Twierdzenie Kroneckera-Capellego i podprzestrzenie w K^n

Ostatnia aktualizacja: 26.11.2021 r.

Celem dzisiejszego wykładu jest zastosowanie pojęć wprowadzonych ostatnio do opisu podprzestrzeni przestrzeni liniowej K^n. Idea jest następująca: dowolne podprzestrzenie w K^n można opisywać jako przestrzenie rozpięte przez pewien układ, czyli lin(...) – co jest prawdą nie tylko dla K^n – lub jako zbiory rozwiązań jednorodnych układów równań – co jest charakterystyczne właśnie dla tej przestrzeni, i co wymaga jeszcze dowodu (owszem – zbiór rozwiązań jest podprzestrzenią, ale nie wiemy jeszcze czy każda podprzestrzeń to zbiór rozwiązań). Obydwa rodzaje opisów podprzestrzeni w K^n opierają sie na pojęciu macierzy. Bazę przestrzeni rozpiętej przez układ wektorów można znaleźć przez „schodkowanie" macierzy, której wiersze stanową kolejne wektory z tego układu. Wymiar owej podprzestrzeni – to oczywiście liczba niezerowych wierszy uzyskanych po schodkowaniu. A jak jest z rozwiązaniami jednorodnych układów równań? Czy i tutaj istnieją niezmienniki wskazujące na związek między własnościami macierzy ich współczynników, a bazą i wymiarem przestrzeni rozwiązań?

Sformułujmy główne twierdzenie tego wykładu.

Twierdzenie 1 (Kroneckera-Capellego)**.** *Niech U będzie układem równań liniowych o współczynnikach w ciele K postaci:*

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

o macierzy współczynników A oraz rozszerzonej macierzy współczynników A_u. Wówczas:

(a) Układ U ma rozwiązanie wtedy i tylko wtedy, gdy $r(A) = r(A_u)$,

(b) Przestrzeń rozwiązań układu jednorodnego odpowiadającego układowi U ma wymiar $n - r(A)$

(c) Jeśli α jest rozwiązaniem układu U, a W jest przestrzenią rozwiązań układu jednorodnego odpowiadającego układowi U, to zbiór rozwiązań układu U jest postaci

$$\alpha + W = \{\alpha + \beta, \mid \beta \in W\}.$$

Przypomnijmy pojęcia użyte w twierdzeniu. Macierze A oraz A_u układu wyżej to odpowiednio:

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

Układ jednorodny odpowiadający układowi U to układ jednorodny o macierzy A.

Zauważmy, że od początku wykładu gromadzimy elementy rozumowań potrzebne do dowodu tego twierdzenia. Punkt (c) wykazaliśmy już na pierwszym wykładzie. Punkty (a) i (b) opisują problem rozwiązywalności i „rozmiaru" zbioru rozwiązań układu równań liniowych w języku wymiaru. Również te fakty są dla nas w zasadzie intuicyjnie jasne. Wiemy bowiem, że układ równań może okazać się sprzeczny jedynie, gdy w wyniku sprowadzania macierzy A_u do postaci zredukowanej pojawi się wiersz postaci $[00\ldots0|1]$. Nietrudno będzie nam formalnie pokazać, na podstawie posiadanej już wiedzy, że sytuacja ta może wystąpić jedynie, gdy $r(A) < r(A_u)$.

Również punkt (b) jest intuicyjnie jasny. Nie pokazaliśmy jeszcze tego w sposób formalny, ale na podstawie wielu przykładów podejrzewamy, że wymiar przestrzeni rozwiązań jednorodnego układu równań równy jest liczbie zmiennych niezależnych tego układu. Ta zaś równa jest liczbie wszystkich zmiennych pomniejszonej o liczbę zmiennych zależnych. Wszystkich zmiennych jest n, zaś zmiennych zależnych jest tyle, co schodków macierzy A po sprowadzeniu jej, za pomocą elementarnych operacji wierszowych, do postaci zredukowanej. Tych schodków jest, jak wiemy, $r(A)$. W dowodzie opiszemy konstrukcję bazy zbioru rozwiązań układu U. Wcześniej jednak zobaczmy ważny przykład.

Ważny przykład. Niech $(a_1, \ldots, a_n) \in K^n$, gdzie $a_1 \neq 0$. Wówczas zbiór rozwiązań równania

$$a_1 x_1 + \ldots + a_n x_n = 0$$

jest podprzestrzenią wymiaru $n-1$ postaci:

$$\operatorname{lin}((-\frac{a_2}{a_1}, 1, 0, \ldots, 0), \ldots, (-\frac{a_n}{a_1}, 0, 0, \ldots, 1)).$$

Rzeczywiście, sprowadzamy macierz układu do postaci zredukowanej:

$$[a_1 \quad a_2 \quad \ldots \quad a_n] \sim \left[1 \quad \frac{a_2}{a_1} \quad \ldots \quad \frac{a_n}{a_1}\right].$$

A zatem rozwiązanie ogólne tego układu zadane jest przez:

$$x_1 = -\frac{a_2}{a_1} x_2 - \frac{a_3}{a_1} x_3 - \ldots - \frac{a_n}{a_1} x_n.$$

Zmienne $x_2, \ldots, x_n$ stanowią $n-1$ parametrów i wszystkie rozwiązania tego układu są postaci:

$$(-\frac{a_2}{a_1} t_2 - \ldots - \frac{a_n}{a_1} t_n, \underbrace{t_2, t_3, \ldots t_n}_{\text{parametry}}), \qquad \text{gdzie } t_2, t_3, \ldots, t_n \in K.$$

W szczególności rozwiązania te są postaci:

$$t_2(-\frac{a_2}{a_1}, 1, 0, \ldots, 0) + t_3(-\frac{a_3}{a_1}, 0, 1, \ldots, 0) + \ldots + t_n(-\frac{a_n}{a_1}, 0, 0, \ldots, 1).$$

Układ wektorów rozpinających zbiór rozwiązań powyższego równania powstał w następujący sposób:

- wektor $(-\frac{a_2}{a_1}, 1, 0, \ldots, 0)$ przez przyjęcie $x_2 = 1$ oraz $x_3 = \ldots = x_n = 0$,
- wektor $(-\frac{a_3}{a_1}, 0, 1, \ldots, 0)$ przez przyjęcie $x_2 = 0$, $x_3 = 1$ oraz $x_4 = \ldots = x_n$,
- ...
- wektor $(-\frac{a_n}{a_1}, 0, 0, \ldots, 1)$ powstał przez przyjęcie $x_2 = \ldots = x_{n-1} = 0$ oraz $x_n = 0$.

Jak się za chwilę okaże, dla każdego układu równań jednorodnych wybór rozwiązań oparty na wstawianiu za jeden z parametrów jedynki, a za pozostałe – zera, prowadzi do uzyskania bazy zbioru rozwiązań.

Dowodzimy twierdzenie Kroneckera-Capellego. Punkt (a) jest jasny. Weźmy $\alpha_1, \ldots, \alpha_n, \beta \in K^n$, które są kolumnami macierzy A_u. Wówczas mamy ciąg równoważnych stwierdzeń:

- $x_1, \ldots, x_n$ jest rozwiązaniem układu U,
- $x_1\alpha_1 + \ldots + x_n\alpha_n = \beta$,
- $\beta \in \operatorname{lin}(\alpha_1, \ldots, \alpha_n)$,
- $\operatorname{lin}(\alpha_1, \ldots, \alpha_n) = \operatorname{lin}(\alpha_1, \ldots, \alpha_n, \beta)$,
- $\dim \operatorname{lin}(\alpha_1, \ldots, \alpha_n) = \dim \operatorname{lin}(\alpha_1, \ldots, \alpha_n, \beta)$,
- $r(A) = r(A_u)$.

Niech U' będzie układem jednorodnym odpowiadającym układowi U. Dowód punktu (b) polega na zauważeniu, że jeśli $r(A) = r$, to macierz A' uzyskana z A przez sprowadzenie do postaci schodkowej ma dokładnie r niezerowych wierszy, a zatem postać ogólna rozwiązania tego układu ma $n-r$ parametrów. Załóżmy, że zmienne zależne to $x_{j_1}, \ldots, x_{j_r}$, a owe parametry to zmienne $x_{t_1}, \ldots, x_{t_{n-r}}$. Innymi słowy, rozwiązanie ogólne układu U' ma, dla pewnych $c_{ij} \in K$, postać:

$$\begin{cases} x_{j_1} & = c_{11}x_{t_1} + c_{12}x_{t_2} + \ldots + c_{1,n-r}x_{t_{n-r}} \\ & \vdots \\ x_{j_r} & = c_{r1}x_{t_1} + c_{r2}x_{t_2} + \ldots + c_{r,n-r}x_{t_{n-r}} \end{cases} \qquad (*)$$

Rozważmy układ $n-r$ wektorów $\alpha_1, \ldots, \alpha_{n-r}$ takich, że α_j jest rozwiązaniem powyższego układu powstałym przez wstawienie za j-ty parametr 1, a za pozostałe parametry – zera. Innymi słowy, jeśli $\alpha_j = (a_{j1}, \ldots, a_{jn})$, to

$$a_{jt_1} = 0, \quad \ldots \quad , a_{jt_j} = 1, \quad \ldots \quad , a_{jt_{n-k}} = 0.$$

Oczywiście wektory $\alpha_1, \ldots, \alpha_{n-r}$ istnieją, bo układ $(*)$ ma jednoznaczne rozwiązane dla każdego z góry zadanego układu parametrów. Twierdzimy, że układ ten jest bazą zbioru rozwiązań układu U'.

Przykład. Rozwiązanie ogólne układu U' zmiennych x_1, x_2, x_3, x_4 o współczynnikach w $\mathbb{R}$ postaci:

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

ma $4-2$ parametry x_3, x_4 (bo rząd macierzy U' to 2). Każde rozwiązanie U' jest postaci:

$$(-s-t, 0, s, t) = s(-1, 0, 1, 0) + t(-1, 0, 0, 1), \quad s, t \in \mathbb{R},$$

i powstaje przed odpowiedni wybór s, t na współrzędnych $t_1 = 3, t_2 = 4$. W zatem w tym przykładzie:

$$\alpha_1 = (a_{11}, a_{12}, a_{13}, a_{14}) = (-1, 0, 1, 0), \quad \alpha_2 = (a_{21}, a_{22}, a_{23}, a_{24}) = (-1, 0, 0, 1).$$

Dowodzimy, że układ $\alpha_1, \ldots, \alpha_{n-r}$ jest liniowo niezależny. Istotnie, jeśli dla $b_1, \ldots, b_{n-r} \in K$ mamy:

$$b_1\alpha_1 + \ldots + b_{n-r}\alpha_{n-r} = 0,$$

to dla i-tych współrzędnych $a_{1i}, \ldots, a_{n-r,i}$ wektorów $\alpha_1, \ldots, \alpha_{n-r}$ mamy:

$$b_1 a_{1i} + \ldots + b_{n-r} a_{n-r,i} = 0,$$

Zatem biorąc i równe kolejno $t_1, \ldots t_{n-r}$ dostajemy $b_1 = \ldots = b_{n-r} = 0$.

Układ $\alpha_1, \ldots, \alpha_{n-r}$ rozpina zbiór rozwiązań układu $(*)$. Jest bowiem jasne, że wektor $\alpha = (a_1, \ldots, a_n)$ spełnia układ $(*)$ wtedy i tylko wtedy, gdy każda z jego współrzędnych jest kombinacją liniową współrzędnych $a_{t_1}, \ldots, a_{t_n}$. Inaczej mówiąc α jest rozwiązaniem układu U' wtedy i tylko wtedy, gdy

$$\alpha = a_{t_1}\alpha_1 + \ldots + a_{t_{n-r}}\alpha_{n-r}.$$

Zatem zbiór rozwiązań układu U' równy jest $\operatorname{lin}(\alpha_1, \ldots, \alpha_{n-r})$, co kończy dowód (b) i całego twierdzenia.

Definicja 1. *Jeśli $V \subseteq K^n$ jest przestrzenią rozwiązań jednorodnego układu równań liniowych U, to mówimy, że przestrzeń V jest* ***opisana układem*** *U.*

Wniosek 1. *Każda podprzestrzeń V przestrzeni K^n jest opisana pewnym jednorodnym układem równań liniowych U. Jeśli $\dim V = k$, to można tak dobrać ten układ U, by składał się z $n-k$ równań. Dla $\dim V = k$ oraz $i < n-k$ nie istnieje złożony z i równań układ równań liniowych opisujący V.*

Dowód tego twierdzenia zawiera w sobie istotny algorytm opisu podprzestrzeni za pomocą układu równań. Ponownie przeprowadzimy jego ilustrację najpierw na przestrzeni rozpiętej przez pojedynczy wektor.

Ważny przykład. Niech $\alpha = (a_1, \ldots, a_n) \in K^n$, gdzie $a_1 \neq 0$. Wówczas przestrzeń $\operatorname{lin}(\alpha)$ opisana jest przez układ równań postaci:

$$\begin{cases} -\frac{a_2}{a_1}x_1 + x_2 &= 0 \\ -\frac{a_3}{a_1}x_1 + x_3 &= 0 \\ & \vdots \\ -\frac{a_n}{a_1}x_1 + x_n &= 0 \end{cases}.$$

Czytelnik dostrzega z pewnością związek pomiędzy wypisanymi równaniami, a wektorami uzyskanymi wcześniej jako rozwiązania równania $a_1x_1 + \ldots + a_nx_n = 0$. Ta dualność nie jest przypadkowa. W istocie, rozwiązać należy ten sam układ, przy czym tym razem szukamy takich n-tek współczynników $(t_1, \ldots, t_n)$, aby równanie liniowe $t_1x_1 + \ldots + t_nx_n = 0$ miało z góry zadane rozwiązanie $a_1, \ldots, a_n$. Zbiór wszystkich takich $(t_1, \ldots, t_n)$ jest podprzestrzenią w K^n, której bazą są wektory $(-\frac{a_2}{a_1}, 1, 0, \ldots, 0), \ldots, (-\frac{a_n}{a_1}, 0, 0, \ldots, 1)$.

Analogiczna sytuacja jest dla układów równań. Jeśli $(s_{11},\dots,s_{1n}),\dots,(s_{m1},\dots,s_{mn})$ są rozwiązaniami układu

$$\begin{cases} a_{11}x_1+\ldots+a_{1n}x_n &= 0\\ \cdots & \\ a_{k1}x_1+\ldots+a_{kn}x_n &= 0 \end{cases},$$

to $(a_{11},\dots,a_{1n}),\dots,(a_{k1},\dots,a_{kn})$ są rozwiązaniami układu:

$$\begin{cases} s_{11}x_1+\ldots+s_{1n}x_n &= 0\\ \cdots & \\ s_{m1}x_1+\ldots+s_{mn}x_n &= 0 \end{cases}.$$

Jeśli $(s_{11},\dots,s_{1n}),\dots,(s_{m1},\dots,s_{mn})$ jest bazą przestrzeni V rozwiązań układu o macierzy schodkowej

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n}\\ \vdots & \vdots & \ddots & \vdots\\ a_{k1} & a_{k2} & \dots & a_{kn} \end{bmatrix},$$

to na mocy tw. Kroneckera-Capellego $m=n-k$, co więcej wektory $(a_{11},\dots,a_{1n}),\dots,(a_{k1},\dots,a_{kn})$ są liniowo niezależne i są rozwiązaniami układu o macierzy rzędu m postaci

$$\begin{bmatrix} s_{11} & s_{12} & \dots & s_{1n}\\ \vdots & \vdots & \ddots & \vdots\\ s_{m1} & s_{m2} & \dots & s_{mn} \end{bmatrix}.$$

Ponownie na mocy tw. Kroneckera-Capellego, powyższy układ ma przestrzeń rozwiązań wymiaru $n-m=n-(n-k)=k$. Czyli jest to $\operatorname{lin}((a_{11},\dots,a_{1n}),\dots,(a_{k1},\dots,a_{kn}))$ – przestrzeń wymiaru k.

Wykazaliśmy zatem, że jeśli $\alpha_1,\dots,\alpha_k$ jest bazą przestrzeni $V\subseteq K^n$ oraz $A\in M_{k\times n}(K)$ jest macierzą o wierszach $\alpha_1,\dots,\alpha_k$, to przestrzeń V można opisać układem dowolnych $n-k$ równań, których współczynniki tworzą bazę przestrzeni rozwiązań układu danego macierzą A. Równań tych nie może być oczywiście mniej, bowiem przestrzeń rozwiązań układu o mniej niż $n-k$ równaniach ma wymiar większy niż $n-(n-k)$, na mocy twierdzenia Kroneckera-Capellego.

Przykład. Rozważmy $V=\operatorname{lin}((1,2,0,1,0),(0,0,1,1,1))\subseteq\mathbb{R}^5$. Rozwiązania układu równań o macierzy

$$\begin{bmatrix} 1 & 2 & 0 & 1 & 0\\ 0 & 0 & 1 & 1 & 1 \end{bmatrix}$$

stanowi przestrzeń współczynników wszystkich równań liniowych, których rozwiązania zawierają V. Wybierając różne bazy tej przestrzeni dostajemy różne (ale równoważne) układy równań opisujące V, na przykład dla bazy $(-2,1,0,0,0),(-1,0,-1,1,0),(0,0,-1,0,1)$ mamy następujący układ opisujący V:

$$\begin{cases} -2x_1+x_2 &= 0\\ -x_1-x_3+x_4 &= 0\\ -x_3+x_5 &= 0 \end{cases}.$$

Wniosek 2. *Układ U ma jednoznaczne rozwiązanie wtedy i tylko wtedy, gdy $r(A_u)=r(A)=n$.*

Dowód. Układ U ma jednoznaczne rozwiązanie wtedy i tylko wtedy, gdy $r(A_u)=r(A)$ oraz $\dim W=0$, co jest równoważne $r(A_u)=r(A)=n$. □

Dokonaliśmy zatem klasyfikacji podprzestrzeni w K^n i umiemy je wyrażać zarówno jako przestrzenie rozpięte przez dany układ wektorów, jak i jako przestrzenie opisane przez określony układ równań. Te dwa opisy są na swój sposób dualne – jak starałem się zasugerować wyżej. Twierdzenie to jest głębsze niż się na początku wydaje. Powiem o tym trochę w dodatku, ale każdego Czytelnika powinna zastanowić następująca myśl: jeśli $W_1\subsetneq W_2$ są podprzestrzeniami K^n to do opisu przestrzeni W_2 potrzeba $n-\dim(W_2)$ jednorodnych równań liniowych i można ten układ uzupełnić $\dim(W_2)-\dim(W_1)$ liniowymi równaniami jednorodnymi tak, by uzyskać układ opisujący W_1.

Uzupełnienie. Kilka naiwnych uwag o prostopadłości

Zbiór rozwiązań liniowego równania jednorodnego $a_1x_1 + a_2x_2 = 0$ w $\mathbb{R}^2$ to prosta przechodząca przez punkt $(0,0)$, o ile tylko $(a_1, a_2) \neq (0,0)$. To jest, proszę zauważyć, twierdzenie Kroneckera-Capellego. Tylko gdy $(a_1, a_2) \neq (0,0)$, macierz tego układu ma rząd 1, a zatem i zbiór rozwiązań ma wymiar 1, co interpretujemy geometrycznie używając pojęcia „prostej". Z punktu widzenia geometrii elementarnej ważne może być stwierdzenie, że zbiór wektorów (x_1, x_2) spełniających powyższe równanie odpowiada zbiorowi wektorów (x_1, x_2), które są **prostopadłe**[1] do wektora (a_1, a_2). Kiedyś w szkole uczono (ale teraz już nie), że prostopadłość wektorów równoważna jest temu, że ich iloczyn skalarny równy jest 0. Natomiast wyrażenie $a_1x_1 + a_2x_2$ opisuje właśnie ów elementarny iloczyn skalarny wektorów (a_1, a_2) oraz (x_1, x_2). Podobnie zdefiniowany elementarny iloczyn skalarny w przestrzeni $\mathbb{R}^3$ pozwala stwierdzić, że jeśli tylko $(a_1, a_2, a_3) \neq (0,0,0)$, to zbiór rozwiązania równania $a_1x_2 + a_2x_2 + a_3x_3 = 0$ jest płaszczyzną przechodzącą przez punkt $(0,0,0)$ i prostopadłą do wektora (a_1, a_2, a_3). Używam określenia „elementarny". bowiem w drugim semestrze pojęcie iloczynu skalarnego zdefiniujemy w sposób aksjomatyczny, dla dowolnej przestrzeni liniowej nad $\mathbb{R}$, a po pewnych umowach i osłabieniu nazwy „iloczyn skalarny" na „funkcjonał dwuliniowy", badać będziemy choćby prostopadłość wektorów także w przestrzeniach nad innymi ciałami. A zatem okaże się, że prostopadłe mogą być grafy, funkcje zespolone, macierze, wielomiany (to akurat będzie dość ważne już na pierwszym roku Analizy) itd.

Możemy na razie nie martwić się abstrakcją, a zastanowić czym miałaby być prostopadłość układu wektorów w znanej przestrzeni K^n. Powiemy mianowicie, że dwa wektory $(a_1, \ldots, a_n)$ oraz $(b_1, \ldots, b_n)$ są prostopadłe, ozn. $(a_1, \ldots, a_n) \perp (b_1, \ldots, b_n)$, jeśli $a_1b_1 + a_2b_2 + \ldots + a_nb_n = 0$. Zbiór wektorów w K^n prostopadłych do ustalonego zbioru wektorów X oznaczać będziemy przez $X^\perp$. Innymi słowy:

$$X^\perp = \{v \in K^n : v \perp x, \forall_{x \in X}\}.$$

Proszę zauważyć, że dla każdego podzbioru $X \in K^n$ zbiór $X^\perp$ jest podprzestrzenią w K^n. Istotnie, jeśli $(v_1, v_2, \ldots, v_n)$ oraz $(w_1, \ldots, w_n)$ są prostopadłe do dowolnego wektora $(x_1, \ldots, x_n) \in X$, to są do niego prostopadłe również wektory $(v_1 + w_1, \ldots, v_n + w_n)$ oraz $(av_1, \ldots, av_n)$. po prostu dlatego, że

$$(v_1 + w_1)x_1 + \ldots + (v_n + w_n)x_n = v_1x_1 + \ldots + v_nx_n + w_1x_1 + \ldots + w_nx_n = 0.$$

Zauważmy dalej, że jeśli $X = \mathrm{lin}(\alpha_1, \ldots, \alpha_n)$, to podprzestrzeń $X^\perp$ opisuje zbiory współczynników wszystkich równań liniowych, których rozwiązaniami są wektory z X. Z drugiej strony: jeśli mamy jednorodny układ równań liniowych o macierzy, której wierszami są wektory $\alpha_1, \ldots, \alpha_n$, to zbiór rozwiązań tego układu równy jest... $\mathrm{lin}(\alpha_1, \ldots, \alpha_n)^\perp$. Czy Czytelnik widzi dualność, którą tu otrzymujemy? Czy Czytelnik widzi co robi tu Twierdzenie Kroneckera-Capellego? Mówi ono po prostu, że dla podprzestrzeni V przestrzeni liniowej K^n mamy (to się w ogólności zepsuje dla pewnych K i pewnych „niestandardowych prostopadłości", ale dla tej „elementarnej" – tzw. standardowej wersji to zawsze jest prawda):

$$(V^\perp)^\perp = V.$$

Zachęcam Czytelnika, by zastanowił się nad innymi konsekwencjami naszkicowanych tu definicji. Oczywiście (choć będą „wyjątki") $\dim V^\perp = n - \dim V$. Oczywiście, jeśli $V \subseteq W$ są podprzestrzeniami K^n, to $W^\perp \subseteq V^\perp$. Jakże istotne jest to odwrócenie kolejności pomiędzy podzbiorami i przestrzeniami prostopadłymi! Jest to bodaj najprostszy przykład odpowiedniości Galois, o której napiszę dalej. To jeszcze nie koniec. Nie wnikając w geometrię warto zauważyć, że prostopadłość jest swego rodzaju „lepszą liniową niezależnością". W przypadku skończonego układu wektorów liniowa niezależność nie wynika z tego, że dowolne dwa elementy układu są liniowo niezależne. Przyjmuje się natomiast następującą definicję.

Definicja 2. *Układ wektorów $X \subseteq K^n$ nazwiemy **prostopadłym** (albo **ortogonalnym**), jeśli $\alpha \perp \beta$, dla każdych $\alpha, \beta \in X$. Układ prostopadły będący bazą przestrzeni V nazywamy **bazą prostopadłą** (albo ortogonalną) przestrzeni V (względem naszego „stamdardowego" iloczynu skalarnego).*

Przykładem bazy prostopadłej jest oczywiście baza standardowa, oczywiście niejedynym. Czytelnikowi zostawiam następujące proste ćwiczenie: dowolny układ prostopadły złożony z niezerowych wektorów jest liniowo niezależny! Konsekwencje tego faktu są bardzo ciekawe, ale na razie nie będziemy eksplorować wątków geometrycznych. Zachęcam Czytelnika do poszukiwania prostopadłości w naszych rozważaniach.

[1] Osoby zainteresowane elementarnymi dowodami tych własności odnoszącymi się do twierdzeń szkolnych zachęcam do zajrzenia do wykładu dra Michała Krycha: *Elementy geometrii analitycznej*, dostępnego na stronie: `https://www.mimuw.edu.pl/~krych/chemia/2016-2017`.

Dodatek. Odpowiedniość Galois i Nullstellensatz Hilberta

Aby jeszcze lepiej i głębiej zrozumieć dlaczego twierdzenie Kroneckera-Capellego jest istotne, zdefiniujmy dwie operacje $\mathcal{R}$ oraz $\mathcal{W}$ na podzbiorach w K^n.

- Dla podzbioru $S \subseteq K^n$ przez $\mathcal{W}(S) \subseteq K^n$ rozumiemy zbiór złożony z n współczynników każdego takiego jednorodnego równania n zmiennych, którego **rozwiązaniem jest każdy element** z S. Innymi słowy, wektor $(a_1, \ldots, a_n) \in K^n$ należy do $\mathcal{W}(S)$ jeśli dla każdego $(s_1, \ldots, s_n) \in S$: zachodzi równość $a_1 s_1 + a_2 s_2 + \ldots + a_n s_n = 0$.
- Dla podzbioru $T \subseteq K^n$ przez $\mathcal{R}(T) \subseteq K^n$ rozumiemy **zbiór rozwiązań wszystkich jednorodnych równań** liniowych n zmiennych, których n-tki współczynników należą do T. Innymi słowy, wektor $(s_1, \ldots, s_n)$ należy do $\mathcal{R}(T)$ jeśli dla każdego $(a_1, \ldots, a_n) \in T$ zachodzi równość: $a_1 s_1 + a_2 s_2 + \ldots + a_n s_n = 0$.

Na przykład $(1, 1, -1) \in \mathcal{W}((2, 1, 3))$, ponieważ $1 \cdot 2 + 1 \cdot 1 + (-1) \cdot 3 = 0$, czyli $(2, 1, 3)$ jest rozwiązaniem równania $1 \cdot x_1 + 1 \cdot x_2 + (-1) \cdot x_3 = 0$. Są oczywiście inne elementy $\mathcal{W}(2, 1, 3)$, na przykład (-2, -2, 2). Weźmy jednak odwrotną sytuację: biorę wektor $(1, 1, -1)$ i interesuje mnie jakiś element $\mathcal{R}((1, 1, -1))$. Oczywiście – jednym z nich jest $(2, 1, 3)$, ale nie jedynym. Co to wszystko znaczy? Po co te komplikacje?

Nietrudno widzieć, że mamy dwie zależności (jest ich więcej):

$$\mathcal{W}(\mathcal{R}(S)) \supseteq S, \quad \mathcal{R}(\mathcal{W}(T)) \supseteq T.$$

Pierwsza z nich mówi, że każdy wektor jest elementem (czasami niejedynym, stąd inkluzja) zbioru rozwiązań równania, którego jest rozwiązaniem, a druga mówi, że jeśli równanie ma określone rozwiązanie, to rozwiązanie to jest jego rozwiązaniem (niekoniecznie jedynym, więc znowu jest inkluzja). Brzmi to niemal banalnie, ale interesujące jest to, że zależności te nie dotyczą jedynie równań liniowych! Zauważmy, że jeśli $S_1 \subseteq S_2$, to $\mathcal{W}(S_1) \supseteq \mathcal{W}(S_2)$, podobnie dla operacji $\mathcal{R}$. Wszystko, co powiedzieliśmy na dzisiejszym wykładzie można w zasadzie streścić prostym i eleganckim stwierdzeniem, że jeśli S jest podprzestrzenią liniową przestrzeni K^n – niezależnie czy rozumianą jako przestrzeń współczynników czy przestrzeń rozwiązań, to mamy $\mathcal{R}(S) = S^\perp$ oraz $\mathcal{W}(S) = S^\perp$, czyli:

$$\mathcal{W}(\mathcal{R}(S)) = S, \quad \mathcal{R}(\mathcal{W}(S)) = S.$$

Możemy też wrócić do wyjściowego przykładu i zapisać wyrażone w nim postulaty w nowym języku. Chcemy znaleźć układ $n-1$ równań, którego zbiorem rozwiązań jest **dokładnie** $\mathrm{lin}(\alpha) \neq 0$. Rzeczywiście:

- $\mathcal{W}(\mathrm{lin}(\alpha))$ jest przestrzenią $n-1$ wymiarową,
- dla $n-1$ liniowo niezależnych elementów $r_1, \ldots, r_{n-1}$ z $\mathcal{W}(\mathrm{lin}(\alpha))$ mamy $\mathcal{R}(r_1, \ldots, r_{n-1}) = \mathrm{lin}(\alpha)$.

Innymi słowy szukane przez nas $n-1$ równań będzie miało współczynniki będące bazą $\mathcal{W}(\mathrm{lin}(\alpha))$.

Operacje tego typu, co $\mathcal{W}$ i $\mathcal{R}$ „rozsiane są" po całej matematyce. Rozważmy jeden ważny przykład: podzbiorom X ciała K przyporządkowujemy zbiór $\mathcal{W}(X)$ wszystkich wielomianów o współczynnikach z $K[x]$, których pierwiastkami są wszystkie elementy zbioru X. I odwrotnie – każdemu zbiorowi wielomianów W można przypisać zbiór $\mathcal{R}(W)$ jego wspólnych pierwiastków w K. Np. dla $K = \mathbb{C}$, $W(\{-i, i\})$ to zbiór wszystkich wielomianów podzielnych przez $(x^2 + 1)$, a $W(\{0, -i, i\})$ równy jest zbiorowi wszystkich wielomianów podzielnych przez $x(x^2 + 1)$. Zauważmy też, że zbiór $\{-i, i\}$ jest zbiorem wspólnych rozwiązań istotnie różnych zbiorów wielomianów, na przykład zbioru wielomianów podzielnych przez $(x^2 + 1)^5$.

Dokładny opis zbioru wielomianów $\mathcal{W}(\mathcal{R}(W))$ nie jest banalny! Jest on treścią słynnego Nullstellensatz – twierdzenia Hilberta o zerach z 1893 roku, które jest uogólnieniem Zasadniczego Twierdzenia Algebry i punktem wyjścia geometrii algebraicznej. Powiedzmy kilka słów o tym twierdzeniu, bez wchodzenia w techniczne detale. Ograniczymy się jedynie do pokazania w jaki sposób twierdzenie to uogólnia Twierdzenie Kroneckera-Capellego. Chodzi mianowicie o rozwiązywanie układów równań, ale wielomianowych. Wychodzimy od następującej sytuacji. Mamy wielomiany $f_1, f_2, \ldots, f_m$ i chcemy coś powiedzieć o zbiorze rozwiązań układu $f_1 = 0, f_2 = 0, \ldots, f_m = 0$. I nie chodzi nam tylko o wielomiany w $K[x]$ tak, by zbiór wspólnych rozwiązań leżał w K. Chodzi nam o tzw. wielomiany n zmiennych i (znowu) o podzbiory K^n.

Definicja 3. ***Wielomianem zmiennych*** $x_1, \ldots, x_n$ ***o współczynnikach z ciała*** K *nazywamy wyrażenie postaci:*

$$\sum_{i_1, \ldots, i_n} a_{i_1 \ldots i_n} x_1^{i_1} x_2^{i_2} \ldots x_n^{i_n},$$

gdzie $i_1, \ldots, i_n$ są liczbami całkowitymi nieujemnymi (suma ta brana jest po wszystkich możliwych układach liczb całkowitych nieujemnych), elementy $a_{i_1 i_2 \ldots i_n} \in K$, przy czym zakładamy, że suma ta jest skończona, czyli współczynniki $a_{i_1 i_2 \ldots i_n}$ są różne od 0 tylko dla skończonej liczby indeksów $i_1, \ldots, i_n$. Zbiór wszystkich wielomianów zmiennych $x_1, \ldots, x_n$ o współczynnikach w ciele K oznaczamy $K[x_1, \ldots, x_n]$.

Stopniem wielomianu *$\sum_{i_1,\ldots,i_n} a_{i_1,\ldots,i_n} x_1^{i_1} x_2^{i_2} \ldots x_n^{i_n} \in K[x_1, \ldots, x_n]$ nazywamy największą z liczb $i_1 + i_2 + \ldots + i_n$, dla których $a_{i_1 \ldots i_n} \neq 0$. Stopień wielomianu f oznaczamy* $\deg f$*. Jeśli f jest* ***wielomianem zerowym*** *– to znaczy $a_{i_1 \ldots i_n} = 0$, dla wszystkich $i_1, \ldots, i_n$, to piszemy* $\deg f = -\infty$.

Zauważmy, że szczególnymi typami wielomianów są wielomiany liniowe, to znaczy wielomiany stopnia 1, np. $x_1, x_1 + \ldots + x_n, 2x_1 + x_3 - x_4$ itd. Rozwiązywanie jednorodnych układów równań liniowych jest z tej perspektywy szczególnym przypadkiem rozwiązywania wielomianowych układów równań. Ich rozwiązaniami są podprzestrzenie liniowe. Rozwiązaniami układów równań wielomianowych są tzw. zbiory algebraiczne. Powiemy o nich więcej na zakończenie drugiego semestru. Istotne jest to, że znamy wiele zbiorów algebraicznych, np. zbiór zer wielomianu dwóch zmiennych $x_1^2 - x_2^2$ to dwie przecinające się proste, a zbiór rozwiązań równania $x_1^2 - x_2$ to parabola (są też sfery, walce, hiperboloidy itd.). Badanie układów równań wielomianowych to punkt wyjścia wielkiego działu matematyki – geometrii algebraicznej. O czym jest zatem[2] Twierdzenie Hilberta? Zacznijmy od „stosunkowo prostej" sytuacji.

Weźmy element $a = (a_1, \ldots, a_n) \in K^n$ i zastanówmy się jak może wyglądać zbiór wielomianów n zmiennych, które zerują się na a, czyli $\mathcal{W}(\{a\})$. W przypadku wielomianów jednej zmiennej jest to po prostu zbiór $(x - a)f$, gdzie $f \in K[x]$. W przypadku wielomianów wielu zmiennych wnioskować należy, że w $\mathcal{W}(\{a\})$ jest każdy wielomian postaci:

$$(x_1 - a_1)f_1 + (x_2 - a_2)f_2 + \ldots + (x_n - a_n)f_n,$$

gdzie $f_1, \ldots, f_n \in K[x_1, \ldots, x_n]$ są dowolne. Można, jak się okazuje pokazać, że powyższy zbiór jest w istocie całym $\mathcal{W}(\{a\})$. Nie jest to bardzo trudne. Trudniej rozwiązywać problem odwrotny.

Twierdzenie Hilberta o zerach podejmuje następujący problem: załóżmy, że startujemy od pewnego zbioru wielomianów n zmiennych nad k, wyznaczamy wszystkie punkty w K^n, na których mogą się one zerować, a potem dla tych punktów wyznaczamy wszystkie wielomiany, które się na nich zerują. Co dostajemy? Innymi słowy, jeśli $X \subseteq K[x_1, \ldots, x_n]$, to czym jest $\mathcal{W}(\mathcal{R}(X))$? Nawet w przypadku wielomianów jednej zmiennej możemy otrzymać nietrywialne odpowiedzi, jak widzieliśmy wyżej. Warto założyć chociaż, że ciało K jest algebraicznie domknięte, żeby nie martwić się zbiorami pustymi. Załóżmy, że X jest skończonym zbiorem złożonym z wielomianów $f_1, \ldots, f_m$. Zbiór $\mathcal{R}(X)$ to zbiór jego wspólnych pierwiastków. Czym jest teraz $\mathcal{W}(\mathcal{R}(X))$? Jakie jeszcze wielomiany zerują się na tym samym zbiorze, co wielomiany $f_1, \ldots, f_m$? Na pewno są to wielomiany postaci: $f_1 g_1 + \ldots + f_m g_m$, gdzie $g_1, \ldots, g_m$ są dowolnymi wielomianami z $K[x_1, \ldots, x_n]$. Co jeszcze? Czasem coś jeszcze, bo np. $(x - 1) \in \mathcal{W}(\mathcal{R}(\{(x-1)^2\}))$. Okazuje się, że jest to jedyny rodzaj „niespodzianki", o czym mówi słynny wynik Hilberta.

Twierdzenie 2 (Hilberta o zerach)**.** *Niech K będzie ciałem algebraicznie domkniętym oraz niech $f_1, \ldots, f_m$ należą do $K[x_1, \ldots, x_n]$. Wówczas jeśli $h \in \mathcal{W}(\mathcal{R}(f_1, \ldots, f_m))$ (tzn. funkcja wielomianowa odpowiadająca h zeruje się na podzbiorze K^n, będącym częścią wspólną zbiorów zer funkcji wielomianowych odpowiadających $f_1, \ldots, f_m$), to istnieje $r \in \mathbb{Z}_+$ oraz wielomiany $g_1, \ldots, g_m$, że:*

$$h^r = f_1 g_1 + \ldots + f_m g_m.$$

W ten sposób wyróżnione zostają zbiory wielomianów X, dla których $\mathcal{W}(\mathcal{R}(X)) = X$. Zbiory te są bowiem w odpowiedniości ze zbiorami rozwiązań wielomianowych układów równań nad ciałem algebraicznie domkniętym, w podobny sposób jak w zależności tej są podprzestrzenie liniowe ze zbiorami rozwiązań jednorodnych liniowych układów równań. Dla zainteresowanych zostawiam jedynie hasło: ideał radykalny.

Olimpijczykom znany może być następujący fakt żyjący pod nazwą „kombinatoryczne Nullstellensatz".

Twierdzenie 3. *Niech p będzie niezerowym wielomianem zmiennych $x_1, \ldots, x_n$ stopnia $\sum_{i=1}^n m_i$, w którym współczynnik przy $x_1^{m_1} \cdots x_n^{m_n}$ jest różny od zera. Wówczas dla dowolnych zbiorów $S_1, \ldots, S_n$ zawartych w $\mathbb{R}$ spełniających warunki $|S_i| > m_i$, dla $1 \leqslant i \leqslant n$, istnieją takie $c_i \in S_i$, że $p(c_1, \ldots, c_n) \neq 0$.*

Zainteresowanych tym twierdzeniem i jego ładnymi elementarnymi aspektami odsyłam na przykład do artykułu Jacka Dymela „O zastosowaniach Combinatorial Nullstellensatz", dostępnego na stronach Delty: `http://www.deltami.edu.pl/temat/matematyka/algebra/2017/06/16/2017-07-delta-dymel.pdf`.

[2] Na motywach tekstu prof. Andrzeja Nowickiego: Afiniczne zbiory algebraiczne (do znalezienia na stronie Profesora) oraz tekstu Hilbert's Nullstellensatz w *The Princeton Companion to Mathematics.*