeniu do wykładu wspominałem, że pojawiające się na nim metody będą miały duże znaczenie w dalszych Państwa studiach. Często poruszane zagadnienia nie będą ściśle związane z algebrą liniową.

W trzeciej części – trivii, poruszam tematy luźno związane z tematem i mające raczej charakter popularyzatorski. Czasem są to sprawy istotne, a czasem – po prostu osobliwe.

W części historycznej będę się starał krótko wskazywać na osoby i wydarzenia związane z rozwojem danego pojęcia. Historia algebry liniowej, jak i cała historia (matematyki), są niezwykle skomplikowane, a ja nie jestem specjalistą w tym zakresie. Osoby zainteresowane pogłębionym studium historii matematyki odsyłam do stosownego wykładu. Nie zawsze będziemy zresztą mówić o historiach bardzo odległych, a czasem część historyczna będzie platformą do przedstawienia jakiegoś znanego problemu matematycznego.

Uzupełnienie. Geometria układów równań (wprowadzenie)

Wróćmy do rozważanego na zajęciach układu równań i rozwiążmy go bez korzystania z metody opisanej na wykładzie.

$$\begin{cases} x + 2y + 3z = 1 \\ 4x + 5y + 6z = 1 \\ 7x + 8y + 9z = 1 \end{cases}. \qquad (\dagger)$$

Naturalną szkolną metodą jest wyznaczenie jednej z niewiadomych z równania pierwszego i dokonanie podstawienia. Mamy zatem $x = 1-2y-3z$, co po wstawieniu do drugiego i trzeciego równania daje nam:

$$\begin{cases} x = 1 - 2y - 3z \\ 4(1-2y-3z) + 5y + 6z = 1 \\ 7(1-2y-3z) + 8y + 9z = 1 \end{cases}.$$

Po uporządkowaniu wyrażeń algebraicznych w drugim i trzecim równaniu dostajemy zatem:

$$\begin{cases} x = 1 - 2y - 3z \\ -3y - 6z = -3 \\ -6y - 12z = -6 \end{cases}.$$

Widzimy, że tak naprawdę rozwiązać mamy układ równań postaci:

$$\begin{cases} x = 1 - 2y - 3z \\ y + 2z = 1 \\ 0 = 0 \end{cases}.$$

Zachowaliśmy trzeci wiersz mając w pamięci metodę macierzową. Możemy wyznaczyć y z drugiego równania i do obydwu stron trzeciego równania dodać z. Otrzymujemy wówczas układ postaci:

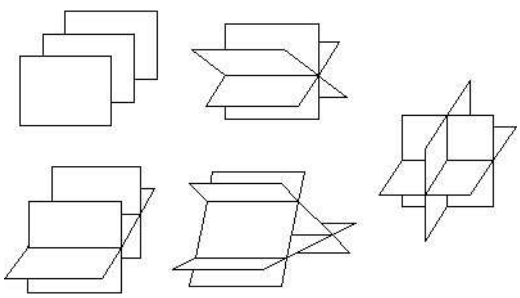
$$\begin{cases} x = 1 - 2(1-2z) - 3z = -1 + z \\ y = 1 - 2z \\ z = z \end{cases}.$$

Jesteśmy teraz w stanie odczytać wszystkie rozwiązania. Są to trójki liczb rzeczywistych postaci:

$$(-1+z, 1-2z, z), \text{ gdzie } z \in \mathbb{R}.$$

Gdybyśmy mieli zobrazować opisany problem w trójwymiarowej przestrzeni kartezjańskiej i na dobry początek wyznaczyć w niej zbiór trójek (x, y, z), których współrzędne spełniają równanie $x+2y+3z = 1$, to co Państwa zdaniem otrzymamy?

Prostą? – daje się usłyszeć nieśmiałe przypuszczenie. Nie, nie otrzymamy prostej! Wiem, że intuicja szkolna jest trudna do przełamania, ale proszę rozważyć ten sam problem dla warunku $x = 0$. Oczywiście zbiór punktów w przestrzeni trójwymiarowej o pierwszej współrzędnej zerowej to płaszczyzna wyznaczona przez osie OY i OZ. Także równanie $x + 2y + 3z = 1$ jest równaniem płaszczyzny w tej przestrzeni.



Intersection of lines and planes by Dan Sunday, http://geomalgorithms.com

A zatem rozumiecie już Państwo geometryczną naturę naszego problemu? Chodziło o wyznaczenie miejsca geometrycznego przecięcia trzech płaszczyzn opisanych równaniami $x+2y+3z = 1$, $4x+5y+6z = 1$ oraz $7x+8y+9z = 1$. Czy widzicie Państwo którą z powyższych konfiguracji reprezentuje sobą zbiór rozwiązań?

Niestety w szkole obecnie rzadko mówi się o postaci parametrycznej prostej, ale zbiór punktów stanowiących rozwiązania układu (†), czyli zbiór trójek $(-1+z, 1-2z, z)$, gdzie $z \in \mathbb{R}$, jest niczym innym tylko prostą. Łatwo to zobaczyć. Prosta ta złożona jest ze wszystkich punktów postaci $(-1,1,0)+t\cdot\vec{v}$, gdzie v jest wektorem postaci $[1,-1,1]$. Opisy tego rodzaju sformalizujemy w drugim semestrze.

Interpretacja geometryczna układu równań zaprezentowana wyżej stanowi swego rodzaju **wierszowy obraz** tego układu – patrzymy na każde równanie osobno, interpretujemy jego rozwiązania jako podzbiory przestrzeni trójwymiarowej i szukamy części wspólnej tych zbiorów. Przejdźmy teraz do odrobinę mniej intuicyjnego, **kolumnowego obrazu** opisującego układ (†). Mianowicie, przepiszmy ten układ w następującej postaci:

$$x\begin{bmatrix}1\\4\\7\end{bmatrix}+y\begin{bmatrix}2\\5\\8\end{bmatrix}+z\begin{bmatrix}3\\6\\9\end{bmatrix}=\begin{bmatrix}1\\1\\1\end{bmatrix}. \qquad (\heartsuit)$$

Widzimy, że naszym celem jest teraz znalezienie takich liczb rzeczywistych x, y, z, aby pewna kombinacja wektorów $(1,4,7)$,(2,5,8), $(3,6,9)$ stała się wektorem $(1,1,1)$. Ta nowa interpretacja geometryczna jest, co może być uznane za zaskoczenie, niezwykle istotna. Dlaczego? Przede wszystkim dlatego, że pozwala nam dużo powiedzieć o strukturze rozwiązań układu (†), o czym przekonamy się dalej. Po drugie, interpretacja ta jest wartościowa, bo pozwala nam zobaczyć na ile istotny jest wybór tego, a nie innego układu współczynników. Zmiana wektora $(1,1,1)$ na inny może całkowicie zmienić odpowiedź na pytanie czy dany układ ma rozwiązanie, czy nie. Na przykład układ postaci:

$$x\begin{bmatrix}1\\4\\7\end{bmatrix}+y\begin{bmatrix}2\\5\\8\end{bmatrix}+z\begin{bmatrix}3\\6\\9\end{bmatrix}=\begin{bmatrix}0\\0\\1\end{bmatrix}.$$

nie może mieć rozwiązania, bo wektor $(0,0,1)$ nie leży na tej samej płaszczyźnie w przestrzeni, co wektory $(1,4,7)$, $(2,5,8)$. $(3,6,9)$. To nie jest oczywiste, bo nie widać wcale, że te trzy wektory muszą leżeć na płaszczyźnie i to takiej, w której nie ma wektora $(0,0,1)$. Jeśli jednak przepiszemy w języku wektorowym operacje wykonane na wyjściowym układzie, wówczas okaże się, że jest to jasne. Zauważmy, że układ

$$\begin{cases} x &= 1-2y-3z\\ y+2z &= 1\\ 0 &= 0\end{cases},$$

równoważny układowi (†) przepisuje się w języku wektorów jako:

$$x\begin{bmatrix}1\\0\\0\end{bmatrix}+y\begin{bmatrix}2\\1\\0\end{bmatrix}+z\begin{bmatrix}3\\2\\0\end{bmatrix}=\begin{bmatrix}1\\1\\0\end{bmatrix}. \qquad (\spadesuit)$$

To, co jest kluczowe to fakt, że lewe strony warunków $(\heartsuit)$ oraz $(\spadesuit)$ opisują ten sam zbiór wektorów. Teraz jest już jasne, że $(0,0,1)$ nie jest jego elementem,

Badanie **kombinacji liniowych** wektorów (definicja na kolejnych wykładach) postaci $xv_1+yv_2+zv_3$ oraz ich geometrycznej natury jest centralnym elementem naszego wykładu. Aby to zrozumieć, na razie na poziomie wprowadzonych na wykładzie pojęć, prześledźmy rozwiązanie następującego zadania.

Zadanie. Ciągi $(1,2,3,4,5)$ oraz $(2,0,0,1,0)$ są rozwiązaniami pewnego jednorodnego układu równań liniowych o współczynnikach rzeczywistych. Opisz wszystkie rozwiązania tego układu zakładając, że jego macierz po sprowadzeniu do postaci schodkowej ma trzy schodki.

W kontekście tego zadania rozważmy trzy pytania:

(1) Które jednorodne równania liniowe mają wśród rozwiązań ciągi $(1,2,3,4,5)$ oraz $(2,0,0,1,0)$? Które ich **układy** zadają macierz o 3 schodkach?

(2) Jakie rozwiązania musi mieć każdy znaleziony **układ**, poza tymi dwoma?

(3) A jakich rozwiązań nie może mieć żaden znaleziony wyżej **układ**?

Odpowiedź na pytanie pierwsze jest całkiem prosta: takie równania można po prostu wypisać. Są to równania postaci $a_1x_1 + a_2x_2 + a_3x_3 + a_4x_4 + a_5x_5 = 0$ takie, że:

$$\begin{cases} a_1 + 2a_2 + 3a_3 + 4a_4 + 5a_5 & = 0 \\ 2a_1 + a_4 & = 0 \end{cases}. \quad (\ddagger)$$

Wniosek: zbiór **wszystkich równań liniowych jednorodnych**, które mają (nie tylko) rozwiązania $(1, 2, 3, 4, 5)$ oraz $(2, 0, 0, 1, 0)$ ma *strukturę* zbioru rozwiązań układu równań, czyli opisany jest przez piątki $(a_1, a_2, a_3, a_4, a_5)$ spełniające ($\ddagger$).

Wystarczy wziąć 3 rozwiązania układu ($\ddagger$) postaci

$$(*, *, *, *, *), \quad (0, *, *, *, *), \quad (0, 0, *, *, *)$$

i dostajemy układ 3 równań tworzący macierz o 3 schodkach, będącą macierzą układu, którego rozwiązaniami są między innymi $(1, 2, 3, 4, 5)$ oraz $(2, 0, 0, 1, 0)$. Pokażemy, że **każdy wybrany w ten sposób układ ma ten sam zbiór rozwiązań!** Oczywiście, są i inne układy o współczynnikach spełniających ($\ddagger$), które mają ten sam zbiór rozwiązań, ale są i takie, które mają *więcej* rozwiązań, np. każdy układ złożony z pojedynczego równania jednorodnego spełniającego warunki z ($\ddagger$).

Przechodzimy do pytania drugiego. Do zbioru rozwiązań **dowolnego** układu równań jednorodnych, który ma rozwiązania $(1, 2, 3, 4, 5)$ oraz $(2, 0, 0, 1, 0)$ należą też:

(a) każda *wielokrotność* $\lambda_1 \cdot (1, 2, 3, 4, 5) = (\lambda_1 \cdot 1, \lambda_1 \cdot 2, \lambda_1 \cdot 3, \lambda_1 \cdot 4, \lambda_1 \cdot 5)$,

(b) każda *wielokrotność* $\lambda_2 \cdot (2, 0, 0, 1, 0) = (\lambda_2 \cdot 2, \lambda_2 \cdot 0, \lambda_2 \cdot 0, \lambda_2 \cdot 1, \lambda_2 \cdot 0)$,

(c) kazda *suma* $\lambda_1 \cdot (1, 2, 3, 4, 5) + \lambda_2 \cdot (2, 0, 0, 1, 0)$ tych *wielokrotności*, czyli

$$(\lambda_1 \cdot 1 + \lambda_2 \cdot 2, \lambda_1 \cdot 2 + \lambda_2 \cdot 0, \lambda_1 \cdot 3 + \lambda_2 \cdot 0, \lambda_1 \cdot 4 + \lambda_2 \cdot 1, \lambda_1 \cdot 5 + \lambda_2 \cdot 0).$$

Uwaga: poza $\lambda_1 = \lambda_2 = 0$, nigdy nie mamy $\lambda_1 \cdot (1, 2, 3, 4, 5) = \lambda_2 \cdot (2, 0, 0, 1, 0)$.

A jakich rozwiązań nie może mieć żaden wyznaczony wcześniej układ ($\ddagger$)? Macierz każdego rozważanego układu ma po sprowadzeniu do postaci schodkowej 3 schodki, a zatem rozwiązanie *zależy* od dwóch parametrów. Czy dowolny taki układ może mieć rozwiązania poza zbiorem:

$$\{\lambda_1 \cdot (1, 2, 3, 4, 5) + \lambda_2 \cdot (2, 0, 0, 1, 0)\,,\ \lambda_1, \lambda_2 \in \mathbb{R}\}?$$

Gdyby takie rozwiązanie $(s_1, s_2, s_3, s_4, s_5)$ istniało, to także każda wielokrotność postaci $\lambda_3(s_1, s_2, \ldots, s_5)$ byłaby rozwiązaniem. A zatem także:

$$\lambda_1 \cdot (1, 2, 3, 4, 5) + \lambda_2 \cdot (2, 0, 0, 1, 0) + \lambda_3(s_1, s_2, \ldots, s_5).$$

Kluczowe – tak być nie może, bo teraz rozwiązanie zależy od 3 parametrów. Celem dalszej części wykładu jest dokładne zrozumienie zdania powyżej.

Niezbyt formalne wnioski z tych rozważań mają charakter wybitnie geometryczny

- zbiór rozwiązań jednorodnego układu równań liniowych jest zamknięty na branie *sumy* i *wielokrotności*,
- zbiór równań jednorodnych o (jakimś ustalonym) wspólnym podzbiorze zbiorów rozwiązań też jest zamknięty na branie *sumy* i *wielokrotności*,
- gdy zbiór rozwiązań ma 1 parametr: rozwiązania zachowują się jak wektory o wspólnym początku, których końce leżą na jednej prostej,
- gdy zbiór rozwiązań ma 2 parametry, rozwiązania zachowują się jak wektory o wspólnym początku, których końce leżą na jednej płaszczyźnie,
- układy równań o wspólnym podzbiorze zbiorów rozwiązań mają *strukturę geometryczną*!

Obserwacje te można uogólnić na niejednorodne układy równań. Zachęcam Czytelnika do pokazania, że jeśli dane są rozwiązania u_1, u_2 niejednorodnego układu równań o współczynnikach w $\mathbb{R}$, to również każdy element postaci $su_1 + (1 - s)u_2, s \in \mathbb{R}$ jest rozwiązaniem tego układu. Z geometrycznego punktu widzenia zbiór punktów $sA + (1 - s)B$ opisuje prostą przechodzącą przez punkty A, B. To jest właśnie *liniowość* układu równań. Zbiory rozwiązań takich układów są *liniowe*, *płaskie* i mają strukturę, którą będziemy powoli poznawać.

Dodatek. Nieliniowe układy równań

Pokazaliśmy, że rozwiązywanie układów równań liniowych polega na zastępowaniu ich przez inne układy o tym samym zbiorze rozwiązań. Zachowanie zbioru rozwiązań gwarantuje stosowanie operacji elementarnych. Czy podobna metoda da się stosować dla innych typów układów równań? Rozważmy układ:

$$\begin{cases} x^2 + yz + x & = 0 \\ y^2 + xz + y & = 0 \\ z^2 + xy + z & = 0 \end{cases}.$$

Układ ten nie jest układem równań liniowych, ale wydaje się, że nic nie stoi na przeszkodzie by wykonywać na nim wszystkie opisane wcześniej operacje elementarne. Czy ich stosowanie zmienia zbiór rozwiązań? Po odjęciu drugiego równania od pierwszego oraz trzeciego równania od drugiego dostajemy (co chyba jasne) układ równoważny postaci:

$$\begin{cases} (x-y)(x+y-z+1) & = 0 \\ (y-z)(y+z-x+1) & = 0 \\ z^2 + xy + z & = 0 \end{cases}.$$

Chwila skrupulatnego rozważania tych warunków (nic trudnego) prowadzi do trójek (x, y, z) rozwiązań:

$$(0,0,0), (-1,0,0), (0,-1,0), (0,0,-1), (-1/2,-1/2,-1/2).$$

Czy rozwiążemy w ten sposób każdy układ równań wielomianowych (na razie rozumianych intuicyjnie)?
Czy ma sens mówienie o macierzy współczynników tego układu? Jak ją określić?
Czy jest jakiś analog postaci schodkowej macierzy związanych z takimi układami?
Czy są jakieś specyficzne dla wielomianów „operacje elementarne”, które nie zmieniają zbioru rozwiązań?
Czym są „zmienne niezależne” w przypadku istnienia nieskończenie wielu rozwiązań układu równań?

Konkretne odpowiedzi na powyższe pytania związane są z działem algebry zwanym teorią pierścieni oraz odkryciem Buchbergera z 1965 roku dotyczącym wyznaczania tak zwanych baz Gröbnera. Kto by chciał poczytać więcej o tym algorytmie i jego związku z algorytmem Gaussa, może zajrzeć pod poniższe adresy:

`https://www.math.usm.edu/perry/Research/GaussToGroebner.pdf`
`https://www.theoremoftheday.org/MathsStudyGroup/Buchberger.pdf`

W tym miejscu damy Czytelnikowi jedną tylko intuicję, która pozwala zobaczyć w nieco innym świetle sam algorytm Gaussa i sens wyznaczania postaci schodkowej (stąd też obecność tego tematu w tym miejscu). Załóżmy, że ograniczamy się do rozważania układów równań postaci:

$$\begin{cases} F_1(x,y) & = 0 \\ & \dots \\ F_s(x,y) & = 0 \end{cases},$$

gdzie $F_i(x, y)$ są, dla $1 \leqslant i \leqslant s$, wielomianami zmiennych x, y, czyli formalnymi sumami jednomianów postaci $c \cdot x^m y^n$, $m, n \geqslant 0$, $c \in \mathbb{R}$, zaś $m, n \in \mathbb{Z}$ (liczby całkowite!). Jeśli każdy z takich jednomianów spełnia warunek $m + n = 1$, otrzymamy po prostu układy równań liniowych (o dwóch niewiadomych).

W przypadku układu równań liniowych zmienne oznaczaliśmy symbolami $x_1, x_2, \dots$ Oznacza to – choć na wykładzie nie zwracaliśmy na to uwagi – że zmienne tworzą **zbiór uporządkowany**. Macierz współczynników układu, postać schodkowa, opis rozwiązań – wszystko to zależy od **kolejności zmiennych**, jaką przyjmujemy (choć nie zmienia się natura geometryczna zbioru rozwiązań). Idea jest następująca: w przypadku wielomianów porządkować będziemy nie tylko same zmienne, ale cały zbiór jednomianów.

Zamiast x pisać będziemy x_1 oraz zamiast y pisać będziemy x_2. Chcemy przez to podkreślić, że x jest w porządku zmiennych „wcześniejsza”. Wprowadzamy następnie zasadę porządkowania jednomianów. Przez stopień $deg(x_1^m x_2^n)$ jednomianu $x_1^m x_2^n$ rozumiemy sumę $m + n$. Ustalamy zasadę, że zawsze „wcześniejszy” (czyli większy) jest jednomian wyższego stopnia. A zatem $x_1^4 x_2^2 > x_1^5$, $x_1 x_2^6 > x_1^2 x_2$ itd. Jeśli natomiast stopnie dwóch jednomianów są identyczne, wcześniejszy jest jednomian z wyższą potęgą przy wcześniejszej zmiennej. A więc, na przykład, $x_1^4 x_2^3 > x_1^3 x_2^4$, chociaż $deg(x_1^4 x_2^3) = deg(x_1^3 x_2^4) = 7$.

Układ równań napisany na początku naszego dodatku jest, przy założeniu porządku $x > y > z$, napisany tak, że kolejne składniki są coraz mniejsze w „porządku jednomianowym". Algorytm Buchbergera proponuje – do pewnego stopnia – wykonywanie podobnej procedury, co algorytm Gaussa. Chodzi o zastępowanie jednych wielomianów innymi i dążenie do prostszego układu. Możemy sobie już wyobrazić macierz układu równań wielomianowych. Co z operacjami elementarnymi? Wprowadza się (pozornie) dodatkową „operację elementarną" na dwóch wielomianach – tak zwany S-wielomian. Jak go określić?

Dla dwóch jednomianów $x_1^{m_1}x_2^{n_1}$ oraz $x_1^{m_2}x_2^{n_2}$ rozważa się tzw. najmniejszą wspólną wielokrotność (zbieżność użytego nazewnictwa z tym występującym w teorii podzielności liczb całkowitych nie jest przypadkowa). Nazwiemy ją $NWW(x_1^{m_1}x_2^{n_1}, x_1^{m_2}x_2^{n_2})$ i jest to po prostu jednomian $x_1^{\max\{m_1,m_2\}}x_2^{\max\{n_1,n_2\}}$.

Dla każdego wielomianu f przez $LT(f)$ oznaczamy największy jednomian – w określonym wyżej porządku – który występuje w f. Na przykład $LT(2x^3y^4 + y^6) = 2x^3y^4$ (uwaga – nie samo x^3y^4, ale $2x^3y^4$).

Dla wielomianów f, g zmiennych x, y, przez $S(f, g)$ rozumiemy wielomian postaci:

$$S(f,g) = \frac{NWW(LT(f), LT(g))}{LT(f_1)} \cdot f_1 - \frac{NWW(LT(f), LT(g))}{LT(f_2)} \cdot f_2.$$

Przykład 2. *Rozważmy przecięcie dwóch elips (proszę uwierzyć mi na słowo, że są to elipsy) postaci:*

$$\begin{cases} 2x^2 + y^2 - 4x - 4y + 3 & = 0 \\ x^2 + 3y^2 - 2x - 12y + 9 & = 0 \end{cases}.$$

Niech $f = 2x^2 + y^2 - 4x - 4y + 3$ oraz $g = x^2 + 3y^2 - 2x - 12y + 9$. Wówczas przyjmując $x > y$ i respektując opisany wyżej porządek jednomianowy dostajemy $LT(f) = 2x^2, LT(g) = x^2$, a zatem widzimy, że:

$$S(f,g) = \frac{-5}{2}y^2 + 10y - \frac{15}{2}.$$

Czy Czytelnik widzi co się stało? Jednomian wiodący S-wielomianu $S(f, g)$ jest mniejszy – w porządku jednomianowym – niż jednomian wiodący każdego z wielomianów f, g. Czy widać tu analogię do uzyskiwania postaci schodkowej przez algorytm Gaussa, zwany często algorytmem **eliminacji** Gaussa? Czy widać, że wykonanie S-wielomianu na dwóch równaniach liniowego układu równań i zastąpienie jednego z równań owym S-wielomianem (dwóch zmiennych – biorąc pod uwagę nasze definicje – ale po odpowiednim uogólnieniu – dowolnym) jest równoważne ze zwykłą operacją elementarną rozważaną na wykładzie?

Czy Czytelnik widzi w jaki sposób stosowanie S-wielomianu prowadzi do zastąpienia wyjściowego układu równań wielomianowych układem prostszym? Jak należy kontynuować ten algorytm i do jakich rozwiązań może on prowadzić? Poszukiwanie odpowiedzi polecam już Państwu. Polecam wskazane wyżej źródła.

Ps. Rozważane elipsy przecinają się w czterech punktach. Dowód wymaga jedynie matematyki szkolnej.

* * *

Zapomniałbym wspomnieć: baz Gröbnera używa się do naprawdę ciekawych rzeczy (choćby w robotyce). Polecam dwa anglojęzyczne źródła (im szybciej Państwo zaczną czytać matematykę w tym języku, tym lepiej). Na razie mogą być one dla Państwa trudne w lekturze, ale za jakiś czas będzie można wrócić do tej lektury. A może warto przejść się do kogoś na konsultacje i poprosić o wyjasnienie trudniejszych fragmentów?

- Antoine Nectoux, *Map colouring and Gröbner Bases*
- Elizabeth Arnold, Stephen Lucas, and Laura Taalman, *Gröbner Basis Representations of Sudoku*

Aby znaleźć te teksty wystarczy wpisać tytuły w Google.

Polecam również dwugodzinny wykład prof. Przemysława Koprowskiego (Uniwersytet Śląski) dotyczący baz Gröbnera i ich zastosowań (https://youtu.be/vdmyrbNqRlY). Jest to znakomite, ścisłe wprowadzenie do tej tematyki, z licznymi przykładami zastosowań.

Trivia. Kwadraty magiczne

Pierwsza trivia dotyczy pięknego tematu, mianowicie kwadratów magicznych. Jest wiele poważnych tekstów, zarówno popularyzatorskich, jak i naukowych dotyczących tych obiektów i problemów z nimi związanych. Znane były one od starożytności, a fascynowały między innymi samego Eulera, który poświęcił im kilka swoich artykułów. Obecnie można znaleźć wiele ich tłumaczeń w tym na język angielski. np.

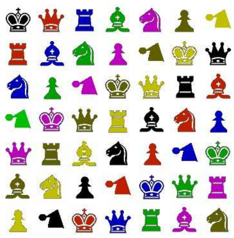
- E530 (http://eulerarchive.maa.org/docs/translations/E530.pdf),
- E795[3] (https://arxiv.org/pdf/math/0408230.pdf).

Czym więc są owe kwadraty? Może zamiast definicji podam problem 36 oficerów, badany przez Eulera. Mamy mianowicie 36 oficerów o sześciu różnych stopniach, wziętych z sześciu różnych oddziałów i próbujemy ustawić ich w kwadracie tak, by w każdym wierszu i każdej kolumnie tego kwadratu stało sześciu oficerów z innego oddziału i różnych stopni.

Oto stosowny obrazek pokazujący oficerów, ustawionych na razie zgodnie z przynależnością do oddziału.

Źródło: http://www.ams.org/publicoutreach/feature-column/fcarc-latinii1

Zachęcam do eksperymentu i próby dokonania żądanego ustawienia. Okaże się, że nie bardzo chce się to udać. Eulerowi też to nie wychodziło. A problem był o tyle frustrujący, że analogiczne problemy 25 i 49 oficerów daje się gładko rozwiązać:



Źródło: http://www.ams.org/publicoutreach/feature-column/fcarc-latinii1

Euler rozpoznaje w badanym zagadnieniu problem algebraiczny, znany zresztą wcześniej. Załóżmy, że każdemu oficerowi nadamy plakietkę a^b, gdzie a oznaczać będzie stopień, a b – numer oddziału, a potem zapomnimy o numerach oddziału i popatrzymy tylko na stopnie, to rozkład wyżej ma postać:

$$\begin{array}{ccccccc}
7 & 6 & 5 & 4 & 3 & 2 & 1 \\
5 & 4 & 3 & 2 & 1 & 7 & 6 \\
3 & 2 & 1 & 7 & 6 & 5 & 4 \\
1 & 7 & 6 & 5 & 4 & 3 & 2 \\
6 & 5 & 4 & 3 & 2 & 1 & 7 \\
4 & 3 & 2 & 1 & 7 & 6 & 5 \\
2 & 1 & 7 & 6 & 5 & 4 & 3
\end{array}$$

Dla Eulera kwadratem magicznym jest tablica liczb rozmiaru $n \times n$ taka, że w każdym wierszu, każdej kolumnie i na obydwu przekątnych sumy wpisanych liczb są równe. A jakie liczby wpisujemy? W przypadku wyżej: siedem zestawów od 1 do 7. Czasem rozważa się kwadraty, w które wpisuje się kolejne liczby naturalne (tzw. normalne kwadraty magiczne), a czasem półmagiczne (bez warunku na przekątne) itd. Problem oficerów rozwiązany został dopiero w 1901 roku przez matematyka-amatora Gastona Tarry'ego.

[3]Pewnie ciekawi Państwa co znaczą te liczby? W latach 1910-1913 szwedzki matematyk Gustav Eneström dokonał gruntownych badań nad dziełami Eulera i z nieprzebranych archiwów publikacji oraz zbiorów notatek wyłonił 866 pozycji: książek, artykułów i istotnych listów, którym przydzielił symbole od E1 do E866.

Aby przybliżyć się nieco do materiału z wykładu zajmijmy się magicznymi kwadratami rozmiaru 3×3 o wyrazach rzeczywistych. Możemy je oczywiście utożsamiać z macierzami. Oto przykłady:

$$\begin{bmatrix} 2 & 9 & 4 \\ 7 & 5 & 3 \\ 6 & 1 & 8 \end{bmatrix}, \quad \begin{bmatrix} -1 & 5/2 & 0 \\ 3/2 & 1/2 & -1/2 \\ 1 & -3/2 & 2 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 & -1 \\ -2 & 0 & 2 \\ 1 & 0 & -1 \end{bmatrix}.$$

Ile jest macierzy magicznych? Powyższe przykłady sugerują, że jest ich nieskończenie wiele. Zauważmy, że wyznaczenie macierzy magicznej równoważne jest problemowi rozwiązania układu równań liniowych. Aby rozstrzygnąć czy macierz:

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

reprezentuje kwadrat magiczny należy nałożyć następujące warunki na jej wyrazy:

$$\begin{cases} a_{21} + a_{22} + a_{23} - a_{11} - a_{12} - a_{13} & = 0 \\ a_{31} + a_{32} + a_{33} - a_{11} - a_{12} - a_{13} & = 0 \\ a_{11} + a_{21} + a_{31} - a_{11} - a_{12} - a_{13} & = 0 \\ a_{12} + a_{22} + a_{32} - a_{11} - a_{12} - a_{13} & = 0 \\ a_{13} + a_{23} + a_{33} - a_{11} - a_{12} - a_{13} & = 0 \\ a_{11} + a_{22} + a_{33} - a_{11} - a_{12} - a_{13} & = 0 \\ a_{31} + a_{22} + a_{13} - a_{11} - a_{12} - a_{13} & = 0 \end{cases}.$$

czyli rozwiązać układ równań, którego macierz ma aż 7 wierszy i 9 kolumn odpowiadających zmiennym

$$a_{11}, a_{12}, a_{13}, a_{21}, a_{22}, a_{23}, a_{31}, a_{32}, a_{33}$$

postaci

$$\left[\begin{array}{ccccccccc|c} -1 & -1 & -1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ -1 & -1 & -1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & -1 & -1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ -1 & 0 & -1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ -1 & -1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & -1 & -1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ -1 & -1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \end{array}\right].$$

Idę o zakład, że „schodkowanie” tej macierzy nie wydałoby się Państwu przyjemnością. Od ilu parametrów zależą rozwiązania tego układu? Nie bardzo widać rozwiązanie. Jest jednak sprytniejsza droga: skorzystamy z twierdzenia o rozwiązaniach układów jednorodnych i niejednorodnych. Proszę zauważyć, że jeśli mamy macierz magiczną (a_{ij}) taką, że suma wyrazów w każdym wierszu, kolumnie na przekątnych to $3S$, to macierz o wyrazach $a_{ij} - S$ również jest magiczna, a nawet 0-magiczna, bo suma wyrazów w każdym jej wierszu, kolumnie i na przekątnych to 0. Intuicja podpowiada zatem, że liczba parametrów potrzebna do opisania wszystkich macierzy magicznych jest o 1 większa niż liczba parametrów służących do opisu macierzy 0-magicznych, bo liczba parametrów potrzebna do opisu macierzy 0-magicznych jest taka sama, jak do opisu macierzy t-magicznych (układ jednorodny vs. niejednorodny). Skoro t się zmienia, to mamy (na poziomie intuicji) dodatkowy parametr. Zobaczmy, że opis macierzy t-magicznych wygląda przejrzyściej. Można go dokonać przez rozwiązanie układu o macierzy:

$$\left[\begin{array}{ccccccccc|c} 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & t \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & t \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & t \\ 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & t \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & t \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & t \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & t \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & t \end{array}\right].$$

Znacznie łatwiej analizować powyższą macierz 8×9, czy ogólnie macierz rozmiaru $(2n+2) \times n^2$, uzyskiwaną dla układu opisującego macierze t-magiczne rozmiaru n. Tu już coś widać. Suma pierwszych trzech wierszy minus suma dwóch kolejnych daje wiersz szósty, więc po wyschodkowaniu jest co najmniej jeden wiersz zerowy... i jak się okazuje nie ma innych. Rozwiązanie zależy od dwóch parametrów. A zatem wszystkie macierze rozmiaru 3×3 opisać można za pomocą trzech parametrów. Jeszcze do tego wrócimy.

Notka historyczna. Kiedy to się zaczęło?

Historia tak szerokiego tematu jak rozwiązywanie układów równań liniowych jest tak szalenie skomplikowana, że w zasadzie nie jest możliwe choćby krótkie jej nakreślenie. Osobom zainteresowanym stosunkowo przystępnym tekstem polecam publicznie dostępny[4] artykuł *How ordinary elimination became Gaussian elimination* autorstwa Josepha Greara (Historia Mathematica 38 (2011), 163-218). Pozwolę sobie przytoczyć tu krótki wstęp, wybierając fragmenty przydatne, mam nadzieję, dla ogólnego spojrzenia na matematykę i jej historię. Niecierpliwym polecam tekst: https://www.fuw.edu.pl/ kostecki/histmat.pdf.

Problemy sprowadzające się do rozwiązywania układów równań znane były już w starożytnej Mezopotamii oraz w Chinach. Według van der Waerdena[5] pierwsza systematyczna metoda rozwiązywania układów trzech równań z trzema niewiadomymi opisana została w rozdziale ósmym monumentalnego chińskiego dzieła „Dziewięć rozdziałów sztuki matematycznej", stanowiącego kompendium ówczesnej wiedzy matematycznej, powstałego pomiędzy 200 r. p. n. e., a 100 r. p. n. e. Przypomina ona metodę eliminacji poznaną przez nas na wykładzie, choć nie opiera się oczywiście o pojęcie macierzy. I oczywiście nie jest systematyczna w sensie, w jakim dziś myślimy o systematyczności. Po prostu jest tam rozwiązany problem praktyczny, sprowadzający się do układu równań, i który rozwiązany jest w sposób systematyczny, sprowadzający się do metody eliminacji. Szczegółowo jest to opisane u van der Waerdena, ale też w innych miejscach. Generalna uwaga odnośnie matematyki egipskiej czy babilońskiej jest taka, że skoncentrowana była ona wokół problemów praktycznych i procedur ich rozwiązywania. Słynny Donald Knuth dopatrywał się w tekstach babilońskich początków informatyki (a na pewno algorytmiki).

Używanie tablic liczbowych do zapisu pewnych układów liczb i związanych z nimi układów równań również znane jest od starożytności, zwłaszcza w kontekście kwadratów magicznych, którym przypisywano magiczną (a przynajmniej ezoteryczną) moc. Motyw owych niezwykłych tablic obecny jest w sztuce, np. w słynnej *Melancholii* Durera (1514) czy ten umieszczony przez Gaudiego na fasadzie Męki Pańskiej w Sagrada Familia (w której znajdziemy jeszcze sporo matematyki – w przyszłości).

Jak to zwykle bywa w historii matematyki, metoda eliminacji prowadząca do rozwiązywania układów równań, którą tradycyjnie przypisuje się Gaussowi[6], sam Książe Matematyków uważał za dobrze znany folklor (podobnie myślał o wielu innych tematach, na przykład o kwaternionach). Idea eliminacji pojawiała się w opublikowanych wbrew woli Newtona notatkach do prowadzonego w latach 1673-1687 wykładu z algebry w Cambridge, wydanych uroczyście w 1707 roku jako *Arithmetica Universalis*. Newton uważał, że publikowanie takich tekstów po 20 latach (a dalej tłumaczenie ich na różne języki) może prowadzić kogoś do wniosku, że są to jego najnowsze wyniki! Tymczasem w Anglii wydano w latach 1650-1750 kilkadziesiąt podręczników do algebry. Klasyczna złośliwość historii sprawiła, że, zgodnie z opinią historyków, ów podręcznik do algebry stał się jednym z najszerzej znanych i wpływowych dzieł matematycznych Newtona. Wielki Euler krytykował metodę eliminacji jako „niepolecaną", jego następca Legendre nazywał „zwyczajną". Gauss potraktował eliminację jak znane wszystkim (common) narzędzie. Praca z 1809 roku dotyczy znacznie bardziej skomplikowanych zagadnień, związanych między innymi z metodą najmniejszych kwadratów oraz zagadnieniami triangulacji istotnymi dla tworzenia map czy problemami astronomicznymi (młodemu Gaussowi sławy współczesnych mu uczonych nie przyniosło odkrycie matematyczne, tylko wyliczenie orbity „zagubionej" przez astronomów planetoidy Ceres).

Pojęcie macierzy zawdzięczamy Brytyjczykom: Arthurowi Cayleyowi i Jamesowi Josephowi Sylvesterowi. Drugi z nich użył tego pojęcia po raz pierwszy w 1850 roku. Pierwszy natomiast, siedem lat później, napisał pierwszą rozprawę o macierzach (*Treatise on the Theory of Matrices*). Jak się nietrudno domyśleć oznacza to, że eliminacja przypisywana wcześniejszym autorom odbyła się bez użycia pojęcia macierzy.

Kto zatem, i dlaczego, przypisał Gaussowi autorstwo metody eliminacji? Stało się to głównie za sprawą rozwoju maszyn liczących. Kluczowe było to, że algorytmy opisane w pracy Gaussa były po prostu niezwykle użyteczne dla wykorzystania praprzodków komputerów. Stosowano je, z nielicznymi modyfikacjami jeszcze do II Wojny Światowej. Notacja „klamrowa" wprowadzona przez Gaussa podkreślała kolejność zmiennych i była bardzo wygodna. Pojęcie „algorytmu eliminacji Gaussa" zastosował po raz pierwszy, jak się wydaje, Alan Turing, który przez dwa tygodnie rozwiązywał, z pomocą „komputera biurkowego" układ 18 równań w roku 1946. Po II WŚ, pojęcie to stopniowo wchodziło do programów nauczania.

[4]Zarówno na ArXiV: https://arxiv.org/pdf/0907.2397.pdf, jak i na stronach Elselviera.
[5]Geometry and Algebra in Ancient Civilizations, Springer-Verlag Berlin Heidelberg, Year: 1983
[6]Gauss, C. F., *Theoria Motus Corporum Coelestium in Sectionibus Conicis Solum Ambientium*, 1809.