Wyznaczniki, wartości i wektory własne, formy kwadratowe

Umiemy więc na razie liczyć pochodne funkcji (i odwzorowań) wielu zmiennych. Będziemy teraz chcieli wykorzystać tę umiejętność do badania funkcji i odwzorowań. Próbka problemów, które będziemy uczyć się rozwiązywać, to:

1. Ekstrema funkcji wielu zmiennych. Musimy w tym celu umieć badać pierwsze i drugie pochodne funkcji.

2. W wielu równaniach różniczkowych w fizyce mamy do czynienia z zamianą zmiennych – np. gdy układ ma symetrię sferyczną, to na ogół jego badanie jest prostsze w układzie współrzędnych sferycznych. Musimy więc nauczyć się testować, czy dana zamiana zmiennych jest 'dobra' (sprecyzujemy to dalej) oraz jeśli jest, to zamieniac zmienne w równaniach różniczkowych.

Niezbędnym wstępem do analizy tych problemów jest część algebraiczna, do które należą:

1. Wyznaczniki, (przy okazji trochę o permutacjach),

2. Wartości i wektory własne,

3. Formy biliniowe i kwadratowe.

Przedstawimy teraz kilka niezbędnych faktów ich dotyczących.

1 Odwzorowania wieloliniowe. Formy wieloliniowe. Wyznaczniki

1.1 1-formy

Zdefiniowaliśmy odwzorowania liniowe $V \to W$, gdzie V, W były przestrzeniami wektorowymi. Teraz rozpatrzmy szczególny przypadek, gdy $W = \mathbb{R}$, tzn. gdy mamy odwzorowanie liniowe $V \to \mathbb{R}$. Takie odwzorowanie nazywamy $1-$*formą.*

Niech $n = \dim V$.

Przykł. $V = \mathbb{R}_n[\cdot]$, $v \in V$. Niech $x_0 \in \mathbb{R}$ – ustalona liczba. Odwzorowanie $\alpha : V \to \mathbb{R}$ określamy następująco: $\alpha(v) = v(x_0)$. Łatwo sprawdza się liniowość α.

Aby w pełni scharakteryzować 1−formę, tzn. umieć podać jej wartość na dowolnym wektorze, trzeba podać jej wartości na n liniowo niezależnych wektorach. $e_1, e_2, \ldots, e_n$. (Gdy mamy mniej wektorów, lub nie są one l.n.z, to nie uda nam się podać jej wartości na *dowolnym* wektorze z V, a jedynie na wektorach należących do powłoki liniowej $\langle e_1, e_2, \ldots, e_n \rangle$. Podanie zaś wartości formy na większej ilości wektorów, niż baza, nie wnosi nowej informacji).

1.2 Formy biliniowe i 2-formy

Def. *Formą biliniową* nazywamy odwzorowanie $b : V \times V \to \mathbb{R}$, liniowe w każdym argumencie, tzn. spełniające dla dowolnych wektorów $u, v, w \in V$, $\alpha, \beta \in \mathbb{R}$:

$$b(\alpha u + \beta v, w) = \alpha b(u, w) + \beta(v, w), \qquad b(w, \alpha u + \beta v) = \alpha b(w, u) + \beta(w, v).$$

Przykł. Iloczyn skalarny jest formą biliniową.

Def. 2−*formą* na przestrzeni V nazywamy formę biliniową ω, która jest *antysymetryczna*, tzn. dla dowolnych $v, w \in V$ spełnia:

$$\omega(v, w) = -\omega(w, v).$$

Przykł. Zorientowane pole powierzchni równoległoboku rozpiętego na dwóch wektorach v, w w przestrzeni $V = \mathbb{R}^3$ jest 2−formą.

Ilu warunków potrzeba, aby w pełni scharakteryzować formę biliniową, bądź 2−formę? Tzn. umieć podać jej wartość na dowolnej parze wektorów. Argumentując analogicznie jak dla 1−form, widzimy, że:

• **Ogólna forma biliniowa**: Musimy podać wartości 2−formy na wszystkich możliwych parach wektorów liniowo niezależnych, co daje nam n^2 warunków.

• **2−forma**: Założenie, że forma jest antysymetryczna, redukuje ilość warunków. I tak, dla dowolnej 2−formy ω i dla dowolnych wektorów e_i, e_j, mamy: $\omega(e_i, e_j) = -\omega(e_j, e_i)$, co redukuje ilość warunków o $n(n-1)/2$. Ponadto, kładąc $i = j$, otrzymujemy $\omega(e_i, e_i) = 0$, przez co wypada kolejnych n warunków. Ostatecznie, *dowolna 2−forma scharakteryzowana jest przez* $n(n-1)/2$ *warunków.*

1.3 Formy $k-$liniowe i k-formy

1.4 $n-$formy na przestrzeniach $n-$wymiarowych i wyznaczniki

1.4.1 $n-$formy na przestrzeniach $n-$wymiarowych

Przejdźmy teraz do form $n-$liniowych na przestrzeni wymiaru n i $n-$form. Mamy, analogicznie jak poprzednio:

Def. *Formą $n-$liniową* nazywamy odwzorowanie $l : V \times V \times \cdots \times V \to \mathbb{R}$, liniowe w każdym argumencie, tzn. spełniające dla dowolnych wektorów $v_1, v_2, \ldots, v_n, w \in V$, $\alpha, \beta \in \mathbb{R}$:

$$l(v_1, v_2, \ldots, \alpha v_k + \beta w, v_{k+1}, \ldots, v_n) =$$

$$= \alpha l(v_1, v_2, \ldots, v_k, v_{k+1}, \ldots, v_n) + \beta l(v_1, v_2, \ldots, w, v_{k+1}, \ldots, v_n).$$

Def. *$n-$formą* na przestrzeni V nazywamy formę $n-$liniową ω, która jest *antysymetryczna w dowolnej parze argumentów*, tzn. spełnia

$$\omega(v_1, \ldots, v_i, \ldots, v_j, \ldots, v_n) = -\omega(v_1, \ldots, v_j, \ldots, v_i, \ldots, v_n).$$

Zapytajmy znów, ilu warunków trzeba, aby w pełni scharakteryzować formę $n-$liniową $G_n(v_1, v_2, \ldots, v_n)$? Oraz $n-$formę?

Dla ogólne $n-$formy odpowiedź brzmi: n^n warunków. (strasznie dużo!). Ale dla dowolnej $n-$formy, mamy $\begin{pmatrix} n \\ n \end{pmatrix}$ – tylko *jeden* warunek!

Przekonajmy się o tym bardziej namacalnie:

Weźmy wartość $n-$formy ω^n na wektorach bazy: $\omega^n(e_{i_1}, e_{i_2}, \ldots, e_{i_n})$. Jeśli *którekolwiek* dwa wektory są równe, to wartość formy jest 0 – z antysymetrii. Zatem wszystkie wektory $e_{i_1}, e_{i_2}, \ldots, e_{i_n}$ muszą być *różne*.

Ile warunków zostało? Mamy (niby) $n!$ (tyle co możliwości przestawień tzn. permutacji zbioru $n-$ elementowego) zamiast n^n – już znaczny postęp w ograniczaniu liczebności. Ale zauważmy, że wartości na poprzestawianych wektorach nie są niezależne! Znając wartość $n-$ formy na *dowolnym* wyborze wektorów bazy, możemy obliczyć jej wartość na *dowolnej innej* permutacji.

Tak naprawdę otrzymujemy więc *jeden* warunek charakteryzujący $n-$formę, tzn. jej wartość na jakiejś dowolnej permutacji wektorów bazy.

Zatem *przestrzeń $n-$form jest jednowymiarowa.*

1.4.2 Forma objętości

Wprowadźmy teraz następującą, pożyteczną $n-$formę, zwaną *formą objętości* vol. Definiujemy ją następująco:

Niech $e_1, e_2, \ldots, e_n$ – baza standardowa w V (tzn. wektor e_k ma $k-$tą składową równą 1, reszta składowych jest 0).

Def. Formę objętości vol definiujemy przez warunek

$$\mathrm{vol}(e_1, e_2, \ldots, e_n) = 1.$$

Stąd już blisko do definicji wyznacznika (macierzy $n \times n$):

Def. Niech $A \in \mathbb{R}^n{}_n$. Wyznacznik $\det A$ macierzy A definiujemy wzorem

$$\det A = \mathrm{vol}(Ae_1, Ae_2, \ldots, Ae_n).$$

Pamiętając o tym, że $Ae_1 = a_1$, $Ae_2 = a_2$, $\ldots$, $Ae_n = a_n$, gdzie a_k – $k-$ty wektor-kolumna macierzy A, możemy zapisać

$$\det A = \mathrm{vol}(a_1, a_2, \ldots, a_n).$$

Definicja ta stanowi jednocześnie przepis, w jaki sposób możemy policzyć wyznacznik dowolnej macierzy $n \times n$. Jest on dość mało praktyczny; później rozwiniemy znacznie szybsze metody liczenia wyznacznika. zobaczmy jak on działa w znanym nam już przypadku macierzy 2×2

Tu liczenie tą metodą wyznacznika macierzy 2 x 2.

A teraz wypiszemy kilka własności wyznacznika, posługując się podaną wyżej definicją.

1.4.3 Kilka własności wyznacznika

Za chwilę dojdziemy do bardziej zwartej formuły na liczenie wyznacznika; ale na razie z powyższej definicji wyciągnijmy kilka pożytecznych własności wyznacznika.

Niech B będzie macierzą jakiegoś odwzorowania liniowego $V \to V$. Niech $a_1, a_2, \ldots, a_n \in V$ – wektory kolumnowe macierzy A. Rozpatrzmy teraz odwzorowanie vol:

$$\mathrm{vol} : \underbrace{V \times V \times \cdots \times V}_{n \text{ razy}} \to \mathbb{R}$$

określone następująco:

$$\underbrace{V \times V \times \cdots \times V}_{n \text{ razy}} \ni (a_1, a_2, \ldots, a_n) \to \mathrm{vol}(Ba_1, Ba_2, \ldots, Ba_n).$$

Widać, że jest to odwzorowanie antysymetryczne w każdej parze argumentów (bo forma vol ma tę własność), jest zatem $n-$formą, czyli *musi być proporcjonalne do formy objętości*:

$$\mathrm{vol}(Ba_1, Ba_2, \ldots, Ba_n) = f(B)\mathrm{vol}(a_1, a_2, \ldots, a_n);$$

$f(B)$ jest czynnikiem proporcjonalności, który zależy tylko od macierzy B, nie zależy zaś od wektorów $a_1, \ldots, a_n$.

Weźmy teraz $a_1 = e_1, \ldots, a_n = e_n$. Mamy wtedy:

$$\mathrm{vol}(Be_1, Be_2, \ldots, Be_n) = f(B)\mathrm{vol}(e_1, e_2, \ldots, e_n) = f(B),$$

co znaczy, że $f(B) = \det B$. Możemy teraz łatwo policzyć $\det(BA)$:

$$\det(BA) = \mathrm{vol}(BAe_1, BAe_2, \ldots, BAe_n) = \mathrm{vol}(Ba_1, Ba_2, \ldots, Ba_n)$$

$$= \det(B) \cdot \mathrm{vol}(a_1, a_2, \ldots, a_n) = \det(B)\det(A).$$

Udowodniony właśnie fakt nazywa się *tw. Cauchy'ego* o wyznaczniku iloczynu macierzy:

Tw. (Cauchy'ego):

$$\det(BA) = \det(B)\det(A).$$

1.4.4 Uzupełnienie: Związek między wymiarami jądra i obrazu

Zanim przejdziemy do sformułowania kolejnej ważnej własności wyznacznika, związanej z (nie)odwracalnością macierzy, musimy podać

Uzupełnienie.

Tw. Niech V, W – przestrzenie wektorowe. Dla dowolnego odwzorowania liniowego $F : V \to W$ zachodzi

$$\dim \mathrm{Ker}\, F + \dim \mathrm{Im}\, F = \dim V. \tag{1}$$

Dow. Załóżmy, że wymiar przestrzeni V jest równy n: $\dim V = n$. Załóżmy, że znaleźliśmy jądro F. Pamiętamy, że jest to podprzestrzeń V. Oznaczmy jej wymiar przez k: $k = \dim \mathrm{Ker}\, F$. Wiemy, że wymiar podprzestrzeni nie może być większy niż wymiar przestrzeni: $k \leqslant n$. Załóżmy na chwilę, że $k > 0$; przypadek $k = 0$ rozpatrzymy oddzielnie za chwilę.

Znajdźmy jakąś bazę $(e_1, \ldots, e_k)$ podprzestrzeni $\mathrm{Ker}\, F$. Następnie uzupełnijmy tę bazę o $n - k$ liniowo niezależnych wektorów $e_{k+1}, \ldots, e_n$ tak, by cały zbiór $(e_1, \ldots, e_k, e_{k+1}, \ldots, e_n)$ stanowił bazę przestrzeni V.

Zauważmy teraz, że następujący układ wektorów z przestrzeni W: $F(e_{k+1}), F(e_{k+2})$, jest układem *liniowo niezależnym*. Przypuśćmy bowiem, że tak nie jest, tzn. że

$F(e_{k+1}), F(e_{k+2}), \ldots, F(e_n)$ *jest* układem liniowo zależnym. Istnieją wtedy liczby

$\lambda_{k+1}, \lambda_{k+2}, \ldots, \lambda_n$, z których przynajmniej jedna jest różna od zera, i które spełniają

$$\lambda_{k+1}F(e_{k+1}) + \lambda_{k+2}F(e_{k+2}) + \cdots + \lambda_n F(e_n) = \mathbf{0}.$$

Ale z liniowości F wynika, że

$$\mathbf{0} = \lambda_{k+1}F(e_{k+1}) + \lambda_{k+2}F(e_{k+2}) + \cdots + \lambda_n F(e_n) =$$

$$F(\lambda_{k+1}e_{k+1} + \lambda_{k+2}e_{k+2} + \cdots + \lambda_n e_n). \qquad (2)$$

Ale wektor $x \equiv \lambda_{k+1}e_{k+1} + \lambda_{k+2}e_{k+2} + \cdots + \lambda_n e_n$ jest *niezerowy*, bo wektory $e_{k+1}, \ldots, e_n$ są z założenia liniowo niezależne. Ponadto wektor $x = \lambda_{k+1}e_{k+1} + \lambda_{k+2}e_{k+2} + \cdots + \lambda_n e_n$ *nie należy do jądra*, ponieważ jądro jest rozpięte przez wektory $e_1, \ldots, e_k$. Skoro niezerowy wektor x nie należy do jądra, to musi zachodzić $F(x) \neq \mathbf{0}$. Zatem *nie może* zachodzić równość (2). Otrzymaliśmy sprzeczność, zatem układ wektorów $F(e_{k+1}), F(e_{k+2}), \ldots, F(e_n)$ jest układem liniowo niezależnym.

Tak więc wiemy, że układ wektorów $F(e_{k+1}), F(e_{k+2}), \ldots, F(e_n)$ rozpina całą podprzestrzeń $\operatorname{Im} F$ (bo wektory $F(e_1), \ldots, F(e_k)$ są zerowe), zatem stanowią bazę $\operatorname{Im} F$. Policzmy teraz wymiary (licząc ilości elementów bazy):

$$\dim V = n, \quad \dim \operatorname{Ker} F = k, \quad \dim \operatorname{Im} F = n - k,$$

zatem zachodzi wzów (1).

Pozostaje jeszcze przyjrzeć się przypadkowi $k = 0$. Ale wtedy jądro jest podprzestrzenią trywialną (tzn. składa się tylko z wektora zerowego) i argumentując jak wyżej widzimy, że *cała* baza V składa się z wektorów, których obrazy $F(e_1), \ldots, F(e_n)$ tworzą układ liniowo niezależny. Wtedy $\dim \operatorname{Im} F = \dim V$ i wzór (1) też zachodzi.

CBDO

(Koniec uzupełnienia)

1.4.5 Niezerowość wyznacznika a odwracalność macierzy

Teraz wypiszmy dwa pożyteczne wnioski z tw. Cauchy'ego.

Wniosek 1. Jeśli macierz A jest odwracalna, to $\det A \neq 0$ i zachodzi

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

Bo: Z tw. Cauchy'ego mamy: $1 = \det(I_n) = \det(AA^{-1}) = \det A \det(A^{-1})$.

CBDO

Wniosek 2. Jeśli $\det A \neq 0$, to A ma liniowo niezależne kolumny.

Dow. (*ad absurdum*) Załóżmy, że zachodzi teza przeciwna, tzn. A ma liniowo zależne kolumny. Wtedy któraś z nich, np. a_i, jest kombinacją liniową pozostałych, tzn. zachodzi równość

$$a_i = \mu^1 a_1 + \cdots + \mu^{i-1} a_{i-1} + \mu^{i+1} a_{i+1} + \ldots \mu^n a_n.$$

Wtedy jednak

$$\begin{gathered}
\det A = \mathrm{vol}\,(a_1, \ldots, a_i, \ldots, a_n) \\
= \mathrm{vol}\,(a_1, \ldots, \mu^1 a_1 + \cdots + \mu^{i-1} a_{i-1} + \mu^{i+1} a_{i+1} + \ldots \mu^n a_n, \ldots, a_n) \\
= \mu^1 \mathrm{vol}\,(a_1, \ldots, a_1, \ldots, a_n) + \mu^2 \mathrm{vol}\,(a_1, a_2, \ldots, a_2, \ldots, a_n) + \ldots \\
+ \mu^{i-1} \mathrm{vol}\,(a_1, \ldots, a_{i-1}, a_{i-1}, \ldots, a_n) + \mu^{i+1} \mathrm{vol}\,(a_1, \ldots, a_{i+1}, a_{i+1}, \ldots, a_n) \\
+ \cdots + \mu^n \mathrm{vol}\,(a_1, \ldots, a_n, \ldots, a_n).
\end{gathered}$$

Gdy dwa argumenty w $n-$formie są takie same, to wartość tej $n-$formy jest równa zeru (z antysymetrii). Tak więc wartość wyznacznika jest równa zeru. Sprzeczność.

CBDO

Zauważmy teraz, że macierz $n \times n$ mająca liniowo niezależne kolumny ma własność: $\dim \mathrm{Im}\, A = n$. Traktując ją jako macierz jakiegoś odwzorowania liniowego $V \to V$, gdzie $\dim V = n$, widzimy, że obrazem A jest cała przestrzeń V. Macierz A jest więc odwzorowaniem *surjektywnym* ("na").

Zauważmy ponadto, że jest *injekcją*, tzn. jeśli $Ax_1 = b$ i $Ax_2 = b$, to $x_1 = x_2$. Bowiem – z liniowości A – mamy: $Ax_1 - Ax_2 = A(x_1 - x_2) = b - b = 0$, tzn. $A(x_1 - x_2) = \mathbf{0}$. Ale z udowodnionego dopiero co wzoru (1) mamy:

$$\dim \mathrm{Im}\, A + \dim \mathrm{Ker}\, A = n + \dim \mathrm{Ker}\, A = n,$$

tzn. $\dim \mathrm{Ker}\, A = 0$, co oznacza, że jedynym rozwiązaniem równania $A(x_1 - x_2) = \mathbf{0}$ jest $x_1 - x_2 = \mathbf{0}$, tzn. $x_1 = x_2$. A zatem A jest injekcją.

Skoro A jest injekcją i surjekcją, zatem jest bijekcją. Jest więc odwracalne.

Powyższe fakty można podsumować, wypowiadając

Tw. Macierz A jest odwracalna wtedy i tylko wtedy, gdy $\det A \neq 0$.

CBDO

Wiemy już conieco o wyznaczniku; jednak warto mieć bardziej poręczne wzory do liczenia go, bo liczenie wprost z definicji zbyt wygodne nie jest. Będzie nam do tego potrzebne pojęcie *permutacji*.

1.4.6 Permutacje. Parzystość permutacji

• Pojęcie permutacji

Niech X – zbiór skończony, zawierający n elementów. Bez zmniejszania ogólności możemy wziąć $X = \{1, 2, \ldots, n\}$.

Def. *Permutacją* zbioru $n-$ elementowego X nazywamy bijekcję $X \to X$.

Przykłady odwzorowań które są i nie są permutacjami

Zapis permutacji. Niech π – permutacja, którą jawnie zapiszmy jako: $1 \to \pi(1) \in X$, $2 \to \pi(2) \in X$, $\ldots$, $n \to \pi(n) \in X$. Permutację π zapisujemy w postaci:

$$\pi = \begin{pmatrix} 1 & 2 & \ldots & n \\ \pi(1) & \pi(2) & \ldots & \pi(n) \end{pmatrix} \tag{3}$$

Przykł. 1: Wypiszmy wszystkie permutacje zbioru 2−elementowego:

$$\begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}, \quad \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$$

Widać, że mamy tu dwie permutacje.

Przykł. 2: oraz wszystkie permutacje zbioru 3−elementowego:

$$\begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix},$$

Tu mamy 6 permutacji.

Nietrudno zobaczyć, że ilość permutacji zbioru $n-$elementowego to $n!$. (Można to sformalizować w dowód indukcyjny).

Zbiór wszystkich permutacji zbioru $n-$elementowego oznaczamy S_n.

Permutacja identycznościowa to taka, która przeprowadza każdy element w siebie, tzn.

$$\mathrm{Id} = \begin{pmatrix} 1 & 2 & \ldots & n \\ 1 & 2 & \ldots & n \end{pmatrix}$$

Złożenie dwóch permutacji też jest permutacją.

Przykł. Składanie permutacji i ich nieprzemienność in general.

$$\begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} \circ \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix},$$

$$\begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} \circ \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}.$$

Dla każdej permutacji π istnieje permutacja *odwrotna* π^{-1}, tzn. taka, że

$$\pi \circ \pi^{-1} = \pi^{-1} \circ \pi = \mathrm{Id}$$

• **Dygresja – pojęcie grupy**

Def. *Grupą* $(G, \circ)$ nazywamy zbiór G z działaniem $\circ$:

$$G \times G \ni (g_1, g_2) \to g_1 \circ g_2 \in G,$$

spełniającym warunki:

1. Działanie w grupie jest *łączne*, tzn. $\forall\ g_1, g_2, g_3 \in G : (g_1 \circ g_2) \circ g_3 = g_1 \circ (g_2 \circ g_3)$;
2. Istnienie elementu jednostkowego e: $\forall\ g \in G : e \circ g = g \circ e = g$;
3. Istnienie elementu odwrotnego: $\forall\ g \in G :\ \exists h \in G :\ g \circ h = h \circ g = e$.

Przykł.

1. S_2, S_3 i ogólnie S_n z działaniem, którym jest składanie permutacji, są grupami.
2. $(\mathbb{R}, +)$ jest grupą. Elementem jednostkowym jest zero.
3. $(\mathbb{R} \setminus \{0\}, \cdot)$ jest grupą.

Uwagi nt. grupy a symetria.

• **Transpozycja, znak permutacji – preludium**

Def. *Transpozycją* (i, j) nazywamy permutację, która zamienia ze sobą elementy i oraz j, a resztę zostawia na swoim miejscu:

$$(i,j) = \begin{pmatrix} 1 & 2 & \dots & i & \dots & j & \dots & n \\ 1 & 2 & \dots & j & \dots & i & \dots & n \end{pmatrix}$$

Transpozycje można uważać za 'cegiełki', z których można zbudować każdą permutację. Mówi o tym następujące

Stw. Każdą permutację można przedstawić (na wiele sposobów) jako złożenie pewnej ilości transpozycji.

Dow. Pokażemy, jak przedstawić dowolną permutację jako złożenie transpozycji, zamieniających wyłącznie *sąsiednie* elementy.

Zapiszmy ogólną permutację jako

$$\pi = \begin{pmatrix} 1 & 2 & \dots & n \\ a & b & \dots & k \end{pmatrix} \tag{4}$$

Roboczo, nazwijmy *tablicą* zespół liczb stanowiących dolny wiersz permutacji (4).

Sposób postępowania jest następujący: Dostatecznie wiele razy stosujemy następującą operację:

Niech i, j oznacza jakąś parę kolejnych elementów.

- Jeśli $i < j$ to nic nie robimy,
- Jeśli $i > j$ to zamieniamy i oraz j.

Nazwijmy tę operację PSE (Porównywanie Sąsiednich Elementów).

Zaczynamy o wyjściowej tablicy $K_0 = (a, b, c, \ldots, k)$.

Pierwszy krok postępowania:

Stosujemy operację PSE do pierwszej pary elementów K_0.

Stosujemy operację PSE do drugiej pary elementów otrzymanej wyżej tablicy.

Itd., łącznie $(n-1)$ razy. Nazwijmy otrzymaną tablicę K_1.

Co otrzymamy jako ostatni element tablicy K_1? Łatwo zauważyć, że jest to liczba n, jako że jest większa od dowolnej innej, i jeśli natrafimy na nią przy stosowaniu operacji PSE, to zawsze jest przesuwana na prawo.

Drugi krok postępowania:

Teraz, zastosujmy opisany wyżej sposób postępowania (tzn. cykl operacji stanowiących 1. krok) do pierwszych $n-1$ elementów K_1 (nazwijmy je *podtablicą* tablicy K_1). Po $n-2$ krokach, otrzymamy tablicę K_2, na końcu której są elementy $n-1, n$.

Itd.,

Po $n-1$ krokach otrzymamy uporządkowaną (w terminologii komputerowej: *posortowaną*) tablicę $K_{n-1} = (1, 2, \ldots, n)$.

CBDO

Uwagi o sortowaniu bąbelkowym

oraz że – względem efektywności – to jak z wołem co był ministrem w bajce Krasickiego 'Wół minister': Algorytm wolny ale skuteczny.

Przykł. Rozważmy permutację:

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 3 & 1 \end{pmatrix}$$

Kolejne kroki to:

$$\textbf{1. krok}: \quad |2,4,3,1| \to |2,4,3,1| \to |2,3,4,1| \to |2,3,1,4|$$

[transpozycja (2,3) oraz (3,4)]

2. krok : $|2,3,1,4| \to |2,3,1,4| \to |2,1,3,4|$ [transpozycja (2,3)]

3. krok : $|2,1,3,4| \to |1,2,3,4|$ [transpozycja (1,2)]

Możemy więc zapisać:

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 3 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 3 & 4 \end{pmatrix} \circ \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 3 & 2 & 4 \end{pmatrix} \circ \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 4 & 3 \end{pmatrix} \circ \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 3 & 2 & 4 \end{pmatrix}$$

A teraz: Policzmy ilość transpozycji które musieliśmy wykonać, aby uporządkować tablicę: Jest ich łącznie 4 (dwie w pierwszym kroku, po jednej w drugim i trzecim). (4 – liczba parzysta).

Oczywiście, porządkowanie tablicy (tzn. przechodzenie do uporządkowania naturalnego) mogliśmy wykonać znacznie ekonomiczniej, jeśli chodzi o liczbę transpozycji – elementów niekoniecznie sąsiednich:

$$|2,4,3,1| \to |2,1,3,4| \to |1,2,3,4|.$$

lub – zapisując bardziej rozwlekle, ale tak, aby trudniej się było pomylić –

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 3 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 3 & 4 \end{pmatrix} \circ \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 4 & 3 & 2 \end{pmatrix}$$

Tu były potrzebne dwie transpozycje, tzn. permutację σ zapisaliśmy jako złożenie *dwóch* transpozycji. (Znów liczba parzysta!)

Permutację σ zapisaliśmy , w obu przypadkach, jako złożenie *parzystej* ilości transpozycji. Nie jest to przypadek, ale przejaw ogólnego

Tw. Zapiszmy jakąś permutację $\pi \in S_n$ jako złożenie k transpozycji. Wówczas parzystość liczby k jest niezależna od sposobu rozkładu π na transpozycje, tzn. zależy tylko od permutacji k.

Twierdzenie to zaraz udowodnimy, ale najsampierw podamy dwa pojęcia, pojawiające się przy permutacjach, tzn. pojęcie *cyklu* oraz *orbity cyklu.*

• Cykle, orbity

Weźmy jakąś permutację $\pi \in S_n$. Weźmy jakąś liczbę j $1 \leqslant j \leqslant n$. Po jednokrotnym zadziałaniu permutacją π, otrzymamy liczbę $\pi(j)$. Po dwukrotnym zadziałaniu π, otrzymamy $\pi(\pi(j)) \equiv \pi^2(j)$, po kolejnym: $\pi^3(j)$ itd. Po którymś kolejnym razie (co najwyżej n) otrzymamy *tę samą liczbę* j, jako że co najwyżej n liczb $j, \pi(j), \pi^2(j), \ldots, \pi^n(j)$ może być różnych. Niech k będzie taką najmniejszą liczbą, że $\pi^k(j) = j$. Powyższą sytuację, działania kolejnych złożeń (iteracji) odwzorowania π, możemy zilustrować jako:

$$j \xrightarrow{\pi} \pi(j) \xrightarrow{\pi} \pi^2(j) \xrightarrow{\pi} \cdots \xrightarrow{\pi} \pi^{k-1}(j) \xrightarrow{\pi} j$$

Def. Powyższy zbiór: $j, \pi(j), \pi^2(j), \ldots, \pi^{k-1}(j)$ nazywamy *orbitą* elementu j pod działaniem permutacji π.

Def. *Cyklem* nazywamy permutację, która na elementach orbity $j, \pi(j), \pi^2(j), \ldots, \pi$ działa następująco:

$$j \to \pi(j); \quad \pi(j) \to \pi^2(j); \quad \ldots \pi^{k-1}(j) \to j$$

a na pozostałych liczbach jest tożsamościowa.

Stw. Dane dwa cykle albo mają te same orbity, albo te orbity są rozłączne.

Bo: Jeśli mają jeden element wspólny, to – z cykliczności – jego kolejne iteracje wypełniają zarówno jedną jak drugą orbitę; więc te orbity muszą się pokrywać.

Korzystając z tego widzimy, że

Stw. Każdą permutację można rozłożyć na cykle, tzn. przedstawić w postaci złożenia cykli o rozłącznych orbitach.

Równoważne zapisy cyklu: $(abcd \ldots jk) = (bcd \ldots jka) = (cd \ldots jkab) = (kabcd \ldots j)$; dopuszczalne są *cykliczne* przestawienia elementów – *ale inne nie!*

Jeszcze jedna malutka

Def. *Cyklem trywialnym* nazywamy cykl odpowiadający orbicie *jednoelementowej*; innymi słowy, cykl trywialny to permutacja *tożsamościowa.*

Uwaga. Zazwyczaj cykle trywialne pomijamy w zapisie rozkładu permutacji na cykle (ze względów ekonomicznych). Ale czasem dopisujemy cykle trywialne.

Przykł. rozkładu permutacji na cykle. Oraz tego jak się mnoży (składa) permutacje zapisane w języku cykli – istotne przy następnym punkcie.

Zauważmy, że rozkładając permutację tożsamościową na cykle, otrzymamy n cykli trywialnych. Jest to też jedyna permutacja spośród należących do S_n składająca się z n cykli.

• Znak permutacji – szkic dowodu jednoznaczności

Przypomnijmy co będziemy (szkicowo) pokazywać:

Tw. Zapiszmy jakąś permutację $\pi \in S_n$ jako złożenie k permutacji. Wówczas parzystość liczby k jest niezależna od sposobu rozkładu π na transpozycje, tzn. zależy tylko od permutacji k.

Dow. Niech permutacja $\pi \in S_n$, której parzystość/nieparzystość będziemy ustalać, składa się z k cykli. Jej (nie)parzystość zbadamy, przekształcając ją w permutację tożsamościową, (składającą się z n cykli) przez składanie jej z jakąś ilością transpozycji (tzn. rozłożymy π^{-1} na transpozycje i zbadamy ich ilość).

Poczyńmy następującą obserwację i zawrzyjmy ją w stwierdzeniu:

Stw. Złożenie dowolnej permutacji π z transpozycją zmniejsza lub zwiększa ilość orbit o 1.

Dow. (**Stw.**) Niech transpozycja będzie (i, l). Musimy rozważyć trzy przypadki:

1. Permutacja π zawiera liczby i oraz l jedynie w cyklach trywialnych:

   $$\pi = ...(i)(l)...$$

   Wtedy złożenie z transpozycją (ij) będzie

   $$\pi \circ (ij) = [(i)(j)] \circ [(ij)] = [(ij)];$$

   z cykli (i) oraz (j) zrobił się (ij) – **zmiana ilości cykli wynosi** -1.

2. Tak naprawdę, to powyższa sytuacja jest szczególnym przypadkiem sytuacji, gdy w π, liczby i oraz l są zawarte w cyklach rozłącznych: $\pi = (ab.\ldots.hi)(AB\ldots Hl)$

   Tu mamy:

   $$\pi \circ (il) = [(ab.\ldots.hi)(AB\ldots Hl)] \circ [(il)] = [(iAB\ldots Hlab\ldots h)];$$

   i znów, z dwóch cykli $(ab.\ldots.hi)$, $(AB\ldots Hl)$ zrobił się jeden $(iAB\ldots Hlab\ldots h)$ – **zmiana ilości cykli wynosi** -1.

3. Liczby i oraz l są zawarte w *jednym* cyklu: $\pi = (ab\ldots hij\ldots kl)$.

   Liczymy:

   $$\pi \circ (il) = (ab\ldots hij\ldots kl) \circ (il) = (iab\ldots h)(lj\ldots k);$$

   z jednego cyklu $(ab\ldots hij\ldots kl)$ zrobiły się dwa $(iab\ldots h)$, $(lj\ldots k)$ – **zmiana ilości cykli wynosi** $+1$.

CBDO (dow. stw.)

Załóżmy teraz, że startujemy z jakiejś permutacji π, posiadającej k cykli, i chcemy dojść do permutacji tożsamościowej, posiadającej n cykli, przez złożenie π z r transpozycjami. Niech r_+ spośród tych transpozycji *zwiększa* ilość cykli o 1, zaś r_- niech *zmniejsza* ilość cykli o 1. Sumaryczna zmiana ilości cykli to $n - k$, co musi być równe $r_+ - r_-$:

$$n - k = r_+ - r_-.$$

Weźmy teraz inny zestaw transpozycji, w ilości r', tak, że złożenie π z tymi r' transpozycjami znów daje permutację identycznościową. Jak poprzednio, niech r'_+ spośród tych transpozycji *zwiększa* ilość cykli o 1, zaś r'_- niech *zmniejsza* ilość cykli o 1. Mamy

$$n - k = r'_+ - r'_-.$$

Niech teraz np. $r'_+ = r_+ + 1$. Wtedy r'_- nie może być dowolna, bo mamy warunek $n - k = r'_+ - r'_-$, co daje $r'_- = r_- + 1$. Zatem $r' = r'_+ - r'_- = r_+ - r_- + 2 = r + 2$!

W ogólnym przypadku, niech $r'_+ = r_+ + m$. Wtedy mamy również $r'_- = r_- + m$, czyli $r' = r + 2m$.

Zatem! W każdym przypadku r oraz r' mają *tę samą parzystość* (tzn. są albo obie parzyste, albo obie nieparzyste).

CBDO

Przykł. Poznajdujmy parzystości wszystkich permutacji S_3:

$$\pi_0 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix} = (1)(2)(3), \;\; \mathrm{sgn}(\pi_0) = +;$$

$$\pi_1 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} = (23)(1), \;\; \mathrm{sgn}(\pi_1) = -;$$

$$\pi_2 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} = (13)(2), \;\; \mathrm{sgn}(\pi_2) = -;$$

$$\pi_3 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} = (12)(3), \;\; \mathrm{sgn}(\pi_3) = -;$$

$$\pi_+ = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} = (123) = [(13)(2)] \circ [(12)(3)], \;\; \mathrm{sgn}(\pi_+) = +;$$

$$\pi_- = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} = (132) = [(12)(3)] \circ [(13)(2)], \;\; \mathrm{sgn}(\pi_-) = +$$

Poręczniejszy wzór na znak permutacji – przez rozkład na cykle nietrywialne i sumowanie po ich długościach

a uprzednio, że każdy cykl długości k rozkłada się na $k - 1$ transpozycji.

1.4.7 Zapis wyznacznika przez permutacje

Teraz rozwiniętą tu wiedzę o permutacjach wykorzystamy do wypisania innej postaci wyznacznika.

Niech A będzie macierzą $n \times n$, którą zapiszemy na 2 sposoby, tak jak to już czyniliśmy:

$$A = \begin{bmatrix} a^1{}_1 & a^1{}_2 & \dots & a^1{}_n \\ a^2{}_1 & a^2{}_2 & \dots & a^2{}_n \\ \vdots & \vdots & & \vdots \\ a^n{}_1 & a^n{}_2 & \dots & a^n{}_n \end{bmatrix} = [a_1, a_2, \dots, a_n];$$

a_k $(k = 1, \dots, n)$ są to wektory kolumnowe, które wyrazimy w bazie standardowej:

$$a_i = a^1{}_i e_1 + a^2{}_i e_2 + \dots + a^n{}_i e_n.$$

Dla wyznacznika $\det A$ otrzymujemy:

$$\begin{aligned} \det A = \mathrm{vol}\,(a_1, a_2, \dots, a_n) &= \mathrm{vol}\,(a^1{}_1 e_1 + a^2{}_1 e_2 + \dots + a^n{}_1 e_n, a_2, a_3, \dots, a_n) \\ &= \sum_{i_1=1}^{n} a^{i_1}{}_1 \mathrm{vol}\,(e_{i_1}, a_2, a_3, \dots, a_n) \\ &= \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} a^{i_1}{}_1 a^{i_2}{}_2 \mathrm{vol}\,(e_{i_1}, e_{i_2}, a_3, \dots, a_n) = \dots \\ &= \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} \cdots \sum_{i_n=1}^{n} a^{i_1}{}_1 a^{i_2}{}_2 \dots a^{i_n}{}_n \mathrm{vol}\,(e_{i_1}, e_{i_2}, \dots, e_{i_n}). \end{aligned} \tag{5}$$

Przyjrzyjmy się ostatniemu czynnikowi, tzn. $\mathrm{vol}\,(e_{i_1}, e_{i_2}, \dots, e_{i_n})$. Wszystkich n argumentów w vol(...) są *wektorami bazy standardowej*, tak więc $\mathrm{vol}\,(e_{i_1}, e_{i_2}, \dots, e_{i_n})$ może przyjmować jedynie trzy wartości: 0 (gdy jakiekolwiek argumenty się powtarzają – z antysymetrii vol), bądź ± 1 – gdy wszystkie n argumentów jest różnych.

Przyjrzyjmy się zatem znów czynnikowi $\mathrm{vol}\,(e_{i_1}, e_{i_2}, \dots, e_{i_n})$ po raz drugi, mając tym razem w pamięci, że wszystkie argumenty są różne. Wszystkich możliwych sytuacji jest tyle, ile przestawień n liczb od 1 do n, czyli $n!$. Przez przestawienia argumentów, wszystkie te sytuacje można doprowadzić do jednej, tzn. $\mathrm{vol}\,(e_1, e_2, \dots, e_n)$ – *z uwzględnieniem odpowiedniego znaku.* A ile wynosi ten znak? Otóż zauważmy, że jest to po prostu *znak permutacji* prowadzącej od $\{i_1, i_2, \dots, i_n\}$ do $\{1, 2, \dots, n\}$, czyli

$$\mathrm{sgn} \begin{pmatrix} 1 & 2 & \dots & n \\ i_1 & i_2 & \dots & i_n \end{pmatrix} \tag{6}$$

Tak więc w sumie (5) sumujemy po *wszystkich* permutacjach. Oznaczmy permutację występującą w (8) jako π:

$$\pi \equiv \begin{pmatrix} 1 & 2 & \dots & n \\ \pi(1) & \pi(2) & \dots & \pi(n) \end{pmatrix} = \begin{pmatrix} 1 & 2 & \dots & n \\ i_1 & i_2 & \dots & i_n \end{pmatrix}$$

Uwzględniając to, przepiszmy zatem na (5) w następującej postaci, dającej równoważne wyrażenie na wyznacznik macierzy A:

$$\det A = \sum_{\pi \in S_n} \operatorname{sgn}(\pi) a^{\pi(1)}{}_1 a^{\pi(2)}{}_2 \dots a^{\pi(n)}{}_n \tag{7}$$

Przykł. Popatrzmy, że powyższe wyrażenie daje znane nam już wzory na wyznaczniki macierzy 2×2 i 3×3.

• • Dla macierzy 2×2, mamy dwie permutacje, dla których wypiszmy odpowiadające im znaki:

$$\operatorname{sgn} \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix} = 1, \qquad \operatorname{sgn} \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} = -1 \tag{8}$$

co prowadzi do następującej jawnej wersji wyrażenia (7) na wyznacznik:

$$\det \begin{bmatrix} a^1{}_1 & a^1{}_2 \\ a^2{}_1 & a^2{}_2 \end{bmatrix} = a^1{}_1 a^2{}_2 - a^1{}_2 a^2{}_1 \tag{9}$$

– dobrze znany nam juz *schemat Sarrusa* dla wyznacznika macierzy 2×2.

• • • Dla macierzy 3×3, mamy sześć permutacji, dla których wypisywaliśmy we wzorze (**??**). Mamy:

$$\det \begin{bmatrix} a^1{}_1 & a^1{}_2 & a^1{}_3 \\ a^2{}_1 & a^2{}_2 & a^2{}_3 \\ a^3{}_1 & a^3{}_2 & a^3{}_3 \end{bmatrix} = a^1{}_1 a^2{}_2 a^3{}_3 + a^3{}_1 a^1{}_2 a^2{}_3 + a^2{}_1 a^3{}_2 a^1{}_3$$

$$- a^2{}_1 a^1{}_2 a^3{}_3 - a^3{}_1 a^2{}_2 a^1{}_3 - a^1{}_1 a^3{}_2 a^2{}_3 \tag{10}$$

– znów *schemat Sarrusa* dla wyznacznika macierzy 3×3.

• • • • Dla macierzy 4×4 już darujemy sobie[1] wypisanie jawnej postaci wyznacznika. Są tu $4! = 24$ wyrazy i bez istotnej potrzeby nie ma po co tego wypisywać.

Zobaczyliśmy właśnie, że już przy macierzy 4×4 wzór (7) na wyznacznik jest mało praktyczny. A stopień pracochłonności przy liczeniu wyznacznika ze wzoru (7) szybko rośnie wraz z n. Już przy rozmiarze wyznacznika rzędu 30 czas jego liczenia na najlepszych komputerach staje się porównywalny z wiekiem Wszechświata. Na szczęście, istnieją bardziej praktyczne metody liczenia wyznaczników. Do poznania ich zaraz przejdziemy.

1.4.8 Wyznacznik macierzy a wyznacznik macierzy transponowanej

Z wzoru (7) pokażemy następującą własność wyznacznika.

Tw. Dla dowolnej macierzy A, wyznacznik macierzy A jest równy wyznacznikowi macierzy transponowanej:

$$\det A = \det(A^T)$$

(dla przypomnienia, przejście od macierzy do macierzy transponowanej polega na zamianie wierszy na kolumny i vice versa).

Dow. Najsampierw zauważmy, że dla dowolnej permutacji π, mamy

$$\operatorname{sgn}(\pi) = \operatorname{sgn}(\pi^{-1})$$

co wynika z własności: $\operatorname{sgn}(\sigma \circ \pi) = \operatorname{sgn}(\sigma) \cdot \operatorname{sgn}(\pi)$, a to wynika wprost z definicji znaku permutacji. (Mamy więc: $\operatorname{sgn}(\pi^{-1} \circ \pi) = \operatorname{sgn}(\mathrm{Id}) = \operatorname{sgn}(\pi^{-1}) \cdot \operatorname{sgn}(\pi)$, czyli znaki π oraz π^{-1} są jednocześnie albo 1, albo -1).

Zapiszmy teraz i poprzekształcajmy wzór (7) na $\det A$:

$$\det A = \sum_{\pi \in S_n} \operatorname{sgn}(\pi) a^{\pi(1)}{}_1 a^{\pi(2)}{}_2 \dots a^{\pi(n)}{}_n$$

[1] Podobnie jak Al – bohater negatywny 'Toy Story 2' – darował sobie prysznic przed wyjazdem do Tokio

$$= \sum_{\pi \in S_n} \operatorname{sgn}(\pi) a^1{}_{\pi^{-1}(1)} a^2{}_{\pi^{-1}(2)} \dots a^n{}_{\pi^{-1}(n)}$$
$$= \sum_{\pi \in S_n} \operatorname{sgn}(\pi^{-1}) a^1{}_{\pi^{-1}(1)} a^2{}_{\pi^{-1}(2)} \dots a^n{}_{\pi^{-1}(n)}$$
$$= \sum_{\sigma^{-1} \in S_n} \operatorname{sgn}(\sigma) a^1{}_{\sigma(1)} a^2{}_{\sigma(2)} \dots a^n{}_{\sigma(n)}$$
$$= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a^1{}_{\sigma(1)} a^2{}_{\sigma(2)} \dots a^n{}_{\sigma(n)}$$

CBDO

1.4.9 Rozwinięcie Laplace'a

Miejmy macierz A, i zapiszmy ją jako zespół n kolumn $\{a_k\}$, $k = 1, \dots, n$:

$$A = [a_1, a_2, \dots, a_n].$$

Weźmy jakąś kolumnę (np. $j-$tą) i dodajmy do niej wielokrotność jakiejś innej kolumny, np. $i-$tej. Jak pamiętamy, wartość wyznacznika się po takiej operacji *nie zmienia*:

$$\det[a_1, a_2, \dots, a_j, \dots, a_n] = \det[a_1, a_2, \dots, a_j + \lambda a_i, \dots, a_n]$$

bo

$$\det[a_1, a_2, \dots, a_i, \dots, a_j + \lambda a_i, \dots, a_n] = \det[a_1, a_2, \dots, a_i, \dots, a_j, \dots, a_n]$$
$$+\lambda \det[a_1, a_2, \dots, a_i, \dots, a_i, \dots, a_n]$$

a wyznacznik tej drugiej macierzy jest równy zeru, bo ma ona dwie takie same kolumny a_i.

Weźmy teraz macierz A i ustalmy i – numer ustalonej $i-$tej kolumny. Mamy:

$$\det A = \operatorname{vol}(a_1, a_2, \dots, a_i, \dots, a_n).$$

Zapiszmy $i-$tą kolumnę jako kombinację liniową wektorów bazy standardowej:

$$a_i = a^1{}_i e_1 + a^2{}_i e_2 + \dots + a^n{}_i e_n$$

wstawmy to rozwinięcie do wyznacznika i skorzystajmy z liniowości wyznacznika względem dowolnego argumentu:

$$\det A = \operatorname{vol}(a_1, a_2, \dots, a^1{}_i e_1 + a^2{}_i e_2 + \dots + a^n{}_i e_n, \dots, a_n) = \sum_{j=1}^{n} a^j{}_i \operatorname{vol}(a_1, a_2, \dots, e_j, \dots$$

(11)

Wypiszmy jawnie $k-$ty składnik powyższej sumy:

$$a^k{}_i \mathrm{vol}\,(a_1, a_2, \ldots, e_k, \ldots, a_n) = a^k{}_i \det \begin{bmatrix} a^1{}_1 & a^1{}_2 & \ldots & 0 & \ldots & a^1{}_n \\ a^2{}_1 & a^2{}_2 & \ldots & 0 & \ldots & a^2{}_n \\ \vdots & \vdots & & \vdots & & \vdots \\ a^k{}_1 & a^k{}_2 & \ldots & 1 & \ldots & a^k{}_n \\ \vdots & \vdots & & \vdots & & \vdots \\ a^n{}_1 & a^n{}_2 & \ldots & 0 & \ldots & a^n{}_n \end{bmatrix}$$

Odejmujmy teraz $k-$tą kolumnę od wszystkich pozostałych, ze współczynnikami a^k_j, aby w całym $i-$ tym wierszu otrzymać zera (z wyjątkiem, oczywiście, $k-$tego miejsca w wierszu). W ten sposób mamy:

$$a^k{}_i \mathrm{vol}\,(a_1, a_2, \ldots, e_k, \ldots, a_n) = a^k{}_i \mathrm{vol}\,(a_1 - a^k{}_1 e_k, a_2 - a^k{}_2 e_k, \ldots, e_k, \ldots, a_n - a^k{}_n e_k)$$

$$= a^k{}_i \det \begin{bmatrix} a^1{}_1 & a^1{}_2 & \ldots & 0 & \ldots & a^1{}_n \\ a^2{}_1 & a^2{}_2 & \ldots & 0 & \ldots & a^2{}_n \\ \vdots & \vdots & & \vdots & & \vdots \\ 0 & 0 & \ldots & 1 & \ldots & 0 \\ \vdots & \vdots & & \vdots & & \vdots \\ a^n{}_1 & a^n{}_2 & \ldots & 0 & \ldots & a^n{}_n \end{bmatrix} = \ldots$$

W otrzymanej powyżej macierzy przestawmy teraz $i-$tą kolumnę na pierwsze miejsce. Musimy przy tym zmienić znak $i-1$ razy. Następnie $k-$ty wiersz na pierwsze miejsce, znów zmieniając znak $k-1$ razy. Mamy więc:

$$\cdots = (-1)^{i-1}(-1)^{k-1} a^k{}_i \det \begin{bmatrix} 1 & 0 & 0 & \ldots & 0 & 0 & \ldots & 0 \\ 0 & a^1{}_1 & a^1{}_2 & \ldots & a^1{}_{i-1} & a^1{}_{i+1} & \ldots & a^1{}_n \\ 0 & a^2{}_1 & a^2{}_2 & \ldots & a^2{}_{i-1} & a^2{}_{i+1} & \ldots & a^2{}_n \\ & & & & & & & \\ 0 & a^{k-1}{}_1 & a^{k-1}{}_2 & \ldots & a^{k-1}{}_{i-1} & a^{k-1}{}_{i+1} & \ldots & a^{k-1}{}_n \\ 0 & a^{k+1}{}_1 & a^{k+1}{}_2 & \ldots & a^{k+1}{}_{i-1} & a^{k+1}{}_{i+1} & \ldots & a^{k+1}{}_n \\ & & & & & & & \\ 0 & a^n{}_1 & a^n{}_2 & \ldots & a^n{}_{i-1} & a^n{}_{i+1} & \ldots & a^n{}_n \end{bmatrix}$$

Łatwo zobaczyć, że ten ostatni wyznacznik jest równy wyznacznikowi nastę-

pującej macierzy rozmiaru $(n-1)\times(n-1)$:

$$\begin{bmatrix} a^1{}_1 & a^1{}_2 & \dots & a^1{}_{i-1} & a^1{}_{i+1} & \dots & a^1{}_n \\ a^2{}_1 & a^2{}_2 & \dots & a^2{}_{i-1} & a^2{}_{i+1} & \dots & a^2{}_n \\ & & & & & & \\ a^{k-1}{}_1 & a^{k-1}{}_2 & \dots & a^{k-1}{}_{i-1} & a^{k-1}{}_{i+1} & \dots & a^{k-1}{}_n \\ a^{k+1}{}_1 & a^{k+1}{}_2 & \dots & a^{k+1}{}_{i-1} & a^{k+1}{}_{i+1} & \dots & a^{k+1}{}_n \\ & & & & & & \\ a^n{}_1 & a^n{}_2 & \dots & a^n{}_{i-1} & a^n{}_{i+1} & \dots & a^n{}_n \end{bmatrix}$$

Uzasadnienie dla niedowiarków,

która – jak widzieliśmy – powstała z wyjściowej macierzy A przez wykreślenie $k-$tego wiersza i $i-$tej kolumny.

Wyznacznik tej macierzy z występującym wyżej znakiem, tzn.

Def. Wyznacznik macierzy powstałej z wyjściowej macierzy A przez wykreślenie $k-$tego wiersza i $i-$tej kolumny, pomnożony przez $(-1)^{k+i}$, nazywamy *dopełnieniem algebraicznym* wyrazu $a^k{}_i$ i oznaczamy $A^i{}_k$.

W ten sposób, $k-$ ty wyraz sumy (11), który tak pracowicie przekształcaliśmy, możemy przepisać jako

$$a^k{}_i A^i{}_k,$$

zaś całą sumę (11), tzn. $\det A$, jako

$$\det A = \sum_{j=1}^{n} a^j{}_i A^i{}_j \tag{12}$$

Wzór, który właśnie otrzymaliśmy, nazywa się **rozwinięcie Laplace'a** (wyznacznika macierzy A względem $i-$tej kolumny).

Z niezmienniczości wyznacznika względem transpozycji macierzy widzimy, iż równie dobrze możemy rozwijać względem *wierszy* jak względem kolumn. Tak więc równoważnie możemy napisać rozwinięcie wyznacznika macierzy A względem $j-$tego wiersza

$$\det A = \sum_{k=1}^{n} a^j{}_k A^k{}_j \tag{13}$$

Uwaga. Wzory (12) i (13) są bliźniaczo podobne, ale *nie identyczne* – zwróćmy uwagę, po czym sumujemy w jednym i drugim przypadku!

Przykł.

Z rozwinięcia Laplace'a ważne i często pożyteczne wzory: wzór na *macierz odwrotną* oraz tzw. *wzory Cramera* dotyczące rozwiązywalności układów równań liniowych.

1.4.10 Wzór na macierz odwrotną

Niech A będzie macierzą *nieosobliwą*, tzn. taką, że $\det A \neq 0$. Wiemy, że wtedy istnieje macierz odwrotna A^{-1}. Oznaczmy przez A^D macierz

$$A^D = \begin{bmatrix} A^1{}_1 & A^1{}_2 & \dots & A^1{}_n \\ A^2{}_1 & A^2{}_2 & \dots & A^2{}_n \\ \vdots & \vdots & & \vdots \\ A^n{}_1 & A^n{}_2 & \dots & A^n{}_n \end{bmatrix} \tag{14}$$

zwaną *macierzą dopełnień algebraicznych*, lub też *macierzą dołączoną*. Obliczmy iloczyn $A\,A^D$ oznaczając go M:

$$M \equiv AA^D = \begin{bmatrix} a^1{}_1 & a^1{}_2 & \dots & a^1{}_n \\ a^2{}_1 & a^2{}_2 & \dots & a^2{}_n \\ \vdots & \vdots & & \vdots \\ a^n{}_1 & a^n{}_2 & \dots & a^n{}_n \end{bmatrix} \begin{bmatrix} A^1{}_1 & A^1{}_2 & \dots & A^1{}_n \\ A^2{}_1 & A^2{}_2 & \dots & A^2{}_n \\ \vdots & \vdots & & \vdots \\ A^n{}_1 & A^n{}_2 & \dots & A^n{}_n \end{bmatrix}$$

Oznaczmy elementy macierzy M jako $m^i{}_j$. $k-$ty element diagonalny to:

$$m^k{}_k = \sum_{j=1}^{n} a^k{}_j A^j{}_k = \det A;$$

równość ta wynika z *rozwinięcia Laplace'a* (w wersji wierszowej). A ile wynoszą wyrazy pozadiagonalne, tzn. $m^k{}_l$ dla $k \neq l$? Otóż:

$$m^k{}_l = \sum_{j=1}^{n} a^k{}_j A^j{}_l;$$

powyższą sumę można interpretować jako wyznacznik macierzy, w której w $k-$ty oraz $l-$ty wiersz są równe $k-$temu wierszowi macierzy A. (rys.) Tak więc!! otrzymujemy:

$$AA^D = \det A \cdot I_n,$$

(I_n – macierz jednostkowa). Analogicznie, wykorzystując rozwinięcie Laplace'a względem kolumn, otrzymujemy

$$A^D A = \det A \cdot I_n,$$

W ten sposób otrzymaliśmy

Tw. Jeśli $\det A \neq 0$, to macierz odwrotna A^{-1} dana jest wzorem

$$A^{-1} = \frac{1}{\det A} A^D. \tag{15}$$

1.4.11 Wzory Cramera

Niech teraz A będzie macierzą układu równań (n równań na n niewiadomych):

$$\left\{\begin{array}{ccccccccc} a^1{}_1x^1 & + & a^1{}_2x^2 & + & \dots & + & a^1{}_nx^n & = & b^1 \\ a^2{}_1x^1 & + & a^2{}_2x^2 & + & \dots & + & a^2{}_nx^n & = & b^2 \\ \vdots & & \vdots & & \dots & & \vdots & & \vdots \\ a^n{}_1x^1 & + & a^n{}_2x^2 & + & \dots & + & a^n{}_nx^n & = & b^n \end{array}\right. \quad \text{tzn.} \quad Ax = b \tag{16}$$

Jeśli wyznacznik $\det A \neq 0$, to istnieje A^{-1} i układ równań ma jednoznaczne rozwiązanie:

$$x = A^{-1}b.$$

Korzystając z wzoru na macierz odwrotną, widzimy, iż $k-$ta składowa wektora x jest równa

$$x^k = \frac{1}{\det A}\sum_{j=1}^{n} A^k{}_l b^l.$$

A teraz! Rozpatrzmy macierz B_k, która powstała z A przez wymianę $k-$tej kolumny na kolumnę wyrazów wolnych b:

$$\left[\begin{array}{ccccccccc} a^1{}_1 & a^1{}_2 & \dots & a^1{}_{k-1} & b^1 & a^1{}_{k+1} & \dots & a^1{}_n \\ a^2{}_1 & a^2{}_2 & \dots & a^2{}_{k-1} & b^2 & a^2{}_{k+1} & \dots & a^2{}_n \\ \vdots & \vdots & & & & & & \vdots \\ a^n{}_1 & a^n{}_2 & \dots & a^n{}_{k-1} & b^n & a^n{}_{k+1} & \dots & a^n{}_n \end{array}\right]$$

Licząc wyznacznik tej macierzy przez rozwinięcie Laplace'a względem $k-$tej kolumny, widzimy, że

$$\sum_{j=1}^{n} A^k{}_l b^l = \det B_k,$$

co daje nam jawny wzór na rozwiązanie układu równań (16):

$$x^k = \frac{\det B_k}{\det A} \tag{17}$$

które to wzory znaliśmy już dla $n = 2$ i $n = 3$ (choć dla $n = 3$ był ten wzór na wykładzie podany 'na wiarę', bez uzasadnienia) – a teraz widzimy, że analogiczne jak dla $n = 2$ i $n = 3$ wzory mają miejsce dla dowolnego rozmiaru układu.

Uwaga. o (nie)praktyczności wzorów Cramera do rozwiązywania ukl. równań z konkretnymi współczynnikami liczbowymi

2 Formy kwadratowe

2.1 Formy biliniowe – przypomnienie i uzupełnienie

Def. *Formą biliniową* nazywamy odwzorowanie $b : V \times V \to \mathbb{R}$, liniowe w każdym argumencie, tzn. spełniające dla dowolnych wektorów $u, v, w \in V$, $\alpha, \beta \in \mathbb{R}$:

$$b(\alpha u+\beta v, w) = \alpha b(u, w)+\beta b(v, w), \qquad b(w, \alpha u+\beta v) = \alpha b(w, u)+\beta b(w, v). \tag{18}$$

Przykł. Standardowy iloczyn skalarny jest formą biliniową.

Przykł. Iloczyn wektorowy na $V = \mathbb{R}^3$ jest formą biliniową (oraz 2−formą – dla przypomnienia see rozdz. o formach bilin. i 2−formach).

Ilu warunków potrzeba, aby w pełni scharakteryzować formę biliniową? Tzn. umieć podać jej wartość na dowolnej parze wektorów.

Konkretnie, zadajmy w przestrzeni V jakąś bazę $e = (e_1, e_2, \ldots, e_n)$. Weźmy dwa wektory

$$\mathbf{x} = \sum_{i=1}^{n} x^i e_i, \qquad \mathbf{y} = \sum_{i=j}^{n} y^j e_j$$

Wartość formy biliniowej b na wektorach $\mathbf{x}, \mathbf{y}$ możemy, korzystając z warunków (18) liniowości w każdym argumencie, zapisać jako

$$b(\mathbf{x}, \mathbf{y}) = b(\sum_{i=1}^{n} x^i e_i, \sum_{i=j}^{n} y^j e_j) = \sum_{i=1}^{n}\sum_{j=1}^{n} b(e_i, e_j)x^i y^j \equiv \sum_{i=1}^{n}\sum_{j=1}^{n} b_{ij}x^i y^j \tag{19}$$

gdzie oznaczyliśmy:

$$b_{ij} = b(e_i, e_j).$$

Aby podać wartość 2−formy na dowolnej parze wektorów, musimy znać wszystkie wartości formy na wszystkich możliwych parach wektorów liniowo niezależnych, tzn. musimy znać macierz b_{ij} – co daje nam n^2 warunków, niezbędnych do scharakteryzowania dowolnej formy biliniowej.

2.2 Formy biliniowe symetryczne i antysymetryczne

Def. Mówimy, że forma biliniowa b jest *symetryczna*, gdy dla dowolnych dwóch wektorów $\mathbf{x}, \mathbf{y} \in V$ zachodzi

$$b(\mathbf{x}, \mathbf{y}) = b(\mathbf{y}, \mathbf{x}) \tag{20}$$

oraz że jest *antysymetryczna*, jeśli dla dowolnych dwóch wektorów $\mathbf{x}, \mathbf{y} \in V$

$$b(\mathbf{x}, \mathbf{y}) = -b(\mathbf{y}, \mathbf{x}) \tag{21}$$

(formę biliniową antysymetryczną nazywamy – jak pamiętamy – 2−formą).

Mamy proste

Stw. Każda forma biliniowa b daje się jednoznacznie przedstawić jako suma pewnej formy *symetrycznej* b_s i *antysymetrycznej* b_a:

$$b(\mathbf{x},\mathbf{y}) = b_s(\mathbf{x},\mathbf{y}) + b_a(\mathbf{x},\mathbf{y}). \tag{22}$$

Dow. (zgaduj zgadula) Zgadnijmy, że:

$$b_s(\mathbf{x},\mathbf{y}) = \frac{1}{2}\left(b(\mathbf{x},\mathbf{y}) + b(\mathbf{y},\mathbf{x})\right), \qquad b_a(\mathbf{x},\mathbf{y}) = \frac{1}{2}\left(b(\mathbf{x},\mathbf{y}) - b(\mathbf{y},\mathbf{x})\right)$$

Bezpośrednim rachunkiem sprawdzamy, że b_s jest symetryczna, zaś b_a – antysymetryczna, oraz że jest spełniony warunek (22).

CBDO

2.3 Jak się zmienia macierz formy biliniowej przy zmianie bazy

Wróćmy do wzoru (19), dającego wartość formy biliniowej b na parze wektorów $\mathbf{x},\mathbf{y}$, i napiszmy ten sam wzór w takiej notacji, aby jawnie zaznaczyć zależność od bazy. Niech w przestrzeni V będzie zadana baza e Oznaczmy:

$$\mathbf{x} \equiv [x]^e, \quad \mathbf{y} \equiv [y]^e, \tag{23}$$

oraz niech $[b]_e$ oznacza macierz, której elementy macierzowe b_{ij} są określone przez

$$b_{ij} = b(e_i, e_j) \tag{24}$$

Potraktujmy powyższe wielkości, tzn. $[x]^e, [y]^e, [b]_e$ jako *macierze*. Zauważmy, że wzór(19), dający wartość formy biliniowej b na parze wektorów $[x]^e, [y]^e$, można zapisać jako mnożenie macierzy:

$$b(\mathbf{x},\mathbf{y}) = ([x]^e)^T \cdot [b]_e \cdot [y]^e. \tag{25}$$

($\cdot$ oznacza mnożenie macierzy). Weźmy teraz inną bazę f w przestrzeni V. Załóżmy, że znamy macierz $[b]_e$ formy b w bazie e oraz macierze przejścia $[\mathrm{Id}]^f{}_e, [\mathrm{Id}]^e{}_f$ pomiędzy bazami e oraz f. Zapytajmy teraz: *Jak znaleźć macierz $[b]_f$, tzn. macierz formy b w bazie f?*

Aby odpowiedzieć na to pytanie, przypomnijmy sobie, jak powiązane są ze sobą przedstawienia wektora $\mathbf{x}$ w bazach e, f:

$$[x]^e = [\mathrm{Id}]^e{}_f \cdot [x]^f, \qquad [y]^e = [\mathrm{Id}]^e{}_f \cdot [y]^f$$

Pamiętajmy też, że *transpozycja odwraca kolejność mnożenia macierzy*: $(AB)^T = B^T A^T$. W odniesieniu do wektora $[x]^e$ mamy:

$$([x]^e)^T = ([\mathrm{Id}\,]^e{}_f \cdot [x]^f)^T = ([x]^f)^T \cdot ([\mathrm{Id}\,]^e{}_f)^T.$$

Mamy więc:

$$b(\mathbf{x},\mathbf{y}) = ([x]^e)^T \cdot [b]_e \cdot [y]^e = ([x]^f)^T \cdot \underbrace{([\mathrm{Id}\,]^e{}_f)^T \cdot [b]_e \cdot [\mathrm{Id}\,]^e{}_f}_{\text{To interpretujemy jako } [b]_f} \cdot [y]^f \equiv ([x]^f)^T \cdot [b]_f \cdot [y]^f$$

Mamy więc szukany wzór transformacyjny:

$$[b]_f = ([\mathrm{Id}\,]^e{}_f)^T \cdot [b]_e \cdot [\mathrm{Id}\,]^e{}_f \tag{26}$$

Wzór ten warto porównać z wzorem na transformację macierzy odwzorowania liniowego przy zmianie bazy. Przypomnijmy to:

Niech T będzie odwzorowaniem liniowym $V \to V$. Pamiętamy wzór na macierze odwzorowań liniowych $[T]^e{}_e$, $[T]^f{}_f$ odwzorowania liniowego w bazach e, e oraz f, f odpowiednio:

$$[T]^f{}_f = [\mathrm{Id}\,]^f{}_e \cdot [T]^e{}_e \cdot [\mathrm{Id}\,]^e{}_f$$

Def. *Rzędem* formy biliniowej nazywamy rząd odpowiadającej jej macierzy – w dowolnej bazie.

Zachodzi bowiem:

Stw. Rząd macierzy formy biliniowej nie zależy od bazy.

Pełnego dowodu tu nie podamy. Zauważmy jedynie, że gdy rząd jest maksymalny (tzn. równy $n = \dim V$), to stwierdzenie jest oczywiste: Z wzoru (26) wynika bowiem, że macierz nieosobliwa przechodzi w macierz nieosobliwą. W ogólnym przypadku trzeba pokazać, że rząd nie zmienia się przy transformacji nieosobliwej (tzn. realizowanej przez macierz nieosobliwą); jest tu trochę więcej gimnastyki i tę część dowodu pominiemy.

2.4 Formy kwadratowe. Wzór polaryzacyjny

Niech b będzie formą biliniową na przestrzeni wektorowej V.

Def. *Formą kwadratową* Q na przestrzeni wektorowej V nazywamy odwzorowanie $Q : V \to \mathbb{K}$, określone jako

$$Q(\mathbf{x}) = b(\mathbf{x},\mathbf{x})$$

Uwagi.

1. Widzimy, że w definicji Q pojawia się forma biliniowa b; z tego względu czasem dopowiadamy, że forma Q jest stowarzyszona z formą biliniową b.

2. Zauważmy, że do określenia Q wchodzi tylko część *symetryczna* formy b, jako że $b_a(\mathbf{x}, \mathbf{x}) \equiv 0$ dla dowolnego wektora $\mathbf{x}$.

Tak więc forma kwadratowa jest wyznaczona przez część symetryczną formy biliniowej b. Zapytajmy, czy na odwrót: Czy można wyznaczyć częśc symetryczną formy biliniowej b, znając formę kwadratową Q? Odpowiedź jest pozytywna. Weźmy dowolny inny wektor $\mathbf{y}$ (niewspółliniowy z $\mathbf{x}$) i wyliczmy:

$$Q(\mathbf{x}+\mathbf{y}) = b_s(\mathbf{x}+\mathbf{y}, \mathbf{x}+\mathbf{y}) = b_s(\mathbf{x}, \mathbf{x})+2b_s(\mathbf{x}, \mathbf{y})+b_s(\mathbf{y}, \mathbf{y}) = Q(\mathbf{x})+Q(\mathbf{y})+2b_s(\mathbf{x}, \mathbf{y}),$$

skąd wynika wzór na 'rekonstrukcję' (części symetrycznej) formy biliniowej z formy kwadratowej:

$$b_s(\mathbf{x}, \mathbf{y}) = \frac{1}{2}\left(Q(\mathbf{x}+\mathbf{y}) - Q(\mathbf{x}) - Q(\mathbf{y})\right) \tag{27}$$

Powyższy wzór nazywamy *wzorem polaryzacyjnym.*

Niech teraz mamy bazę e w V. Z wzoru (19) mamy:

$$Q(\mathbf{x}) = \sum_{i=1}^{n}\sum_{j=1}^{n} q_{ij}x^i x^j \tag{28}$$

gdzie

$$q_{ij} = (b_s)_{ij}$$

jest symetryczną macierzą nazywaną *macierzą formy kwadratowej* (w bazie e).

Def. *Rzędem* formy kwadratowej nazywamy rząd macierzy tej formy (w jakiejkolwiek bazie – rząd nie zależy od bazy).

2.5 Przykłady form kwadratowych

1. Rozważmy jednorodny[2] wielomian drugiego stopnia od n zmiennych $x_1, x_2, \ldots, x_n$:

   $$Q(x_1, x_2, \ldots, x_n) = \sum_{i=1}^{n}\sum_{j=1}^{n} q_{ij}x^i x^j;$$

   widać, że jest to dokładnie prawa strona wzoru (28). Tak więc wielomian jednorodny 2. stopnia możemy uważać za formę kwadratową.

[2]tzn. sumaryczna potęga każdego członu jest równa 2

2. Rozważmy dwa sprzężone jednowymiarowe oscylatory harmoniczne. **RYS.** x_1, x_2 są wychyleniami z położenia równowagi. Energię potencjalną takiego układu możemy zapisać jako:

$$E(x_1, x_2) = \frac{1}{2}\left(k_1 x_1^2 + k_2 x_2^2 + k_{12}(x_1 - x_2)^2\right) = \frac{1}{2}\left((k_1 + k_{12})x_1^2 + (k_2 + k_{12})x_2^2 -$$

a więc znowu jest to forma kwadratowa. Ogólnie, dla n sprzężonych oscylatorów harmonicznych, energię kinetyczną możemy zapisać jako

$$E(x_1, x_2, \ldots, x_n) = \sum_{i=1}^{n}\sum_{j=1}^{n} k_{ij} x^i x^j;$$

występującą tu (symetryczną) macierz k_{ij} nazywa się często *macierzą stałych siłowych*.

2.6 Metoda Lagrange'a diagonalizacji formy kwadratowej

Def. Formę kwadratową nazywamy *diagonalną*, jeśli macierz tej formy jest diagonalna, tzn. jedyne jej niezerowe elementy występują na przekątnej.

Tw. (Lagrange'a). Każda forma kwadratowa ma bazę diagonalizującą, tzn. taką, że macierz formy w tej bazie jest diagonalna.

Dow. Zajmujmy się formą we współrzędnych, tzn. $Q(\mathbf{x}) = \sum_{i=1}^{n}\sum_{j=1}^{n} q_{ij} x^i x^j$. Zamiana bazy będzie się odbywała poprzez transformacje *składowych* wektora.

Rozważmy dwa przypadki: **i)** Macierz formy posiada chociaż jeden element diagonalny różny od zera, **ii)** macierz formy ma wszystkie elementy diagonalne równe zeru.

i) Załóżmy, że elementem różnym od zera jest q_{11} (gdy tak nie jest, to możemy przenumerować składowe).

Zamiana bazy będzie się odbywała w kolejnych krokach (ilości co najwyżej n) i będzie polegała na 'uzupełnieniu do pełnego kwadratu'.

Krok 1. Zapiszmy formę w następującej postaci, wydzielając jawnie wskaźnik 1 oraz pozostałe, oraz korzystając z tego, że macierz jest symetryczna, co w szczególności daje $q_{1i} = q_{i1}$:

$$Q(\mathbf{x}) \equiv Q(x^1, x^2, \ldots, x^n) = \sum_{i=1}^{n}\sum_{j=1}^{n} q_{ij} x^i x^j = q_{11} x^1 x^1 + 2\sum_{i=2}^{n} q_{1i} x^1 x^i + \sum_{i=2}^{n}\sum_{j=2}^{n} q_{ij} x^i x^j$$

$$= q_{11}\left(x^1 + \frac{1}{q_{11}}\sum_{i=2}^{n} q_{1i} x^i\right)^2 - \frac{1}{q_{11}}\left(\sum_{i=2}^{n} q_{1i} x^i\right)^2 + \sum_{i=2}^{n}\sum_{j=2}^{n} q_{ij} x^i x^j = \ldots$$

Nazwijmy teraz: $X^1 = x^1 + \frac{1}{q_{11}} \sum_{i=2}^n q_{1i} x^i$; możemy więc formę zapisać jako

$$Q(X^1, x^2, \ldots, x^n) = q_{11} X^1 X^1 - \frac{1}{q_{11}} \left(\sum_{i=2}^n q_{1i} x^i \right)^2 + \sum_{i=2}^n \sum_{j=2}^n q_{ij} x^i x^j.$$

Zauważmy, że w sumach powyżej *nie występuje już* x^1 – są tam jedynie współrzędne o numerach 2 i wyższych. Możemy więc zapisać

$$Q(X^1, x^2, \ldots, x^n) = q_{11} X^1 X^1 + \sum_{i=2}^n \sum_{j=2}^n c_{ij} x^i x^j$$

dla pewnej macierzy c_{ij}. (Można wyrazić c_{ij} przez q_{ij}, ale nie jest to potrzebne, więc sobie darujemy).

Widzimy więc, że forma Q jest diagonalna w zmiennej X^1.

Krok 2. Jeśli macierz c_{ij} posiada chociaż jeden element diagonalny różny od zera, to do formy kwadratowej $\sum_{i=2}^n \sum_{j=2}^n c_{ij} x^i x^j$ stosujemy znów sposób postępowania analogiczny jak w kroku 1. Jeśli nie – to stosujemy sposób postępowania opisany poniżej w **ii)**. Itd;

...

ii). Trzeba teraz powiedzieć, co robimy, gdy forma kwadratowa nie ma ani jednego niezerowego elementu diagonalnego. Macierz formy musi mieć wtedy przynajmniej jeden niezerowy wyraz niediagonalny (gdyby miała wszystkie elementy niediagonalne równe zeru, to byłaby formą totalnie zerową, i jako taka byłaby diagonalna). Załóżmy, że elementem niediagonalnym macierzy jest q_{12} (gdyby tak nie było, zawsze możemy przez przenumerowanie bazy osiągnąć to, aby $q_{12} \neq 0$). Forma kwadratowa ma więc postać:

$$Q(x^1, x^2, \ldots, x^n) = q_{12} x^1 x^2 + \ldots.$$

Zamieńmy współrzędne następująco: $(x^1, x^2) \to (X^1, X^2)$:

$$x^1 = X^1 + X^2, \qquad x^2 = X^1 - X^2.$$

Wtedy forma przybiera postać:

$$Q(X^1, X^2, x^3, \ldots, x^n) = q_{12} X^1 X^1 - q_{12} X^2 X^2 + \ldots$$

Widać, że przez powyższą zamianę zmiennych doprowadziliśmy do tego, że forma ma wyraz diagonalny. Teraz stosujemy metodę diagonalizacji z **i)**.

Cała procedura diagonalizacji zakończy się w co najwyżej n krokach.

CBDO

Przykł.

1. Rozpatrzmy formę:

$$Q(x^1, x^2, x^3) = x^1x^1 + 2x^1x^2 + x^2x^2 + 4x^2x^3 + 5x^3x^3$$

Postępując jak w **i)**, weźmy:

$$X^1 = x^1 + x^2, \quad \Longrightarrow \quad x^1 = X^1 - x^2$$

i mamy:

$$Q(X^1, x^2, x^3) = X^1X^1 - 2X^1x^2 + x^2x^2 + 2X^1x^2 - 2x^2x^2 + 4x^2x^3 + 5x^3x^3$$

$$= X^1X^1 + -x^2x^2 + 4x^2x^3 + 5x^3x^3$$

tzn. w pierwszym kroku, w nowej zmiennej X^1 mamy wyraz wyłącznie diagonalny.

Dalej: Weźmy:

$$X^2 = x^2 - 2x^3, \quad \Longrightarrow \quad x^2 = X^2 + 2x^3.$$

Liczymy:

$$Q(X^1, X^2, x^3) = X^1X^1 - X^2X^2 - 4X^2x^3 - 4x^3x^3 + 4X^2x^3 + 8x^3x^3 + 5x^3x^3$$

$$= X^1X^1 - X^2X^2 + 9x^3x^3$$

czyli otrzymaliśmy formę w postaci diagonalnej.

Ważna w zastosowaniach okazuje się być ilość plusów i minusów na diagonali, która tu wynosi: dwa plusy i jeden minus.

2. Weźmy teraz formę, której macierz posiada wszystkie elementy diagonalne równe zeru:

$$Q(x^1, x^2, x^3) = x^1x^2 + x^1x^3 + x^2x^3.$$

Zgodnie z **ii)**, wprowadzamy:

$$x^1 = X^1 + X^2, \qquad x^2 = X^1 - X^2.$$

i mamy

$$Q(X^1, X^2, x^3) = X^1X^1 - X^2X^2 + X^1x^3 + X^2x^3 + X^1x^3 - X^2x^3$$

$$= X^1X^1 - X^2X^2 + 2X^1x^3.$$

Mamy tu już samodzielny wyraz diagonalny $-X^2X^2$. Zamieniamy więc zmienne w parze X^1, x^3 wg procedury z **i)**: Wprowadźmy $Y = X^1 + x^3$, tzn. $X^1 = Y - x^3$. Mamy:

$$Q(Y, X^2, x^3) = -X^2X^2 + (Y - x^3)(Y - x^3) + 2(Y - x^3)x^3$$

$$= -X^2X^2 + YY - 2Yx^3 + x^3x^3 + 2Yx^3 - 2x^3x^3 = -X^2X^2 + YY - x^3x^3.$$

Widzimy więc, że we współrzędnych (Y, X^2, x^3) forma Q przyjęła postać diagonalną. Znów ważna (w zastosowaniach) jest ilość plusów i minusów na diagonali: Tu dwa minusy i jeden plus.

2.7 Postać kanoniczna i sygnatura formy kwadratowej. Twierdzenie Sylvester'a

Po diagonalizacji formy otrzymamy postać, zawierającą wyłącznie wyrazy kwadratowe. Forma ma więc postać:

$$Q(y^1, \ldots, y^n) = \sum_{i=1}^{r} a_i (y^i)^2, \qquad a_i \neq 0 \ \text{ dla } \ i = 1, 2, \ldots, r. \tag{29}$$

Liczba r niezerowych wyrazów na diagonali, jak łatwo spostrzec, jest *rzędem* formy (bowiem macierz formy, posiadającej na diagonali r niezerowych wyrazów, a poza tym same zera, ma rząd r).

Można jeszcze nieco uprościć postać formy. Spotykamy tu dwie różne sytuacje zależnie od tego, czy mamy formę na przestrzeni V nad $\mathbb{R}$, czy na $\mathbb{C}$. Przypadek $\mathbb{C}$ okazuje się prostszy.

2.7.1 Formy nad $\mathbb{C}$

Przeskalujmy zmienne y^i w postaci formy (31) następująco:

$$Y^i = \sqrt{a_i} y^i$$

Można zapytać, jaki znak pierwiastka tu wybieramy? Pamiętamy bowiem, że wyciąganie pierwiastka kwadratowego nad $\mathbb{C}$ jest zawsze operacją *dwuznaczną!*. Wszystko jedno. W ten sposób osiągamy postać

$$Q(Y^1, \ldots, Y^n) = \sum_{i=1}^{r} (Y^i)^2, \quad \text{dla } \ i = 1, 2, \ldots, r. \tag{30}$$

Postać (32) nazywamy *postacią kanoniczną formy kwadratowej nad* $\mathbb{C}$.

2.7.2 Formy nad $\mathbb{R}$

Postać formy kwadratowej po diagonalizacji metodą Lagrange'a nie jest wyznaczona jednoznacznie: Mogliśmy startować od dowolnego wyrazu diagonalnego i nie ma powodu, że dostaniemy dokładnie tę samą postać niezależnie od tego, który wyraz diagonalny weźmiemy na początku. Na razie unormujmy współczynniki przy pełnych kwadratach. Napiszmy formę w postaci

$$Q(y^1,\dots,y^n)=\sum_{i=1}^{p} a_i(y^i)^2-\sum_{i=p+1}^{r} a_i(y^i)^2,\qquad a_i>0 \text{ dla } i=1,2,\dots,r. \tag{31}$$

Po unormowaniu analogicznym jak dla przypadku nad $\mathbb{C}$:

$$Y^i=\sqrt{a_i}y^i$$

(możemy tak zrobić, bo wszystkie $a_i>0$ w zapisie (31)) osiągamy postać:

$$Q(Y^1,\dots,Y^n)=\sum_{i=1}^{p}(Y^i)^2-\sum_{i=p+1}^{r}(Y^i)^2 \tag{32}$$

Powróćmy do problemu jednoznaczności takiej postaci formy. Wiemy, że rząd musi być taki sam niezależnie od sposobu diagonalizacji. Ale co z ilością plusów i minusów? Czy przez inny sposób diagonalizacji nie osiągnęlibyśmy postaci z innymi liczbami plusów i minusów? Okazuje się, że *nie*. Zachodzi bowiem

Tw. . (Sylvestera; zwane też tw. Sylvestera i Jacobiego, lub tw. o bezwładności form kwadratowych). Jeśli formę kwadratową nad $\mathbb{R}$ sprowadzimy za pomocą dwóch różnych przekształceń do postaci kanonicznych, to obie formy kanoniczne mają te same ilości współczynników dodatnich oraz współczynników ujemnych.

Uwaga. Zestaw tych dwu liczb, tzn. ilości plusów oraz ilości minusów, nazywamy *sygnaturą* formy kwadratowej. Tw. Sylvestera możemy więc wypowiedzieć w ten sposób, że *sygnatura rzeczywistej formy kwadratowej jest wyznaczona jednoznacznie.*

Dow. . Miejmy formę kwadratową $Q(\mathbf{x})\equiv Q(x^1,x^2,\dots,x^n)$. Niech dwa przekształcenia

$$x^j=\sum_{k=1}^{n} b^j{}_k y^k \quad \text{oraz} x^j=\sum_{k=1}^{n} c^j{}_k z^k \tag{33}$$

przekształcają formę $Q(\mathbf{x})$ odpowiednio w postaci diagonalne

$$Q(\mathbf{y})=a_1(y^1)^2+\dots+a_k(y^k)^2-a'_1(y^{k+1})^2-\dots-a'_l(y^{k+l})^2, \tag{34}$$

gdzie wszystkie współczynniki a_i, a'_i są większe od zera, oraz

$$Q(\mathbf{z}) = b_1(z^1)^2 + \cdots + b_m(z^m)^2 - b'_1(z^{m+1})^2 - \cdots - b'_l(z^{m+p})^2, \tag{35}$$

gdzie również wszystkie współczynniki b_i, b'_i są większe od zera.

Suma ilości niezerowych wyrazów diagonalnych dodatnich i ujemnych musi być równa rzędowi formy:

$$r = k + l = m + p. \tag{36}$$

Należy dowieść, że

$$m = k \quad \text{oraz} \quad p = l. \tag{37}$$

Przypuśćmy (dow. a.a.) że tak nie jest, i że np.

$$k > m; \tag{38}$$

z (36) wynika, że $l < p$.

Rozpatrzmy przekształcenia odwrotne do (33), tzn. wyraźmy składowe y^j, z^j wektorów $\mathbf{y}$ i $\mathbf{z}$ liniowo przez składowe $x^1, x^2, \ldots, x^n$.

Rozpatrzmy dalej następujący układ $n - k + m$ równań liniowych jednorodnych o n niewiadomych $x^1, x^2, \ldots, x^n$:

$$z^1 = 0, \;\; z^2 = 0, \;\; \ldots, \;\; z^m = 0, y^{k+1} = 0, \;\; y^{k+2} = 0, \;\; \ldots, y^n = 0. \tag{39}$$

Ponieważ $n - k + m < n$, (tzn. ilość równań jest mniejsza niż ilość niewiadomych), to układ powyższy ma rozwiązanie niezerowe. Oznaczmy jakieś z tych niezerowych rozwiązań jako $\mathbf{x}_0$:

$$\mathbf{x}_0 = [x_0^1, x_0^2, \ldots, x_0^n].$$

Wstawmy je do $Q(\mathbf{x})$ na dwa sposoby: z postaci (34) otrzymujemy

$$Q(\mathbf{x}_0) = a_1(y^1)^2 + \cdots + a_k(y^k)^2 \geqslant 0, \tag{40}$$

a z postaci (35) dostaniemy

$$Q(\mathbf{x}_0) = -b'_1(z^{m+1})^2 - \cdots - b'_p(z^{m+p})^2 \leqslant 0. \tag{41}$$

Zatem $Q(\mathbf{x}_0) = 0$. Stąd, na mocy dodatniości współczynników a_i oraz (40) wynika, że

$$y^1 = 0, \;\; y^2 = 0, \ldots, y^k = 0 \tag{42}$$

Ale wyżej widzieliśmy, p. (39), że pozostałe składow $y^{k+1}, \ldots, y^n$ też są zerowe, razem więc

$$y^1 = 0, \;\; y^2 = 0, \ldots, y^n = 0. \tag{43}$$

a stąd, na mocy pierwszej z definicji (33) widzimy, że

$$x_0^1 = 0, \quad x_0^2 = 0, \ldots, x_0^n = 0, \tag{44}$$

a to znaczy, że rozwiązanie $\mathbf{x}_0$ musi być *zerowe*. Sprzeczność.

Analogicznie jest z przypuszczeniem, że zachodzi nierówność $k < m$, czyli $p > l$.

CBDO

Sztuczka mnemotechniczna do zapamiętania tw. Sylvestera: Molesta, 'Xeroboj'.

2.8 Kryterium wyznacznikowe dodatniej określoności form kwadratowych

Def. Mówimy, że forma Q jest *dodatnio (nieujemnie) określona*, gdy dla dowolnego wektora $\mathbf{x}$ zachodzi $Q(\mathbf{x}) > 0$ ($Q(\mathbf{x}) \geqslant 0$); i bliźniaczo

Mówimy, że forma Q jest *ujemnie (niedodatnio) określona*, gdy dla dowolnego wektora $\mathbf{x}$ zachodzi $Q(\mathbf{x}) < 0$ ($Q(\mathbf{x}) \leqslant 0$).

W zastosowaniach często ważne jest stwierdzenie, czy dana forma kwadratowa jest dodatnio, bądź ujemnie określona. Niedługo zobaczymy, że badanie (lokalne) maksimów/minimów funkcji f wielu zmiennych sprowadza się do przetestowania, czy forma kwadratowa, będąca rozwinięciem Taylora f do 2. rzędu jest dodatnio/ujemnie określona.

Problem dodatniej/ujemnej określoności formy kwadratowej oczywiście można rozstrzygnąć przy pomocy metody Lagrange'a diagonalizacji formy. Rozpowszechnione jest jednak inne kryterium, które teraz podamy bez dowodu.

Tw. (kryterium dodatniej/ujemnej określoności formy kwadratowej). Niech $q = \{q_{ij}\}$ będzie macierzą formy kwadratowej Q. Utwórzmy ciąg minorów:

$$\Delta_1 = q_{11}; \quad \Delta_2 = \begin{vmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{vmatrix}; \quad \ldots, \quad \Delta_n = \det q.$$

Wtedy:

1. Jeżeli $\Delta_1 > 0, \Delta_2 > 0, \ldots, \Delta_n > 0$, to forma jest dodatnio określona.
2. Jeżeli $\Delta_1 < 0, \Delta_2 > 0, \Delta_3 < 0, \ldots$, i ogólnie dla każdego $k \leqslant n$ zachodzi: $\operatorname{sgn}(\Delta_k) = (-1)^k$, to forma jest ujemnie określona.

Dow. . można znaleźć np. w materiałach K. Grabowskiej do Matematyki II sprzed trzech lat, lub w większości książek poświęconych algebrze liniowej.

Tutaj też to udowodnimy, ale nieco później, gdy pojawi się pojęcie *wektorów* i *wartości własnych*.

Uwaga. Metoda (pozornie) nie działa, gdy któryś z minorów jest równy zeru. Jednak wtedy forma nie jest ani dodatnio ani ujemnie określona.

Ilustracja dla form, które były jako przykłady.