§ 4. FORMY DWULINIOWE I KWADRATOWE

1. Odwzorowania wieloliniowe. Punkt ten można opuścić przy pierwszym czytaniu. Pojęcie kowektora (formy liniowej na przestrzeni V), które udowodniło już swoją przydatność, ma daleko idące uogólnienie:

Rozpatrzmy przestrzenie liniowe $V_1, \ldots, V_p, U$ nad ciałem $\mathfrak{K}$. Odwzorowanie

$$f : V_1 \times V_2 \times \ldots \times V_p \to U$$

nazywamy *wieloliniowym* (w danym wypadku *p-liniowym*), jeśli dla każdego $i = 1, \ldots, p$ i dla dowolnych wektorów $\mathbf{a}_j \in V_j$, gdzie $1 \leqslant j \leqslant p$ i $j \neq i$, odwzorowanie

$$f_i : \mathbf{v} \mapsto f(\mathbf{a}_1, \ldots, \mathbf{a}_{i-1}, \mathbf{v}, \mathbf{a}_{i+1}, \ldots, \mathbf{a}_p)$$

jest liniowe, tj.

$$f_i(\alpha \mathbf{x} + \beta \mathbf{y}) = \alpha f_i(\mathbf{x}) + \beta f_i(\mathbf{y}) \qquad \forall \mathbf{x}, \mathbf{y} \in V_i, \ \alpha, \beta \in \mathfrak{K}. \tag{1}$$

Badaniem takich odwzorowań będziemy się zajmować w rozdziale 2; tutaj ograniczymy się do spostrzeżenia, że — podobnie jak w przypadku form liniowych — kombinacja liniowa dwóch odwzorowań p-liniowych jest odwzorowaniem p-liniowym. Pozwala to rozpatrywać zbiór $\mathcal{L}(V_1, \ldots, V_p; U)$ wszystkich odwzorowań p-liniowych $V_1 \times \ldots \times V_p \to U$ jako przestrzeń liniową nad $\mathfrak{K}$.

Najprostszy przykład otrzymamy, biorąc $V_1 = \ldots = V_p = U = \mathfrak{K}$ (przestrzenie jednowymiarowe) i przyjmując

$$f(v_1, \ldots, v_p) = v_1 \ldots v_p.$$

Ogólniej, dowolne odwzorowanie p-liniowe $V_1 \times \ldots \times V_p \to \mathfrak{K}$ nazywamy też *formą p-liniową* na $V_1 \times \ldots \times V_p$. Jeśli np. $l^i : \mathbf{v}_i \mapsto l^i(\mathbf{v}_i)$ dla $i = 1, \ldots, p$ są formami liniowymi na V_i, to funkcja określona wzorem

$$f(\mathbf{v}_1, \ldots, \mathbf{v}_p) = l^1(\mathbf{v}_1) \ldots l^p(\mathbf{v}_p)$$

jest formą p-liniową na $V_1 \times \ldots \times V_p$, zwaną ***iloczynem tensorowym*** form liniowych $l^1, \ldots, l^p$. Oznaczamy ją przez $f = l^1 \otimes \ldots \otimes l^p$ (porządek czynników jest istotny).

Można wykazać, że każda forma p-liniowa na $V_1 \times \ldots \times V_p$ jest sumą iloczynów tensorowych form liniowych, ale na razie fakt ten nie będzie nam potrzebny. Jeśli $V_1 = \ldots = V_p = V$, to piszemy $V^p = V \times \ldots \times V$ (iloczyn kartezjański p egzemplarzy przestrzeni V). W tym wypadku wygodnie jest przyjąć oznaczenie

$$\mathcal{L}_p(V, \mathfrak{K}) = \mathcal{L}(V, \ldots, V; \mathfrak{K}).$$

Formę wieloliniową f na $V^p \times V^{*q}$ będziemy w przyszłości nazywali tensorem typu (p, q). Jak zauważyliśmy w § 3, refleksywność przestrzeni V oznacza, że tensory typu $(0, 1)$ (czyli formy liniowe) na V^* można uważać za wektory z V.

Formę p-liniową f na V^p (czyli tensor typu $(p, 0)$) nazywamy *symetryczną*, jeśli

$$f(\mathbf{v}_{\pi(1)}, \ldots, \mathbf{v}_{\pi(p)}) = f(\mathbf{v}_1, \ldots, \mathbf{v}_p)$$

dla dowolnych $\mathbf{v}_1, \ldots, \mathbf{v}_p \in V$ i każdej permutacji $\pi \in S_p$. Jeśli natomiast

$$f(\mathbf{v}_{\pi(1)}, \ldots, \mathbf{v}_{\pi(p)}) = \varepsilon_\pi f(\mathbf{v}_1, \ldots, \mathbf{v}_p),$$

gdzie ε_π oznacza znak permutacji π, to formę f nazywamy *antysymetryczną*. W przypadku $\dim V = p$ znamy już przykład formy antysymetrycznej — jest to wyznacznik macierzy A, rozpatrywany jako funkcja jej wierszy lub kolumn.

Celem tych ogólnych definicji jest ułożenie w pewien wspólny schemat wielu szczególnych pojęć, z którymi się już spotkaliśmy i z którymi będziemy mieć jeszcze do czynienia. Tensorami w ogólnym przypadku będziemy się zajmować dopiero w rozdziale 6.

2. Formy dwuliniowe. Ograniczymy się teraz do przypadku $V_1 = V_2 = V$ i będziemy mówić o formach *dwuliniowych* ($p = 2$) na przestrzeni V (choć należałoby właściwie mówić o formach na V^2). Zgodnie z ogólną definicją forma dwuliniowa f na przestrzeni V nad $\mathfrak{K}$ charakteryzuje się spełnieniem równości

$$\begin{aligned} f(\alpha\mathbf{u} + \beta\mathbf{v}, \mathbf{w}) &= \alpha f(\mathbf{u}, \mathbf{w}) + \beta f(\mathbf{v}, \mathbf{w}), \\ f(\mathbf{w}, \alpha\mathbf{u} + \beta\mathbf{v}) &= \alpha f(\mathbf{w}, \mathbf{u}) + \beta f(\mathbf{w}, \mathbf{v}) \end{aligned} \tag{2}$$

dla dowolnych $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$ oraz $\alpha, \beta \in \mathfrak{K}$. Zauważmy, że na ogół $f(\mathbf{u}, \mathbf{v}) \neq f(\mathbf{v}, \mathbf{u})$.

Załóżmy teraz, że przestrzeń V jest skończenie wymiarowa. Jeśli wybierzemy w V bazę $(\mathbf{e}_1, \ldots, \mathbf{e}_n)$ i zapiszemy wektory $\mathbf{x}, \mathbf{y} \in V$ w tej bazie:

$$\mathbf{x} = x_1\mathbf{e}_1 + \ldots + x_n\mathbf{e}_n, \qquad y = y_1\mathbf{e}_1 + \ldots + y_n\mathbf{e}_n,$$

to własności (2) pozwalają na wyrażenie wartości $f(\mathbf{u}, \mathbf{v})$ poprzez n^2 skalarów $f(\mathbf{e}_i, \mathbf{e}_j)$:

$$\begin{aligned} f(\mathbf{x}, \mathbf{y}) &= f\Big(\sum_i x_i\mathbf{e}_i, \sum_j y_j\mathbf{e}_j\Big) = \sum_i x_i f\Big(\mathbf{e}_i, \sum_j y_j\mathbf{e}_j\Big) \\ &= \sum_i x_i \sum_j y_j f(\mathbf{e}_i, \mathbf{e}_j) = \sum_{i,j} f_{ij} x_i y_j, \quad \text{gdzie} \quad f_{ij} = f(\mathbf{e}_i, \mathbf{e}_j). \end{aligned} \tag{3}$$

Macierz $F = (f_{ij})$ nazywamy *macierzą formy dwuliniowej f w bazie* $(\mathbf{e}_1, \ldots, \mathbf{e}_n)$. Jeśli oznaczymy przez $X = [x_1, \ldots, x_n]$ i $Y = [y_1, \ldots, y_n]$ kolumny współrzędnych wektorów $\mathbf{x}$ i $\mathbf{y}$ (będące macierzami $n \times 1$), a przez ${}^tX = (x_1, \ldots, x_n)$ macierz transponowaną do X (wymiarów $1 \times n$, czyli wiersz), to równość (3) można przepisać w postaci

$$f(\mathbf{x}, \mathbf{y}) = {}^tX \cdot F \cdot Y, \tag{4}$$

gdzie kropka oznacza mnożenie macierzy. Wystarczy się w tym celu posłużyć znanymi regułami mnożenia macierzy o wymiarach $1 \times n$, $n \times n$, $n \times 1$.

Na odwrót, mając macierz kwadratową $F = (f_{ij})$, można określić wzorem (4) lub (3) formę dwuliniową f na V, spełniającą warunek $f(\mathbf{e}_i, \mathbf{e}_j) = f_{ij}$. Jeśli więc w przestrzeni liniowej V nad $\mathfrak{K}$ wybraliśmy bazę $(\mathbf{e}_1, \ldots, \mathbf{e}_n)$, to istnieje odpowiedniość wzajemnie jednoznaczna pomiędzy macierzami kwadratowymi stopnia

n nad $\mathfrak{K}$ i formami dwuliniowymi na V. Odpowiedniość ta jest w istocie izomorfizmem przestrzeni liniowej $\mathcal{L}_2(V, \mathfrak{K})$ wszystkich form dwuliniowych $V \times V \to \mathfrak{K}$ (jeśli $f, g \in \mathcal{L}_2(V, \mathfrak{K})$, to łatwo sprawdzić, że $\alpha f + \beta g \in \mathcal{L}_2(V, \mathfrak{K})$ dla dowolnych $\alpha, \beta \in \mathfrak{K}$) na przestrzeń liniową $M_n(\mathfrak{K})$ wszystkich macierzy kwadratowych stopnia n nad $\mathfrak{K}$. Istotnie, jeśli

$$f(\mathbf{x}, \mathbf{y}) = {}^tXFY, \qquad g(\mathbf{x}, \mathbf{y}) = {}^tXGY,$$

to

$$\alpha f(\mathbf{x}, \mathbf{y}) + \beta g(\mathbf{x}, \mathbf{y}) = {}^tX \cdot (\alpha F + \beta G) \cdot Y.$$

3. Transformacja macierzy formy dwuliniowej. Aksjomatyczna definicja formy dwuliniowej poprzez własności (2) nie zależy od wyboru jakiejkolwiek bazy w V. Aby zapis macierzowy formy f przedstawiał rzeczywistą wartość, trzeba uzupełnić odpowiedniość $f \mapsto F$ przez podanie reguły transformacji macierzy F przy przejściu od jednej bazy do drugiej. Przypuśćmy, że w przestrzeni V oprócz bazy $(\mathbf{e}_1, \ldots, \mathbf{e}_n)$ dana jest jeszcze baza $(\mathbf{e}'_1, \ldots, \mathbf{e}'_n)$ o macierzy przejścia $A = (a_{ij})$:

$$\mathbf{e}'_j = \sum_{i=1}^{n} a_{ij}\mathbf{e}_i, \qquad j = 1, \ldots, n.$$

Jeśli $x_1\mathbf{e}_1 + \ldots + x_n\mathbf{e}_n = \mathbf{x} = x'_1\mathbf{e}'_1 + \ldots + x'_n\mathbf{e}'_n$, to kolumny współrzędnych X i X' związane są zależnością $X = A \cdot X'$. Niech teraz $F = (f_{ij})$ będzie macierzą formy dwuliniowej f w bazie $(\mathbf{e}_i)$, a $F' = (f'_{ij})$ — macierzą tej samej formy w bazie $\mathbf{e}'_i$, tj. $f_{ij} = f(\mathbf{e}_i, \mathbf{e}_j)$ oraz $f'_{ij} = f(\mathbf{e}'_i, \mathbf{e}'_j)$. Ponieważ ${}^t(AX') = {}^t(X') \cdot {}^tA$ i wartość $f(\mathbf{x}, \mathbf{y})$ nie zależy od wyboru bazy, więc

$$\begin{aligned} {}^tX' \cdot F' \cdot Y' = f(\mathbf{x}, \mathbf{y}) &= {}^tX \cdot F \cdot Y = {}^t(AX') \cdot F \cdot (AY') \\ &= {}^tX' \cdot {}^tA \cdot F \cdot A \cdot Y' \end{aligned}$$

dla dowolnych wektorów X', Y', skąd otrzymujemy

TWIERDZENIE 1. *Macierze F i F' formy dwuliniowej f na przestrzeni V w bazach odpowiednio $(\mathbf{e}_i)$ i $(\mathbf{e}'_i)$ związane są zależnością*

$$F' = {}^tA \cdot F \cdot A, \tag{5}$$

gdzie A jest macierzą przejścia od $(\mathbf{e}_i)$ do $(\mathbf{e}'_i)$. ■

DEFINICJA 1. *Rzędem* formy dwuliniowej f (oznaczenie: $\operatorname{rank} f$) nazywamy rząd macierzy F odpowiadającej formie f w jakiejkolwiek bazie $(\mathbf{e}_i)$. Jeśli macierz F jest nieosobliwa, to formę f nazywamy *niezdegenerowaną*.

WNIOSEK. *Rząd jest niezmiennikiem formy dwuliniowej, tzn. nie zależy od wyboru bazy.*

Dowód. Rząd macierzy nie ulega zmianie przy pomnożeniu jej przez macierz nieosobliwą (część I, rozdz. 2, § 3, wniosek 1 z twierdzenia 5); wystarczy zastosować ten fakt do macierzy F, F' związanych zależnością (5). ∎

Inny dowód stwierdzenia o niezmienniczości rzędu formy dwuliniowej można otrzymać w następujący sposób. Oznaczmy przez L_f zbiór tych wszystkich $\mathbf{x} \in V$, dla których $f(\mathbf{x}, \mathbf{y}) = 0$ dla dowolnego $\mathbf{y} \in V$ (krótko piszemy: $f(\mathbf{x}, V) = 0$). Łatwo sprawdzić, że L_f jest podprzestrzenią liniową w V. Nazywamy ją *jądrem lewostronnym* formy dwuliniowej f. Liczba $\dim L_f$ zależy oczywiście tylko od f. Niech $(\mathbf{e}_1, \ldots, \mathbf{e}_n)$ będzie bazą w V. Warunek $\mathbf{x} \in L_f$ jest równoważny temu, że

$$f(\mathbf{x}, \mathbf{e}_1) = 0, \quad \ldots, \quad f(\mathbf{x}, \mathbf{e}_n) = 0.$$

Jest to układ liniowy jednorodny wyznaczony przez formy liniowe $\mathbf{x} \mapsto f_j(\mathbf{x}) = f(\mathbf{x}, \mathbf{e}_j)$ dla $j = 1, \ldots, n$. Współrzędnymi formy f_j w bazie dualnej $(e^1, \ldots, e^n)$ przestrzeni V^* są skalary $f_j(\mathbf{e}_i)$, czyli wyrazy $f(\mathbf{e}_i, \mathbf{e}_j) = f_{ij}$ j-tej kolumny macierzy F. Tak więc rząd układu form liniowych $f_1, \ldots, f_n \in V^*$ jest rzędem macierzy F, i jeśli wynosi on r, to zachodzi równość $\dim L_f = n - r$ (§ 3, twierdzenie 4). Inaczej mówiąc,

$$r = \dim V - \dim L_f,$$

a ta wielkość nie zależy od wyboru bazy.

4. Formy symetryczne i antysymetryczne. Zgodnie z punktem 1 formę dwuliniową $f : V \times V \to \mathfrak{K}$ nazywamy *symetryczną*, jeśli $f(\mathbf{x}, \mathbf{y}) = f(\mathbf{y}, \mathbf{x})$ dla dowolnych $\mathbf{x}, \mathbf{y} \in V$, i *antysymetryczną*, jeśli $f(\mathbf{x}, \mathbf{y}) = -f(\mathbf{y}, \mathbf{x})$. Terminologia ta jest zgodna z pojęciami funkcji symetrycznych i antysymetrycznych (część I), a także z pojęciami *macierzy symetrycznej* $A = (a_{ij})$, czyli spełniającej warunek ${}^tA = A$, oraz *antysymetrycznej*, czyli takiej, że ${}^tA = -A$. Istotnie, ponieważ $f(\mathbf{y}, \mathbf{x}) = {}^t[f(\mathbf{y}, \mathbf{x})]$ (transpozycja macierzy 1×1, czyli skalara), więc z równości $f(\mathbf{x}, \mathbf{y}) = \varepsilon f(\mathbf{y}, \mathbf{x})$, gdzie $\varepsilon = \pm 1$, oraz z zależności (4) wynika, że

$$\begin{aligned} {}^tX \cdot F \cdot Y = f(\mathbf{x}, \mathbf{y}) &= \varepsilon f(\mathbf{y}, \mathbf{x}) = \varepsilon \cdot {}^tf(\mathbf{y}, \mathbf{x}) \\ &= \varepsilon \cdot {}^t({}^tY \cdot F \cdot X) = \varepsilon \cdot {}^tX \cdot {}^tF \cdot Y \end{aligned}$$

dla dowolnych X, Y, a stąd ${}^tF = \varepsilon F$. Na odwrót, jeśli ${}^tF = \varepsilon F$ i $\varepsilon = \pm 1$, to forma dwuliniowa f o macierzy F spełnia warunek $f(\mathbf{x}, \mathbf{y}) = \varepsilon f(\mathbf{y}, \mathbf{x})$.

Tak więc forma dwuliniowa f jest symetryczna lub antysymetryczna dokładnie wtedy, gdy macierz F tej formy w dowolnej bazie jest odpowiednio symetryczna lub antysymetryczna.

TWIERDZENIE 2. *Jeśli* $\operatorname{char} \mathfrak{K} \neq 2$, *to przestrzeń* $\mathcal{L}_2(V, \mathfrak{K})$ *form dwuliniowych jest sumą prostą*

$$\mathcal{L}_2(V, \mathfrak{K}) = \mathcal{L}_2^+(V, \mathfrak{K}) \oplus \mathcal{L}_2^-(V, \mathfrak{K})$$

podprzestrzeni $\mathcal{L}_2^+(V, \mathfrak{K})$ *form symetrycznych i* $\mathcal{L}_2^-(V, \mathfrak{K})$ *form antysymetrycznych.*

Dowód. Jeśli $f \in \mathcal{L}_2^+(V, \mathfrak{K}) \cap \mathcal{L}_2^-(V, \mathfrak{K})$, to dla dowolnych $\mathbf{x}, \mathbf{y} \in V$ zachodzi równość

$$f(\mathbf{x}, \mathbf{y}) = f(\mathbf{y}, \mathbf{x}) = -f(\mathbf{x}, \mathbf{y}),$$

czyli $2f(\mathbf{x}, \mathbf{y}) = 0$, a stąd $f(\mathbf{x}, \mathbf{y}) = 0$ wobec założenia, że $\operatorname{char} \mathfrak{K} \neq 2$. Tak więc suma algebraiczna $\mathcal{L}_2^+ + \mathcal{L}_2^-$ jest prosta.

Z drugiej strony, równość

$$f(\mathbf{x}, \mathbf{y}) = \frac{1}{2}\{f(\mathbf{x}, \mathbf{y}) + f(\mathbf{y}, \mathbf{x})\} + \frac{1}{2}\{f(\mathbf{x}, \mathbf{y}) - f(\mathbf{y}, \mathbf{x})\}$$

lub też odpowiadająca jej równość macierzowa

$$F = \frac{1}{2}(F + {}^tF) + \frac{1}{2}(F - {}^tF)$$

pokazują, że każdą formę dwuliniową f można przedstawić w postaci sumy formy symetrycznej i antysymetrycznej. ■

Dla $\mathfrak{K} = Z_2$ każda macierz symetryczna jest jednocześnie antysymetryczna, a więc teza twierdzenia przestaje zachodzić, ponieważ np. macierz $\left(\begin{smallmatrix} 0 & 1 \\ 0 & 1 \end{smallmatrix}\right)$ nie jest symetryczna. Istnieje jeszcze pojęcie *formy alternującej*, tj. formy dwuliniowej spełniającej warunek $f(\mathbf{u}, \mathbf{u}) = 0$ dla każdego $\mathbf{u} \in V$. Gdy $\operatorname{char} \mathfrak{K} \neq 2$, pojęcie to jest jednak identyczne z pojęciem formy antysymetrycznej (Czytelnik powinien to sprawdzić).

W dalszym ciągu tego rozdziału zakładamy milcząco, że $\operatorname{char} \mathfrak{K} \neq 2$.

5. Formy kwadratowe. Z formami dwuliniowymi symetrycznymi wiąże się ważne pojęcie formy kwadratowej, pojawiające się w naturalny sposób w różnych działach matematyki.

DEFINICJA 2. *Formą kwadratową* na przestrzeni liniowej V nad ciałem $\mathfrak{K}$ nazywamy każdą funkcję $q : V \to \mathfrak{K}$ o następujących własnościach:

(i) $q(-\mathbf{v}) = q(\mathbf{v})$ dla każdego $\mathbf{v} \in V$;

(ii) odwzorowanie $f : V \times V \to \mathfrak{K}$ określone wzorem

$$f(\mathbf{x}, \mathbf{y}) = \frac{1}{2}\{q(\mathbf{x} + \mathbf{y}) - q(\mathbf{x}) - q(\mathbf{y})\}, \tag{6}$$

jest symetryczną formą dwuliniową na V. Jeśli przestrzeń V jest skończenie wymiarowa, to rząd formy f nazywamy *rzędem* formy kwadratowej q: $\operatorname{rank} q = \operatorname{rank} f$. Mówimy, że symetryczna forma dwuliniowa f powstaje przez *polaryzację* formy kwadratowej q lub też jest *formą biegunową* formy q.

Niech teraz f będzie dowolną symetryczną formą dwuliniową na V. Przyjmując

$$q_f(\mathbf{x}) = f(\mathbf{x}, \mathbf{x}), \tag{7}$$

otrzymujemy funkcję $q_f : V \to \mathfrak{K}$ spełniającą warunki (i) i (ii) w definicji formy kwadratowej, ponieważ $f(-\mathbf{x}, -\mathbf{x}) = f(\mathbf{x}, \mathbf{x})$ oraz

$$f(\mathbf{x}, \mathbf{y}) = \frac{1}{2}\{f(\mathbf{x}+\mathbf{y}, \mathbf{x}+\mathbf{y}) - f(\mathbf{x}, \mathbf{x}) - f(\mathbf{y}, \mathbf{y})\}. \tag{8}$$

Można by pomyśleć, że q_f to forma kwadratowa specjalnego typu; okazuje się jednak, że każda forma kwadratowa jest tej postaci (poniższy prosty dowód można pominąć bez szkody dla rozumienia dalszego tekstu):

TWIERDZENIE 3. *Każda forma kwadratowa q jest wyznaczona jednoznacznie przez swoją formę biegunową f; inaczej mówiąc, $q = q_f$.*

Dowód. Przyjmując we wzorze (6) $\mathbf{y} = -\mathbf{x}$, otrzymujemy

$$-f(\mathbf{x}, \mathbf{x}) = \frac{1}{2}\{q(\mathbf{0}) - q(\mathbf{x}) - q(-\mathbf{x})\},$$

a stąd

$$q(\mathbf{x}) = f(\mathbf{x}, \mathbf{x}) + \frac{1}{2}q(\mathbf{0}).$$

Ponieważ f jest formą dwuliniową, więc $f(\mathbf{0}, \mathbf{0}) = 0$. Zatem dla $\mathbf{x} = \mathbf{0}$ otrzymujemy $q(\mathbf{0}) = \frac{1}{2}q(\mathbf{0})$, czyli $q(\mathbf{0}) = 0$. Stąd $q(\mathbf{x}) = f(\mathbf{x}, \mathbf{x})$. ∎

DEFINICJA 3. *Macierzą formy kwadratowej $q = q_f$* w bazie $(\mathbf{e}_1, \ldots, \mathbf{e}_n)$ przestrzeni V nazywamy macierz F formy dwuliniowej f, biegunowej dla q.

Tak więc $F = (f_{ij})$, gdzie

$$f_{ij} = \frac{1}{2}\{q(\mathbf{e}_i + \mathbf{e}_j) - q(\mathbf{e}_i) - q(\mathbf{e}_j)\}, \qquad i, j = 1, \ldots, n.$$

Na odwrót, każdej macierzy symetrycznej $F = (f_{ij})$ odpowiada forma kwadratowa q, określona wzorem

$$q(\mathbf{x}) = {}^tX \cdot F \cdot X = \sum_{i,j} f_{ij} x_i x_j. \tag{9}$$

Widzimy więc, że zgodnie ze swoją nazwą forma kwadratowa jest formą jednorodną stopnia 2 względem współrzędnych $x_1, \ldots, x_n$ wektora $\mathbf{x} = x_1\mathbf{e}_1 + \ldots + x_n\mathbf{e}_n$. Przypadek, gdy macierz F jest diagonalna, zasługuje na specjalną nazwę.

DEFINICJA 4. Mówimy, że forma kwadratowa q ma w bazie $(\mathbf{e}_1, \ldots, \mathbf{e}_n)$ przestrzeni V *postać diagonalną* (lub *kanoniczną*), jeśli dla każdego wektora $\mathbf{x} = \sum x_i \mathbf{e}_i$ wartość $q(\mathbf{x})$ wyraża się wzorem

$$q(\mathbf{x}) = \sum_i f_{ii} x_i^2.$$

Bazę $(\mathbf{e}_i)$ nazywamy wówczas również *bazą kanoniczną* dla q.

Analogiczną terminologię stosujemy dla formy dwuliniowej f, biegunowej dla q:

$$f(\mathbf{x},\mathbf{y}) = \sum_i f_{ii} x_i y_i.$$

Postać (ani baza) kanoniczna nie jest określona jednoznacznie — np. każda permutacja bazy kanonicznej jest też oczywiście bazą kanoniczną.

Zauważmy, że w bazie kanonicznej rząd $\operatorname{rank} q_f = \operatorname{rank} f$ jest po prostu liczbą niezerowych współczynników f_{ii}. Z drugiej strony, zgodnie z uwagą w końcu punktu 3, $\operatorname{rank} q = \dim V - \dim L_q$, gdzie $L_q = L_f$ jest jądrem formy f (lewostronnym lub prawostronnym, na mocy symetryczności f). Podprzestrzeń $L_q \subset V$, zwana też *podprzestrzenią zerową* (lub *izotropową*) formy q, ma następującą definicję w terminach q:

$$L_q = \{\mathbf{u} \in V \mid q(\mathbf{u}+\mathbf{v}) = q(\mathbf{u}) + q(\mathbf{v})\ \forall \mathbf{v} \in V\}.$$

Rząd formy q jest więc wielkością niezmienniczą.

6. Postać kanoniczna formy kwadratowej. Możliwość wyboru bazy, w której dana forma kwadratowa miałaby najprostszą (tj. kanoniczną) postać, ma ważne znaczenie teoretyczne i praktyczne.

TWIERDZENIE 4. *Dla każdej symetrycznej formy dwuliniowej f na przestrzeni n-wymiarowej V istnieje baza kanoniczna.*

Dowód. Stosujemy indukcję względem n; przypadek $n = 1$ jest oczywisty. Jeśli $f(\mathbf{x},\mathbf{y}) = 0$ dla dowolnych $\mathbf{x},\mathbf{y} \in V$ (tj. $f = 0$), to teza jest oczywista: każda baza jest kanoniczna. Jeśli $f \neq 0$, to odpowiadająca f forma kwadratowa q też nie jest zerem (wzory (6), (8) lub twierdzenie 3). Niech $\mathbf{e}_1$ będzie takim wektorem, że $f(\mathbf{e}_1,\mathbf{e}_1) = q(\mathbf{e}_1) \neq 0$. Wtedy forma liniowa $f_1 : \mathbf{x} \mapsto f(\mathbf{x},\mathbf{e}_1)$ jest niezerowa ($f_1(\mathbf{e}_1) \neq 0$). Wobec tego podprzestrzeń

$$L = \operatorname{Ker} f_1 = \{\mathbf{x} \in V \mid f_1(\mathbf{x}) = 0\}$$

ma wymiar $n-1$, czyli jest hiperpłaszczyzną (§ 3, twierdzenie 4). Na mocy założenia indukcyjnego w L istnieje baza $(\mathbf{e}_2,\ldots,\mathbf{e}_n)$, w której macierz formy f obciętej do L jest diagonalna, tj.

$$f(\mathbf{e}_i,\mathbf{e}_j) = 0 \quad \text{dla } i \neq j, \quad i,j = 2,\ldots,n.$$

Ponieważ na mocy konstrukcji $f(\mathbf{e}_i,\mathbf{e}_1) = 0$ dla $i = 2,\ldots,n$, więc otrzymaliśmy równości $f(\mathbf{e}_i,\mathbf{e}_j) = 0$ dla $i \neq j$, charakteryzujące bazę kanoniczną, o ile tylko wykażemy, że układ $\mathbf{e}_1,\ldots,\mathbf{e}_n$ jest liniowo niezależny. Gdyby tak nie było, to w dowolnej zależności liniowej

$$\alpha_1\mathbf{e}_1 + \alpha_2\mathbf{e}_2 + \ldots + \alpha_n\mathbf{e}_n = \mathbf{0}$$

współczynnik α_1 byłby różny od zera, gdyż $(\mathbf{e}_2, \ldots, \mathbf{e}_n)$ jest bazą w L. Wówczas jednak $\mathbf{e}_1 = \sum_{i>1} \beta_i \mathbf{e}_i$ oraz

$$0 \neq f_1(\mathbf{e}_1) = f_1\Big(\sum_{i>1} \beta_i \mathbf{e}_i\Big) = \sum_{i>1} \beta_i f_1(\mathbf{e}_i) = 0.$$

Sprzeczność ta kończy dowód twierdzenia. ■

Wniosek 1. *Niech q będzie formą kwadratową rzędu r na przestrzeni liniowej V wymiaru n. Wtedy w V istnieje baza $(\mathbf{e}_i)$, w której q przyjmuje postać kanoniczną*

$$q(\mathbf{x}) = \lambda_1 x_1^2 + \lambda_1 x_2^2 + \ldots + \lambda_r x_r^2, \qquad \lambda_1 \ldots \lambda_r \neq 0. \; ■ \tag{10}$$

Wniosek 2. *Dla każdej macierzy symetrycznej F istnieje macierz nieosobliwa A taka, że ${}^t\!AFA$ jest macierzą diagonalną tego samego rzędu co F.* ■

Rozpatrzona powyżej indukcyjna metoda sprowadzania form dwuliniowych (a więc i kwadratowych) do postaci kanonicznej pochodzi od Lagrange'a (1736–1813). W praktyce, ma się rozumieć, metodę tę stosuje się do zapisu formy we współrzędnych, działając w nieco innej kolejności. Wychodzimy od postaci (9), interpretowanej jako wielomian jednorodny stopnia 2 względem n zmiennych niezależnych:

$$q(x_1, \ldots, x_n) := q(\mathbf{x}) = \sum_{i,j=1}^{n} f_{ij} x_i x_j.$$

Będziemy usuwali człony mieszane $x_i x_j$ $(i \neq j)$ znaną już w starożytnej Babilonii metodą uzupełniania do pełnego kwadratu. Wydzielmy wszystkie człony zawierające współrzędną x_1:

$$q(x_1, \ldots, x_n) = f_{11} x_1^2 + 2f_{12} x_1 x_2 + 2f_{13} x_1 x_3 + \ldots + 2f_{1n} x_1 x_n + \sum_{i,j \neq 1} f_{ij} x_i x_j$$

(powinny wystąpić sumy $f_{1j} x_1 x_j + f_{j1} x_j x_1$, ale $f_{j1} = f_{1j}$ i stąd się biorą podwojone iloczyny). Załóżmy najpierw, że $f_{11} \neq 0$, i uzupełnijmy do pełnego kwadratu, modyfikując współczynniki przy wyrazach niezawierających x_1:

$$q(x_1, \ldots, x_n) = \frac{1}{f_{11}} (f_{11} x_1 + f_{12} x_2 + \ldots + f_{1n} x_n)^2 + \sum_{i,j \neq 1} f'_{ij} x_i x_j.$$

Jeśli przyjmiemy teraz

$$x'_1 = f_{11} x_1 + f_{12} x_2 + \ldots + f_{1n} x_n, \qquad x'_i = x_i, \qquad i > 1,$$

to doprowadziliśmy formę q do postaci

$$q(x_1, \ldots, x_n) = \frac{1}{f_{11}} (x'_1)^2 + q'(x'_2, \ldots, x'_n),$$

gdzie $q'(x_2', \ldots, x_n') = \sum_{i,j=2}^{n} f_{ij}' x_i' x_j'$ jest formą kwadratową mniejszej liczby zmiennych. Jeśli $f_{22}' \neq 0$, to przepisujemy formę q' w postaci

$$\begin{aligned} q'(x_2', \ldots, x_n') &= f_{22}'(x_2')^2 + f_{23}' x_2' x_3' + \ldots + f_{2n}' x_2' x_n' + \sum_{i,j>2} f_{ij}' x_i' x_j' \\ &= \frac{1}{f_{22}'}(f_{22}' x_2' + f_{23}' x_3' + \ldots + f_{2n}' x_n')^2 + \sum_{i,j>2} f_{ij}'' x_i' x_j' \end{aligned}$$

(przejście od f_{ij}' do f_{ij}'' wiąże się z wydzieleniem nowego pełnego kwadratu). Kolejna zmiana zmiennych

$$x_1'' = x_1', \qquad x_2'' = f_{22}' x_2' + f_{23}' x_3' + \ldots + f_{2n}' x_n', \qquad x_i'' = x_i', \qquad i > 2,$$

prowadzi do wyrażenia

$$q(\mathbf{x}) = \frac{1}{f_{11}}(x_1'')^2 + \frac{1}{f_{22}'}(x_2'')^2 + q''(x_3'', \ldots, x_n''),$$

gdzie $q''(x_3'', \ldots, x_n'') = \sum_{i,j=3}^{n} f_{ij}'' x_i'' x_j''$ jest formą jeszcze mniejszej liczby zmiennych.

Procedura ta kończy się oczywiście na zapisaniu $q(\mathbf{x})$ w postaci kombinacji liniowej $r = \operatorname{rank} q$ kwadratów. Zamiany zmiennych dokonywane po drodze są nieosobliwe i odpowiadają przejściu do nowych baz. Jeden szczegół wymaga komentarza: założenia $f_{11} \neq 0$, $f_{22} \neq 0$, ... mogą się wydać ograniczające. Jeśli jednak $f_{11} = 0$, ale $f_{kk} \neq 0$ dla pewnego k, to wystarczy zamienić miejscami zmienne x_1, x_k (lub – co na jedno wychodzi – inaczej ponumerować wektory bazowe). Jeśli zaś forma $q \neq 0$ nie zawiera żadnego kwadratu, tj. $f_{kk} = 0$ dla każdego k, to można przyjąć, że np. $2f_{12}x_1x_2 \neq 0$; w tym wypadku zamiana zmiennych

$$x_1 = x_1' + x_2', \qquad x_2 = x_1' - x_2', \qquad x_k = x_k', \qquad k > 2,$$

prowadzi do pojawienia się członu $2f_{12}((x_1')^2 - (x_2')^2)$, co pozwala rozpocząć opisaną procedurę.

7. Formy kwadratowe rzeczywiste. Ogólnie rzecz biorąc, jeśli działamy nad dowolnym ciałem $\mathfrak{K}$ (spełniającym warunek $\operatorname{char} \mathfrak{K} \neq 2$), to prostszej postaci formy kwadratowej niż postać diagonalna nie da się już osiągnąć. Jeśli jednak $\mathfrak{K} = \mathbb{R}$, to możemy doprowadzić do sytuacji, w której wszystkie współczynniki w (10) są równe ± 1. Istotnie, po odpowiednim przestawieniu wektorów bazowych można uważać, że pierwsze s współczynników $\lambda_1, \ldots, \lambda_s$ w (10) jest dodatnich, a pozostałe są ujemne. Wówczas po zamianie zmiennych

$$x_i' = \sqrt{\lambda_i} \cdot x_i, \qquad i = 1, \ldots, s; \qquad x_i' = \sqrt{-\lambda_i} \cdot x_i, \qquad i = s+1, \ldots, r,$$
$$x_i' = x_i, \qquad i = r+1, \ldots, n,$$

otrzymamy $q(\mathbf{x}) = \sum_{i=1}^{s}(x_i')^2 - \sum_{i=s+1}^{r}(x_i')^2$.

DEFINICJA 5. Forma kwadratowa q jest *w postaci normalnej*, jeśli

$$q(\mathbf{x}) = x_1^2 + \ldots + x_s^2 - x_{s+1}^2 - \ldots - x_r^2. \tag{11}$$

Dla $\mathfrak{K} = \mathbb{R}$ mamy więc silniejszy wariant wniosku 1 z twierdzenia 4:

WNIOSEK 1′. *Każdą formę kwadratową na rzeczywistej przestrzeni liniowej można sprowadzić do postaci normalnej.* ∎

W ten sposób, oprócz rzędu, pojawiła się jeszcze jedna charakterystyka liczbowa formy kwadratowej na przestrzeni rzeczywistej — liczba s współczynników równych 1 w postaci normalnej. Okazuje się, że liczba ta również nie zależy od sposobu sprowadzenia formy do postaci normalnej.

TWIERDZENIE 5 (prawo bezwładności form kwadratowych). *Niech q będzie formą kwadratową na n-wymiarowej przestrzeni liniowej V nad $\mathbb{R}$. Wówczas liczby całkowite r i s ($s \leqslant r \leqslant n$) występujące w postaci normalnej* (11) *zależą tylko od formy q.*

Dowód. Niezmienniczość r jest nam już znana; trzeba się tylko przekonać o niezmienniczości s (tj. niezależności od wyboru bazy). Przypuśćmy, że w pewnej innej bazie $(\mathbf{e}'_1, \ldots, \mathbf{e}'_n)$ forma q przyjmuje postać normalną

$$q(\mathbf{x}) = (x'_1)^2 + \ldots + (x'_t)^2 - (x'_{t+1})^2 - \ldots - (x'_r)^2 \tag{11′}$$

o t członach dodatnich ($\mathbf{x} = \sum_{i=1}^n x_i \mathbf{e}_i = \sum_{i=1}^n x'_i \mathbf{e}'_i$), przy czym $t \neq s$. Możemy przyjąć, że $t < s$.

Rozważmy w V podprzestrzenie

$$L = \langle \mathbf{e}_1, \ldots, \mathbf{e}_s \rangle, \qquad L' = \langle \mathbf{e}'_{t+1}, \ldots, \mathbf{e}'_n \rangle.$$

Ponieważ $\dim(L + L') \leqslant \dim V = n$, więc z twierdzenia 6 w § 2 wynika, że

$$\begin{aligned}\dim(L \cap L') &= \dim L + \dim L' - \dim(L + L') \\ &\geqslant s + (n - t) - n = s - t > 0.\end{aligned}$$

Istnieje więc niezerowy wektor $\mathbf{a} \in L \cap L'$:

$$\mathbf{0} \neq \mathbf{a} = a_1\mathbf{e}_1 + \ldots + a_s\mathbf{e}_s = a'_{t+1}\mathbf{e}'_{t+1} + \ldots + a'_n\mathbf{e}'_n.$$

Na mocy (11) mamy

$$q(\mathbf{a}) = a_1^2 + \ldots + a_s^2 > 0.$$

Ale z (11′) wynika, że

$$q(\mathbf{a}) = -(a'_{t+1})^2 - \ldots - (a'_r)^2 \leqslant 0$$

(może się zdarzyć, że $r < n$ i $a'_{t+1} = \ldots = a'_r = 0$). Otrzymana sprzeczność dowodzi, że $t = s$. ∎

W związku z twierdzeniem 5 niezmienniki liczbowe formy kwadratowej mają specjalne nazwy.

DEFINICJA 6. Liczbę s kwadratów ze znakiem plus w postaci normalnej nazywamy *dodatnim indeksem bezwładności* rzeczywistej formy kwadratowej, a liczbę $r-s$ kwadratów ze znakiem minus — *ujemnym indeksem bezwładności*. Przez *sygnaturę* formy kwadratowej rozumie się albo parę $(s, r-s)$, albo różnicę $2s-r$ liczby kwadratów ze znakiem plus i kwadratów ze znakiem minus.

Prawo bezwładności form kwadratowych, pochodzące od Jamesa Sylvestera (1814–1897), ma swoje źródło w mechanice. Dla form kwadratowych na przestrzeniach zespolonych pojęcia dodatniego i ujemnego indeksu bezwładności tracą oczywiście sens, ponieważ wszystkie niezerowe współczynniki λ_i w postaci diagonalnej (10) można uczynić równymi 1 lub wszystkie równymi -1.

8. Formy i macierze dodatnio określone. Niech V będzie rzeczywistą przestrzenią liniową. Formę kwadratową q na V nazywamy *niezdegenerowaną*, jeśli $\operatorname{rank} q = \dim_{\mathbb{R}} V$, czyli rząd formy jest równy wymiarowi przestrzeni.

DEFINICJA 7. Niezdegenerowaną formę kwadratową $q : V \to \mathbb{R}$ nazywamy *dodatnio określoną* (odpowiednio *ujemnie określoną*) lub po prostu *dodatnią* (odpowiednio *ujemną*), jeśli $q(\mathbf{x}) > 0$ (odpowiednio $q(\mathbf{x}) < 0$) dla każdego wektora $\mathbf{x} \neq \mathbf{0}$. Formę q nazywamy *dodatnio półokreśloną* (lub *nieujemną*), jeśli $q(\mathbf{x}) \geqslant 0$ dla każdego $\mathbf{x} \in V$. Wreszcie forma q jest *nieokreślona*, jeśli przyjmuje zarówno wartości dodatnie, jak i ujemne.

Warto zwrócić uwagę, że pojęcia te nie zależą od wyboru bazy. Odpowiednie postaci normalne dla $n = \dim_{\mathbb{R}} V$ są następujące:

- $x_1^2 + x_2^2 + \ldots + x_n^2$ dla formy dodatnio określonej;
- $-x_1^2 - x_2^2 - \ldots - x_n^2$ dla formy ujemnie określonej;
- $x_1^2 + x_2^2 + \ldots + x_r^2$, gdzie $r \leqslant n$, dla formy dodatnio półokreślonej;
- $x_1^2 + x_2^2 + \ldots + x_s^2 - x_{s+1}^2 - \ldots - x_r^2$, gdzie $0 < s < r$, dla formy nieokreślonej.

Jest oczywiste, że jeśli $q(\mathbf{x}) = \lambda_1 x_1^2 + \lambda_2 x_2^2 + \ldots + \lambda_n x_n^2$ jest jakąkolwiek postacią kanoniczną formy kwadratowej w bazie $(\mathbf{e}_i)$, to forma ta jest dodatnio określona wtedy i tylko wtedy, gdy wszystkie współczynniki λ_i są dodatnie: wystarczy zauważyć, że $\lambda_i = q(\mathbf{e}_i)$.

O formie biegunowej odpowiadającej dodatnio określonej formie kwadratowej również mówimy, że jest dodatnio określona. Terminologia ta przenosi się też na macierze: np. rzeczywistą macierz symetryczną F nazywamy *dodatnio określoną*, jeśli odpowiada ona dodatnio określonej formie kwadratowej. Ponieważ macierzą dodatnio określonej formy kwadratowej w postaci normalnej jest macierz jednostkowa, więc na mocy twierdzenia 1 zachodzi

TWIERDZENIE 6. *Rzeczywista macierz kwadratowa F jest dodatnio określona wtedy i tylko wtedy, gdy jest postaci*

$$F = {}^tA \cdot A, \tag{12}$$

gdzie A jest rzeczywistą macierzą nieosobliwą. ∎

Często zachodzi konieczność ustalenia, czy forma kwadratowa jest dodatnio określona, bezpośrednio na podstawie jej macierzy.

Przykład. Niech $\varphi(x, y)$ będzie funkcją rzeczywistą klasy C^2 dwóch zmiennych rzeczywistych w otoczeniu początku układu współrzędnych [1]. Pochodne cząstkowe będziemy oznaczali symbolami φ'_x, φ'_y itd. Załóżmy, że punkt $(0, 0)$ jest *punktem krytycznym* (lub *stacjonarnym*) funkcji φ, tzn. $\varphi'_x(0, 0) = 0 = \varphi'_y(0, 0)$. Wtedy na podstawie wzoru Taylora

$$\varphi(x, y) = \varphi(0, 0) + \frac{1}{2}\{ax^2 + 2bxy + cy^2\} + \ldots,$$

gdzie $a = \varphi''_{xx}(0, 0)$, $b = \varphi''_{xy}(0, 0)$ i $c = \varphi''_{yy}(0, 0)$, a wielokropek oznacza człony wyższego rzędu. W dostatecznie małym otoczeniu zera człony te można zaniedbać, tzn. wartość funkcji φ jest w przybliżeniu równa stałej $\varphi(0, 0)$ plus $\frac{1}{2}q(\mathrm{v})$ $(\mathrm{v} = x\mathrm{e}_1 + y\mathrm{e}_2)$, gdzie

$$q(\mathrm{v}) = ax^2 + 2bxy + cy^2$$

jest formą kwadratową na przestrzeni $V = \langle \mathrm{e}_1, \mathrm{e}_2 \rangle$. Jeśli $\operatorname{rank} q = 2$, to punkt krytyczny $(0, 0)$ nazywamy *niezdegenerowanym*. Jeśli forma q jest dodatnio określona, to φ ma oczywiście w punkcie $(0, 0)$ *minimum lokalne*. Maksimum odpowiada formie ujemnie określonej. Jeśli natomiast sygnatura formy q jest równa $(1, 1)$, to w punkcie $(0, 0)$ nie ma ani minimum, ani maksimum — $(0, 0)$ jest *punktem siodłowym* funkcji φ.

Zapisując $q(\mathrm{v})$ w postaci

$$q(\mathrm{v}) = a\left(x + \frac{by}{a}\right)^2 + \left(c - \frac{b^2}{a}\right)y^2, \qquad a \neq 0,$$

(lub w analogicznej postaci, gdy $a = 0$ i $c \neq 0$), widzimy, że układ nierówności

$$a > 0, \qquad \begin{vmatrix} a & b \\ b & c \end{vmatrix} > 0$$

stanowi prosty warunek konieczny i dostateczny dodatniej określoności formy q, a więc warunek dostateczny istnienia minimum lokalnego w punkcie $(0, 0)$.

---

[1] Szczegóły analityczne związane z tym przykładem można znaleźć w wielu podręcznikach analizy, np. W. Kołodziej, *Analiza matematyczna*, PWN, Warszawa 1978, s. 262 (*przyp. tłum.*).

W powyższych nierównościach występują wyznaczniki, którym w przypadku n-wymiarowym odpowiadają *minory główne*

$$\Delta_1 = f_{11}, \quad \Delta_2 = \begin{vmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{vmatrix}, \ \ldots, \ \Delta_k = \begin{vmatrix} f_{11} & f_{12} & \ldots & f_{1k} \\ f_{21} & f_{22} & \ldots & f_{2k} \\ \ldots & \ldots & \ldots & \ldots \\ f_{k1} & f_{k2} & \ldots & f_{kk} \end{vmatrix}, \ \ldots \tag{13}$$

macierzy $F = (f_{ij})$. W szczególności $\Delta_n = \det F$. Dodatkowo dla wygody przyjmujemy $\Delta_0 = 1$. Rolę tych minorów ilustruje pewien specjalny sposób sprowadzania formy kwadratowej do postaci kanonicznej.

TWIERDZENIE 7 (metoda Jacobiego). *Niech q będzie formą kwadratową o macierzy F, której wszystkie minory główne* (13) *są różne od zera. Wtedy istnieje baza $(\mathbf{e}'_1, \ldots, \mathbf{e}'_n)$ przestrzeni V, w której forma $q(\mathbf{x})$ przybiera postać kanoniczną*

$$q(\mathbf{x}) = \frac{\Delta_0}{\Delta_1}(x'_1)^2 + \frac{\Delta_1}{\Delta_2}(x'_2)^2 + \ldots + \frac{\Delta_{n-1}}{\Delta_n}(x'_n)^2. \tag{14}$$

Dowód. Niech $(\mathbf{e}_1, \ldots, \mathbf{e}_n)$ będzie bazą, w której forma q ma macierz F. Rozważmy $(n-1)$-wymiarową podprzestrzeń $L = \langle \mathbf{e}_1, \ldots, \mathbf{e}_{n-1} \rangle$ oraz obcięcie $\overline{q} = q|_L$ formy q do L. Macierz $\overline{F}$ formy $\overline{q}$ otrzymuje się z macierzy F formy q przez skreślenie ostatniego wiersza i ostatniej kolumny, więc jej minorami głównymi są $\overline{\Delta}_1 = \Delta_1, \ldots, \overline{\Delta}_{n-1} = \Delta_{n-1}$. Na mocy założenia wszystkie te minory są różne od zera. Rozumując przez indukcję względem n, wybieramy w L bazę $(\mathbf{e}'_1, \ldots, \mathbf{e}'_{n-1})$, w której forma $\overline{q}(\overline{\mathbf{x}})$ ($\overline{\mathbf{x}} \in L$) ma postać

$$\overline{q}(\overline{\mathbf{x}}) = q(\overline{\mathbf{x}}) = \frac{\Delta_0}{\Delta_1}(x'_1)^2 + \ldots + \frac{\Delta_{n-2}}{\Delta_{n-1}}(x'_{n-1})^2.$$

Wyraźmy ten fakt w terminach formy dwuliniowej f, biegunowej dla q:

$$f(\mathbf{e}'_i, \mathbf{e}'_i) = \frac{\Delta_{i-1}}{\Delta_i}, \qquad f(\mathbf{e}'_i, \mathbf{e}'_j) = 0, \qquad 1 \leqslant i \neq j \leqslant n-1.$$

Układ

$$f(\mathbf{x}, \mathbf{e}'_1) = 0, \ \ldots, \ f(\mathbf{x}, \mathbf{e}'_{n-1}) = 0, \qquad \mathbf{x} \in V,$$

$n-1$ równań liniowych na n-wymiarowej przestrzeni V ma rozwiązanie niezerowe $\mathbf{x}$; przyjmijmy $\mathbf{e}'_n = \mathbf{x}$. Łatwo zauważyć, że $(\mathbf{e}'_1, \ldots, \mathbf{e}'_n)$ jest bazą w V. Ponieważ wektor $\mathbf{e}'_n$ jest określony z dokładnością do czynnika liczbowego, unormujemy go tak, by macierz A przejścia od bazy $(\mathbf{e}_i)$ do $(\mathbf{e}'_i)$ miała wyznacznik

$$\det A = (\Delta_n)^{-1} = (\det F)^{-1}.$$

Niech F' będzie macierzą formy f w bazie $(\mathbf{e}'_i)$. Wtedy $f(\mathbf{e}'_i, \mathbf{e}'_j) = 0$ dla $i \neq j$

oraz

$$\frac{f(\mathbf{e}'_n,\mathbf{e}'_n)}{\Delta_{n-1}} = \frac{\Delta_0}{\Delta_1}\cdot\frac{\Delta_1}{\Delta_2}\cdots\frac{\Delta_{n-2}}{\Delta_{n-1}}\cdot f(\mathbf{e}'_n,\mathbf{e}'_n) = \prod_{i=1}^{n} f(\mathbf{e}'_i,\mathbf{e}'_i)$$

$$= \det F' = \det({}^tA\cdot F\cdot A) = (\det A)^2\det F = \frac{1}{\Delta_n},$$

skąd

$$f(\mathbf{e}'_n,\mathbf{e}'_n) = \frac{\Delta_{n-1}}{\Delta_n}.$$

Forma q ma więc w bazie $(\mathbf{e}'_i)$ żądaną postać (14). ∎

Łatwo się przekonać, że macierz A w powyższym rozumowaniu jest trójkątna:

$$\begin{aligned}
\mathbf{e}'_1 &= a_{11}\mathbf{e}_1,\\
\mathbf{e}'_2 &= a_{12}\mathbf{e}_1 + a_{22}\mathbf{e}_2,\\
&\cdots\cdots\cdots\cdots\cdots\cdots\\
\mathbf{e}'_n &= a_{n1}\mathbf{e}_1 + a_{n2}\mathbf{e}_2 + \ldots + a_{nn}\mathbf{e}_n,
\end{aligned}$$

choć fakt ten nie będzie przez nas wykorzystywany.

WNIOSEK. *Jeśli rzeczywista forma kwadratowa q ma macierz F, której wszystkie minory główne Δ_i dla $i = 1,\ldots,n$ są różne od zera, to ujemny indeks bezwładności formy q jest równy liczbie zmian znaku w ciągu*

$$1 = \Delta_0, \Delta_1, \ldots, \Delta_n.$$

Jeśli w szczególności

$$\Delta_1 > 0,\ \ldots,\ \Delta_n > 0,$$

to forma kwadratowa q jest dodatnio określona. ∎

Jak się okazuje, stwierdzenie zawarte w tym wniosku można odwrócić:

TWIERDZENIE 8 (kryterium Sylvestera). *Forma kwadratowa q na n-wymiarowej rzeczywistej przestrzeni liniowej V jest dodatnio określona wtedy i tylko wtedy, gdy wszystkie minory główne $\Delta_1,\ldots,\Delta_n$ jej macierzy F są dodatnie.*

Dowód. Zgodnie z powyższym wnioskiem nierówności $\Delta_i > 0$ dla $i = 1,\ldots,n$ implikują dodatnią określoność formy q. Aby udowodnić implikację odwrotną, podobnie jak w dowodzie twierdzenia 7 zastosujemy indukcję względem n, rozpatrując obcięcie $\overline{q} = q|_U$ formy q do $(n-1)$-wymiarowej podprzestrzeni $U = \langle\mathbf{e}_1,\ldots,\mathbf{e}_{n-1}\rangle \subset V = \langle\mathbf{e}_1,\ldots,\mathbf{e}_n\rangle$ (tutaj $(\mathbf{e}_i)$ jest bazą, w której forma q ma macierz F).

Minorami głównymi macierzy $\overline{F}$ formy $\overline{q}$ są oczywiście $\overline{\Delta_1} = \Delta_1, \ldots, \overline{\Delta_{n-1}} = \Delta_{n-1}$. Ponieważ zakładamy teraz, że forma q jest dodatnio określona, więc i $\overline{q}$

ma tę własność, a stąd na mocy założenia indukcyjnego $\Delta_1 > 0, \ldots, \Delta_{n-1} > 0$. Musimy jeszcze wykazać, że $\Delta_n > 0$. Z twierdzenia 6 wiemy, że $F = {}^tAA$ dla pewnej macierzy nieosobliwej A. Wobec tego

$$\Delta_n = \det F = \det {}^tA \cdot \det A = (\det A)^2 > 0. \blacksquare$$

9. Postać kanoniczna formy antysymetrycznej. Poświęciwszy główną uwagę formom kwadratowym i odpowiadającym im formom dwuliniowym symetrycznym, zajmiemy się teraz, kierując się twierdzeniem 2, przestrzenią $\mathcal{L}_2^-(V, \mathfrak{K})$ form dwuliniowych antysymetrycznych. Niech więc $f : V \times V \to \mathfrak{K}$ spełnia warunek

$$f(\mathbf{x}, \mathbf{y}) = -f(\mathbf{y}, \mathbf{x}) \quad \forall \mathbf{x}, \mathbf{y} \in V.$$

Podobnie jak dla form symetrycznych, *jądrem* formy f nazwiemy podprzestrzeń liniową

$$V_0 = \operatorname{Ker} f = \{\mathbf{v} \in V \mid f(\mathbf{v}, \mathbf{x}) = 0 \ \forall \mathbf{x} \in V\}.$$

Jeśli V_1 jest dowolną podprzestrzenią dopełniającą dla V_0 w V, czyli $V = V_0 \oplus V_1$, to obcięcie $f|_{V_1}$ jest formą antysymetryczną niezdegenerowaną, tj. $\operatorname{Ker} f|_{V_1} = \{0\}$. Istotnie, jeśli $\mathbf{a} \in V_1 \setminus \{\mathbf{0}\}$ i $f(\mathbf{a}, \mathbf{x}_1) = 0$ dla każdego $\mathbf{x}_1 \in V_1$, to dla każdego wektora $\mathbf{x} = \mathbf{x}_0 + \mathbf{x}_1 \in V$ (gdzie $\mathbf{x}_0 \in V_0$) mamy

$$f(\mathbf{a}, \mathbf{x}) = f(\mathbf{a}, \mathbf{x}_0 + \mathbf{x}_1) = f(\mathbf{a}, \mathbf{x}_0) + f(\mathbf{a}, \mathbf{x}_1) = -f(\mathbf{x}_0, \mathbf{a}) = 0$$

(skorzystaliśmy z antysymetryczności formy f), czyli $\mathbf{a} \in V_0$, wbrew założeniu.

W ten sposób sprowadziliśmy badanie formy f do przypadku formy niezdegenerowanej. Załóżmy więc od początku, że $f : V \times V \to \mathfrak{K}$ jest niezdegenerowaną formą antysymetryczną. Jeśli

$$V = \langle \mathbf{e}_1, \ldots, \mathbf{e}_n \rangle, \qquad \mathbf{x} = \sum_i x_i \mathbf{e}_i, \qquad \mathbf{y} = \sum_j y_j \mathbf{e}_j,$$

to

$$f(\mathbf{x}, \mathbf{y}) = \sum_{i,j=1}^{n} f_{ij} x_i y_j = {}^tXFY, \qquad f_{ij} = f(\mathbf{e}_i, \mathbf{e}_j),$$

gdzie F jest macierzą antysymetryczną: ${}^tF + F = 0$. Tak więc

$$f(\mathbf{x}, \mathbf{y}) = \sum_{1 \leqslant i < j \leqslant n} f_{ij}(x_i y_j - x_j y_i). \tag{15}$$

Z części I (rozdz. 3, § 2) wiemy, że wyznacznik macierzy antysymetrycznej F stopnia n spełnia warunek $[1 + (-1)^{n-1}] \det F = 0$, czyli nierówność $\det F \neq 0$ (warunek niezdegenerowania formy f) może być spełniona tylko dla parzystych n. Otrzymamy teraz ten sam rezultat na innej drodze, sprowadzając jednocześnie formę f do postaci kanonicznej.

W tym celu wprowadzimy pojęcie *płaszczyzny hiperbolicznej* (lub *symplektycznej*) w V, zdefiniowanej jako dowolna podprzestrzeń dwuwymiarowa $W \subset V$ spełniająca warunek $f|_W \neq 0$. Podprzestrzeń taka zawsze istnieje, ponieważ z tego, że forma f jest niezdegenerowana, wynika, że dla każdego wektora $\mathbf{e}'_1 \neq \mathbf{0}$ istnieje wektor $\mathbf{e}'_2$ taki, że $f(\mathbf{e}'_1, \mathbf{e}'_2) \neq 0$. Mnożąc $\mathbf{e}'_2$ przez odpowiedni skalar, możemy przyjąć, że $f(\mathbf{e}'_1, \mathbf{e}'_2) = 1$; poza tym oczywiście $f(\mathbf{e}'_1, \mathbf{e}'_1) = 0 = f(\mathbf{e}'_2, \mathbf{e}'_2)$.

TWIERDZENIE 9. *Niech f będzie niezdegenerowaną formą antysymetryczną na przestrzeni skończenie wymiarowej V. Wtedy* $\dim V$ *jest liczbą parzystą i jeśli* $\dim V = 2m$, *to V jest sumą prostą m płaszczyzn hiperbolicznych, parami skośnie ortogonalnych względem f* ([1]).

Dowód. Stosujemy indukcję względem $n = \dim V$. Na mocy powyższej uwagi istnieje płaszczyzna hiperboliczna $W = \langle \mathbf{e}'_1, \mathbf{e}'_2 \rangle \subset V$. Jeśli $n > 2$, to rozpatrujemy *dopełnienie skośnie ortogonalne*

$$W^{\perp} = \{\mathbf{x} \in V \mid f(\mathbf{e}'_i, \mathbf{x}) = 0 \text{ dla } i = 1, 2\}.$$

Uzupełnijmy $\mathbf{e}'_1, \mathbf{e}'_2$ do bazy $(\mathbf{e}'_1, \ldots, \mathbf{e}'_n)$ przestrzeni V. Jeśli $\mathbf{x} = x'_1\mathbf{e}'_1 + \ldots + x'_n\mathbf{e}'_n$, to układ liniowy jednorodny

$$\begin{aligned} f(\mathbf{e}'_1, \mathbf{x}) &= f'_{12}x'_2 + f'_{13}x'_3 + \ldots + f'_{1n}x'_n = 0, \\ f(\mathbf{e}'_2, \mathbf{x}) &= f'_{21}x'_1 + f'_{23}x'_3 + \ldots + f'_{2n}x'_n = 0 \end{aligned}$$

ma rząd 2, gdyż wiersze macierzy F są liniowo niezależne. Oznacza to, że przestrzeń rozwiązań $\langle \mathbf{e}'_1, \mathbf{e}'_2 \rangle^{\perp}$ tego układu ma wymiar $n - 2$. Ponieważ

$$\langle \mathbf{e}'_1, \mathbf{e}'_2 \rangle \cap \langle \mathbf{e}'_1, \mathbf{e}'_2 \rangle^{\perp} \subseteq \operatorname{Ker} f = \{\mathbf{0}\},$$

otrzymujemy rozkład

$$V = \langle \mathbf{e}'_1, \mathbf{e}'_2 \rangle \oplus \langle \mathbf{e}'_1, \mathbf{e}'_2 \rangle^{\perp},$$

przy czym obcięcie f do $\langle \mathbf{e}'_1, \mathbf{e}'_2 \rangle^{\perp}$ jest niezdegenerowaną formą antysymetryczną. Na mocy założenia indukcyjnego przestrzeń $\langle \mathbf{e}'_1, \mathbf{e}'_2 \rangle^{\perp}$ jest parzystowymiarowa i rozkłada się na sumę prostą płaszczyzn hiperbolicznych, parami skośnie ortogonalnych. Wynika stąd, że $n = \dim V = 2m$ dla pewnego m, przy czym V ma bazę $(\mathbf{e}''_i)$ taką, że $\mathbf{e}''_1 = \mathbf{e}'_1$, $\mathbf{e}''_2 = \mathbf{e}'_2$ oraz

$$V = \langle \mathbf{e}''_1, \mathbf{e}''_2 \rangle \oplus \langle \mathbf{e}''_3, \mathbf{e}''_4 \rangle \oplus \ldots \oplus \langle \mathbf{e}''_{2m-1}, \mathbf{e}''_{2m} \rangle,$$

gdzie

$$f(\alpha\mathbf{e}''_{2i-1} + \beta\mathbf{e}''_{2i}, \gamma\mathbf{e}''_{2j-1} + \delta\mathbf{e}''_{2j}) = 0, \qquad i \neq j,$$

$$f(\mathbf{e}''_{2i-1}, \mathbf{e}''_{2i}) = 1. \ \blacksquare$$

([1]) Jeśli $f : V \times V \to \mathfrak{K}$ jest formą antysymetryczną i $f(\mathbf{x}, \mathbf{y}) = 0$, to wektory $\mathbf{x}, \mathbf{y} \in V$ nazywamy *skośnie ortogonalnymi* względem f (*przyp. tłum.*).

Wniosek. *Dla każdej nieosobliwej macierzy antysymetrycznej F stopnia $2m$ istnieje macierz nieosobliwa A taka, że ${}^tAFA = J$, gdzie*

$$J = \begin{Vmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ \multicolumn{4}{c}{\dotfill} \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{Vmatrix}.$$

Dowód. Teza wynika natychmiast z twierdzenia 9 — wystarczy przypomnieć, jak zmienia się macierz formy dwuliniowej przy przejściu do nowej bazy. ■

Uwaga. Twierdzenie 9 i wniosek z niego zachodzą dla dowolnego ciała skalarów o charakterystyce różnej od 2. Niekiedy za standardową macierz antysymetryczną przyjmuje się

$$J_0 = \begin{Vmatrix} 0 & -E_m \\ E_m & 0 \end{Vmatrix}.$$

Przejście od J do J_0 polega na przenumerowaniu wektorów bazowych.

Jeśli w praktyce chcemy sprowadzić daną formę antysymetryczną (15) do postaci kanonicznej, możemy znów zastosować metodę Lagrange'a. Zakładając, że $f \neq 0$ (w przeciwnym razie nie ma nic do roboty) i przenumerowując ewentualnie wektory bazowe, możemy przyjąć, że $f_{12} \neq 0$. Wydzielmy w (15) wszystkie człony zawierające x_1 lub y_1:

$$x_1(f_{12}y_2 + \ldots + f_{1n}y_n) - (f_{12}x_2 + \ldots + f_{1n}x_n)y_1.$$

Wprowadzając nowe zmienne

$$x_2' = f_{12}x_2 + \ldots + f_{1n}x_n, \qquad y_2' = f_{12}y_2 + \ldots + f_{1n}y_n$$

$(x_1, y_1, x_3, y_3, \ldots, x_n, y_n$ pozostają bez zmian), wydzielamy z f człony zawierające x_2' i y_2':

$$\begin{aligned} f(\mathbf{x}, \mathbf{y}) &= (x_1 + f_{32}'x_3 + \ldots + f_{n2}'x_n)y_2' - x_2'(y_1 + f_{32}'y_3 + \ldots + f_{n2}'y_n) + \ldots \\ &= x_1'y_2' - x_2'y_1' + \ldots, \end{aligned}$$

gdzie

$$x_1' = x_1 + f_{32}'x_3 + \ldots + f_{n2}'x_n, \qquad y_1' = y_1 + f_{32}'y_3 + \ldots + f_{n2}'y_n,$$

a wielokropek oznacza człony zawierające tylko $x_3, y_3, \ldots, x_n, y_n$. Z tymi członami postępujemy analogicznie, otrzymując ostatecznie postać kanoniczną

$$f(\mathbf{x}, \mathbf{y}) = (x_1'y_2' - x_2'y_1') + \ldots + (x_{2m-1}'y_{2m}' - x_{2m}'y_{2m-1}'). \tag{16}$$

10. Pfaffian. Zgodnie z wnioskiem z twierdzenia 9, dla każdej nieosobliwej macierzy antysymetrycznej F stopnia $2m$ można znaleźć macierz nieosobliwą A

taką, że

$$
{}^tAFA = \begin{Vmatrix} 0 & -E_m \\ E_m & 0 \end{Vmatrix} = J_0.
$$

Stąd $(\det A)^2 \cdot \det F = 1$, czyli $\det F$ jest kwadratem w ciele skalarów $\mathfrak{K}$. Sugeruje to, by rozpatrzyć ciało ułamków

$$
\mathfrak{K} = \mathbb{Q}(t) = \mathbb{Q}(t_{12}, t_{13}, \ldots, t_{n-1,n})
$$

pierścienia wielomianów

$$
\mathbb{Z}[t] = \mathbb{Z}[t_{12}, t_{13}, \ldots, t_{n-1,n}]
$$

$n(n-1)/2$ zmiennych niezależnych oraz macierz antysymetryczną $T = (t_{ij})$, czyli $t_{ji} = -t_{ij}$ dla $i \leqslant j$. Wiemy, że $\det T$ jest kwadratem w ciele $\mathbb{Q}(t)$. Z drugiej strony, $\det T$ to wielomian z pierścienia $\mathbb{Z}[t]$. Oznacza to (korzystamy tu z jednoznaczności rozkładu w $\mathbb{Z}[t]$, zob. części I i III), że $\det T$ jest kwadratem pewnego wielomianu z $\mathbb{Z}[t]$:

$$
\det T = P_n(t)^2.
$$

Unormujmy $P_n(t) = P_n(t_{12}, t_{13}, \ldots)$ w taki sposób, że $P_n(t_{12}^0, t_{13}^0, \ldots) = 1$ dla tych wartości $t_{ij}^0 = 0, \pm 1$ zmiennych t_{ij}, dla których $T_0 = (t_{ij}^0) = J_0$. Otrzymujemy w ten sposób jednoznacznie określony wielomian $\mathrm{Pf}_n(t)$ zwany *pfaffianem* stopnia n. Na przykład

$$
\mathrm{Pf}_2(t) = t, \qquad \mathrm{Pf}_4(t) = t_{12}t_{34} - t_{13}t_{24} + t_{14}t_{23},
$$

co łatwo uzasadnić, obliczając wyznaczniki macierzy

$$
\begin{Vmatrix} 0 & -t \\ t & 0 \end{Vmatrix}, \qquad \begin{Vmatrix} 0 & -t_{12} & -t_{13} & -t_{14} \\ t_{12} & 0 & -t_{23} & -t_{24} \\ t_{13} & t_{23} & 0 & -t_{34} \\ t_{14} & t_{24} & t_{34} & 0 \end{Vmatrix}.
$$

Symbolem $\mathrm{Pf}(F)$ będziemy oznaczać rezultat podstawienia w $\mathrm{Pf}_n(t_{ij})$ wyrazów f_{ij} macierzy antysymetrycznej F w miejsce t_{ij} (zamieniając wszędzie $\mathbb{Q}$ na ciało proste Z_p, rozszerzamy nasze rozważania na przypadek ciał dowolnej charakterystyki $\neq 2$).

TWIERDZENIE 10. *Jeśli F jest macierzą antysymetryczną stopnia n, to*

$$
\det F = \mathrm{Pf}(F)^2.
$$

Ponadto

$$
\mathrm{Pf}({}^tAFA) = \det A \cdot \mathrm{Pf}(F)
$$

dla dowolnej macierzy kwadratowej A stopnia n.

Dowód. Związek $\det F = \mathrm{Pf}(F)^2$ wyraża znane nam już własności macier antysymetrycznych i pfaffianu. Niech teraz $U = (u_{ij})$ będzie dowolną macie

$\times n$ o algebraicznie niezależnych elementach u_{ij}, a T — macierzą antysyme-yczną rozpatrywaną powyżej. Wówczas

$$\mathrm{Pf}({}^tUTU)^2 = \det({}^tUTU) = (\det U)^2 \det T = (\det U)^2\,\mathrm{Pf}(T)^2,$$

stąd

$$\mathrm{Pf}({}^tUTU) = \pm(\det U)\,\mathrm{Pf}(T).$$

›śli teraz nadamy u_{ij} i t_{ij} takie wartości, że U stanie się macierzą jednostkową, T — standardową macierzą antysymetryczną J_0, to po lewej stronie otrzymamy f$(J_0) = 1$, a po prawej $\pm\,\mathrm{Pf}(J_0)$, tj. należy wziąć znak plus. Oznacza to, że ,dana równość zachodzi także dla konkretnych macierzy $U = A$, $T = F$. ∎

Zauważmy na koniec, że dla $T = (t_{ij})_{i,j=1}^{2m}$ (${}^tT = -T$) pfaffian $\mathrm{Pf}(T)$ jest ielomianem uniwersalnym, będącym formą jednorodną stopnia n o współczynkach całkowitych lub należących do ciała prostego.

WICZENIA

Niech $\Delta_1, \ldots, \Delta_n = \det F$ będą minorami głównymi rzeczywistej formy kwadratowej q o macierzy F. Udowodnić, że q i F są ujemnie określone dokładnie wtedy, gdy $(-1)^k\Delta_k > 0$ dla $k = 1, \ldots, n$.

Podać przykład:

(a) macierzy dodatnio określonej $A = (a_{ij})$, w której $a_{ij} < 0$ dla pewnych i, j;
(b) macierzy $A = (a_{ij})$, która nie jest dodatnio określona, ale w której $a_{ij} > 0$ dla wszystkich i, j.

Dla jakich $\lambda, \mu \in \mathbb{R}$ macierze

$$\begin{Vmatrix} 1 & \lambda & \lambda \\ \lambda & 1 & \lambda \\ \lambda & \lambda & 1 \end{Vmatrix}, \qquad \begin{Vmatrix} 1 & 1 & \mu \\ 1 & \mu & 1 \\ \mu & 1 & 1 \end{Vmatrix}$$

są dodatnio określone?

Niech $\mathbf{x} = [x_1, x_2, x_3] \in \mathbb{C}^3$ i $Q(\mathbf{x}) = x_1^3 + x_2^3 + x_3^3 - 3x_1x_2x_3$. Wykorzystując rozkład

$$Q(\mathbf{x}) = (x_1 + x_2 + x_3)(x_1 + \varepsilon x_2 + \varepsilon^2 x_3)(x_1 + \varepsilon^2 x_2 + \varepsilon x_3),$$

gdzie ε jest pierwiastkiem pierwotnym z jedynki stopnia 3, wykazać, że $Q(\mathbf{x})Q(\mathbf{y}) = Q(\mathbf{z})$, gdzie $\mathbf{z} = [z_1, z_2, z_3]$, przy czym $z_i(\mathbf{x}, \mathbf{y}) = \sum_{j,k} a_{jk}^{(i)} x_j y_k$ są symetrycznymi formami dwuliniowymi. Znaleźć ich jawną postać.

Niech A będzie dowolną rzeczywistą macierzą symetryczną, a $\varepsilon = \varepsilon(A)$ — dostatecznie małą liczbą dodatnią. Udowodnić, że macierz $B = E + \varepsilon A$ jest dodatnio określona.