Macierz i wyznacznik Grama. Kryterium Sylvestera.

Na ostatnim wykładzie omówione zostały podstawowe metody wyznaczania bazy ortogonalnej przestrzeni euklidesowej. Dziś przekonamy się, że z iloczynem skalarnym i układem wektorów przestrzeni euklidesowej $(V, \langle\, ,\, \rangle)$ związać można pewne bardzo istotne macierze, tzw. macierze Grama. W języku wyznaczników związanych z tymi macierzami sformułujemy kryterium wyróżniające iloczyny skalarne spośród funkcji $f : V \times V \to \mathbb{R}$ liniowych ze względu na każdą ze zmiennych i symetrycznych, pochodzące od Sylvestera.

Definicja 1. *Mówimy, że macierz $A = (a_{ij}) \in M_{n\times n}(K)$ jest* `symetryczna`*, jeśli $a_{ij} = a_{ji}$ dla każdych $i, j = 1, \ldots, n$. Inaczej mówiąc A jest symetryczna jeśli $A = A^T$.*

Ważnym narzędziem do badania układów wektorów przestrzeni euklidesowej jest wyznacznik pewnej macierzy symetrycznej zwanej macierzą Grama.

Definicja 2. *Niech $\alpha_1, \ldots, \alpha_k$ będzie układem wektorów przestrzeni euklidesowej $(V, \langle\, ,\, \rangle)$.* `Macierzą Grama` *układu $\alpha_1, \ldots, \alpha_k$ nazywamy macierz $G(\alpha_1, \ldots, \alpha_k) \in M_{k\times k}(\mathbb{R})$, która w i-tym wierszu i j-tej kolumnie ma element $\langle \alpha_i, \alpha_j \rangle$, to znaczy:*

$$G(\alpha_1, \ldots, \alpha_k) = \begin{bmatrix} \langle \alpha_1, \alpha_1 \rangle & \langle \alpha_1, \alpha_2 \rangle & \ldots & \langle \alpha_1, \alpha_k \rangle \\ \langle \alpha_2, \alpha_1 \rangle & \langle \alpha_2, \alpha_2 \rangle & \ldots & \langle \alpha_2, \alpha_k \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle \alpha_k, \alpha_1 \rangle & \langle \alpha_k, \alpha_2 \rangle & \ldots & \langle \alpha_k, \alpha_k \rangle \end{bmatrix}$$

`Wyznacznikiem Grama` *układu $\alpha_1, \ldots, \alpha_k$ nazywamy liczbę $W(\alpha_1, \ldots, \alpha_k) = \det G(\alpha_1, \ldots, \alpha_k) \in \mathbb{R}$.*

Zauważmy, że rozmiar macierzy Grama nie musi mieć nic wspólnego z wymiarem przestrzeni V. W n-wymiarowej przestrzeni możemy rozpatrywać układ złożony na przykład z dwóch wektorów i wyznaczyć jego macierz Grama rozmiarów 2×2. Możemy też rozwazać układ złożony z $2n$ wektorów i jego macierz Grama będzie rozmiarów $2n \times 2n$. Z symetryczności iloczynu skalarnego $\langle \alpha_i, \alpha_j \rangle = \langle \alpha_j, \alpha_i \rangle$ dla każdych $1 \leqslant i, j \leqslant k$ widzimy, że macierz Grama jest symetryczna. Zobaczmy przykłady:

- W przestrzeni euklidesowej $\mathbb{R}^3$ ze standardowym iloczynem skalarnym rozpatrzmy układ wektorów $\alpha_1 = (1, 0, 1), \alpha_2 = (1, 1, 2)$. Wówczas $\langle \alpha_1, \alpha_1 \rangle = 2$, $\langle \alpha_1, \alpha_2 \rangle = \langle \alpha_2, \alpha_1 \rangle = 3$, $\langle \alpha_2, \alpha_2 \rangle = 6$. Stąd:

$$G(\alpha_1, \alpha_2) = \begin{bmatrix} 2 & 3 \\ 3 & 6 \end{bmatrix}, \quad W(\alpha_1, \alpha_2) = \begin{vmatrix} 2 & 3 \\ 3 & 6 \end{vmatrix} = 3.$$

- W $\mathbb{R}^2$ z iloczynem skalarnym zadanym wzorem $\langle (x_1, x_2), (y_1, y_2) \rangle = 2x_1y_1 + x_1x_2 + x_2x_1 + x_2y_2$ rozpatrzmy układ wektorów $\alpha_1 = (1, 0), \alpha_2 = (0, 1), \alpha_3 = (1, 1)$. Wówczas:

$$G(\alpha_1, \alpha_2, \alpha_3) = \begin{bmatrix} \langle \alpha_1, \alpha_1 \rangle & \langle \alpha_1, \alpha_2 \rangle & \langle \alpha_1, \alpha_3 \rangle \\ \langle \alpha_2, \alpha_1 \rangle & \langle \alpha_2, \alpha_2 \rangle & \langle \alpha_2, \alpha_3 \rangle \\ \langle \alpha_3, \alpha_1 \rangle & \langle \alpha_3, \alpha_2 \rangle & \langle \alpha_3, \alpha_3 \rangle \end{bmatrix} = \begin{bmatrix} 2 & 1 & 3 \\ 1 & 1 & 2 \\ 3 & 2 & 5 \end{bmatrix}, \quad W(\alpha_1, \alpha_2, \alpha_3) = 0.$$

  Zwróćmy uwagę na to, że $\alpha_1 + \alpha_2 = \alpha_3$, oraz dla macierzy $G(\alpha_1, \alpha_2, \alpha_3)$ pierwszy wiersz + drugi wiersz = trzeci wiersz oraz pierwsza kolumna + druga kolumna = trzecia kolumna. Widać dlaczego?

Uwaga 1. *Macierz Grama $G(\alpha_1, \ldots, \alpha_k)$ układu wektorów w $(V, \langle\, ,\, \rangle)$ wyznacza wartości iloczynu skalarnego na $\mathrm{lin}(\alpha_1, \ldots, \alpha_k)$: dla $v = a_1\alpha_1 + \ldots + a_k\alpha_k$ oraz $w = b_1\alpha_1 + \ldots + b_k\alpha_k$ mamy:*

$$\langle v, w \rangle = \langle a_1\alpha_1 + \ldots + a_k\alpha_k, b_1\alpha_1 + \ldots + b_k\alpha_k \rangle = \begin{bmatrix} a_1 & \ldots & a_k \end{bmatrix} \cdot G(\alpha_1, \ldots, \alpha_k) \cdot \begin{bmatrix} b_1 \\ \vdots \\ b_k \end{bmatrix}.$$

Uwaga 2. *Dla układu wektorów $\alpha_1, \ldots, \alpha_k$ przestrzeni euklidesowej $(V, \langle\, ,\, \rangle)$ oraz dla bazy ortonormalnej $\beta_1, \ldots, \beta_n$ tej przestrzeni niech A będzie macierzą rozmiaru $n \times k$ mającą w j-tej kolumnie współrzędne wektora α_j w bazie $\beta_1, \ldots, \beta_n$, to znaczy: $A = (a_{ij})$, gdzie $a_{ij} = \langle \alpha_j, \beta_i \rangle$, dla $i = 1, \ldots, k$, $j = 1, \ldots, n$. Wówczas:*

$$G(\alpha_1, \ldots, \alpha_k) = A^T A.$$

W szczególności dla $k = n$ mamy:

$$W(\alpha_1, \ldots, \alpha_n) = (\det A)^2.$$

Dowód. Z definicji iloczynu macierzy musimy pokazać, że wyraz $\langle \alpha_i, \alpha_j \rangle$ znajdujący się w i-tym wierszu i j-tej kolumnie macierzy $G(\alpha_1, \ldots, \alpha_k)$ równy jest wyrazowi w i-tym wierszu i j-tej kolumnie macierzy $A^T A$, czyli jedynemu wyrazowi iloczynu i-tego wiersza macierzy A^T traktowanej jako macierz $1 \times k$ oraz j-tej kolumny macierzy A, traktowanej jako macierz $k \times 1$. Z definicji macierzy transponowanej i-ty wiersz macierzy A^T ma kolejne wyrazy takie same jak kolejne wyrazy i-tej kolumny macierzy A. Czyli zawiera on współrzędne wektora α_i w bazie $\beta_1, \ldots, \beta_n$. Traktując go jako macierz $1 \times k$ mamy:

$$\begin{bmatrix} \langle \alpha_i, \beta_1 \rangle & \langle \alpha_i, \beta_2 \rangle & \ldots & \langle \alpha_i, \beta_n \rangle \end{bmatrix}.$$

Z drugiej strony j-ta kolumna A ma współrzędne wektora α_j w bazie $\beta_1, \ldots, \beta_n$, czyli: $\langle \alpha_j, \beta_1 \rangle, \ldots, \langle \alpha_j, \beta_n \rangle$. A zatem wyraz w i-tym wierszu i j-tej kolumnie macierzy $A^T A$ to jedyny wyraz macierzy 1×1 postaci:

$$\begin{aligned}\begin{bmatrix} \langle \alpha_i, \beta_1 \rangle & \ldots & \langle \alpha_i, \beta_n \rangle \end{bmatrix} \cdot \begin{bmatrix} \langle \alpha_j, \beta_1 \rangle \\ \vdots \\ \langle \alpha_j, \beta_n \rangle \end{bmatrix} &= [\langle \alpha_i, \beta_1 \rangle \cdot \langle \alpha_j, \beta_1 \rangle + \ldots + \langle \alpha_i, \beta_n \rangle \cdot \langle \alpha_j, \beta_n \rangle] = \\ &= [a_{i1}a_{j1} + \ldots + a_{in}a_{jn}].\end{aligned}$$

Zauważmy teraz, że element stojący w i-tym wierszu i j-tej kolumnie macierzy $G(\alpha_1, \ldots, \alpha_n)$ czyli $\langle \alpha_i, \alpha_j \rangle$, możemy przedstawić w innej postaci korzystając z formuł na współrzędne wektorów α_i w bazie ortonormalnej $\beta_1, \ldots, \beta_n$. Dokładniej mówiąc: $\alpha_i = a_{i1}\beta_1 + \ldots + a_{in}\beta_n$, bo $a_{ij} = \langle \alpha_i, \beta_j \rangle$. A zatem mamy:

$$\begin{aligned}\langle \alpha_i, \alpha_j \rangle &= \langle a_{i1}\beta_1 + \ldots + a_{in}\beta_n, a_{j1}\beta_1 + \ldots + a_{jn}\beta_n \rangle = \\ &= a_{i1}\langle \beta_1, a_{j1}\beta_1 + \ldots + a_{jn}\beta_n \rangle \qquad + \ldots + a_{in}\langle \beta_n, a_{j1}\beta_1 + \ldots + a_{jn}\beta_n \rangle = \\ &= a_{i1}a_{ji}\langle \beta_1, \beta_1 \rangle + \ldots + a_{i1}a_{jn}\langle \beta_1, \beta_n \rangle + \ldots + a_{in}a_{jn}\langle \beta_1, \beta_n \rangle + \ldots + a_{in}a_{jn}\langle \beta_n, \beta_n \rangle.\end{aligned}$$

Skoro baza $\beta_1, \ldots, \beta_n$ jest ortonormalna to $\langle \beta_i, \beta_j \rangle$ równe jest 0, dla $i \neq j$, oraz 1, dla $i = j$. A zatem ostatnie wyrażenie wynosi dokładnie $a_{i1}a_{j1} + \ldots + a_{in}a_{jn}$. To dowodzi, że $G(\alpha_1, \ldots, \alpha_k) = A^T A$. Druga część tezy wynika natychmiast z faktu, że $\det A = \det A^T$, gdy $k = n$, a zatem z twierdzenia Cauchy'ego o wyznacznikach mamy $\det G(\alpha_1, \ldots, \alpha_k) = \det A^T \det A = (\det A)^2$. Dowód jest zakończony. □

Dygresja. *Miasteczko Nieparzystów liczy sobie n mieszkańców. Ich głównym zajęciem jest formowanie rozmaitych klubów, co grozi wielkim politycznym chaosem. Rada Nieparzystowa rozporządziła jednak, aby każdy klub miał nieparzystą liczbę członków. Co więcej każde dwa kluby muszą mieć parzystą liczbę wspólnych członków. Pokazać, że nie jest możliwe stworzenie więcej klubów, niż wynosi ludność Nieparzystowa.*

Liczba potencjalnych klubów, czyli podzbiorów zbioru n-elementowego wynosi 2^n, a więc więcej niż n. Ile wśród nich jest klubów o nieparzystej liczbie członków? Wydaje się, że również możliwe jest, żeby było ich więcej niż n. Gdy $n = 3$ mamy możliwe trzy kluby o jednym członku i jeden klub o trzech członkach, a więc więcej niż 3. Łatwo jednak sprawdzić, że konfiguracja taka nie może wystąpić w przypadku, gdy każde dwa kluby muszą mieć parzystą liczbę wspólnych członków. Jak to pokazać dla każdego n?

Dowód. Ponumerujmy mieszkańców liczbami $1, 2, \ldots, n$, a kluby nazwijmy $C_1, \ldots, C_m$. Definiujemy macierz $A \in M_{n \times m}(\mathbb{R})$, w której kolumnach stoją „listy przynależności" do klubów. Dokładniej $A = (a_{ij})$:

$$a_{ij} = \begin{cases} 1, & \text{jeśli mieszkaniec } i \text{ należy do } C_j \\ 0, & \text{jeśli mieszkaniec } i \text{ nie należy do } C_j. \end{cases}$$

Przyjrzyjmy się teraz iloczynowi $A^T A$. Jest to macierz rozmiarów $m \times m$, której wyraz w wierszu i-tym oraz w kolumnie j-tej dostajemy ze wzoru: $\sum_{k=1}^{n} a_{ki}a_{kj}$. Co ta liczba oznacza? Zauważmy, że i-ty wiersz macierzy A^T liczy sobie dokładnie n wyrazów i wyraz a_{ki} koduje przynależność k-tego mieszkańca do klubu C_i. Wyraz ten wynosi 1, jeśli k-ty mieszkaniec należy do klubu C_i oraz 0, jeśli nie należy. Mnożenie i-tego wiersza A^T oraz j-tego wiersza A to po prostu standardowy iloczyn skalarny „list przynależności" do klubów C_i i C_j. Proszę zauważyć, że iloczyn ten ma składniki równe 0 lub 1. Składnik $a_{ki}a_{kj}$ jest równy 1 wtedy i tylko wtedy gdy k-mieszkaniec jest członkiem zarówno klubu C_i jak i klubu C_j. A zatem wyraz macierzy $A^T A$ stojący w i-tym wierszu i j-tej kolumnie będący sumą takich $a_{ki}a_{kj}$ opisuje dokładnie liczbę wspólnych członków klubu i-tego i j-tego! To jest dokładnie macierz Grama układu wektorów opisujących „listy przynależności" do klubów od 1 do m. Zauważmy jeszcze, że można mówić o „prostopadłości" wektorów z listami klubów. Oznacza ona po prostu, że kluby nie mają wspólnych członków.

Co to oznacza dla naszego problemu? Prawo w Nieparzystowie implikuje, że na przekątnej naszej macierzy Grama $A^T A$ stoi nieparzysta liczba członków i-tego klubu (bo i-ty klub ma wszystkich członków wspólnych z i-tym klubem), a wszędzie indziej stoją liczby parzyste. Nietrudno pokazać, że macierz $A^T A$ rozmiaru $m \times m$ jest pełnego rzędu m. Wynika to stąd, że wyznacznik macierzy Grama jest liczbą całkowitą (wszystkie wyrazy są całkowite) zaś wyznacznik macierzy o współczynnikach całkowitych ma resztę z dzielenia przez 2 taką samą jak wyznacznik macierzy powstałej przez podmiany pierwotnych wyrazów przez ich reszty modulo 2.[1] A więc dla zbadania parzystości wyznacznika macierzy całkowitoliczbowej:

$$\begin{vmatrix} 102495 & 550428 & 873298 & 660694 \\ 370628 & 909093 & 127450 & 925600 \\ 835044 & 601178 & 624653 & 263392 \\ 663780 & 487252 & 292276 & 593107 \end{vmatrix}$$

wystarczy wziąć każdy ze współczynników modulo 2 i dostajemy macierz I_4, której wyznacznik jest nieparzysty. To samo dzieje się w przypadku naszej „klubowej" macierzy Grama. Po wzięciu jej współczynników modulo 2 dostajemy I_m, a więc macierz $m \times m$ o niezerowym wyznaczniku 1. Zatem $r(A^T A) = m$. Na pierwszym semestrze wszyscy robiliśmy zapewne zadanie mówiące, że $r(XY) \leqslant \min(r(X), r(Y))$. To wynika szybko stąd, że wiersze XY są kombinacjami liniowymi wierszy macierzy X, zaś kolumny XY są kombinacjami liniowymi kolumn macierzy Y. Macierz A jest rozmiarów $n \times m$, więc nie może mieć rzędu większego niż n. Stąd $m = r(A^T A) \leqslant r(A) = r(A^T) \leqslant n$. Klubów jest nie więcej, niż mieszkańców. □

Sprytne, prawda? Argumenty w zadaniu o Nieparzystowie dotyczyły zasadniczo rzędu macierzy symetrycznej i na końcu kluczowa okazała się kwestia odwracalności macierzy Grama. Nie jest to przypadek.

Uwaga 3. *Dla każdego układu $\alpha_1, \ldots, \alpha_k$ wektorów przestrzeni euklidesowej $(V, \langle\, , \rangle)$ mamy nierówność $W(\alpha_1, \ldots, \alpha_k) \geqslant 0$. Co więcej następujące warunki są równoważne:*

(a) $\alpha_1, \ldots, \alpha_k$ *jest liniowo niezależny,*

(b) $W(\alpha_1, \ldots, \alpha_k) > 0$.

Dowód. Dowód pierwszej części tezy jest jasny, bo $W(\alpha_1, \ldots, \alpha_k)$ jest kwadratem liczby rzeczywistej. Rozważmy podprzestrzeń W przestrzeni V rozpiętą przez wektory $\alpha_1, \ldots, \alpha_k$. Oczywiście jest to przestrzeń euklidesowa z iloczynem skalarnym $\langle\, , \rangle|_W$. Niech $\beta_1, \ldots, \beta_m$ będzie bazą ortonormalną przestrzeni W. Oczywiście dla każdego $j = 1, 2, \ldots, k$ możemy napisać $a_j = a_{1j}\beta_1 + a_{2j}\beta_2 + \ldots + a_{mj}\beta_m$, dla pewnych $a_{ij} \in \mathbb{R}$. Z Lematu Steinitza wynika też oczywiście, że $m \leqslant k$ (układ m wektorów liniowo niezależnych należy do przestrzeni rozpiętej przez k wektorów).

Dowodzimy (a) $\Rightarrow$ (b). Skoro układ $\alpha_1, \ldots, \alpha_k$ jest liniowo niezależny, to jest to minimalny układ rozpinający W, a więc baza tej przestrzeni. Stąd $k = m$. W szczególności wyraz a_{ij} macierzy $A \in M_{k\times k}(\mathbb{R})$ zawiera j-tą współrzędną w bazie $\mathcal{B} = (\beta_1, \ldots, \beta_k)$ przestrzeni W i-tego wektora w bazie $\mathcal{A} = (\alpha_1, \ldots, \alpha_k)$. W rezultacie $A = M(id)_{\mathcal{A}}^{\mathcal{B}}$, czyli A jest odwracalna i $\det A \neq 0$. Zatem $(\det A)^2 = W(\alpha_1, \ldots, \alpha_k) \neq 0$. W połączeniu z pierwszą częścią tezy mamy (b).

Dowodzimy (b) $\Rightarrow$ (a). Załóżmy, że (a) nie jest prawdą, czyli układ $\alpha_1, \ldots, \alpha_k$ jest liniowo zależny. Oznacza to, że kolumny macierzy A są liniowo zależne (one zawierają współrzędne wektorów α_j w bazie $\mathcal{B}$). Jak wspomnieliśmy już w rozwiązaniu zadania o Nieparzystowie kolumny macierzy $A^T A$ są kombinacjami liniowymi kolumn macierzy A. A zatem i one są liniowo zależne. W szczególności mamy $\det(A^T A) = 0$, co jest zaprzeczeniem (b). □

Zwieńczeniem rozważań w tej części wykładu jest niezwykle istotne (i popularne) kryterium Sylvestera. Mówi ono o tym kiedy symetryczna macierz kwadratowa może być macierzą Grama. Przyczynę owej tajemniczej „popularności" zrozumiemy później. Kilka zdań wprowadzenia: rozpatrzmy n wymiarową przestrzeń liniową V nad ciałem $\mathbb{R}$ i jej bazę $\alpha_1, \ldots, \alpha_n$. Funkcja $f : V \times V \to \mathbb{R}$ zadana dla każdych wektorów $x_1\alpha_1 + \ldots + x_n\alpha_n, y_1\alpha_1 + \ldots + y_n\alpha_n$ wzorem: $f(x_1\alpha_1 + \ldots + x_n\alpha_n, y_1\alpha_1 + \ldots + y_n\alpha_n) = \sum_{i,j=1}^{n} a_{ij}x_i y_j$ jest liniowa ze względu na każdą ze zmiennych oraz $f(\alpha, \beta) = f(\beta, \alpha)$, dla każdych $\alpha, \beta \in V$. Pytanie: co musimy wiedzieć o współczynnikach a_{ij}, aby powiedzieć, że spełniony jest warunek $f(\alpha, \alpha) > 0$, dla $\alpha \neq 0$? Okazuje się, że kryterium jest bardzo czytelne, ale jego dowód jest dość techniczny.

Definicja 3. *Dla macierzy $A \in M_{n\times n}(K)$ oraz dla $i = 1, \ldots, n$ niech $A^{(i)} \in M_{i\times i}(K)$ oznacza macierz powstałą z A przez usunięcie ostatnich $n - i$ wierszy i ostatnich $n - i$ kolumn.*

[1] Można to pokazać na wiele sposobów. Natychmiast wynika to na przykład z wzoru permutacyjnego na wyznacznik.

Twierdzenie 1 (Kryterium Sylvestera). *Niech V będzie przestrzenią liniową nad $\mathbb{R}$ i niech $\mathcal{A} = (\alpha_1, \ldots, \alpha_n)$ będzie jej bazą. Dla macierzy symetrycznej $A = [a_{ij}] \in M_{n\times n}(\mathbb{R})$ rozpatrujemy funkcję $f : V \times V \to \mathbb{R}$ zadaną dla każdych wektorów postaci $x_1\alpha_1 + \ldots + x_n\alpha_n$, $y_1\alpha_1 + \ldots + y_n\alpha_n$ wzorem:*

$$f(x_1\alpha_1 + \ldots + x_n\alpha_n, y_1\alpha_1 + \ldots + y_n\alpha_n) = \sum_{i,j=1}^{n} a_{ij}x_iy_j.$$

Wówczas następujące warunki są równoważne.

(i) funkcja f jest iloczynem skalarnym na przestrzeni V,

(ii) $\det A^{(i)} > 0$*, dla $i = 1, \ldots, n$.*

Dowód. Dowodzimy (i) $\Rightarrow$ (ii). Jeśli f jest iloczynem skalarnym na V, to oczywiscie A jest macierzą Grama układu $(\alpha_1, \ldots, \alpha_n)$. Co więcej ograniczenie $f|_{W_i}$ do podprzestrzeni $W_i = \mathrm{lin}(\alpha_1, \ldots, \alpha_i)$ jest iloczynem skalarnym na przestrzeni W_i, dla każdego $i = 1, \ldots, n$. Stąd macierz $A^{(i)} = G(\alpha_1, \ldots, \alpha_i)$ ma dodatni wyznacznik na podstawie wcześniejszej uwagi.

Dowodzimy (ii) $\Rightarrow$ (i) przez indukcję po wymiarze n przestrzeni V. Oczywiście f spełnia trzy warunki: liniowości ze względu na każdą zmienną oraz warunek symetryczności (bo macierz A jest symetryczna) występujące w definicji iloczynu skalarnego. Trzeba jedynie pokazać, że f jest dodatnio określona, tzn. $f(\alpha, \alpha) > 0$ dla każdego niezerowego wektora $\alpha \in V$. Dla $n = 1$ teza jest jasna, bo $V = \mathrm{lin}(\alpha_1)$ oraz $A = [a_{11}]$, przy czym $a_{11} > 0$ (bo to jest $\det A^{(1)}$). A zatem dla każdego niezerowego $a\alpha_1 \in V$ zachodzi $f(a\alpha_1, a\alpha_1). = a^2a_{11} > 0$. Zatem f jest iloczynem skalarnym. Załóżmy, że dowodzona implikacja jest prawdziwa dla $n - 1$. Dowodzimy dla n. Niech $W = \mathrm{lin}(\alpha_1, \ldots, \alpha_{n-1})$. Z założenia indukcyjnego $(W, f|_W)$ jest przestrzenią euklidesową. Niech $\beta_1, \ldots, \beta_{n-1}$ będzie bazą prostopadłą W otrzymaną z bazy $\alpha_1, \ldots, \alpha_{n-1}$ metodą Grama-Schmidta. W szczególnosci $f(\beta_i, \beta_j) = 0$, dla $i \neq j$. Definiujemy wektor β_n postaci:

$$\beta_n = \alpha_n - \sum_{i=1}^{n-1} \frac{f(\alpha_n, \beta_i)}{f(\beta_i, \beta_i)}\beta_i.$$

Wówczas nowy wektor spełnia $f(\beta_n, \beta_i) = 0$, dla wszystkich $i = 1, \ldots, n - 1$ (sprawdzamy to podobnie jak w dowodach z poprzedniego wykładu, po prostu stosując liniowość ze względu na każdą zmienną oraz symetryczność f). Rozpatrzmy macierz $B = [b_{ij}] \in M_{n\times n}(\mathbb{R})$, gdzie $b_{ij} = f(\beta_i, \beta_j)$, dla $i, j = 1, \ldots, n$. Jest to macierz:

$$B = \begin{bmatrix} f(\beta_1, \beta_1) & 0 & \ldots & 0 \\ 0 & f(\beta_2, \beta_2) & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & f(\beta_n, \beta_n) \end{bmatrix}.$$

Teraz trzeba udowodnić coś intuicyjnie jasnego, ale technicznego w dowodzie. Mianowicie: w istocie macierz B powstaje z macierzy A przez wykonanie operacji elementarnych typu (1) na wierszach i kolumnach (dodawanie wiersza pomnożonego przez stałą do innego wiersza, analogicznie dla kolumn). Aby to zobaczyć po pierwsze zauważmy, że macierz $X_{\mathcal{B}}$ mająca w kolejnych wierszach wektory[2] $\beta_1, \ldots, \beta_n$ powstaje z macierzy $X_{\mathcal{A}}$ mającej w wierszach wektory $\alpha_1, \ldots, \alpha_n$ bazy $\mathcal{A}$ poprzez operacje typu (1) na wierszach.[3]

Pokażmy przez indukcję ze względu na m, że sprowadzenie pierwszych m wierszy macierzy $X_{\mathcal{A}}$ do postaci pierwszych m wierszy macierzy $X_{\mathcal{B}}$ wymaga stosowania jedynie operacji (1) na pierwszych m wierszach. Dla $m = 1$ mamy $\beta_1 = \alpha_1$, więc teza jest jasna, bo nie trzeba wykonywać żadnych operacji. Natomiast jeśli wiemy, że jest tak dla $m-1$, to bierzemy macierz $X_{\mathcal{A}}$. Zamieniamy pierwsze $m-1$ wierszy z założenia indukcyjnego na $\beta_1, \ldots, \beta_{m-1}$ operacjami (1) wykonywanymi w obrębie $m-1$ pierwszych wierszy. Dostajemy macierz $X_{\mathcal{A}'}$ której pierwsze $m-1$ wierszy to $\beta_1, \ldots, \beta_{m-1}$, a kolejne to $\alpha_m, \ldots, \alpha_n$. Z twierdzenia o ortogonalizacji oraz z definicji wektora β_n (gdyby $m = n$) wiemy, że $\beta_m = \alpha_m - c_1\beta_1 - c_2\beta_2 - \ldots - c_{m-1}\beta_{m-1}$, gdzie $c_i = f(\alpha_m, \beta_i)/f(\beta_i, \beta_i)$. A zatem od wiersza m-tego $X_{\mathcal{A}'}$ odejmujemy c_1 razy pierwszy wiersz. Od uzyskanego m-tego wiersza odejmujemy c_2 razy drugi wiersz i tak dalej, na końcu odejmując c_{m-1} razy wiersz $m - 1$ od wiersza m-tego. W rezultacie dostajemy macierz $X_{\mathcal{B}'}$, która w pierwszych m wierszach ma wektory $\beta_1, \ldots, \beta_m$, uzyskane przez stosowanie operacji (1) na pierwszych m wierszach, a dalej (jeśli są) wektory $\alpha_{m+1}, \ldots, \alpha_n$. To kończy dowód indukcyjny, że $X_{\mathcal{B}}$ można uzyskać z $X_{\mathcal{A}}$ przez operacje (1).

[2] Mała delikatność: jeszcze nie wiemy czy układ $\beta_1, \ldots, \beta_n$ jest bazą V, bo nie wiemy czy f to iloczyn skalarny na V.

[3] Intuicja: wykonujemy na kolejnych wierszach $X_{\mathcal{A}}$ te same operacje, co przy ortogonalizacji G-S, choć nie wiemy czy f to iloczyn skalarny. Na i-tym wierszu stosujemy $i - 1$ operacji (1), co zamienia α_i na $\beta_i = \alpha_i - c_1\beta_1 - \ldots - c_{i-1}\beta_{i-1}$.

Niech A' będzie f-macierzą Grama układu $\beta_1, \ldots, \beta_{m-1}, \alpha_m, \ldots, \alpha_n$ oraz B' niech będzie f-macierzą Grama układu $\beta_1, \ldots, \beta_m, \alpha_{m+1}, \ldots, \alpha_n$. Pisząc f-macierz Grama układu wektorów $v_1, \ldots, v_k$ mamy tu na myśli macierz o tej samej definicji, co macierz Grama, ale dla funkcji $f : V \times V \to \mathbb{R}$, spełniającej tylko warunek (b), czyli macierz o wyrazach $f(v_i, v_j)$, dla $1 \leqslant i, j \leqslant k$. Do dowodu, że B jest uzyskiwana z A przez operacje elementarne wierszowe i kolumnowe typu (1) wystarczy pokazać, że B' można uzyskać z A' przez operacje typu (1) zmieniające jedynie wiersze i kolumny o indeksie m. Wynika to z faktu, że macierze A' oraz B' mają identyczne wyrazy poza m-tym wierszem i m-tą kolumną, bo układy $\beta_1, \ldots, \beta_{m-1}, \alpha_m, \ldots, \alpha_n$ oraz $\beta_1, \ldots, \beta_m, \alpha_{m+1}, \ldots, \alpha_n$ różnią się tylko m-tym elementem. Zobaczmy:

$$A' = \begin{bmatrix} & \cdots & f(\beta_1, \beta_{m-1}) & f(\beta_1, \color{red}{\alpha_m}) & \cdots & \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ f(\beta_{m-1}, \beta_1) & \cdots & f(\beta_{m-1}, \beta_{m-1}) & f(\beta_{m-1}, \color{red}{\alpha_m}) & \vdots & f(\beta_{m-1}, \alpha_n) \\ f(\color{red}{\alpha_m}, \beta_1) & \cdots & f(\color{red}{\alpha_m}, \beta_{m-1}) & f(\color{red}{\alpha_m}, \color{red}{\alpha_m}) & \vdots & f(\color{red}{\alpha_m}, \alpha_n) \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ & \cdots & f(\alpha_n, \beta_{m-1}) & f(\alpha_n, \color{red}{\alpha_m}) & \cdots & \end{bmatrix}$$

$$B' = \begin{bmatrix} & \cdots & f(\beta_1, \beta_{m-1}) & f(\beta_1, \color{blue}{\beta_m}) & \cdots & \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ f(\beta_{m-1}, \beta_1) & \cdots & f(\beta_{m-1}, \beta_{m-1}) & f(\beta_{m-1}, \color{blue}{\beta_m}) & \vdots & f(\beta_{m-1}, \alpha_n) \\ f(\color{blue}{\beta_m}, \beta_1) & \cdots & f(\color{blue}{\beta_m}, \beta_{m-1}) & f(\color{blue}{\beta_m}, \color{blue}{\beta_m}) & \vdots & f(\color{blue}{\beta_m}, \alpha_n) \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ & \cdots & f(\alpha_n, \beta_{m-1}) & f(\alpha_n, \color{blue}{\beta_m}) & \cdots & \end{bmatrix}$$

Wiemy, że $\beta_m = \alpha_m - c_1\beta_1 - c_2\beta_2 - \ldots - c_{m-1}\beta_{m-1}$, dla znanych c_i. Macierz $X_{\mathcal{B}'}$ uzyskiwaliśmy z $X_{\mathcal{A}'}$ przez wykonanie $m-1$ operacji odjęcia od m-tego wiersza wielokrotności c_i wiersza i-tego. Aby uzyskać macierz B' z A' trzeba wykonać $2m-2$ operacji typu (1). Pierwsze dwie to: odjęcie od m-tego wiersza c_1 razy wiersz pierwszy, potem od kolumny m-tej trzeba odjąć c_1 razy kolumnę pierwszą. Oto rezultat:

$$\begin{bmatrix} & \cdots & f(\beta_1, \beta_{m-1}) & f(\beta_1, \color{red}{\alpha_m} - \color{cyan}{c_1\beta_1}) & \cdots & \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ f(\beta_{m-1}, \beta_1) & \cdots & f(\beta_{m-1}, \beta_{m-1}) & f(\beta_{m-1}, \color{red}{\alpha_m} - \color{cyan}{c_1\beta_1}) & \vdots & f(\beta_{m-1}, \alpha_n) \\ f(\color{red}{\alpha_m} - \color{olive}{c_1\beta_1}, \beta_1) & \cdots & f(\color{red}{\alpha_m} - \color{olive}{c_1\beta_1}, \beta_{m-1}) & f(\color{red}{\alpha_m} - \color{olive}{c_1\beta_1}, \color{red}{\alpha_m} - \color{cyan}{c_1\beta_1}) & \vdots & f(\color{red}{\alpha_m} - \color{olive}{c_1\beta_1}, \alpha_n) \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ & \cdots & f(\alpha_n, \beta_{m-1}) & f(\alpha_n, \color{red}{\alpha_m} - \color{cyan}{c_1\beta_1}) & \cdots & \end{bmatrix}$$

Następne dwie operacje to: od m-tego wiersza odjąć c_2 razy drugi wiersz i potem od m-tej kolumny odjąć c_2 razy kolumnę drugą, i tak dalej. W ten sposób, korzystając z liniowości f względem każdej ze zmiennych dostajemy żądaną postać B'. Zauważmy wszak, że m-ty wiersz macierzy B' to kombinacja liniowa c_1 razy pierwszy wiersz + c_2 razy drugi wiersz + $\ldots + c_{m-1}$ razy $m-1$-wszy wiersz. Podobnie dla kolumn. Zatem pokazaliśmy, że A można dostać z B operacjami elementarnymi typu (1).

Po co nam był ten wysiłek? Otóż operacje elementarne typu (1) nie zmieniają wyznacznika. Więc $\det A = \det B$. Wiemy też, że B jest diagonalna i ma na diagonali wyrazy $f(\beta_1, \beta_1), \ldots, f(\beta_n, \beta_n)$. Ich iloczyn to wyznacznik macierzy B. Wiemy z założenia indukcyjnego, że $f(\beta_1, \beta_1), \ldots, f(\beta_{n-1}, \beta_{n-1})$ są dodatnie, bo $f|_W$ jest iloczynem skalarnym i korzystamy z (i) $\Rightarrow$ (ii). Ostatni element diagonali $f(\beta_n, \beta_n)$ też musi być jednak dodatni, bo z (ii) mamy $\det A^{(n)} = \det A > 0$. To oznacza, że f jest iloczynem skalarnym, bo biorąc $0 \neq \alpha = x_1\beta_1 + \ldots + x_n\beta_n$ i korzystając z $f(\beta_i, \beta_j) = 0$, dla $i \neq j$, mamy $f(\alpha, \alpha) = \sum_{i=1}^{n} x_i^2 f(\beta_i, \beta_i) > 0$. $\square$

Przyznacie Państwo, że dowód był stosunkowo trudny. Stosowanie kryterium Sylvestera znacznie przekracza samą algebrę liniową. Przekonamy się za jakiś czas, że jest ono równoważne z kryterium tak zwanej dodatniej określoności macierzy. Jest to bardzo ważny warunek choćby w wielowymiarowej analizie matematycznej (np. przy szukaniu ekstremów funkcji wielu zmiennych), ale też w rachunku prawdopodobieństwa, kombinatoryce itd. Na kolejnym wykładzie powiemy o algebraicznym (prostym) podejściu do fundamentalnych w matematyce tematów n-wymiarowej objętości oraz orientacji przestrzeni n-wymiarowej.