5 Równania liniowe, macierze

Przypomnijmy, że odwzorowanie $S : X \to Y$ między dwiema przestrzeniami wektorowymi nazywamy odwzorowaniem (= operatorem) liniowym, gdy $\forall_{u,w\in X}\forall_{t,r\in\mathbb{R}} S(ru + tw) = rS(u) + tS(w)$. Metodą indukcji łatwo wnioskujemy, że analogiczna równość zachodzi dla wartości S na sumie skończonej ilości wektorów pomnożonych przez skalary: $S(t_1u_1 + t_2u_2 + \cdots + t_ku_k) = t_1S(u_1) + t_2S(u_2) + \cdots + t_kS(u_k)$ Zbiór wszystkich odwzorowań liniowych z przestrzeni wektorowej X do Y oznaczamy symbolem $\mathcal{L}(X, Y)$. Ponadto niech $\mathcal{L}(X) := \mathcal{L}(X, X)$. Operatory liniowe o wartościach skalarnych nazywamy funkcjonałami liniowymi, ich zbiór, czyli $\mathcal{L}(X, \mathbb{R})$ oznaczamy symbolem X^*.

Stwierdzenie. Wartości dowolnego operatora liniowego $S \in \mathcal{L}(X, Y)$ na wektorach przestrzeni X są jednoznacznie określone poprzez wartości S na wektorach dowolnie ustalonej bazy. I na odwrót: gdy dla dowolnego układu wektorów $(y_j)_{j\in J}$ istnieje dokładnie jeden operator $S \in \mathcal{L}(X, Y)$ taki, że $S(e_j) = y_j$, to można wykazać, że układ $(e_j)_{j\in J}$ jest bazą przestrzeni wektorowej X.

Zarys dowodu. Jeśli bazę w przestrzeni X stanowią wektory $e_1, e_2, \dots, e_d$, to dowolny wektor $v \in X$ ma jednoznacznie wyznaczone współczynniki $c_1, c_2, \dots, c_d \in \mathbb{R}$, dla których $v = \sum_{n=1}^{d} c_je_j$. Korzystając z liniowości S możemy napisać

$$S(v) = S\left(\sum_{n=1}^{d} c_je_j\right) = \sum_{n=1}^{d} c_jS(e_j). \tag{1}$$

To dowodzi pierwszej tezy. Drugą można sprawdzać metodą nie wprost. Gdyby układ (e_j) nie był liniowo niezależny, to pewien jego element e_k byłby liniową kombinacją pozostałych, np gdy $k = 1$, a układ ma d elementów, powinny istnieć liczby $\alpha_2, \dots \alpha_d \in \mathbb{R}$ dla kórych $e_1 = \sum_{j=2}^{d} \alpha_je_j$. Jeżeli wybierzemy y_1 różne od $\sum_{j=2}^{d} \alpha_jy_j$, to układ wartości $y_1, y_2, \dots, y_d$ nie będzie układem wartości $S(e_j)$ dla żadnego liniowego operatora S.

Z kolei, gdyby układ e_j nie rozpinał całej przestrzeni X, to wraz z pewnym wektorem e_{d+1} byłby nadal liniowo niezależny, ale dla różnych wartości $y_{d+1} = S(e_{d+1})$ mielibyśmy różne rozszerzenia przy tych samych wartościach $S(e_j)$ dla $j = 1, 2, \dots, d$. □

Gdy $X = \mathbb{R}^d$, $Y = \mathbb{R}^m$, to mamy takich d wektorów $S(\varepsilon_1), S(\varepsilon_2), \dots, S(\varepsilon_d)$. Tu $\varepsilon_1 = (1, 0, \dots, 0), \varepsilon_2 = (0, 1, 0, \dots, 0), \dots \varepsilon_d = (0, \dots, 0, 1)$ jest kanoniczną bazę zero-jedynkową w $\mathbb{R}^d$. Oznaczmy przez $\tilde{\varepsilon_j}, j = 1, 2, \dots, m$ kanoniczną bazę zerojedynkową w $\mathbb{R}^m$. Każdy z tych wektorów $S(\varepsilon_k)$ zapisujemy w postaci kolumny m liczb, otrzymując macierz odwzorowania S, oznaczmy ją M_S, złożoną z m wierszy i d kolumn. **Wyraz leżący w j -tym wierszu i k -tej kolumnie oznaczmy a_{jk}.** Jest to współczynnik dla wektora $S(\varepsilon_k)$ stojący przy $\tilde{\varepsilon_j}$. Obliczamy go jako iloczyn skalarny tego wektora$S(\varepsilon_k)$ przez $\tilde{\varepsilon_j}$ Tak więc k-tą kolumną macierzy M_S jest

$$S(\varepsilon_k) = \begin{pmatrix} a_{1k} \\ a_{2k} \\ \vdots \\ a_{mk} \end{pmatrix}, \qquad \texttt{gdzie} \quad a_{jk} = \langle S(\varepsilon_k), \tilde{\varepsilon_j}\rangle. \tag{2}$$

Cała macierz, czyli zestawienie tych d kolumn ma postać

$$M_S = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1d} \\ a_{21} & a_{22} & \dots & a_{2d} \\ \cdots & \cdots & \cdots & \cdots \\ & & \ddots & \\ \dots & \dots & & \dots \\ a_{m1} & a_{m2} & \dots & a_{md} \end{pmatrix}.$$

Na przykład, gdy $X = Y = \mathbb{R}^d$, operator identyczności I ma macierz, kóra na głównej przekątnej ma same jedynki, poza przekątną- zera, czyli $\forall_{k\leqslant d}\, a_{kk} = 1$, natomiast dla $j \neq k$ jest $a_{jk} = 0$. Można się zastanawiać, jakiemu operatorowi $S \in \mathcal{L}(\mathbb{R}^m, \mathbb{R}^d)$ odpowiada macierz M^T o wyrazach a_{kj}, gdy $S \in \mathcal{L}(\mathbb{R}^d, \mathbb{R}^m)$ ma macierz $M = \big(a_{jk}\big)$. Taką macierz nazywamy macierzą transponowaną dla macierzy M i oznaczmy M^T. W przypadku najważniejszym -macierzy kwadratowych, transponowanie, to symetria względem głównej przekątnej. Na przykład,

$$\begin{pmatrix} 2 & 7 & 5 \\ 1 & 2 & 3 \\ 0 & 4 & 6 \end{pmatrix}^T = \begin{pmatrix} 2 & 1 & 0 \\ 7 & 2 & 4 \\ 5 & 3 & 6 \end{pmatrix}, \qquad \begin{pmatrix} 7 & 5 \\ 2 & 3 \\ 4 & 6 \end{pmatrix}^T = \begin{pmatrix} 7 & 2 & 4 \\ 5 & 3 & 6 \end{pmatrix}.$$

Macierzą transponwaną do M_S -macierzy operatora S jest macierz operatora S^* zwanego operatorem sprzężonym do $S \in \mathcal{L}(X, Y)$. Jest to jedyny operator $S^* \in \mathcal{L}(Y, X)$ spełniający relację

$$\forall_{x \in X, y \in Y} \, \langle Sx, y\rangle = \langle x, S^* y\rangle.$$

Trzeba podkreślić, że dotyczy to przestrzeni nad ciałem skalarów $\mathbb{R}$.

Nieco trudniej jest opisać macierz M_{ST} odpowiadającą złożeniu ST operatorów $T \in \mathcal{L}(X, Y)$ oraz $S \in \mathcal{L}(Y, Z)$. Jest to tak zwany iloczyn macierzy $A \cdot B$ dla $M_S = A$, $M_T = B$. Wektor $\vec{x} = (x_1, \ldots, x_d) \in \mathbb{R}^d$ może być utożsamiany z odwzorowaniem $\mathbb{R} \ni t \mapsto t\vec{x} \in \mathbb{R}^d$, macierzą tego odwzorowania jest kolumna o d wierszach. Wektory traktowane jako macierze zapisujemy więc w formacie kolumnowym, zaś dla $S \in \mathcal{L}(\mathbb{R}^d, \mathbb{R}^m)$ wektor $S(\vec{x})$ -czyli wynik działania operatora S na wektorze $\vec{x}$, to iloczyn macierzy M_S przez macierz $\vec{x} = (x_1, x_2, \ldots, x_d)^T$ stojącą (jako kolumna) po prawej stornie. Wynikiem jest wektor w $\mathbb{R}^m$, który dzięki wozom (1),(2) ma współrzędne $y_j = \sum_{k=1}^{d} a_{jk} x_k$. Jest to iloczyn skalarny m-tego wiersza macierzy M_S przez $\vec{x}$. Jeżeli zamiast 1 kolumny mamy po prawej stronie jakąś macierz $B = M_T$ odwzorowania T, macierz o wyrazach b_{ik}, to mnożenie $M_S \cdot M_T$ tych macierzy jest wykonalne tylko w takim przypadku, gdy złożenie operatorów $S \circ T$ jest wykonalne, czyli przestrzeń do której działa T jest taka sama jak dziedzina S. Dla macierzy oznacza to, że ilość kolumn macierzy $A = M_S$ jest taka sama, jak ilość wierszy $B = M_T$. Wynikiem będzie macierz $A \cdot B$. Jest to macierz o wyrazach

$$c_{ij} = \sum_{k=1}^{d} a_{ik} b_{kj}. \tag{3}$$

Na przykład, gdy B jest macierzą odwzorowania identyczności I_d w $\mathbb{R}^d$, to $b_{kj} = 0$ dla $k \neq j$ oraz $b_{kk} = 1$. Suma (3) ma więc tylko jeden niezerowy składnik, dla $k = j$ i jest równa a_{ij}. Czyli $C = A$, albo $AI_d = A$. Jest to więc "zasada obróconego T" -w kształcie $-|$ (bo c_{ij} = wynik mnożenia skalarnego i-tego wiersza z lewej strony przez j-tą kolumnę z prawej strony).

Mnożenie macierzy nie jest przemienne, ale jest łączne. Iloczyn dwu macierzy kwadratowych $d \times d$ jest macierzą kwadratową takiego samego rozmiaru. Jeżeli dla macierzy kwadratowych A, B ich iloczynem jest macierz identyczności, to mówimy, że macierz A jest odwracalna (lub nieosobliwa), zaś B jest jej odwrotnością: $B = A^{-1}$.

Jak znaleźć A^{-1}, a przede wszystkim jak określić, kiedy A jest odwracalna? Kluczowe dla tej drugiej (jak i pierwszej) odpowiedzi będzie pojęcie wyznacznika macierzy A, jest to liczba oznaczana $\det(A)$. Innym oznaczeniem jest zastąpienie nawiasów okrągłych przez kreski pionowe po bokach macierzy. Łatwo określić ją dla macierzy kwadratowej 2×2. Przypuśćmy, że $a_{11} = a, a_{12} = b, a_{21} = c, a_{22} = d$. Mamy

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc. \tag{4}$$

Dla macierzy 3×3 wyznacznik możemy wyliczyć przy użyciu wzoru Sarrusa

$$\det(A) = (a_{11} \cdot a_{22} \cdot a_{33} + a_{21} \cdot a_{32} \cdot a_{13} + a_{31} \cdot a_{12} \cdot a_{23}) +$$

$$-(a_{13} \cdot a_{22} \cdot a_{31} + a_{23} \cdot a_{32} \cdot a_{11} + a_{33} \cdot a_{12} \cdot a_{21}).$$

UWAGA: dla macierzy wyższych wymiarów nie można stosować tego wzoru. Dość prosty przepis na tworzenie obydwu sum jest zawarty w schemacie graficznym, w którym pod macierzą dopisujemy dwa pierwsze wiersze, następnie tworzymy iloczyny wzdłuż przekątnych. Ze znakiem plus dla kierunku \, za znakiem minus (drugi nawias) -dla kierunku /. Rysunek zrobię na wykładzie. Dla macierzy $d \times d$ mamy wzór Laplace'a pomagający liczyć wyznacznik przy użyciu wyznaczników macierzy $(d-1) \times (d-1)$, który podamy poniżej. Macierz o wyrazach a_{ij} nazwiemy macierzą diagonalną jeżeli $\forall_{k \neq j} a_{kj} = 0$. Wtedy jej wyznacznikiem jest iloczyn wszystkich wyrazów na przekątnej głównej: $a_{11} \cdot a_{22} \cdot \cdots \cdot a_{dd}$. Taki sam wzór na wyznacznik obowiązuje dla macierzy trójkątnych górnych, czyli takich, w których wszystkie wyrazy pod przekątną główną są zerowe.

Definicja. Minorem $M_{kj}(A)$ macierzy kwadratowej A nazywamy wyznacznik macierzy kwadratowej powstałej po wykreśleniu z tej macierzy k-tego wiersza oraz j-tej kolumny. Dopełnieniem algebraicznym A_{kj} (wyrazu a_{kj}) tej macierzy A nazywamy liczbę $(-1)^{k+j} M_{kj}(A)$. (Angielska nazwa = *cofactor*)

Twierdzenie. (Rozwinięcie Laplace'a): Wyznacznik $\det(A)$ macierzy $A = (a_{jp})_{j,p \leqslant d}$ jest dla każdych $n, k \in \{1, 2, \ldots, d\}$ równy liczbie

$$D_{row}(n) := \sum_{k=1}^{d} a_{nk} A_{nk} \qquad \texttt{(rozwinięcie wg } n\texttt{-tego wiersza)}$$

oraz równy liczbie

$$D_{col}(k) := \sum_{n=1}^{d} a_{nk} A_{nk} \qquad \texttt{(rozwinięcie wg } k\texttt{-tej kolumny)}$$

UWAGA: Formalna definicja wyznacznika jest dość skomplikowana, więc możemy poprzestać na wzorach Laplace'a i krok po kroku zwiększać rozmiar macierzy. Jest tylko jeden problem: Każdy wzór ogólny na wyznacznik macierzy $d \times d$ wymaga wykonania ponad $d!$ obliczeń. Np. $3! = 6$ i tyle składników jest we wzorze Sarrusa. Ale już dla macierzy 20×20 będzie $20!$, czyli około $2,4 \times 10^{18}$ mnożeń. Dla komputera o miliardzie operacji na sekundę to zajmie aż 77 lat! Dlatego stosuje się raczej rozwiązania typu *metoda eliminacji Gaussa* do rozwiązywania dużych układów równań.Ich "koszt obliczeniowy" jest nieporównanie niższy. Metoda ta polega na sprowadzaniu poprzez operacje niezmieniające wyznacznika (por.(iv) poniżej) do postaci trójkątnej górnej.

Twierdzenie.

(i) Dla macierzy kwadratowych jest $\det(A \cdot B) = \det(A) \cdot \det(B)$.

(ii) Dla macierzy transponowanej $\det(A^T) = \det(A)$.

(iii) Dla stałej $c \in \mathbb{R}$ i dla macierzy $d \times d$ mamy $\det(cA) = c^d \det(A)$.

(iv) $\det(A) = \det(B)$, jeśli macierz B powstaje przez dodanie do jednego z wierszy macierzy A innego jej wiersza pomnożonego przez pewną stałą.

(v) $\det(A) = -\det(B)$, jeśli macierz B powstaje przez zamianę kolejności 2 wierszy (lub 2 kolumn) macierzy A .

(vi) Macierz kwadratowa A jest nieosobliwa (czyli odwracalna) wtedy i tylko wtedy, gdy jej wyznacznik jest niezerowy.

(vii) Macierz jest nieosobliwa wtedy i tylko wtedy, gdy jej kolumny tworzą układ wektorów liniowo niezależnych

(viii) Gdy $\det(A) \neq 0$, to niech C = macierz dopełnień algebraicznych A_{jk}. Wówczas
$$A^{-1} = \frac{1}{\det(A)} C^T.$$

Odwracanie macierzy wiąże się z rozwiązywaniem układów równań liniowych. Zacznijmy od równań operatorowych postaci

$$Tx = y, \quad T \in \mathcal{L}(X,Y),\, y \in Y - \texttt{dany wektor},$$

gdzie niewiadomą jest szukany wektor x. Gdy $y = 0$, mówimy o równaniu jednorodnym. Może ono mieć albo tylko jedno rozwiązanie: $x = 0$, albo nieskończenie wiele rozwiązań, które tworzą podprzestrzeń wektorową zwaną jądrem operatora T:

$$\ker(T) := \{x \in X : \; T(x) = 0\}$$

Zauważmy, że faktycznie jest to podprzestrzeń wektorowa w X: gdy $T(x) = 0$ oraz $T(z) = 0$,zaś $\alpha \in \mathbb{R}$, to $T(\alpha x) = \alpha T(x) = 0, T(x+z) = 0+0 = 0$. Ponadto mamy następujące proste obserwacje:

Lemat. Odwzorowanie liniowe T jest injekcją $\Leftrightarrow$ gdy $\ker T = \{0\}$.

Faktycznie, $T(x) = T(z) \Leftrightarrow T(x-z) = 0 \Leftrightarrow x - z \in \ker(T)$.

Lemat. Gdy $z \in X$ jest rozwiązaniem równania liniowego niejednorodnego $T(z) = y$, gdzie $T \in \mathcal{L}(X,Y)$, to zbiór wszystkich rozwiązań tego równania jednorodnego jest postaci

$$z + \ker(T) = \{z + x : \; x \in \ker(T)\}.$$

Zasada ta ma istotne zastosowanie dla równań różniczkowych liniowych.

Rząd operatora liniowego T, oznaczany $\mathrm{rk}(T)$ – to wymiar obrazu przestrzeni X przez T. Ten obraz, czyli zbiór $T(X) = \{T(x) : \; x \in X\}$ jest oczywiście podprzestrzenią wektorową w Y gdy $T \in \mathcal{L}(X,Y)$. Mamy wówczas ważną równość:

$$\dim X = \texttt{rk}(T) + \dim(\ker(T)).$$

W przypadku gdy $X = Y$, czyli dla $T \in \mathcal{L}(X)$ bijektywność, surjektywność i injektywność T -są sobie równoważne. Faktycznie, injektywność jest równoważna temu, że $\dim(\ker(T)) = 0$, a to zachodzi $\Leftrightarrow$ $\mathrm{rk}(T) = d$, dzięki tej równości. Te trzy warunki są jeszcze równoważne nieosobliwości macierzy M_T, czyli warunkowi $\det(M_T) \neq 0$. Odwzorowanie odwrotne do liniowej bijekcji jest liniowe (to łatwo sprawdzić). Jedynym rozwiązaniem równania niejednorodnego $Tx = y$ w przypadku bijektywnego T jest $x = T^{-1}y$, co wynika wprost z definicji odwzorowania odwrotnego. Ponieważ macierzą odwzorowania T^{-1} jest odwrotna macierz do M_T, możemy w ten sposób rozwiązywać równania liniowe macierzowe, czyli układy zwykłych równań liniowych.

Konkretnie, gdy A jest macierzą $d \times d$ oraz mamy wektor danych $y = (y_1, y_2, \ldots, y_d)$, to równanie macierzowe $Ax = y$, gdzie $x = (x_1, x_2, \ldots, x_d)$ jest szukanym rozwiązaniem, można interpretować jako układ d równań z d niewiadomymi $x_k, k \leqslant d$, a mianowicie

$$a_{j1}x_1 + \cdots + a_{jk}x_k + \cdots + a_{jd}x_d = y_j, \quad j \in \{1, 2, \ldots, d\}.$$

Jego rozwiązanie możemy zapisać wektorowo w postaci $x = A^{-1}y$. Mamy też wzory Cramera używające wyznacznika $D = \det(A)$ oraz wyznaczników D_k macierzy utworzonych poprzez zastąpienie k-tej kolumny macierzy A przez kolumnę wektora y. Np.

$$D_1 = \begin{vmatrix} y_1 & a_{12} & \ldots & a_{1d} \\ y_2 & a_{22} & \ldots & a_{2d} \\ \ldots & \ldots & \ldots & \ldots \\ y_d & a_{d2} & \ldots & a_{dd} \end{vmatrix}, \qquad D_2 = \begin{vmatrix} a_{11} & y_1 & \ldots & a_{1d} \\ a_{21} & y_2 & \ldots & a_{2d} \\ \ldots & \ldots & \ldots & \ldots \\ a_{d1} & y_d & \ldots & a_{dd} \end{vmatrix}, \ldots$$

Twierdzenie. Dla macierzy nieosobliwej rozwiązaniem powyższego układu równań liniowych jest wektor $(x_1, x_2, \ldots, x_d)$, gdzie $x_j = \frac{D_j}{D}$ dla $j = 1, 2, \ldots, d$.