O sumach kwadratów

Andrzej SŁADEK, Katowice

1. Tożsamości dwuliniowe

Jedną z ważniejszych umiejętności, jaką powinien posiąść student w trakcie nauki geometrii przestrzeni euklidesowej, jest znajdowanie wektorów prostopadłych do danych wektorów. W szczególności ważne jest znajdowanie bazy prostokątnej zawierającej niezerowy, z góry ustalony wektor. Jeżeli przestrzeń ma wymiar 2, to można posłużyć się metodą „uczniowską" i do danego wektora $[y_1, y_2]$ po prostu dopisać wektor $[-y_2, y_1]$. Zastanówmy się, czy w przestrzeniach o wyższych wymiarach nie można z jednego wektora otrzymać bazy prostokątnej (lub przynajmniej „dużego" zbioru wektorów wzajemnie prostopadłych) permutując współrzędne wyjściowego wektora i zmieniając gdzieniegdzie znak na przeciwny. Łatwo zauważyć, że w przestrzeniach o nieparzystym wymiarze metoda ta jest całkowicie bezużyteczna. A jak to wygląda w przypadku parzystego wymiaru? W przestrzeni czterowymiarowej, startując od dowolnego niezerowego wektora $[y_1, y_2, y_3, y_4]$, możemy szybko znaleźć bazę prostokątną dopisując wektory $[y_2, -y_1, y_4, -y_3]$, $[y_3, -y_4, -y_1, y_2]$ $[y_4, y_3, -y_2, -y_1]$. W przestrzeni ośmiowymiarowej zadanie również ma rozwiązanie. Wektory $[y_1, y_2, y_3, y_4, y_5, y_6, y_7, y_8]$, $[-y_2, y_1, -y_4, y_3, -y_6, y_5, y_8, -y_7]$, $[-y_3, y_4, y_1, -y_2, -y_7, -y_8, y_5, y_6]$, $[-y_4, -y_3, y_2, y_1, -y_8, y_7, -y_6, y_5]$, $[-y_5, y_6, y_7, y_8, y_1, -y_2, -y_3, -y_4]$, $[-y_6, -y_5, y_8, -y_7, y_2, y_1, y_4, -y_3]$, $[-y_7, -y_8, -y_5, y_6, y_3, -y_4, y_1, y_2]$ oraz $[-y_8, y_7, -y_6, -y_5, y_4, y_3, -y_2, y_1]$ tworzą bazę prostokątną (sprawdź!).

Sformułujmy nasz problem bardziej precyzyjnie.

Zadanie: Niech $y_1, \ldots, y_n$ będą niezależnymi zmiennymi oraz niech S będzie zbiorem wszystkich wektorów postaci $[\mp y_{\sigma(1)}, \ldots, \mp y_{\sigma(n)}]$, $\sigma \in S_n$. Zbiór S ma $2^n n!$ elementów. Jaka jest możliwie największa liczba elementów podzbioru, w którym każde dwa różne elementy są prostopadłe (w sensie zwykłego iloczynu skalarnego)?

Wskazówka: Niech T będzie podzbiorem zbioru S, którego każde dwa różne elementy są prostopadłe. Oznaczmy przez r liczbę elementów zbioru T, natomiast przez A macierz, której kolumnami są wektory należące do T. Zatem $A = y_1A_1 + \ldots + y_nA_n$, gdzie $A_1, \ldots, A_n$ są macierzami wymiaru $n \times s$ o współrzędnych równych $0, \mp 1$, oraz $A^tA = (y_1^2 + \ldots + y_n^2)I_r$. Pomnóżmy powyższą równość z lewej strony przez $X = [x_1, \ldots, x_r]$ oraz z prawej strony przez X^t. Otrzymujemy $XA^tAX^t = (y_1^2 + \ldots + y_n^2)XX^t$. Oznaczmy teraz $Z = [z_1, \ldots, z_n] = AX^t$. Wtedy zachodzi równość $Z^tZ = (y_1^2 + \ldots + y_n^2)XX^t$, czyli

$$(1) \qquad (x_1^2 + \ldots + x_r^2)(y_1^2 + \ldots + y_n^2) = z_1^2 + \ldots + z_n^2,$$

gdzie $z_1, \ldots, z_n$ są formami dwuliniowymi zmiennych niezależnych $X = [x_1, \ldots, x_r]$ oraz $Y = [y_1, \ldots, y_n]$.

Równość (1) sugeruje następującą definicję.

Definicja. Tożsamością dwuliniową typu (r, s, n) nad ciałem F nazywamy tożsamość

$$(x_1^2 + \ldots + x_r^2)(y_1^2 + \ldots + y_s^2) = z_1^2 + \ldots + z_n^2,$$

gdzie $z_1, \ldots, z_n$ są formami dwuliniowymi (nad F) zmiennych niezależnych $X = [x_1, \ldots, x_r]$ oraz $Y = [y_1, \ldots, y_s]$.

Przykłady znanych tożsamości:

(1,1,1): $$x_1^2y_1^2 = (x_1y_1)^2$$

(2,2,2): $$(x_1^2 + x_2^2)(y_1^2 + y_2^2) = (x_1y_1 + x_2y_2)^2 + (x_1y_2 - x_2y_1)^2$$

(Diofantos, ok.250 n.e.)

$$(4,4,4): \quad (x_1^2+x_2^2+x_3^2+x_4^2)(y_1^2+y_2^2+y_3^2+y_4^2) =$$
$$= (x_1y_1+x_2y_2+x_3y_3+x_4y_4)^2+(x_1y_2-x_2y_1+x_3y_4-x_4y_3)^2+$$
$$+(x_1y_3-x_2y_4-x_3y_1+x_4y_2)^2+(x_1y_4+x_2y_3-x_3y_2-x_4y_1)^2$$

(Euler, 1748)

Zauważmy, że z ostatnich dwóch tożsamości można odczytać rozwiązanie naszego zadania dla $n = 2, 4$. W roku 1845 Cayley podał tożsamość typu $(8,8,8)$. Nie będziemy jej tutaj przytaczać, gdyż jest ona sporych rozmiarów. Czytelnik łatwo może ją odtworzyć z rozwiązania naszego zadania dla $n = 8$ lub znaleźć w książce W. Sierpińskiego, *Teoria liczb*, PWN 1950, str. 97.
O swego rodzaju kompletności powyższej listy przykładów świadczy następujące twierdzenie.

Twierdzenie (Hurwitz, 1898). Tożsamość dwuliniowa typu (n,n,n) nad ciałem F istnieje wtedy i tylko wtedy, gdy $n = 1, 2, 4, 8$.

Po udowodnieniu tego twierdzenia Hurwitz postawił pytanie: dla jakich liczb r, s oraz n istnieją tożsamości typu (r,s,n)?

Przykład:
$$(x_1^2+x_2^2+x_3^2)(y_1^2+y_2^2+y_3^2+y_4^2+y_5^2) =$$
$$= z_1^2+z_2^2+z_3^2+z_4^2+(x_1y_5)^2+(x_2y_5)^2+(x_3y_5)^2+(x_4y_5)^2,$$
gdzie z_1, z_2, z_3, z_4 są takie jak w tożsamości Eulera (dla $x_4 = 0$). Istnieje zatem tożsamość typu $(3,5,7)$.

Zauważmy, że teraz nasze zadanie można sformułować inaczej: dla jakiego r istnieje nad ciałem $\mathbf{R}$ tożsamość dwuliniowa typu (r,n,n) (spójrz na (1))? Rozwiązanie wynika z następującego twierdzenia.

Twierdzenie (Hurwitz, Radon, 1923). Tożsamość dwuliniowa typu (r,n,n) nad ciałem F istnieje wtedy i tylko wtedy, gdy $r \leq \rho(n)$, gdzie ρ jest funkcją określoną następująco:

$$\text{jeżeli } n = 2^m n_0,\ n_0 \notin 2\mathbf{Z},\ \text{to } \rho(n) = \begin{cases} 2m+1, & \text{gdy } m \equiv 0 \\ 2m, & \text{gdy } m \equiv 1 \\ 2m, & \text{gdy } m \equiv 2 \\ 2m+2, & \text{gdy } m \equiv 3 \end{cases} \pmod 4.$$

Funkcja ρ nazywana jest funkcją Hurwitza–Radona. Z tw. Hurwitza–Radona wynika tw. Hurwitza, bo $\rho(n) = n$ wtedy i tylko wtedy, gdy $n = 1, 2, 4, 8$.

Przykłady. $\rho(12) = 4$, $\rho(16) = 9$, $\rho(32) = 10$. Istnieją zatem tożsamości typu $(4,12,12)$, $(9,16,16)$ oraz $(10,32,32)$. Za chwilę poznamy metodę ich znajdowania.

Postępując podobnie jak we wskazówce do naszego zadania można wykazać, że tożsamość dwuliniowa typu (r,s,n) nad ciałem F istnieje wtedy i tylko wtedy, gdy istnieją macierze $A_1, \ldots, A_s$ nad F wymiaru $n \times s$ i takie, że macierz $A = y_1A_1 + \ldots + y_sA_s$ spełnia równość $A^tA = (y_1^2 + \ldots + y_s^2)I_r$. Tą ostatnią równość możemy zamienić na układ równań macierzowych (zwanym układem macierzowym Hurwitza)

$$(2) \qquad \begin{cases} A_i^tA_i = I_r & \text{dla } 1 \leq i \leq s, \\ A_i^tA_j + A_j^tA_i = 0 & \text{dla } 1 \leq i,j \leq s,\ i \neq j. \end{cases}$$

Metoda podwajania: Jeżeli $A_1, \ldots, A_s$ spełniają (2) dla $r = n$ (tzn. istnieje tożsamość typu (n,s,n)), to macierze

$$C_1 = \begin{bmatrix} A_1 & 0 \\ 0 & A_1 \end{bmatrix}, \quad C_j = \begin{bmatrix} A_j & 0 \\ 0 & -A_j \end{bmatrix}, \quad 2 \leq j \leq s, \quad C_{s+1} = \begin{bmatrix} 0 & A_1 \\ -A_1 & 0 \end{bmatrix},$$

również spełniają układ macierzowy Hurwitza i pozwalają napisać tożsamość typu $(2n, s+1, 2n)$ (lub, co wychodzi na to samo, tożsamość typu $(s+1, 2n, 2n)$). Zatem z tożsamości typu $(8,8,8)$ otrzymać można tożsamości typu $(9,16,16)$ oraz $(10,32,32)$.

2. Funkcyjna wersja tożsamości dwuliniowych

Zdefiniujmy odwzorowanie $\| \|^2 : F^k \to F$ określone wzorem: $\|X\|^2 = x_1^2 + \ldots + x_k^2$ dla $X = [x_1, \ldots, x_k] \in F^k$. Nietrudno zauważyć, że tożsamość typu (r, s, n) nad ciałem F istnieje wtedy i tylko wtedy, gdy istnieje odwzorowanie dwuliniowe $f : F^r \times F^s \to F^n$ spełniające warunek „normy": $\|f(X,Y)\|^2 = \|X\|^2 \cdot \|Y\|^2$, dla $X \in F^r$, $Y \in F^s$. Przypuśćmy, że istnieje tożsamość dwuliniowa typu (n, n, n). Wtedy przestrzeń wektorową F^n można wyposażyć w działanie mnożenia $*$, mianowicie $X * Y := f(X,Y)$, otrzymując tzw. algebrę kompozycyjną. W ten sposób dla $F = \mathbf{R}$ otrzymujeny ciało $\mathbf{R}$ $(n = 1)$, ciało $\mathbf{C}$ $(n = 2)$, algebrę kwaternionów Hamiltona $(n = 4)$ oraz algebrę Cayleya $(n = 8)$. Przypuśćmy teraz, że istnieje tożsamość dwuliniowa typu (r, s, n) oraz $F = \mathbf{R}$. Wtedy odwzorowanie f jest regularne (tzn. jeżeli $X \neq 0$ oraz $Y \neq 0$, to $f(X,Y) \neq 0$) i pełni ważną rolę w topologii algebraicznej. Hopf użył go do konstrukcji odwzorowania pomiędzy sferami jednostkowymi $H : S^{r+s-1} \to S^n$, $H(x,y) = (\|x\|^2 - \|y\|^2, 2f(x,y))$. Dla $r = s = n = 2, 4, 8$ otrzymujemy odwzorowania $S^3 \to S^2$, $S^7 \to S^4$, $S^{15} \to S^8$.

Definicja. $r *_F s := \min\{n :$ istnieje tożsamość typu (r, s, n) nad $F\}$.

Twierdzenie (Stiefel, Hopf, 1940). Jeżeli istnieje odwzorowanie dwuliniowe $f : \mathbf{R}^r \times \mathbf{R}^s \to \mathbf{R}^n$ spełniające warunek „normy" (tzn. $r *_{\mathbf{R}} s \leq n$), to spełniony jest warunek $\mathcal{H}(r, s, n)$: współczynniki Newtona $\binom{n}{k}$ są parzyste, gdy $n - r < k < s$.

Przykład. Z istnienia tożsamości typu $(3, 5, 7)$ (jeden z poprzednich przykładów) wynika nierówność $3 *_{\mathbf{R}} 5 \leq 7$, natomiast z powyższego twierdzenia mamy $6 < 3 *_{\mathbf{R}} 5$, bo $\binom{6}{4} = 15$. Zatem $3 *_{\mathbf{R}} 5 = 7$.

Szereg interesujących informacji na temat tożsamości dwuliniowych oraz znanych wartości $r *_F s$ znaleźć można w przeglądowym artykule D. Shapiro [1].

Oryginalne dowody twierdzeń Hurwitza oraz Hurwitza–Radona dotyczyły tożsamości nad $\mathbf{C}$. Ich uogólnienie na dowolne ciało nie stwarzało jednak żadnych trudności. Wiele innych znanych tożsamości dwuliniowych (jak np. tożsamość typu $(3, 5, 7)$) również nie zależy od ciała F. Naturalnym zatem wydaje się pytanie o te cechy ciała, które decydują o istnieniu lub nieistnieniu tożsamości danego typu. Odpowiedź na to pytanie nie jest znana.

Piękne uogólnienie twierdzenia Stiefela–Hopfa znalazł M. Szyjewski.

Twierdzenie (D. Shapiro, M. Szyjewski [2]). Załóżmy, że ρ, σ, τ są formami kwadratowymi nad ciałem F wymiarów odpowiednio r, s oraz n. Oznaczmy $r_0 = r - w(\rho)$, $s_0 = s - w(\sigma)$, $n_0 = n - w(\tau)$, gdzie $w(\rho), w(\sigma), w(\tau)$ są indeksami Witta (wymiarami maksymalnych całkowicie izotropowych podform) form ρ, σ, τ. Jeżeli istnieje dwuliniowe odwzorowanie $f : F^r \times F^s \to F^n$ spełniające warunek „normy":

$$\rho(X)\sigma(Y) = \tau(f(X,Y)) \text{ dla } X \in F^r,\ Y \in F^s,$$

to spełniony jest warunek $\mathcal{H}(r_0, s_0, n_0)$.

Przykład. Weźmy $r = 5$, $s = 10$, $n = 12$, F ciało algebraicznie domknięte. Wtedy formy ρ, σ oraz τ można potraktować jako sumy kwadratów i $r_0 = 3$, $s_0 = 5$, $n_0 = 6$. Ponieważ $\binom{6}{4} = 15$, więc tożsamość dwuliniowa typu $(5, 10, 12)$ nad F nie istnieje. Skoro każde ciało zawiera się w ciele algebraicznie domkniętym, więc tożsamość typu $(5, 10, 12)$ nie jest możliwa nad żadnym ciałem.

3. Tożsamości wymierne

Zależność pomiędzy warunkiem Hopfa $\mathcal{H}$, a iloczynami kwadratów zmiennych niezależnych (nad dowolnym ciałem F) opisuje następujące twierdzenie.

Twierdzenie (A. Pfister, J.W.S. Cassels, T.Y. Lam, A. Wadsworth, zob. [1]). Jeżeli F jest dowolnym ciałem charakterystyki $\neq 2$, to warunek $\mathcal{H}(r, s, n)$ jest

spełniony wtedy i tylko wtedy, gdy istnieje tożsamość

$$(3) \qquad (x_1^2 + \ldots + x_r^2)(y_1^2 + \ldots + y_s^2) = z_1^2 + \ldots + z_n^2,$$

gdzie $z_1, \ldots, z_n$ są **funkcjami wymiernymi** (nad F) zmiennych niezależnych $X = [x_1, \ldots, x_r]$ oraz $Y = [y_1, \ldots, y_s]$.

Definicja. Tożsamość (3) nazywamy tożsamością wymierną typu (r, s, n) nad ciałem F.

Jeżeli spojrzymy na trójkąt Pascala to zauważymy, że jeżeli n jest potęgą liczby 2, to wszystkie współczynniki Newtona $\binom{n}{k}$ są parzyste. Zatem, jeżeli $n = 2^m$, to tożsamość wymierna typu (n, n, n) jest możliwa nad dowolnym ciałem F. I tak na przykład tożsamość wymierna typu $(16, 16, 16)$ istnieje nad dowolnym ciałem podczas, gdy na mocy twierdzenia Hurwitza tożsamość dwuliniowa tego samego typu nie istnieje nad żadnym ciałem.

Wcześniejsze twierdzenie oraz uwaga uczyniona powyżej prowadzą do następującego twierdzenia.

Twierdzenie (A. Pfister, 1965). Oznaczmy $D_F(n)$ zbiór niezerowych sum n kwadratów elementów ciała F. Jeżeli $n = 2^m$, to zbiór $D_F(n)$ jest podgrupą multyplikatywnej grupy ciała F.

Definicja. Niezmiennikiem Pfistera ciała F nazywamy liczbę

$$s(F) := \min\{n : -1 \in D_F(n)\}$$

lub ∞, jeżeli -1 nie jest sumą kwadratów elementów ciała F.

Przykłady. $s(\mathbb{C}) = 1$, $s(\mathbb{F}_5) = 1$, $s(\mathbb{F}_3) = 2$, $s(\mathbb{Q}_2) = 4$, $s(\mathbb{Q}) = s(\mathbb{R}) = \infty$. Pfister wykazał, że dla dowolnego m istnieje ciało F takie, że $s(F) = 2^m$.

Powyższe przykłady sugerują, że niezmiennik Pfistera jest potęgą liczby 2. I tak rzeczywiście jest.

Twierdzenie (A. Pfister, 1965). Jeżeli $s(F) < \infty$, to $s(F) = 2^m$ dla pewnej nieujemnej liczby całkowitej m.

Dowód: Przypuśćmy, że $s(F) = k$. Zatem $-1 = a_1^2 + \ldots + a_k^2$ dla pewnych $a_1, \ldots, a_k \in F$. Wybierzmy m tak, aby $2^m \leq k < 2^{m+1}$. Po prostych przekształceniach otrzymujemy równość

$$-1 = (a_1^2 + \ldots + a_{2^m}^2)(1 + a_{2^m+1}^2 + \ldots + a_k^2)^{-1}.$$

Na podstawie poprzedniego twierdzenia $-1 \in D_F(2^m)$. Stąd (oraz z definicji $s(F)$) otrzymujemy $k = 2^m$. ∎

Na zakończenie pragniemy podkreślić, że w artykule rozważaliśmy sumy kwadratów przy założeniu, że F jest ciałem. Równie ciekawy jest problem, gdy F jest dowolnym pierścieniem (niekoniecznie przemiennym). Jest to jednak temat na odrębny artykuł.

Literatura

[1] D.B. Shapiro, *Product of sums of squares*, Expo. Math. 2 (1984), 235–261.

[2] D.B. Shapiro, M. Szyjewski, *Product formulas for quadratic forms* (w druku), oraz artykuły tam cytowane.