ROZDZIAŁ 1

Przestrzenie liniowe

Zaczniemy od przypomnienia definicji i własności najprostszych struktur algebraicznych: grupy i ciała.

1.1 Definicje i własności grupy i ciała

Definicja grupy:
Grupą $(G, \oplus)$ nazywamy zbiór G wraz z jednoznacznym działaniem dwuargumentowym $\oplus$: $G \times G \to G$ o następujących własnościach:

- łączność, tzn. dla dowolnych $a, b, c \in G$ zachodzi $(a \oplus b) \oplus c = a \oplus (b \oplus c)$
- istnieje element neutralny $e \in G$ taki, że dla każdego $a \in G$ zachodzi $a \oplus e = e \oplus a = a$
- każdy element $a \in G$ ma element przeciwny $b \in G$, tzn. $a \oplus b = e$

Jeśli działanie jest przemienne dla wszystkich elementów G, tzn. $a \oplus b = b \oplus a$, to mówimy, że grupa jest przemienna (lub abelowa).

Zauważmy, że zbiór G może być zupełnie dowolny, np. zbiór kwiatów, słów, idei, liczb itd. i działania $\oplus$ nie należy utożsamiać z dodawaniem. Działanie to oznacza jedynie tyle, że każdym dwóm elementom ze zbioru G jest przyporządkowany jakiś element tego zbioru. Ilustruje to poniższy

Przykład:
Niech G będzie zbiorem dwuelementowym $G = \{a, b\}$. Elementy a, b mogą reprezentować cokolwiek, np. przyjmiemy, że a to aster, a b to bratek. W tym zbiorze określamy działanie dwuargumentowe $\oplus$ w następujący sposób:

$$a \oplus a = a, \qquad b \oplus b = a, \qquad a \oplus b = b \oplus a = b$$

tzn. parze (aster,aster) jest przyporządkowany aster, parze (aster,bratek) jest przyporządkowany bratek itp. Łatwo sprawdzić bezpośrednim rachunkiem, że działanie $\oplus$ jest łączne, elementem neutralnym jest a (aster) i elementem odwrotnym do a jest a, natomiast elementem odwrotnym do b jest b. Tak zdefiniowana para $(G, \oplus)$ jest grupą przemienną.

Twierdzenie:

W dowolnej grupie $(G, \oplus)$ równanie $x \oplus a = d$ ma rozwiązanie $x = d \oplus b$ dla zadanych a, d, gdzie b jest elementem odwrotnym do a.

Dowód: Korzystając wyłącznie z definicji grupy i własności relacji równości tzn., zwrotności $(a = a)$, symetrii (jeśli $a = b$ to $b = a$) oraz przechodniości (jeśli $a = b$ i $b = c$ to $a = c$) mamy

$$x = x \oplus e = x \oplus (a \oplus b) = (x \oplus a) \oplus b = d \oplus b$$

Z drugiej strony, niech $x = d \oplus b$. Wstawiając do równania $x \oplus a = d$ sprawdzamy, że jest to rzeczywiście rozwiązanie tego równania. cbdo

---

Z tego twierdzenia wynika prawo "skracania stronami", tzn. równość $c \oplus a = d \oplus a$ pociąga za sobą $c = d$.

Zauważmy, że w dowodzie twierdzenia skorzystaliśmy jedynie z tego, że istnienieje element odwrotny do a bez zakładania, że taki element jest tylko jeden. Możemy teraz wykazać, że każdy element na tylko jeden element odwrotny. Gdyby bowiem były dwa różne elementy odwrotne b, c do a to mamy $a \oplus b = e$ i $a \oplus c = e$. Ale z powyższego twierdzenia rozwiązanie drugiej równości możemy zapisać w postaci $c = e \oplus b = b$, czyli $b = c$. Istnieje również tylko jeden element neutralny, bowiem gdyby były dwa e_1 i e_2, to z równości $e_1 \oplus e_2 = e_1$ (bo e_2 neutralny) i $e_1 \oplus e_2 = e_2$ (bo e_1 neutralny) i przechodniości relacji równości wynika, że $e_1 = e_2$.

Ponieważ w każdej grupie jest tylko jeden element neutralny i dla każdego elementu istnieje tylko jeden element odwrotny, wprowadzimy wygodną notację: element odwrotny do a będziemy oznaczali przez $-a$, a element neutralny e przez 0 (zero). Jest to tylko **oznaczenie**, bo w zbiorze G może nie być żadnych liczb. Np. odwołując się do powyższego przykładu, 0 to aster, a element odwrotny do bratka $-b$ to bratek b , czyli $b = -b$. I b nie jest zerem!

Jeśli przyjmiemy taką notację, to własności przemiennego działania $\oplus$ są formalnie takie same, jak dla dodawania, bo np. $x \oplus a = d \quad \Rightarrow x = d \oplus (-a)$ i tę ostatnią równość będziemy zapisywać jak $d - a$. Podkreślmy jeszcze raz, że $d - a$ oznacza jedynie $d \oplus (-a)$ i nie ma nic wspólnego z odejmowaniem.

Powyższa notacja jest addytywna. Alternatywnie można przyjąć notację multiplikatywną dla dla działania grupowego, np. $\odot$ zamiast $\oplus$, w której element neutralny oznacza się przez 1 (jedynka), a element odwrotny do a oznacza się przez a^{-1}. Taka notacja jest zgodna z "mnożeniem", bo wtedy np. $a \odot a^{-1} = 1$, ale znowu jest to tylko notacja i nic więcej.

Definicja ciała:

Ciałem $(\mathbb{K}, \uplus, \odot)$ nazywamy zbiór $\mathbb{K}$ z dwoma dwuargumentowymi działaniami $\uplus, \odot : \quad \mathbb{K} \times \mathbb{K} \to \mathbb{K}$ o następujących własnościach:

- $(\mathbb{K}, \uplus)$ jest grupą przemienną; element neutralny względem $\oplus$ oznaczymy przez 0, a element odwrotny względem $\uplus$ do a przez $-a$;
- $(\mathbb{K} - \{0\}, \odot)$ jest grupą przemienną; element neutralny względem $\odot$ oznaczymy przez 1, a element odwrotny względem $\odot$ do a przez a^{-1};
- dla dowolnych $a, b, c \in \mathbb{K}$ zachodzi

$$a \odot (b \uplus c) = (a \odot b) \uplus (a \odot c)$$

tzn. "mnożenie" $\odot$ lewostronnie jest rozdzielne względem "dodawania" $\uplus$.

Pytanie, czy zachodzi $(a \uplus b) \odot c = (a \odot c) \uplus (b \odot c)$ wygląda na trywialne, jeśli traktować $\uplus$ i $\odot$ jako zwykłe dodawanie i mnożenie. Ale dla abstrakcyjnych operacji $\uplus$ i $\odot$ odpowiedź na to pytanie wymaga formalnego dowodu:

$$\begin{aligned}
&(a \uplus b) \odot c \\
&= c \odot (a \uplus b) \qquad (\quad \text{przemienność } \odot) \\
&= (c \odot a) \oplus (c \odot b) \qquad (\quad \text{trzeci punkt definicji}) \\
&= (a \odot c) \oplus (b \odot c) \qquad (\quad \text{przemienność } \odot)
\end{aligned}$$

Podobnie wykazujemy, że $a \odot 0 = 0$. "Mnożąc" lewostronnie $a = a \uplus 0$ przez a dostajemy

$$a \odot a = a \odot (a \uplus 0) = (a \odot a) \uplus a \odot 0$$

Korzystając więc z udowodnionego wcześniej twierdzenia dostajemy rzeczywiście $a \odot 0 = 0$.

W podobny sposób pokazujemy formalnie, że

$$a \odot (-b) = (-a) \odot b = -(a \odot b), \qquad (-a) \odot (-b) = a \odot b$$

W szczególności zaczynamy wreszcie rozumieć, dlaczego $(-1) \odot (-1) = 1$. I to nawet wtedy, gdy elementy zbioru $\mathbb{K}$, w tym (-1), nie są liczbami!!

1.2 Przestrzenie wektorowe

Co łączy ze sobą zbiór rozważanych dotychczas przez nas wektorów, tj. odcinków skierowanych (strzałek) w przestrzeni $\mathbb{R}^3$ z np. wielomianami zmiennej rzeczywistej Zarówno wektory, jak i wielomiany możemy do siebie dodawać, odejmować, mnożyć przez liczby rzeczywiste, itd. Tych analogii jest zresztą więcej. Oba te przykłady są szczególnymi przypadkami pewnej struktury zwanej przestrzenią wektorową.

Definicja przestrzeni wektorowej:
Przestrzenią wektorową V nad ciałem $\mathbb{K}$ nazywamy strukturę $(V, \oplus; \mathbb{K}, \uplus, \odot; \otimes)$ o następujących własnościach:

- $(V, \oplus)$ jest grupą abelową,
- $(\mathbb{K}, \uplus, \odot)$ jest ciałem,
- ostatnie działanie $\otimes$ zwane mnożeniem jest takim odwzorowaniem $\mathbb{K} \times V \longrightarrow V$, że jeśli $a, b \in \mathbb{K}$, $\alpha, \beta \in V$ i 1 jest jednością w $\mathbb{K}$ (ze względu na mnożenie $\odot$) to zachodzi:
  1. $a \otimes (\alpha \oplus \beta) = a \otimes \alpha \oplus a \otimes \beta$
  2. $(a \uplus b) \otimes \alpha = a \otimes \alpha \oplus b \otimes \alpha$
  3. $a \otimes (b \otimes \alpha) = (a \odot b) \otimes \alpha$
  4. $1 \otimes \alpha = \alpha$

Często zamiast przestrzeń wektorowa mówimy przestrzeń liniowa, lub liniowa przestrzeń wektorowa. Czasem też sam zbiór V nazywany jest przestrzenią wektorową, jeśli spełnione są powyższe warunki. Elementy V zwane wektorami często będziemy oznaczali literami greckimi. Warto wspomnieć, że w literaturze np. fizycznej często używa się na wektory oznaczenia $\vec{v}$, $\vec{F}$, $\vec{r}$ lub oznacza się je tłustym drukiem. Symbole mnożenia zwykle się opuszcza.

Patrz plik dodatek1.pdf zawierający przypomnienie definicji i podstawowych własności grupy i ciała.

W definicji przestrzeni wektorowej dla działań $\oplus, \uplus$ użyliśmy notacji addytywnej, a dla działań $\otimes, \odot$ multiplikatywnej. W Dodatku 1 (dodatek1.pdf) wykazaliśmy, że dla takiej notacji obowiązują formalnie te same reguły, jak dla zwykłego dodawania i mnożenia. Dlatego też od tego momentu nie będziemy stosować tej szczególnej notacji tylko zwykłego dodawania i mnożenia (jeśli nie będzie wyraźnie zaznaczone, że jest inaczej). W wielu rozpatrywanych przypadkach rozważać będziemy przestrzenie wektorowe zbudowane z obiektów liczbowych nad ciałem liczb i wtedy te działania sprowadzają się do zwykłego dodawania i mnożenia. Należy jednak pamiętać, że w ogólności są to abstrakcyjne działania i charakter tych działań wynika z kontekstu.

Przykład 1:
Rozważana wcześniej przestrzeń $\mathbb{R}^3$ stanowi przykład przestrzeni wektorowej nad ciałem liczb rzeczywistych.

Przykład 2:
Ciało $\mathbb{K}$ jest przestrzenią wektorową nad sobą samym.

Przykład 3:
Przestrzeń wektorowa n- wymiarowa: $\mathbb{K}^n = \underbrace{\mathbb{K} \times \mathbb{K} \times \ldots \times \mathbb{K}}_{\text{n razy}}$ nad ciałem $\mathbb{K}$. Składa się ona z wektorów typu:

$$w = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix}, \quad v = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}, \quad u = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix},$$

gdzie w_i, v_i, u_i (składowe wektora) są elementami $\mathbb{K}$. Niech $a \in \mathbb{K}$, wówczas mnożenie wektora przez liczbę oraz dodawanie wektorów definiujemy jako

$$aw = \begin{bmatrix} aw_1 \\ aw_2 \\ \vdots \\ aw_n \end{bmatrix}, \qquad w + v = \begin{bmatrix} w_1 + v_1 \\ w_2 + v_2 \\ \vdots \\ w_n + v_n \end{bmatrix}.$$

Wektor zerowy to oczywiście wektor, w którym wszystkie składowe są równe zero. Szczególnymi przykładami $\mathbb{K}^n$ są $\mathbb{R}^n$ oraz $\mathbb{C}^n$. Przestrzenie te dla $n = 1, 2, 3$ już znamy. Dla $n > 3$ uogólnienie jest bardzo proste. Należy dodać, że przestrzenie te nie stanowią jakiś wyimaginowanych przykładów akademickich, ale maja praktyczne znaczenie i zastosowanie np. w fizyce. W mechanice zamiast rozważać na przykład ruch n punktów materialnych w przestrzeni $\mathbb{R}^3$ wygodnie jest rozważyć ruch jednego punktu (tzw. punktu-obrazu) w $\mathbb{R}^{3n}$ zwanej przestrzenią konfiguracyjną. Współrzędne punktu-obrazu x_j $(j = 1, 2, ..., 3n)$ tworzymy ze współrzędnych x_i, y_i, z_i wektorów położeń $\vec{\mathbf{r}}_i$ $(i = 1, 2, 3)$ punktów materialnych utożsamiając x_i, y_i, z_i z $x_{3i-2}, x_{3i-1}, x_{3i}$. Inny przykład, to tzw. przestrzeń fazowa w formalizmie kanonicznym, w której w każdej chwili czasu podajemy współrzędne położenia i składowe pędu, co jak łatwo widać jest przestrzenią sześciowymiarową.

Przykład 4:

Przestrzeń macierzy o wymiarze $n \times m$ nad $\mathbb{R}$ jest przestrzenią wektorową $n \cdot m$-wymiarową. Działaniami w tej przestrzeni są dodawanie macierzy i mnożenie macierzy przez liczbę. Jeśli $n = 1$ to otrzymujemy przestrzeń podaną w przykładzie 3., natomiast jeśli $m = 1$ to otrzymujemy przestrzeń wektorów transponowanych.

Przykład 5:

Zbiór wszystkich wielomianów względem ciała $\mathbb{K}$ z działaniami dodawania wielomianów oraz mnożenia współczynników przez liczby z tego ciała jest przestrzenią wektorową.

Przykład 6:

Zbiór wszystkich funkcji ciągłych rzeczywistych $(f : \mathbb{R} \longrightarrow \mathbb{R})$ z działaniami - dodawaniem funkcji i mnożeniem ich przez liczby rzeczywiste jest przestrzenią wektorową nad ciałem $\mathbb{R}$.

Udowodnimy teraz proste
Twierdzenie:

Jeśli V jest przestrzenią wektorową nad ciałem $\mathbb{K}$, to dla każdego $a \in \mathbb{K}$ i $w \in V$ mamy:

1. $0\, w = 0$ (pierwsze zero należy do $\mathbb{K}$, drugie do V),
2. $a0 = 0$,
3. $(-a)\, w = a(-w) = -(a\, w)$.

Dowód: Te równości mogą się w pierwszej chwili wydać zupełnie oczywiste. Jednak wymagają one formalnego dowodu. Zauważmy najpierw, że z definicji działania w przestrzeni wektorowej wynika, iż $(w, v \in V)$

$$(a-b)\,w = (a-b)\,w + b\,w - b\,w = (a-b+b)\,w - b\,w = a\,w - b\,w$$

oraz

$$a(w-v) = a(w-v) + a\,v - a\,v = a(w-v+v) - a\,v = a\,w - a\,v$$

Teraz już możemy dowieść równości
1.: $0\,w = (a-a)w = a\,w - a\,w = 0$. Podobnie postępujemy w przypadku równości
2.: $a0 = a(w-w) = a\,w - a\,w = 0$,.
Dowód ostatniej równości wymaga skorzystania z własności 4. zawartej w definicji przestrzeni wektorowej oraz zauważenia, że $(-1)\,w + w = (1-1)\,w = 0$. Ponieważ również $w - w = 0$, wiec musi zachodzić $(-1)w = -w$. Stąd już łatwo widać, że $a(-w) = (-1a)w = a((-1)w) = a(-w)$.

cbdo

Definicja podprzestrzeni wektorowej:

V_1 nazywamy podprzestrzenią wektorową przestrzeni wektorowej V, jeśli

1. $V_1 \subset V$
2. V_1 oraz V są przestrzeniami wektorowymi nad ciałem $\mathbb{K}$
3. działania na wektorach w V_1 są identyczne z działaniami na tych samych wektorach w V.

Przykład 1:

Każda przestrzeń wektorowa ma dwie trywialne podprzestrzenie wektorowe: siebie samą oraz przestrzeń złożoną z jednego tylko wektora 0.

Przykład 2:

$\mathbb{R}^1$ jest podprzestrzenią w $\mathbb{R}^2$ i $\mathbb{R}^3$. $\mathbb{R}^2$ jest podprzestrzenią w $\mathbb{R}^3$.

Przykład 3:

Podprzestrzenią przestrzeni wektorowej wszystkich wielomianów może być zbiór wielomianów stopnia niewiększego niż n, gdzie n jest liczbą naturalną. Zauważmy również, że zbiór wielomianów określonego stopnia (na przykład n) nie jest przestrzenią wektorową, ponieważ dodawanie wielomianów może wyprowadzić poza ten zbiór.

Przykład 4:

Podprzestrzenią wszystkich funkcji ciągłych rzeczywistych może być przestrzeń funkcji różniczkowalnych.

Przykład 5:

Niech V będzie płaszczyzną w $\mathbb{R}^3$ przechodzącą przez środek układu współrzędnych. Punkty tej płaszczyzny tworzą przestrzeń wektorową, która jest podprzestrzenią $\mathbb{R}^3$. Aby się o tym przekonać wystarczy wykazać, że działania dodawania wektorów i mnożenia wektora przez liczbę nie wyprowadzają poza ten zbiór. Płaszczyzna przechodząca przez początek układu współrzędnych jest opisana równaniem

$$ax + by + cz = 0. \tag{1.1}$$

Po pierwsze widać, że punkt $0 \equiv (0,0,0)$ należy do V. Jest to wektor zerowy. Następnie zauważamy, że jeśli wektor w należy do V to $-w$ też należy do V (jeśli w spełnia równanie (1.1), to $-w$ też spełnia to równanie). Ponadto $k\,w$ należy do V. Pozostaje jeszcze wykazać, że jeżeli $w \in V$ oraz $v \in V$ to $w + v \in V$. Niech $w = (a_1, a_2, a_3)$ oraz $v = (b_1, b_2, b_3)$, wówczas muszą być spełnione równania $a_1x+a_2y+a_3z = 0$ oraz $b_1x+b_2y+b_3z = 0$. Dodając te równania stronami otrzymujemy $(a_1 + b_1)x + (a_2 + b_2)y + (a_3 + b_3)z = 0$, a zatem suma wektorów $w + v$ też należy do V.

1.3 Liniowa niezależność wektorów

Niech $e_1, e_2, \ldots, e_n$ będą wektorami z przestrzeni wektorowej V nad ciałem $\mathbb{K}$, $a_1, a_2, \ldots, a_n \in \mathbb{K}$. Wektor

$$\sum_{j=1}^{n} a_j\, e_j \in V \tag{1.2}$$

nazywamy kombinacją liniową wektorów $e_1, e_2, \ldots, e_n$, a liczby $a_1, a_2, \ldots, a_n$ współczynnikami tej kombinacji.

Definicja liniowej niezależności wektorów:

Wektory $e_1, e_2, \ldots, e_n \in V$ nad ciałem $\mathbb{K}$ nazywamy liniowo niezależnymi wtedy i tylko wtedy, gdy z równości

$$\sum_{j=1}^{n} a_j\, w_j = 0, \qquad\qquad (a_j \in \mathbb{K})$$

wynika, że

$$a_1 = a_2 = \ldots = a_n = 0.$$

W przeciwnym przypadku mówimy, że wektory $\{e_i\}$ są liniowo zależne. Innymi słowy kombinacja liniowo niezależnych wektorów znika wtedy i tylko wtedy, gdy wszystkie współczynniki są równe zero.

Uwaga: $\{0\}$ jest zbiorem wektorów liniowo zależnych.

Twierdzenie:

$$\left(\begin{array}{l} \text{Wektory } e_1, e_2, \ldots, e_n \in V \\ \text{są liniowo zależne} \end{array} \right) \Longleftrightarrow \left(\bigvee_j \quad e_j = \sum_{s \neq j}^{n} a_s\, e_s \right).$$

Dowód:

$\Longrightarrow$ Ponieważ wektory są liniowo zależne, więc ich kombinacja liniowa $\sum a_j\, e_j$ znika nawet wówczas, kiedy nie wszystkie współczynniki a_j są równe zero. Niech na przykład $a_j \neq 0$. Wówczas
$e_j = -(a_1/a_j)e_1 - \ldots - (a_{j-1}/a_j)e_{j-1} - (a_{j+1}/a_j)e_{j+1} \ldots - (a_n/a_j)e_n$.
$\Longleftarrow$ Niechaj

$$\bigvee_j \quad e_j = \sum_{s\neq j}^{n} a_s\, e_s,$$

wówczas przyjmując $a_j = -1$ otrzymujemy $\sum_{s=1}^{n} a_s\, e_s = 0$. cbdo

---

Definicja przestrzeni rozpiętej przez zbiór wektorów:

Niech V będzie przestrzenią liniową nad ciałem $\mathbb{K}$, a S jej dowolnym podzbiorem. Zbiór $< S > \equiv \{w \in V : w$ jest kombinacją liniową wektorów z S$\}$, który jest (z konstrukcji) podprzestrzenią wektorową V, nazywamy podprzestrzenią rozpiętą (generowaną) przez zbiór S. Jeżeli S składa się z wektorów $e_1, e_2, \ldots, e_n$, to $< S >$ oznaczamy $\{e_1, e_2, \ldots, e_n\}$.

1.4 Baza

Definicja bazy:

Układ wektorów $e_1, e_2, \ldots, e_n$ nazywa się bazą przestrzeni wektorowej V, jeżeli

1. $e_j \in V \;\; (j = 1, 2, \ldots, n)$,
2. wektory $e_1, e_2, \ldots, e_n$ są liniowo niezależne
3. $V = \{e_1, e_2, \ldots, e_n\}$, tzn. wektory e_i rozpinają przestrzeń V.

Uwaga Ostatni punkt powyższej definicji oznacza, że każdy wektor przestrzeni V daje się przedstawić jako liniowa kombinacja wektorów bazy. Przestrzeń $\{0\}$ nie ma bazy, bo nie zawiera wektorów liniowo niezależnych.

Twierdzenie:

Jeśli $e_1, e_2, \ldots, e_n$ tworzą bazę w przestrzeni wektorowej V, to każdy wektor tej przestrzeni posiada jednoznaczne przedstawienie w postaci

$$w = \sum_{j=1}^{n} w_j\, e_j.$$

Dowód: (Nie wprost) Niech będą dwa różne przedstawienia wektora α:

$$w = \sum_{j=1}^{n} w_j\, e_j = \sum_{j=1}^{n} w'_j\, e_j.$$

Przedstawienia te są różne jeśli dla $(j = 1, 2, \ldots, n)$ nie wszystkie liczby $c_j = a_j - a'_j$ znikają równocześnie. Ale wówczas odejmując stronami oba przedstawienia wektora α

$$0 = \sum_{j=1}^{n} c_j\, e_j = \sum_{j=1}^{n} (a_j - a'_j)\, e_j.$$

otrzymujemy znikającą kombinację liniową wektorów bazy $e_1, e_2, \ldots, e_n$, ze współczynnikami c_j, z których nie wszystkie znikają. A zatem wektory bazy są liniowo zależne, co pozostaje w sprzeczności z założeniem twierdzenia. cbdo

---

Baza w przestrzeni wektorowej nie jest ustalona jednoznacznie. Możemy wybrać w danej przestrzeni nieskończenie wiele różnych baz. Zachodzi twierdzenie mówiące o tym, że wszystkie te bazy mają tę samą liczbę elementów.

Uwaga: W fizyce powszechnie używa się notacji Diraca, w której wektory oznaczane są przez tzw. kety (od połówki angielskiego słowa *bracket*), tzn. wektory oznaczane są przez

$$|w\rangle,\ |e_i\rangle \qquad \text{itp.}$$

Rozkład wektora na kombinację liniową wektorów bazy zapisuje się

$$|w\rangle = w_1|e_i\rangle + w_2|e_2\rangle + \ldots w_n|e_n\rangle.$$

Notacja ta jest wygodna, gdyż wyraźnie odróżnia wektory od współczynników liczbowych.

Twierdzenie:

Jeśli przestrzeń wektorowa V ma n-elementową bazę, to każda inna baza tej przestrzeni ma również n elementów.

Aby udowodnić to twierdzenie wykażemy najpierw następujący

Lemat:

Jeśli $|e_1\rangle, |e_2\rangle, \ldots, |e_n\rangle$ stanowią bazę w przestrzeni wektorowej, to każdy układ wektorów $|f_1\rangle$, $|f_2\rangle, \ldots, |f_m\rangle$, gdzie $m > n$ w tej przestrzeni jest liniowo zależny.

Dowód lematu: Jeśli $|e_1\rangle, |e_2\rangle, \ldots, |e_n\rangle$ jest bazą, to każdy wektor tej przestrzeni, a w szczególności wszystkie wektory $|f_j\rangle$ dla $(j = 1, 2, \ldots, m)$ dają się przedstawić jako kombinacje liniowe wektorów bazy

$$|f_s\rangle = \sum_{j=1}^{n} a_{sj}|e_j\rangle \qquad\qquad (s = 1, 2, \ldots, m).$$

Wykorzystamy te wzory, aby pokazać, że istnieje znikająca kombinacja wektorów $|f_s\rangle$ z nieznikającymi współczynnikami. Będzie to znaczyło, że $|f_s\rangle$ $(s = 1, 2, \ldots, m)$ tworzą kombinację liniowo zależną. Rzeczywiście, rozważmy znikającą kombinację

$$0 = b_1|f_1\rangle + b_2|f_2\rangle + \ldots + b_m|f_m\rangle = b_1 \sum_{j=1}^{n} a_{1j}|e_j\rangle + b_2 \sum_{j=1}^{n} a_{2j}|e_j\rangle + \ldots + b_m \sum_{j=1}^{n} a_{mj}|e_j\rangle =$$

$$= \left(\sum_{l=1}^{m} a_{l1}b_l\right)|e_1\rangle + \left(\sum_{l=1}^{m} a_{l2}b_l\right)|e_2\rangle + \ldots + \left(\sum_{l=1}^{m} a_{ln}b_l\right)|e_n\rangle = 0$$

Wektory $|e_1\rangle, |e_2\rangle, \ldots, |e_n\rangle$ z założenia stanowią bazę, zatem są liniowo niezależne. Muszą być zatem spełnione warunki

$$\left(\sum_{l=1}^{m} a_{lj}b_l\right) = 0 \qquad\qquad (j = 1, 2, \ldots, n).$$

Mamy tu do czynienia z układem n równań liniowych jednorodnych na m niewiadomych. Ponieważ niewiadomych jest więcej niż równań, istnieją rozwiązania niezerowe na $(b_s;\ s = 1, 2, \ldots, m)$. Otrzymaliśmy więc tezę naszego lematu.

Dowód twierdzenia: Z udowodnionego lematu twierdzenie równej liczby elementów dowolnej bazy przestrzeni wektorowej V wynika już bezpośrednio. Bowiem jeśli jedna z baz ma n elementów, a inna m to musi zachodzić jednocześnie $m \leq n$ oraz $n \leq m$, czyli $m = n$. cbdo

---

Definicja wymiaru przestrzeni wektorowej:

Przestrzeń wektorowa V, mająca bazę złożoną z n elementów ma wymiar równy liczbie elementów jej bazy. (dim V=n). Udowodnione powyżej twierdzenie zapewnia nam jednoznaczność tej definicji. Jeśli przestrzeń wektorowa ma bazę składającą się z nieskończonej liczby elementów to przyjmujemy dim V=∞. Zachodzi także dim $\{0\}$=0.

Przykład 1:

Przestrzeń $\mathbb{K}^n$ nad ciałem $\mathbb{K}$ jest n wymiarowa, tj. dim $\mathbb{K}^n = n$. Jako bazę można na przykład wybrać zbiór wektorów (tzw. baza kanoniczna)

$$|e_1\rangle = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad |e_2\rangle = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \ldots, |e_n\rangle = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}.$$

Jest to istotnie baza $\mathbb{K}^n$, ponieważ wszystkie wektory α_i należą do tej przestrzeni; są liniowo niezależne, ponieważ znikanie liniowej kombinacji

$$\sum_{j=1}^{n} a_j|e_j\rangle = \begin{bmatrix} a_1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ a_2 \\ \vdots \\ 0 \end{bmatrix} + \ldots + \begin{bmatrix} 0 \\ 0 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} = 0 = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix},$$

pociąga za sobą znikanie wszystkich współczynników a_j, a dowolny wektor $|w\rangle$ z przestrzeni wektorowej $\mathbb{K}^n$ daje się zapisać jako kombinacja wektorów $|e_j\rangle$ ponieważ

$$|w\rangle = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} = b_1 \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + b_2 \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix} + \ldots + b_n \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix} = \sum_{j=1}^{n} b_j|e_j\rangle$$

Przykład 2:

Szczególnym przykładem $\mathbb{K}^n$ mogą być a) $\mathbb{R}^n$ lub b) $\mathbb{C}^n$. W przykładzie b) jako ciało, nad którym jest skonstruowana przestrzeń wektorowa możemy wybrać $\mathbb{C}$ bądź $\mathbb{R}$, a wówczas wymiar będzie wynosił n lub $2n$. W szczególności zbiór liczb zespolonych, jako przestrzeń wektorowa nad ciałem liczb rzeczywistych jest przestrzenią dwuwymiarową, a jako wektory bazy możemy wybrać $[1]$ oraz $[i]$, które w oczywisty sposób są liniowo niezależne.

Przykład 3:

Jako bazę przestrzeni wektorowej nad $\mathbb{R}$ zbudowanej z wielomianów stopnia $\leq n$ względem ciała $\mathbb{K}$ można wybrać zbiór wielomianów $x^0 = 1, x^1 = x, x^2, \ldots, x^n$. Wymiar tej przestrzeni jest $n + 1$.

Przykład 4:

Przestrzeń zbudowana z funkcji rzeczywistych ($f : \mathbb{R} \to \mathbb{R}$) nad ciałem $\mathbb{R}$ jest nieskończenie wymiarowa.

Przykład 5:

Przestrzeń wektorowa złożona z macierzy $m \times n$, $V = M_{m\times n}(\mathbb{K})$ ze standardowymi działaniami dodawania macierzy i mnożenia macierzy przez skalar z ciała $\mathbb{K}$ ma wymiar mn. Bazą mogą być macierze posiadające tylko jeden element różny od zera (na przykład równy 1) w miejscu przecięcia się wiersza m z kolumną n .

Jak sprawdzić, czy dany zbiór wektorów stanowi układ liniowo niezależny. Na przykład mamy zbiór trzech wektorów z $\mathbb{R}^3$: $|\alpha_1\rangle = (1, -2, 3)$, $|\alpha_2\rangle = (5, 6, -1)$, $|\alpha_3\rangle = (3, 2, 1)$. Zbadamy czy znikanie kombinacji liniowej

$$a_1|\alpha_1\rangle + a_2|\alpha_2\rangle + a_3|\alpha_3\rangle$$

pociąga za sobą znikanie współczynników a_1, a_2, a_3. Wstawiamy do powyższego równania wektory otrzymując

$$\begin{aligned} &a_1(1, -2, 3) + a_2(5, 6, -1) + a_3(3, 2, 1) = \\ &(a_1 + 5a_2 + 3a_3, -2a_1 + 6a_2 + 2a_3, 3a_1 - a_2 + a_3) = \\ &(0, 0, 0) \end{aligned}$$

Równość ta prowadzi do układu trzech jednorodnych równań liniowych o niewiadomych a_j

$$\begin{aligned} a_1 + 5a_2 + 3a_3 &= 0 \\ -2a_1 + 6a_2 + 2a_3 &= 0 \\ 3a_1 - a_2 + a_3 &= 0 \end{aligned}$$

Jeżeli otrzymane równania mają nietrywialne rozwiązania (niezerowe), to wektory $\alpha_1, \alpha_2, \alpha_3$ tworzą układ liniowo zależny. Układ n równań liniowych jednorodnych o n niewiadomych (tutaj $n = 3$) ma niezerowe rozwiązania, gdy wyznacznik układu znika. W naszym przykładzie

$$\det A = \det \begin{bmatrix} 1 & 5 & 3 \\ -2 & 6 & 2 \\ 3 & -1 & 1 \end{bmatrix} = 0$$

(trzecie równanie jest różnicą pierwszego i drugiego). Oznacza to, że wektory $|\alpha_1\rangle, |\alpha_2\rangle, |\alpha_3\rangle$ tworzą układ liniowo zależny.

Ustaliliśmy już, że w danej bazie, np. $|e_1\rangle, |e_2\rangle, \ldots, |e_n\rangle$, wektory $|\alpha\rangle$ przestrzeni V ($\dim V{=}n$) mają jednoznaczne przedstawienie w postaci liniowej kombinacji wektorów bazy

$$|\alpha\rangle = a_1|e_1\rangle + a_2|e_2\rangle + \ldots + a_n|e_n\rangle = \sum_{s=1}^{n} a_s|e_s\rangle.$$

Innymi słowy, każdemu wektorowi można przypisać jednoznacznie współczynniki a_i jego rozwinięcia w danej bazie:

$$|\alpha\rangle \leftrightarrow \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}$$

Pamiętać jednak należy, że wektor reprezentowany przez a_i zależy nie tylko od tych współczynników, ale też od obranej bazy (zmiana bazy pociąga za sobą zmianę tych współczynników, ale tym zajmiemy się później).

Dodawaniu wektorów w V i ich mnożeniu przez liczby z $\mathbb{K}$ odpowiada dodawanie kolumn – wektorów z $\mathbb{K}^n$ (takich, jak napisana wyżej) i ich mnożenie przez liczby, tak jak to się działo w $\mathbb{K}^n$. Oznacza to, że zamiast pracować z wektorami w n-wymiarowej przestrzeni V możemy pracować z odpowiadającymi im wektorami-kolumnami w $\mathbb{K}^n$. Np. ustalenie, czy k wektorów z przestrzeni wektorowej V

$$\begin{aligned} |\beta_1\rangle &= b_{11}|e_1\rangle + b_{12}|e_2\rangle + \ldots + b_{1n}|e_n\rangle \\ |\beta_2\rangle &= b_{21}|e_1\rangle + b_{22}|e_2\rangle + \ldots + b_{2n}|e_n\rangle \\ &\vdots \\ |\beta_k\rangle &= b_{k1}|e_1\rangle + b_{k2}|e_2\rangle + \ldots + b_{kn}|e_n\rangle \end{aligned}$$

stanowi układ liniowo niezależny sprowadzamy do badania liniowej niezależności wektorów

$$\begin{bmatrix} b_{11} \\ b_{12} \\ \vdots \\ b_{1n} \end{bmatrix}, \begin{bmatrix} b_{21} \\ b_{22} \\ \vdots \\ b_{2n} \end{bmatrix}, \ldots, \begin{bmatrix} b_{k1} \\ b_{k2} \\ \vdots \\ b_{kn} \end{bmatrix}$$

z przestrzeni $\mathbb{K}^n$. Zajmiemy się teraz bardziej szczegółowo tą sprawą.

1.5 Badanie liniowej niezależności wektorów w $\mathbb{K}^n$

Jak już powiedzieliśmy, badanie niezależności liniowej wektorów $\{|\beta_i\rangle\}$ sprowadza się w pewnej konkretnej bazie $\{|e_j\rangle\}$ do sprawdzenia, czy wektory-kolumny $\{b_{ij}\}$ reprezentujące te wektory są liniowo niezależne. Jeśli te wektory-kolumny ustawimy w macierz, to badając rząd tak utworzonej macierzy znajdziemy liczbę liniowo niezależnych kolumn. Przypomnijmy, że *rząd macierzy jest równy maksymalnej liczbie kolumn liniowo niezależnych.* Ponieważ rząd macierzy jest równy rzędowi macierzy transponowanej, to do badania liniowej niezależności wektorów nie jest istotne, czy wektory-kolumny $\{b_{ij}\}$ tworzą kolumny, czy wiersze badanej macierzy. Jeśli liczba liniowo niezależnych wektorów jest mniejsza od liczby badanych wektorów, to chcielibyśmy też określić, które wektory możemy wybrać jako niezależne. Najpierw wprowadzimy kilka pojęć.

Minor bazowy:

Rozważmy macierz należącą do $M(\mathbb{K})$. Minor macierzy nazywamy bazowy, jeśli jest on różny od zera i albo nie ma minorów stopnia wyższego, albo wszystkie minory stopnia wyższego zerują się. Kolumny (i wiersze) macierzy przecinające macierz minora bazowego nazywamy kolumnami (i wierszami) bazowymi.

Lemat o minorze bazowym:

Kolumny (wiersze) bazowe są liniowo niezależne. Dowolna kolumna (wiersz) jest liniową kombinacją kolumn (wierszy) bazowych.

Dowód: Dla ustalenia uwagi zakładamy, że minorowi bazowemu odpowiada podmacierz w lewym górnym rogu macierzy $A \in M_{m\times n}(\mathbb{K})$ (to założenie nie zmniejsza ogólności rozważań). Stopień minora bazowego oznaczymy przez r, a jego wartość przez D. Oczywiście $D \neq 0$ (minor bazowy!). W poniższej równości czerwoną czcionką wyróżniliśmy elementy podmacierzy (0 wymiarze $r \times r$) minora bazowego.

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1r} & \dots & a_{1n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{r1} & a_{r2} & \dots & a_{rr} & \dots & a_{rn} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mr} & \dots & a_{mn} \end{bmatrix}$$

Weźmy teraz dowolne indeksy i, k $(1 \leq i \leq m,\ 1 \leq k \leq n)$ i obliczmy wyznacznik macierzy stopnia $r+1$ utworzonej przez dopisanie do podmacierzy minora bazowego kolumny i wiersza w następujący sposób

$$\Delta_{ik} = \begin{vmatrix} \mathbf{a_{11}} & \mathbf{a_{12}} & \dots & \mathbf{a_{1r}} & a_{1k} \\ \dots & \dots & \dots & \dots & \dots \\ \mathbf{a_{r1}} & \mathbf{a_{r2}} & \dots & \mathbf{a_{rr}} & a_{rk} \\ a_{i1} & a_{i2} & \dots & a_{ir} & a_{ik} \end{vmatrix}$$

Są trzy możliwości: (1) $i \leq r$, (2) $k \leq r$ oraz (3) $i, k > r$. W każdym z tych przypadków wyznacznik Δ_{ik} się zeruje. W pierwszym lub drugim przypadku wynika to z faktu, że w powyższej macierzy powtarza się odpowiednio wiersz lub kolumna. Gdyby $\Delta_{ik} \neq 0$ w trzecim przypadku,

to minor bazowy A byłby wyższego stopnia niż r, co przeczy założeniu lematu.

Rozwiniemy teraz Δ_{ik} względem ostatniego wiersza oznaczając odpowiednie dopełnienie algebraiczne przez A_j

$$\Delta_{ik} = A_1 a_{i1} + \ldots + A_r a_{ir} + A_{r+1} a_{ik} = 0$$

Zauważmy, że $A_{r+1} = D \neq 0$. Pozwala to podzielić rozważaną równość przez D i zapisać ją w postaci

$$a_{ik} = (-\frac{A_1}{D}) a_{i1} + \ldots + (-\frac{A_r}{D}) a_{ir}$$

Ustalmy teraz k i zmieniajmy i $(i = 1, 2, \ldots, m)$. Prowadzi to do m równości, które możemy zebrać w odpowiednią równość kolumn

$$\begin{bmatrix} a_{1k} \\ \vdots \\ a_{mk} \end{bmatrix} = (-\frac{A_1}{D}) \begin{bmatrix} a_{11} \\ \vdots \\ a_{m1} \end{bmatrix} + \ldots + (-\frac{A_r}{D}) \begin{bmatrix} a_{1r} \\ \vdots \\ a_{mr} \end{bmatrix}, \quad (k = 1, \ldots, n)$$

Jest to przedstawienie k-tej kolumny macierzy $A \in M_{m\times n}(\mathbb{K})$ w postaci kombinacji liniowej kolumn bazowych. Ponieważ k może być dowolnie wybrane spośród liczb $1, 2, \ldots, n$, więc stwierdzenie to jest słuszne dla wszystkich kolumn. cbdo

---

Podobnie dowodzimy analogiczny lemat dla wierszy.

Twierdzenie:

Wyznacznik macierzy równy jest zeru wtedy i tylko wtedy, gdy pomiędzy kolumnami tej macierzy istnieje liniowa zależność.
Identyczne twierdzenie prawdziwe jest dla wierszy.

Dowód: Niech kolumny macierzy (traktujemy je jak wektory w $\mathbb{K}^n$) stanowią układ liniowo zależny. To znaczy, można znaleźć taką kolumnę, którą da się przedstawić jako kombinację liniową pozostałych. Z własności wyznacznika wynika, że musi on znikać.

Dowodząc w drugą stronę chcemy pokazać, że ze znikania wyznacznika wynika liniowa zależność kolumn. Jest tak istotnie, bo gdy wyznacznik znika, wówczas minor bazowy jest stopnia niższego niż badana macierz i z lematu o minorze bazowym wynika, że przynajmniej jedna z kolumn macierzy jest liniową kombinacją innych (bazowych). Oznacza to, że kolumny tworzą układ liniowo zależny. cbdo

---

Z definicji rzędu macierzy i jej minora bazowego wynika, że rząd macierzy równy jest stopniowi jej minora bazowego. Wszystkie kolumny przechodzące przez minor bazowy są liniowo niezależne, a wszystkie pozostałe dają się przedstawić jako ich kombinacja liniowa. Z równości $\det A^T = \det A$ wynika, że to, co powiedziano o kolumnach słuszne jest w przypadku wierszy. Wynika stąd

Wniosek praktyczny:

Aby ustalić liniową niezależność układu wektorów z $\mathbb{K}^n$ należy utworzyć z nich macierz ustawiając je obok siebie (gdy wektory z $\mathbb{K}^n$ zapisane są w postaci kolumn) albo ustawiając je jeden nad drugim ((gdy wektory z $\mathbb{K}^n$ zapisane są w postaci wierszy). Następnie badamy rząd tak powstałej macierzy. Wektory liniowo niezależne to te, które przecinają minor bazowy. Pozostałe są ich liniowymi kombinacjami.

Przykład:

Zbadać, czy następujący układ wektorów z $\mathbb{K}^4$ jest liniowo niezależny

$$|\gamma_1\rangle = (1,0,0,2,5), \qquad |\gamma_2\rangle = (0,1,0,3,4),$$
$$|\gamma_3\rangle = (0,0,1,4,7), \qquad |\gamma_4\rangle = (2,-3,4,11,12)$$

Z wektorów $|\gamma_1\rangle, |\gamma_2\rangle, |\gamma_3\rangle, |\gamma_4\rangle$ układamy macierz o wymiarze 4×5

$$A = \begin{bmatrix} 1 & 0 & 0 & 2 & 5 \\ 0 & 1 & 0 & 3 & 4 \\ 0 & 0 & 1 & 4 & 7 \\ 2 & -3 & 4 & 11 & 12 \end{bmatrix}$$

i badamy jej rząd. Oczywiście $\mathrm{rz}A \leq 4$. Jeśli $\mathrm{rz}A = 4$, to wówczas wszystkie wektory $|\gamma_i\rangle$ tworzyłyby układ liniowo niezależny. Aby obliczyć $\mathrm{rz}A$ odbierzemy następującą drogę: (1) od 4-tego wiersza odejmujemy 1-wszy pomnożony przez 2, (2) do powstałego wiersza 4 dodajemy 2-gi pomnożony przez 3 i (3) od tak otrzymanego wiersza odejmujemy 3-ci pomnożony przez 4. Krótko mówiąc, w macierzy A zastępujemy 4-ty wiersz nim samym plus odpowiednią kombinacją liniową pozostałych wierszy, co nie zmienia rzędu macierzy A

$$\mathrm{rz}A = \mathrm{rz}\begin{bmatrix} 1 & 0 & 0 & 2 & 5 \\ 0 & 1 & 0 & 3 & 4 \\ 0 & 0 & 1 & 4 & 7 \\ 0 & 0 & 0 & 0 & -14 \end{bmatrix} = 4$$

Aby szybko przekonać się o tym, że $\mathrm{rz}A = 4$, wystarczy np. z otrzymanej macierzy utworzyć minor bazowy wykreślając czwartą kolumnę.

Ostatecznie wnioskujemy, że $|\gamma_1\rangle, |\gamma_2\rangle, |\gamma_3\rangle, |\gamma_4\rangle$ tworzą układ liniowo niezależny.

Przykład:

Znaleźć dowolną bazę układu wektorów

$$|\gamma_1\rangle = (5,2,-3,1), \qquad |\gamma_2\rangle = (4,1,-2,3),$$
$$|\gamma_3\rangle = (1,1,-1,-2), \qquad |\gamma_4\rangle = (3,4,-1,2)$$

i wyrazić pozostałe wektory (nie wchodzące w skład bazy) jako kombinacje liniowe wektorów tej bazy.

Łatwo zauważyć, że $|\gamma_1\rangle = |\gamma_2\rangle + |\gamma_3\rangle$. Wektory $|\gamma_1\rangle, \ldots, |\gamma_4\rangle$ nie są więc liniowo niezależne. Gdybyśmy tego nie zauważyli, to doszlibyśmy do tego badając rząd macierzy utworzonej z tych wektorów: byłby on mniejszy niż 4. Zbadajmy liniową niezależność układu

$|\gamma_2\rangle, |\gamma_3\rangle, |\gamma_4\rangle$. Uwaga: Moglibyśmy postąpić inaczej, np. stwierdzając, że $|\gamma_2\rangle = |\gamma_3\rangle - |\gamma_1\rangle$, wyrzucić $|\gamma_2\rangle$ i badając liniową niezależność układu $|\gamma_1\rangle, |\gamma_3\rangle, |\gamma_4\rangle$. Wróćmy do wcześniejszego wyboru. Badamy rząd macierzy utworzonej z $|\gamma_2\rangle, |\gamma_3\rangle, |\gamma_4\rangle$. Okazuje się, że

$$\mathrm{rz}\left[\begin{array}{ccc|c} \boxed{4} & \boxed{1} & \boxed{-2} & 3 \\ \boxed{1} & \boxed{1} & \boxed{-1} & -2 \\ \boxed{3} & \boxed{4} & \boxed{-1} & 2 \end{array}\right] = 3$$

Aby się o tym przekonać, można wybrać minor bazowy taki, jaki zaznaczono ramką w powyższym wzorze. Wniosek: wektory $|\gamma_2\rangle, |\gamma_3\rangle, |\gamma_4\rangle$ są liniowo niezależne i wektor $|\gamma_1\rangle$ daje się przedstawić jako wcześniej wspomniana kombinacja liniowa wektorów $|\gamma_2\rangle$ i $|\gamma_3\rangle$.

Inne sformułowanie tego wniosku: wektory $|\gamma_1\rangle, |\gamma_2\rangle, |\gamma_3\rangle, |\gamma_4\rangle$, jak każdy inny układ wektorów rozpinają pewną przestrzeń wektorową $V = \{|\gamma_1\rangle, |\gamma_2\rangle, |\gamma_3\rangle, |\gamma_4\rangle\}$. Bazą w tej przestrzeni mogą być np. wektory $|\gamma_2\rangle, |\gamma_3\rangle, |\gamma_4\rangle$. $\dim V = 3$.

Rozwiążemy jeszcze raz powyższy problem używając tym razem *metody eliminacji Gaussa-Jordana*. Zapytamy o jakąkolwiek bazę przestrzeni $V = \{|\gamma_1\rangle, |\gamma_2\rangle, |\gamma_3\rangle, |\gamma_4\rangle\}$ i $\dim V$. Ogólnie mówiąc, metoda Gaussa-Jordana polega na zastępowaniu wierszy macierzy przez nie same plus odpowiednią kombinację liniową innych wierszy. Te wiersze, które były kombinacjami pozostałych udaje się wyzerować, pozostałe nie. Ostatecznie zostają tylko te wiersze, które tworzą (maksymalny) układ liniowo niezależny. Wypowiedzenie tego w języku wektorów (wierszy) oznacza, że znajdujemy jakąś bazę w przestrzeni rozpiętej przez zadane na początku wektory i z liczby jej elementów wnioskujemy o wymiarze tej przestrzeni.

Konstruujemy z wektorów $|\gamma_1\rangle, |\gamma_2\rangle, |\gamma_3\rangle, |\gamma_4\rangle$ macierz stawiając $|\gamma_3\rangle$ na pierwsze miejsce (dla wygody, bo pierwszy wyraz to 1). Dostajemy macierz w lewej strony w równaniu poniżej. Teraz wykonamy szereg operacji na tej macierzy, żeby ją maksymalnie uprościć. Krok A: wiersz 2-gi (W_2) zastępujemy przez niego samego minus 5 razy wiersz pierwszy, tzn. $W_2 \to W_2 - 5W_1$. Podobnie $W_3 \to W_3 - 4W_1$ oraz $W_4 \to W_4 - 3W_1$. Krok B: ponownie dla wygody przestawimy wiersz ostatni na drugie miejsce. Krok C: oznaczając teraz wiersze tyldą, dokonujemy następujących przekształceń $\tilde{W}_3 \to \tilde{W}_3 + 3\tilde{W}_2$ oraz $\tilde{W}_4 \to \tilde{W}_4 + 3\tilde{W}_2$. Sekwencja tych operacji jest pokazana poniżej:

$$\left[\begin{array}{cccc} 1 & 1 & -1 & -2 \\ 5 & 2 & -3 & 1 \\ 4 & 1 & -2 & 3 \\ 3 & 4 & -1 & 2 \end{array}\right] \stackrel{A}{\to} \left[\begin{array}{cccc} 1 & 1 & -1 & -2 \\ 0 & -3 & 2 & 11 \\ 0 & -3 & 2 & 11 \\ 0 & 1 & 2 & 8 \end{array}\right] \stackrel{B}{\to} \left[\begin{array}{cccc} 1 & 1 & -1 & -2 \\ 0 & 1 & 2 & 8 \\ 0 & -3 & 2 & 11 \\ 0 & -3 & 2 & 11 \end{array}\right] \stackrel{C}{\to} \left[\begin{array}{cccc} 1 & 1 & -1 & -2 \\ 0 & 1 & 2 & 8 \\ 0 & 0 & 8 & 35 \\ 0 & 0 & 8 & 35 \end{array}\right]$$

Kolejny krok D: od czwartego wiersza odejmujemy trzeci, a wiersz trzeci dzielimy przez 8. Krok E: teraz od drugiego wiersza odejmujemy 2 razy wiersz trzeci, a na miejscu pierwszego piszemy sumę pierwszego i trzeciego. W końcowym kroku F od pierwszego wiersza odejmujemy drugi. Ta sekwencja kroków pokazana jest poniżej:

$$\stackrel{D}{\to} \left[\begin{array}{cccc} 1 & 1 & -1 & -2 \\ 0 & 1 & 2 & 8 \\ 0 & 0 & 1 & 35/8 \\ 0 & 0 & 0 & 0 \end{array}\right] \stackrel{E}{\to} \left[\begin{array}{cccc} 1 & 1 & 0 & 19/8 \\ 0 & 1 & 0 & -3/4 \\ 0 & 0 & 1 & 35/8 \\ 0 & 0 & 0 & 0 \end{array}\right] \stackrel{F}{\to} \left[\begin{array}{cccc} 1 & 0 & 0 & 25/8 \\ 0 & 1 & 0 & -3/4 \\ 0 & 0 & 1 & 35/8 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Otrzymujemy ostatecznie macierz w postaci kanonicznej Gaussa, w której nad i pod jedynkami rozpoczynającymi niezerowe wiersze stoją wyłącznie jedynki. Widać teraz, że układ był liniowo zależny. Jeden z wektorów był liniową kombinacją pozostałych. Przestrzeń rozpięta przez $|\gamma_1\rangle, |\gamma_2\rangle, |\gamma_3\rangle, |\gamma_4\rangle$ ma wymiar 3. Jedną z możliwych baz tworzą wektory $|\alpha_1\rangle = (1, 0, 0, 25/8)$, $|\alpha_2\rangle = (0, 1, 0, -3/4)$, $|\alpha_3\rangle = (0, 0, 1, 35/8)$. *Sprawdzenie*: widać, że $|\alpha_1\rangle, |\alpha_2\rangle, |\alpha_3\rangle$ tworzą układ liniowo niezależny, gdyż ze znikania kombinacji liniowej $a_1|\alpha_1\rangle + a_2|\alpha_2\rangle + a_3|\alpha_3\rangle = (a_1, a_2, a_3, 25a_1/8 - 3a_2/4 + 35a_3/8) = (0, 0, 0, 0)$ wynika zerowanie a_1, a_2, a_3.

Przykład:

Udowodnić liniową niezależność układu funkcji

$$\sin x, \sin 2x, \sin 3x, \ldots, \sin nx$$

w przestrzeni wektorowej funkcji rzeczywistych, gdzie dodawanie funkcji i mnożenie przez liczby rzeczywiste zdefiniowane jest w sposób naturalny.

Rozwiązanie przez indukcję. Dla $n = 1$ twierdzenie jest oczywiste. Funkcja sin nie jest funkcja zerową, tj. nie znika tożsamościowo. Z $a_1 \sin x = 0$ wynika $a_1 = 0$, a więc jest liniowo niezależna.

Niech twierdzenie zachodzi dla n, tj. niech z faktu, że dla każdej liczby $x \in \mathbb{R}$

$$a_1 \sin x + a_2 \sin 2x + \ldots + a_n \sin nx = 0$$

wynika, że $a_1 = a_2 = \ldots = a_n = 0$.

Rozpatrzmy teraz (dla każdego $x \in \mathbb{R}$)

$$b_1 \sin x + b_2 \sin 2x + \ldots + b_n \sin nx + b_{n+1} \sin(n+1)x = 0$$

gdzie b_j $(j = 1, 2, \ldots, n+1)$ są rzeczywistymi współczynnikami tej kombinacji liniowej. Należy pokazać, że wszystkie b_j znikają. Gdy $b_{n+1} = 0$, wówczas na mocy założenia indukcyjnego (przypadek n) pozostałe $b_1, b_2, \ldots, b_n$ też znikają. Ale b_{n+1} nie może być różne od zera. Gdyby jednak $b_{n+1} \neq 0$, to ze znikania badanej kombinacji liniowej po jej podzieleniu przez b_{n+1} i wprowadzeniu oznaczenia $c_s = -\frac{b_s}{b_{n+1}}$ otrzymalibyśmy, że

$$\sin(n+1)x = c_1 \sin x + c_2 \sin 2x + \ldots + c_n \sin nx.$$

Różniczkując to równanie dwukrotnie po x i uporządkowaniu dostajemy

$$\sin(n+1)x = \frac{1^2 c_1}{(n+1)^2} \sin x + \frac{2^2 c_2}{(n+1)^2} \sin 2x + \ldots + \frac{n^2 c_n}{(n+1)^2} \sin nx$$

Odejmując stronami otrzymujemy

$$0 = c_1(1 - \frac{1^2}{(n+1)^2}) \sin x + c_2(1 - \frac{2^2}{(n+1)^2}) \sin 2x + \ldots + c_n(1 - \frac{n^2}{(n+1)^2}) \sin nx$$

Ponieważ na mocy założenia indukcyjnego twierdzenie zachodzi dla n, to wszystkie współczynniki tej n-elementowej kombinacji znikają. Ale $1 - \frac{k^2}{(n+1)^2}) > 0$ dla $k \leq n$, więc $c_1 = c_2 = \ldots = c_n$ muszą znikać, co oznacza, że $b_1 = b_2 = \ldots = b_n = 0$. Gdy istotnie znikają współczynniki b_j aż do $j = n$, to z badanej na wstępie kombinacji pozostaje tylko

$$b_{n+1} \sin(n+1)x = 0$$

tożsamościowo (tj. dla każdego x), co oznacza, że jednak $b_{n+1} = 0$.

1.6 Przestrzenie unitarne

W podrozdziale tym $\mathbb{K}$ oznaczać będzie albo ciało liczb rzeczywistych $\mathbb{R}$, albo zespolonych $\mathbb{C}$.

Rozważając przestrzeń wektorową $\mathbb{R}^3$ zdefiniowaliśmy iloczyn skalarny. Zastanowimy się teraz nad uogólnieniem tego bardzo ważnego pojęcia na dowolne przestrzenie wektorowe. W ogólnych przestrzeniach wektorowych iloczyn skalarny określamy aksjomatycznie.

Definicja iloczynu skalarnego:

Niech V będzie przestrzenią wektorową nad ciałem $\mathbb{K}$. Iloczynem skalarnym (oznaczonym przez $\langle\ |\ \rangle$) nazywamy odwzorowanie, które parze wektorów $|\alpha\rangle$, $|\beta\rangle$ przyporzadkowuje liczbę z ciała $\mathbb{K}$:

$$V \times V \ni (|\alpha\rangle\,, |\beta\rangle) \longrightarrow \langle\alpha|\beta\rangle \in \mathbb{K},$$

które ma następujące własności:

1. $\langle\alpha|\beta\rangle = \langle\beta|\alpha\rangle^*$
2. $\langle\alpha|\beta+\gamma\rangle = \langle\alpha|\beta\rangle + \langle\alpha|\gamma\rangle$
   $\langle\alpha|a\beta\rangle = a\langle\alpha|\beta\rangle$
3. $\langle\alpha|\alpha\rangle > 0$, gdy $\alpha \neq 0$

Inne, często spotykane w literaturze oznaczenia iloczynu skalarnego, to $(\alpha|\beta)$, $\vec{\alpha}\cdot\vec{\beta}$ lub po prostu $\alpha\beta$, jeśli z kontekstu wiadomo, że α i β sa wektorami. Na wykładzie będziemy czasami oznaczali wektory pisząc kety $|\alpha\rangle$, ale też często pisząc po prostu literę α.

Własności iloczynu skalarnego wynikające bezpośrednio z definicji:

$a)$ $\langle 0|\beta\rangle = \langle\beta|0\rangle = 0$
$b)$ $\langle\alpha+\beta|\gamma\rangle = \langle\alpha|\beta\rangle + \langle\alpha|\gamma\rangle$
$c)$ $\langle b\alpha|\beta\rangle = b^*\langle\alpha|\beta\rangle$

Udowodnimy np. własność a).

$$\langle\beta|0\rangle = \langle\beta|\alpha-\alpha\rangle = \langle\beta|\alpha\rangle - \langle\beta|\alpha\rangle = 0$$

oraz

$$\langle 0|\beta\rangle = \langle\beta|0\rangle^* = 0^* = 0$$

Definicja przestrzeni unitarnej:

Przestrzenią unitarną nad ciałem $\mathbb{K}$ nazywamy przestrzeń wektorową V nad ciałem $\mathbb{K}$, w której określony jest iloczyn skalarny. Pełne oznaczenie przestrzeni unitarnej ma postać $(V, \langle\ |\ \rangle)$. Czasami piszemy $\langle\cdot|\cdot\rangle$, gdzie kropki oznaczają miejsca, w które wstawiamy argumenty (tutaj wektory) iloczynu skalarnego.

Przykłady:

1. Znanym już nam przykładem iloczynu skalarnego jest zdefiniowany wcześniej iloczyn $\vec{u} \cdot \vec{v}$ w przestrzeni $\mathbb{R}^3$.

2. Niech $V = \mathbb{C}^n$, a wektory $|\alpha\rangle$ i $|\beta\rangle$ w pewnej bazie mają postać $|\alpha\rangle = (\alpha_1, \alpha_2, \ldots, \alpha_n)$, $|\beta\rangle = (\beta_1, \beta_2, \ldots, \beta_n)$. Iloczyn skalarny definiujemy w następujący sposób

$$\langle\alpha|\beta\rangle = \alpha_1^*\beta_1 + \alpha_2^*\beta_2 + \ldots + \alpha_n^*\beta_n = \sum_{j=1}^{n} \alpha_j^*\beta_j.$$

3. Niech V będzie przestrzenią funkcji ciągłych na $[0, 1] \subset \mathbf{R}$ przyjmujących wartości zespolone. Iloczyn skalarny definiujemy jako

$$\langle f|g\rangle = \int_0^1 f^*(x)g(x)\mathrm{d}x.$$

Nie są to oczywiście jedyne możliwe definicje iloczynu skalarnego w tych przestrzeniach.

Definicja normy:

Normą w przestrzeni unitarnej V nazywamy odwzorowanie

$$V \ni \alpha \longrightarrow \|\alpha\| \equiv \sqrt{\langle\alpha|\alpha\rangle} \in \mathbb{R}$$

Zauważmy, że $\langle\alpha|\alpha\rangle$jest zawsze rzeczywiste i nieujemne, co wynika z definicji iloczynu skalarnego. $\|\alpha\|$, tj. normę wektora α, nazywamy też długością wektora α. Gdy $\|\xi\| = 1$ to mówimy, że wektor ξ jest unormowany. Inne często spotykane oznaczenie normy wektorów, to $|\xi|$.

Definicja kąta między wektorami:

Jeśli V jest przestrzenią wektorową nad $\mathbb{R}$, to możemy określić kąt między wektorami. Kątem miedzy wektorami $\alpha, \beta \in V$ nazywamy taką liczbę $\phi \in [0, \pi]$, że

$$\cos\phi = \frac{\langle\alpha|\beta\rangle}{\|\alpha\|\,\|\beta\|}.$$

Zauważmy, że tak określony $\cos\phi$ jest liczbą rzeczywistą, gdyż $\langle\alpha|\beta\rangle \in \mathbb{R}$, bo przestrzeń V zbudowana jest nad ciałem liczb rzeczywistych.

Pokażemy jeszcze, że zgodnie z oczekiwaniami $|\cos\phi| \leq 1$. Wynika to z nierówności Schwarza.

Nierówność Schwarza:

Niech $\alpha, \beta \in V$ nad $\mathbb{R}$ lub nad $\mathbb{C}$. Zachodzi

$$|\langle\alpha|\beta\rangle| \leq \|\alpha\|\,\|\beta\|.$$

Dowód: Gdy $\langle\alpha|\beta\rangle = 0$ nierówność jest oczywista. Niech teraz $\langle\alpha|\beta\rangle \neq 0$. Zauważmy, że

$$0 \leq \|\alpha - \frac{\|\alpha\|^2}{\langle\alpha|\beta\rangle}\beta\|^2 = \langle\alpha - \frac{\|\alpha\|^2}{\langle\alpha|\beta\rangle}\beta|\alpha - \frac{\|\alpha\|^2}{\langle\alpha|\beta\rangle}\beta\rangle$$
$$= \|\alpha\|^2 - \frac{\|\alpha\|^2}{\langle\alpha|\beta\rangle^*}\langle\beta|\alpha\rangle - \frac{\|\alpha\|^2}{\langle\alpha|\beta\rangle}\langle\alpha|\beta\rangle + \frac{\|\alpha\|^4\|\beta\|^2}{|\langle\alpha|\beta\rangle|^2}$$
$$= -\|\alpha\|^2 + \frac{\|\alpha\|^4\|\beta\|^2}{|\langle\alpha|\beta\rangle|^2} = \frac{\|\alpha\|^2}{|\langle\alpha|\beta\rangle|^2}(\|\alpha\|^2\|\beta\|^2 - |\langle\alpha|\beta\rangle|^2)$$

Ponieważ $\|\alpha\|^2/|\langle\alpha|\beta\rangle|^2 > 0$ więc nawias w powyższym równaniu musi być dodatni, co kończy dowód. cbdo

---

Przykład:

Niech V będzie przestrzenią unitarną nad $\mathbb{R}$. Obliczyć kwadrat normy sumy wektorów $\alpha + \beta$.
Rozwiązanie: Obliczamy korzystając z własności iloczynu skalarnego

$$\|\alpha + \beta\|^2 = \langle\alpha + \beta|\alpha + \beta\rangle = \langle\alpha|\alpha\rangle + 2\langle\alpha|\beta\rangle + \langle\beta|\beta\rangle = \|\alpha\|^2 + 2\langle\alpha|\beta\rangle + \|\beta\|^2$$

Wstawiając do tej równości $\langle\alpha|\beta\rangle = \|\alpha\|\,\|\beta\|\cos\phi$ otrzymujemy wzór Carnota (twierdzenie cosinusów)

$$\|\alpha + \beta\|^2 = \|\alpha\|^2 + 2\|\alpha\|\,\|\beta\|\cos\phi + \|\beta\|^2.$$

Aby wprowadzić pojęcie kąta, musieliśmy ograniczyć się do przestrzeni unitarnej nad ciałem liczb rzeczywistych. W dowolnej przestrzeni unitarnej możemy wprowadzić jedynie pojęcie ortogonalności (prostopadłości wektorów).

Definicja ortogonalności:

Mówimy, że wektory α, β należące do przestrzeni unitarnej są ortogonalne (prostopadłe do siebie), gdy $\langle\alpha|\beta\rangle = 0$. Piszemy wówczas $\alpha\perp\beta$ lub $\beta\perp\alpha$.

Twierdzenie (o rozkładzie wektora):

Niech V będzie przestrzenią unitarną, a α i β jej wektorami, przy czym $\beta \neq 0$. Wektor

$$\alpha_\parallel = \frac{\langle\beta|\alpha\rangle}{\|\beta\|^2}\beta = \langle\frac{\beta}{\|\beta\|}|\alpha\rangle\frac{\beta}{\|\beta\|}$$

jest jedynym wektorem posiadającym następujące własności:

1. $\alpha_\parallel$ należy do podprzestrzeni rozpiętej przez β, tj. $\alpha_\parallel = a\beta$, $a \in \mathbb{K}$,
2. Różnica $\alpha - \alpha_\parallel$ jest wektorem ortogonalnym do β.

Innymi słowy, α można zapisać jednoznacznie jako sumę

$$\alpha = \alpha_\parallel + (\alpha - \alpha_\parallel) = \alpha_\parallel + \alpha_\perp$$

dwóch składników: równoległego i prostopadłego do β.

Dowód: To, że $\alpha_\parallel = a\beta$ wynika natychmiast z definicji $\alpha_\parallel$. Wiemy, że $a = \langle\alpha|\beta\rangle/\|\beta\|^2$. Prostopadłość $\alpha_\perp$ do β wykazujemy bezpośrednim rachunkiem:

$$\langle\alpha - \alpha_\parallel|\beta\rangle = \langle\alpha|\beta\rangle - \langle\alpha_\parallel|\beta\rangle = \langle\alpha|\beta\rangle - \frac{\langle\alpha|\beta\rangle}{\|\beta\|^2}\langle\beta|\beta\rangle = \langle\alpha|\beta\rangle - \langle\alpha|\beta\rangle = 0$$

Pozostaje nam wykazać jednoznaczność rozkładu. Niech będzie wektor γ proporcjonalny do β, tj. $\gamma = b\beta$ o tej własności, że $(\alpha-\gamma)\perp\beta$. Pokażemy, że $\gamma = \alpha_{\|}$, tzn. że $b = \langle\alpha|\beta\rangle/\|\beta\|^2$. W tym celu zbadajmy $\langle\alpha-\gamma|\beta\rangle$, co z założenia jest równe 0

$$0 = \langle\alpha-\gamma|\beta\rangle = \langle\alpha|\beta\rangle - \langle\gamma|\beta\rangle = \langle\alpha|\beta\rangle - b\langle\beta|\beta\rangle = \langle\alpha|\beta\rangle - b\|\beta\|^2.$$

Wynika stąd, że rzeczywiście

$$b = \frac{\langle\alpha|\beta\rangle}{\|\beta\|^2}.$$

cbdo

Wniosek: Gdy V jest przestrzenią wektorową nad $\mathbb{R}$, to

$$\alpha_{\|} = \frac{\langle\alpha|\beta\rangle}{\|\beta\|^2}\beta = \frac{\|\alpha\|\,\|\beta\|\cos\phi}{\|\beta\|^2}\beta = \|\alpha\|\cos\phi\frac{\beta}{\|\beta\|},$$

gdzie $\cos\phi = \cos(\angle(\alpha,\beta))$.

Twierdzenie:

Ciąg $w_1, w_2, \ldots, w_n$ niezerowych wektorów wzajemnie ortogonalnych jest liniowo niezależny.

Dowód: Budujemy kombinację liniową

$$\gamma = \sum_{k=1}^{n} a_k\, w_k, \qquad a_1, a_2, \ldots, a_n \in \mathbb{K}.$$

Mamy pokazać, że ze znikania tej kombinacji ($\gamma = 0$) wynika znikanie wszystkich jej współczynników $a_1, a_2, \ldots, a_n$. W tym celu zbadajmy $\langle\gamma|w_s\rangle = 0$, $s = 1, 2, \ldots, n$.

$$0 = \langle\gamma|w_s\rangle = \sum_{k=1}^{n} a_k\langle w_k|w_s\rangle = \sum_{k=1}^{n} a_k\delta_{ks}\|w_s\|^2 = a_s\|w_s\|^2.$$

Ponieważ wektory w_j są niezerowe, więc $\|w_s\| > 0$. Wynika stąd, że $a_s = 0$, $s = 1, 2, \ldots, n$, a więc wszystkie współczynniki kombinacji liniowej γ znikają. Oznacza to, że wektory $w_1, w_2, \ldots, w_n$ tworzą układ liniowo niezależny. cbdo

---

Definicja bazy ortogonalnej (ortonormalnej):

Bazę $e_1, e_2, \ldots, e_n$ w przestrzeni unitarnej V nazywamy bazą ortogonalną (ortonormalną, tj. ortogonalną i unormowaną), gdy jej wektory są wzajemnie ortogonalne (ortonormalne).

Wniosek: Baza ortonormalna ma tę własność, że $\langle e_i|e_j\rangle = \delta_{ij}$.

Przykłady:

1. Rozważana przez nas baza $\hat{\mathbf{i}}$, $\hat{\mathbf{j}}$, $\hat{\mathbf{k}}$ w $\mathbb{R}^3$ jest bazą ortonormalną.

2. Niech V będzie przestrzenią nad $\mathbb{R}$ złożoną z wielomianów rzeczywistych stopnia ≤ 1 określonych na $[-1,1] \in \mathbb{R}$. Dodawanie wektorów (wielomianów) i mnożenie ich przez liczbę określone są w sposób naturalny. Iloczyn skalarny wielomianów $p(x)$ i $q(x)$ określone jest wzorem

$$\langle p|q\rangle = \int_{-1}^{1} p(x)q(x)\mathrm{d}x.$$

Pokażemy, że wielomiany $w_1 = 1$, $w_2 = x$ tworzą bazę ortogonalną w V, co oznacza, że dim$V$=2.

Widać, że $V = \{1, x\}$, tj. wspomniane wielomiany rozpinają przestrzeń V. Jeśli pokażemy, że są one ortogonalne, to z ostatniego twierdzenia wynikać będzie, że sa liniowo niezależne. To wszystko w sumie doprowadzi nas do wniosku, że jest to baza ortogonalna w V. W tym celu obliczmy iloczyn skalarny

$$\langle w_1|w_2\rangle = \int_{-1}^{1} 1 \cdot x\mathrm{d}x = \frac{1}{2}x^2\Big|_{-1}^{1} = 0.$$

Wynika stąd, że $w_1 \perp w_2$ w sensie iloczynu skalarnego. Baza ta nie jest ortonormalna, bo $\|w_1\| \neq 1$ i $\|w_2\| \neq 1$. Obliczmy

$$\|w_1\|^2 = \langle w_1|w_1\rangle = \int_{-1}^{1} 1 \cdot 1\mathrm{d}x = x|_{-1}^{1} = 2$$
$$\|w_2\|^2 = \langle w_2|w_2\rangle = \int_{-1}^{1} \cdot x\mathrm{d}x = \frac{1}{3}x^3\Big|_{-1}^{1} = \frac{2}{3}.$$

Przykładem bazy ortonormalnej w tej przestrzeni może być baza złożona z

$$e_1 = \frac{w_1}{\|w_1\|} = \frac{1}{\sqrt{2}}, \qquad e_2 = \frac{w_2}{\|w_2\|} = \sqrt{\frac{3}{2}}\, x.$$

Istnienie bazy ortonormalnej w (skończenie wymiarowej) przestrzeni unitarnej nie jest rzeczą niezwykłą. W istocie, za pomocą procedury ortonormalizacyjnej Schmidta, z każdej bazy można utworzyć bazę ortonormalną. Bazę ortonormalną będziemy oznaczać przez $e_1, e_2, \ldots, e_n$ i przypomnijmy, że baza ortonormalna to taka, że $\langle e_i|e_j\rangle = \delta_{ij}$.

Procedura Schmidta:

Konstrukcja przebiega w następujący sposób: Zaczynamy od dowolnej bazy $\{f_i : \quad i = 1, 2, \ldots, n\}$. Definiujemy rekurencyjnie

$$e_1 = \frac{f_1}{\|f_1\|}$$
$$e_2 = \frac{f_2 - \langle e_1|f_2\rangle e_1}{\|f_2 - \langle e_1|f_2\rangle e_1\|}$$

Oczywiście, z konstrukcji $\|e_1\| = \|e_2\| = 1$. Łatwo również przekonać się, że e_1 i e_2 są ortogonalne

$$\langle e_1|e_2\rangle = \frac{\langle e_1|f_2\rangle - \langle e_1|f_2\rangle}{\|f_2 - \langle e_1|f_2\rangle e_1\|} = 0$$

Zakładamy teraz, zgodnie z procedurą rekurencyjną, że skonstruowaliśmy już wektory e_i od $i = 1$ do pewnego k. $k+1$ wektor bazy ortonormalnej definiujemy jako

$$e_{k+1} = \frac{f_{k+1} - \langle e_1|f_{k+1}\rangle e_1 - \ldots - \langle e_k|f_{k+1}\rangle e_k}{\|f_{k+1} - \langle e_1|f_{k+1}\rangle e_1 - \ldots - \langle e_k|f_{k+1}\rangle e_k\|}$$

Łatwo możemy wykazać, że e_{k+1} jest ortogonalny do wektorów e_j, $j = 1, 2, \ldots, k$

$$\langle e_j|e_{k+1}\rangle = \frac{\langle e_j|f_{k+1}\rangle - \langle e_j|f_{k+1}\rangle}{\|f_{k+1} - \langle e_1|f_{k+1}\rangle e_1 - \ldots - \langle e_k|f_{k+1}\rangle e_k\|} = 0$$

W ten sposób postępując konstruujemy pełną bazę ortonormalną.

Wniosek: Przestrzeń unitarna skończonego wymiaru posiada bazę ortonormalną. cbdo

---

W dalszym ciągu udowodnimy parę stwierdzeń, które pokażą, że nasze intuicje wyniesione z rozważań nad przestrzeniami $\mathbb{K}^n$ ($\mathbb{R}^n$ lub $\mathbb{C}^n$) przenoszą się na dowolne przestrzenie unitarne.

Stwierdzenie 1:

Niech V będzie przestrzenią unitarną, α dowolnym wektorem i $\{e_i\}$ bazą ortonormalną. Wówczas wektor α daje się przedstawić jako

$$\alpha = \sum_{i=1}^{n} \langle e_i|\alpha\rangle e_i$$

Dowód: Wiadomo, że każdy wektor można zapisać w bazie $\alpha = \sum_{i=1}^{n} a_i e_i$, gdzie a_i są współrzędnymi tego wektora. Obliczając iloczyn skalarny α z e_i dostajemy $a_i = \langle e_i|\alpha\rangle$. cbdo

---

Stwierdzenie 2:

Niech $\{e_i\}$ będzie bazą ortonormalną w V. Iloczyn skalarny dwóch wektorów $\alpha = \sum_{i=1}^{n} a_i e_i$ i $\beta = \sum_{i=1}^{n} b_i e_i$ dany jest wzorem:

$$\langle \alpha|\beta\rangle = \sum_{i=1}^{n} a_i^* b_i$$

natomiast norma wektora

$$\|\alpha\| = \sqrt{\sum_{i=1}^{n} |a_i|^2}$$

Dowód: Korzystając z $\langle \alpha|b\beta\rangle = b\langle \alpha|\beta\rangle$ oraz z liniowości iloczynu skalarnego i ortonormalności bazy mamy:

$$\langle \alpha|\beta\rangle = \langle \sum_{i=1}^{n} a_i e_i| \sum_{i=1}^{n} b_i e_i\rangle = \sum_{i=1}^{n} \langle a e_i|b e_i\rangle = \sum_{i=1}^{n} a_i^* b_i.$$

cbdo

Definicja odległości między dwoma wektorami:
W przestrzeni unitarnej V odległość d między wektorami definiujemy następująco:

$$d(\alpha,\beta) = \|\alpha - \beta\|, \qquad \alpha,\beta \in V.$$

Definicja ta spełnia wszystkie warunki wynikające z aksjomatycznej definicji odległości podanej w podrozdziale pt. *Przestrzenie Metryczne* w podrozdziale 1.8. Niektóre z tych warunków wynikają bezpośrednio z definicji normy. Pozostaje jeszcze pokazać słuszność nierówności trójkąta

$$d(\alpha,\beta) \leq d(\alpha,\gamma) + d(\gamma,\beta)$$

dla tak zdefiniowanej odległości, czyli

$$\|\alpha - \beta\| \leq \|\alpha - \gamma\| + \|\gamma - \beta\|.$$

Ta ostatnia nierówność wynika z nierówności Minkowskiego

$$\|\alpha + \beta\| \leq \|\alpha\| + \|\beta\|,$$

w której zamiast β wstawiamy $-\beta$ i piszemy

$$\|\alpha - \beta\| = \|\alpha - \gamma + \gamma - \beta\| \leq \|\alpha - \gamma\| + \|\gamma - \beta\|$$

czyli

$$d(\alpha,\beta) \leq d(\alpha,\gamma) + d(\gamma,\beta).$$

Oczywiście γ jest dowolnie wybranym wektorem. Pozostaje nam jeszcze do udowodnienia **nierówność Minkowskiego**. Dowód przeprowadzamy następująco:

$$\begin{aligned}
&\|\alpha+\beta\|^2 = \langle\alpha+\beta|\alpha+\beta\rangle = \|\alpha\|^2 + \langle\alpha|\beta\rangle + \langle\beta|\alpha\rangle + \|\beta\|^2 \\
&= \|\alpha\|^2 + 2\mathrm{Re}\langle\alpha|\beta\rangle + \|\beta\|^2 \leq \\
&\|\alpha\|^2 + 2|\langle\alpha|\beta\rangle| + \|\beta\|^2 \leq \|\alpha\|^2 + 2\|\alpha\|\,\|\beta\| + \|\beta\|^2 = (\|\alpha\| + \|\beta\|)^2.
\end{aligned}$$

1.7 Przestrzenie unormowane

To, że w przestrzeni unitarnej możemy wprowadzić normę (długość wektora) wydaje się naturalne (czy też wręcz trywialne). Ale czy na prawdę pojęcie normy i odległości wiąże się wyłącznie z przestrzeniami unitarnymi?

Jak już wspomnieliśmy, normę można wprowadzić bez pojęcia iloczynu skalarnego definiując odwzorowanie przyporządkowujące wektorowi α liczbę nieujemną oznaczaną przez $\|\alpha\|$:

$$V \ni \alpha \longrightarrow \|\alpha\| \in \mathbb{R}_+$$

spełniające następujące warunki:

a. $\|\alpha\| = 0 \Leftrightarrow \alpha = 0$
b. $\|a\alpha\| = |a|\,\|\alpha\|, \quad a \in \mathbb{K}$
c. $\|\alpha+\beta\| \leq \|\alpha\| + \|\beta\|$ nierówność Minkowskiego.

W ten sposób otrzymujemy przestrzeń unormowaną $(V, \| \ \|)$. Widać, że badana przez nas wcześniej norma wektora w przestrzeni unitarnej, zdefiniowanej przez iloczyn skalarny, spełnia powyższe warunki. Oznacza to, że przestrzeń unitarna jest unormowana.

W przestrzeni unormowanej można z kolei wprowadzić pojęcie metryki (odległości):

$$d(\alpha, \beta) = \|\alpha - \beta\|$$

Z powyższej definicji wynikają następujące własności (oczekiwane w naturalny sposób od odległości):

1^o $d(\alpha, \beta) = d(\beta, \alpha)$
gdyż $d(\beta, \alpha) = \|\beta - \alpha\| = \|(-1)(\alpha - \beta)\| = \|\alpha - \beta\|$

2^o $d(\alpha, \beta) \leq d(\alpha, \gamma) + d(\gamma, \beta)$
gdyż $\|\alpha - \beta\| = \|(\alpha - \gamma) + (\gamma - \beta)\| \leq \|\alpha - \gamma\| + \|\gamma - \beta\|$

gdzie skorzystaliśmy z własności $c.$ definicji normy.

1.8 Przestrzenie metryczne

Matematycy wprowadzają pojęcie przestrzeni metrycznej nawet, jeśli V nie jest przestrzenią liniową. Wystarczy, że istnieje przekształcenie spełniające cechy oczekiwane od metryki.

Definicja metryki:
Metryką nazywamy odwzorowanie

$$\rho : \ V \times V \longrightarrow \mathbb{R}$$

spełniające następujące warunki:

1^o $\rho(\alpha, \beta) \geq 0, \qquad \rho(\alpha, \beta) = 0 \ \Leftrightarrow \alpha = \beta$

2^o $\rho(\alpha, \beta) = \rho(\beta, \alpha)$

3^o $\rho(\alpha, \beta) \leq \rho(\alpha, \gamma) + \rho(\gamma, \beta)$

Przykłady:

Przykład 1:
$V = \mathbb{R}^n$ z metryką ρ zdefiniowaną w następujący sposób

$$\rho(\alpha, \beta) = |a_1 - b_1| + |a_2 - b_2| + \ldots + |a_n - b_n|,$$

gdzie $\alpha = (a_1, a_2, \ldots, a_n)$, $\beta = (b_1, b_2, \ldots, b_n)$. Pierwsze dwa aksjomaty metryki spełnione są automatycznie. Trzeci jest konsekwencją nierówności Schwarza zastosowaną do modułu w przestrzeni $\mathbb{R}^1$ (jeśli zauważymy, że każda ze współrzędnych spełnia tę nierówność to oczywiście suma po współrzędnych także).
Metryka ta nazywana jest metryką miejską.

Przykład 2:

$V = \mathbb{R}^n$ z metryką $\tilde{\rho}$ zdefiniowaną w następujący sposób

$$\tilde{\rho}(\alpha,\beta) = \begin{cases} \rho(\alpha,\beta) & \text{gdy punkty } \alpha,\ \beta,\ 0 \text{ leżą na jednej prostej} \\ \rho(\alpha,0)+\rho(0,\beta) & \text{w przeciwnym wypadku,} \end{cases}$$

gdzie ρ jest "naturalną" metryką. Metryka $\tilde{\rho}$ nazywana jest metryką węzła kolejowego.

Przykład:

Dowolna przestrzeń V z metryką zdefiniowaną następująco

$$\rho(\alpha,\beta) = \begin{cases} 1 & \text{gdy} \quad \alpha \neq \beta \\ 0 & \text{gdy} \quad \alpha = \beta \end{cases}$$

Taką metrykę nazywamy dyskretną.

Każda unormowana przestrzeń wektorowa jest jednocześnie przestrzenią metryczną. Przypomnijmy, że pojęcie metryki jest niezbędne do zdefiniowania w abstrakcyjnych przestrzeniach takich pojęć, jak otoczenie punktu, zbiory otwarte i domknięte, sfera, kula, sześcian itp. Jest też niezbędnym narzędziem do badania zbieżności szeregów.

Przestrzeń Banacha:

Jeżeli wszystkie ciągi zbieżne złożone z elementów danej przestrzeni mają granicę w tej przestrzeni, to mówimy, że przestrzeń jest zupełna. Unormowana i zupełna przestrzeń wektorowa nazywa się przestrzenią Banacha .
