Wykład 2.

Matematyka 2, semestr letni 2010/2011

Kontynuujemy przygotowanie do studiowania rachunku różniczkowego funkcji wielu zmiennych. Wiedzą już państwo, że pochodna funkcji $f : \mathbb{R}^n \to \mathbb{R}$ w punkcie $x_0 \in \mathbb{R}^n$ jest odwzorowaniem liniowym działającym na wektorach zaczepionych w x_0 o wartościach w $\mathbb{R}$. Jest to obiekt nieco innego rodzaju niż wyjściowa funkcja. Podobnie rzecz się ma z drugą pochodną. Okazuje się, że druga pochodna funkcji wielu zmiennych jest formą dwuliniową na tej samej przestrzeni wektorowej, na której działa pochodna. Zanim zatem na dobre zajmiemy się różniczkowaniem funkcji wielu zmiennych, powinniśmy dowiedzieć się kilku rzeczy o formach dwuliniowych. Podobnie jak w rachunku różniczkowym funkcji jednej zmiennej, badanie drugiej pochodnej (a czasami i wyższych) jest potrzebne na przykład do określania rodzaju punktów ekstremalnych.

Definicja 1. Niech V będzie przestrzenią wektorową nad ciałem $\mathbb{R}$. *Formą dwuliniową* na przestrzeni V nazywamy odwzorowanie

$$Q : V \times V \to \mathbb{R}$$

które jest liniowe ze względu na każdy ze swoich argumentów, tzn. dla wszystkich $v_1, v_2, v \in V$ oraz $\lambda^1, \lambda^2 \in \mathbb{R}$ spełnia warunki

$$Q(\lambda^1 v_1 + \lambda^2 v_2,\, v) = \lambda^1 Q(v_1,\, v) + \lambda^2 Q(v_2,\, v) \qquad Q(v,\, \lambda^1 v_1 + \lambda^2 v_2) = \lambda^1 Q(v,\, v_1) + \lambda^2 Q(v,\, v_2).$$

Mówimy, że forma jest *symetryczna*, jeśli ponadto dla wszystkich $v_1, v_2 \in V$

$$Q(v_1,\, v_2) = Q(v_2,\, v_1),$$

a *antysymetryczna*, jeśli

$$Q(v_1,\, v_2) = -Q(v_2,\, v_1).$$

Przykład 1. Formą dwuliniową jest *iloczyn skalarny* $(\cdot|\cdot)$, o którym mówili państwo w poprzednim semestrze. Niektóre przestrzenie wektorowe, na przykład $\mathbb{R}^n$, wyposażone są w kanoniczny iloczyn skalarny. Jeśli

$$v = \begin{bmatrix} x^1 \\ x^2 \\ \vdots \\ x^n \end{bmatrix}, \qquad w = \begin{bmatrix} y^1 \\ y^2 \\ \vdots \\ y^n \end{bmatrix} \quad \text{to} \quad (v\,|\,w) = [x^1 \;\; x^2 \;\; \cdots \;\; x^n] \begin{bmatrix} y^1 \\ y^2 \\ \vdots \\ y^n \end{bmatrix} = \sum_{k=1}^{n} x^k y^k$$

♣

Przykład 2. W ruchu obrotowym bryły sztywnej związek między prędkością kątową a energią kinetyczną bryły wyraża się wzorem

$$E = \frac{1}{2} I_{(p,\overrightarrow{n})} \omega^2,$$

gdzie E oznacza energię, ω wartość prędkości kątowej, a $I_{(p,\overrightarrow{n})}$ tak zwany *moment bezwładności* względem osi o wektorze kierunkowym $\overrightarrow{n}$ przechodzącej przez punkt p. Moment bezwładności związany jest z rozkładem masy wewnątrz bryły, zależy również od położenia osi obrotu względem bryły. Należy obliczać go oddzielnie dla każdej osi obrotu. Okazuje się jednak, że informację o momentach bezwładności danej bryły względem osi o dowolnym kierunku, przechodzącej przez ustalony punkt przestrzeni p przedstawić można w postaci formie dwuliniowej symetrycznej nazywanej *tensorem bezwładności*. Zamiast wielkości $I_{(p,\overrightarrow{n})}\omega^2$ do wzoru na energię wstawić należy $I_p(\overrightarrow{\omega}, \overrightarrow{\omega})$, gdzie tym razem I_p oznacza odpowiednią formę dwuliniową. Przykład konkretnej formy dla konkretnej bryły sztywnej pojawi się w dalszej części wykładu, kiedy poznamy już sposoby zapisywania form dwuliniowych. ♣

Macierz formy dwuliniowej. Wybierzmy bazę $e = (e_1, e_2, \ldots, e_n)$ w przestrzeni wektorowej V. Zauważmy, że z warunku dwuliniowości wynika, że forma Q jest *jednoznacznie określona przez swoje wartości na bazie.* Istotnie, jeśli

$$v = \lambda^1 e_1 + \lambda^2 e_2 + \cdots + \lambda^n e_n, \qquad w = \mu^1 e_1 + \mu^2 e_2 + \cdots + \mu^n e_n$$

to

$$\begin{aligned}
Q(v, w) = Q(\lambda^1 e_1 + \lambda^2 e_2 + \cdots + \lambda^n e_n,\ \mu^1 e_1 + \mu^2 e_2 + \cdots + \mu^n e_n) = \\
\lambda^1 Q(e_1, \mu^1 e_1 + \mu^2 e_2 + \cdots + \mu^n e_n) + \lambda^2 Q(e_2, \mu^1 e_1 + \mu^2 e_2 + \cdots + \mu^n e_n) + \cdots + \\
+ \lambda^n Q(e_n, \mu^1 e_1 + \mu^2 e_2 + \cdots + \mu^n e_n) = \\
\sum_{i=1}^{n} \lambda^i Q(e_i, \mu^1 e_1 + \mu^2 e_2 + \cdots + \mu^n e_n) = \\
\sum_{i=1}^{n} \lambda^i \Big(\mu^1 Q(e_i, e_1) + \mu^2 Q(e_i, e_2) + \cdots \mu^n Q(e_i, e_n) \Big) = \sum_{i=1}^{n} \sum_{k=1}^{n} \lambda^i \mu^k Q(e_i, e_k)
\end{aligned}$$

Widzimy więc, że żeby wyznaczyć wartość $Q(v, w)$ trzeba znać współrzędne wektorów v i w w bazie oraz wartości Q na wektorach bazowych: $Q(e_i, e_k)$. Forma Q chrakteryzowana jest więc przez n^2 liczb. Jeśli wartości $Q(e_i, e_k)$ oznaczymy Q_{ik} możemy napisać:

$$Q(v, w) = \sum_{i=1}^{n} \sum_{k=1}^{n} \lambda^i \mu^k Q_{ik}$$

Liczby Q_{ik} wygodnie jest ustawić w macierz $n \times n$:

$$[Q]_e = \begin{bmatrix} Q_{11} & Q_{12} & \cdots & Q_{1n} \\ Q_{21} & Q_{22} & \cdots & Q_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ Q_{n1} & Q_{n2} & \cdots & Q_{nn} \end{bmatrix}$$

Jeśli forma jest symetryczna, jej macierz jest symetryczna względem głównej przekątnej, tzn $Q_{ik} = Q_{ki}$, jeśli jest antysymetryczna, to wyrazy na głównej przekątnej muszą być równe zero oraz macierz jest być antysymetryczna ze względu na główną przekątną.

Zwróćmy uwagę na to, że „element macierzowy” Q_{ik} ma oba wskaźniki na dole, a nie, jak było w przypadku macierzy odwzorowania liniowego, jeden wskaźnik na górze a drugi na dole. Różnica ta odzwierciedla różnicę między obiektami geometrycznymi reprezentowanymi przez macierz. Mimo, że macierz odwzorowania i macierz formy wyglądają jednakowo, to jednak działają inaczej i mają różne własności. Wzór

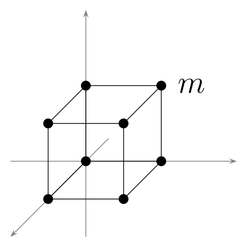
$$Q(v, w) = \sum_{i=1}^{n} \sum_{k=1}^{n} \lambda^i \mu^k Q_{ik}$$

w zapisie macierzowym wyglądałby tak:

$$(1)\qquad Q(v,w) = \begin{bmatrix} \lambda^1 & \lambda^2 & \cdots & \lambda^n \end{bmatrix} \begin{bmatrix} Q_{11} & Q_{12} & \cdots & Q_{1n} \\ Q_{21} & Q_{22} & \cdots & Q_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ Q_{n1} & Q_{n2} & \cdots & Q_{nn} \end{bmatrix} \begin{bmatrix} \mu^1 \\ \mu^2 \\ \vdots \\ \mu^n \end{bmatrix} = ([v]^e)^T [Q]_e [w]^e.$$

Przykład 3. Macierz kanonicznego iloczynu skalarnego na $\mathbb{R}^n$ względem bazy standardowej jest macierzą jednostkową. ♣

Przykład 4. Rozważamy bryłę sztywną B składającą się z ośmiu jednakowych mas m umieszczonych w wierzchołkach sześcianu rozpiętego na wektorach bazy standardowej w $\mathbb{R}^3$.



Wyrazy macierzowe momentu bezwładności wyznaczamy zgodnie ze wzorem:

$$I_{ii} = \sum_{\alpha=1}^{8} m\left[(x^1_\alpha)^2 + (x^3_\alpha)^2 + (x^3_\alpha)^2 - (x^i_\alpha)^2\right], \quad I_{ij} = -\sum_{\alpha=1}^{8} m x^i_\alpha x^j_\alpha \quad \text{dla} \;\; i \neq j,$$

gdzie α numeruje wierzchołki sześcianu. Łatwo sprawdzić, że

$$I = m \begin{bmatrix} 8 & -2 & -2 \\ -2 & 8 & -2 \\ -2 & -2 & 8 \end{bmatrix}.$$

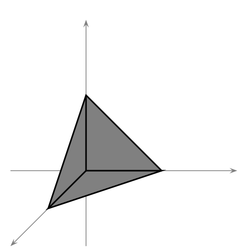
Korzystając z macierzy $[I]_e$ obliczyć możemy moment bezwładności w obrocie względem dowolnej osi przechodzącej przez początek układu współrzędnych. Na przykład względem osi, której wektorem kierunkowym jest

$$\vec{n} = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \qquad I_{\vec{n}} = m\frac{1}{3}[1 \;\; 1 \;\; 1] \begin{bmatrix} 8 & -2 & -2 \\ -2 & 8 & -2 \\ -2 & -2 & 8 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = m\frac{1}{3}[1 \;\; 1 \;\; 1] \begin{bmatrix} 4 \\ 4 \\ 4 \end{bmatrix} = 4m$$

♣

Przykład 5. Macierz momentu bezwładności jednorodnego czworościanu B o masie M rozpiętego na wektorach bazy standardowej względem początku układu współrzędnych jest postaci

$$[I]_e = \frac{M}{10} \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}.$$



Elementy macierzowe I wyznaczamy zgodnie ze wzorem

$$I_{ik} = \int_B \rho(x^1, x^2, x^3)\left((x^1)^2 + (x^2)^2 + (x^3)^2)\delta^{ik} - x^i x^k\right) \mathsf{d}x^1 \mathsf{d}x^2 \mathsf{d}x^3,$$

gdzie ρ jest gęstością bryły. Tego rodzaju całki obliczać będziemy w drugiej części semestru. Korzystając z macierzy $[I]_e$ obliczyć możemy moment bezwładności w obrocie względem dowolnej osi przechodzącej przez początek układu współrzędnych. Na przykład względem osi, której wektorem kierunkowym jest

$$\overrightarrow{n} = \frac{1}{\sqrt{3}}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \qquad I_{\overrightarrow{n}} = \frac{1}{3}\frac{M}{10}[1\ \ 1\ \ 1]\begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \frac{M}{30}[1\ \ 1\ \ 1]\begin{bmatrix} 4 \\ 4 \\ 4 \end{bmatrix} = \frac{M}{30}12 = \frac{2M}{5}$$

♣

Do rozwiązania mamy teraz następujące zadanie:

Zadanie 1. Niech Q będzie formą dwuliniową na przestrzeni wektorowej V. Znając macierz formy $[Q]_e$ w bazie e i macierze przejścia $[\mathrm{id}\,]^f{}_e$, $[\mathrm{id}\,]^e{}_f$ między bazami e i f znaleźć macierz formy $[Q]_f$ w bazie f

Rozwiązanie: Startujemy ze wzoru (1)

$$Q(v, w) = ([v]^e)^T\ [Q]_e\ [w]^e.$$

Wiemy w jaki sposób dokonuje się zmiany reprezentacji wektora w bazie za pomocą macierzy przejścia:

$$[v]^e = [\mathrm{id}\,]^e{}_f\ [v]^f, \qquad [w]^e = [\mathrm{id}\,]^e{}_f\ [w]^f.$$

Musimy też pamiętać, że transpozycja zamienia kolejność mnożenia macierzy, to znaczy

$$([v]^e)^T = ([\mathrm{id}\,]^e{}_f\ [v]^f)^T = ([v]^f)^T\ ([\mathrm{id}\,]^e{}_f)^T.$$

Wstawiamy zgodnie z kolorem:

$$Q(v, w) = ([v]^e)^T\ [Q]_e\ [w]^e = ([v]^f)^T\ ([\mathrm{id}\,]^e{}_f)^T\ [Q]_e\ [\mathrm{id}\,]^e{}_f\ [w]^f$$

W ostatnim wzorze warto wyróżnić następujący fragment:

$$Q(v, w) = ([v]^f)^T\ ([\mathrm{id}\,]^e{}_f)^T\ [Q]_e\ [\mathrm{id}\,]^e{}_f\ [w]^f$$

i porównać to z

$$Q(v, w) = ([v]^f)^T\ [Q]_f\ [w]^f.$$

Wniosek:

$$[Q]_f = ([\mathrm{id}\,]^e{}_f)^T\ [Q]_e\ [\mathrm{id}\,]^e{}_f \tag{2}$$

Wzór (2) warto porównać z odpowiednim wzorem dla macierzy odwzorowań liniowych. Niech F będzie odwzorowaniem z przestrzeni V do przestrzeni V. Wtedy do zapisania odwzorowania w postaci macierzy wystarczy jedna baza, której możemy użyć dwa razy: w dziedzinie i w obrazie, otrzymując macierz $[F]^e{}_e$. Przejście od macierzy $[F]^e{}_e$ do macierzy $[F]^f{}_f$ odbywa się według wzoru

$$[F]^f{}_f = [\mathrm{id}\,]^f{}_e\ [F]^e{}_e\ [\mathrm{id}\,]^e{}_f. \tag{3}$$

Zwróćmy uwagę na różnicę:

$$[Q]_f = ([\mathrm{id}\,]^e{}_f)^T\ [Q]_e\ [\mathrm{id}\,]^e{}_f \qquad [F]^f{}_f = [\mathrm{id}\,]^f{}_e\ [F]^e{}_e\ [\mathrm{id}\,]^e{}_f.$$

Z prawej strony starą macierz mnożymy przez tę samą macierz przejścia w obu przypadkach, ale z lewej jest macierz transponowana dla formy i odwrotna dla macierzy odwzorowania. O tej różnicy trzeba pamiętać! ♣

Przykład 6. Na przestrzeni $\mathbb{R}_2[\cdot]$ wielomianów stopnia nie większego niż 2 rozważmy teraz formę dwuliniową daną wzorem

$$Q(v,w) = 6\int_0^1 v'(t)w(t)\mathsf{d}t. \tag{4}$$

Użyjemy bazy standardowej do zapisania macierzy tej formy:

$$e_1(t) = 1, \qquad e_2(t) = t, \qquad e_3(t) = t^2.$$

Wyznaczamy wyrazy macierzowe: na przykład

$$Q(e_3, e_2) = 6\int_0^1 e_3'(t)e_2(t) = 6\int_0^1 (2t)(t)\mathsf{d}t = 6\int_0^1 2t^2\mathsf{d}t = 6\left[\frac{2}{3}t^3\right]_0^1 = 4.$$

Podobnie:

$$\begin{aligned} &Q(e_1,e_1) = 0, \quad Q(e_1,e_2) = 3, \quad Q(e_1,e_3) = 2, \\ &Q(e_2,e_1) = 6, \quad Q(e_2,e_2) = 2, \quad Q(e_2,e_3) = 2, \\ &Q(e_3,e_1) = 6, \quad Q(e_3,e_2) = 4, \quad Q(e_3,e_3) = 3. \end{aligned}$$

Macierz tej formy ma więc postać

$$[Q]_e = \begin{bmatrix} 0 & 3 & 2 \\ 6 & 2 & 2 \\ 6 & 4 & 3 \end{bmatrix}. \tag{5}$$

Używając współrzędnych (c, b, a) związanych z bazą e, tzn takich, że

$$v(t) = at^2 + bt + c = ae_3(t) + be_2(t) + ce_1(t), \qquad [v]^e = \begin{bmatrix} c \\ b \\ a \end{bmatrix},$$

możemy zapisać

$$Q(v,w) = [c_v \;\; b_v \;\; a_v] \begin{bmatrix} 0 & 3 & 2 \\ 6 & 2 & 2 \\ 6 & 4 & 3 \end{bmatrix} \begin{bmatrix} c_w \\ b_w \\ a_w \end{bmatrix} = [c_v \;\; b_v \;\; a_v] \begin{bmatrix} 3b_w + 2a_w \\ 6c_w + 2b_w + 2a_w \\ 6c_w + 4b_w + 3a_w \end{bmatrix} = \tag{6}$$

$$3c_vb_w + 2c_va_w + 6b_vc_w + 2b_vb_w + 2b_va_w + 6a_vc_w + 4a_vb_w + 3a_va_w.$$

Powyższe wzory (4), (5), (6) przedstawiają *trzy sposoby opisania tej samej formy dwuliniowej*: przy pomocy abstrakcyjnego wzoru (4), przy pomocy macierzy w wybranej bazie (5), przy pomocy wzoru z użyciem współrzędnych (6). ♣

O ogólnych formach dwuliniowych to wszystko. Dalej będziemy zajmować się tylko formami dwuliniowymi symetrycznymi, ponieważ takie właśnie formy pojawiają się w rachunku różniczkowym funkcji wielu zmiennych jako drugie pochodne funkcji.

Formy dwuliniowe symetryczne i formy kwadratowe. W przykładach (2), (4) widzieliśmy, że formy dwuliniowej używamy często wstawiając ten sam wektor w oba miejsca przeznaczone na argument. Odwzorowanie, które wówczas dostajemy:

$$V \ni v \longmapsto q(v) = Q(v,v) \tag{7}$$

nazywane jest *formą kwadratową*. Nazwa bierze się stąd, że jeśli wektor pomnożymy przez stałą λ, wartość formy pomnoży się przez λ^2:

$$q(\lambda v) = Q(\lambda v, \lambda v) = \lambda^2 Q(v,v) = \lambda^2 q(v).$$

Forma kwadratowa związana z iloczynem skalarnym to kwadrat długości wektora:

$$(v|v) = \|v\|^2.$$

Forma dwuliniowa symetryczna definiuje formę kwadratową wzorem (7). Okazuje się, że jeśli znamy jedynie formę kwadratową, możemy odtworzyć formę dwuliniową od której pochodzi. Istotnie, obliczmy

$$q(v+w) = Q(v+w, v+w) = Q(v,v) + Q(v,w) + Q(w,v) + Q(w,w) = \\ Q(v,v) + 2Q(v,w) + Q(w,w) = q(v) + q(w) + 2Q(v,w).$$

Możemy zatem wyznaczyć $Q(v,w)$ używając jedynie formy kwadratowej q:

$$Q(v,w) = \frac{1}{2}\left(q(v+w) - q(v) - q(w)\right). \tag{8}$$

Wzór (8) nosi nazwę formuły polaryzacyjnej. Formy kwadratowe i formy dwuliniowe symetryczne są we wzajemnie jednoznacznej odpowiedniości. Macierzą formy kwadratowej nazywamy macierz formy dwuliniowej od której forma pochodzi. Przyjrzyjmy się jak wygląda forma kwadratowa zapisana we współrzędnych. Jeśli forma symetryczna Q i wektor v w bazie e mają postać

$$[Q]_e = \begin{bmatrix} Q_{11} & Q_{12} & \cdots & Q_{1n} \\ Q_{21} & Q_{22} & \cdots & Q_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ Q_{n1} & Q_{n2} & \cdots & Q_{nn} \end{bmatrix}, \qquad [v]^e = \begin{bmatrix} \lambda^1 \\ \lambda^2 \\ \vdots \\ \lambda^n \end{bmatrix},$$

to we współrzędnych

$$q(v) = [\lambda^1 \ \ \lambda^2 \ \ \ldots \ \ \lambda^n] \begin{bmatrix} Q_{11} & Q_{12} & \cdots & Q_{1n} \\ Q_{21} & Q_{22} & \cdots & Q_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ Q_{n1} & Q_{n2} & \cdots & Q_{nn} \end{bmatrix} \begin{bmatrix} \lambda^1 \\ \lambda^2 \\ \vdots \\ \lambda^n \end{bmatrix} = \\ \sum_{i,j=1}^{n} \lambda^i \lambda^j Q_{ij} = \sum_{i=1}^{n} Q_{ii} (\lambda^i)^2 + \sum_{i<j} 2Q_{ij} \lambda^i \lambda^j.$$

W ostatniej równości wykorzystaliśmy symetrię macierzy $[Q]_e$, co spowodowało pojawienie się czynników 2.

Przykład 7. Forma kwadratowa odpowiadająca tensorowi bezwładności z przykładu (4) we współrzędnych (x^1, x^2, x^3) związanych z bazą kanoniczną ma postać:

$$\imath(x^1, x^2, x^3) = m[x^1 \ \ x^2 \ \ x^2 \ \ ] \begin{bmatrix} 8 & -2 & -2 \\ -2 & 8 & -2 \\ -2 & -2 & 8 \end{bmatrix} \begin{bmatrix} x^1 \\ x^2 \\ x^3 \end{bmatrix} = \\ m\left(8(x^1)^2 + 8(x^2)^2 + 8(x^3)^2 - 4x^1x^2 - 4x^2x^3 - 4x^1x^3\right).$$

♣

Diagonalizacja. Niech Q będzie formą dwuliniową symetryczną na przestrzeni V. Mówimy, że baza e *diagonalizuje* formę Q jeśli macierz $[Q]_e$ jest macierzą diagonalną.

Twierdzenie 1 (Lagrange)**.** *Każda forma dwuliniowa symetryczna ma bazę diagonalizującą.*

Szkic dowodu: Dowód twierdzenia (1) jest konstruktywny, to znaczy polega na skonstruowaniu przykładowej bazy diagonalizującej. W tym kontekście wygodniej jest pracować z formą kwadratową zamiast z formą dwuliniową. Zacznijmy od konkretnego przykładu - rozważmy formę $\imath$ z przykładu (7). (dla wygody przyjmijmy $m=1$):

$$\imath(x^1,x^2,x^3)=8(x^1)^2+8(x^2)^2+8(x^3)^2-4x^1x^2-4x^2x^3-4x^1x^3.$$

Spróbujemy doprowadzić formę $\imath$ do takiej postaci, aby we wzorze występowały tylko wyrazy kwadratowe, bez mieszanych:

$$\begin{aligned}\imath(x^1,x^2,x^3)&=8(x^1)^2+8(x^2)^2+8(x^3)^2-4x^1x^2-4x^2x^3-4x^1x^3=\\&8(x^1)^2-4x^1(x^2+x^3)+8(x^2)^2+8(x^3)^2-4x^2x^3=\\&8\left[(x^1)^2-\frac{1}{2}x^1(x^2+x^3)\right]+8(x^2)^2+8(x^3)^2-4x^2x^3\end{aligned}$$

Wyrażenie w nawiasie kwadatowym traktujemy jak początek rozwinięcia pełnego kwadratu:

$$\begin{aligned}&8\left[(x^1-\frac{1}{4}x^2-\frac{1}{4}x^3)^2-\frac{1}{16}(x^2+x^3)^2\right]+8(x^2)^2+8(x^3)^2-4x^2x^3=\\&8(x^1-\frac{1}{4}x^2-\frac{1}{4}x^3)^2-\frac{1}{2}(x^2+x^3)^2+8(x^2)^2+8(x^3)^2-4x^2x^3=\\&8(x^1-\frac{1}{4}x^2-\frac{1}{4}x^3)^2+\frac{1}{2}\left(-(x^2+x^3)^2+16(x^2)^2+16(x^3)^2-8x^2x^3\right)\end{aligned}$$

Zauważmy, że część zaznaczona na czerwono nie zawiera w ogóle współrzędnej x^1. Po uporządkowaniu możemy „poddać ją” takiej samej operacji sprowadzania do pełnych kwadratów. Dla ułatwienia rachunków zajmiemy się na razie tylko częścią czerwoną:

$$\begin{aligned}&-(x^2+x^3)^2+16(x^2)^2+16(x^3)^2-8x^2x^3=\\&-(x^2)^2-2x^2x^3-(x^3)^2+16(x^2)^2+16(x^3)^2-8x^2x^3=\\&15(x^2)^2+15(x^3)^2-10x^2x^3=15\left[(x^2)^2-\frac{2}{3}x^2x^3\right]+15(x^3)^2=\\&15\left[(x^2+\frac{1}{5}x^3)^2-\frac{1}{9}(x^3)^2\right]+15(x^3)^2=\\&15(x^2-\frac{1}{3}x^3)^2-\frac{5}{3}(x^3)^2+15(x^3)^2=15(x^2-\frac{1}{3}x^3)^2-\frac{40}{3}(x^3)^2\end{aligned}$$

Otrzymany wynik wstawiamy do wyrażenia na $\imath$:

$$\imath(x^1,x^2,x^3)=8(x^1-\frac{1}{4}x^2-\frac{1}{4}x^3)^2+\frac{15}{2}(x^2-\frac{1}{3}x^3)^2-\frac{20}{3}(x^3)^2.$$

Nowe współrzędne oznaczamy literą α:

$$\begin{aligned}\alpha^1&=x^1-\frac{1}{4}x^2-\frac{1}{4}x^3,\\\alpha^2&=x^2-\frac{1}{3}x^3,\\\alpha^3&=x^3.\end{aligned}\tag{9}$$

Jakiej bazie odpowiadają te współrzędne? Na razie jeszcze nie wiemy. Jeśli jednak tę nieznaną bazę oznaczymy literą f, to wiemy, że

$$[I]_f = \begin{bmatrix} 8 & 0 & 0 \\ 0 & \frac{15}{2} & 0 \\ 0 & 0 & \frac{20}{3} \end{bmatrix}.$$

Jak znaleźć wektory bazowe mając odpowiadające im współrzędne? Potrzebna nam będzie zależność odwrotna do (9):

$$\begin{aligned} x^1 &= \alpha^1 + \frac{1}{4}\alpha^2 + \frac{1}{3}\alpha^3, \\ x^2 &= \alpha^2 + \frac{1}{3}\alpha^3, \\ x^3 &= \alpha^3. \end{aligned} \tag{10}$$

Wektor v, który w bazie e zapisuje się jako

$$v = x^1 e_1 + x^2 e_2 + x^3 e_3$$

można teraz łatwo zapisać w nowej bazie, której odpowiadają współrzędne α:

$$v = x^1 e_1 + x^2 e_2 + x^3 e_3 = \left(\alpha^1 + \frac{1}{4}\alpha^2 + \frac{1}{3}\alpha^3\right) e_1 + \left(\alpha^2 + \frac{1}{3}\alpha^3\right) e_2 + \alpha^3 e_3 =$$
$$\alpha_1 e_1 + \alpha^2 \left(\frac{1}{4}e_1 + e_2\right) + \alpha^3 \left(\frac{1}{3}e_1 - \frac{1}{3}e_2 + e_3\right).$$

Wektory bazy f to wektory

$$\begin{aligned} f^1 &= e_1 \\ f^2 &= \frac{1}{4}e_1 + e_2, \\ f^3 &= \frac{1}{3}e_1 - \frac{1}{3}e_2 + e_3. \end{aligned} \tag{11}$$

Znaleźliśmy jakąś bazę diagonalizującą dla formy I (oraz $\imath$). Oczywiście taka baza nie jest jedyna. Nie będziemy podawać ogólnego przepisu na diagonalizację metodą Lagranża. Z przeprowadzonych dopiero co rachunków na konkretnym przykładzie wynika także algorytm dla dowolnej formy. Pewien problem mogą nam sprawić tylko formy w których nie ma żadnych wyrazów kwadratowych - nie ma więc od czego zacząć algorytmu. W takim przypadku należy wybrać wyraz mieszany $x^i x^j$ i podstawić

$$x^i = \alpha + \beta, \qquad x^j = \alpha - \beta,$$

wtedy

$$x^i x^j = (\alpha + \beta)(\alpha - \beta) = \alpha^2 - \beta^2$$

i już mamy potrzebne wyrażenia kwadratowe. $\Box$

Wspomnieliśmy już, że baza diagonalizująca nie jest wyznaczona jednoznacznie. Współczynniki, które pojawiają się na diagonali macierzy formy też mogą być bardzo różne (zależą oczywiście od bazy). Okazuje się jednak, że jest coś co od wyboru bazy diagonalizującej nie zależy. Mówi o tym następujące twierdzenie

Twierdzenie 2 (Sylvestera o bezwładności form)**.** *Jeśli* $(\alpha^1, \ldots, \alpha^n)$ *i* $(\beta^1 \ldots, \beta^n)$ *są układami współrzędnych odpowiadających dwum bazom diagonalizującym formę kwadratową* q *i jeśli*

$$q(\alpha^1, \ldots, \alpha^n) = \sum_{k=1}^{n} a_i(\alpha^i)^2, \qquad q(\beta^1, \ldots, \beta^n) = \sum_{k=1}^{n} b_i(\beta^i)^2$$

to w zbiorach

$$A = \{a_1, \alpha_2, \ldots a_n\}, \qquad B = \{b_1, b_2, \ldots b_n\}$$

jest tyle samo współczynników dodatnich, tyle samo ujemnych i tyle samo zerowych.

Dowód: Zauważmy na początku, że wartość bezwzględną współczynników a_i można „wciągnąć” do definicji współrzędnych (i oczywiście tym samym bazy diagonalizującej). Możemy więc założyć, że zbiory A i B składają się z plus i minus jedynek. Współrzędne można też uporządkować tak, aby pierwsze były te ze współczynnikami $+1$, dalej -1 a następnie te, które mają współczynnik zero. Jeśli współrzędne w układzie α wektora v oznaczymy $\alpha^i(v)$, podobnie ze współrzędnymi β, to forma kwadratowa q we współrzędnych odpowiadających obu bazom ma postać:

$$q(v) = \sum_{i=1}^{p_\alpha}(\alpha^i(v))^2 - \sum_{i=1}^{r_\alpha}(\alpha^{p_\alpha+i}(v))^2 = \sum_{i=1}^{p_\beta}(\beta^i(v))^2 - \sum_{i=1}^{r_\beta}(\beta^{p_\beta+i}(v))^2.$$

Naszym zadaniem jest pokazać, że $p_\alpha = p_\beta$ i $r_\alpha = r_\beta$. Jasne jest, że $p_\alpha + r_\alpha = p_\beta + r_\beta$. Jeśli jądrem q nazwiemy podprzestrzeń zdefiniowaną warunkiem

$$\ker q = \ker Q = \{v \in V : \ \ \forall w \in V \ Q(v, w) = 0\},$$

to

$$p_\alpha + r_\alpha = p_\beta + r_\beta = n - \dim\ker q.$$

Załóżmy teraz, że $p_\beta < p_\alpha$, tzn.

$$p_\beta - p_\alpha < 0 \quad \Longrightarrow \quad n + p_\beta - p_\alpha < n$$

Warunki

$$\beta^i(v) = 0 \ \text{ dla } \ i = 1, \ldots, p_\beta, \qquad \alpha^j(v) = 0 \ \text{ dla } \ j = p_\alpha + 1, \ldots, n \tag{12}$$

zapisane w jednym z układów współrzędnych (α lub β) przyjmą postać układu równań liniowych, który składa się z mniej niż n równań. Ma więc niezerowe rozwiązanie (układ równań liniowych, jednorodny, liczba równań mniejsza od wymiaru przestrzeni). Oznaczmy to rozwiązanie v_0. Wartość formy na wektorze v_0 w układzie współrzędnych α to:

$$q(v) = \sum_{i=1}^{p_\alpha}(\alpha^i(v_0))^2 - \sum_{i=1}^{r_\alpha}(\alpha^{p_\alpha+i}(v_0))^2 = \sum_{i=1}^{p_\alpha}(\alpha^i(v_0))^2 \geqslant 0 \tag{13}$$

A wartość formy na wektorze v_0 w układzie współrzędnych α to:

$$q(v) = \sum_{i=1}^{p_\beta}(\beta^i(v))^2 - \sum_{i=1}^{r_\beta}(\beta^{p_\beta+i}(v))^2 = -\sum_{i=1}^{r_\beta}(\beta^{p_\beta+i}(v))^2 \leqslant 0 \tag{14}$$

Ze wzorów (13) i (14) wynika, że

$$q(v_0) = 0, \tag{15}$$

Równanie (15) w połączeniu z (13) oznacza, że

$$\sum_{i=1}^{p_\alpha}(\alpha^i(v_0))^2 = 0 \quad \text{czyli} \quad \alpha^i(v_0) = 0 \quad \text{dla} \quad i = 1 \ldots p_\alpha.$$

Pamiętamy jednak (układ równań (15)), że także dla pozostałych indeksów $\alpha^i(v_0) = 0$. Wygląda więc na to, że wszystkie współrzędne wektora v_0 w układzie α są równe zero. To jednak oznacza, że $v_0 = 0$. Doprowadziliśmy do sprzeczności. Okazuje się, że nie może być $p_\beta < p_\alpha$. Ponieważ żaden z układów współrzędnychnie jest wyróżniony, nie może być także $p_\beta > p_\alpha$. Pozostaje zatem $p_\beta = p_\alpha$, a to już gwarantuje, że $r_\beta = r_\alpha$. $\square$

Jeśli liczbę współczynników dodatnich w zbiorze A oznaczymy p a ujemnych r, to parę (p, q) nazywamy *sygnaturą* formy Q (lub q) a liczbę $r = p + q$ *rzędem* formy. Znajdowanie sygnatury formy jest ważną umiejętnością potrzebną przy badaniu ekstremów funkcji wielu zmiennych.

Istnieje jeszcze jedna metoda znajdowania sygnatury, nie wymagająca diagonalizowania formy. Najpierw jednak trzeba przyzwoicie nauczyć się co to jest wyznacznik!