Odwzorowania n-liniowe; formy n-liniowe

Definicja 1 Niech $V_1, \dots, V_n, U$ będą przestrzeniami liniowymi nad ciałem $\mathbb{K}$. Odwzorowanie $G \colon V_1 \times \cdots \times V_n \to U$ nazywamy *n-liniowym*, jeśli dla każdego $k \in [n]$ i wszelkich $x_1 \in V_1, \dots, x_{k-1} \in V_{k-1}$, $x_{k+1} \in V_{k+1}, \dots, x_n \in V_n$ odwzorowanie

$$V_k \ni x \mapsto G(x_1, \dots, x_{k-1}, x, x_{k+1}, \dots, x_n)$$

jest liniowe.

Zbiór wszystkich odwzorowań n-liniowych z $V_1 \times \cdots \times V_n$ w U oznaczamy $L(V_1, \dots, V_n; U)$. Zbiór ten ze zwykłym dodawaniem funkcji i mnożeniem ich przez skalary jest przestrzenią liniową. Każdy element G przestrzeni $L(V_1, \dots, V_n; \mathbb{K})$ nazywamy *formą n-liniową*.

Jeśli $f_1 \in V_1^*, \dots, f_n \in V_n^*$, gdzie V_i^* oznacza przestrzeń sprzężoną z V_i, to odwzorowanie

$$G(v_1, \dots, v_n) := f_1(v_1) \cdot \cdots \cdot f_n(v_n)$$

jest formą n-liniową. Oznaczamy je $f_1 \otimes \cdots \otimes f_n$ i nazywamy *iloczynem tensorowym funkcjonałów* $f_1, \dots, f_n$.

Zadanie 1 Dano przestrzenie liniowe $V_1, \dots, V_n$. Udowodnić, że jeśli $f_{i1}, \dots, f_{im_i}$ jest bazą przestrzeni V_i^*, $i = 1, \dots n$, to układ $f_{1k_1} \otimes \cdots \otimes f_{nk_n}$, $k_i \in [m_i]$, $i = 1, \dots, n$, tworzy bazę przestrzeni $L(V_1, \dots, V_n; \mathbb{K})$.

Formy dwuliniowe

Definicja 2 Niech V będzie przestrzenią liniową nad ciałem $\mathbb{K}$. Formę dwuliniową $F \in L(V, V; \mathbb{K})$ nazywamy *symetryczną*, jeśli $F(x, y) = F(y, x)$, dla każdej pary x, $y \in V$. Forma F jest *antysymetryczna*, jeśli $F(x, y) = -F(y, x)$, dla każdej pary x, $y \in V$.

Oczywistym przykładem formy dwuliniowej symetrycznej jest iloczyn skalarny. Forma $F \in L(\mathbb{R}^n, \mathbb{R}^n; \mathbb{R})$ dana wzorem

$$F(x, y) = \sum_{i<j} (x_i y_j - x_j y_i),$$

gdzie symbol $\sum_{i<j}$ oznacza, że sumowanie odbywa się po wszystkich parach wskaźników i, j spełniających dodatkowe ograniczenie $i < j$, jest antysymetryczna.

Jak łatwo zauważyć, zbiór form symetrycznych (podobnie antysymetrycznych) stanowi podprzestrzeń przestrzeni $L(V, V, \mathbb{K})$. Oznaczamy ją $\mathrm{Sym}(V)$.

Zadanie 2 Niech V będzie przestrzenią liniową nad $\mathbb{R}$. Udowodnić, że forma dwuliniowa $F \in L(V, V; \mathbb{R})$ jest antysymetryczna wtedy i tylko wtedy, gdy $F(x, x) = 0$, dla kżdego $x \in V$.

Każda forma dwuliniowa G o wartościach w ciele $\mathbb{K}$ charakterystyki różnej od 2 może zostać rozłożona na sumę form – symetrycznej G_+ i antysymetrycznej G_-. W tym celu wystarczy przyjąć

$$G_+(x, y) = \frac{G(x, y) + G(y, x)}{2}, \qquad G_-(x, y) = \frac{G(x, y) - G(y, x)}{2}.$$

Od tej chwili zakładamy, że rozważane przez nas przestrzenie liniowe są przestrzeniami skończonego wymiaru nad ciałem liczb rzeczywistych.

Definicja 3 Niech V będzie przestrzenią liniową i niech $F \in \mathrm{Sym}(V)$. Jeśli $\mathbf{e} = (e_1, \ldots, e_m)$ jest bazą uporządkowaną (reperem) w V, to macierz $M_F = [f_{ij}]$ o współczynnikach wyrażających się wzorem $f_{ij} = F(e_i, e_j)$ nazywamy *macierzą formy F w bazach* $(\mathbf{e}, \mathbf{e})$.

Dwuliniowość F prowadzi do wzoru

$$F(x, y) = \sum_{i=1}^{m} \sum_{j=1}^{m} x_i y_j f_{ij},$$

gdzie x_i, y_j to współrzędne wektorów x, y w układzie współrzędnych wyznaczonym przez reper $\mathbf{e}$. Zauważmy, że wyrazy macierzy M_F mogą być interpretowane jako współrzędne formy F w układzie współrzędnych wyznaczonym przez bazy $(\mathbf{e}, \mathbf{e})$. Co więcej, F jest symetryczna wtedy i tylko wtedy, gdy M_F jest macierzą symetryczną.

Definicja 4 Forma dwuliniowa symetryczna $F \in \mathrm{Sym}(V)$ *diagonalizuje się* w bazie $e_1, \ldots, e_m$ jeśli macierz M_F formy F w bazach $(\mathbf{e}, \mathbf{e})$ ma postać diagonalną, to znaczy, $f_{ij} = 0$, o ile tylko $i \neq j$. Mówimy też, że F ma postać *kanoniczną*, w układzie współrzędnych wyznaczonym przez $\mathbf{e}$. O bazie $\mathbf{e}$ mówimy, że jest *bazą kanoniczną* formy F.

Twierdzenie 1 (o diagonalizacji formy) *Niech V – przestrzeń liniowa nad ciałem $\mathbb{R}$ wymiaru $m < \infty$. Dla każdej formy $F \in \mathrm{Sym}(V)$ istnieje baza $e_1, \ldots, e_m$, w której F diagonalizuje się.*

Dowód. Niech $\mathbf{e}' = (e'_1, \dots, e'_m)$ – jakakolwiek baza (dokładniej reper) przestrzeni V. Niech $M'_F = [f'_{ij}]$ – macierz formy F w bazach $(\mathbf{e}', \mathbf{e}')$. Dla dowolnej pary $x = x'_1 e'_1 + \dots + x'_m e'_m$, $y = y'_1 e'_1 + \dots + y'_m e'_m$ określmy

$$\langle x, y \rangle = x'_1 y'_1 + \dots + x'_m y'_m \tag{1}$$

Odwzorowanie $(x, y) \mapsto \langle x, y \rangle$ jest iloczynem skalarnym w V, zaś $e'_1, \dots e'_m$ – bazą ortonormalną względem tego iloczynu. Określmy $T \in \mathrm{End}(V)$ wzorem

$$Te'_i = \sum_{k=1}^{m} f'_{ik} e'_k, \qquad i = 1, \dots, m.$$

Zauważmy, że

$$\langle Te'_i, e'_j \rangle = \langle \sum_k f'_{ik} e'_k, e'_j \rangle = f'_{ij} \text{ oraz } \langle e'_i, Te'_j \rangle = \langle e'_i, \sum_k f'_{jk} e'_k \rangle = f'_{ij}.$$

Stąd, biorąc pod uwagę symetryczność F, która z kolei pociąga symetryczność macierzy M'_F, dostaniemy

$$\langle Te'_i, e'_j \rangle = \langle e'_i, Te'_j \rangle.$$

Co oznacza, że T jest odwzorowaniem samosprzężonym. Z twierdzenia o diagonalizacji wynika teraz istnienie bazy ortonormalnej $\mathbf{e} = (e_1, \dots, e_m)$, względem wprowadzonego iloczynu skalarnego, oraz liczb $\lambda_1, \dots, \lambda_m$, że $Te_i = \lambda_i e_i$ dla $i = 1, \dots, m$. Z określenia odwzorowania T wnioskujemy, że dla każdej pary $x, y \in V$ mamy

$$\langle Tx, y \rangle = \sum_i \sum_j \langle Te'_i, e'_j \rangle x'_i x'_j = \sum_i \sum_j f'_{ij} x'_i x'_j = F(x, y)$$

W konsekwencji,

$$F(x, y) = \langle T(\sum_i x_i e_i), \sum_j y_j e_j \rangle = \langle \sum_i \lambda_i x_i e_i, \sum_j y_j e_j \rangle = \sum_i \sum_j \lambda_i x_i y_j \langle e_i, e_j \rangle.$$

Wobec ortonormalności bazy $\mathbf{e}$ mamy $\langle e_i, e_j \rangle = \delta_{ij}$, co prowadzi do zależności

$$F(x, y) = \sum_i \lambda_i x_i y_i.$$

F diagonalizuje się zatem w bazie $\mathbf{e}$, co więcej macierz M_F w bazach $(\mathbf{e}, \mathbf{e})$ jest identyczna z macierzą odwzorowania T w tych bazach. □

Uwaga 1 Fakt, że forma $F \in \mathrm{Sym}(V)$ diagonalizuje się w bazie $e_1, \dots, e_m$ można wyrazić w następujący sposób: Niech $e^1, \dots, e^m$ oznacza bazę dualną do $e_1, \dots, e_m$. Wówczas istnieją skalary $\lambda_1, \dots, \lambda_m$, że F daje się zapisać w postaci

$$F = \sum_i^m \lambda_i e^i \otimes e^i.$$

Prawo bezwładności formy dwuliniowej symetrycznej

Twierdzenie dyskutowane w tym paragrafie nazywane jest imieniem jego odkrywcy, angielskiego matematyka Sylvestra. Mówi się o nim także jako o *prawie bezwładności formy dwuliniowej symetrycznej.*

Twierdzenie 2 (Sylvestra prawo bezwładności) *Niech przestrzeń liniowa V nad $\mathbb{R}$ będzie skończonego wymiaru m. Niech forma $F \in \mathrm{Sym}(V)$ ma postać kanoniczną w układzie współrzędnych wyznaczonym przez reper $\mathbf{e} = (e_1, \dots, e_m)$, to znaczy,*

$$F(x, y) = \lambda_1 x_1 y_1 + \dots + \lambda_m x_m y_m. \tag{2}$$

Niech $d := |\{i \colon \lambda_i > 0\}|$, $u := |\{i \colon \lambda_i < 0\}|$ oraz $z := m - d - u = |\{i \colon \lambda_i = 0\}|$. Liczby u, d, z nie zależą od wyboru bazy kanonicznej formy F; to znaczy, dla każdej takiej bazy są identyczne.

Dowód. Wyróżnijmy podprzestrzenie V_0, V_- oraz V_+ przestrzeni V:

$$V_0 := \{x \colon x_i \neq 0 \Rightarrow \lambda_i = 0, \text{dla każdego } i \in [m]\},$$

$$V_- := \{x \colon x_i \neq 0 \Rightarrow \lambda_i \leqslant 0, \text{dla każdego } i \in [m]\},$$

$$V_+ := \{x \colon x_i \neq 0 \Rightarrow \lambda_i \geqslant 0, \text{dla każdego } i \in [m]\}.$$

Jak łatwo zauważyć, podprzestrzenie te wchodzą w następujące związki:

$$V_+ + V_- = V, \tag{3}$$

$$V_+ \cap V_- = V_0. \tag{4}$$

Niech $\mathbf{e}' = (e'_1, \dots, e'_m)$ będzie inną bazą kanoniczną. Wyróżnione względem tej bazy podprzestrzenie oraz odpowiedniki liczb d, u, z, oznaczmy przez „dołożenie primów". Zauważmy jednak, że

$$V_0 := \{x \colon F(x, y) = 0, \text{dla każdego } y \in V\}.$$

W takim razie V_0 nie zależy od wyboru bazy kanonicznej, więc $V_0 = V_0'$, a stąd $z' = z$.

Przypomnijmy, że jeśli X i Y są podprzestrzeniami (skończonego wymiaru) dowolnej podprzestrzeni liniowej Z, to

$$\dim(X+Y) = \dim X + \dim Y - \dim(X \cap Y). \tag{5}$$

By zakończyć dowód twierdzenia musimy wykazać, że $d = d'$ oraz $u = u'$. Przypuśćmy, że tak nie jest. Wtedy przynajmniej jedna z równości nie zachodzi. Jako że wszystkie przypadki rozpatruje się w identyczny sposób, możemy przyjąć, że nie zachodzi pierwsza z nich i że ponadto $d' > d$. Na podstawie (5), (3) i (4):

$$m \geqslant \dim(V_- + V_+') = \dim(V_-) + \dim(V_+') - \dim(V_- \cap V_+') =: k,$$

$$m = \dim(V_- + V_+) = \dim(V_-) + \dim(V_+) - \dim(V_0) =: l.$$

Oczywiście $k - l \leqslant 0$, skąd już łatwo wywnioskować, że

$$\dim(V_- \cap V_+') - \dim V_0 \geqslant \dim(V_+') - \dim(V_+) = d' - d > 0.$$

Ponieważ $V_- \cap V_+' \supseteq V_0$, więc ostatnia nierówność pozwala stwierdzić, że istnieje element x, iż $x \in V_- \setminus V_0$, a także $x \in V_+' \setminus V_0$. Fakt, że $x \in V_- \setminus V_0$ pociąga istnienie takich zbiorów J i K zawartych w $[m]$, że $J \neq \emptyset$ oraz $\lambda_i < 0$ dla $i \in J$ i $\lambda_i = 0$ dla $i \in K$, a nadto przedstawienie x w bazie $e_1, \ldots, e_m$ ma postać $x = \sum_{i \in J} x_i e_i + \sum_{i \in K} x_i e_i$, gdzie współczynniki x_i, $i \in J$, są niezerowe. W takim razie

$$F(x,x) = \sum_{i=1}^{m} \lambda_i x_i^2 = \sum_{i \in J} \lambda_i x_i^2 < 0.$$

W podobny sposób przynależność x do $V_+' \setminus V_0$ prowadzi do wniosku, że $F(x,x) > 0$. Otrzymaliśmy więc sprzeczność. □

Układ liczb (d, u, z) określonych w twierdzeniu Sylvestra nazywamy *sygnaturą* formy F

Zadanie 3 Określmy odwzorowanie $F\colon \mathsf{M}_{n\times n} \times \mathsf{M}_{n\times n} \to \mathbb{R}$ wzorem $F(A,B) = \operatorname{tr}(AB)$, gdzie $\operatorname{tr} C := \sum_{i=1}^{n} c_{ii}$ oznacza *ślad* macierzy C. Udowodnić, że F jest formą dwuliniową symetryczną. Znaleźć jej sygnaturę.

Formy kwadratowe Niech V będzie przestrzenią liniową (nad $\mathbb{R}$) i niech $F \in \mathrm{Sym}(V)$. Określmy odwzorowanie $q_F \colon V \to \mathbb{R}$ wzorem

$$q_F(x) = F(x,x).$$

Odwzorowanie to nazywamy *formą kwadratową stowarzyszoną z F*. Znajomość q_F pozwala na odtworzenie F:

$$\begin{aligned}\frac{q_F(x+y) - q_F(x) - q_F(y)}{2} &= \frac{F(x+y,x+y) - F(x,x) - F(y,y)}{2} \\ &= \frac{F(x,y) + F(y,x)}{2} = F(x,y),\end{aligned}$$

gdzie w ostatnim przejściu wykorzystaliśmy fakt, że odtwarzana forma ma być symetryczna. Formuła, którą posłużyliśmy się, nosi nazwę *formuły polaryzacyjnej.*

Definicja 5 Odwzorowanie q z przestrzeni liniowej V w $\mathbb{R}$ nazywamy formą kwadratową, jeśli istnieje $F \in \mathrm{Sym}(V)$, że $q = q_F$. Zbiór wszystkich form kwadratowych $q : V \to \mathbb{R}$ oznaczamy $Q(V)$.

Na mocy powyższej definicji, odwzorowanie $\mathrm{Sym}(V) \ni F \mapsto q_F \in Q(V)$ jest surjekcją. Z drugiej strony, jak wykazaliśmy z pomocą formuły polaryzacyjnej, jest ono także injekcją. Wprost z definicji tego odwzorowania wynika, że jeśli $F, G \in \mathrm{Sym}(V)$ oraz $\alpha, \beta \in \mathbb{R}$, to

$$q_{\alpha F + \beta G} = \alpha q_F + \beta q_G$$

W rezultacie $Q(V)$ jest przestrzenią liniową zaś $F \mapsto q_F$ ustala izomorfizm przestrzeni liniowych $\mathrm{Sym}(V)$ i $Q(V)$. Oznacza to, że twierdzenia sformułowane dla form dwuliniowych symetrycznych można wyrazić w języku form kwadratowych.

W analizie matematycznej czy w geometrii analitycznej mówi się o formach kwadratowych n-zmiennych rzeczywistych $x_1, \ldots, x_n$. Taka forma, to dowolna funkcja $f \colon \mathbb{R}^n \to \mathbb{R}$ zadana wzorem

$$f(x_1, \ldots, x_n) = a_1 x_1^2 + \ldots + a_n x_n^2 + \sum_{i<j} a_{ij} x_i x_j,$$

Ustalone liczby a_i, $i = 1, \ldots, n$, a_{ij}, $j = 1, \ldots, n$, $i = 1, \ldots, j-1$ nazywamy współczynnikami formy f. Określmy macierz $[f_{ij}] \in \mathsf{M}_{n\times n}(\mathbb{R})$ w następujący sposób:

$$f_{ii} = a_i,\ i = 1, \ldots, n; \quad f_{ij} = \frac{a_{ij}}{2}, \text{ jeśli } i < j; \quad f_{ij} = \frac{a_{ji}}{2}, \text{ jeśli } i > j.$$

Z pomocą tej macierzy określmy formę $F \in \mathrm{Sym}(V)$

$$F(x, y) = \sum_{i=1}^{n}\sum_{j=1}^{n} x_i y_j f_{ij}.$$

Oczywiście $f = q_F$. Chodzi więc o te same formy kwadratowe, które zdefiniowaliśmy wcześniej.

Twierdzenie o diagonalizacji i twierdzenie Sylvestra zbierzemy teraz w jedno twierdzenie dotyczące form kwadratowych.

Twierdzenie 3 (Sylvester) *Jeśli $q \in Q(V)$, to istnieje baza $e_1, \ldots, e_m$ przestrzeni V, w której forma q diagonalizuje się, to znaczy, przyjmuje postać*

$$q(x) = q(x_1e_1 + \cdots + x_ne_n) = a_1x_1^2 + \cdots + a_nx_n^2.$$

Liczby $d := |\{i\colon a_i > 0\}|$ oraz $u := |\{i\colon a_i < 0\}|$ nie zależą od wyboru bazy $e_1, \ldots, e_m$ diagonalizującej q.

Podobnie jak w przypadku form symetrycznych, układ liczb (d, u, z), gdzie $z = m - d - u$ nazywamy *sygnaturą* formy kwadratowej q.

Definicja 6 Formę kwadratową $q \in Q(V)$ nazywamy:

- *dodatnio określoną*, jeśli $q(x) > 0$, dla każdego $x \neq 0$;
- *nieujemnie określoną*, jeśli $q(x) \geqslant 0$, dla każdego x;
- *niedodatnio określoną*, jeśli $q(x) \leqslant 0$, dla każdego x;
- *ujemnie określoną*, jeśli $q(x) > 0$, dla każdego elementu x;
- *nieokreśloną*, jeśli przyjmuje ona tak wartości ujemne, jak i dodatnie.

Rzecz jasna q jest

- dodatnio określona jedynie wówczas, gdy $d = m$;
- nieujemnie określona jedynie wówczas, gdy $u = 0$;
- niedodatnio określona jedynie wówczas, gdy $d = 0$;
- ujemnie określona jedynie wówczas, gdy $u = m$;
- nieokreślona jedynie wówczas, gdy $d > 0$ i $u > 0$.

Jeśli natomiast $z = 0$, to mówimy, że forma q jest *niezdegenerowana*; w przeciwnym razie forma jest *zdegenerowana.*

Na koniec zauważmy, że q jest dodatnio określona wtedy i tylko wtedy, gdy jest stowarzyszona z formą $F \in \mathrm{Sym}(V)$ (to znaczy $q = q_F$) będącą iloczynem skalarnym w V.

Zależność między macierzami form dwuliniowych w różnych bazach

Zacznijmy od przypomnienia. Niech $\mathbf{e}$ oraz $\mathbf{e}'$ – dwa repery przestrzeni V. Niech $A = [a_{ij}]$ – macierz przejścia od bazy $\mathbf{e}$ do bazy $\mathbf{e}'$, to znaczy

$$e_k = \sum_i a_{ik} e'_i$$

Wtedy dla każdego $v = x_1 e_1 + \cdots + x_m e_m$ mamy

$$v = \sum_k x_k (\sum_i a_{ik} e'_i) = \sum_i (\sum_k a_{ik} x_k) e'_i = \sum_i x'_i e'_i.$$

Stąd

$$x' = \begin{bmatrix} x'_1 \\ \vdots \\ x'_n \end{bmatrix} = Ax = \begin{bmatrix} a_{11} & \dots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mm} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}.$$

(We wprowadzonych wcześniej oznaczeniach $x = v_{\mathbf{e}}$ oraz $x' = v_{\mathbf{e}'}$.) Ustalmy teraz zależność pomiędzy macierzą M_F formy F w bazach $(\mathbf{e}, \mathbf{e})$, a macierzą M'_F tej samej formy w bazach $(\mathbf{e}', \mathbf{e}')$:

$$f_{ij} = F(e_i, e_j) = F(\sum_k a_{ki} e'_k, \sum_l a_{lj} e'_l) = \sum_k \sum_l a_{ki} a_{lj} F(e'_k, e'_l),$$

stąd

$$f_{ij} = \sum_k \sum_l a_{ki} a_{lj} f'_{kl} = \sum_k \sum_l a^{\mathsf{T}}_{ik} f' kl a_{lj} = (A^{\mathsf{T}} M'_F A)_{ij}$$

W konsekwencji

$$M_F = A^\mathsf{T} M'_F A.$$

Łatwo teraz zrozumieć sformułowanie twierdzenia Sylvestra w języku macierzy:

Twierdzenie 4 (Sylvestra dla macierzy) *Niech $A, B, C \in \mathsf{M}_{n\times n}(\mathbb{R})$. Jeśli macierze A i B są odwracalne oraz macierze $A^\mathsf{T}CA$ i $B^\mathsf{T}CB$ są diagonalne, to mają one na przekątnych tę samą liczbę wyrazów ujemnych, zer i wyrazów dodatnich.*

Uwaga 2 Każdą formę $F \in L(\mathbb{R}^m, \mathbb{R}^m; \mathbb{R})$ można zapisać w postaci iloczynu macierzy $F(x, y) = x^\mathsf{T}Cy$, gdzie $C = M_F$, a M_F odniesiona jest do bazy standardowej w $\mathbb{R}^m$. Ten rodzaj zapisu jest często stosowany w programowaniu matematycznym i naukach inżynierskich.