

PRZYPIS I

WYZNACZNIKI, RÓWNANIA LINIOWE I FORMY KWADRATOWE

§ 1. Wyznaczniki

1. Permutacje. Dwa różne przedmioty a i b możemy ustawić na dwóch miejscach dwoma różnymi sposobami:

$$ab, \quad ba.$$

Dla trzech różnych przedmiotów a, b i c jest sześć różnych ustawień na trzech miejscach:

$$\begin{array}{ccc} abc, & bac, & cab, \\ acb, & bca, & cba. \end{array}$$

Łatwo sprawdzić, że na czterech miejscach można cztery przedmioty ustawić 24 sposobami.

Każde ustawienie n różnych przedmiotów na n miejscach nazywamy *permutacją* tych n przedmiotów.

Liczbę różnych permutacji n przedmiotów oznaczamy przez P_n .

Udowodnimy, że

$$(1) \quad P_n = 1 \cdot 2 \cdot 3 \cdot \dots \cdot n.$$

Istotnie, dla $n=1$ jest

$$(2) \quad P_1 = 1,$$

a dla $n>1$ zachodzi wzór

$$(3) \quad P_n = n P_{n-1}.$$

Gdy bowiem na pierwszym miejscu ustawimy jakiś przedmiot, to pozostaje $n-1$ miejsc, na których można pozostałych $n-1$ przedmiotów ustawić P_{n-1} sposobami. Ponieważ na pierwszym miejscu

możemy ustawić po kolei każdy z n przedmiotów, więc otrzymujemy łącznie $n P_{n-1}$ różnych permutacji. Wobec (2) otrzymujemy z (3)

$$P_2 = 2 \cdot 1, \quad P_3 = 3 P_2 = 3 \cdot 2 \cdot 1;$$

ogólnie przez indukcję dostajemy (1).

Iloczyn $1 \cdot 2 \cdot 3 \cdot \dots \cdot n$ oznacza się krótko przez $n!$, co czyta się *n silnia*.

Tak więc wzór (1) możemy napisać w postaci

$$(4) \quad P_n = n!$$

W dalszym ciągu będziemy rozpatrywali tylko permutacje liczb $1, 2, \dots, n$. Niech

$$(5) \quad a_1, a_2, \dots, a_n$$

oznacza dowolną permutację tych liczb. Jeśli

$$a_i > a_k \quad \text{dla} \quad i < k,$$

to mówimy, że para a_i, a_k tworzy w permutacji (5) *inwersję*.

Np. w permutacji

$$(6) \quad \underline{4}, \underline{1}, \underline{6}, \underline{2}, \underline{5}, \underline{3}$$

pary 4,1 i 6,5 tworzą inwersję, para zaś 1,3 nie tworzy inwersji. Wszystkich inwersji jest tu 7.

2. Określenie macierzy i wyznacznika. Niech każdej parze uporządkowanej liczb naturalnych i, k , gdzie $1 \leq i \leq m$ i $1 \leq k \leq n$, będzie przyporządkowana liczba a_{ik} . Przyporządkowanie to jest więc funkcją dwóch zmiennych i, k , której wartościami są liczby a_{ik} .

Funkcję taką nazywa się *macierzą* i oznacza przez

$$(a_{ik}),$$

dopisując w razie potrzeby zakresy zmienności:

$$i=1, 2, \dots, m, \quad k=1, 2, \dots, n.$$

Macierze oznaczamy również literami niemieckimi $\mathfrak{A}, \mathfrak{B}, \mathfrak{C}, \dots$. Wygodnie też bywa zapisywać macierze w postaci

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix},$$

gdzie w i -tym wierszu stoją wyrazy o tym samym pierwszym wskaźniku i , a w k -tej kolumnie — wyrazy o tym samym drugim wskaźniku k . Oto przykłady macierzy 1-, 2- i wielowyrzowych:

$$\begin{pmatrix} 1/2 \end{pmatrix}, \quad \begin{pmatrix} 2/3 & -2 \end{pmatrix}, \quad \begin{pmatrix} 3 & 1 \\ 2 & 4 \end{pmatrix}, \quad \begin{pmatrix} 3 & 2 \\ -1 & 5 \\ 2 & 0 \end{pmatrix}.$$

Gdy $m=n$, macierz nazywa się *kwadratową*, a n jej *stopniem*. Weźmy pod uwagę macierz kwadratową

$$(7) \quad \mathfrak{A} = (a_{ik}) = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}$$

Niech $a_1, a_2, \dots, a_n$ będzie dowolną permutacją liczb $1, 2, \dots, n$, a I ilość inwersji w tej permutacji. Utwórzmy iloczyn

$$(8) \quad (-1)^I a_{1a_1} a_{2a_2} \dots a_{na_n}.$$

Iloczyn (8) zawiera wśród swych czynników dokładnie 1 wyraz z każdego wiersza i dokładnie 1 wyraz z każdej kolumny macierzy (7). Zesumujmy iloczyny (8), utworzone dla wszystkich $n!$ permutacji liczb $1, 2, \dots, n$. Otrzymaną sumę oznacza się znakiem

$$(9) \quad \sum (-1)^I a_{1a_1} a_{2a_2} \dots a_{na_n}$$

lub znakami

$$(10) \quad |\mathfrak{A}|, \quad |a_{ik}|, \quad \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}$$

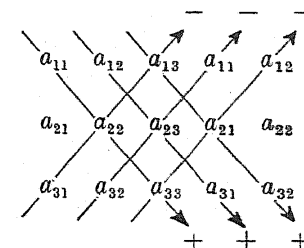
i nazywa się *wyznacznikiem* macierzy (7).

Z tej definicji wyznacznika wynika w szczególności, że

$$(11) \quad \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11} a_{22} - a_{12} a_{21},$$

$$(12) \quad \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11} a_{22} a_{33} - a_{13} a_{22} a_{31} + a_{12} a_{23} a_{31} - \\ - a_{11} a_{23} a_{32} + a_{13} a_{21} a_{32} - a_{12} a_{21} a_{33}.$$

Znany jest prosty sposób mnemotechniczny obliczania wyznacznika 3-go stopnia, mianowicie tzw. *schemat Sarrusa*: dopisuje się z prawej strony do trzech kolumn jeszcze raz pierwszą i drugą kolumnę macierzy. Powstaje w ten sposób schemat



według którego się oblicza.

3. Własności wyznaczników. Jeżeli jakiś wiersz lub kolumna wyznacznika składa się z samych zer, to wyznacznik jest równy 0.

Każdy bowiem składnik sumy (9) zawiera czynnik z każdego wiersza i każdej kolumny.

Utwórzmy z macierzy kwadratowej $\mathfrak{A}$ nową macierz $\mathfrak{B}$, zamieniając w (7) wiersze na kolumny. Otrzymujemy macierz

$$\mathfrak{B} = (b_{ik})$$

gdzie $b_{ik} = a_{ki}$.

Taką macierz $\mathfrak{B}$ nazywa się *macierzą przestawioną* macierzy $\mathfrak{A}$ i oznacza się przez $\mathfrak{A}^+$.

Wyznaczniki macierzy i macierzy przestawionej są równe:

$$(13) \quad |\mathfrak{A}| = |\mathfrak{A}^+|.$$

Np. dla wyznaczników 2-go stopnia jest na mocy (11)

$$|\mathfrak{A}^+| = \begin{vmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{vmatrix} = a_{11} a_{22} - a_{12} a_{21} = |\mathfrak{A}|;$$

dla wyznaczników 3-go stopnia jest na mocy (12)

$$|\mathfrak{A}^+| = \begin{vmatrix} a_{11} & a_{21} & a_{31} \\ a_{12} & a_{22} & a_{32} \\ a_{13} & a_{23} & a_{33} \end{vmatrix} = a_{11} a_{22} a_{33} - a_{31} a_{22} a_{13} + a_{21} a_{32} a_{13} - \\ - a_{11} a_{32} a_{23} + a_{31} a_{12} a_{23} - a_{21} a_{12} a_{33} = |\mathfrak{A}|.$$

Gdy macierz $\mathfrak{B}$ powstaje z macierzy $\mathfrak{A}$ przez przestawienie dwóch kolumn lub dwóch wierszy, to

$$(14) \quad |\mathfrak{B}| = -|\mathfrak{A}|.$$

Np. dla wyznaczników 2-go stopnia przez przestawienie kolumn otrzymujemy

$$\begin{vmatrix} a_{12} & a_{11} \\ a_{22} & a_{21} \end{vmatrix} = a_{12}a_{21} - a_{11}a_{22} = -|\mathfrak{A}|,$$

a gdy w wyznaczniku 3-go stopnia przestawimy pierwszy i trzeci wiersz, dostaniemy wyznacznik

$$\begin{vmatrix} a_{31} & a_{32} & a_{33} \\ a_{21} & a_{22} & a_{23} \\ a_{11} & a_{12} & a_{13} \end{vmatrix} = a_{31}a_{22}a_{13} - a_{11}a_{22}a_{33} + a_{32}a_{23}a_{11} - \\ - a_{12}a_{23}a_{31} + a_{33}a_{21}a_{13} - a_{13}a_{21}a_{32} = -|\mathfrak{A}|.$$

Wyznacznik macierzy, w której dwa wiersze lub dwie kolumny są jednakowe, jest równy 0.

Przestawiając bowiem te dwa wiersze (lub kolumny), dostajemy ten sam wyznacznik $|\mathfrak{A}|$, a stąd wynika w myśl poprzedniego twierdzenia, że

$$|\mathfrak{A}| = -|\mathfrak{A}|$$

czyli

$$\mathfrak{A} = 0.$$

Utwórzmy macierz $(n-1)$ -go stopnia $\mathfrak{B}$ przez wykreślenie i -tego wiersza i k -tej kolumny z macierzy n -tego stopnia $\mathfrak{A}$.

Liczbę

$$A_{ik} = (-1)^{i+k} |\mathfrak{B}|$$

nazywamy *dopełnieniem algebraicznym wyrazu a_{ik}* .

$$\text{Np. gdy } \mathfrak{A} = \begin{pmatrix} 5 & 0 & -1 \\ 2 & 0 & 3 \\ 4 & 2 & 0 \end{pmatrix}, \text{ to } A_{23} = (-1)^{2+3} \begin{vmatrix} 5 & 0 \\ 4 & 2 \end{vmatrix} = -10.$$

Zachodzą następujące równości:

$$(15) \quad |\mathfrak{A}| = a_{i1}A_{i1} + a_{i2}A_{i2} + \dots + a_{in}A_{in} \quad \text{dla } i=1,2,\dots,n$$

$$(16) \quad |\mathfrak{A}| = a_{1k}A_{1k} + a_{2k}A_{2k} + \dots + a_{nk}A_{nk} \quad \text{dla } k=1,2,\dots,n.$$

Są to, tzw. *rozwinęcia wyznacznika według i -tego wiersza oraz według k -tej kolumny*; zawdzięczamy je Laplace'owi.

Dla wyznaczników 2-go stopnia wzory (15) i (16) są oczywiste.

Dla wyznaczników 3-go stopnia, porządkując (12) według wyrazów drugiego wiersza, otrzymujemy

$$|\mathfrak{A}| = a_{21}(a_{13}a_{32} - a_{12}a_{33}) + a_{22}(a_{11}a_{33} - a_{13}a_{31}) + a_{23}(a_{12}a_{31} - a_{11}a_{32}) = \\ = -a_{21} \begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} + a_{22} \begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix} - a_{23} \begin{vmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{vmatrix}$$

czyli rozwiniecie (15) dla $i=2$. Podobnie, porządkując (12) według wyrazów innego wiersza lub kolumny, otrzymuje się rozwiniecie Laplace'a dla pozostałych wartości i i k . Ogólnie, można dowieść twierdzenia Laplace'a dla wyznaczników stopni $n > 3$, np. przez indukcję. Wzory (15) i (16) dają praktyczny sposób ich obliczania, które według samej tylko definicji wyznacznika byłoby żmudne.

Dla przykładu obliczmy wyznacznik 4-go stopnia

$$A = \begin{vmatrix} 1 & 0 & 3 & -1 \\ 2 & 3 & 0 & 1 \\ 0 & 1 & 0 & 2 \\ 2 & 0 & 4 & 0 \end{vmatrix}.$$

Przez rozwiniecie według trzeciej kolumny otrzymujemy

$$A = 3 \begin{vmatrix} 2 & 3 & 1 \\ 0 & 1 & 2 \\ 2 & 0 & 0 \end{vmatrix} - 4 \begin{vmatrix} 1 & 0 & -1 \\ 2 & 3 & 1 \\ 0 & 1 & 2 \end{vmatrix}.$$

Obliczając obydwie wyznaczniki 3-go stopnia np. za pomocą schematu Sarrusa, otrzymujemy łatwo

$$A = 3 \cdot 10 - 4 \cdot 3 = 18.$$

Jeżeli macierz $\mathfrak{B}$ powstaje z macierzy $\mathfrak{A}$ przez pomnożenie wyrazów jakiegoś wiersza lub kolumny przez stałą s , to

$$(17) \quad |\mathfrak{B}| = s|\mathfrak{A}|.$$

Aby tego dowieść, zauważmy, że gdy w wyznaczniku zmienić w dowolny sposób wyrazy jakiegokolwiek wiersza lub kolumny, to nie ulegną zmianie dopełnienia algebraiczne tych wyrazów. Jeśli więc w szczególności pomnożymy elementy i -tego wiersza przez s , to stosując do otrzymanej macierzy $\mathfrak{B}$ rozwiniecie Laplace'a według tego wiersza, otrzymamy

$$|\mathfrak{B}| = (sa_{i1})A_{i1} + (sa_{i2})A_{i2} + \dots + (sa_{in})A_{in} = \\ = s(a_{i1}A_{i1} + a_{i2}A_{i2} + \dots + a_{in}A_{in}) = s|\mathfrak{A}|.$$

a więc

$$x = \frac{1}{7} \begin{vmatrix} 5 & -1 \\ 8 & 2 \end{vmatrix} = \frac{18}{7}, \quad y = \frac{1}{7} \begin{vmatrix} 2 & 5 \\ 3 & 8 \end{vmatrix} = \frac{1}{7}.$$

Podobnie, dla układu trzech równań

$$2x - y + z = 3, \quad 3x + 2y + 5z = 1, \quad x + 2y - z = -1$$

jest

$$|\mathcal{M}| = \begin{vmatrix} 2 & -1 & 1 \\ 3 & 2 & 5 \\ 1 & 2 & -1 \end{vmatrix} = -28 \neq 0,$$

a więc

$$x = -\frac{1}{28} \begin{vmatrix} 3 & -1 & 1 \\ 1 & 2 & 5 \\ -1 & 2 & -1 \end{vmatrix} = 1, \quad y = -\frac{1}{28} \begin{vmatrix} 2 & 3 & 1 \\ 3 & 1 & 5 \\ 1 & -1 & -1 \end{vmatrix} = -1,$$

$$z = -\frac{1}{28} \begin{vmatrix} 2 & -1 & 3 \\ 3 & 2 & 1 \\ 1 & 2 & -1 \end{vmatrix} = 0.$$

W przypadku, gdy

$$b_1 = b_2 = \dots = b_n = 0,$$

równania (1) oraz ich układ nazywają się *jednorodnymi*.

Wówczas $B_k = 0$, a więc wobec (3) jest $x_k = 0$.

Zatem *układ jednorodny o wyznaczniku różnym od 0 ma jedyne rozwiązanie*

$$x_1 = x_2 = \dots = x_n = 0.$$

Rozwiązanie takie nazywa się *zerowe*.

W szczególności więc, *jeśli układ jednorodny kwadratowy ma rozwiązanie niezerowe, to $|\mathcal{M}| = 0$.*

2. Układ ogólniejszy. Zbadamy obecnie układ m równań o n niewiadomych

$$(4) \quad \begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{1m}x_m + a_{1m+1}x_{m+1} + \dots + a_{1n}x_n &= b_1, \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2m}x_m + a_{2m+1}x_{m+1} + \dots + a_{2n}x_n &= b_2, \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mm}x_m + a_{mm+1}x_{m+1} + \dots + a_{mn}x_n &= b_m, \end{aligned}$$

gdzie $m < n$.

Niech $\mathcal{M}$ będzie macierzą prostokątną, utworzoną ze współczynników układu (4).

Założmy, że jeden z wyznaczników m -tego stopnia, jakie można utworzyć z kolumn macierzy $\mathcal{M}$, jest różny od zera.

Przez zmianę numeracji niewiadomych można osiągnąć to, że

$$(5) \quad \Delta = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mm} \end{vmatrix} \neq 0.$$

Układ (4) jest równoważny układowi

$$(6) \quad \begin{aligned} a_{11}x_1 + \dots + a_{1m}x_m &= b_1 - a_{1m+1}x_{m+1} - \dots - a_{1n}x_n, \\ a_{21}x_1 + \dots + a_{2m}x_m &= b_2 - a_{2m+1}x_{m+1} - \dots - a_{2n}x_n, \\ &\vdots \\ a_{m1}x_1 + \dots + a_{mm}x_m &= b_m - a_{mm+1}x_{m+1} - \dots - a_{mn}x_n. \end{aligned}$$

Jeśli obecnie za $x_{m+1}, x_{m+2}, \dots, x_n$ podstawić do (6) dowolne wartości, to prawe strony tych równań staną się liczbami stałymi i (6) stanie się wobec (5) układem Cramera o niewiadomych $x_1, x_2, \dots, x_m$. Wynika stąd, że *układ (4) ma nieskończenie wiele rozwiązań*, w których za $x_{m+1}, \dots, x_n$ możemy przyjąć dowolne wartości, a następnie wyznaczyć $x_1, \dots, x_m$.

W szczególności, gdy układ (4) jest jednorodny, a więc gdy $b_1 = b_2 = \dots = b_m = 0$, ma on nieskończenie wiele rozwiązań różnych od zerowego. Np. dla układu równań

$$(7) \quad x + 2y - 5z + 8t = 0, \quad 2x + 4y + z - t = 0$$

jest

$$\Delta = \begin{vmatrix} 2 & -5 \\ 4 & 1 \end{vmatrix} = 22 \neq 0.$$

Napiszmy układ równań (7) w postaci

$$2y - 5z = -x - 8t, \quad 4y + z = -2x + t.$$

Podstawiając $x = a$ i $t = b$, gdzie a i b są dowolnymi liczbami, otrzymujemy za pomocą wzorów Cramera

$$y = \frac{1}{22} \begin{vmatrix} -a-8b & -5 \\ -2a+b & 1 \end{vmatrix} = -\frac{1}{22} (11a + 3b),$$

$$z = \frac{1}{22} \begin{vmatrix} 2 & -a-8b \\ 4 & -2a+b \end{vmatrix} = \frac{17}{11} b.$$

Każde rozwiązanie układu (7) można więc napisać w postaci

$$x=a, \quad y=-\frac{1}{22}(11a+3b), \quad z=\frac{17}{11}b, \quad t=b.$$

Często spotyka się układ równań

$$(8) \quad \begin{aligned} A_1x + B_1y + C_1z &= 0, \\ A_2x + B_2y + C_2z &= 0, \end{aligned}$$

gdzie np.

$$\begin{vmatrix} A_1 & B_1 \\ A_2 & B_2 \end{vmatrix} \neq 0.$$

Podstawiając tu $z=c$, gdzie c jest liczbą dowolną, dostajemy

$$x = -\frac{c}{\Delta} \begin{vmatrix} C_1 & B_1 \\ C_2 & B_2 \end{vmatrix}, \quad y = -\frac{c}{\Delta} \begin{vmatrix} A_1 & C_1 \\ A_2 & C_2 \end{vmatrix}, \quad z=c.$$

Przyjmując oznaczenie

$$\varrho = \frac{c}{\Delta},$$

otrzymujemy więc

$$x=\varrho \begin{vmatrix} B_1 & C_1 \\ B_2 & C_2 \end{vmatrix}, \quad y=\varrho \begin{vmatrix} C_1 & A_1 \\ C_2 & A_2 \end{vmatrix}, \quad z=\varrho \begin{vmatrix} A_1 & B_1 \\ A_2 & B_2 \end{vmatrix}.$$

Jest to najogólniejszy kształt rozwiązań układu równań (8). W szczególności, dla $\varrho \neq 0$ mamy

$$(9) \quad x:y:z = \begin{vmatrix} B_1 & C_1 \\ B_2 & C_2 \end{vmatrix} : \begin{vmatrix} C_1 & A_1 \\ C_2 & A_2 \end{vmatrix} : \begin{vmatrix} A_1 & B_1 \\ A_2 & B_2 \end{vmatrix}.$$

3. Rząd macierzy. Skreślmy w dowolnej macierzy prostokątnej lub kwadratowej $\mathfrak{A}$ pewną ilość wierszy i kolumn tak, aby pozostałe wiersze i kolumny utworzyły macierz kwadratową $\mathfrak{M}$. Wyznacznik macierzy $\mathfrak{M}$ nazywa się *minorem* macierzy $\mathfrak{A}$. Np. wyznaczniki

$$2, \quad \begin{vmatrix} 1 & 3 \\ 0 & 4 \end{vmatrix} = 4, \quad \begin{vmatrix} 1 & 3 & 4 \\ 2 & 0 & 3 \\ 0 & 4 & 1 \end{vmatrix} = 14$$

są minorami 1-go, 2-go i 3-go stopnia macierzy

$$(10) \quad \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & -1 & 0 & 3 \\ 0 & 1 & 4 & 1 \end{pmatrix};$$

nie są to oczywiście wszystkie minory tej macierzy.

Rzędem macierzy nazywa się najwyższy ze stopni tych jej minorów, które są różne od 0.

Np. macierz (10) jest rzędu 3, a macierz

$$\begin{pmatrix} 1 & 2 & -3 & 0 \\ 0 & 1 & 3 & 4 \\ 1 & 3 & 0 & 4 \end{pmatrix}$$

jest rzędu 2, ponieważ wszystkie jej minory 3-go stopnia są równe 0 i istnieje w niej minor 2-go stopnia różny od zera: mianowicie, jak łatwo obliczyć,

$$\begin{vmatrix} 1 & 2 \\ 0 & 1 \end{vmatrix} \neq 0.$$

Rząd macierzy $\mathfrak{A}$ oznaczamy przez $R(\mathfrak{A})$.

$R(\mathfrak{A})=0$ znaczy więc, że wszystkie wyrazy macierzy $\mathfrak{A}$ są równe 0.

Rząd macierzy nie ulega zmianie, gdy:

1° *przestawiamy wiersze lub kolumny,*

2° *mnożymy wiersze lub kolumny przez dowolną liczbę $s \neq 0$,*

3° *dodajemy do wiersza (kolumny) inny wiersz pomnożony (kolumnę pomnożoną) przez dowolną liczbę $s \neq 0$.*

Własność 1° wynika stąd, że przestawienie kolumn lub wierszy wpływa jedynie na znak minorów macierzy (p. § 1, str. 493), a więc każdy minor macierzy przekształconej albo jest równy odpowiedniemu minorowi macierzy pierwotnej, albo różni się od niego tylko znakiem.

Własność 2° wynika stąd, że każdy minor macierzy przekształconej jest równy albo pewnemu minorowi pierwotnej, albo jego iloczynowi przez s .

Aby dowieść własności 3^o, założmy np., że do pierwszej kolumny dodaliśmy drugą kolumnę pomnożoną przez s . Niech więc

$$\mathfrak{A} = \begin{pmatrix} a_{11} & a_{12} & \dots \\ a_{21} & a_{22} & \dots \\ \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots \end{pmatrix}, \quad \mathfrak{B} = \begin{pmatrix} a_{11} + sa_{12} & a_{12} & \dots \\ a_{21} + sa_{22} & a_{22} & \dots \\ \dots & \dots & \dots \\ a_{m1} + sa_{m2} & a_{m2} & \dots \end{pmatrix}.$$

Udowodnimy naprzód, że

$$(11) \quad R(\mathfrak{B}) \leq R(\mathfrak{A}).$$

Niech dla krótkości

$$(12) \quad r = R(\mathfrak{A})$$

i niech M będzie dowolnym minorem $(r+1)$ -go stopnia macierzy $\mathfrak{B}$.

Jeśli minor M nie zawiera pierwszej kolumny, to jest minorem $(r+1)$ -go stopnia macierzy $\mathfrak{A}$, a więc wobec (12) jest równy 0.

Jeśli zaś minor M zawiera pierwszą kolumnę, np. zawiera pierwszych $r+1$ wierszy, to

$$M = \begin{vmatrix} a_{11} + sa_{12} & a_{13} & \dots \\ a_{21} + sa_{22} & a_{23} & \dots \\ \dots & \dots & \dots \\ a_{r+11} + sa_{r+12} & a_{r+13} & \dots \end{vmatrix} = \begin{vmatrix} a_{11} & a_{13} & \dots \\ a_{21} & a_{23} & \dots \\ \dots & \dots & \dots \\ a_{r+11} & a_{r+13} & \dots \end{vmatrix} + s \begin{vmatrix} a_{12} & a_{13} & \dots \\ a_{22} & a_{23} & \dots \\ \dots & \dots & \dots \\ a_{r+12} & a_{r+13} & \dots \end{vmatrix}.$$

Pierwszy ze składników jest minorem $(r+1)$ -go stopnia macierzy $\mathfrak{A}$, a więc na mocy (12) równy jest 0. Drugi albo jest minorem $(r+1)$ -go stopnia macierzy $\mathfrak{A}$, albo ma pierwsze dwie kolumny identyczne, a więc także jest równy 0. Zatem w każdym przypadku jest $M=0$, a wobec tego wzór (11) jest dowiedziony.

Zauważmy teraz, że i na odwrót, macierz $\mathfrak{A}$ można otrzymać z macierzy $\mathfrak{B}$ przez dodanie do pierwszej kolumny macierzy $\mathfrak{B}$ drugiej kolumny tejże macierzy pomnożonej przez liczbę $-s$. Jest więc na mocy wzoru (11), zastosowanego do odwróconych ról macierzy $\mathfrak{A}$ i $\mathfrak{B}$,

$$(13) \quad R(\mathfrak{A}) \leq R(\mathfrak{B}).$$

Wzory (11) i (13) dają równość

$$R(\mathfrak{A}) = R(\mathfrak{B});$$

tym samym własność 3^o została także dowiedziona.

Wyznaczanie rzędu macierzy jest bardzo często potrzebne zarówno w algebrze jak w geometrii. Udowodnione twierdzenie je ułatwia, dając możność powiększenia w macierzy ilości wyrazów równych 0 przez wykonywanie na jej wierszach lub kolumnach trzech działań, o których mowa w tym twierdzeniu. Np. niech

$$\mathfrak{A} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & 1 & 2 & -1 & 3 \\ 1 & 3 & 5 & 3 & 8 \\ 2 & 5 & 8 & 7 & 13 \end{pmatrix};$$

odejmując od wierszy trzeciego i czwartego wiersz drugi, dostajemy

$$\mathfrak{A}_1 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & 1 & 2 & -1 & 3 \\ 1 & 2 & 3 & 4 & 5 \\ 2 & 4 & 6 & 8 & 10 \end{pmatrix}.$$

Odejmując teraz wiersz pierwszy od trzeciego i podwojony wiersz pierwszy od czwartego, otrzymujemy

$$\mathfrak{A}_2 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & 1 & 2 & -1 & 3 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Widzimy, że

$$R(\mathfrak{A}) = R(\mathfrak{A}_1) = R(\mathfrak{A}_2) = 2.$$

4. Twierdzenie Capelliego i Kroneckera. Oznaczmy dla krótkości literą x zespół n zmiennych $x_1, x_2, \dots, x_n$ i weźmy pod uwagę m form liniowych tych zmiennych:

$$(14) \quad f_k(x) \equiv a_{k1}x_1 + \dots + a_{kn}x_n.$$

Niech $\mathfrak{A} = (a_{ik})$ i niech

$$(15) \quad R(\mathfrak{A}) = r.$$

Udowodnimy następujące twierdzenie pomocnicze:

Przy założeniu (15) można tak wybrać r spośród form (14), np. formy $f_1(x), \dots, f_r(x)$, że

$$(16) \quad f_j(x) = c_{j1}f_1(x) + c_{j2}f_2(x) + \dots + c_{jr}f_r(x),$$

gdzie $j = r+1, r+2, \dots, m$ i gdzie liczby c_{jk} są stałe.

Innymi słowy: gdy macierz układu form liniowych ma rząd r , to można tak wybrać r form, że pozostałe m form są ich kombinacjami liniowymi o stałych współczynnikach.

Dla dowodu zauważmy wpierw, że macierz

$$\mathcal{A}_1 = \begin{pmatrix} a_{11} & \dots & a_{1n} & f_1(x) \\ a_{21} & \dots & a_{2n} & f_2(x) \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & \dots & a_{mn} & f_m(x) \end{pmatrix}$$

ma ten sam rząd co macierz

$$\mathfrak{A}_2 = \begin{pmatrix} a_{11} & \dots & a_{1n} & 0 \\ a_{21} & \dots & a_{2n} & 0 \\ \dots & \dots & \dots & \dots \\ a_{m1} & \dots & a_{mn} & 0 \end{pmatrix},$$

gdyż macierz $\mathfrak{U}_1$ powstaje z macierzy $\mathfrak{U}_2$ przez dodanie do jej ostatniej kolumny poprzednich kolumn, pomnożonych kolejno przez $x_1, x_2, \dots, x_n$. Lecz $R(\mathfrak{U}_2)=r$, gdyż każdy minor macierzy $\mathfrak{U}$ jest minorem macierzy $\mathfrak{U}_2$, pozostałe zaś minory macierzy $\mathfrak{U}_2$ mają kolumnę złożoną z samych zer. Ponieważ $R(\mathfrak{U}_1)=R(\mathfrak{U}_2)$, więc

$$(17) \quad R(\mathfrak{U}_1) = r.$$

Możemy założyć, że minorem r -tego stopnia macierzy $\mathfrak{A}_1$ różnym od zera jest minor

$$(18) \quad \Delta = \begin{vmatrix} a_{11} & \dots & a_{1r} \\ \vdots & \ddots & \vdots \\ a_{r1} & \dots & a_{rr} \end{vmatrix} \neq 0.$$

Gdyby bowiem był nim inny minor r -tego stopnia macierzy $\mathfrak{A}_1$, wystarczyłoby zmienić odpowiednio numerację form i zmiennych, by uzyskać (18).

Następujące minory $(r+1)$ -go stopnia macierzy $\mathfrak{A}_1$ są równe zeru na mocy (17):

$$\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1r} & f_1(x) \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{r1} & a_{r2} & \dots & a_{rr} & f_r(x) \\ a_{j1} & a_{j2} & \dots & a_{jr} & f_j(x) \end{vmatrix} = 0, \quad \text{gdzie } j = r+1, \dots, m.$$



Rozwijając je według ostatniej kolumny, otrzymujemy

$$\Delta f_j(x) + d_{j1}f_1(x) + \dots + d_{jr}f_r(x) = 0,$$

stąd zaś wobec (18) wynika (16), c. b. d. o.

Udowodnione twierdzenie pomocnicze nie ulega oczywiście zmianie, gdy formy jednorodne $f_k(x)$ zastąpić przez niejednorodne

$$F_k(x) = a_{k1}x_1 + \dots + a_{kn}x_n + b_k,$$

a macierz $\mathcal{U}$ przez macierz

$$\mathcal{N}_0 = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{pmatrix}.$$

Wystarczy bowiem udowodnione twierdzenie zastosować do form

$$a_{k1}x_1 + \dots + a_{kn}x_n + b_kx_{n+1},$$

a następnie podstawić $x_{n+1}=1$.

Weźmy teraz pod uwagę układ równań liniowych w postaci najogólniejszej:

$$(19) \quad \begin{aligned} F_1(x) &= a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n + b_1 = 0, \\ F_2(x) &= a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n + b_2 = 0, \\ &\vdots \\ F_m(x) &= a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n + b_m = 0. \end{aligned}$$

Niech

$$(20) \quad R(\mathfrak{U}) = r > 0$$

oraz

$$\mathfrak{B} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \dots & \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{pmatrix}.$$

Macierz $\mathfrak{B}$ powstaje więc z macierzy $\mathfrak{A}$ przez dopisanie kolumny wyrazów wolnych układu równań (19).

Zachodzi następujące twierdzenie Capelliego i Kroneckera:
Ażby układ równań (19) miał rozwiązanie, potrzeba i wystar-
cza, by

$$(21) \quad R(\mathfrak{U}) = R(\mathfrak{B}).$$

Dowód. Na podstawie twierdzeń już nam znanych (p. str. 505) zachodzi równość

$$(22) \quad R(\mathfrak{B}) = R \begin{pmatrix} a_{11} & \dots & a_{1n} & F_1(x) \\ a_{21} & \dots & a_{2n} & F_2(x) \\ \dots & \dots & \dots & \dots \\ a_{m1} & \dots & a_{mn} & F_m(x) \end{pmatrix}$$

dla każdego układu liczb $x_1, x_2, \dots, x_n$. W szczególności więc, jeśli ten układ liczb jest rozwiązaniem układu równań (19), to

$$R(\mathfrak{B}) = R \begin{pmatrix} a_{11} & \dots & a_{1n} & 0 \\ a_{21} & \dots & a_{2n} & 0 \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & \dots & a_{mn} & 0 \end{pmatrix} = R(\mathfrak{A}).$$

W ten sposób udowodniliśmy konieczność warunku (21).

Na odwrót, załóżmy (21). Wówczas na podstawie twierdzenia pomocniczego jest

$$F_j(x) = c_{j1}F_1(x) + c_{j2}F_2(x) + \dots + c_{jr}F_r(x) \quad \text{dla} \quad j = r+1, r+2, \dots, n.$$

Stąd wynika, że każde rozwiązanie układu pierwszych r równań

$$(23) \quad F_1(x)=0, \quad F_2(x)=0, \quad \dots, \quad F_r(x)=0,$$

jest zarazem rozwiązaniem układu wszystkich m równań (19) i na
odwrot.

Jest oczywiście $r \leq n$. Jeśli $r = n$, to układ (23) jest układem Cramera, a więc ma dokładnie jedno rozwiązanie. Jeśli zaś $r < n$, to otrzymujemy układ równań (4), o którym dowiedliśmy (p. str. 501), że ma nieskończenie wiele rozwiązań. Tym samym została dowiedziona również dostateczność warunku (21).

Z równoważności układów (19) i (23), stwierdzonej w toku dowodu dostateczności warunku (21), wynika, że *gdy macierz współczynników układu m równań liniowych jest rzędu r i Δ jest jej minorem r -tego stopnia różnym od zera, to odrzucając równania, których współczynniki nie są wyrazami minora Δ , dostajemy układ r równań, równoważny układowi (19).*

Rozważmy jeszcze układ równań liniowych jednorodnych kwadratowy:

[illegible]

Wówczas $R(\mathfrak{A}) = r < n$, możemy więc odrzucić $n - r$ równań. Pozostaje układ $r < n$ równań, który ma nieskończenie wiele rozwiązań, a w szczególności rozwiązania niezerowe. Zatem:

Na to, by układ równań liniowych jednorodnych kwadratowy miał rozwiązanie niezerowe, potrzeba i wystarcza, żeby $|\mathfrak{A}|=0$.

§ 3. Przekształcenia liniowe i mnożenie macierzy

1. Mnożenie macierzy kwadratowych. Weźmy pod uwagę wzory

$$(1) \quad \begin{aligned} y_1 &= a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n, \\ y_2 &= a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n, \\ &\vdots \\ y_n &= a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n. \end{aligned}$$

gdzie

$$(2) \quad |\mathfrak{A}| = |a_{ik}| \neq 0.$$

Każdemu ciągowi n wyrazów $x_1, x_2, \dots, x_n$ wzory (1) przyporządkowują ciąg tyluż wyrazów $y_1, y_2, \dots, y_n$. Jeżeli liczby $y_1, y_2, \dots, y_n$ są dane, to wobec (2) możemy wyznaczyć liczby $x_1, x_2, \dots, x_n$ za pomocą wzorów Cramera. Wzory (1) określają zatem pewne przekształcenie zbioru ciągów n -wyrazowych w ten sam zbiór.

Nazywamy je *przekształceniem liniowym*.

Niech będą dane dwa przekształcenia: (1) i

$$(3) \quad \begin{aligned} z_1 &= b_{11}y_1 + b_{12}y_2 + \dots + b_{1n}y_n, \\ z_2 &= b_{21}y_1 + b_{22}y_2 + \dots + b_{2n}y_n, \\ &\vdots \\ z_n &= b_{n1}y_1 + b_{n2}y_2 + \dots + b_{nn}y_n. \end{aligned}$$

Podstawiając wartości z (1) do (3), dostajemy wzory

$$z_i = c_{i1}x_1 + c_{i2}x_2 + \dots + c_{in}x_n \quad \text{dla} \quad i=1,2,\dots,n,$$

gdzie — jak łatwo sprawdzić —

$$(4) \quad c_{ik} = a_{i1}b_{1k} + a_{i2}b_{2k} + \dots + a_{in}b_{nk}.$$

Macierz

$$\mathbb{C} = (c_{ik})$$

nazywa się *iloczynem macierzy* $\mathbb{A}$ i $\mathbb{B}$ i oznacza się przez $\mathbb{AB}$.

Poczyn macierzy jest więc określony tylko dla macierzy kwadratowych tego samego stopnia.

Wyraz c_{ik} macierzy $\mathbb{AB}$ powstaje w ten sposób, że elementy i -tego wiersza macierzy $\mathbb{B}$ mnożymy przez odpowiednie elementy k -tej kolumny macierzy $\mathbb{A}$ i dodajemy. Np. dla

$$\mathbb{A} = \begin{pmatrix} 0 & 1 & 0 \\ 2 & 1 & 0 \\ 1 & 1 & 0 \end{pmatrix}, \quad \mathbb{B} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

otrzymujemy

$$\mathbb{AB} = \begin{pmatrix} 1 \cdot 0 + 0 \cdot 2 + 1 \cdot 1 & 1 \cdot 1 + 0 \cdot 1 + 1 \cdot 1 & 1 \cdot 0 + 0 \cdot 0 + 1 \cdot 0 \\ 0 \cdot 0 + 1 \cdot 2 + 0 \cdot 1 & 0 \cdot 1 + 1 \cdot 1 + 0 \cdot 1 & 0 \cdot 0 + 1 \cdot 0 + 0 \cdot 0 \\ 0 \cdot 0 + 0 \cdot 2 + 2 \cdot 1 & 0 \cdot 1 + 0 \cdot 1 + 2 \cdot 1 & 0 \cdot 0 + 0 \cdot 0 + 2 \cdot 0 \end{pmatrix} =$$

$$= \begin{pmatrix} 1 & 2 & 0 \\ 2 & 1 & 0 \\ 2 & 2 & 0 \end{pmatrix}.$$

Natomiast $\mathbb{AB} = \begin{pmatrix} 0 & 1 & 0 \\ 2 & 1 & 2 \\ 1 & 1 & 1 \end{pmatrix}.$

Z powyższych dwóch przykładów jest widoczne, że *iloczyn macierzy nie jest przemienny*.

Można bezpośrednim rachunkiem sprawdzić *łączność mnożenia macierzy* czyli wzór

$$(5) \quad (\mathbb{AB})\mathbb{C} = \mathbb{A}(\mathbb{BC}).$$

Dla wyznacznika iloczynu dwóch macierzy zachodzi następujące *twierdzenie Cauchy'ego*

$$(6) \quad |\mathbb{AB}| = |\mathbb{A}| \cdot |\mathbb{B}|.$$

Twierdzenie to sprawdzimy tu dla wyznaczników 2-go stopnia (dla wyznaczników 3-go stopnia dowód jest analogiczny):

$$\begin{vmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{vmatrix} =$$

$$= \begin{vmatrix} a_{11}b_{11} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} & a_{21}b_{12} + a_{22}b_{22} \end{vmatrix} + \begin{vmatrix} a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{vmatrix} =$$

$$= b_{11} \begin{vmatrix} a_{11} & a_{11}b_{12} \\ a_{21} & a_{21}b_{12} \end{vmatrix} + b_{11} \begin{vmatrix} a_{11} & a_{12}b_{22} \\ a_{21} & a_{22}b_{22} \end{vmatrix} + b_{21} \begin{vmatrix} a_{12} & a_{11}b_{12} \\ a_{22} & a_{21}b_{12} \end{vmatrix} + b_{21} \begin{vmatrix} a_{12} & a_{12}b_{22} \\ a_{22} & a_{22}b_{22} \end{vmatrix};$$

wyciągając teraz przed każdy wyznacznik wspólny czynnik każdej kolumny i odrzucając wyznaczniki o dwóch jednakowych kolumnach, otrzymujemy

$$b_{11}b_{22} \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} + b_{21}b_{12} \begin{vmatrix} a_{12} & a_{11} \\ a_{22} & a_{21} \end{vmatrix} =$$

$$= \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} (b_{11}b_{22} - b_{21}b_{12}) = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \cdot \begin{vmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{vmatrix}.$$

W podobny sposób udowadnia się następujące uogólnienie twierdzenia Cauchy'ego:

Każdy minor p -tego stopnia macierzy $\mathbb{AB}$ daje się przedstawić w postaci $\Sigma A_j B_j$, gdzie A_j jest minorem p -tego stopnia macierzy $\mathbb{A}$, a B_j — macierzy $\mathbb{B}$.

W szczególności wynika stąd, że gdy każdy minor p -tego stopnia macierzy $\mathbb{A}$ jest równy 0, to i każdy minor p -tego stopnia macierzy $\mathbb{AB}$ jest równy 0. Wobec tego

$$(7) \quad R(\mathbb{AB}) \leq R(\mathbb{A}), \quad R(\mathbb{AB}) \leq R(\mathbb{B}).$$

2. Macierz odwrotna. Jeśli za pomocą wzorów Cramera wyrazić $x_1, x_2, \dots, x_n$ przez $y_1, y_2, \dots, y_n$, to ze wzorów (1) otrzymuje się wzory:

$$x_1 = \frac{1}{A} \begin{vmatrix} y_1 & a_{12} & \dots & a_{1n} \\ y_2 & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ y_n & a_{n2} & \dots & a_{nn} \end{vmatrix}$$

i analogiczne dla $x_2, \dots, x_n$. Ogólnie więc dla $i=1, 2, \dots, n$

$$(8) \quad x_i = b_{i1}y_1 + b_{i2}y_2 + \dots + b_{in}y_n,$$

gdzie

$$b_{ik} = \frac{A_{ki}}{A}.$$

Wzory (8) również określają przekształcenie liniowe.

Nazywa się je *przekształceniem odwrotnym* przekształcenia (1).

Macierz przekształcenia (8) nazywa się *macierzą odwrotną* macierzy $\mathfrak{U}=(a_{ik})$ i oznacza się przez $\mathfrak{U}^{-1}$. Wobec tego

$$(9) \quad \mathfrak{U}^{-1} = \left(\frac{A_{ki}}{A} \right).$$

Przyjmujemy oznaczenie

$$(10) \quad \mathfrak{J} = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 \end{pmatrix}.$$

Łatwo widzieć, że dla każdej macierzy $\mathfrak{U}$ zachodzą równości

$$(11) \quad \mathfrak{U}\mathfrak{J} = \mathfrak{J}\mathfrak{U} = \mathfrak{U},$$

a także na mocy wzorów (19) i (20) z § 1, str. 498, równości

$$(12) \quad \mathfrak{U}\mathfrak{U}^{-1} = \mathfrak{J}, \quad \mathfrak{U}^{-1}\mathfrak{U} = \mathfrak{J}.$$

Oczywiście

$$(\mathfrak{U}^{-1})^{-1} = \mathfrak{U},$$

ponieważ przekształceniem odwrotnym do (8) jest (1). Z (12) wynika dla wyznacznika macierzy odwrotnej, że

$$|\mathfrak{U}| \cdot |\mathfrak{U}^{-1}| = |\mathfrak{J}| = 1$$

czyli

$$(13) \quad |\mathfrak{U}^{-1}| = \frac{1}{|\mathfrak{U}|} = |\mathfrak{U}|^{-1}.$$

Macierz *przestawioną* $\mathfrak{U}^+$ macierzy $\mathfrak{U}$ nazwaliśmy macierz powstającą z $\mathfrak{U}$ przez zamianę wierszy na kolumny (p. § 1, str. 495). Łatwo sprawdzić, że

$$(14) \quad (\mathfrak{U}\mathfrak{B})^+ = \mathfrak{B}^+\mathfrak{U}^+, \quad (\mathfrak{U}^+)^{-1} = (\mathfrak{U}^{-1})^+.$$

Jeśli $|\mathfrak{U}| \neq 0$, to każde z równań macierzowych

$$(15) \quad \mathfrak{U}\mathfrak{X} = \mathfrak{B}, \quad \mathfrak{Y}\mathfrak{U} = \mathfrak{B}$$

ma dokładnie jedno rozwiązanie:

$$(16) \quad \mathfrak{X} = \mathfrak{U}^{-1}\mathfrak{B}, \quad \mathfrak{Y} = \mathfrak{B}\mathfrak{U}^{-1}.$$

Udowodnimy to np. dla pierwszego z równań (15). Mnożąc lewostronnie przez $\mathfrak{U}^{-1}$, otrzymujemy

$$\mathfrak{U}^{-1}\mathfrak{B} = \mathfrak{U}^{-1}(\mathfrak{U}\mathfrak{X}) = (\mathfrak{U}^{-1}\mathfrak{U})\mathfrak{X} = \mathfrak{J}\mathfrak{X} = \mathfrak{X};$$

na odwrót, podstawiając $\mathfrak{X}$ z (16) do (15), otrzymujemy

$$\mathfrak{U}(\mathfrak{U}^{-1}\mathfrak{B}) = (\mathfrak{U}\mathfrak{U}^{-1})\mathfrak{B} = \mathfrak{J}\mathfrak{B} = \mathfrak{B}.$$

Jeśli $|\mathfrak{U}| \neq 0$, to dla każdej macierzy $\mathfrak{B}$ zachodzi równość

$$(17) \quad R(\mathfrak{U}\mathfrak{B}) = R(\mathfrak{B}).$$

W istocie, przyjmując oznaczenie

$$\mathfrak{C} = \mathfrak{U}\mathfrak{B},$$

stwierdzamy na mocy (7), że

$$(18) \quad R(\mathfrak{C}) \leq R(\mathfrak{B}),$$

a ponieważ

$$\mathfrak{B} = \mathfrak{U}^{-1}\mathfrak{C},$$

więc stosując ponownie (7), mamy wobec nierówności $|\mathfrak{U}|^{-1} \neq 0$, wynikającej z (13),

$$R(\mathfrak{B}) \leq R(\mathfrak{C}).$$

Stąd $R(\mathfrak{C}) = R(\mathfrak{B})$ czyli, wobec (18), równość (17).

3. Mnożenie macierzy prostokątnych. Niechaj

$$\mathfrak{U} = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \dots & \dots & \dots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \quad i \quad \mathfrak{B} = \begin{pmatrix} b_{11} & \dots & b_{1s} \\ \dots & \dots & \dots \\ b_{n1} & \dots & b_{ns} \end{pmatrix}.$$

Macierz $\mathfrak{U}$ ma więc tyle kolumn, ile macierz $\mathfrak{B}$ ma wierszy.

Możemy zatem utworzyć wyrażenia

$$c_{ik} = a_{i1}b_{1k} + a_{i2}b_{2k} + \dots + a_{in}b_{nk}, \quad \text{gdzie } i=1, \dots, m \text{ i } k=1, \dots, s.$$

Otrzymujemy w ten sposób macierz

$$\mathfrak{C} = (c_{ik}),$$

która ma tyleż wierszy co macierz $\mathfrak{U}$ i tyleż kolumn co macierz $\mathfrak{B}$.

Macierz $\mathfrak{C}$ nazywamy *iloczynem macierzy* $\mathfrak{A}$ i $\mathfrak{B}$ i oznaczamy przez $\mathfrak{AB}$.

Mnożenie macierzy kwadratowych, określone na str. 510, jest szczególnym przypadkiem obecnie określonego mnożenia macierzy prostokątnych.

Częsty jest przypadek, gdy macierz $\mathfrak{B}$ składa się z jednej tylko kolumny:

$$(19) \quad \mathfrak{B} = (x) = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

Wówczas

$$(20) \quad \mathfrak{AB} = \mathfrak{A}(x) = \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n \\ \dots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n \end{pmatrix}.$$

Widzimy więc, że wzory (1), określające przekształcenie liniowe, dają się napisać krótko w postaci

$$(y) = \mathfrak{A}(x).$$

Stąd też

$$(x) = \mathfrak{A}^{-1}(y).$$

W tym znakowaniu układ równań liniowych można napisać jednym równaniem macierzowym

$$\mathfrak{A}(x) = (b),$$

gdzie

$$(b) = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}.$$

Macierz przestawiona macierzy (19) jest złożona z jednego tylko wiersza

$$(x)^+ = (x_1 \ x_2 \ \dots \ x_n).$$

Z (20) wynika równość

$$(21) \quad (y)^+ \mathfrak{A}(x) = (f),$$

gdzie f jest wielomianem

$$(22) \quad \begin{aligned} f &\equiv a_{11}x_1y_1 + a_{12}x_1y_2 + \dots + a_{1n}x_1y_n + \dots \\ &\dots \\ &\dots + a_{n1}x_ny_1 + a_{n2}x_ny_2 + \dots + a_{nn}x_ny_n. \end{aligned}$$

4. Macierze ortogonalne. Macierz kwadratowa nazywa się *ortogonalną*, gdy

$$(23) \quad \mathfrak{A}^+ \mathfrak{A} = \mathfrak{J}.$$

Tak np. macierzą ortogonalną 2-go stopnia jest dla dowolnego φ macierz

$$\begin{pmatrix} \cos \varphi & \sin \varphi \\ \sin \varphi & -\cos \varphi \end{pmatrix};$$

macierzą ortogonalną 3-go stopnia jest macierz

$$\begin{pmatrix} \frac{3}{2} & -\frac{1}{3} & \frac{2}{3} \\ -\frac{11}{15} & -\frac{2}{15} & \frac{2}{3} \\ \frac{2}{15} & \frac{14}{15} & \frac{1}{3} \end{pmatrix}.$$

Warunek (23) jest równoważny każdej z równości

$$(24) \quad \mathfrak{A}^+ = \mathfrak{A}^{-1}, \quad \mathfrak{A} \mathfrak{A}^+ = \mathfrak{J}.$$

Wyznacznik macierzy ortogonalnej jest równy ± 1 ; istotnie, na mocy (23) z $|\mathfrak{A}^+| = |\mathfrak{A}|$ wynika, że

$$|\mathfrak{A}|^2 = 1 \quad \text{i} \quad |\mathfrak{A}| = \pm 1.$$

Niech $\mathfrak{A}^+ = (b_{ik})$. Wówczas $b_{ik} = a_{ki}$. Ponieważ $\mathfrak{A}^+ \mathfrak{A} = (c_{ik})$, gdzie

$$c_{ik} = b_{i1}a_{1k} + b_{i2}a_{2k} + \dots + b_{in}a_{nk} = a_{i1}a_{1k} + a_{i2}a_{2k} + \dots + a_{in}a_{nk},$$

więc z (23) wynika, że na to, by macierz $\mathfrak{A}$ była ortogonalna, potrzeba i wystarczy, żeby dla kolumn macierzy $\mathfrak{A}$ zachodziły równości

$$(25) \quad a_{i1}a_{1k} + a_{i2}a_{2k} + \dots + a_{in}a_{nk} = \begin{cases} 1 & \text{dla } i=k, \\ 0 & \text{dla } i \neq k. \end{cases}$$

Podobnie, stosując drugą z równości (24), otrzymujemy jako warunek ortogonalności macierzy $\mathfrak{A}$ następujące równości dla wierszy:

$$(26) \quad a_{i1}a_{k1} + a_{i2}a_{k2} + \dots + a_{in}a_{kn} = \begin{cases} 1 & \text{dla } i=k \\ 0 & \text{dla } i \neq k. \end{cases}$$

Iloczyn dwóch macierzy ortogonalnych jest macierzą ortogonalną; istotnie, z równości

$$\mathfrak{U}^+ = \mathfrak{U}^{-1}, \quad \mathfrak{B}^+ = \mathfrak{B}^{-1},$$

wynika, że

$$(\mathfrak{U}\mathfrak{B})^+ = \mathfrak{B}^+\mathfrak{U}^+ = \mathfrak{B}^{-1}\mathfrak{U}^{-1} = (\mathfrak{U}\mathfrak{B})^{-1}.$$

Macierz odwrotna macierzy ortogonalnej jest ortogonalna; istotnie, gdy zachodzi (23), to wobec (14) jest

$$(\mathfrak{U}^{-1})^+\mathfrak{U}^{-1} = (\mathfrak{U}^+)^+\mathfrak{U}^{-1} = \mathfrak{U}\mathfrak{U}^{-1} = \mathfrak{I}.$$

§ 4. Formy dwuliniowe i kwadratowe

1. Określenia. Wielomian postaci

$$a_{11}x_1y_1 + a_{12}x_1y_2 + \dots + a_{n-1n}x_{n-1}y_n + a_{nn}x_ny_n$$

nazywa się *formą dwuliniową* zmiennych $x_1, x_2, \dots, x_n, y_1, y_2, \dots, y_n$.

Oznaczamy go znakiem $f(x|y)$.

Jest to wielomian drugiego stopnia względem tych zmiennych. Dla ustalonych $x_1, x_2, \dots, x_n$ jest on stopnia pierwszego względem $y_1, y_2, \dots, y_n$ i, tak samo, dla ustalonych $y_1, y_2, \dots, y_n$ jest on pierwszego stopnia względem $x_1, x_2, \dots, x_n$.

Forma dwuliniowa jest wyznaczona przez macierz $\mathfrak{U}$ jej współczynników. Formie dwuliniowej o macierzy $\mathfrak{U}$ odpowiada wzajemnie jednoznacznie iloczyn postaci $(y)^+\mathfrak{U}(x)$ (por. wzór (20), str. 514); dlatego będziemy używali również nazwy *forma* $(y)^+\mathfrak{U}(x)$.

Formą kwadratową nazywa się wielomian jednorodny drugiego stopnia.

Każdą formę kwadratową zmiennych $x_1, x_2, \dots, x_n$ można napisać jednym i tylko jednym sposobem w postaci

$$(1) \quad \sum_{i,k=1}^n a_{ik}x_ix_k,$$

przyjmując, że

$$(2) \quad a_{ik} = a_{ki}$$

(znakiem $\sum_{i,k=1}^n a_{ik}x_ix_k$ oznacza się — jak wiadomo — sumę wyrazów postaci $a_{ik}x_ix_k$, gdzie wskaźniki i, k przebiegają wszystkie możliwe pary wartości spośród liczb $1, 2, \dots, n$).

Jeśli bowiem po redukcji wyrazów podobnych iloczyn x_ix_k występuje ze współczynnikiem a , to warunkiem koniecznym i dostatecznym równości (2) jest, żeby

$$a_{ik} = a_{ki} = \begin{cases} a/2 & \text{dla } i \neq k, \\ a & \text{dla } i = k. \end{cases}$$

Postać (1) z warunkiem (2) jest powszechnie przyjętym sposobem pisanie formy kwadratowej.

Np. forma kwadratowa

$$(3) \quad x^2 - 3xy + z^2 + 6zx,$$

daje się napisać w tej postaci, jak następuje:

$$x^2 - \frac{3}{2}xy + 0y^2 - \frac{3}{2}yx + z^2 + 0yz + 0zy + 3zx + 3xz.$$

Macierz $\mathfrak{U} = (a_{ik})$ nazywa się *macierzą formy* (1), a wyznacznik $|\mathfrak{U}|$ — jej *wyróżnikiem*.

Np. macierz i wyróżnik formy kwadratowej (3) są następujące:

$$\mathfrak{U} = \begin{pmatrix} 1 & -3/2 & 3 \\ -3/2 & 0 & 0 \\ 3 & 0 & 1 \end{pmatrix}, \quad |\mathfrak{U}| = -\frac{9}{4}.$$

Macierz spełniająca warunek (2) nazywa się *symetryczną*.

Z każdej formy dwuliniowej otrzymuje się formę kwadratową przez podstawienie

$$(4) \quad y_1 = x_1, \quad y_2 = x_2, \quad \dots, \quad y_n = x_n.$$

Macierzą tak otrzymanej formy kwadratowej nie musi jednak być macierz pierwotnej formy dwuliniowej, ponieważ macierz formy pierwotnej może nie być symetryczna. Przez podstawienie wartości (4) do wzoru (22) z § 3, str. 515, otrzymujemy formę kwadratową

$$\sum_{i,k=1}^n a_{ik}x_ix_k,$$

której macierzą jest $\mathfrak{B} = (b_{ik})$, gdzie

$$b_{ik} = \frac{a_{ik} + a_{ki}}{2}.$$

Formie kwadratowej o macierzy $\mathfrak{A}$ odpowiada wzajemnie jednoznacznie iloczyn $(x)^+\mathfrak{A}(x)$; dlatego będziemy również używali nazwy *forma* $(x)^+\mathfrak{A}(x)$.

Formą biegunową formy kwadratowej (1) nazywa się forma dwuliniowa

$$(5) \quad \sum_{i,k=1}^n a_{ik} x_i x_k.$$

Formy biegunowe różnią się od pozostałych form dwuliniowych tym, że mają macierze symetryczne.

Zgodnie z tym, że dla formy dwuliniowej (5) przyjęliśmy znak $f(x|y)$, formę kwadratową (1) możemy oznaczać znakiem $f(x|x)$.

2. Przekształcenia liniowe form kwadratowych. Niech będzie dana forma

$$(6) \quad f(x|x) = \sum_{i,k=1}^n a_{ik} x_i x_k.$$

Wprowadzamy podstawienie liniowe

$$(7) \quad x_i = b_{i1} u_1 + b_{i2} u_2 + \dots + b_{in} u_n, \quad \text{gdzie } i=1, 2, \dots, n;$$

zatem

$$(8) \quad (x) = \mathfrak{B}(y),$$

gdzie

$$(9) \quad |\mathfrak{B}| \neq 0.$$

Podstawiając wartości (7) do (6), dostajemy nową formę kwadratową $g(u|u)$ zmiennych $u_1, u_2, \dots, u_n$. Aby znaleźć jej macierz i wyróżnik, napiszmy (7) krótko w postaci

$$x_i = \sum_{s=1}^n b_{is} u_s, \quad \text{gdzie } i=1, 2, \dots, n$$

i podstawmy do (6). Otrzymujemy

$$\begin{aligned} g(u|u) &= \sum_{i,k=1}^n a_{ik} \left(\sum_{s=1}^n b_{is} u_s \right) \left(\sum_{p=1}^n b_{kp} u_p \right) = \\ &= \sum_{i,k=1}^n \sum_{s=1}^n \sum_{p=1}^n a_{ik} b_{is} b_{kp} u_s u_p = \\ &= \sum_{s,p=1}^n (u_s u_p \sum_{i,k=1}^n a_{ik} b_{is} b_{kp}), \end{aligned}$$

stąd zaś

$$(10) \quad g(u|u) = \sum_{s,p=1}^n c_{sp} u_s u_p,$$

gdzie

$$(11) \quad c_{sp} = \sum_{i,k=1}^n a_{ik} b_{is} b_{kp}.$$

Dla macierzy $(c_{sp}) = \mathfrak{C}$ zachodzi równość

$$(12) \quad \mathfrak{C} = \mathfrak{B}^+ \mathfrak{A} \mathfrak{B},$$

wynikająca bezpośrednio z (11) na mocy definicji mnożenia macierzy prostokątnych.

Jeśli poddamy przekształceniu (8) macierz $(x)^+\mathfrak{A}(x)$ odpowiadającą formie (6), to dojdziemy do macierzy odpowiadającej formie przekształconej (10); istotnie

$$(x)^+\mathfrak{A}(x) = [\mathfrak{B}(u)]^+ \mathfrak{A} [\mathfrak{B}(u)] = (u)^+ \mathfrak{B}^+ \mathfrak{A} \mathfrak{B}(u) = (u)^+ \mathfrak{C}(u),$$

gdzie $\mathfrak{C}$ jest macierzą określoną wzorem (12).

Z założenia (9) wynika na mocy wzoru (17) z § 3, str. 513, że

$$R(\mathfrak{C}) = R(\mathfrak{A}).$$

A zatem: *rzęd macierzy formy kwadratowej jest niezmiennikiem przekształceń liniowych.*

Jeżeli forma kwadratowa i przekształcenie liniowe mają współczynniki rzeczywiste, to stosując twierdzenie Cauchy'ego o mnożeniu wyznaczników, otrzymujemy

$$|\mathfrak{C}| = |\mathfrak{B}^+ \cdot \mathfrak{A} \cdot \mathfrak{B}| = |\mathfrak{B}|^2 |\mathfrak{A}|.$$

A zatem: *znak wyróżnika formy kwadratowej o współczynnikach rzeczywistych jest niezmiennikiem przekształceń liniowych o współczynnikach rzeczywistych.*

Udowodnimy jeszcze, że przez przekształcenie liniowe forma biegunowa $f(x|y)$ formy kwadratowej $f(x|x)$ przechodzi w formę biegunową $g(u|v)$ formy przekształconej $g(u|u)$.

Dla krótkości użyjemy macierzy. Formie $f(x|y)$ odpowiada macierz jednowyrazowa

$$(f(x|y)) = (x)^+ \mathfrak{A}(y).$$

Po przekształceniu liniowym otrzymujemy

$$(x) = \mathfrak{B}(u), \quad (y) = \mathfrak{B}(v),$$

więc

$$(x)^+ \mathfrak{A}(y) = (u)^+ \mathfrak{B}^+ \mathfrak{A} \mathfrak{B}(v) = (u)^+ \mathfrak{C}(v) = (g(u|v)),$$

gdzie $\mathfrak{C}$ jest macierzą określoną wzorem (12).

3. Sprowadzenie formy kwadratowej do sumy kwadratów.

Udowodnimy, że istnieje przekształcenie liniowe formy kwadratowej (6) w formę

$$(13) \quad a_1 u_1^2 + a_2 u_2^2 + \dots + a_n u_n^2.$$

Dowód przeprowadzimy, za Lagrangem z pomocą indukcji względem ilości zmiennych n .

Dla $n=1$ twierdzenie jest prawdziwe, bo sama forma (6) ma postać

$$f(x|x) = a_{11} x_1^2.$$

Założmy, że twierdzenie jest prawdziwe dla $n-1$ zmiennych. Rozróżnimy dwa przypadki:

1° Istnieje i , dla którego $a_{ii} \neq 0$. Niech np.

$$(14) \quad a_{11} \neq 0.$$

Zbierając osobno wyrazy zawierające x_1 , otrzymujemy

$$\begin{aligned} f(x|x) &= a_{11} x_1^2 + 2a_{12} x_1 x_2 + \dots + 2a_{1n} x_1 x_n + \sum_{i,k=2}^n a_{ik} x_i x_k = \\ &= \frac{1}{a_{11}} (a_{11}^2 x_1^2 + 2a_{12} a_{11} x_1 x_2 + \dots + 2a_{1n} a_{11} x_1 x_n) + \sum_{i,k=2}^n a_{ik} x_i x_k. \end{aligned}$$

Wielomian w nawiasie zawiera wszystkie wyrazy kwadratu

$$(a_{11} x_1 + a_{12} x_2 + \dots + a_{1n} x_n)^2,$$

w których występuje zmienna x_1 . A więc

$$(15) \quad f(x|x) = \frac{1}{a_{11}} (a_{11} x_1 + \dots + a_{1n} x_n)^2 + f_1(x_2, x_3, \dots, x_n),$$

gdzie f_1 jest formą kwadratową $n-1$ zmiennych. Poddajmy formę (15) przekształceniu

$$(16) \quad v_1 = a_{11} x_1 + \dots + a_{1n} x_n, \quad v_2 = x_2, \quad \dots, \quad v_n = x_n,$$

czyli podstawieniu

$$(17) \quad x_1 = \frac{1}{a_{11}} v_1 - \frac{a_{12}}{a_{11}} v_2 - \dots - \frac{a_{1n}}{a_{11}} v_n, \quad x_2 = v_2, \quad \dots, \quad x_n = v_n.$$

Wyznacznik tego przekształcenia, jako równy $1/a_{11}$, jest różny od 0. Na mocy (15)

$$(18) \quad f(x|x) = g(v|v) = \frac{1}{a_{11}} v_1^2 + f_1(v_2, v_3, \dots, v_n).$$

Ponieważ f_1 jest formą $n-1$ zmiennych, więc w myśl przyjętego w indukcji założenia istnieje przekształcenie

$$v_2 = a_{22} u_2 + \dots + a_{2n} u_n,$$

$$\dots$$

$$v_n = a_{n2} u_2 + \dots + a_{nn} u_n,$$

mające wyznacznik różny od 0, które formę f_1 przekształca w formę

$$g_1(u|u) = a_2 u_2^2 + \dots + a_n u_n^2.$$

Wówczas przekształcenie

$$v_1 = u_1,$$

$$v_2 = a_{22} u_2 + \dots + a_{2n} u_n,$$

$$\dots$$

$$v_n = a_{n2} u_2 + \dots + a_{nn} u_n$$

ma również wyznacznik różny od 0 i przeprowadza formę (18) w formę

$$h(u|u) = \frac{1}{a_{11}} u_1^2 + a_2 u_2^2 + \dots + a_n u_n^2$$

czyli w formę kształtu (13).

Sprowadzenie formy (6) do formy (13) za pomocą dwóch kolejnych kroków (17) i (19) jest równoważne sprowadzeniu formy (6) do formy (13) za pomocą jednego przekształcenia, które jest iloczynem przekształceń (17) i (19).

2° Dla każdego $i=1, 2, \dots, n$ jest $a_{ii} = 0$. Wówczas niech np.

$$a_{12} \neq 0$$

czyli

$$f(x|x) = 2a_{12} x_1 x_2 + \dots,$$

gdzie każdy z wyrazów wykropkowanych zawiera co najmniej jedną ze zmiennych $x_3, \dots, x_n$. Zatem

$$f(x|x) = \frac{1}{2} a_{12} (x_1 + x_2)^2 - \frac{1}{2} a_{12} (x_1 - x_2)^2 + \dots$$

Przez przekształcenie liniowe

$$v_1 = x_1 + x_2, \quad v_2 = x_1 - x_2, \quad v_3 = x_3, \quad \dots, \quad v_n = x_n$$

o wyznaczniku różnym od 0 otrzymujemy

$$f(x|x) = h(v|v) = \frac{1}{2} a_{12} v_1^2 - \frac{1}{2} a_{12} v_2^2 + \dots$$

Każdy z wyrazów wykropkowanych zawiera co najmniej jedną ze zmiennych $u_3, \dots, u_n$, a więc te wyrazy nie mogą się zredukować

z wyrazem u_1^2 . Ponieważ w nowej formie występują kwadraty, więc można ją dalej sprowadzić do postaci (13) jak w przypadku 1^o.

Twierdzenie zostało zatem dowiedzione w obu przypadkach.

W dowodzie użyliśmy wyłącznie *przekształceń, w których współczynniki są funkcjami wymiernymi współczynników formy pierwotnej*. Dowiedliśmy więc zarazem, że jeżeli forma $f(x|x)$ ma same współczynniki rzeczywiste, to istnieje przekształcenie o współczynnikach rzeczywistych, które sprowadza tę formę do sumy kwadratów.

Formę kwadratową (6) można oczywiście sprowadzić do postaci (13) nieskończenie wieloma sposobami, gdyż przekształcenie liniowe

$$u_1 = a_1 v_1, \quad u_2 = a_2 v_2, \quad \dots, \quad u_n = a_n v_n$$

o współczynnikach dowolnych, byleby różnych od 0, przeprowadza formę kształtu (6) w formę kształtu (13).

W szczególności gdy forma (13) ma współczynniki rzeczywiste, można ją przekształcić tak, by wszystkie jej współczynniki różne od zera były równe ± 1 . W tym celu wystarczy poddać formę (13) przekształceniu

$$v_1 = \sqrt{|a_1|} u_1, \quad v_2 = \sqrt{|a_2|} u_2, \quad \dots, \quad v_n = \sqrt{|a_n|} u_n;$$

otrzymujemy

$$h(v|v) = \pm v_1^2 \pm v_2^2 \pm \dots$$

Podobnie, jeśli dopuścimy przekształcenia o współczynnikach zespolonych, to zawsze będziemy mogli osiągnąć to, żeby współczynniki różne od zera były równe 1; wystarczy użyć przekształcenia

$$v_1 = \sqrt{a_1} u_1, \quad v_2 = \sqrt{a_2} u_2, \quad \dots, \quad v_n = \sqrt{a_n} u_n.$$

Wreszcie, przy każdym przekształceniu formy kwadratowej (6) w formę (13) ilość współczynników różnych od zera jest ta sama; jest ona równa rzędowi macierzy formy.

Istotnie, rząd tej macierzy jest niezmiennikiem przekształceń liniowych (p. str. 519). Macierzą formy (13) jest

$$\mathfrak{A} = \begin{pmatrix} a_1 & 0 & \dots & 0 \\ 0 & a_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & a_n \end{pmatrix},$$

a więc rząd $R(\mathfrak{A})$ jest równy ilości liczb $a_i \neq 0$.

4. Twierdzenie o bezwładności form kwadratowych. Twierdzenie to, odkryte i dowiedzione niezależnie przez Sylwestera i Jacobiego, brzmi:

Jeśli formę kwadratową (6) o współczynnikach rzeczywistych sprowadzić do form kształtu (13) za pomocą dwóch różnych przekształceń o współczynnikach rzeczywistych, to obie otrzymane formy mają zarówno te same ilości współczynników dodatnich jak i te same ilości współczynników ujemnych.

Istotnie, niech przekształcenie

$$(20) \quad x_i = b_{i1} u_1 + b_{i2} u_2 + \dots + b_{in} u_n, \quad \text{gdzie} \quad i=1,2,\dots,n,$$

o współczynnikach b_{ik} rzeczywistych i takich, że $|b_{ik}| \neq 0$, sprowadza formę $f(x|x)$ do formy

$$(21) \quad g(u|u) = a_1 u_1^2 + \dots + a_k u_k^2 - a'_1 u_{k+1}^2 - \dots - a'_l u_{k+l}^2,$$

gdzie

$$(22) \quad \alpha_j > 0 \quad \text{i} \quad \alpha'_j > 0.$$

Niech drugie przekształcenie

$$(23) \quad x_i = c_{i1} v_1 + c_{i2} v_2 + \dots + c_{in} v_n, \quad \text{gdzie} \quad i=1,2,\dots,n,$$

o współczynnikach c_{ik} rzeczywistych i takich, że $|c_{ik}| \neq 0$, sprowadza tę samą formę $f(x|x)$ do formy

$$(24) \quad h(v|v) = \beta_1 v_1^2 + \dots + \beta_m v_m^2 - \beta'_1 v_{m+1}^2 - \dots - \beta'_p v_{m+p}^2,$$

gdzie

$$(25) \quad \beta_j > 0 \quad \text{i} \quad \beta'_j > 0.$$

Zachodzi równość

$$(26) \quad m + p = k + l,$$

ponieważ wspólna wartość jej obu stron jest równa rzędowi macierzy formy $f(x|x)$, który jest niezmiennikiem przekształceń liniowych (p. str. 519).

Należy dowieść, że

$$(27) \quad m = k \quad \text{i} \quad p = l.$$

Przypuśćmy, że tak nie jest, np. że

$$(28) \quad k > m.$$

Stąd wynika wobec (26), że

$$l < p.$$

Weźmy pod uwagę przekształcenia odwrotne do (20) i (23):

$$(29) \quad \begin{aligned} u_i &= b'_{i1}x_1 + b'_{i2}x_2 + \dots + b'_{in}x_n, \\ v_i &= c'_{i1}x_1 + c'_{i2}x_2 + \dots + c'_{in}x_n, \end{aligned} \quad \text{gdzie} \quad i=1, 2, \dots, n.$$

Tutaj u_i i v_i są formami liniowymi zmiennych $x_1, x_2, \dots, x_n$.

Weźmy dalej pod uwagę następujący układ $n-k+m$ równań liniowych jednorodnych o n zmiennych $x_1, x_2, \dots, x_n$:

$$(30) \quad \begin{aligned} v_1 &= 0, & v_2 &= 0, & \dots, & v_m &= 0, \\ u_{k+1} &= 0, & u_{k+2} &= 0, & \dots, & u_n &= 0. \end{aligned}$$

Ponieważ $n-k+m < n$ wobec (26), więc układ (30) ma rozwiązanie niezerowe (p. str. 501 i 508). Niech

$$(31) \quad x_1^0, x_2^0, \dots, x_n^0$$

będzie jednym z tych rozwiązań. Podstawiając te wartości do $f(x|x)$, otrzymujemy wobec (22) i (30)

$$(32) \quad f(x^0|x^0) = a_1 u_1^2 + \dots + a_k u_k^2 \geq 0,$$

a wobec (25) i (30)

$$f(x^0|x^0) = -\beta_1 v_{m+1}^2 - \dots - \beta_p v_{m+p}^2 \leq 0.$$

Zatem $f(x^0|x^0) = 0$, skąd wynika na mocy (22) i (32), że

$$u_1 = 0, \quad u_2 = 0, \quad \dots, \quad u_k = 0.$$

Widzimy więc, że dla układu liczb (31) jest

$$u_1 = u_2 = \dots = u_n = 0,$$

a stąd wynika wobec (20), że

$$x_1^0 = x_2^0 = \dots = x_n^0 = 0,$$

czyli że rozwiązanie (31) jest zerowe. Przypuszczenie, że zachodzi nierówność (28), prowadzi zatem do sprzeczności. Tak samo jest z przypuszczeniem, że zachodzi nierówność $k < m$ czyli $p > l$.

Tak więc równości (27) są dowiedzione, c. b. d. o.

Sprowadźmy formę kwadratową rzeczywistą przez przekształcenie rzeczywiste do postaci

$$(33) \quad u_1^2 + u_2^2 + \dots + u_k^2 - u_{k+1}^2 - \dots - u_r^2,$$

gdzie wyrazy o współczynnikach dodatnich poprzedzają pozostałe.

Z twierdzenia o bezwładności wynika, że dla każdej formy istnieje tylko jedna postać (33), do której ta forma daje się sprowadzić.

Nazywa się ona *postacią kanoniczną* tej formy.

Tak np. dla form dwóch zmiennych istnieje pięć postaci kanonicznych:

$$u_1^2 + u_2^2, \quad -u_1^2 - u_2^2, \quad u_1^2 - u_2^2, \quad u_1^2, \quad -u_1^2,$$

a dla form trzech zmiennych — te same oraz cztery postaci

$$u_1^2 + u_2^2 + u_3^2, \quad u_1^2 + u_2^2 - u_3^2, \quad u_1^2 - u_2^2 - u_3^2, \quad -u_1^2 - u_2^2 - u_3^2,$$

czyli razem dziewięć.

Forma kwadratowa rzeczywista nazywa się *określoną dodatnio*, gdy jej postać kanoniczna ma tylko współczynniki 0 i 1. Jeśli nie występuje w niej współczynnik 0, czyli gdy ma ona postać kanoniczną

$$u_1^2 + u_2^2 + \dots + u_n^2,$$

nazywa się ona *istotnie dodatnią*, w przeciwnym zaś razie — *nie-istotnie dodatnią*. Podobnie definiuje się formy *określone ujemnie*.

Na to, by forma $f(x|x)$ była określona dodatnio, potrzeba i wystarczy, żeby dla każdego układu liczb rzeczywistych $x_1, x_2, \dots, x_n$ zachodziła nierówność

$$(34) \quad f(x|x) \geq 0.$$

Na to, by forma $f(x|x)$ była istotnie dodatnią, potrzeba i wystarczy, żeby dla każdego układu niezerowego liczb rzeczywistych $x_1, x_2, \dots, x_n$ zachodziła nierówność

$$(35) \quad f(x|x) > 0.$$

Aby dowieść konieczności tych warunków, sprowadźmy formę $f(x|x)$ do postaci kanonicznej przez przekształcenie

$$(36) \quad x_i = b_{i1}u_1 + \dots + b_{in}u_n, \quad \text{gdzie} \quad i=1, 2, \dots, n.$$

Jeśli forma jest określona dodatnio, to

$$f(x|x) = u_1^2 + u_2^2 + \dots + u_r^2,$$

gdzie $r = R(\mathfrak{A})$. Stąd wynika (34).

Jeśli forma jest istotnie dodatnia, to $r = n$. Zatem równość $f(x|x) = 0$ pociąga za sobą równość

$$u_1 = u_2 = \dots = u_n = 0,$$

skąd $x_1 = x_2 = \dots = x_n = 0$ na mocy (36). Stąd dla układów niezerowych wynika (35).

Aby dowieść, że nierówność (34) dla wszelkich $x_1, x_2, \dots, x_n$ jest warunkiem dostatecznym na to, by forma $f(x|x)$ była określona dodatnio, przypuścimy, że w jej postaci kanonicznej występuje współczynnik -1 , a więc że

$$f(x|x) = u_1^2 + \dots - u_{k+1}^2 - \dots$$

Można tak dobrać układ $x_1, x_2, \dots, x_n$ liczb rzeczywistych, aby

$$u_1 = \dots = u_k = u_{k+2} = \dots = u_n = 0, \quad u_{k+1} = 1.$$

Dla tego układu jest

$$f(x|x) = -1$$

wbrew (34). Podobnie dowodzi się dostateczności warunku (35).

§ 5. Przekształcenia ortogonalne form kwadratowych

Przekształcenie liniowe rzeczywiste

$$(1) \quad x_i = b_{i1}u_1 + \dots + b_{in}u_n, \quad \text{gdzie } i=1, 2, \dots, n,$$

nazywa się *przekształceniem ortogonalnym*, gdy macierz $\mathcal{B} = (b_{ik})$ jest ortogonalna.

Na to, by przekształcenie (1) było ortogonalne, potrzeba i wystarcza, żeby przekształcało ono formę

$$(2) \quad x_1^2 + x_2^2 + \dots + x_n^2$$

w formę

$$(3) \quad u_1^2 + u_2^2 + \dots + u_n^2.$$

Istotnie, macierzą formy (2) jest $\mathcal{U} = \mathcal{I}$. Jak wiemy z § 4 (str. 519), przekształcenie (1) przekształca formę (2) w formę o macierzy

$$\mathcal{C} = \mathcal{B}^+ \mathcal{U} \mathcal{B} = \mathcal{B}^+ \mathcal{I} \mathcal{B} = \mathcal{B}^+ \mathcal{B}.$$

Ponieważ macierzą formy (3) również jest $\mathcal{I}$, więc na to, aby forma (2) przeszła w (3), potrzeba i wystarcza, żeby $\mathcal{B}^+ \mathcal{B} = \mathcal{I}$, czyli żeby macierz $\mathcal{U}$ była ortogonalna, c. b. d. o.

Zbadamy obecnie, jakie własności form kwadratowych rzeczywistych nie ulegają zmianie przy przekształceniach ortogonalnych.

Jednym takim *niezmiennikiem* jest rząd macierzy formy kwadratowej, gdyż jest to niezmiennik wszelkich przekształceń liniowych (§ 4, str. 519).

Dalszym *niezmiennikiem przekształceń ortogonalnych* jest wyróżnik $|\mathcal{U}|$ formy kwadratowej, gdyż dla wyróżnika $\mathcal{C}$ formy przekształconej jest (p. tamże)

$$|\mathcal{C}| = |\mathcal{B}^+ \mathcal{U} \mathcal{B}| = |\mathcal{B}^{-1} \mathcal{U} \mathcal{B}| = |\mathcal{B}^{-1}| |\mathcal{U}| |\mathcal{B}| = |\mathcal{U}|.$$

Niech

$$(4) \quad \mathcal{U}(\lambda) = \begin{pmatrix} a_{11} - \lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} - \lambda & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} - \lambda \end{pmatrix}.$$

Wyznacznik

$$(5) \quad w(\lambda) = |\mathcal{U}(\lambda)|$$

nazywa się *wielomianem charakterystycznym* formy $\sum a_{ik} x_i x_k$.

Jest to wielomian stopnia n :

$$(6) \quad w(\lambda) = (-1)^n (\lambda^n - A_1 \lambda^{n-1} + \dots + (-1)^n A_n).$$

Np. dla formy (2) jest

$$w(\lambda) = (1 - \lambda)^n.$$

Udowodnimy, że *niezmiennikami przekształceń ortogonalnych formy kwadratowej* są:

1° dla każdego λ rząd macierzy $\mathcal{U}(\lambda)$;

2° współczynniki wielomianu charakterystycznego.

Aby tego dowieść, weźmy pod uwagę przekształcenie ortogonalne T o macierzy $\mathcal{B}$, przekształcające formę kwadratową $f(x|x)$ w formę kwadratową $g(u|u)$. Przekształca więc ono formę

$$(7) \quad f(x|x) - \lambda(x_1^2 + x_2^2 + \dots + x_n^2)$$

w formę

$$(8) \quad g(u|u) - \lambda(u_1^2 + u_2^2 + \dots + u_n^2)$$

(por. str. 526). Niech $\mathcal{U}$ będzie macierzą formy $f(x|x)$, a

$$\mathcal{C} = \mathcal{B}^+ \mathcal{U} \mathcal{B}$$

macierzą formy $g(u|u)$. Wówczas $\mathcal{U}(\lambda)$ jest macierzą formy (7), a $\mathcal{C}(\lambda)$ macierzą formy (8). Ponieważ T przekształca (7) w (8), więc rzędy macierzy $\mathcal{U}(\lambda)$ i $\mathcal{C}(\lambda)$ są równe. Stąd niezmienniczość rzędu macierzy, czyli 1°.

Niezmienniczość zaś wielomianów charakterystycznych $|\mathfrak{A}(\lambda)|$ i $|\mathfrak{C}(\lambda)|$ wynika stąd, że te wielomiany są wyróżnikami form (7) i (8), a więc też są równe dla każdego λ . Dwa wielomiany jednak, które dla każdej wartości zmiennej przybierają te same wartości, mają te same współczynniki. Stąd 2^o.

Z udowodnionego twierdzenia nie wynika, że rząd macierzy $\mathfrak{A}(\lambda)$ nie zależy od λ . Zmieniając λ , możemy zmienić rząd; ustalwszy jednak λ , ustalamy ten rząd niezależnie od przekształceń ortogonalnych.

Pierwiastki wielomianu charakterystycznego nazywają się *pierwiastkami charakterystycznymi formy*.

Wskutek niezmienniczości współczynników tego wielomianu również *pierwiastki charakterystyczne formy są niezmiennikami przekształceń ortogonalnych*.

Pierwiastki charakterystyczne macierzy rzeczywistej i symetrycznej są liczbami rzeczywistymi.

Istotnie, niech λ spełnia równanie

$$(5) \quad \begin{vmatrix} a_{11}-\lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22}-\lambda & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn}-\lambda \end{vmatrix} = 0.$$

Mamy dowieść, że λ jest liczbą rzeczywistą, czyli że

$$(6) \quad \lambda = \bar{\lambda}$$

(przypominamy, że $\bar{\lambda}$ oznacza liczbę zespoloną sprzężoną z λ i że gdy $\lambda = a + bi$, to $\bar{\lambda} = a - bi$, oraz że $\lambda_1 + \lambda_2 = \overline{\lambda_1 + \lambda_2}$ i $\lambda_1 \lambda_2 = \overline{\lambda_1 \lambda_2}$).

Wyznacznik po lewej stronie równania (5) jest wyznacznikiem układu kwadratowego równań liniowych jednorodnych

$$(7) \quad \begin{aligned} (a_{11}-\lambda)x_1 + a_{12}x_2 + \dots + a_{1n}x_n &= 0, \\ a_{21}x_1 + (a_{22}-\lambda)x_2 + \dots + a_{2n}x_n &= 0, \\ \dots & \\ a_{n1}x_1 + a_{n2}x_2 + \dots + (a_{nn}-\lambda)x_n &= 0. \end{aligned}$$

Wobec (5) układ ten ma co najmniej jedno rozwiązanie niezerowe $x_1, x_2, \dots, x_n$ (na razie nie wiadomo, czy rzeczywiste); zatem

$$(8) \quad |x_1|^2 + |x_2|^2 + \dots + |x_n|^2 > 0.$$

Z (7) wynika, że

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n &= \lambda x_1, \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n &= \lambda x_2, \\ \dots & \\ a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n &= \lambda x_n. \end{aligned}$$

Mnożąc pierwszą z tych równości przez $\bar{x}_1$ (liczbę sprzężoną z x_1), drugą przez $\bar{x}_2$ itd., a następnie dodając je wszystkie stronami, otrzymujemy

$$(9) \quad \begin{aligned} a_{11}x_1\bar{x}_1 + a_{12}x_2\bar{x}_1 + a_{21}x_1\bar{x}_2 + \dots + a_{nn}x_n\bar{x}_n &= \\ = \lambda(x_1\bar{x}_1 + x_2\bar{x}_2 + \dots + x_n\bar{x}_n). \end{aligned}$$

Przejdźmy teraz po obu stronach do liczb sprzężonych. Ponieważ współczynniki a_{ik} są rzeczywiste, więc $\bar{a}_{ik} = a_{ik}$, a ponieważ macierz $\mathfrak{A}$ jest symetryczna, więc $a_{ik} = a_{ki}$. Zatem

$$\begin{aligned} a_{11}\bar{x}_1x_1 + a_{21}\bar{x}_2x_1 + a_{12}\bar{x}_1x_2 + \dots + a_{nn}\bar{x}_nx_n &= \\ = \bar{\lambda}(\bar{x}_1x_1 + \bar{x}_2x_2 + \dots + \bar{x}_nx_n). \end{aligned}$$

Stąd i z (9) wynika, że

$$\lambda(x_1\bar{x}_1 + x_2\bar{x}_2 + \dots + x_n\bar{x}_n) = \bar{\lambda}(\bar{x}_1x_1 + \bar{x}_2x_2 + \dots + \bar{x}_nx_n).$$

Ponieważ $x_k\bar{x}_k = |x_k|^2$, więc wobec (8) sumy w nawiasach są różne od zera, a że są one identyczne, więc musi zachodzić równość (6), c. b. d. o.

Udowodnimy teraz następujące twierdzenie pomocnicze:

Dla wszelkich n liczb rzeczywistych $b_{11}, b_{12}, \dots, b_{1n}$, spełniających warunek

$$b_{11}^2 + b_{12}^2 + \dots + b_{1n}^2 = 1,$$

istnieje macierz ortogonalna, której pierwszy wiersz jest złożony z tych n liczb.

Należy udowodnić istnienie n wierszy spełniających warunek

$$(10) \quad b_{i1}b_{j1} + b_{i2}b_{j2} + \dots + b_{in}b_{jn} = \begin{cases} 1 & \text{dla } i=j, \\ 0 & \text{dla } i \neq j. \end{cases}$$

W myśl założenia istnieje wiersz spełniający warunek (10) dla $j=i=1$. Załóżmy, że istnieje $k < n$ wierszy złożonych z liczb rzeczy-

wistych, spełniających warunek (10). Układ $k < n$ równań jednorodnych

$$(11) \quad \begin{aligned} b_{11}x_1 + b_{12}x_2 + \dots + b_{1n}x_n &= 0, \\ b_{21}x_1 + b_{22}x_2 + \dots + b_{2n}x_n &= 0, \\ \dots & \\ b_{k1}x_1 + b_{k2}x_2 + \dots + b_{kn}x_n &= 0 \end{aligned}$$

ma rozwiązanie niezerowe rzeczywiste $x_1, x_2, \dots, x_n$, dla którego

$$x_1^2 + x_2^2 + \dots + x_n^2 > 0.$$

Niech

$$(12) \quad \begin{aligned} b_{k+11} &= \frac{x_1}{\sqrt{x_1^2 + x_2^2 + \dots + x_n^2}}, \\ b_{k+12} &= \frac{x_2}{\sqrt{x_1^2 + x_2^2 + \dots + x_n^2}}, \\ \dots & \\ b_{k+1n} &= \frac{x_n}{\sqrt{x_1^2 + x_2^2 + \dots + x_n^2}}. \end{aligned}$$

Liczby (12) stanowią również rozwiązanie układu równań (11), a więc

$$b_{i1}b_{k+11} + b_{i2}b_{k+12} + \dots + b_{in}b_{k+1n} = 0 \quad \text{dla} \quad i=1, 2, \dots, k$$

oraz

$$b_{k+11}^2 + b_{k+12}^2 + \dots + b_{k+1n}^2 = 1.$$

Znaleźliśmy zatem $k+1$ wierszy spełniających warunek (10). Stąd przez indukcję wynika istnienie n takich wierszy, c. b. d. o.

Oczywiście, w dowiedzionym twierdzeniu pomocniczym możemy słowo *wiersz* zastąpić słowem *kolumna*.

W dowodzie następującego twierdzenia oprzemy się także na tym, że iloczyn dwu przekształceń ortogonalnych i przekształcenie odwrotne do ortogonalnego są znowu ortogonalne. Wynika to z prawdziwości twierdzeń dla macierzy, dowiedzionych w § 3 (p. str. 516).

Każdą formę kwadratową rzeczywistą można sprowadzić za pomocą przekształcenia ortogonalnego do sumy kwadratów

$$(13) \quad \lambda_1 x_1^2 + \lambda_2 x_2^2 + \dots + \lambda_n x_n^2;$$

liczby λ_k są tu pierwiastkami charakterystycznymi formy.

Istotnie, niech λ_1 będzie dowolnym pierwiastkiem charakterystycznym formy. Zatem równanie (5) jest spełnione, gdy podstawimy $\lambda = \lambda_1$. Tym samym dla tej wartości zmiennej λ układ równań (7) ma rozwiązanie niezerowe rzeczywiste $x_1, x_2, \dots, x_n$.

Niech

$$(14) \quad \begin{aligned} b_{11} &= \frac{x_1}{\sqrt{x_1^2 + x_2^2 + \dots + x_n^2}}, \\ b_{21} &= \frac{x_2}{\sqrt{x_1^2 + x_2^2 + \dots + x_n^2}}, \\ \dots & \\ b_{n1} &= \frac{x_n}{\sqrt{x_1^2 + x_2^2 + \dots + x_n^2}}. \end{aligned}$$

Stąd

$$(15) \quad b_{11}^2 + b_{21}^2 + \dots + b_{n1}^2 = 1$$

i liczby (14) też stanowią rozwiązanie niezerowe układu równań (7).

Wobec tego

$$(16) \quad a_{i1}b_{11} + a_{i2}b_{21} + \dots + a_{in}b_{n1} = \lambda_1 b_{i1}.$$

Z (14) wynika na mocy dowiedzonego twierdzenia pomocniczego, że istnieje macierz ortogonalna $\mathfrak{B}$, której pierwszą kolumną jest układ liczb (14). Utwórzmy iloczyn

$$\mathfrak{C} = \mathfrak{B} + \mathfrak{A}\mathfrak{B}.$$

Mamy tu $\mathfrak{A}\mathfrak{B} = (d_{jk})$, gdzie $d_{jk} = \sum_{p=1}^n a_{jp}b_{pk}$. Zatem $\mathfrak{C} = (c_{il})$, gdzie

$$c_{il} = \sum_{q=1}^n b_{qi}d_{ql} = \sum_{q=1}^n \left(b_{qi} \sum_{p=1}^n a_{qp}b_{pl} \right) = \sum_{p=1}^n b_{pl} \left(\sum_{q=1}^n b_{qi}a_{qp} \right).$$

W szczególności

$$c_{11} = \sum_{p=1}^n \left(b_{p1} \sum_{q=1}^n b_{q1}a_{qp} \right).$$

Wobec (16) jest

$$c_{11} = \lambda_1 \sum_{p=1}^n b_{p1}b_{p1}.$$

Ponieważ macierz $\mathfrak{B}$ jest ortogonalna, otrzymujemy stąd

$$c_{11} = \lambda_1, \quad c_{12} = \dots = c_{1n} = 0,$$

a wobec tego, że macierz $\mathbb{C}$ jest symetryczna, jest również

$$c_{21} = \dots = c_{2n} = 0.$$

Tym samym

$$\mathbb{C} = \left(\begin{array}{c|ccc} \lambda_1 & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & \mathbb{A}_1 & \\ 0 & & & \end{array} \right),$$

gdzie $\mathbb{A}_1$ jest macierzą symetryczną stopnia $n-1$. Wynika stąd, że przekształcenie ortogonalne o macierzy $\mathbb{B}$ sprowadza formę kwadratową $f(x|x)$ do postaci

$$f_1(u|u) = \lambda_1 u_1^2 + g(u_2, u_3, \dots, u_n).$$

Stąd zaś wynika przez indukcję, tak jak w dowodzie twierdzenia Lagrange'a (p. § 4, str. 520), że istnieje przekształcenie ortogonalne, które sprowadza formę $f(x|x)$ do postaci (13).

Należy jeszcze dowieść, że współczynniki λ_k w (13) są pierwiastkami charakterystycznymi formy $f(x|x)$. Otóż są one pierwiastkami charakterystycznymi formy przekształconej, gdyż dla niej

$$w(\lambda) = \begin{vmatrix} \lambda_1 - \lambda & 0 & \dots & 0 \\ 0 & \lambda_2 - \lambda & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \lambda_n - \lambda \end{vmatrix} = (\lambda_1 - \lambda)(\lambda_2 - \lambda) \dots (\lambda_n - \lambda).$$

Z niezmienniczości jednak tych pierwiastków względem przekształceń ortogonalnych (p. str. 528) wynika, że są one zarazem pierwiastkami charakterystycznymi formy pierwotnej, c. b. d. o.

Wśród pierwiastków charakterystycznych formy kwadratowej ilość tych, które są różne od zera, jest równa rzędowi formy, gdyż macierz formy (13) ma na głównej przekątnej liczby λ_k , a poza nią same zera.

Nie wszystkie pierwiastki charakterystyczne muszą być różne. Zbierając pierwiastki równe, możemy wielomian charakterystyczny napisać w postaci

$$(17) \quad w(\lambda) = (\lambda - \lambda_1)^s (\lambda - \lambda_2)^u \dots (\lambda - \lambda_k)^w.$$

Liczba s nazywa się *krotnością pierwiastka* λ_1 .

Dowodzimy, że *gdy λ_1 ma krotność s , to rząd macierzy $\mathbb{A}(\lambda_1)$ jest równy liczbie $n-s$ (twierdzenie Weierstrassa).*

Ponieważ $R[\mathbb{A}(\lambda_1)]$ jest niezmiennikiem przekształceń ortogonalnych (p. str. 527), więc wystarczy udowodnić to twierdzenie dla form sprowadzonych do postaci (13). Forma taka ma wówczas kształt

$$\lambda_1 u_1^2 + \lambda_1 u_2^2 + \dots + \lambda_1 u_s^2 + \lambda_{s+1} u_{s+1}^2 + \dots + \lambda_n u_n^2.$$

Zatem

$$\mathbb{A}(\lambda) = \begin{pmatrix} \lambda_1 - \lambda & & & \\ & \ddots & & \\ & & \lambda_1 - \lambda & \\ & & & \lambda_{s+1} - \lambda & \\ & & & & \ddots & \\ & & & & & \lambda_n - \lambda \end{pmatrix}.$$

Wszystkie wyrazy poza przekątną główną są tu zerami. Wobec tego

$$\mathbb{A}(\lambda_1) = \begin{pmatrix} 0 & & & \\ & \ddots & & \\ & & 0 & \\ & & & \lambda_{s+1} - \lambda_1 & \\ & & & & \ddots & \\ & & & & & \lambda_n - \lambda_1 \end{pmatrix}.$$

Ponieważ $\lambda_1 \neq \lambda_k$ dla $k \geq s+1$, więc rząd macierzy $\mathbb{A}(\lambda_1)$ jest równy $n-s$, c. b. d. o.

Zbadamy jeszcze zależność między znakami pierwiastków charakterystycznych a współczynnikami wielomianu charakterystycznego dla macierzy 3-go stopnia. Wówczas

$$(18) \quad w(\lambda) = -\lambda^3 + A_1 \lambda^2 - A_2 \lambda + A_3.$$

Ponieważ

$$w(\lambda) = (\lambda_1 - \lambda)(\lambda_2 - \lambda)(\lambda_3 - \lambda) = -\lambda^3 + (\lambda_1 + \lambda_2 + \lambda_3)\lambda^2 - (\lambda_1\lambda_2 + \lambda_1\lambda_3 + \lambda_2\lambda_3)\lambda + \lambda_1\lambda_2\lambda_3,$$

więc

$$(19) \quad A_1 = \lambda_1 + \lambda_2 + \lambda_3, \quad A_2 = \lambda_1\lambda_2 + \lambda_1\lambda_3 + \lambda_2\lambda_3, \quad A_3 = \lambda_1\lambda_2\lambda_3.$$

Rozróżnimy 6 przypadków.

1° Wszystkie pierwiastki są dodatnie:

$$(20) \quad \lambda_1 > 0, \quad \lambda_2 > 0, \quad \lambda_3 > 0.$$

Wynika stąd na mocy (19), że

$$(21) \quad A_1 > 0, \quad A_2 > 0, \quad A_3 > 0.$$

Na odwrót, z (21) wynika (20), gdyż dla $\lambda \leq 0$ jest wobec (21)

$$-\lambda^3 \geq 0, \quad A_1 \lambda^2 \geq 0, \quad -A_2 \lambda \geq 0, \quad A_3 > 0,$$

a zatem $w(\lambda) > 0$; wynika stąd, że liczba $\lambda \leq 0$ nie może być pierwiastkiem charakterystycznym, a więc zachodzi (20).

2° Wszystkie pierwiastki są ujemne:

$$(22) \quad \lambda_1 < 0, \quad \lambda_2 < 0, \quad \lambda_3 < 0.$$

Jak w poprzednim przypadku, okazuje się, że na to, aby zachodziły nierówności (22), potrzeba i wystarcza, żeby zachodziły nierówności

$$A_1 < 0, \quad A_2 > 0, \quad A_3 < 0.$$

Tak więc na to, by wszystkie trzy pierwiastki były tego samego znaku, potrzeba i wystarcza, żeby

$$(23) \quad A_2 > 0 \quad \text{i} \quad A_1 A_3 > 0.$$

3° Wszystkie pierwiastki są różne od zera, lecz nie mają zgodnych znaków. Aby tak było, potrzeba i wystarcza, żeby spełnione było zaprzeczenie warunku (23), czyli żeby

$$(24) \quad A_2 \leq 0 \quad \text{lub} \quad A_1 A_3 \leq 0.$$

4° Jeden z pierwiastków jest zerem, a dwa — różne od zera. Zachodzi to wtedy i tylko wtedy, gdy

$$A_2 \neq 0 \quad \text{i} \quad A_3 = 0.$$

Gdy $A_2 > 0$, oba pierwiastki różne od zera są tego samego znaku, a gdy $A_2 < 0$, mają one znaki przeciwne.

5° Dwa pierwiastki są równe zeru, a jeden różny od zera. Zachodzi to wtedy i tylko wtedy, gdy

$$A_1 \neq 0 \quad \text{i} \quad A_2 = A_3 = 0.$$

6° Wszystkie trzy pierwiastki są równe 0. Zachodzi to wtedy i tylko wtedy, gdy rząd macierzy $\mathfrak{A}$ jest zerem, a więc gdy $a_{ik} = 0$ dla wszystkich i, k .

PRZYPIS II

ILOCZYN WEKTOROWY

1. Określenie. W § 3 określiliśmy iloczyn skalarowy dwóch wektorów jako pewną liczbę, przyporządkowaną parze wektorów (p. str. 16). Obecnie określimy inne działanie, zwane *mnożeniem wektorowym*, które każdej parze uporządkowanej wektorów przyporządkowuje nie liczbę, lecz wektor.

W tym celu wyróżnimy i ustalmy jedną z dwóch możliwych orientacji trójek wektorów w przestrzeni i nazwijmy ją orientacją dodatnią (p. § 11, str. 56).

Niech wektory a i b tworzą kąt φ . Jeśli jeden z nich jest wektorem zerowym o lub jeśli oba są równoległe, to ich *iloczynem wektorowym* nazywamy wektor o . W przeciwnym razie ich iloczynem wektorowym nazywamy wektor c , spełniający następujące trzy warunki:

- 1° wektor c jest prostopadły do obu wektorów a i b ;
- 2° trójka uporządkowana a, b, c ma orientację dodatnią;
- 3° zachodzi równość

$$(1) \quad |c| = |a| |b| \sin \varphi.$$

Tak określony iloczyn wektorowy oznacza się przez $a \times b$.

2. Najprostsze własności. Z warunku 3° wynika, że długość iloczynu wektorowego $a \times b$ jest równa polu równoległoboku zbudowanego na wektorach a i b .

Z warunku 2° wynika, że iloczyn wektorowy $a \times b$ zależy od porządku czynników; orientacja bowiem trójki wektorów zmienia się, gdy zmieniamy porządek dwu z nich.

Dokładniej: jeśli zamienimy ze sobą wektory a i b , to żeby zachować orientację, musimy zmienić tylko zwrot wektora c (gdyż