Rozdział 10

Formy dwuliniowe i kwadratowe

10.1 Formy dwuliniowe

10.1.1 Definicja i przykłady

Niech $\mathcal{X}_{|\mathbf{K}}$ będzie przestrzenią liniową nad ciałem $\mathbf{K}$, $\dim(\mathcal{X}_{|\mathbf{K}}) = n$.

Definicja 10.1 *Przekształcenie* $\varphi : \mathcal{X} \times \mathcal{X} \to \mathbf{K}$ *nazywamy* formą dwuliniową *na przestrzeni* $\mathcal{X}_{|\mathbf{K}}$ *jeśli*

(i) $\forall x, y_1, y_2 \in \mathcal{X}$, $\forall \alpha_1, \alpha_2 \in \mathbf{K}$

$$\varphi(x, y_1 * \alpha_1 + y_2 * \alpha_2) = \varphi(x, y_1) * \alpha_1 + \varphi(x, y_2) * \alpha_2$$

(liniowość ze względu na drugą zmienną),

(ii) $\forall x, y \in \mathcal{X} \quad \varphi(x, y) = \varphi(y, x)$ *(forma zwykła)*
albo
$\forall x, y \in \mathcal{X} \quad \varphi(x, y) = \overline{\varphi(y, x)}$ *(forma* hermitowska*).*

Oczywiście, o formach hermitowskich możemy mówić tylko wtedy gdy $\mathbf{K} \subseteq \mathbf{C}$. Dalej, dla uproszczenia, będziemy rozpatrywać jedynie formy hermitowskie.

Zauważmy, że $\forall x_1, x_2, y \in \mathcal{X}$, $\forall \beta_1, \beta_2 \in \mathbf{K}$,

$$\begin{aligned}\varphi(x_1 * \beta_1 + x_2 * \beta_2, y) &= \overline{\varphi(y, x_1 * \beta_1 + x_2 * \beta_2)} \\ &= \overline{\varphi(y, x_1)} * \overline{\beta}_1 + \overline{\varphi(y, x_2)} * \overline{\beta}_2 \\ &= \varphi(x_1, y) * \overline{\beta}_1 + \varphi(x_2, y) * \overline{\beta}_2.\end{aligned}$$

Dość oczywistym jest fakt, że zbiór wszystkich form dwuliniowych na $\mathcal{X}_{|\mathbf{K}}$ jest przestrzenią liniową nad $\mathbf{R}$ (ale nie nad $\mathbf{C}$!) z naturalnymi działaniami:

$$\begin{aligned}(\alpha * \varphi)(x, y) &:= \alpha * \varphi(x, y), \\ (\varphi_1 + \varphi_2)(x, y) &:= \varphi_1(x, y) + \varphi_2(x, y).\end{aligned}$$

Przykładami form dwuliniowych na $\mathcal{X}_{|\mathbf{K}} = \mathbf{K}^n_{|\mathbf{K}}$ ($\mathbf{K} \subseteq \mathbf{C}$) są:

$$\begin{aligned}\varphi(\vec{x}, \vec{y}) &= \sum_{i=1}^{n} \overline{x}_i * y_i * \rho_i, \quad \text{gdzie } \rho_i \in \mathbf{R},\ 1 \le i \le n, \\ \varphi(\vec{x}, \vec{y}) &= \vec{x}^H * A * \vec{y}, \quad \text{gdzie } A \in \mathbf{K}^{n,n},\ A = A^H,\end{aligned}$$

a na $\mathcal{P}^n_{|\mathbf{R}}$:

$$\begin{aligned}\varphi(p, q) &= \sum_{i=1}^{n} \overline{p^{(i)}(t_i)} \cdot q^{(i)}(t_i) \cdot \rho_i, \qquad \rho_i \in \mathbf{R},\ 1 \le i \le n, \\ \varphi(p, q) &= \int_0^1 \overline{p(t)} \cdot q(t) \cdot \rho(t)\, dt, \qquad \rho : \mathbf{R} \to \mathbf{R}.\end{aligned}$$

10.1.2 Macierz formy dwuliniowej

Dalej wygodnie nam będzie rozszerzyć działanie danej formy dwuliniowej $\varphi : \mathcal{X} \times \mathcal{X} \to \mathbf{K}$ na $\varphi : \mathcal{X}^{1,s} \times \mathcal{X}^{1,t} \to \mathbf{K}^{s,t}$ w następujący sposób. Niech $\mathbb{A} = [x_1, \ldots, x_s]$ i $\mathbb{B} = [y_1, \ldots, y_t]$. Wtedy

$$\varphi(\mathbb{A}, \mathbb{B}) := \left(\varphi(x_i, y_j)\right)_{i,j} \in \mathbf{K}^{s,t}.$$

W szczególności, macierz $\varphi(\mathbb{A}, \mathbb{A}) = (\varphi(x_i, x_j))_{i,j}$ jest kwadratowa i hermitowska, $\varphi(\mathbb{A}, \mathbb{A}) \in \mathrm{Herm}^{n,n}$. Mamy też

$$\begin{aligned}\forall \varphi\ \forall \alpha \in \mathbf{R} \qquad (\alpha * \varphi)(\mathbb{A}, \mathbb{B}) &= \alpha * \varphi(\mathbb{A}, \mathbb{B}), \\ \forall \varphi, \psi \qquad (\varphi + \psi)(\mathbb{A}, \mathbb{B}) &= \varphi(\mathbb{A}, \mathbb{B}) + \psi(\mathbb{A}, \mathbb{B}).\end{aligned}$$

Pożyteczne będą też następujące wzory rachunkowe:

$$\begin{array}{lrcl} \forall \vec{b} \in \mathbf{K}^t & \varphi(\mathbb{A}, \mathbb{B} * \vec{b}) & = & \varphi(\mathbb{A}, \mathbb{B}) * \vec{b}, \\ \forall \vec{a} \in \mathbf{K}^s & \varphi(\mathbb{A} * \vec{a}, \mathbb{B}) & = & \vec{a}^H * \varphi(\mathbb{A}, \mathbb{B}). \end{array}$$

Rzeczywiście,

$$\varphi(\mathbb{A}, \mathbb{B} * \vec{b}) \;=\; \varphi\Big(\mathbb{A}, \sum_{j=1}^t y_j * \beta_j\Big) \;=\; \sum_{j=1}^t \varphi(\mathbb{A}, y_j) * \beta_j \;=\; \varphi(\mathbb{A}, \mathbb{B}) * \vec{b},$$

gdzie $\vec{b} = [\beta_1, \ldots, \beta_t]^T$, oraz

$$\varphi(\mathbb{A} * \vec{a}, \mathbb{B}) \;=\; (\varphi(\mathbb{B}, \mathbb{A} * \vec{a}))^H \;=\; \vec{a}^H * (\varphi(\mathbb{B}, \mathbb{A}))^H \;=\; \vec{a}^H * \varphi(\mathbb{A}, \mathbb{B}).$$

Uogólniając te wzory mamy

$$\begin{array}{lrcl} \forall B \in \mathbf{K}^{t,r} & \varphi(\mathbb{A}, \mathbb{B} * B) & = & \varphi(\mathbb{A}, \mathbb{B}) * B, \\ \forall A \in \mathbf{K}^{s,r} & \varphi(\mathbb{A} * A, \mathbb{B}) & = & A^H * \varphi(\mathbb{A}, \mathbb{B}). \end{array}$$

Mamy bowiem

$$\begin{array}{rcl} \varphi(\mathbb{A}, \mathbb{B} * B) & = & \varphi(\mathbb{A}, [\mathbb{B} * \vec{b}_1, \ldots, \mathbb{B} * \vec{b}_r]) \\ & = & [\varphi(\mathbb{A}, \mathbb{B} * \vec{b}_1), \ldots, \varphi(\mathbb{A}, \mathbb{B} * \vec{b}_r)] \\ & = & [\varphi(\mathbb{A}, \mathbb{B}) * \vec{b}_1, \ldots, \varphi(\mathbb{A}, \mathbb{B}) * \vec{b}_r] \\ & = & \varphi(\mathbb{A}, \mathbb{B}) * B, \end{array}$$

gdzie $B = [\vec{b}_1, \ldots, \vec{b}_r]$, oraz

$$\begin{array}{rcl} \varphi(\mathbb{A} * A, \mathbb{B}) & = & (\varphi(\mathbb{B}, \mathbb{A} * A))^H \;=\; (\varphi(\mathbb{B}, \mathbb{A}) * A)^H \\ & = & A^H * (\varphi(\mathbb{B}, \mathbb{A}))^H \;=\; A^H * \varphi(\mathbb{A}, \mathbb{B}). \end{array}$$

Definicja 10.2 *Niech* $\mathbb{A} = [x_1, \ldots, x_n]$ *będzie bazą* $\mathcal{X}$*, a* $\varphi : \mathcal{X} \times \mathcal{X} \to \mathbf{K}$ *formą dwuliniową na* $\mathcal{X}$*. Macierz hermitowską*

$$\Phi_{\mathbb{A}} \;:=\; \varphi(\mathbb{A}, \mathbb{A}) \;=\; (\varphi(x_i, x_j))_{i,j=1}^n$$

nazywamy macierzą formy φ *w bazie* $\mathbb{A}$.

Znaczenie macierzy formy wynika z następującej równości. Niech $x = \mathbb{A} * \vec{a}$ i $y = \mathbb{A} * \vec{b}$. Wtedy

$$\begin{aligned} \varphi(x, y) &= \varphi(\mathbb{A} * \vec{a}, \mathbb{A} * \vec{b}) = \vec{a}^H * \varphi(\mathbb{A}, \mathbb{A}) * \vec{b} \\ &= \vec{a}^H * \Phi_{\mathbb{A}} * \vec{b} = (\mathbb{A}^{-1} \cdot x)^H * \Phi_{\mathbb{A}} * (\mathbb{A}^{-1} \cdot y). \end{aligned}$$

Przy ustalonej bazie $\mathbb{A}$, każdej formie hermitowskiej $\varphi : \mathcal{X} \times \mathcal{X} \to \mathbf{K}$ można przyporządkować jej macierz $\Phi_{\mathbb{A}} = \varphi(\mathbb{A}, \mathbb{A})$, która jest hermitowska. Ale też odwrotnie, każda macierz hermitowska Φ definiuje formę hermitowską zgodnie ze wzorem $\varphi(x, y) = (\mathbb{A}^{-1} \cdot x)^H * \Phi * (\mathbb{A}^{-1} \cdot y)$. Mamy przy tym, że jeśli $\gamma = \varphi + \psi$ to $\Gamma_{\mathbb{A}} = \Phi_{\mathbb{A}} + \Psi_{\mathbb{A}}$ oraz jeśli $\gamma = \alpha * \varphi$, $\alpha \in \mathbf{R}$, to $\Gamma_{\mathbb{A}} = \alpha * \Phi_{\mathbb{A}}$. Stąd przestrzeń wszystkich form hermitowskich nad $\mathbf{R}$ jest izomorficzna z przestrzenią macierzy hermitowskich nad $\mathbf{R}$. W przypadku $\mathbf{K} = \mathbf{C}$ jej wymiar wynosi n^2.

10.2 Twierdzenie Sylwester'a

Definicja 10.3 *Powiemy, że macierz* $A \in \mathbf{K}^{n,n}$ przystaje *do macierzy* $B \in \mathbf{K}^{n,n}$ *gdy istnieje macierz nieosobliwa* $C \in \mathbf{K}^{n,n}$ *taka, że*

$$B = C^H * A * C.$$

Niech $\mathbb{A}$ i $\mathbb{B}$ będą dwiema bazami $\mathcal{X}_{|\mathbf{K}}$. Niech $C = \mathbb{A}^{-1} \cdot \mathbb{B} \in \mathbf{K}^{n,n}$ będzie macierzą zmiany bazy z $\mathbb{A}$ na $\mathbb{B}$ tak, że

$$\mathbb{B} = \mathbb{A} * C.$$

Jeśli $\Phi_{\mathbb{A}}$ jest macierzą danej formy $\varphi : \mathcal{X} \times \mathcal{X} \to \mathbf{K}$ w bazie $\mathbb{A}$ to macierz φ w bazie $\mathbb{B}$ można wyrazić wzorem

$$\begin{aligned} \Phi_{\mathbb{B}} &= \varphi(\mathbb{B}, \mathbb{B}) = \varphi(\mathbb{A} * C, \mathbb{A} * C) \\ &= C^H * \varphi(\mathbb{A}, \mathbb{A}) * C = C^H * \Phi_{\mathbb{A}} * C. \end{aligned}$$

Stąd, w klasie macierzy hermitowskich $\mathrm{Herm}^{n,n}$ macierz A przystaje do B gdy obie są macierzami tej samej formy (ale być może w różnych bazach).

Relacja przystawania macierzy jest zwrotna (bo $A = I^H * A * I$), symetryczna (bo jeśli $B = C^H * A * C$ to $A = (C^{-1})^H * B * C^{-1}$) oraz przechodnia (bo jeśli $A_2 = C_1^H * A_1 * C_1$ i $A_3 = C_2^H * A_2 * C_2$ to $A_3 = (C_1 * C_2)^H * A_1 * (C_1 * C_2)$).

Jest to więc relacja równoważności. A jeśli tak, to zbiór wszystkich macierzy hermitowskich można przedstawić jako rozłączną sumę macierzy do siebie wzajemnie przystających (klas abstrakcji relacji przystawania, albo jeszcze inaczej, macierzy tej samej formy, ale w różnych bazach).

Ile jest klas abstrakcji relacji przystawania w klasie macierzy hermitowskich? Odpowiedź daje natępujące twierdzenie, które podajemy bez dowodu.

Twierdzenie 10.1 (Sylwester'a)
Dla dowolnej macierzy hermitowskiej $A = A^H \in \mathbf{K}^{n,n}$ istnieje macierz nieosobliwa $C \in \mathbf{K}^{n,n}$ taka, że

$$C^H * A * C = \operatorname{diag}(I_\pi, -I_\nu, 0_\xi),$$

gdzie wymiary π, ν, ξ $(\pi + \nu + \xi = n)$ są wyznaczone jednoznacznie.

Stąd klas abstrakcji relacji przystawania jest tyle ile macierzy diagonalnych z elementami na diagonali kolejno $1, -1, 0$, czyli

$$\sum_{k=0}^{n}(k+1) = \frac{(n+1)(n+2)}{2}.$$

Z twierdzenia Sylwester'a wynika również następujący ważny wniosek.

Wniosek 10.1 *Dla dowolnej formy dwuliniowej $\varphi : \mathcal{X} \times \mathcal{X} \to \mathbf{K}$ istnieje baza $\mathbb{A}$ w $\mathcal{X}$, w której forma ma postać*

$$\varphi(x, y) = \sum_{k=1}^{\pi} \overline{a}_k * b_k - \sum_{k=\pi+1}^{\pi+\nu} \overline{a}_k * b_k,$$

gdzie $x = \mathbb{A} * \vec{a}$, $y = \mathbb{A} * \vec{b}$.

10.3 Formy kwadratowe

10.3.1 Określoność formy kwadratowej

Każdej formie dwuliniowej $\varphi : \mathcal{X} \times \mathcal{X} \to \mathbf{K}$ odpowiada *forma kwadratowa* $h : \mathcal{X} \to \mathbf{R}$ zdefiniowana wzorem

$$h(x) = \varphi(x, x) \qquad x \in \mathcal{X}.$$

Jeśli dla wszystkich $x \neq \mathbf{0}$ mamy $h(x) = \varphi(x, x) > 0$ to formę kwadratową h (i odpowiednio formę dwuliniową φ) nazywamy *dodatnio określoną* i piszemy $h > 0$ (odpowiednio $\varphi > 0$). Podobnie, forma h jest określona

- ujemnie, gdy $h(x) < 0 \ \forall x \neq 0 \quad (h < 0)$,
- niedodatnio, gdy $h(x) \leq 0 \ \forall x \quad (h \leq 0)$,
- nieujemnie, gdy $h(x) \geq 0 \ \forall x \quad (h \geq 0)$.

We wszystkich pozostałych przypadkach forma jest *nieokreślona.*

Z równości

$$h(x) = \vec{a}^H * \Phi_{\mathbb{A}} * \vec{a} \qquad (x = \mathbb{A} * \vec{a})$$

wynika, że określoność formy jest taka sama jak określoność jej macierzy (w dowolnej bazie!). W szczególności, stosując notację z twierdzenia Sylwester'a mamy:

$$h > 0 \iff \pi = n, \qquad h \geq 0 \iff \nu = 0,$$
$$h < 0 \iff \nu = n, \qquad h \leq 0 \iff \pi = 0.$$

10.3.2 Kryterium Sylwester'a

Twierdzenie 10.2 *Niech* $A = A^H = (a_{i,j})_{i,j=1}^n \in \mathrm{Herm}^{n,n}$ *oraz* $A^{(k)} = (a_{i,j})_{i,j=1}^k$, $1 \leq k \leq n$, *będą odpowiednimi macierzami kątowymi. Wtedy*

(i) A jest dodatnio określona $\iff \det(A^{(k)}) > 0$ *dla* $1 \leq k \leq n$,

(ii) A jest ujemnie określona $\iff (-1)^k \cdot \det(A^{(k)}) > 0$ *dla* $1 \leq k \leq n$.

Dowód. Przypomnijmy (twierdzenie 7.5), że dla macierzy o nieosobliwych macierzach kątowych (a takimi są macierze dodatnio/ujemnie określone) można przeprowadzić eliminację Gaussa bez przestawień wierszy/kolumn. Dlatego A można przedstawić jako

$$A = L * R = L * D * L^H,$$

gdzie $L \in \mathrm{TRIL}^{n,n}$, $l_{i,i} = 1 \ \forall i$, $D = \mathrm{diag}(r_{1,1}, \ldots, r_{n,n})$. Podstawiając $\vec{y} := L^H * \vec{x}$, mamy

$$\begin{aligned} \vec{x}^H * A * \vec{x} &= \vec{x}^H * L * D * L^H * \vec{x} = (L^H * \vec{x})^H * D * (L^H * \vec{x}) \\ &= \vec{y}^H * D * \vec{y} = \sum_{i=1}^n |y_i|^2 \cdot r_{i,i}. \end{aligned}$$

Stąd $A > 0$ wtedy i tylko wtedy gdy $r_{i,i} > 0\ \forall i$, oraz $A < 0$ wtedy i tylko wtedy gdy $r_{i,i} < 0\ \forall i$.

Dowód uzupełnia spostrzeżenie, że

$$A^{(k)} = L^{(k)} * R^{(k)} \ = \ L^{(k)} * D^{(k)} * (L^{(k)})^H$$

oraz

$$\det(A^{(k)}) = |\det(L^{(k)})|^2 \cdot \det(D^{(k)}) = \prod_{i=1}^{k} r_{i,i}.$$