4. Układy równań liniowych

Definicja 4.1. *Układem m równań liniowych n zmiennych nad ciałem k* nazywamy układ

$$(u) = \begin{cases} a_{11} \cdot x_1 + \cdots + a_{1n} \cdot x_n &= b_1 \\ a_{21} \cdot x_1 + \cdots + a_{2n} \cdot x_n &= b_2 \\ \vdots & \vdots \\ a_{m1} \cdot x_1 + \cdots + a_{mn} \cdot x_n &= b_m \end{cases}$$

gdzie $a_{ij} \in k$ dla każdego i, j.

Definicja 4.2. *Macierzą stowarzyszoną z układem* (u) nazywamy macierz

$$A_{(u)} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

zaś *macierzą rozszerzoną układu* (u) macierz

$$\widetilde{A}_{(u)} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} & b_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_{mn} \end{bmatrix}.$$

Zauważmy, że jeśli oznaczymy przez $X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ oraz $B = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$, to układ równań (u) możemy równoważnie zapisać w postaci

$$A_{(u)} \cdot X = B,$$

gdzie „$\cdot$” oznacza mnożenie macierzowe. Zbiór rozwiązań układu (u) będziemy oznaczać symbolem $\mathrm{rozw}(u)$. Oczywiście może się zdarzyć, że $\mathrm{rozw}(u) = \emptyset$.

Definicja 4.3. Niech $C = [c_{ij}]$ będzie macierzą. Element c_{ij} nazywamy *wiodącym*, o ile $c_{ij} \neq 0$ oraz $c_{is} = 0$ dla każdego $s < j$.

Przykład 4.4. Dla macierzy $\begin{bmatrix} 1 & 2 & 0 & -3 \\ 0 & 0 & e & 16 \\ 0 & \pi & 0 & 0 \end{bmatrix}$ elementami wiodącymi są 1, e i π.

Definicja 4.5. Macierz $C = [c_{ij}]$ o wymiarach $m \times n$ nazywamy *górno-schodkową*, o ile dla każdego współczynnika wiodącego c_{ij} i dla dowolnych $i \leqslant k \leqslant m$, $1 \leqslant l \leqslant j$ mamy $c_{kl} = 0$ (poza c_{ij} oczywiście).

Przykład 4.6. Macierzą górno-schodkową jest $\begin{bmatrix} 1 & \zeta & 0 & -3 \\ 0 & 0 & 1 & 16 \\ 0 & 0 & 0 & 1 \end{bmatrix}$, natomiast macierz z przykładu 4.4 nie jest górno-schodkowa.

Definicja 4.7. Macierz górno-schodkową nazywamy *zredukowaną* (odp. *całkowicie zredukowaną*), o ile wszystkie jej współczynniki wiodące są równe 1 (odp. wszystkie jej współczynniki wiodące są równe 1 oraz jeśli c_{ij} jest współczynnikiem wiodącym, to $c_{kj} = 0$ dla każdego $1 \leqslant k < i$.).

Przykład 4.8. Macierz z przykładu 4.6 jest zredukowana, lecz nie jest całkowicie zredukowana, natomiast macierz $\begin{bmatrix} 1 & \zeta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$ jest całkowicie zredukowana.

Definicja 4.9. *Operacjami elementarnymi* w macierzy M nazywamy następujace operacje:

1. zamiana wierszy miejscami w macierzy M;
2. przemnożenia jednego wiersza macierzy M przez skalar, dodanie go do drugiego i wstawienie wyniku w miejsce drugiego;
3. przemnożenie danego wiersza przez niezerowy skalar.

Twierdzenie 4.10. (o eliminacji Gaussa) Każdą macierz o współczynnikach w ciele k można sprowadzić do macierzy górno-schodkowej za pomocą operacji elementarnych.

Definicja 4.11. Jeśli M jest macierzą górno-schodkową, to symbolem $r(M)$ oznaczamy liczbę jej współczynników wiodących. Jeśli M jest dowolną macierzą, to na podstawie twierdzenia o eliminacji Gaussa można ją sprowadzić do postaci górno-schodkowej M'. Wówczas definiujemy $r(M) := r(M')$. Można pokazać, że liczba $r(M)$ nie zależy od wyboru macierzy M' i jest równa rzędowi macierzy M.

Twierdzenie 4.12. (o rozwiązywaniu układów równań liniowych) Przypuśćmy, że (u) jest układem m równań liniowych z n niewiadomymi, $\widetilde{A}_{(u)}$ macierzą rozszerzoną układu (u) oraz B całkowicie zredukowaną macierzą górno-schodkową stowarzyszoną z $\widetilde{A}_{(u)}$. Wówczas:

1. Układ (u) ma dokładnie jedno rozwiązanie ($|\text{rozw}(u)| = 1$) wtedy i tylko wtedy gdy $n = r(B)$ oraz żaden współczynnik wiodący nie stoi w $n+1$-wszej kolumnie macierzy.
2. Układ (u) ma nieskończenie wiele rozwiązań ($|\text{rozw}(u)| = \infty$) wtedy i tylko wtedy gdy $n > r(B)$ oraz żaden współczynnik wiodący nie stoi w $n+1$-wszej kolumnie.
3. Układ (u) nie posiada rozwiązań ($|\text{rozw}(u)| = 0$) wtedy i tylko wtedy, gdy ostatnia kolumna macierzy B zawiera współczynnik wiodący.

Twierdzenie 4.13. (Kronecker-Capelli) Układ (u) ma rozwiązanie ($\text{rozw}(u) \neq \emptyset$) wtedy i tylko, gdy $r(A_{(u)}) = r(\widetilde{A}_{(u)})$.

Metody rozwiązywania układów liniowych - eliminacja Gaussa.

Przypuśćmy, że (u) jest układem równań liniowych. Eliminacja Gaussa polega na sprowadzeniu macierzy $\widetilde{A}_{(u)}$ do postaci całkowicie zredukowanej macierzy górno-schodkowej. Równanie w tej postaci ma oczywiste rozwiązania.

Niech C będzie dowolną macierzą o m wierszach i n kolumnach. Pokażemy indukcyjnie (ze względu na liczbę wierszy) w kilku krokach jak sprowadzić macierz C do postaci górno-schodkowej. Oczywiście dla $m = 1$ sytuacja jest oczywista. Przypuśćmy, że $m > 1$.

1. Niech $C(j_0)$ oznacza pierwszą niezerową kolumnę macierzy C oraz $c_{i_0 j_0}$ pierwszy niezerewy element w tej kolumnie. Dokonujemy zamiany wierszy o nr. 1 i nr. i_0. Otrzymujemy w ten sposób macierz
$$\begin{bmatrix} 0 & 0 & c_{i_0 j_0} & \text{coś} \\ 0 & 0 & \text{coś} & \text{coś} \\ 0 & 0 & \ddots & \vdots \\ 0 & 0 & \vdots & \text{coś} \end{bmatrix}$$
2. Jeśli $k > 1$, to do k-tego wiersza dodajemy pierwszy wiersz przemnożony przez $c_{kj_0} \cdot c_{i_0 j_0}^{-1}$. Wówczas otrzymujemy macierz:
$$C' = \begin{bmatrix} 0 & 0 & c_{i_0 j_0} & \text{coś} \\ 0 & 0 & 0 & \\ 0 & 0 & 0 & C'' \\ 0 & 0 & 0 & \end{bmatrix}.$$

Możemy teraz zastosować założenie indukcyjne dla macierzy C'', co po m krokach daje nam macierz górno-schodkową. Zauważmy, że do tej pory wykorzystaliśmy tylko dwie operacje elementarne. Pokażemy, że przy użyciu trzeciej każdą macierz górno-schodkową można sprowadzić do postaci całkowicie zredukowanej. Przypuśćmy, że B

jest macierzą górno-schodkową. Postępujemy znowu indukcyjnie:

Jeśli

$$B = \begin{bmatrix} 0 & a & b & \text{coś} \\ 0 & 0 & c & \text{coś} \\ 0 & 0 & 0 & \text{coś} \\ 0 & 0 & 0 & \text{coś} \end{bmatrix},$$

to pierwszy wiersz mnożymy przez a^{-1}. Wówczas otrzymujemy macierz

$$B' = \begin{bmatrix} 0 & 1 & ba^{-1} & \text{coś} \\ 0 & 0 & c & \text{coś} \\ 0 & 0 & 0 & \text{coś} \\ 0 & 0 & 0 & \text{coś} \end{bmatrix}.$$

Następnie mnożymy drugi wiersz przez $-ba^{-1}c^{-1}$ i dodajemy do pierwszego. W ten sposób otrzymujemy macierz

$$B'' = \begin{bmatrix} 0 & 1 & 0 & \text{coś} \\ 0 & 0 & -ba^{-1} & \text{coś} \\ 0 & 0 & 0 & \text{coś} \\ 0 & 0 & 0 & \text{coś} \end{bmatrix}.$$

Indukcyjnie postępujemy dla kolejnych wierszy.

Metody rozwiązywania układów liniowych - wyznacznik.

Przypuśćmy, że mamy dany układ równań liniowych (u).

1. Na mocy twierdzenia Kroneckera-Capelliego układ (u) ma rozwiązanie wtedy i tylko wtedy, gdy $r(A_{(u)}) = r(\widetilde{A}_{(u)})$, więc w pierwszym kroku sprawdzamy ten warunek. Niech $r = r(A_{(u)}) = r(\widetilde{A}_{(u)})$.

2. Przypomnijmy, że minorem w macierzy M nazywamy taką macierz kwadratową A, że A powstaje z M poprzez usunięcie ustalonych wierszy i kolumn. W kroku drugim szukamy $r \times r$ minora M w macierzy $A_{(u)}$. Jak go znajdziemy, to rozważamy wiersze macierzy $\widetilde{A}_{(u)}$ wyznaczone przez M. Zmienne nie występujące w M przerzucamy na drugą stronę równania i traktujemy jako parametry.

3. Otrzymaliśmy równanie (macierzowe) $M \cdot X = B'$, gdzie M jest macierzą kwadratową, o niezerowym wyznaczniku (układ Cramera). Wówczas rozwiązaniem jest

$$X = M^{-1} \cdot B'.$$

5. Iloczyn skalarny i wektorowy.

Definicja 5.0. Niech k będzie ciałem, zaś V przestrzenią liniową nad k. *Formą dwuliniową* nazywamy dowolne odwzorowanie $f : V \times V \to k$ takie, że $f(a \cdot v + b \cdot w, u) = a \cdot f(v, u) + b \cdot f(w, u)$ oraz $f(v, a \cdot w + b \cdot u) = a \cdot f(v, w) + b \cdot f(v, u)$ dla wszystkich $a, b \in k$ oraz $u, v, w \in V$. Ponadto powiemy, że forma dwuliniowa $f : V \times V \to k$ jest *symetryczna*, o ile $f(v, w) = f(w, v)$ dla wszystkich $v, w \in V$.

Definicja 5.1. Niech V będzie przestrzenią liniową nad ciałem liczb rzeczywistych $\mathbb{R}$. Symetryczną formę dwuliniową $\varphi : V \times V \to \mathbb{R}$ nazywamy *dodatniookreśloną*, o ile dla każdego $v \in V \backslash \{0\}$ zachodzi $\varphi(v, v) > 0$. Wówczas piszemy $\varphi > 0$. Macierz kwadratową $M \in \mathbb{M}_n(\mathbb{R})$ nazywamy *dodatniookreśloną*, o ile dla każdego (niezerowego) wektora $x \in \mathbb{R}^n$ zachodzi $x^t M x > 0$. Wówczas piszemy $M > 0$.

Zauważmy, że jeśli $\dim V < \infty$ (co od tego momentu będziemy wciąż zakładać), to $\varphi > 0$ wtedy i tylko wtedy, gdy $M > 0$, gdzie M jest macierzą stowarzyszoną z φ, tzn. przy ustalonej bazie $\{e_1, \dots, e_n\}$ przestrzeni V mamy $m_{ij} = \varphi(e_i, e_j)$. Ponadto własność ta nie zależy od wyboru bazy przestrzeni V.

Jeśli $A \in \mathbb{M}_n(\mathbb{R})$, to definiujemy $\varphi_A : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ wzorem $\varphi(x, y) = x^t A y$. Wówczas $A > 0$ wtedy i tylko wtedy, gdy $\varphi > 0$.

Definicja 5.2. Dowolną symetryczną, dwuliniową formę dodatniookreśloną $\varphi : V \times V \to \mathbb{R}$ nazywamy *iloczynem skalarnym* na V i zwyczajowo piszemy $\langle x, y \rangle := \varphi(x, y)$. Parę $(V, \langle -, - \rangle)$ nazywamy *przestrzenią Euklidesową*, o ile $\dim V < \infty$.

Definicja 5.3. Dla dowolnej przestrzeni Euklidesowej, jeśli $v, w \in V$, to powiemy, że v jest prostopadłe do w (i piszemy $v \perp w$), jeśli $\langle v, w \rangle = 0$. Jeśli obydwa wektory są niezerowe, to są liniowo niezależne. Ponadto symbolem $\|v\| = \langle v, v \rangle$ oznaczamy „długość” wektora v. Ponadto, jeśli W jest podprzestrzenią V, to $W^\perp = \{v \in V \mid \forall_{w \in W}\ v \perp w\}$ jest podprzestrzenią V.

Lemat 5.4. Dla dowolnej podprzestrzeni W przestrzeni Euklidesowej V zachodzi $V = W \oplus W^\perp$.

Lemat 5.5. Dla dowolnej przestrzeni Euklidesowej V istnieje baza $\mathcal{B} = \{e_1, \dots, e_n\}$ taka, że $\langle e_i, e_j \rangle = 0$ (tzn. $\mathcal{B}$ jest otrogonalna) oraz $\|e_i\| = 1$ dla wszystkich $i, j, i \neq j$ (w sumie te dwa warunki oznaczają, iż baza $\mathcal{B}$ jest ortonormalna).

Ortogonalizacja Grama-Schmidta

Przypuśćmy, że V jest przestrzenią Euklidesową, zaś $\mathcal{B} = \{e_1, \dots, e_n\}$ jest bazą V. Jak z $\mathcal{B}$ otrzymać bazę ortonormalną? Będziemy konstruować tę bazę indukcyjnie.

1. Definiujemy $u_1 = e_1$;

2. Przypuśćmy, że $u_1, \ldots, u_{k-1}$ są już dane. Wówczas określamy

$$u_k = e_k + \sum_{i=1}^{k=1} \alpha_i \cdot e_i,$$

gdzie $\alpha_i = -\frac{\langle e_i, u_i \rangle}{\|u_i\|}$

Zadaje to nam wektory $\{u_1, \ldots, u_n\}$ i można pokazać, że tworzą one bazę ortonormalną.

Własności geometryczne iloczynu skalarnego

Twierdzenie 5.6. **(Nierówność Schwartza)** Dla dowolnych wektorów $x, y \in V$ zachodzi

$$|\langle x, y \rangle| \leqslant \|v\| \cdot \|w\|.$$

W szczególności zachodzi nierówność trójkąta, tzn.

$$\|x + y\| \leqslant \|x\| + \|y\|.$$

Definicja 5.7. Niech $x, y \in V$ będą niezerowe. *Kątem niezorientowanym* między x i y nazywamy (jedyną) liczbę rzeczywistą $\alpha \in [0, \pi)$ taką, że

$$\cos \alpha = \frac{\langle x, y \rangle}{\|x\| \cdot \|y\|}.$$

Zauważmy, że wówczas $x \perp y$ wtedy i tylko wtedy, gdy $\cos \alpha = 0$.

Twierdzenie 5.8. **(Pitagoras)** Dwa niezerowe wektory x, y są prostopadłe wtedy i tylko wtedy, gdy

$$\|x + y\|^2 = \|x\|^2 + \|y\|^2.$$

Iloczyn wektorowy

Definicja 5.9. Niech $x, y \in \mathbb{R}^3$. *Iloczynem wektorowym* wektora x przez y nazywamy wektor

$$x \times y = \begin{vmatrix} i & j & k \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{vmatrix} = (x_2 y_3 - x_3 y_2)i - (x_1 y_3 - x_3 y_1)j + (x_1 y_2 - x_2 y_1)k,$$

gdzie $i = e_1$, $j = e_2$, $k = e_3$ jest bazą standardową w $\mathbb{R}^3$

Zauważmy, że dla każdego wektora $x \in \mathbb{R}^3$ zachodzi $x \times x = 0$. Ponadto mamy:

Lemat 5.10. Dla dowolnych wektorów $a, b, c \in \mathbb{R}^3$ oraz dla $\zeta \in \mathbb{R}$ zachodzi:

$$a \times b = -b \times a;$$

$$(a + b) \times c = a \times c + b \times c;$$

$$a \times (b + c) = a \times b + a \times c;$$

$$(\zeta a) \times b = a \times (\zeta b) = \zeta(a \times b);$$

$$\langle a \times b, c\rangle = \langle a, b \times c\rangle;$$

$$(a \times b) \times c = \langle a, c\rangle b - \langle b, c\rangle a;$$

$$\|a \times b\|a\| \cdot \|b\| \cdot \cos\alpha,$$

gdzie α jest kątem niezorientowanym między a i b.

Zastosowanie w geometrii

Przypuśćmy, że $A, B \in \mathbb{R}^3$. Określamy $\overrightarrow{AB} = (x_B - x_A, y_B - y_A, z_B - z_A)$. Jeśli $v = (v_1, v_2, v_3)$ jest punktem, zaś $A = (z_A, y_A, z_A)$ jest wektorem, to prosta wyznaczona przez v i A jest zadana wzorem

$$l = \{(x_A + tv_1, y_A + tv_2, z_A + tv_3) \mid t \in \mathbb{R}\}.$$

Jest to postać parametryczna. Prosta wyznaczona przez punkty A, B jest prostą wyznaczoną przez A i $\overrightarrow{AB}$. Odległoś punktu A od B jest zadana wzorem $\|\overrightarrow{AB}\|$.

Jeśli l, k są prostymi przecinającymi się w punkcie C, zaś $A \in l$, $B \in k$ są różnymi punktami, to mówimy, że proste l i k są prostopadłe, o ile wektory $\overrightarrow{CA}$ i $\overrightarrow{CB}$ są prostopadłe.

Jeśli $P = (x_0, y_0, z_0)$ jest punktem, zaś $n = (a, b, c)$ wektorem, to prosta prostopadła do n przechodząca przez P jest zadana wzorem

$$\Pi = \{(x, y, z) \mid a(x - x_0) + b(y - y_0) + c(z - z_0) = 0\}.$$

Jeśli mamy daną płaszczyznę wyznaczoną przez 3 punkty A, B, C, to wektor prostopadły do niej jest zadany wzorem

$$n = \overrightarrow{AB} \times \overrightarrow{AC}.$$

Zauważmy, że płaszczyzna jest wyznaczona przez 3 punkty, o ile nie są one współliniowe.

Lemat 5.11. (kryterium współliniowości) Trzy punkty A, B, C są współliniowe wtedy i tylko wtedy, gdy $\overrightarrow{AB} \times \overrightarrow{AC} = 0$, a to jest wtedy i tylko wtedy, gdy wektory $\overrightarrow{AB}$, $\overrightarrow{AC}$ są równoległe.

Lemat 5.12. (objętość równoległościanu i czworościanu) Przypuśćmy, że mamy dane wektory $\overrightarrow{a}, \overrightarrow{b}, \overrightarrow{c}$ zaczepione w tym samym punkcie, niewspółliniowe. Wówczas wektory te wyznaczają równoległościan R lub czworościan C. Objętości brył wyrażają się wzorami:

$$V_R = |\langle \overrightarrow{a} \times \overrightarrow{b}, \overrightarrow{a} \rangle|;$$

$$V_C = \frac{1}{6}|\langle \overrightarrow{a} \times \overrightarrow{b}, \overrightarrow{c} \rangle|.$$

Wyliczymy odległość punktu od prostej. Przypuśćmy, że mamy daną prostą l wyznaczoną przez punkty A, B i punkt $P = (x_0, y_0, z_0) \notin l$ w $\mathbb{R}^3$. Szukamy punktu $X \in l$ takiego, że $\overrightarrow{PX} \perp \overrightarrow{AB}$. Wówczas odległością jest $\|\overrightarrow{PX}\|$. Przypomnijmy, że warunek prostopadłości równoważny jest warunkow $\langle \overrightarrow{PX}, \overrightarrow{AB} \rangle = 0$. Ponieważ w punkcie X mamy jedną zmienną (prosta l ma postać parametryczną z parametrem t), więc otrzymujemy równanie liniowe jednej zmiennej t. Pozostaje je rozwiązać i wyliczyć długość wektora $\overrightarrow{PX}$.