Mathematics is central to the intellectual history of all societies

Algebra Liniowa z Geometrią Analityczną

Aleksander Strasburger,
Instytut Matematyki i Kryptologii WAT

Semestr Letni 2019

Kontakt z prowadzącym

Dr hab. Aleksander Strasburger
Pokój: 3/43A, bud. 34
Termin Konsultacji: (do uzgodnienia)
E-mail: aleksander_strasburger@sggw.pl

Motto

Algebra, with its emphasis on questions of structure and operations, deals with mathematical abstractions which are not too far removed from problems of computer organization and programming.

R. E. Kalman

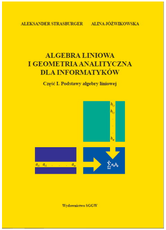
Mechanika kursu

- Przedmiot rozwija tematykę kursu „Podstawy matematyki wyższej".
- Matematyka jest „przedmiotem interaktywnym" - wymaga zrozumienia idei i systematycznej praktyki (ćwiczenia) kształcącej to zrozumienie.
- Kurs realizowany w formie wykładu i ćwiczeń. Zaliczenie wspólne. Ćwiczenia prowadzą: dr Alina Jóźwikowska, dr Diana Dziewa-Dawidczyk i A. S.
- Dwa kolokwia w trakcie semestru i egzamin kończący.
- Konsultacje — termin do ustalenia.

Monografie i podręczniki I

Podręcznik podstawowy dla wykładu.

- A. Strasburger, A. Jóźwikowska, *Algebra liniowa i geometria analityczna dla informatyków*, Cz. I. *Podstawy algebry liniowej,* wydanie II, Wydawnictwo SGGW, Warszawa 2017



Podręczniki uzupełniające dla wykładu.

- P. Kajetanowicz, J. Wierzejewski, *Algebra z geometrią analityczną*, PWN, Warszawa 2008
- R. Leitner, *Zarys matematyki wyższej*, WNT, Warszawa 1995 (wiele wydań)

(Pre)-historia

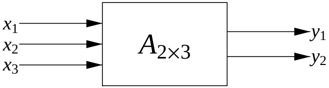
Pierwszy odnotowany problem prowadzący do układu równań liniowych pochodzi prawdopodobnie z okresu 200 B.C. i jest zawarty w chińskim traktacie *Nine Chapters on the Mathematical Art.*

> *Three sheafs of a good crop, two sheafs of a mediocre crop, and one sheaf of a bad crop are sold for 39 dou. Two sheafs of good, three mediocre, and one bad are sold for 34 dou; and one good, two mediocre, and three bad are sold for 26 dou. What is the price received for each sheaf of a good crop, each sheaf of a mediocre crop, and each sheaf of a bad crop?*

We współczesnym języku problem wyraża się przez układ równań

$$
\begin{aligned}
3x + 2y + z &= 39,\\
2x + 3y + z &= 34,\\
1x + 2y + 3z &= 26.
\end{aligned}
$$

Układy równań liniowych — schemat zastosowania



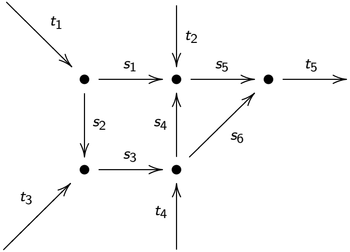
$$y_1 = a_{11}x_1 + a_{12}x_2 + a_{13}x_3$$

$$y_2 = a_{21}x_1 + a_{22}x_2 + a_{23}x_3$$

Układ wejście-wyjście (input-output)
jest reprezentowany przez równanie macierzowe

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} .$$

Realny diagram i jego macierz



$$
\begin{array}{rcrcrcrcrcrcl}
s_1 & + & s_2 & & & & & & & & & = & t_1, \\
-s_1 & & & & & - & s_4 & + & s_5 & & & = & t_2, \\
 & - & s_2 & + & s_3 & & & & & & & = & t_3, \\
 & & & - & s_3 & + & s_4 & & & + & s_6 & = & t_4, \\
 & & & & & & & & s_5 & + & s_6 & = & t_5.
\end{array}
$$

Graf skierowany z pięcioma wierzchołkami i sześcioma krawędziami.

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

Tak zwana *macierz incydencji grafu.*

Układy równań liniowych I

W równaniach można oznaczać każdą z niewiadomych osobną literą, ale wygodniej dla niewiadomych używać notacji indeksowej — $x_1, x_2, \ldots$:

Można tak

$$\begin{aligned} x - 2y + 2z - w &= 1, \\ -2x + 4y - z + 2w &= 1, \\ x + 0y + 3z + w &= 3, \end{aligned}$$

Tak wygodniej

$$\begin{aligned} x_1 - 2x_2 + 2x_3 - x_4 &= 1, \\ -2x_1 + 4x_2 - x_3 + 2x_4 &= 1, \\ x_1 + 0x_2 + 3x_3 + x_4 &= 3. \end{aligned}$$

Większe korzyści przynosi użycie rachunku macierzowego, który pozwala zapisać lewe strony układu za pomocą iloczynu macierzy

$$\begin{bmatrix} 1 \cdot x_1 - 2 \cdot x_2 + 2 \cdot x_3 - 1 \cdot x_4 \\ -2 \cdot x_1 + 4 \cdot x_2 - 1 \cdot x_3 + 2 \cdot x_4 \\ 1 \cdot x_1 + 0 \cdot x_2 + 3 \cdot x_3 + 1 \cdot x_4 \end{bmatrix} = \begin{bmatrix} 1 & -2 & 2 & -1 \\ -2 & 4 & -1 & 2 \\ 1 & 0 & 3 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 3 \end{bmatrix}.$$

Układy równań liniowych II

Układ równań liniowych (krótko *układ liniowy*) o n niewiadomych zapisujemy symbolicznie

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

lub

$$\begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

Współczynniki a_{ij} występujące przy niewiadomych nazywamy współczynnikami układu, a $b_1, \ldots, b_m$ po prawej stronie znaku równości nazywamy wyrazami wolnymi, (lub prawymi stronami) układu.

Układy liniowe — notacja macierzowa

Ze współczynników układu powstaje macierz układu A, a po dołączeniu kolumny złożonej z wyrazów wolnych tzw. macierz rozszerzona układu $[\,A \mid b\,]$.

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \dots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}; \quad [\,A \mid b\,] = \left[\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \vdots & \vdots & \dots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{array}\right]$$

Macierz układu, **Macierz rozszerzona układu.**

W formie tych tablic (macierzy) przechowywany jest (także w pamięci komputera) komplet danych jednoznacznie wyznaczających układ równań liniowych — macierzy rozszerzonej odpowiada układ (1)
Jeśli wszystkie wyrazy wolne są zerami, $b_1 = b_2 = \ldots = b_m = 0$, to mówimy, że układ (1) jest *jednorodny*, w przeciwnym przypadku — *niejednorodny*.

Co to znaczy rozwiązać układ

Definicja (Rozwiązanie układu)

Rozwiązaniem układu liniowego

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

nazywamy każdy taki ciąg n liczb $r_1, ..., r_n$, które po podstawieniu $x_1 = r_1, x_2 = r_2, ..., x_n = r_n$ na miejsce niewiadomych przekształcają każde z równań tego układu w tożsamość, tzn. spełniają równość

$$\begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{bmatrix} \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} \tag{2}$$

Liczby $r_1, ..., r_n$ nazywamy składowymi rozwiązania.

Przykład

Przykład

Dla układu

$$\begin{aligned} x_1 - 2x_2 + 2x_3 - x_4 &= 1,\\ -2x_1 + 4x_2 - x_3 + 2x_4 &= 1,\\ x_1 + 0x_2 + 3x_3 + x_4 &= 3. \end{aligned}$$

ciag $(0, \frac{1}{2}, 1, 0)$ *jest rozwiązaniem (jednym z wielu). Istotnie*

$$\begin{bmatrix} 1 & -2 & 2 & -1 \\ -2 & 4 & -1 & 2 \\ 1 & 0 & 3 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ \frac{1}{2} \\ 1 \\ 0 \end{bmatrix} = 0 \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} + \frac{1}{2} \begin{bmatrix} -2 \\ 4 \\ 0 \end{bmatrix} + 1 \begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix} + 0 \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 3 \end{bmatrix}.$$

Natomiast ciąg z przestawionymi elementami $(1, \frac{1}{2}, 0, 0)$ *nie jest rozwiązaniem tego układu*

$$\begin{bmatrix} 1 & -2 & 2 & -1 \\ -2 & 4 & -1 & 2 \\ 1 & 0 & 3 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ \frac{1}{2} \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} + \frac{1}{2} \begin{bmatrix} -2 \\ 4 \\ 0 \end{bmatrix} + 0 \begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix} + 0 \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \neq \begin{bmatrix} 1 \\ 1 \\ 3 \end{bmatrix}.$$

Układ liniowy w nowej perspektywie

Dla ogólnego układu $Ax = b$ podstawienie $x = (r_1, \ldots, r_n)$ daje

$$Ar = \sum_{i=1}^{n} r_i A_i = b, \tag{3}$$

i w postaci rozwiniętej jako

$$\sum_{i=1}^{n} r_i A_i = r_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} + r_2 \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} + \ldots + r_n \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}, \tag{4}$$

To pokazuje, że kolumna wyrazów wolnych jest kombinacją liniową kolumn macierzy A o współczynnikach będących składowymi rozwiązania. Stąd wniosek

Wniosek

Jeśli układ (1) *ma rozwiązanie, to kolumna prawych stron jest kombinacją liniową kolumn macierzy współczynników.*

Układy liniowe — jednorodne i niejednorodne

Z każdym ogólnym układem liniowym związany jest jednorodny układ liniowy o tej samej macierzy współczynników i wszystkimi wyrazami wolnymi równymi zeru

Układ ogólny

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

Układ jednorodny

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= 0, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n &= 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots&\ldots\ldots\ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n &= 0. \end{aligned}$$

który będziemy nazywać *układem jednorodnym odpowiadającym temu układowi (stowarzyszonym z nim)*.

Klasyfikacja układów liniowych

Układ może być:

- sprzeczny — taki układ, który nie ma żadnego rozwiązania;
- niesprzeczny — taki układ, który ma przynajmniej jedno rozwiązanie;
- oznaczony — taki układ, który ma dokładnie jedno rozwiązanie;
- nieoznaczony — układ niesprzeczny, który ma więcej niż jedno rozwiązanie;

W szczególności

Stwierdzenie

Każdy układ jednorodny jest niesprzeczny.

Rzeczywiście, nadając każdej z niewiadomych $x_1, x_2, \ldots, x_n$ w układzie (1) wartość zero otrzymujemy rozwiązanie układu jednorodnego.

Centralny problem: Dla danego układu liniowego opisać jego zbiór rozwiązań.

Wielkość zbioru rozwiązań układu liniowego

TYP UKŁADU LINIOWEGO				
			$m < n$	
	Ogólny	Jednorodny	Niejednorodny	Jednorodny
Możliwa liczba rozwiązań	$0,\ 1,\ \infty$	1 lub ∞	0 lub ∞	∞

Metoda Gaussa eliminacji niewiadomych I

Pokażemy, jak systematyczne zastosowanie metody eliminacji kolejnych niewiadomych prowadzi do wyznaczenia rozwiązania układu równań liniowych. Metoda ta jest podstawą szeroko stosowanego ***algorytmu Gaussa eliminacji zmiennych***. Rozważymy poprzedni przykład w obu zapisach:

$$\begin{aligned} x_1 - 2x_2 + 2x_3 - x_4 &= 1,\\ -2x_1 + 4x_2 - x_3 + 2x_4 &= 1,\\ x_1 + 0x_2 + 3x_3 + x_4 &= 3, \end{aligned} \qquad \begin{bmatrix} 1 & -2 & 2 & -1\\ -2 & 4 & -1 & 2\\ 1 & 0 & 3 & 1 \end{bmatrix} \begin{bmatrix} x_1\\ x_2\\ x_3\\ x_4 \end{bmatrix} = \begin{bmatrix} 1\\ 1\\ 3 \end{bmatrix}.$$

Dokonajmy następujących przekształceń układu: Dodajmy do równania II podwojone rówanie I, a następnie od równania III odejmijmy równanie I. Na końcu zamieńmy miejscami otrzymane równania II i III. Przy tych przekształceniach rozwiązania układu nie ulegają zmianie, a sam układ i jego macierz rozszerzona przyjmą postać:

$$\begin{aligned} x_1 - 2x_2 + 2x_3 - x_4 &= 1,\\ 0x_1 + 2x_2 + 1x_3 + 2x_4 &= 2,\\ 0x_1 + 0x_2 + 3x_3 + 0x_4 &= 3, \end{aligned} \qquad \left[\begin{array}{cccc|c} 1 & -2 & 2 & -1 & 1\\ 0 & 2 & 1 & 2 & 2\\ 0 & 0 & 3 & 0 & 3 \end{array}\right].$$

Metoda Gaussa eliminacji niewiadomych II

Jak poprzednio, współczynniki przekształconych równań zajmują miejsca w kolejnych wierszach przekształconej macierzy współczynników. Rozwiązanie układu wyznaczamy rekurencyjnie, zaczynając od ostatniego równania, w którym jest najmniejsza liczba niewiadomych. Rozwiązaniem równania

$$3x_3 + 0x_4 = 3,$$

jest każdy ciąg $(x_3, x_4) = (1. t)$, gdzie t oznacza dowolną liczbę rzeczywistą, $t \in \mathbb{R}$. Po podstawieniu tych wartości do drugiego równania

$$0x_1 + 2x_2 + 1x_3 + 2x_4 = 2,$$

pozostaje do wyznaczenia tylko jedna niewiadoma, x_2, dla której dostajemy

$$x_2 = \frac{1}{2}(2 - x_3 - 2x_4) = \frac{1}{2} - t.$$

Analogicznie z pierwszego równania $x_1 - 2x_2 + 2x_3 - x_4 = 1$ otrzymamy

$$x_1 = 1 + 2x_2 - 2x_3 + x_4 = -t.$$

Przekształcenia elementarne układu liniowego

Metoda eliminacji Gaussa polega na sekwencyjnym stosowaniu operacji elementarnych.

Definicja (Operacje elementarne na równaniach układu)

Operacjami elementarnymi na równaniach układu będziemy nazywać następujące przekształcenia układu równań liniowych (1):
typ (I) — zamiana miejscami dwóch równań bez zmiany pozostałych;
typ (II) — dodanie do jednego z równań *innego* równania pomnożonego przez stałą;
typ (III) — pomnożenie równania przez niezerową stałą.
Odpowiadające tym operacjom na równaniach układu operacje na wierszach macierzy współczynników układu polegające na:
typ (I) — zamianie miejscami wierszy macierzy układu;
typ (II) — dodaniu do jednego wiersza innego wiersza pomnożonego przez stałą;
typ (III) — pomnożeniu wiersza przez niezerową stałą,
będą nazywane *operacjami elementarnymi na wierszach macierzy*.
Po dokonaniu operacji elementarnych na układzie liniowym zbiór jego rozwiązań nie ulega zmianie — mówimy, że *układy są równoważne*.

Schemat blokowy gaussowskiej eliminacji

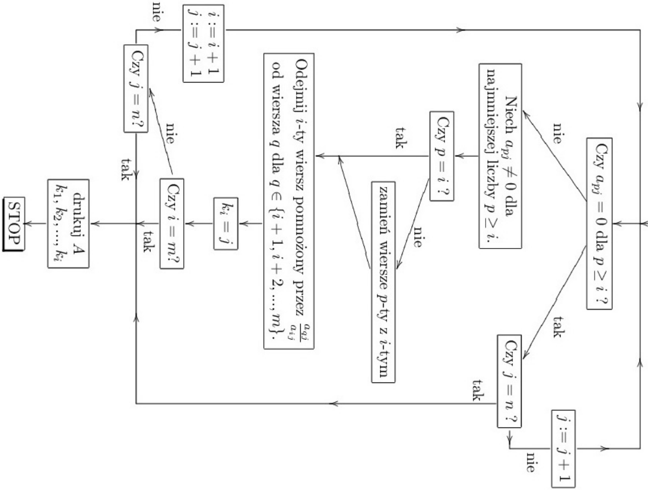


Tabela 2.1: Schemat blokowy gaussowskiej eliminacji

Algorytm eliminacji Gaussa

Stwierdzenie

Każdy układ liniowy o n niewiadomych postaci (1) *z co najmniej jednym* $a_{ij} \neq 0$ *można sprowadzić do postaci zredukowanej, inaczej schodkowej, przez wykonanie na nim skończonej liczby operacji elementarnych.*

$$
\begin{aligned}
\overline{a}_{1k}x_k + \ldots + \overline{a}_{1l}x_l + \ldots + \overline{a}_{1q}x_q + \ldots + \overline{a}_{1s}x_s + \ldots + \overline{a}_{1n}x_n &= \overline{b}_1, \\
\overline{a}_{2l}x_l + \ldots + \overline{a}_{2q}x_q + \ldots + \overline{a}_{2s}x_s + \ldots + \overline{a}_{2n}x_n &= \overline{b}_2, \\
\overline{a}_{3q}x_q + \ldots + \overline{a}_{3s}x_s + \ldots + \overline{a}_{3n}x_n &= \overline{b}_3, \\
\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
\overline{a}_{rs}x_s + \ldots + \overline{a}_{rn}x_n &= \overline{b}_r, \\
0 &= \overline{b}_{r+1}, \\
\vdots &= \vdots \\
0 &= \overline{b}_m,
\end{aligned}
\tag{5}
$$

ze współczynnikami $\overline{a}_{1k} \neq 0$, $\overline{a}_{2l} \neq 0$, $\overline{a}_{3q} \neq 0$, $\overline{a}_{rs} \neq 0$.

Liczba r niezerowych równań w postaci zredukowanej jest zawsze taka sama i nazywa się ją rzędem macierzy układu.

Schemat macierzy w postaci schodkowej I

Macierz dołączona układu (5) ma postać

$$[\,\overline{A}\mid\overline{b}\,]=\left[\begin{array}{ccccccccccc|c}
0 & \dots & \overline{a}_{1k} & \dots & \overline{a}_{1l} & \dots & \overline{a}_{1q} & \dots & \overline{a}_{1s} & \dots & \overline{a}_{1n} & \overline{b}_1\\
0 & \dots & 0 & \dots & \overline{a}_{2l} & \dots & \overline{a}_{2q} & \dots & \overline{a}_{2s} & \dots & \overline{a}_{2n} & \overline{b}_2\\
0 & \dots & 0 & \dots & 0 & \dots & \overline{a}_{3q} & \dots & \overline{a}_{3s} & \dots & \overline{a}_{3n} & \overline{b}_3\\
\multicolumn{11}{c|}{\dotfill} & \dots\\
0 & \dots & 0 & \dots & 0 & \dots & 0 & \dots & \overline{a}_{rs} & \dots & \overline{a}_{rn} & \overline{b}_r\\
0 & \dots & 0 & \dots & 0 & \dots & 0 & \dots & 0 & \dots & 0 & \overline{b}_{r+1}\\
0 & \dots & 0 & \dots & 0 & \dots & 0 & \dots & 0 & \dots & 0 & 0\\
\multicolumn{11}{c|}{\dotfill} & \dots\\
0 & \dots & 0 & \dots & 0 & \dots & 0 & \dots & 0 & \dots & 0 & 0
\end{array}\right].$$

Schemat macierzy w postaci schodkowej II

Warto zapamiętać ogólny schemat macierzy w postaci schodkowej, gdzie niezerowe elementy reprezentowane są przez symbol $\blacksquare$, a $*$ reprezentują dowolne (być może również zera) elementy macierzy.

$$\begin{bmatrix} 0 & \blacksquare & * & * & * & * & * & * & * & * \\ 0 & 0 & 0 & \blacksquare & * & * & * & * & * & * \\ 0 & 0 & 0 & 0 & \blacksquare & * & * & * & * & * \\ 0 & 0 & 0 & 0 & 0 & \blacksquare & * & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \blacksquare & * \end{bmatrix}.$$

Wykład 2 (7.03.2019)

Zaczynamy pracę

Kartezjańska przestrzeń n-wymiarowa

Zbiór n-elementowych ciągów $u = (u_1, \ldots, u_n)$ o wyrazach rzeczywistych oznaczamy symbolem $\mathbb{R}^n$;

$$\mathbb{R}^n = \{ (u_1, \ldots, u_n) \mid u_i \in \mathbb{R},\ i = 1, \ldots, n \}.$$

Tradycyjnie oznaczamy ciąg obejmując jego wyrazy okrągłymi nawiasami, dla odróżnienia od oznaczenia zbioru za pomocą nawiasów klamrowych. Zapis $u = (u_1, \ldots, u_n)$ oznacza więc ciąg nazywany u, którego wyrazami są liczby $u_1, \ldots, u_n$. Ciąg złożony z samych zer będziemy oznaczać krótko symbolem 0,

$$0 = (0, \ldots, 0).$$

W zbiorze $\mathbb{R}^n$ wprowadzamy „działania"

dodawanie: $$(u_1, \ldots, u_n) + (v_1, \ldots, v_n) = (u_1 + v_1, \ldots u_n + v_n); \tag{6}$$

mnożenie przez liczbę $a \in \mathbb{R}$: $$a(u_1, \ldots, u_n) = (au_1, \ldots, au_n). \tag{7}$$

Zbiór $\mathbb{R}^n$ rozpatrywany wspólnie z tymi działaniami będziemy nazywali *n-wymiarową rzeczywistą przestrzenią kartezjańską*.

W analogiczny sposób definiujemy *n-wymiarową zespoloną przestrzeń kartezjańską*, oznaczaną $\mathbb{C}^n$.

Notacja macierzowa

Macierze o rozmiarach $1 \times n$ i wyrazach z $\mathbb{R}$, zapisywane jako $u = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}$, możemy utożsamiać z n-elementowymi ciągami o wyrazach z $\mathbb{R}$ i analogicznie w przypadku macierzy $m \times 1$, $v = \begin{bmatrix} v_1 & v_2 & \dots & v_m \end{bmatrix}$.

Ściśle mówiąc, odwzorowujemy bijektywnie przestrzeń $\mathbf{M}_{1\times n}(\mathbb{R})$ i odpowiednio $\mathbf{M}_{m\times 1}(\mathbb{R})$, na przestrzeń $\mathbb{R}^n$, odpowiednio $\mathbb{R}^m$. Przy tej bijekcji działania na macierzach określone znanymi wzorami dokładnie odpowiadają działaniom na ciągach określonym wzorami (6) i (7) — w każdym przypadku działania te wykonywane są „po współrzędnych".

Działania te mają własności: dodawanie — *łączności, przemienności, istnienie* 0 *i odwracalności*;

Mnożenie przez liczby (skalary) — *łączności, unitarności i rozdzielności względem dodawania*.

Dla $j \leqslant n$ będziemy oznaczać przez $e_j = (0, \dots, 0, 1, 0, \dots, 0)$ taki ciąg, którego jedyny element różny od 0 jest równy 1 i występuje na miejscu j.

Model — dlaczego wektory?

Przykład (Zaopatrzenie)

Wektor $u = (u_1, \ldots, u_m)$ opisuje stan zaopatrzenia hipermarketu w sprzedawane tam towary —zamiast nazwy towaru używamy numeru identyfikującego (np. kod kreskowy).
Wektor $b = (b_1, \ldots, b_m)$ to zestaw towarów sprzedanych w ciągu tygodnia.
Towary mogą być dostarczane do hipermarketu w pakietach, każdy pakiet to $z^{(1)} = (z_1^{(1)}, \ldots, z_m^{(1)})$, $z^{(2)} = (z_1^{(2)}, \ldots, z_m^{(2)})$, ... Dostawa zawierająca x_1 pakietów $z^{(1)}$, x_2 pakietów $z^{(2)}$ itd. uzupełnia zapas towarów w magazynie $S = (s_1, \ldots, s_n)$ o

$$d = x_1 z^{(1)} + x_2 z^{(2)} + x_3 z^{(3)} + \ldots + x_n z^{(n)}$$

W wygodnym (oszczędnym) zapisie kolumnowym

$$= \begin{bmatrix} x_1 z_1^{(1)} + x_2 z_1^{(2)} + \ldots + x_n z_1^{(n)} \\ x_1 z_2^{(1)} + x_2 z_2^{(2)} + \ldots + x_n z_2^{(n)} \\ \vdots \\ x_1 z_m^{(1)} + x_2 z_m^{(2)} + \ldots + x_n z_m^{(n)} \end{bmatrix} = \begin{bmatrix} z_1^{(1)} & z_1^{(2)} & \ldots & z_1^{(n)} \\ z_2^{(1)} & z_2^{(2)} & \ldots & z_2^{(n)} \\ \vdots & & & \\ z_m^{(1)} & z_m^{(2)} & \ldots & z_m^{(n)} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

Dlaczego wektory II

Zatem dla uzupełnienia sprzedaży potrzebujemy spełnienia relacji

$$\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} = d = \begin{bmatrix} z_1^{(1)} & z_1^{(2)} & \dots & z_1^{(n)} \\ z_2^{(1)} & z_2^{(2)} & \dots & z_2^{(n)} \\ \vdots & & & \\ z_m^{(1)} & z_m^{(2)} & \dots & z_m^{(n)} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

Problem

Czy istnieje rozwiązanie — niesprzeczność układu;
czy jest jednoznaczne — oznaczoność układu?

Przestrzeń liniowa — definicja i przykłady

Definicja (Przestrzeń liniowa, I — dodawanie wektorów)

Przestrzenią liniową nad ciałem $\mathbb{K}$ nazywamy zbiór V, wraz z odwzorowaniami, nazywanymi odpowiednio *dodawaniem* i *mnożeniem przez liczby z ciała* $\mathbb{K}$,

$$V \times V \ni (v, w) \mapsto v + w \in V, \qquad \mathbb{K} \times V \ni (\alpha, v) \mapsto \alpha v \in V, \tag{8}$$

spełniające następujące warunki:

(I) **Własności dodawania**

Dla każdych $u, v, w \in V$

$$u + (v + w) = (u + v) + w \qquad \textbf{łączność dodawania}; \tag{W1}$$

$$u + v = v + u \qquad \textbf{przemienność dodawania}; \tag{W2}$$

Istnieje element $0 \in V$ spełniający dla każdego $v \in V$

$$0 + v = v \qquad \textbf{neutralność } 0 \textbf{ wzg. dodawania}; \tag{W3}$$

Dla każdego $v \in V$ istnieje taki element $-v \in V$, że

$$v + (-v) = 0 \qquad \textbf{odwracalność dodawania}; \tag{W4}$$

Element $-v$ jest jedyny i nazywa się *przeciwnym* do v.

Definicja (Przestrzeń liniowa, II — mnożenie przez skalary)

(II) **Własności mnożenia przez skalary**

Ponadto, dla każdych v, $w \in V$ i każdych dwóch skalarów λ, $\mu \in \mathbb{K}$ zachodzą równości

$$\lambda(\mu v) = (\lambda\mu)v \qquad \textbf{łączność}; \tag{W5}$$

$$(\lambda + \mu)v = \lambda v + \mu v \qquad \textbf{prawo rozdzielności mnożenia}; \tag{W6}$$

$$\lambda(v + w) = \lambda v + \lambda w \qquad \textbf{prawo rozdzielności dodawania}; \tag{W7}$$

$$1v = v \qquad \textbf{unitarność mnożenia przez skalary}. \tag{W8}$$

Elementy przestrzeni liniowej będziemy nazywać ogólnie *wektorami*([1]), a elementy ciała $\mathbb{K}$ ($\mathbb{K} = \mathbb{R}$ lub $\mathbb{C}$) — *skalarami*.

Dla $\alpha \in \mathbb{K}$ i $v \in V$ wektor αv nazywamy *skalarną wielokrotnością* wektora v.

[1] Z tego punktu widzenia macierz jest także „wektorem".

Ważny przykład — przestrzeń liniowa funkcji

Definicja (Algebraiczne działania na funkcjach)

Dla dowolnych funkcji $f, g \in \mathcal{F}(X, \mathbb{K})$ i dowolnej liczby $\alpha \in \mathbb{K}$ są określone:

$$\left.\begin{array}{ll} \textbf{suma funkcji:} & (f+g)(x) = f(x) + g(x); \\ \textbf{iloczyn funkcji przez liczbę:} & (\alpha f)(x) = \alpha f(x), \end{array}\right\} \quad \forall\, x \in X. \tag{9}$$

Na mocy tej definicji i własności działań w ciele $\mathbb{K}$ odwzorowania

$$\mathcal{F}(X, \mathbb{K}) \times \mathcal{F}(X, \mathbb{K}) \ni (f, g) \mapsto f + g \in \mathcal{F}(X, \mathbb{K}), \tag{10}$$

$$\mathbb{K} \times \mathcal{F}(X, \mathbb{K}) \ni (\alpha, f) \mapsto \alpha f \in \mathcal{F}(X, \mathbb{K}) \tag{11}$$

spełniają reguły W1–W8. W szczególności elementem neutralnym względem dodawania jest funkcja równa tożsamościowo zeru — będziemy ją oznaczać symbolem 0, a elementem przeciwnym do funkcji f jest funkcja $-f$, określona wzorem $(-f)(x) = -f(x)$ dla każdego $x \in X$.

Zbiór $\mathcal{F}(X, \mathbb{K})$ ma strukturę przestrzeni liniowej nad ciałem $\mathbb{K}$.

Menażeria przestrzeni liniowych (Podręcznik, **4.1, 4.4**)

Przykłady

I. Przestrzenie liniowe macierzy

- *Przestrzeń* $\mathbf{M}_{m\times n}(\mathbb{R})$*, której elementami są macierze*

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \dots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix};$$

- *Przestrzenie* $\mathbf{M}_{n\times 1}(\mathbb{R})$ *i* $\mathbf{M}_{1\times n}(\mathbb{R})$ *wektorów kolumnowych i wierszowych.*
- *Przestrzeń* $\mathbf{M}_n(\mathbb{R}) = \mathbf{M}_{n\times n}(\mathbb{R})$ *macierzy kwadratowych stopnia* n*;*
- *Przestrzeń* $\mathbf{T}_n(\mathbb{R}) \subset \mathbf{M}_n(\mathbb{R})$ *macierzy górno-trójkątnych postaci*

$$T = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ 0 & a_{22} & a_{23} & \dots & a_{2n} \\ 0 & 0 & \ddots & \dots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & a_{nn} \end{bmatrix}.$$

Menażeria przestrzeni liniowych, II

Przykłady

II. Przestrzenie wielomianów, ciągów nieskończonych i funkcji

- *Przestrzeń $\mathcal{P}^n[t]$ funkcji wielomianowych stopnia nie większego niż n, tj. funkcji postaci*

$$f(t) = f_0 + f_1 t + f_2 t^2 + \ldots + f_n t^n, \qquad t \in \mathbb{R},\ f_j \in \mathbb{R}.$$

- *Przestrzeń ciągów nieskończonych $(a_n)_{n=1}^{\infty}$ o wyrazach $a_n \in \mathbb{R}$.*
- *Przestrzeń funkcji ciągłych o wartościach rzeczywistych na osi liczbowej $\mathbb{R}$, lub na odcinku $[0, 1] \subset \mathbb{R}$.*
- *Przestrzeń funkcji na osi liczbowej $\mathbb{R}$, lub na odcinku $[0, 1] \subset \mathbb{R}$ przyjmujących wartości w przestrzeni liniowej V nad ciałem $\mathbb{R}$.*

Twierdzenie

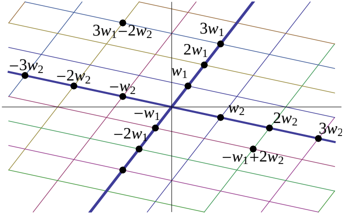
Każdy z wymienionych powyżej zbiorów wyposażony w naturalne działania dodawania i mnożenia przez liczby rzeczywiste (skalary) spełnia własności W1–W8 sformułowane w definicji przestrzeni liniowej.

Kombinacje liniowe

Definicja (Kombinacja liniowa układu wektorów)

Przy danych wektorach $v, w \in V$ ich kombinacją liniową nazywamy dowolny wektor postaci $\alpha v + \beta w$, gdzie $\alpha, \beta \in \mathbb{K}$.
Dla wektorów $x = (x_1, x_2),\ y = (y_1, y_2) \in \mathbb{R}^2$ konstrukcję ilustruje rysunek.

$$\alpha \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \beta \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} \alpha x_1 + \beta y_1 \\ \alpha 2_1 + \beta y_2 \end{bmatrix}.$$



Kombinacje liniowe na płaszczyźnie

Ogólnie, kombinacją liniową układu wektorów $\{v_1, \ldots, v_k\}$ ze współczynnikami $\{\alpha_1, \ldots, \alpha_k\}$ nazywamy wektor

$$w = \alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_k v_k = \sum_{j=1}^{k} \alpha_j v_j \qquad (12)$$

Kombinacje liniowe kolumn macierzy

Macierz $A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \vdots & \vdots & \dots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{bmatrix} = \begin{bmatrix} A_1 & A_2 & \dots & A_m \end{bmatrix} \in \mathbf{M}_{n\times m}(\mathbb{R})$

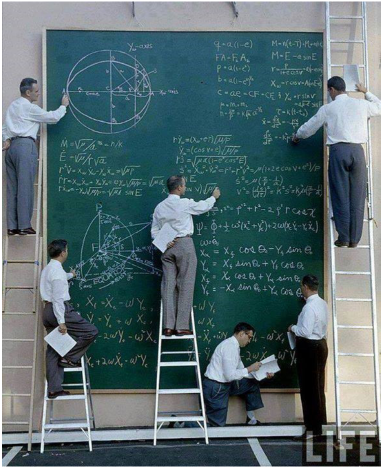
traktujemy jako ciąg kolumn A_i, $i \in \{1, \dots, m\}$, gdzie $A_i = \begin{bmatrix} a_{1i} \\ \vdots \\ a_{ni} \end{bmatrix} \in \mathbb{R}^n$ itd.

Liczby $w_1, \dots, w_m \in \mathbb{R}$ traktujmy jako składowe wektora kolumnowego $w = \begin{bmatrix} w_1 \\ \vdots \\ w_m \end{bmatrix} \in \mathbb{R}^m$. Wówczas wektor

$$Aw = \sum_{i=1}^{m} w_i A_i = w_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{bmatrix} + w_2 \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{n2} \end{bmatrix} + \dots + w_m \begin{bmatrix} a_m \\ a_{2m} \\ \vdots \\ a_{nm} \end{bmatrix} \tag{13}$$

jest kombinacją kolumn macierzy A.

Wykład 3 (14.03.2019)



Zaczynamy pracę

Podprzestrzeń przestrzeni liniowej

Definicja (Podprzestrzeń liniowa)

Niepusty podzbiór W przestrzeni liniowej V nad ciałem $\mathbb{K}$ będziemy nazywać *podprzestrzenią liniową przestrzeni* V, jeśli jest zamknięty względem kombinacji liniowych, tzn.
Dla każdej pary $v, w \in W$ i każdej pary skalarów $\alpha, \beta \in \mathbb{K}$, kombinacja liniowa $\alpha v + \beta w$ także jest elementem W.

Stwierdzenie

$W \subset V$ jest podprzestrzenią liniową przestrzeni V wtedy i tylko wtedy, gdy dla każdych $v, w \in V$ i każdego $\alpha \in \mathbb{K}$

$$v, w \in W \quad \Longrightarrow \quad v + w \in W; \qquad v \in W \quad \Longrightarrow \quad \alpha v \in W.$$

Jeśli $W \subset V$ jest podprzestrzenią liniową przestrzeni V, to ograniczenia do W odwzorowań dodawania wektorów i mnożenia ich przez liczby

$$W \times W \ni (v, w) \mapsto v + w \in W, \qquad \mathbb{K} \times W \ni (\alpha, v) \mapsto \alpha v \in W,$$

spełniają aksjomaty W1–W8 przestrzeni liniowej.

Podprzestrzenie przestrzeni liniowej

Stwierdzenie

Każda podprzestrzeń liniowa W przestrzeni V, rozpatrywana samoistnie, jest przestrzenią liniową względem działań przeniesionych (odziedziczonych) z przestrzeni V.

W szczególności:

a) $0 \in W$;

b) $v \in W \implies \mathbb{K}\, v = \{\, \alpha v \mid \alpha \in \mathbb{K} \,\} \subset W$.

Przykłady

- **Podprzestrzeń trywialna** Zbiór złożony z samego wektora zerowego przestrzeni liniowej V jest podprzestrzenią liniową przestrzeni V.
- **Podprzestrzeń generowana przez wektor** Dla każdego niezerowego wektora v przestrzeni liniowej V nad ciałem $\mathbb{K}$ zbiór $\{\, \alpha v \mid \alpha \in \mathbb{K} \,\}$ skalarnych wielokrotności wektora v jest podprzestrzenią liniową przestrzeni V.
- Podzbiór $Z = \{\, (x_1, \ldots, x_k, x_{k+1}, \ldots, x_n) \in \mathbb{R}^n \mid x_{k+1} = \ldots = x_n = 0 \,\}$ jest podprzestrzenią liniową przestrzeni $\mathbb{R}^n$.
- Podzbiór $D = \left\{ \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \mid a_{12} = a_{21} = 0 \right\}$ jest podprzestrzenią liniową przestrzeni $\mathbf{M}_2(\mathbb{R})$ macierzy kwadratowych stopnia 2.
- Podzbiór przestrzeni $\mathcal{P}^3[t]$ wielomianów stopnia nie większego niż 3 złożony z wielomianów $p(t)$ spełniających warunek $p(0) = p(1) = 0$ jest podprzestrzenią liniową.

Przestrzenie liniowe związane z układami liniowymi I

Niech $A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & \vdots & \dots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \in \mathbf{M}_{m\times n}(\mathbb{R})$.

Zbiór $\{\, x \in \mathbb{R}^n \mid Ax = 0 \,\}$ rozwiązań jednorodnego układu równań z macierzą współczynników A, będziemy nazywali *przestrzenią zerową*, lub *jądrem* macierzy A i oznaczali $\mathcal{N}(A)$.
Zbiór $\{\, y \in \mathbb{R}^m \mid y = Ax \,\}$ będziemy nazywali *przestrzenią kolumn*, lub *obrazem* macierzy A i oznaczali $\mathcal{R}(A)$.

Twierdzenie

Przestrzeń zerowa macierzy A — $\mathcal{N}(A) \subset \mathbb{R}^n$ oraz przestrzeń $\mathcal{R}(A) \subset \mathbb{R}^m$ są zamknięte względem kombinacji liniowych, tzn. są podprzestrzeniami liniowymi odpowiednio $\mathbb{R}^n$ i $\mathbb{R}^m$.
Zatem $\mathcal{N}(A)$, $\mathcal{R}(A)$ są przestrzeniami liniowymi nad ciałem $\mathbb{R}$.

Przestrzenie liniowe związane z układami liniowymi II

Dla wykazania tego twierdzenia zauważamy, że jeśli v, $w \in \mathcal{N}(A)$, to dla dowolnych α, $\beta \in \mathbb{R}$ mamy

$$A(\alpha v + \beta w) = \alpha Av + \beta Aw = 0$$

więc $(\alpha v + \beta w) \in \mathcal{N}(A)$.
Podobnie, jeśli v, $w \in \mathcal{R}(A)$, więc $v = Au$, $w = Az$ dla pewnych u, $z \in \mathbb{R}^n$, to dla dowolnych α, $\beta \in \mathbb{R}$

$$\alpha v + \beta w = \alpha Au + \beta Az = A(\alpha u + \beta z),$$

więc $(\alpha v + \beta w) \in \mathcal{R}(A)$.
$\mathcal{R}(A)$ nazywamy przestrzenią kolumn macierzy A ponieważ

$$\mathcal{R}(A) = \{\, t_1 A_1 + t_2 A_2 + \ldots + t_n A_n \mid t_1,\, t_2,\, \ldots,\, t_n \in \mathbb{R} \,\},$$

gdzie $A_1, \ \ldots, \ A_n$ oznaczają kolumny macierzy A (będące wektorami z $\mathbb{R}^m$);

$$A = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ \vdots & \vdots & \ldots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{bmatrix} = \begin{bmatrix} A_1 & A_2 & \ldots & A_n \end{bmatrix}$$

Podprzestrzenie liniowe, rozpinanie podprzestrzeni

Stwierdzenie (Powłoka liniowa układu wektorów)

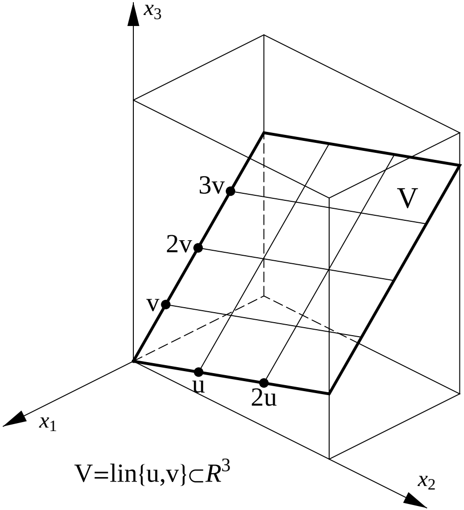
Dla każdego układu $v_1, \ldots, v_k$ elementów przestrzeni liniowej V nad $\mathbb{K}$ zbiór wszystkich kombinacji liniowych wektorów tego układu

$$\{v = \sum_{j=1}^{k} \lambda_j v_j \in V \mid \lambda_j \in \mathbb{K}, j = 1, \ldots, k\} = \mathrm{lin}_{\mathbb{K}}\{v_1, \ldots, v_k\} = \mathrm{lin}\{v_1, \ldots, v_k\}$$

jest zamknięty względem kombinacji liniowych. Zbiór ten jest nazywany powłoką liniową (nad ciałem $\mathbb{K}$) układu $v_1, \ldots, v_k$.

A zatem zbiór $\mathrm{lin}_{\mathbb{K}}\{v_1, \ldots, v_k\}$ jest podprzestrzenią liniową przestrzeni V — nazywamy ją *przestrzenią rozpiętą* lub *generowaną* nad ciałem $\mathbb{K}$ przez układ $v_1, \ldots, v_k$.
W szczególności, każdy taki zbiór zawiera wektor zerowy (element neutralny względem dodawania) i wraz z każdym elementem $v \in V$ zawiera także jego element przeciwny $-v$.

Układ wektorów rozpinający podprzestrzeń



$V = \mathrm{lin}\{u, v\} \subset R^3$

Podprzestrzeń $\mathrm{lin}\{u, v\}$ jest rozpięta przez wektory u, v

Liniowe relacje w przestrzeni liniowej

Definicja (Liniowo zależne i liniowo niezależne układy wektorów)

Układ $v_1, \ldots, v_k$ wektorów przestrzeni liniowej V nazywamy układem *liniowo zależnym*, jeśli istnieją takie liczby $\lambda_1, \ldots, \lambda_k$, nie wszystkie równe zeru, że

$$\lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_k v_k = 0. \tag{14}$$

Związek postaci $\lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_k v_k = 0$, w którym przynajmniej jeden ze współczynników λ_j jest różny od zera, będziemy nazywali *relacją liniową* między wektorami układu $v_1, \ldots, v_k$.
Jeśli układ nie jest liniowo zależny, to nazywamy go *liniowo niezależnym*.

Stwierdzenie

Układ $v_1, \ldots, v_k$ wektorów przestrzeni liniowej V jest liniowo niezależny wtedy i tylko wtedy, gdy dla każdego układu liczb $\lambda_1, \ldots, \lambda_k$ jest spełniona równoważność

$$\lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_k v_k = 0 \quad \Longleftrightarrow \quad \lambda_1 = \lambda_2 = \ldots = \lambda_k = 0. \tag{15}$$

Układy o małej liczbie wektorów

Stwierdzenie

a) Układ zawierający wektor 0 *jest liniowo zależny.*
b) Układ złożony z jednego różnego od zera wektora jest liniowo niezależny.
c) Jeśli układ dwóch wektorów $v, w \in V$ *jest liniowo zależny, to albo przynajmniej jeden z nich jest wektorem zerowym, albo istnieje różny od zera element* $\lambda \in \mathbb{R}$*, taki że* $v = \lambda w$*. W tym ostatnim przypadku mówimy, że wektory* v, w *są współliniowe.*

Wykażemy *c)*. W przypadku gdy $v \neq 0$ i $w \neq 0$ oba współczynniki kombinacji liniowej $\alpha v + \beta w = 0$ są różne od zera (dlaczego?) i wówczas

$$v = -\frac{\beta}{\alpha} w, \qquad \text{i także} \qquad w = -\frac{\alpha}{\beta} v.$$

Pozostała część wypowiedzi wynika bezpośrednio z punktu *a)*.

Liniowa niezależność wektorów przestrzeni kartezjańskiej

Stwierdzenie

Układ $v_1, \ldots, v_k$ wektorów przestrzeni liniowej $\mathbb{R}^m$ jest liniowo niezależny wtedy i tylko wtedy, gdy przestrzeń zerowa macierzy $A = \begin{bmatrix} v_1 & v_2 & \ldots & v_k \end{bmatrix}$ o kolumnach $v_1, \ldots, v_k$ zawiera tylko wektor zerowy, $\mathcal{N}(A) = \{0\}$.

Zapiszmy wektory $v_1, \ldots, v_k$ za pomocą współrzędnych w postaci kolumnowej,

$$v_1 = \begin{bmatrix} v_{11} \\ v_{21} \\ \vdots \\ v_{m1} \end{bmatrix}, \; v_2 = \begin{bmatrix} v_{12} \\ v_{22} \\ \vdots \\ v_{m2} \end{bmatrix}, \; \ldots v_k = \begin{bmatrix} v_{1k} \\ v_{2k} \\ \vdots \\ v_{mk} \end{bmatrix} \quad \text{zatem} \quad A = \begin{bmatrix} v_{11} & \ldots & v_{1k} \\ v_{21} & \ldots & v_{2k} \\ \vdots & & \vdots \\ v_{m1} & \ldots & v_{mk} \end{bmatrix},$$

skąd wynika (por. wcześniejszy wzór (13))

$$\lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_k v_k = A\,z, \qquad \text{gdzie} \quad z = \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_k \end{bmatrix}. \tag{16}$$

Liniowa zależność układu wektorów w $\mathbb{R}^n$

Stwierdzenie

Jeśli $k > n$, to układ wektorów $v_1, \ldots, v_k$ należących do przestrzeni $\mathbb{R}^n$ jest liniowo zależny.

Rzeczywiście, z warunku (16) wynika, że współczynniki λ_j zerowej kombinacji liniowej wektorów układu spełniają układ równań

$$\lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_k v_k = \begin{bmatrix} v_{11} & \ldots & v_{1k} \\ v_{21} & \ldots & v_{2k} \\ \vdots & & \vdots \\ v_{m1} & \ldots & v_{mk} \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_k \end{bmatrix} = 0,$$

w którym jest więcej niewiadomych niż równań, a taki układ ma zawsze rozwiązania niezerowe.

Konsekwencje liniowej zależności układu wektorów

Jeśli układ wektorów $v_1, \ldots, v_k$ przestrzeni liniowej V jest liniowo zależny i $\lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_k v_k = 0$, to dla pewnego wskaźnika $1 \leqslant j \leqslant k$ mamy $\lambda_j \neq 0$, skąd $\lambda_j v_j = -\sum_{i=1,\, i\neq j}^{k} \lambda_i v_i$ i dalej

$$v_j = -\sum_{i=1,\, i\neq j}^{k} \frac{\lambda_i}{\lambda_j} v_i$$

Stwierdzenie

Układ wektorów $v_1, \ldots, v_k$ przestrzeni V jest liniowo zależny wtedy i tylko wtedy, gdy jeden z wektorów układu jest kombinacją liniową pozostałych.

Wniosek

Z każdego liniowo zależnego układu wektorów $v_1, \ldots, v_k$ przestrzeni V można wybrać (pod-)układ liniowo niezależny rozpinający tę samą podprzestrzeń.

Kombinacje liniowe układu liniowo niezależnego

Stwierdzenie

Jeśli układ wektorów $v_1, \ldots, v_k$ przestrzeni liniowej V jest liniowo niezależny, to współczynniki dowolnej kombinacji liniowej tego układu są jednoznacznie wyznaczone;

$$\lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_k v_k = \mu_1 v_1 + \mu_2 v_2 + \ldots + \mu_k v_k \implies \lambda_1 = \mu_1, \ldots, \lambda_k = \mu_k$$

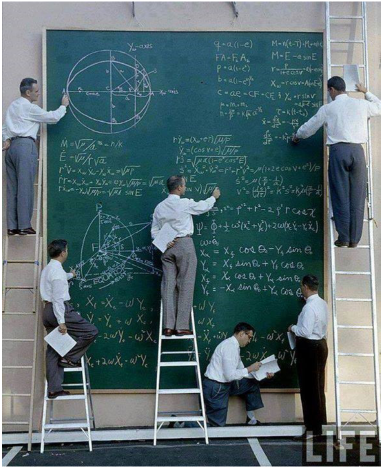
Definicja (Baza przestrzeni liniowej)

Układ wektorów $v_1, \ldots, v_k$ przestrzeni liniowej V nazywa się bazą V, jeśli każdy wektor $v \in V$ jest kombinacją liniową wektorów tego układu z jednoznacznie wyznaczonymi współczynnikami;

$$v = \lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_k v_k = \sum_{j=1}^{k} \lambda_j v_j.$$

Współczynniki λ_j tej kombinacji liniowej nazywamy współrzędnymi wektora v względem bazy $v_1, \ldots, v_k$.

Wykład 4 (21.03.2019)



Zaczynamy pracę

Baza przestrzeni wyznacza jej własności algebraiczne

Twierdzenie (Bazy i współrzędne)

Niech $v_1, \ldots, v_k$ będzie bazą przestrzeni liniowej V.
a) Każdy wektor przestrzeni liniowej V jest jednoznacznie wyznaczony przez swoje współrzędne względem tej bazy:

$$v = \sum_{i=1}^{k} x_i v_i = \sum_{i=1}^{k} z_i v_i \quad \Longleftrightarrow \quad x_i = z_i, \quad i = 1, \ldots, k.$$

b) Dodawanie wektorów, mnożenie przez liczby i liniowe kombinacje wykonuje się „po współrzędnych" względem wybranej bazy :

$$v = \sum_{i=1}^{k} x_i v_i; \quad w = \sum_{i=1}^{k} z_i v_i \quad \Longrightarrow \alpha v + \beta w = \sum_{i=1}^{k} (\alpha x_i + \beta z_i) v_i.$$

c) Przestrzenie liniowe V i W, których bazy $v_1, \ldots, v_k$ i $w_1, \ldots, w_k$ są równoliczne, są pod względem własności algebraicznych jednakowe (izomorficzne).

Przestrzeń $\mathcal{P}^n[t]$ wielomianów stopnia $\leqslant n$ o współczynnikach rzeczywistych jest izomorficzna z przestrzenią $\mathbb{R}^{n+1}$.

Wymiar przestrzeni liniowej

Twierdzenie (Równoliczność baz)

Każde dwie bazy przestrzeni liniowej V mają tę samą liczbę elementów.

Definicja (Wymiar przestrzeni liniowej)

Liczba elementów bazy przestrzeni liniowej V nazywa się wymiarem przestrzeni. Oznacza się ją symbolem $\dim V$.

Stwierdzenie (Konkretne bazy)

Podane układy są bazami w wymienionych przestrzeniach:
a) $e_1 = (1, 0, \ldots, 0), \quad e_2 = (0, 1, 0, \ldots, 0), \ldots, e_n = (0, 0, \ldots, 1)$ *— baza standardowa w przestrzeni* $\mathbb{K}^n$*;*
b) $\{E_{jk} = [\delta_{jk}],\ 0 \leqslant j \leqslant m,\ 0 \leqslant k \leqslant n\}$ *— baza standardowa w przestrzeni macierzy prostokątnych* $\mathbf{M}_{m\times n}(\mathbb{K})$*;*
c) $\{m_k(t) = t^k,\ 0 \leqslant k \leqslant n\}$ *— baza jednomianowa w przestrzeni* $\mathcal{P}^n[t]$*.*

Wymiary

Twierdzenie (Wymiary podstawowych przestrzeni liniowych)

Przestrzenie z Przykładu 1 mają następujące wymiary:

$\dim \mathbb{R}^n = n$;

$\dim \mathbf{M}_{m\times n}(\mathbb{R}) = m \cdot n$, *w szczególności* $\dim \mathbf{M}_{1\times n}(\mathbb{R}) = \dim \mathbf{M}_{n\times 1}(\mathbb{R}) = n$;

$\dim \mathbf{M}_n(\mathbb{R}) = n^2$;

$\dim \mathbf{T}_n(\mathbb{R}) = \dfrac{n(n+1)}{2}$

$\dim \mathcal{P}^n[t] = n + 1$.

Wymiary przestrzeni wierszy (kolumn) macierzy

Stwierdzenie (Rząd macierzy i wymiar przestrzeni wierszy)

Niech $\overline{A}$ będzie macierzą w postaci schodkowej:

$$\overline{A} = \begin{bmatrix} \overline{a}_{11} & \dots & \overline{a}_{1l} & \dots & \overline{a}_{1q} & \dots & \overline{a}_{1s} & \dots & \overline{a}_{1n} \\ 0 & \dots & \overline{a}_{2l} & \dots & \overline{a}_{2q} & \dots & \overline{a}_{2s} & \dots & \overline{a}_{2n} \\ 0 & \dots & 0 & \dots & \overline{a}_{3q} & \dots & \overline{a}_{3s} & \dots & \overline{a}_{3n} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & \dots & 0 & \dots & \overline{a}_{rs} & \dots & \overline{a}_{rn} \\ 0 & \dots & 0 & \dots & 0 & \dots & 0 & \dots & 0 \end{bmatrix}$$

Niezerowe wiersze macierzy w postaci schodkowej są liniowo niezależne i rozpinają przestrzeń liniową wymiaru r ($r = r(A)$ jest rzędem macierzy A).

Twierdzenie (Rząd macierzy)

Wymiar przestrzeni rozpiętej przez kolumny macierzy jest równy wymiarowi przestrzeni rozpiętej przez wiersze macierzy. Tę wspólną wartość nazywamy rzędem macierzy;

$$\dim \mathcal{R}(A) = \dim \mathcal{W}(A) = r(A)$$

Przestrzeń zerowa macierzy i jej wymiar

Twierdzenie (Wymiar przestrzeni zerowej macierzy)

Niech $\mathcal{N}(A) = \{ x \in \mathbb{R}^n \mid Ax = 0 \}$ *będzie przestrzenią zerową macierzy* $A \in \mathbf{M}_{m \times n}(\mathbb{R})$*, tj. zbiorem rozwiązań jednorodnego układu liniowego* $Ax = 0$*.*
Zachodzi równość

$$\dim \mathcal{N}(A) = n - r(A).$$

Jako bazę przestrzeni $\mathcal{N}(A)$ *można przyjąć układ* $k = n - r(A)$ *wektorów* $u_1, \ldots, u_k$*, skonstruowanych następująco.*
Niewiadome wolne w układzie $Ax = 0$*, ustalone w dowolnie wybranej kolejności, nazwiemy* $y_1, y_2, \ldots, y_k$*.*
Rozwiążemy kolejno układ $Ax = 0$ *przyjmując* 1 *jako wartość dla jednej niewiadomej wolnej, a* 0 *dla pozostałych. Każde z takich rozwiązań jest jednoznacznie wyznaczone i daje wektor, który oznaczymy przez* u_j*, gdzie* j *jest aktualnie wybranym numerem niezerowej zmiennej wolnej.*
Można się przekonać, że te wektory tworzą bazę przestrzeni $\mathcal{N}(A)$*.*

Wybór bazy w przestrzeni kolumn macierzy I

Rozważmy układ

$$\begin{array}{rcrcrcrcrcl} x_1 & + & 3x_2 & + & x_3 & - & 3x_4 & + & x_5 & = & 3, \\ -2x_1 & - & 6x_2 & - & 3x_3 & + & 5x_4 & + & 2x_5 & = & 3, \\ x_1 & + & 3x_2 & - & 4x_3 & - & 8x_4 & - & x_5 & = & 4. \end{array}$$

i odpowiadającą mu macierz rozszerzoną

$$[A|b] = \left[\begin{array}{rrrrr|r} \mathbf{1} & 3 & 1 & -3 & 1 & 3 \\ -2 & -6 & -3 & 5 & 2 & 3 \\ 1 & 3 & -4 & -8 & -1 & 4 \end{array}\right]$$

Po sprowadzeniu do postaci schodkowej otrzymamy

$$\left[\begin{array}{rrrrr|r} \mathbf{1} & 3 & 0 & -4 & 0 & 2 \\ 0 & 0 & \mathbf{1} & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 & \mathbf{1} & 2 \end{array}\right]$$

Które kolumny tworzą bazę w przestrzeni $\mathcal{R}(A)$ kolumn macierzy współczynników A?

Wybór bazy w przestrzeni kolumn macierzy II

Układem jednorodnym odpowiadającym temu układowi jest

$$\begin{array}{rcrcrcrcrcl} x_1 & + & 3x_2 & + & x_3 & - & 3x_4 & + & x_5 & = & 0, \\ -2x_1 & - & 6x_2 & - & 3x_3 & + & 5x_4 & + & 2x_5 & = & 0, \\ x_1 & + & 3x_2 & - & 4x_3 & - & 8x_4 & - & x_5 & = & 0, \end{array}$$

z macierzą zredukowaną postaci $\begin{bmatrix} \mathbf{1} & 3 & 0 & -4 & 0 \\ 0 & 0 & \mathbf{1} & 1 & 0 \\ 0 & 0 & 0 & 0 & \mathbf{1} \end{bmatrix}$.

Przestrzenią zerową (przestrzenią rozwiązań jednorodnego układu) $\mathcal{N}(A) = \{\, x \in \mathbb{R}^5 \mid Ax = 0 \,\}$ jest

$$\mathcal{N}(A) = \{\, x = [-3u + 4v,\ u,\ -v,\ v,\ 0]^t \mid v,\ u \in \mathbb{R} \,\}.$$

Jako bazę przestrzeni zerowej $\mathcal{N}(A)$ można wziąć układ złożony z wektorów

$$n_1 = [-3,\ 1,\ 0,\ 0,\ 0]^t, \qquad n_2 = [4,\ 0,\ -1,\ 1,\ 0]^t.$$

Zauważmy

$$\dim \mathcal{N}(A) + \dim \mathcal{R}(A) = 2 + 3 = \dim \mathbb{R}^5.$$

Jawna postać bazy przestrzeni zerowej macierzy

$$u_1 = \begin{bmatrix} x_1 \\ \vdots \\ x_r \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \qquad u_2 = \begin{bmatrix} x_1' \\ \vdots \\ x_r' \\ 0 \\ 1 \\ 0 \\ \vdots \end{bmatrix}, \qquad \dots, \qquad u_k = \begin{bmatrix} x_1'' \\ \vdots \\ x_r'' \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}.$$

Przyjęto tutaj dla uproszczenia zapisu, że niewiadomymi wolnymi w układzie są niewiadome $x_{r+1}, \dots, x_n$.

Przykład (Przypadek jednego równania)

Dla $A = \begin{bmatrix} a_1 & a_2 & \dots & a_n \end{bmatrix} \in \mathbf{M}_{1\times n}(\mathbb{R})$, $a_1 \neq 0$, przestrzeń rozwiązań równania

$$Ax = a_1x_1 + a_2x_2 + \dots + a_nx_n = 0$$

jest rozpięta przez wektory $u_j \in \mathbb{R}^n$ postaci

$$u_1 = \begin{bmatrix} -\frac{a_2}{a_1} & 1 & 0 & \dots & 0 \end{bmatrix}^t, \quad u_2 = \begin{bmatrix} -\frac{a_3}{a_1} & 0 & 1 & 0\dots & 0 \end{bmatrix}^t, \quad \dots$$

Struktura podprzestrzeni liniowych w przestrzeni $\mathbb{R}^n$

Twierdzenie (O uzupełnianiu do bazy)

Jeśli $W \subset V$ jest właściwą ($W \neq V$) podprzestrzenią liniową przestrzeni V i układ $\{w_1, \ldots, w_k\}$ jest bazą przestrzeni W, to w przestrzeni V istnieje taki układ liniowo niezależny $\{v_1, \ldots, v_m\}$, że $\{w_1, \ldots, w_k, v_1, \ldots, v_m\}$ jest bazą przestrzeni V.

Twierdzenie (O konstruowaniu podprzestrzeni)

a) Każda podprzestrzeń liniowa wymiaru $k < n$ przestrzeni $\mathbb{R}^n$ jest przestrzenią zerową pewnej macierzy $A \in \mathbf{M}_{m\times n}(\mathbb{R})$ o rzędzie równym $n - k$.
b) Część wspólna $V \cap W$ podprzestrzeni liniowych $V, W \subset \mathbb{R}^n$ jest podprzestrzenią liniową o wymiarze

$$\dim V \cap W \leqslant \min(\dim V, \dim W).$$

c) Dla dowolnych podprzestrzeni liniowych $V, W \subset \mathbb{R}^n$ istnieje najmniejsza podprzestrzeń liniowa zawierająca obie te podprzestrzenie. Jest ona nazywana sumą podprzestrzeni V i W i oznaczana przez $V + W$, a jej wymiar spełnia

$$\dim(V + W) = \dim V + \dim W - \dim V \cap W.$$

Przykłady

Przykład (Rozszerzanie bazy podprzestrzeni do pełnej bazy)

Z poprzedniego przykładu wiemy, że przestrzeń zerowa

$\mathcal{N}(A) = \{ x \in \mathbb{R}^5 \mid Ax = 0 \}$ macierzy $A = \begin{bmatrix} \mathbf{1} & 3 & 0 & -4 & 0 \\ 0 & 0 & \mathbf{1} & 1 & 0 \\ 0 & 0 & 0 & 0 & \mathbf{1} \end{bmatrix}$

jest równa podprzestrzeni $W = \{ x = [-3u+4v,\ u,\ -v,\ v,\ 0]^t \mid v,\ u \in \mathbb{R} \} \subset \mathbb{R}^5$.

Jako bazę tej podprzestrzeni obierzemy układ

$$u_1 = [-3,\ 1,\ 0,\ 0,\ 0]^t, \qquad u_2 = [4,\ 0,\ -1,\ 1,\ 0]^t.$$

Jako układ uzupełniający bazę $\{u_1,\ u_2\}$ podprzestrzeni W możemy przyjąć układ złożony z wektorów

$$v_1 = [0,\ 0,\ 0,\ 0,\ 1]^t, \quad v_2 = [0,\ 0,\ 1,\ 1,\ 0]^t, \quad v_3 = [0,\ 1,\ 0,\ 0,\ 0]^t,$$

w oczywisty sposób będący liniowo niezależnym układem nie należącym do W. Wobec $\dim \operatorname{lin}\{v_1,\ v_2,\ v_3\} = 3$ układ $\{u_1,\ u_2,\ v_1,\ v_2,\ v_3\}$ jest bazą $\mathbb{R}^5$,

Wykład 5 (28.03.2019)

Zaczynamy pracę

Mierzenie długości i kątów w przestrzeni kartezjańskiej

Definicja (Norma, odległość i iloczyn skalarny w $\mathbb{R}^n$)

Dla wektorów z przestrzeni $\mathbb{R}^n$ używamy tutaj zapisu kolumnowego.

Iloczynem skalarnym w przestrzeni $\mathbb{R}^n$ nazywamy funkcję par wektorów

$$\mathbb{R}^n \times \mathbb{R}^n \ni (x,\, y) \mapsto x \cdot y = x^t y = \sum_{j=1}^{n} x_j y_j \in \mathbb{R}. \tag{17}$$

Normą euklidesową w przestrzeni $\mathbb{R}^n$ nazywa się funkcję wektora

$$\mathbb{R}^n \ni x \mapsto |x| = \sqrt{x_1^2 + \ldots + x_n^2} = \left(\sum_{j=1}^{n} x_j^2\right)^{\frac{1}{2}} \in \mathbb{R}. \tag{18}$$

Metryką (odległością) euklidesową w $\mathbb{R}^n$ nazywa się funkcję par wektorów

$$\mathbb{R}^n \times \mathbb{R}^n \ni (x,\, y) \mapsto d(x,\, y) = |x - y| = \left(\sum_{j=1}^{n} (x_j - y_j)^2\right)^{\frac{1}{2}} \in \mathbb{R}. \tag{19}$$

Iloczyn skalarny w $\mathbb{R}^n$

Stwierdzenie (Formalne własności iloczynu skalarnego w $\mathbb{R}^n$)

(a) *Odwzorowanie*

$$\mathbb{R}^n \times \mathbb{R}^n \ni (x, y) \mapsto x \cdot y = \sum_{j=1}^{n} x_j y_j \in \mathbb{R}, \tag{20}$$

spełnia dla każdych $x, y, v \in \mathbb{R}^n$ *i* $\alpha, \beta \in \mathbb{R}$ *warunki:*

(i)	$(\alpha x + \beta y) \cdot v = \alpha(x \cdot v) + \beta(y \cdot v),$	**(liniowość wzg. każdego**
	$v \cdot (\alpha x + \beta y) = \alpha(v \cdot x) + \beta(v \cdot y),$	**argumentu z osobna)**
(ii)	$x \cdot y = y \cdot x,$	**(symetria)**
(iii)	$x \cdot x \geqslant 0,$ *oraz* $x \cdot x = 0 \iff x = 0,$	**(dodatnia określoność)**.

(b) *Dla dowolnych wektorów* $x, y \in \mathbb{R}^n$ *jest spełniona nierówność Schwarza*

$$|x \cdot y| \leqslant |x||y|, \tag{21}$$

a w nierówności (21) *równość zachodzi wtedy i tylko wtedy, gdy* x *i* y *są współliniowe, tj.* $\lambda x + \mu y = 0$ *dla pewnych* $\lambda, \mu \in \mathbb{R}$.

Iloczyn skalarny ogólnie

Definicja (Iloczyn skalarny w przestrzeni liniowej nad ciałem $\mathbb{R}$)

Dla ogólnej przestrzeni liniowej V nad ciałem $\mathbb{R}$ iloczynem skalarnym nazywamy każdą funkcję pary wektorów

$$V \times V \ni (x, y) \mapsto \langle x, y \rangle \in \mathbb{R}, \tag{22}$$

spełniającą warunki

$$\begin{array}{lll}
\text{(IL1)} & \langle (\alpha x + \beta y), v \rangle = \alpha \langle x, v \rangle + \beta \langle y, v \rangle, & \textbf{(liniowość wzg. każdego} \\
 & \langle v, (\alpha x + \beta y) \rangle = \alpha \langle v, x \rangle + \beta \langle v, y \rangle, & \textbf{argumentu z osobna)} \\
\text{(IL2)} & \langle x, y \rangle = \langle y, x \rangle, & \textbf{(symetria)} \\
\text{(IL3)} & \langle x, x \rangle \geqslant 0, \ \text{oraz} \ \langle x, x \rangle = 0 \iff x = 0, & \textbf{(dodatnia określoność)}.
\end{array}$$

Jak poprzednio definiujemy normę wektora wzorem

$$\|x\| = \sqrt{\langle x, x \rangle}.$$

W tym kontekście także obowiązuje fundamentalna nierówność Schwarza

$$\langle x, y \rangle \leqslant \|x\| \, \|y\|, \tag{23}$$

wraz z analogicznym warunkiem dla zachodzenia we wzorze równości.

Ortogonalność

Definicja (Wektory ortogonalne)

Wektory $v, w \in V$ nazywamy *ortogonalnymi* (prostopadłymi), gdy $\langle v, w \rangle = 0$. Ortogonalność wektorów v, w zapisujemy symbolicznie $v \perp w$.

Zauważmy, że relacja ortogonalności wektorów jest symetryczna; $v \perp w \iff w \perp v$ oraz że wektor 0 (i tylko on) jest ortogonalny do każdego z wektorów przestrzeni V.

Przykład (Ortogonalność wektorów w $\mathbb{R}^2$)

Dla wektorów $v = (v_1, v_2) \neq 0$ i $w = (w_1, w_2) \neq 0$ przestrzeni $\mathbb{R}^2$ mamy $v \perp w \iff v \cdot w = v_1 w_1 + v_2 w_2 = 0$. A zatem

$$w \perp v \qquad \iff \qquad w = \lambda(-v_2, v_1), \quad \text{dla pewnego } \lambda \in \mathbb{R}.$$

Stwierdzenie

Zbiór wektorów ortogonalnych do różnego od zera wektora $v \in V$ jest podprzestrzenią liniową przestrzeni V nazywaną ortogonalnym dopełnieniem wektora v i oznaczaną $v^\perp$.

W przypadku $V = \mathbb{R}^n$ zachodzi $\qquad \dim v^\perp = n - 1.$

Iloczyn skalarny w przestrzeni funkcji ciągłych

Przykłady (Iloczyn skalarny w przestrzeniach funkcji)

a) Dla przestrzeni $V = \mathcal{C}([a, b])$ funkcji ciągłych na odcinku $[a, b] \subset \mathbb{R}$ zazwyczaj definiuje się iloczyn skalarny jako odwzorowanie

$$V \times V \ni (f, g) \mapsto \langle f, g \rangle = \frac{1}{b-a} \int_a^b f(t)g(t)\, dt \in \mathbb{R}.$$

Sprawdzenie warunków **IL1–IL2** jest banalne, natomiast prawdziwość warunku **IL3** wynika z głębszych własności całki, o których mówi wykład Analizy Matematycznej.

b) Szczególnie często wykorzystywany jest iloczyn skalarny w przestrzeni wielomianów rzeczywistych zdefiniowany wzorem

$$(f \mid g) = \int_{-1}^{1} f(t)g(t)dt. \tag{24}$$

W tej sytuacji każdy wielomian parzystego stopnia jest ortogonalny do wielomianu nieparzystego stopnia (jaka jest tego przyczyna?), co ułatwia pewne rozumowania.

Funkcje trygonometryczne

Dla przestrzeni wielomianów trygonometrycznych, lub przestrzeni $\mathcal{C}([0, 2\pi])$ funkcji ciągłych na odcinku $[0, 2\pi]$, rozważa się iloczyn skalarny

$$(f \mid g) = \int_0^{2\pi} f(t)g(t)dt. \tag{25}$$

Wówczas układ funkcji

$$\frac{1}{\sqrt{2\pi}},\ \frac{1}{\sqrt{\pi}}\sin t,\ \frac{1}{\sqrt{\pi}}\cos t,\ \ldots,\ \frac{1}{\sqrt{\pi}}\sin nt,\ \frac{1}{\sqrt{\pi}}\cos nt, \ldots \tag{26}$$

jest układem ortonormalnym, nazywanym *(unormowanym) układem trygonometrycznym*. Rzeczywiście, dla wszystkich k, $n \in \mathbb{N}$

$$\int_0^{2\pi} \sin(kt)\cos(nt)\,dt = \int_0^{2\pi} \sin(nt)dt = \int_0^{2\pi} \cos(nt)\,dt = 0$$

$$\int_0^{2\pi} \sin(kt)\sin(nt)\,dt = \int_0^{2\pi} \cos(kt)\cos(nt)\,dt = \begin{cases} 0, & k \neq n, \\ \pi, & k = n \neq 0, \end{cases}$$

$$\int_0^{2\pi} dt = 2\pi,$$

Podstawowe zależności

- Długość wektora $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$

$$|x| = \sqrt{x \cdot x} = \left(\sum_{j=1}^{n} x_j^2\right)^{1/2}.$$

Obowiązują także w dowolnej przestrzeni z iloczynem skalarnym

- Długość sumy (różnicy) wektorów ze wzoru

$$\begin{aligned} |x \pm y|^2 &= (x \pm y) \cdot (x \pm y) = x \cdot (x \pm y) \pm y \cdot (x \pm y) = x \cdot x + y \cdot y \pm 2x \cdot y \\ &= |x|^2 + |y|^2 \pm x \cdot y. \end{aligned}$$

- Jeśli $x \perp y$, to „PITAGORAS”

$$|x + y|^2 = |x|^2 + |y|^2. \tag{27}$$

- Prawo równoległoboku

$$|x + y|^2 + |x - y|^2 = 2(|x|^2 + |y|^2).$$

- nierówność trójkąta

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

Kąt między wektorami w $\mathbb{R}^n$

Definicja (Kąt między wektorami w $\mathbb{R}^n$)

Jeśli żaden z wektorów v, $w \in \mathbb{R}^n$ nie jest zerem, to nierówność Schwarza

$$|v \cdot w| \leqslant |v||w| \quad \text{można przepisać jako} \quad -1 \leqslant \frac{v \cdot w}{|v||w|} \leqslant 1.$$

Jednoznacznie wyznaczoną liczbę θ z odcinka $[0, \pi]$, taką że

$$\cos\theta = \frac{v \cdot w}{|v|\,|w|}, \quad \text{równoważnie}, \quad v \cdot w = \cos\theta\,|v|\,|w| \tag{28}$$

nazywamy kątem między wektorami v, w.
Zgodnie z tą definicją kąt między niezerowymi ortogonalnymi wektorami jest równy $\pi/2$, co wyjaśnia użycie tej terminologii.

Przykłady (Wektory ortogonalne)

a) Wektor 0 jest ortogonalny do każdego wektora przestrzeni.
b) Wektory standardowej bazy jednostkowej $\{e_j\}$ w $\mathbb{R}^n$ są parami ortogonalne;

$$e_j \cdot e_k = 0, \qquad j \neq k.$$

Kąt między wektorami na płaszczyźnie $\mathbb{R}^2$

Przykład (Współrzędne biegunowe na płaszczyźnie $\mathbb{R}^2$)

W geometrii płaszczyzny tradycyjnie nazywa się wektory $e_1 = (1, 0)$ i $e_2 = (0, 1)$ wersorami osi. Niezerowy wektor $v = (x, y) \in \mathbb{R}^2$ tworzy z wersorami osi kąty φ_1 i φ_2 dane wzorami

$$\cos\varphi_1 = \frac{x}{\sqrt{x^2+y^2}}; \qquad \cos\varphi_2 = \frac{y}{\sqrt{x^2+y^2}}.$$

Ponieważ

$$\left(\frac{x}{\sqrt{x^2+y^2}}\right)^2 + \left(\frac{y}{\sqrt{x^2+y^2}}\right)^2 = 1,$$

to zauważamy, że kąty φ_1 i φ_2 są kątami dopełniającymi, $\varphi_1 + \varphi_2 = \pi/2$. Oznaczając $\varphi = \varphi_1$ otrzymamy

$$v = (x, y) = |v|(\cos\varphi, \sin\varphi)$$

Jest to rzeczywista forma zapisu wzoru na przedstawienie trygonometryczne liczby zespolonej $z = x + iy$.

Kąt między wektorami na płaszczyźnie $\mathbb{R}^2$, cd.

Przykład (Jeden centralny wzór z trygonometrii)

a) Niech $x = (x_1, x_2)$ i $y = (y_1, y_2)$ tworzą z osią Ox_1 odpowiednio kąty φ_1 i φ_2, tak że

$$x = (x_1, x_2) = |x|(\cos\varphi_1, \sin\varphi_1);$$
$$y = (y_1, y_2) = |y|(\cos\varphi_2, \sin\varphi_2).$$

gdzie $|x| = \sqrt{x_1^2 + x_2^2}$, $|y| = \sqrt{y_1^2 + y_2^2}$.
Wówczas

$$x \cdot y = |x||y|\left(\cos\varphi_1\cos\varphi_2 + \sin\varphi_1\sin\varphi_2\right)$$

i po wykorzystaniu wzoru

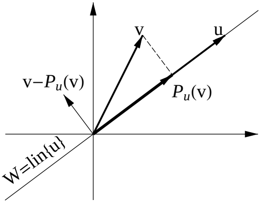
$$\cos\alpha\cos\beta + \sin\alpha\sin\beta = \cos(\alpha - \beta)$$

otrzymamy

$$x \cdot y = |x||y|\cos(\varphi_1 - \varphi_2).$$

Rzut prostopadły na kierunek wektora

Dla $u \neq 0$ wyznaczamy podprzestrzeń $W = \mathbb{R}\, u = \mathrm{lin}\{u\} \subset \mathbb{R}^n$ (kierunek wektora u) i jej dopełnienie ortogonalne $W^\perp = \{\, x \in \mathbb{R}^n \mid x \perp u \,\} = \{u\}^\perp$. Oznaczamy $u_n = |u|^{-1} u$ unormowany wektor w kierunku u.



Rzut ortogonalny v na kierunek wektora u

Twierdzenie (Formuła dla rzutu ortogonalnego)

Dla każdego wektora $v \in \mathbb{R}^n$ wektor v_0 zdefiniowany równoznacznymi wzorami

$$v_0 = \frac{v \cdot u}{|u|^2} u = (v \cdot u_n) u_n = \cos\theta \, |v| \, u_n$$

jest jedynym wektorem $v_0 \in \mathrm{lin}\{u\}$, takim że $v_\perp = v - v_0 \in \mathrm{lin}\{u\}^\perp$.

Zatem $\qquad v = v_0 + v_\perp, \qquad v_0 \in \mathrm{lin}\{u\},\ v_\perp \in \mathrm{lin}\{u\}^\perp.$

Własność ekstremalności rzutu ortogonalnego

Definicja (Rzutowanie ortogonalne na kierunek wektora)

Wektor $v_0 = (v \cdot u_n)u_n$ będziemy nazywać rzutem prostopadłym lub ortogonalnym wektora v na kierunek wektora u, a odwzorowanie

$$\mathbb{R}^n \ni v \mapsto P_u(v) = (v \cdot u_n)u_n \in \mathbb{R}^n \tag{29}$$

nazwiemy rzutowaniem ortogonalnym lub rzutowaniem prostopadłym przestrzeni $\mathbb{R}^n$ na kierunek wektora u.

Twierdzenie (Rzut ortogonalny minimalizuje odległość)

Dla dowolnie ustalonego wektora $v \notin \mathrm{lin}\{u\}$ funkcja liczbowa

$$\mathrm{lin}\{u\} \ni w \mapsto d(v, w) = |v - w| \in \mathbb{R}$$

przyjmuje minimum w punkcie $v_0 = P_u(v)$,

$$\min\{d(v, w) \mid w \in \mathrm{lin}\{u\}\} = |v - P_u(v)|.$$

Przykłady

Dane są $W = \mathrm{lin}\{u\} \subset \mathbb{R}^n$ i niezerowy wektor $v \notin W$. Przypomnijmy: ortogonalne dopełnienie wektora u jest oznaczane $u^\perp$ i zdefiniowane jako $u^\perp = \{\, x \in \mathbb{R}^n \mid u \cdot x = 0 \,\}$.

Przykłady (Rzuty prostopadłe)

a) Obierzmy $u = e_n = (0, \ldots, 0, 1)$, $v = (v_1, \ldots, v_{n-1}, v_n)$. Wówczas

$$P_u(v) = (v \cdot e_n)e_n = (0, \ldots, 0, v_n), \qquad v_\perp = (v_1, \ldots, v_{n-1}, 0).$$

b) Niech $n = 3$ i $u = (1, 1, 1) \in \mathbb{R}^3$. Zauważmy, że $|u|^2 = 3$ oraz $u \cdot x = x_1 + x_2 + x_3$, więc

$$\frac{u \cdot x}{|u|^2} = \bar{x} \qquad \text{średnia wartość } x_1,\, x_2,\, x_3.$$

Zatem rzut wektora $v = (v_1, v_2, v_3)$ na kierunek u wyraża się wzorem (29)

$$P_u(v) = \frac{v_1 + v_2 + v_3}{3}(1, 1, 1) = \bar{v}(1, 1, 1),$$

a prostopadła do u składowa $v_\perp$ wyraża się przez

$$v_\perp = (v_1 - \bar{v},\, v_2 - \bar{v},\, v_3 - \bar{v}).$$

Układy i bazy ortonormalne

Definicja (Układy ortogonalne i ortonormalne)

Układ wektorów $\{v_1, v_2, \ldots, v_k\} \subset \mathbb{R}^n$ będziemy nazywać:

- układem ortogonalnym (w skrócie u. o.), jeśli dowolne dwa wektory tego układu (o różnych wskaźnikach) są ortogonalne, tj. dla każdej pary i, j zachodzi $\quad i \neq j \quad \Longrightarrow \quad v_i \cdot v_j = 0$;
- jeśli ponadto $v_i \cdot v_i = 1$ dla każdego i, tj. norma każdego wektora układu jest równa 1, to powiemy, że jest on układem ortonormalnym (w skrócie u. o. n.);
- Bazę $\{v_1, v_2, \ldots, v_n\} \subset \mathbb{R}^n$ przestrzeni $\mathbb{R}^n$, która jest układem ortonormalnym będziemy nazywać bazą ortonormalną (w skrócie b. o. n.) przestrzeni $\mathbb{R}^n$.

Stwierdzenie (O własnościach układów ortogonalnych)

Układ ortonormalny $\{v_1, v_2, \ldots, v_k\}$ jest układem liniowo niezależnym.
Układ ortonormalny $\{v_1, v_2, \ldots, v_k\}$ w przestrzeni V jest bazą przestrzeni V wtedy i tylko wtedy, gdy $\operatorname{lin}\{v_1, v_2, \ldots, v_k\} = V$.
W szczególności: układ ortonormalny $\{v_1, v_2, \ldots, v_k\}$ w przestrzeni $\mathbb{R}^n$ jest bazą tej przestrzeni wtedy i tylko wtedy, gdy $k = n$.

Wektory unormowane

Odwzorowanie $v \mapsto \frac{1}{\|v\|} v$ określone dla wektorów $v \in V \setminus \{0\}$ nazywamy unormowaniem (normalizacją) wektora v. Ogólnie, wektorami *unormowanymi* będziemy nazywać wektory o długości 1.

Wektory unormowane w przestrzeni $\mathbb{R}^n$ tworzą zbiór

$$S^{n-1} = \{\, x \in \mathbb{R}^n \mid |x| = \sqrt{\textstyle\sum_{j=1}^n x_j^2} = 1 \,\},$$

który jest odpowiednikiem „sfery jednostkowej" w przestrzeni $\mathbb{R}^3$.

Dla $n = 3$ wektory unormowane możemy przedstawić w formie

$$(\sin\theta\cos\phi,\ \sin\theta\sin\phi,\ \cos\theta), \quad 0 \leqslant \phi \leqslant 2\pi,\ 0 \leqslant \theta \leqslant \pi.$$

Jeśli $\{v_1, \ldots, v_k\}$ jest u. o. w przestrzeni $\mathbb{R}^n$ i nie zawiera 0, to wektory unormowane

$$u_1 = \frac{1}{|v_1|} v_1, \quad u_2 = \frac{1}{|v_2|} v_2, \quad \ldots, \quad u_k = \frac{1}{|v_1|} v_k$$

tworzą układ ortonormalny, rozpinający tę samą przestrzeń co $\{v_1, \ldots, v_k\}$;

$$\mathrm{lin}\{v_1, \ldots, v_k\} = \mathrm{lin}\{u_1, \ldots, u_k\}.$$

Przykłady układów ortogonalnych i ortonormalnych

Przykład

a) Jeśli $0 \neq w = (w_1,\, w_2)$ i przyjąć $w' = (-w_2,\, w_1)$, to układ $\{w,\, w'\}$ jest ortogonalny

$$w \cdot w' = (w_1,\, w_2) \cdot (-w_2,\, w_1) = -w_1 w_2 + w_2 w_1 = 0.$$

Na przykład, wektory $(\cos\phi,\, \sin\phi)$ i $(-\sin\phi,\, \cos\phi)$ są ortogonalne i unormowane, co w zapisie kolumnowym wyraża się przez równości

$$\begin{bmatrix}\cos\phi\\ \sin\phi\end{bmatrix}^t \begin{bmatrix}-\sin\phi\\ \cos\phi\end{bmatrix} = 0, \quad \begin{bmatrix}\cos\phi\\ \sin\phi\end{bmatrix}^t \begin{bmatrix}\cos\phi\\ \sin\phi\end{bmatrix} = 1, \quad \begin{bmatrix}-\sin\phi\\ \cos\phi\end{bmatrix}^t \begin{bmatrix}-\sin\phi\\ \cos\phi\end{bmatrix} = 1.$$

Łącznie te trzy równości są równoważne równości macierzowej

$$\begin{bmatrix}\cos\phi & \sin\phi\\ -\sin\phi & \cos\phi\end{bmatrix} \begin{bmatrix}\cos\phi & -\sin\phi\\ \sin\phi & \cos\phi\end{bmatrix} = \begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}.$$

Wykład 7 (11.04.2019)

Najważniejsze zagadnienia

- Współrzędne wektora w bazie o.n.
- Uogólniony wzór Pitagorasa
- Rozkład wektora w bazie o.n.
- Konstrukcja rzutu ortogonalnego na podprzestrzeń
- Ekstremalne własności rzutu i jego zastosowania
- Przybliżone rozwiązania układu równań liniowych
- Klasyczna metoda najmniejszych kwadratów

Szczegóły w podręczniku, Rozdz. 7.

Współrzędne wektora w bazie ortogonalnej

$\{v_1, \ldots, v_k \in \mathbb{R}^n\}$ układ ortogonalny nie zawierający wektora zerowego,
$V = \mathrm{lin}\{v_1, \ldots, v_k\}$ podprzestrzeń rozpięta na tych wektorach.

Stwierdzenie (Rozkład wektora w bazie ortogonalnej)

a) $\{v_1, \ldots, v_k\}$ *jest bazą ortogonalną przestrzeni* V*,*
zaś $\{u_1, \ldots, u_k\}$*, gdzie* $u_i = \frac{1}{|v_i|} v_i$*, jest jej bazą ortonormalną.*

b) Każdy $x \in V = \mathrm{lin}\{v_1, \ldots, v_k\}$ *jest sumą wzajemnie ortogonalnych składowych*

$$x = \sum_{j=1}^{k} \frac{(x \cdot v_j)}{|v_j|^2} v_j = \sum_{j=1}^{k} (x \cdot u_j) u_j. \tag{30}$$

c) Dla dowolnych $x, y \in V = \mathrm{lin}\{v_1, \ldots, v_k\}$ *mamy*

$$x \cdot y = \sum_{j=1}^{k} (x \cdot u_j)(u_j \cdot y), \tag{31}$$

oraz „uogólniony wzór Pitagorasa”

$$|x|^2 = \sum_{j=1}^{k} |x \cdot u_j|^2 \tag{32}$$

Rzut ortogonalny na podprzestrzeń

Współczynniki kombinacji liniowej po prawej stronie wzoru (30) nazywamy współrzędnymi wektora x względem układu ortonormalnego $\{u_1, u_2, \ldots, u_k\}$, a sam wzór (30) będziemy nazywali rozwinięciem wektora x względem układu ortonormalnego $\{u_j\}$.

Rozdzielając składniki sumy (30) na dwie części, $\sum_{j=1}^{m}(x \cdot u_j)u_j$ dla $m < k$ i pozostałe, otrzymujemy wektory o szczególnych własnościach. Oznaczymy

$$P_m(x) = \sum_{j=1}^{m}(x \cdot u_j)u_j, \qquad Q_m(x) = \sum_{j=m+1}^{k}(x \cdot u_j)u_j,$$

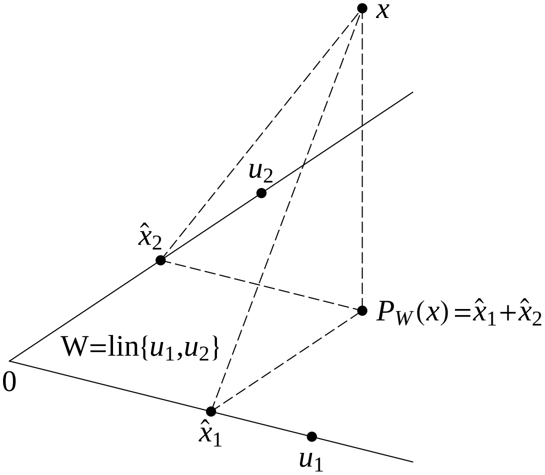
co pozwoli zapisać

$$x = P_m(x) + Q_m(x), \tag{33}$$

$$\text{gdzie} \qquad P_m(x) \in \operatorname{lin}\{u_1, \ldots, u_m\}, \qquad Q_m(x) \in \operatorname{lin}\{u_1, \ldots, u_m\}^{\perp}. \tag{34}$$

Składnik $P_m(x)$ w sumie (35) nazywamy rzutem ortogonalnym wektora x na podprzestrzeń $\operatorname{lin}\{u_1, \ldots, u_m\}$.

Rzut ortogonalny w przestrzeni $\mathbb{R}^3$



$$P_W(x) = \hat{x}_1 + \hat{x}_2 = (x \cdot u_1)u_1 + (x \cdot u_2)u_2$$

Rozkład ortogonalny przestrzeni $\mathbb{R}^n$

Twierdzenie (O rozkładzie ortogonalnym przestrzeni)

Niech $W \subset \mathbb{R}^n$ będzie podprzestrzenią liniową wymiaru k. Wówczas zbiór

$$W^\perp = \{\, x \in \mathbb{R}^n \mid x \perp w \text{ dla każdego } w \in W \,\},$$

nazywany ortogonalnym dopełnieniem przestrzeni W, jest podprzestrzenią liniową wymiaru $n-k$. Każdy wektor $x \in \mathbb{R}^n$ można zapisać jednoznacznie w postaci sumy

$$x = w + w^\circ, \qquad \text{gdzie} \quad w \in W,\ w^\circ \in W^\perp, \tag{35}$$

przy czym występujący w tym rozkładzie wektor $w \in W$ jest liniową funkcją wektora x. Inaczej mówiąc, $x \mapsto w \in W$ jest odwzorowaniem liniowym, oznaczanym symbolicznie $x \mapsto P_W(x)$ i nazywanym rzutem ortogonalnym na przestrzeń W.

Ekstremalność rzutu ortogonalnego w przestrzeni $\mathbb{R}^n$

Twierdzenie (Rzut ortogonalny minimalizuje odległość)

Niech $\{u_1, u_2, \ldots, u_n\}$ będzie bazą o.n w przestrzeni $\mathbb{R}^n$;
$W = \mathrm{lin}\{u_1, u_2, \ldots, u_m\}$, $m < n$, podprzestrzenią rozpiętą na pierwszych m wektorach tej bazy.
a) $P_W : \mathbb{R}^n \to \mathbb{R}^n$ odwzorowanie rzutu ortogonalnego na W jest dane wzorem

$$P_W(x) = \sum_{j=1}^{m} (x \cdot u_j) u_j, \qquad x \in \mathbb{R}^n.$$

b) Dla dowolnie ustalonego wektora $v \notin W$ funkcja liczbowa

$$W \ni w \mapsto d(v, w) = |v - w| \in \mathbb{R}$$

przyjmuje minimum w punkcie $v_0 = P_W(v)$,

$$\min\{d(v, w) \mid w \in W\} = |v - P_W(v)|.$$

c) Wektor $v_0 = P_W(v)$ jest wyznaczony warunkami: $v_0 \in W$ i $v - v_0 \in W^\perp$.

Bazy ortogonalne i macierze, I

Wektory z przestrzeni $\mathbb{R}^n$ zapisujemy w notacji kolumnowej. Iloczyn skalarny wyraża się wzorem — wskaźnik górny ^{t} oznacza transpozycję macierzy;

$$x \cdot y = x^t\, y = \begin{bmatrix} x_1 & x_2 & \dots & x_n \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \sum_{j=1}^{n} x_j\, y_j .$$

Układ wektorów $\{u_1, \dots, u_m\}$ z przestrzeni $\mathbb{R}^n$ można przedstawić jako macierz $[u_1, \dots, u_m]$ o wymiarach $n \times m$, dla $m = n$ jako macierz kwadratową stopnia n.

Stwierdzenie (Macierzowe kryterium ortonormalności)

Układ $\{u_1, \dots, u_m\}$ wektorów z przestrzeni $\mathbb{R}^n$ jest układem ortonormalnym wtedy i tylko wtedy, gdy macierz U, której kolumnami są wektory u_j spełnia

$$U^t\, U = \mathbf{I}_m, \qquad \mathbf{I}_m \text{— } \textit{macierz jednostkowa stopnia } m.$$

Bazy ortogonalne i macierze, II

Stwierdzenie (Bazy ortonormalne i macierze)

Układ $\{u_1, \ldots, u_n\}$ wektorów z przestrzeni $\mathbb{R}^n$ jest bazą ortonormalną wtedy i tylko wtedy, gdy macierz U, której kolumnami są wektory u_j spełnia

$$U^t\, U = \mathbf{I}_n, \qquad \mathbf{I}_n \text{— macierz jednostkowa stopnia } n.$$

Definicja (Macierz ortogonalna)

Macierz kwadratową $U \in \mathbf{M}_n(\mathbb{R})$, która spełnia równość

$$U^t\, U = \mathbf{I}_n$$

nazywa się macierzą ortogonalną.

Wniosek

Macierz ortogonalna U jest macierzą odwracalną, zatem

$$U^t\, U = U\, U^t = \mathbf{I}_n; \qquad U^t = U^{-1}; \qquad \det U = \pm 1,$$

Obliczenia we współrzędnych

Niech U będzie macierzą ortogonalną stopnia n, $\{u_1, \dots, u_n\}$ bazą ortonormalną przestrzeni $\mathbb{R}^n$ utworzoną z kolumn macierzy U. Współrzędne $(y_1, y_2, \dots, y_n)$ wektora $x \in \mathbb{R}^n$ w bazie o. n. $\{u_1, \dots, u_n\}$ wyznaczone przez $x = \sum_{j=1}^{n} y_j u_j$ są dane wzorami (w zapisie kolumnowym)

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} u_1 \cdot x \\ u_2 \cdot x \\ \vdots \\ u_n \cdot x \end{bmatrix} = \begin{bmatrix} u_{11} & u_{12} & \dots & u_{1n} \\ u_{11} & u_{12} & \dots & u_{1n} \\ \vdots & \vdots & \dots & \vdots \\ u_{n1} & u_{n2} & \dots & u_{nn} \end{bmatrix}^t \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}. \tag{36}$$

Stwierdzenie

Odwzorowanie $\mathbb{R}^n \ni x \mapsto y = U^t\, x \in \mathbb{R}^n$ zachowuje długość wektorów i kąty miedzy nimi:

$$|y| = |U^t\, x| = |x|, \qquad v \cdot w = U^t\, v \cdot U^t\, w, \quad v,\, w \in \mathbb{R}^n.$$

Odwzorowania zachowujące długości wektorów nazywamy izometriami.

Metoda ortogonalizacji Grama–Schmidta

Metoda G–S prowadzi od układu liniowo niezależnego $v_1, \dots, v_k$ do układu ortonormalnego $u_1, \dots, u_k$ z zachowaniem kolejnych podprzestrzeni rozpinanych przez te układy:
$\mathrm{lin}\{v_1\} = \mathrm{lin}\{u_1\}$,
$\mathrm{lin}\{v_1, v_2\} = \mathrm{lin}\{u_1, u_2\}$,
$\mathrm{lin}\{v_1, v_2, v_3\} = \mathrm{lin}\{u_1, u_2, u_3\}$,
itd.
Mamy (współczynniki α_j zapewniają unormowanie wektorów w_j)

$$
\begin{array}{ll}
\alpha_1 u_1 = v_1, & \alpha_1 = |v_1|; \\
\alpha_2 u_2 = v_2 - (v_2 \cdot u_1)u_1, & \alpha_2 = |v_2 - (v_2 \cdot u_1)u_1|; \\
\alpha_3 u_3 = v_3 - (v_3 \cdot u_1)u_1 - (v_3 \cdot u_2)u_2, & \alpha_3 = |v_3 - (v_3 \cdot u_1)u_1 - (v_3 \cdot u_2)u_2|; \\
\alpha_4 u_4 = v_4 - (v_4 \cdot u_1)u_1 - (v_4 \cdot u_2)u_2 - (v_4 \cdot u_3)u_3; & \dots \\
\dots &
\end{array}
$$

Warto zauważyć, że po prawych stronach znaku równości występują rzuty kolejnych wektorów v_j na podprzestrzenie rozpięte przez $u_1, u_2, \dots, u_{j-1}$.

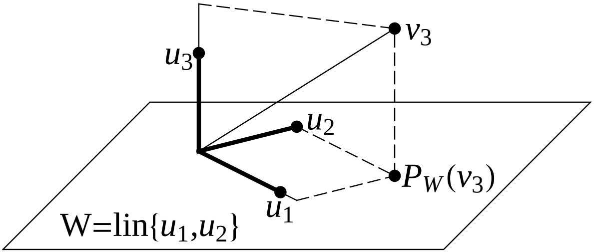
Ilustracja konstrukcji Grama-Schmidta

Przykład (Konstrukcja G.–S. dla układu $\{v_1, v_2, v_3\}$.)

Mamy

$$u_1 = |v_1|^{-1} v_1;$$

$$w_2 = v_2 - (v_2 \cdot u_1) u_1 = v_2 - \frac{(v_2 \cdot v_1)}{|v_1|^2} v_1, \qquad u_2 = |w_2|^{-1} w_2.$$



Trzeci krok konstrukcji układu o. n. $\{u_1, u_2, u_3\}$

Przykład konstrukcji Grama–Schmidta I

Bazę przestrzeni $\mathbb{R}^3$ złożoną z wektorów $v_1 = (1,\, 1,\, 1)$, $v_2 = (1,\, 1,\, 0)$, $v_3 = (1,\, 0,\, 0)$ przeprowadzimy na bazę ortonormalną metodą Grama–Schmidta. W pierwszym kroku przyjmiemy $u_1 = v_1 = (1,\, 1,\, 1)$, a dla wyznaczenia wektora należącego do przestrzeni $\mathrm{lin}\{v_1,\, v_2\}$ i ortogonalnego do u_1 odejmiemy od wektora v_2 jego rzut na kierunek $u_1\,(= v_1)$,

$$u_2 = v_2 - \lambda u_1, \quad \text{gdzie} \quad \lambda = \frac{u_1 \cdot v_2}{|u_1|^2}. \quad \text{Po obliczeniu} \quad \lambda = \frac{2}{3};$$

co daje (bez normowania wektora)

$$u_2 = (1,\, 1,\, 0) - \frac{2}{3}(1,\, 1,\, 1) = \frac{1}{3}(1,\, 1,\, -2)$$

W następnym kroku przyjmujemy z podobnych względów

$$u_3 = v_3 - \alpha_1 u_1 - \alpha_2 u_2, \quad \text{gdzie } \alpha_1 = \frac{v_3 \cdot u_1}{|u_1|^2} = \frac{1}{3}, \quad \alpha_2 = \frac{v_3 \cdot u_2}{|u_2|^2} = \frac{1}{2};$$

$$u_3 = (1,\, 0,\, 0) - \frac{1}{3}(1,\, 1,\, 1) - \frac{1}{6}(1,\, 1,\, -2) = \frac{1}{2}(1,\, -1,\, 0).$$

Przykład konstrukcji Grama–Schmidta II

Otrzymany układ wektorów $\{u_1, u_2, u_3\}$ jest ortogonalny, ale nie unormowany. Obliczymy długości otrzymanych wektorów

$$|u_1| = \sqrt{3}, \quad |u_2| = \sqrt{\frac{2}{3}}, \quad |u_3| = \frac{1}{\sqrt{2}},$$

i przez normalizację

$$w_1 = |u_1|^{-1} u_1 = \frac{1}{\sqrt{3}}(1, 1, 1), \quad w_2 = |u_2|^{-1} u_2 = \frac{1}{\sqrt{6}}(1, 1, -2),$$
$$w_3 = |u_3|^{-1} u_3 = \frac{1}{\sqrt{2}}(1, -1, 0),$$

otrzymamy układ $\{w_1, w_2, w_3\}$ będący poszukiwaną bazą ortonormalną.

Przykład konstrukcji Grama–Schmidta III

Relacje między wektorami wyjściowej bazy $\{v_1, v_2, v_3\}$ i otrzymanej bazy ortonormalnej $\{w_1, w_2, w_3\}$ można także wyrazić równościami

$$\begin{aligned}
v_1 &= |v_1| w_1 = \sqrt{3} w_1, \\
v_2 &= (w_1 \cdot v_2) w_1 + |u_2| w_2 = \frac{2}{\sqrt{3}} w_1 + \sqrt{\frac{2}{3}} w_2, \\
v_3 &= (w_1 \cdot v_3) w_1 + (w_2 \cdot v_3) w_2 + |u_3| w_3 = \frac{1}{\sqrt{3}} w_1 + \frac{1}{\sqrt{6}} w_2 + \frac{1}{\sqrt{2}} w_3.
\end{aligned}$$

Przepiszmy trzy ostatnie równości stosując kolumnowy zapis wektorów

$$v_1 = \sqrt{3} \frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}; \qquad v_2 = \sqrt{\frac{2}{3}} \frac{1}{\sqrt{6}} \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix} + \frac{2}{\sqrt{3}} \frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix};$$

$$v_3 = \frac{1}{\sqrt{3}} \frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + \frac{1}{\sqrt{6}} \frac{1}{\sqrt{6}} \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix} + \frac{1}{\sqrt{2}} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}.$$

Przykład konstrukcji Grama–Schmidta IV

Oznaczając przez $W = \begin{bmatrix} w_1 & w_2 & w_3 \end{bmatrix}$ macierz kwadratową stopnia 3, której kolumnami są (w tej kolejności) w_1, w_2, w_3, można wyrazić poprzednie związki jako rozkład macierzy V w postaci iloczynu dwóch szczególnych macierzy: ortogonalnej macierzy W i górno-trójkątnej macierzy T

$$V = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & -\frac{2}{\sqrt{6}} & 0 \end{bmatrix} \begin{bmatrix} \sqrt{3} & \frac{2}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ 0 & \sqrt{\frac{2}{3}} & \frac{1}{\sqrt{6}} \\ 0 & 0 & \frac{1}{\sqrt{2}} \end{bmatrix} = W\,T.$$

Rozkłady tego typu mają wiele zastosowań — o niektórych powiemy później.

Macierz Grama układu wektorów

Rozważamy układ wektorów $v_1, \ldots, v_n$ należących do $\mathbb{R}^m$ i macierz $V \in \mathbf{M}_{m\times n}(\mathbb{R})$,

$$V = \begin{bmatrix} v_{11} & v_{12} & \ldots & v_{1n} \\ \vdots & \vdots & \ldots & \vdots \\ v_{m1} & v_{m2} & \ldots & v_{mn} \end{bmatrix} = \begin{bmatrix} v_1 & v_2 & \ldots & v_n \end{bmatrix}$$

traktowaną jako ciąg kolumn. Macierzą Grama układu $\{v_1, \ldots, v_n\}$ jest macierz

$$V^t V = \begin{bmatrix} v_1 \cdot v_1 & v_1 \cdot v_2 & \ldots & v_1 \cdot v_n \\ \ldots & & & \\ v_n \cdot v_1 & v_n \cdot v_2 & \ldots & v_n \cdot v_n \end{bmatrix}$$

Wniosek

Układ wektorów $v_1, \ldots, v_n$ przestrzeni $\mathbb{R}^m$ jest układem ortonormalnym wtedy i tylko wtedy, gdy macierz Grama tego układu jest macierzą jednostkową

$$V^t V = \mathbf{I}_n. \tag{37}$$

Wykład 8 (9.05.2019)

Najważniejsze zagadnienia

- Powtórzenie — konstrukcja rzutu ortogonalnego na podprzestrzeń
- Powtórzenie — ekstremalne własności rzutu i jego zastosowania
- Przybliżone rozwiązania układu równań liniowych
- Klasyczna metoda najmniejszych kwadratów

Szczegóły w podręczniku, Rozdz. 7.

Przybliżone rozwiązania nadokreślonych układów równań liniowych

Układ równań liniowych $\quad Ax = b, \quad A \in \mathbf{M}_{m\times n}(\mathbb{R}),\ b \in \mathbb{R}^m, \quad$ gdzie $\ m > n$, nazywa się układem nadokreślonym. Taki układ ma więcej równań niż niewiadomych, czyli przestrzeń $K(A)$ kolumn macierzy A jest właściwą podprzestrzenią $\mathbb{R}^m$(!) Twierdzenie K.–C. orzeka, że układ jest niesprzeczny wtedy i tylko wtedy, gdy $b \in K(A)$ — to znaczące ograniczenie wektora b.

$$\text{Funkcja} \quad \mathbb{R}^n \ni x \mapsto \delta(x) = |Ax - b| = \sqrt{\sum_{j=1}^{m}(b_j - (Ax)_j)^2} \in \mathbb{R}$$

przyjmuje wartość 0 dla rozwiązania x układu i jest miarą oddalenia od rozwiązania dla innych elementów.

Definicja (Rozwiązanie przybliżone układu równań liniowych)

Rozwiązaniem przybliżonym układu $Ax = b$ będziemy nazywać taki $x_0 \in \mathbb{R}^n$, że $\delta(x_0) = \min\{\delta(x) \mid x \in \mathbb{R}^n\}$.

Wyznaczenie przybliżonego rozwiązania opiera się na metodzie *rzutu ortogonalnego*, nazywanej w tym kontekście metodą najmniejszych kwadratów.

Równania normalne

Odległość $\delta(x) = |Ax - b|$ oblicza się ze wzoru

$$\delta(x)^2 = |Ax - b|^2 = (Ax - b) \cdot (Ax - b) = |Ax|^2 - 2(Ax) \cdot b + |b|^2$$

Minimum tej funkcji można wyznaczyć *analitycznie*, wyznaczając punkt krytyczny, w którym jej gradient znika, co daje

$$\operatorname{grad} \delta(x)^2 = 2A^t Ax - 2A^t b = 0, \qquad \Longleftrightarrow \qquad A^t Ax = A^t b$$

To ostatnie równanie (układ równań liniowych) nazywa się *układem równań normalnych*. W przypadku gdy $A^t A$ jest odwracalna (np. jeśli kolumny macierzy A są liniowo niezależne, tj. rząd A jest maksymalny), mamy bardzo znaną formułę

$$x_{NK} = (A^t A)^{-1} A^t b. \tag{38}$$

dla przybliżonego rozwiązania (metodą Najmniejszych Kwadratów).

Ten sam wzór wynika z zastosowania charakteryzacji punktu minimum funkcji $\delta(x)^2$ jako rzutu ortogonalnego b na przestrzeń kolumn macierzy A.

Spojrzenie algebraiczne

Minimum funkcji $\delta(x) = |Ax - b|$ jest osiągane w takim punkcie $\widehat{x}$, że

$$(A\widehat{x} - b) \cdot Az = 0, \quad \text{dla każdego} \quad z \in \mathbb{R}^n$$

Ze względu na równość $x \cdot Az = A^t x \cdot z$ mamy

$$(A\widehat{x} - b) \cdot Az = (A^t A\widehat{x} - A^t b) \cdot z = 0, \quad \text{dla każdego} \quad z \in \mathbb{R}^n$$

co jest spełnione wtedy i tylko wtedy, gdy $A^t A\widehat{x} - A^t b = 0$, co daje znowu równania normalne $\widehat{x} = x_{NK} = (A^t A)^{-1} A^t b$.
Własności przybliżonego rozwiązania x_{NK}:

- $x_{NK} = A^{-1}b$, jeżeli A jest kwadratowa i odwracalna;
- $Ax_{NK} = A(A^t A)^{-1} A^t b = b$, jeżeli $b \in K(A)$;
- Macierz $A^\dagger = (A^t A)^{-1} A^t$ spełnia relację $A^\dagger A = \mathbf{I}$;
- Odwzorowanie $\mathbb{R}^m \ni y \mapsto A(A^t A)^{-1} A^t y = Ax_{NK}$ jest rzutem ortogonalnym na $K(A)$.

Dopasowanie prostej metodą najmniejszych kwadratów I

W naukach empirycznych często ma się do czynienia z następującym problemem — przy badaniu zależności dwóch (lub więcej) wielkości otrzymuje się ciąg punktów pomiarowych $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$. Postulowana jest liniowa zależność tych wielkości $y = ax + b$, a parametry $a, b \in \mathbb{R}$ tej zależności mają zostać wyznaczone na podstawie zgromadzonych danych pomiarowych. Na ogół jednak na skutek nieuniknionych błędów przy pomiarze wartości x_i, y_i nie można tak dobrać współczynników $a, b \in \mathbb{R}$, żeby spełnione były wszystkie równania

$$y_i = ax_i + b, \qquad i = 1, \ldots, n. \tag{39}$$

Inaczej mówiąc, ten układ równań potraktowany jako układ dla niewiadomych a, b nie ma rozwiązań (jest sprzeczny). Szuka się więc przybliżonego rozwiązania, a jako miarę przybliżenia przyjmuje się sumę kwadratów odchyleń wyrażoną wzorem

$$f(a, b) = (y_1 - ax_1 - b)^2 + \ldots + (y_n - ax_n - b)^2 = \sum_{i=1}^{n} (y_i - ax_i - b)^2.$$

Dopasowanie prostej metodą najmniejszych kwadratów II

Po zapisaniu układu (39) w postaci wektorowej,

$$(y_1, y_2, \ldots, y_n) = a(x_1, x_2, \ldots, x_n) + (b, b, \ldots, b)$$

widzimy, że zadanie to sprowadza się do zagadnienia opisanego wcześniej.
Dla danych wektorów

$$Y = (y_1, \ldots, y_n), \qquad X = (x_1, \ldots, x_n)$$

wyznaczyć liczby $a, b \in \mathbb{R}$, takie by odległość wektora Y od kombinacji liniowej $aX + bC$, gdzie $C = (1, \ldots, 1)$ jest wektorem stałym, była najmniejsza.
W poprzedzających rozważaniach zadanie takie rozwiązaliśmy wykazując, że taką własnością *minimalizacji odległości* jest scharakteryzowany rzut prostopadły wektora Y na płaszczyznę rozpiętą przez wektory X, C. Dla wyznaczenia tego rzutu posłużymy się warunkiem ortogonalności wektora różnicy $Y - (aX + bC)$ do wektorów X i C (por. warunek **c)** Twierdzenia 13).
Stąd otrzymujemy równania

$$\big(Y - (aX + bC)\big) \cdot X = 0; \qquad \big(Y - (aX + bC)\big) \cdot C = 0$$

Dopasowanie prostej metodą najmniejszych kwadratów III

które po wykonaniu obliczeń prowadzą do układu równań o niewiadomych a, b

$$\begin{aligned} a(X\cdot X)+b(C\cdot X) &= Y\cdot X,\\ a(X\cdot C)+b(C\cdot C) &= Y\cdot C. \end{aligned} \tag{40}$$

Jest to wprowadzony wcześniej układ równań normalnych dla metody najmniejszych kwadratów. Zauważmy, że macierzą współczynników układu równań normalnych jest macierz Grama układu wektorów $\{X,\, C\}$.
Współczynniki układu równań (40) są dane przez

$$(C\cdot C)=n,\quad (X\cdot C)=(C\cdot X)=\sum_{i=1}^{n} x_i,\quad (X\cdot X)=\sum_{i=1}^{n} x_i^2,$$

$$Y\cdot C=\sum_{i=1}^{n} y_i,\quad (Y\cdot X)=\sum_{i=1}^{n} y_i x_i,$$

a dla wyznacznika macierzy układu mamy

$$D=\det\begin{pmatrix} X\cdot X & C\cdot X\\ X\cdot C & C\cdot C\end{pmatrix} = n\sum_{i=1}^{n} x_i^2-\left(\sum_{i=1}^{n} x_i\right)^2.$$

Wzory dla współczynników prostej regresji

Po rozwiązaniu układu (40) (wzorami Cramera lub inaczej) dostajemy **Wzory dla współczynników prostej** $y = ax + b$ **dopasowanej metodą najmniejszych kwadratów do danych** $Y = (y_1, \ldots, y_n)$, $X = (x_1, \ldots, x_n)$.

$$a = \frac{(C \cdot C)(X \cdot Y) - (C \cdot X)(C \cdot Y)}{D} = \frac{n \sum_{i=1}^{n} x_i y_i - \left(\sum_{i=1}^{n} x_i\right)\left(\sum_{i=1}^{n} y_i\right)}{n \sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2}$$

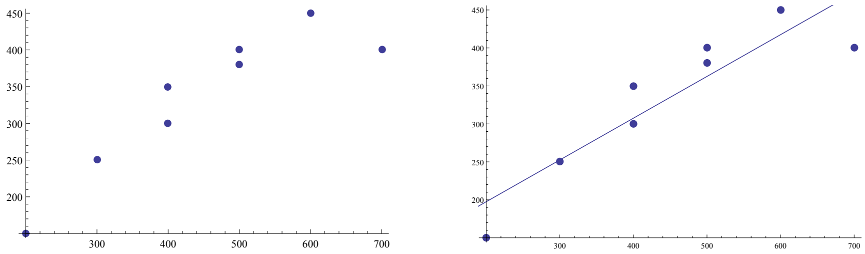
$$b = \frac{(X \cdot X)(C \cdot Y) - (C \cdot X)(X \cdot Y)}{D} = \frac{\left(\sum_{i=1}^{n} x_i^2\right)\left(\sum_{i=1}^{n} y_i\right) - \left(\sum_{i=1}^{n} x_i\right)\left(\sum_{i=1}^{n} x_i y_i\right)}{n \sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2}$$

W szczególnym przypadku, gdy w funkcji liniowej $y = ax + b$ przyjmiemy $b = 0$, układ (40) sprowadza się do formuły dla rzutu ortogonalnego wektora Y na jednowymiarową podprzestrzeń $\mathbb{R}\,X$

$$a\,X = \frac{X \cdot Y}{X \cdot X}\,X$$

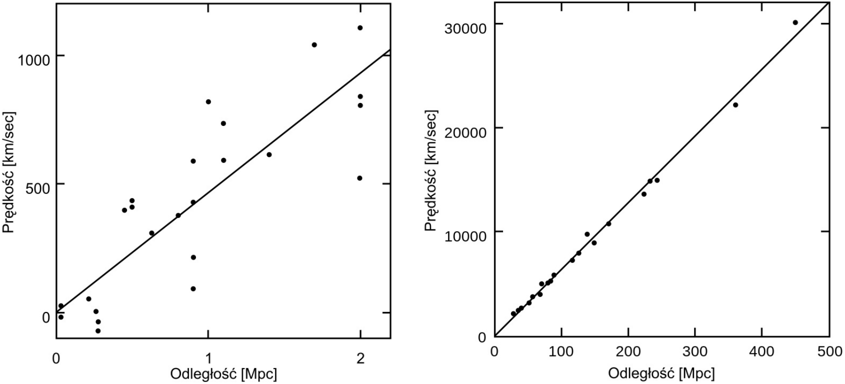
Ilustracja metody najmniejszych kwadratów

x	200	300	400	400	500	500	600	700
y	150	250	300	350	400	380	450	400



Zbiór danych i prosta regresji liniowej

Prawo Hubble — zastosowanie metody najmniejszych kwadratów



Prędkość ucieczki galaktyk rośnie z odległością