Fizyka dla firm — Matematyka

26. Iloczyny skalarne, normy i metryki

P. F. Góra

http://th-www.if.uj.edu.pl/zfs/gora/

24 marca 2021

Iloczyn skalarny ☺

Niech $\mathcal{V}$ będzie przestrzenią liniową nad ciałem liczb zespolonych, $\mathbb{C}$. Iloczynem skalarnym nazywam odwzorowanie $\mathcal{V} \times \mathcal{V} \to \mathbb{C}$ spełniające ($\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathcal{V}$, $\alpha, \beta \in \mathbb{C}$, $(\cdot)^*$ oznacza sprzężenie zespolone):

$$\mathbf{x} \circ \mathbf{y} = (\mathbf{y} \circ \mathbf{x})^* \tag{1a}$$

$$\mathbf{x} \circ (\alpha \mathbf{y} + \beta \mathbf{z}) = \alpha\, \mathbf{x} \circ \mathbf{y} + \beta\, \mathbf{x} \circ \mathbf{z} \tag{1b}$$

$$(\alpha \mathbf{x} + \beta \mathbf{y}) \circ \mathbf{z} = \alpha^*\, \mathbf{x} \circ \mathbf{z} + \beta^*\, \mathbf{y} \circ \mathbf{z} \tag{1c}$$

$$\mathbf{x} \circ \mathbf{x} \geqslant 0\,; \quad \mathbf{x} \circ \mathbf{x} = 0 \Leftrightarrow \mathbf{x} = 0 \tag{1d}$$

Ostatni warunek wymaga pewnego komentarza: W ogólności iloczyn skalarny dwóch wektorów zespolonych jest zespolony, ale *kwadrat skalarny* jednego wektora jest rzeczywisty i nieujemny. Tylko kwadrat skalarny wektora zerowego równa się zero.

Przestrzeń $\mathbb{R}^n$

Przestrzeń $\mathbb{R}^n$ jest szczególnym przypadkiem przestrzeni $\mathcal{V}$, o jakiej mowa była na poprzednim slajdzie. Jest to przestrzeń nad ciałem liczb rzeczywistych, więc wektory i skalary są rzeczywiste, a sprzężenie zespolone jest operacją identycznościową. W tej przestrzeni iloczynem skalarnym jest

$$\mathbf{x} \circ \mathbf{y} = \mathbf{x}^T \mathbf{y} = \sum_{i=1}^{n} x_i y_i \tag{2}$$

Innym, ale równie dobrze zdefiniowanym iloczynem skalarnym jest

$$\mathbf{x} \circ \mathbf{y} = \mathbf{x}^T \mathbf{A} \mathbf{y} \tag{3}$$

gdzie $\mathbf{A}$ jest macierzą symetryczną, rzeczywistą i *dodatnio określoną*: $\forall \mathbf{x} \neq 0 : \mathbf{x}^T \mathbf{A} \mathbf{x} > 0$.

Przestrzeń $\mathbb{C}^n$

Przestrzeń $\mathbb{C}^n$ jest innym przykładem przestrzeni $\mathcal{V}$. W tej przestrzeni iloczynem skalarnym jest

$$\mathbf{x} \circ \mathbf{y} = \mathbf{x}^\dagger \mathbf{y} = \sum_{i=1}^{n} x_i^* y_i \,. \tag{4}$$

Innym, ale również dobrze zdefiniowanym iloczynem skalarnym jest

$$\mathbf{x} \circ \mathbf{y} = \mathbf{x}^\dagger \mathbf{A} \mathbf{y} \tag{5}$$

gdzie $\mathbf{A}$ jest macierzą hermitowską, $\mathbf{A}^\dagger = \mathbf{A}$, i *dodatnio określoną*: $\forall \mathbf{x} \neq 0 : \mathbf{x}^\dagger \mathbf{A} \mathbf{x} > 0$.

Ponieważ istnieje nieskończenie wiele różnych macierzy hermitowskich $\mathbb{C}^{n\times n}$ dodatnio określonych, w $\mathbb{C}^n$ można, formalnie rzecz biorąc, zdefiniować nieskończenie wiele iloczynów skalarnych. *Niektóre* z nich są użyteczne.

Przestrzeń unormowana

Niech $\mathcal{V}$ będzie przestrzenią liniową nad ciałem liczb zespolonych, $\mathbb{C}$. Normą wektora nazywam odwzorowanie $\mathcal{V} \to \mathbb{R}$ spełniające
Nieujemność:

$$\forall \mathbf{x} \in \mathcal{V} : \|\mathbf{x}\| \geqslant 0 \tag{6a}$$

Norma wektorów niezerowych jest dodatnia:

$$\|\mathbf{x}\| = 0 \Leftrightarrow \mathbf{x} = 0 \tag{6b}$$

Mnożenie przez skalar:

$$\forall \mathbf{x} \in \mathcal{V}\,,\ \alpha \in \mathbb{C} : \|\alpha \cdot \mathbf{x}\| = |\alpha| \cdot \|\mathbf{x}\| \tag{6c}$$

Nierówność trójkąta:

$$\forall \mathbf{x}, \mathbf{y} \in \mathcal{V} : \|\mathbf{x} + \mathbf{y}\| \leqslant \|\mathbf{x}\| + \|\mathbf{y}\| \tag{6d}$$

Przestrzeń liniową (wektorową), na której określono normę, nazywam przestrzenią unormowaną.

Norma zadana przez iloczyn skalarny

Twierdzenie: Jeżeli $\mathbf{x} \circ \mathbf{y}$ jest iloczynem skalarnym, wówczas

$$\|\mathbf{x}\| = \sqrt{\mathbf{x} \circ \mathbf{x}} \tag{7}$$

jest normą. Normę tę nazywam normą zadaną przez iloczyn skalarny.

Przykład 1

W $\mathbb{R}^n$ normę zadaje iloczyn skalarny (2)

$$\|\mathbf{x}\|_2 = \sqrt{\sum_{i=1}^{n} x_i^2} \tag{8}$$

Normę tę nazywam normą euklidesową. Jest to najczęściej używana norma, wobec czego zapis $\|\cdot\|$ często oznacza właśnie normę euklidesową, choć *powinno się* pisać $\|\cdot\|_2$.

Przykład 2

Uogólnieniem normy euklidesowej na przestrzeń $\mathbb{C}^n$ jest

$$\|\mathbf{x}\|_2 = \sqrt{\sum_{i=1}^{n} |x_i|^2} \tag{9}$$

Przykład 3

Istnieją normy niezadawane przez iloczyn skalarny. Na przykład w $\mathbb{C}^n$ (i odpowiednio w $\mathbb{R}^n$ istnieją normy

Norma Manhattan albo taksówkowa:

$$\|\mathbf{x}\|_1 = \sum_{i=1}^{n} |x_i| \tag{10a}$$

czy też norma maksimum lub *worst offender*:

$$\|\mathbf{x}\|_\infty = \max_{i=1,\dots,n} |x_i| \tag{10b}$$

W przypadku tych norm na ogół *nie* unika się pisania znaczników dolnych, $\|\cdot\|_1$, $\|\cdot\|_\infty$.

Norma indukowana macierzy

Niech $\mathbf{A} \in \mathbb{R}^{n\times n}$ lub $\mathbf{A} \in \mathbb{C}^{n\times n}$. Macierze takie tworzą przestrzeń liniową nad $\mathbb{R}^n$ lub, odpowiednio, $\mathbb{C}^n$ (macierze są tam “wektorami niegeometrycznymi”).

Niech $\|\cdot\|$ będzie normą wektorów w $\mathbb{R}^n$ lub, odpowiednio, $\mathbb{C}^n$ — najczęściej, choć nie wyłącznie, mamy na myśli normę euklidesową. Poniżej będę pomijał zapis “lub $\mathbb{C}^n$”, gdyż uogólnienie na przypadek zespolony jest oczywiste.

Definicja: Normą macierzy indukowaną przez normę wektorów nazywam wielkość

$$\|\mathbf{A}\| = \max_{\mathbf{x}\neq 0} \frac{\|\mathbf{A}\mathbf{x}\|}{\|\mathbf{x}\|} = \max_{\|\mathbf{x}\|=1} \|\mathbf{A}\mathbf{x}\| \tag{11}$$

Po prawej stronie (11) stoją normy wektorów, a więc wielkości znane. Ponieważ $\forall\, \mathbf{x} \neq 0\ \ \exists \mathbf{e} : \mathbf{x} = \|\mathbf{x}\| \cdot \mathbf{e}\ \wedge\ \|\mathbf{e}\| = 1$

$$\frac{\|\mathbf{A}\mathbf{x}\|}{\|\mathbf{x}\|} = \frac{\|\mathbf{x}\| \cdot \|\mathbf{A}\mathbf{e}\|}{\|\mathbf{x}\| \cdot \|\mathbf{e}\|} = \|\mathbf{A}\mathbf{e}\| \tag{12}$$

i druga równość w (11) staje się oczywista.

Można pokazać, że wyrażenie (11) jest normą w $\mathbb{R}^{n\times n}$.

Można też definiować inne normy w tej przestrzeni, nieindukowane przez normy wektorów, ale mają one mniejsze znaczenie.

Przykład 4

Niech $\mathbf{A} \in \mathbb{R}^{n\times n}$ będzie macierzą symetryczną, rzeczywistą. Obliczmy jej normę indukowaną przez normę euklidesową wektorów.

Niech $\{\mathbf{e}_i\}_{i=1}^{n}$ będą unormowanymi wektorami własnymi macierzy $\mathbf{A}$, a $\{\lambda_i\}$ odpowiadającymi im wartościami własnymi. Wiadomo, że wektory $\{\mathbf{e}_i\}$ stanowią bazę ortonormalną w $\mathbb{R}^n$. Zatem $\forall\, \mathbf{x} \in \mathbb{R}^n$ istnieje rozkład

$$\mathbf{x} = \sum_{i=1}^{n} \alpha_i \mathbf{e}_i \tag{13a}$$

$$\begin{aligned} \|\mathbf{x}\|_2^2 &= \mathbf{x}^T\mathbf{x} = \left(\sum_{i=1}^{n} \alpha_i \mathbf{e}_i^T\right)\left(\sum_{j=1}^{n} \alpha_j \mathbf{e}_j\right) \\ &= \sum_{i,j=1}^{n} \alpha_i\alpha_j\, \mathbf{e}_i^T\mathbf{e}_j = \sum_{i,j=1}^{n} \alpha_i\alpha_j\, \delta_{ij} = \sum_{i=1}^{n} \alpha_i^2 \end{aligned} \tag{13b}$$

A zatem $\|\mathbf{x}\|_2 = 1 \Leftrightarrow \|\mathbf{x}\|_2^2 = 1 \Leftrightarrow \sum\limits_{i=1}^{n} \alpha_i^2 = 1$ gdzie α_i są współczynnikami rozkładu (13a).

Niech $\|\mathbf{x}\|_2 = 1$. Obliczam

$$
\begin{aligned}
\|\mathbf{Ax}\| &= \left\|\mathbf{A}\sum_{i=1}^{n}\alpha_i\mathbf{e}_i\right\|_2 \\
&= \left\|\sum_{i=1}^{n}\alpha_i\mathbf{A}\mathbf{e}_i\right\|_2 = \left\|\sum_{i=1}^{n}\alpha_i\lambda_i\mathbf{e}_i\right\|_2 \\
&= \sqrt{\sum_{i=1}^{n}\alpha_i^2\lambda_i^2} \leqslant \sqrt{\sum_{i=1}^{n}\alpha_i^2 \max_{j=1,\dots,n}\lambda_j^2} \\
&= \sqrt{\left(\max_{j=1,\dots,n}\lambda_j^2\right)\cdot\underbrace{\left(\sum_{i=1}^{n}\alpha_i^2\right)}_{=1}} = \max_{j=1,\dots,n}|\lambda_j| \qquad (13c)
\end{aligned}
$$

A więc

$$\|\mathbf{A}\| = \max_{j=1,\dots,n} |\lambda_j| \tag{13d}$$

Ten wynik można bardzo łatwo uogólnić na przypadek macierzy hermitowskich.

Przykład 5

Wartościami własnymi macierzy

$$\mathbf{B} = \begin{bmatrix} 4 & 1 & 1 \\ 1 & 4 & 1 \\ 1 & 1 & 4 \end{bmatrix} \tag{14}$$

są liczby $\lambda_1 = 6$ i $\lambda_2 = 3$ (podwójnie zdegenerowana). Wobec tego norma macierzy $\mathbf{B}$ indukowana przez normę euklidesową w $\mathbb{R}^3$ wynosi $\|\mathbf{B}\| = 6$.

Przestrzeń metryczna

Przestrzenią metryczną nazywam zbiór $\mathcal{V}$ wraz z określoną na nim metryką, czyli funkcją $d : \mathcal{V} \times \mathcal{V} \to \mathbb{R}$, spełniającą

$$d(A,A) = 0 \tag{15a}$$

$$A \neq B \Rightarrow d(A,B) > 0 \tag{15b}$$

$$d(A,B) = d(B,A) \tag{15c}$$

$$d(A,B) \leqslant d(A,C) + d(C,B) \tag{15d}$$

Jeżeli $\mathcal{V}$ jest przestrzenią liniową z zadaną normą, wówczas

$$d(\mathbf{x},\mathbf{y}) = \|\mathbf{x} - \mathbf{y}\| \tag{16}$$

jest metryką. O metryce takiej mówię, że jest zadana przez normę.

Przykład 6

Norma euklidesowa w $\mathbb{R}^n$ zadaje metrykę euklidesową:

$$d(\mathbf{x},\mathbf{y}) = \|\mathbf{x}-\mathbf{y}\|_2 = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2} \tag{17}$$

Przykład 7

Norma Manhattan zadaje metrykę Manhattan, norma maksimum zadaje metrykę maksimum (zwaną niekiedy w tym kontekście metryką Czebyszewa).

Istnieją metryki, które **nie** są zadawane przez normę, na przykład metryka dyskretna:

$$d(A,B) = \begin{cases} 0 & A = B \\ 1 & A \neq B \end{cases} \tag{18}$$

Kula otwarta

Niech $\mathcal{V}$ będzie przestrzenią metryczną z metryką d. *Kulą otwartą* o środku w punkcie P i promieniu r nazywam zbiór punktów

$$\{X \in \mathcal{V} : d(P, X) < r\} \tag{19}$$

Przykład 8

W przestrzeni $\mathbb{R}^2$ kulą otwartą o środku w punkcie $P(x_0, y_0)$ i promieniu r jest wnętrze okręgu

$$(x - x_0)^2 + (y - y_0)^2 = r \tag{20}$$

W tej samej przestrzeni z metryką Manhattan kulą otwartą są punkty spełniające

$$|x - x_0| + |y - y_0| < r \tag{21}$$

Przestrzeń L_2

Rozważmy pewien przedział $D \subset \mathbb{R}$ — domknięty lub niewłaściwy. Interesować będą nas trzy przypadki:

$$D = \begin{cases} [-1, 1] & \text{lub} \\ [0, +\infty) & \text{lub} \\ (-\infty, +\infty) & \end{cases} \tag{22}$$

Zauważmy, że za pomocą liniowej zmiany zmiennych każdy przedział domknięty można przeprowadzić w przedział $[-1, 1]$, a każdy przedział jednostronnie nieskończony w przedział $[0, +\infty)$.

Niech będzie określona funkcja $w(x)$ o tej własności, że $\forall x \in D : w(x) > 0$. Rozważmy teraz zbiór funkcji $f : D \to \mathbb{C}$ (funkcje zespolone argumentu rzeczywistego). Dodatkowo utożsammy ze sobą wszystkie funkcje różniące się pomiędzy sobą co najwyżej dla przeliczalnej liczby argumentów. Definiuję

$$L_2(D,w) = \left\{ f : D \to \mathbb{C} : \int\limits_D w(x)|f(x)|^2\, dx < +\infty \right\} \tag{23}$$

Jest to zbiór funkcji *całkowalnych z kwadratem* z wagą $w(x)$ na przedziale D.

Twierdzenie: Przestrzeń $L_2(D, w)$ jest przestrzenią liniową.

Jedynym nietrywialnym punktem jest stwierdzenie, że kombinacja liniowa funkcji całkowalnych z kwadratem jest całkowalna z kwadratem. Istotnie,

$$
\begin{aligned}
\int\limits_D w(x)|\alpha f(x) + \beta g(x)|^2 \, dx &\leqslant \int\limits_D w(x) \left(|\alpha f(x)|^2 + |\beta g(x)|^2 \right) dx \\
&= |\alpha|^2 \int\limits_D w(x)|f(x)|^2 \, dx \\
&+ |\beta|^2 \int\limits_D w(x)|g(x)|^2 \, dx < +\infty \qquad (24)
\end{aligned}
$$

Iloczyn skalarny w L_2

Niech $f, g \in L_2(D, w)$. Wówczas operacja

$$\langle f|g\rangle = \int\limits_D w(x) f^*(x) g(x)\, dx \tag{25}$$

jest iloczynem skalarnym.

Iloczyn skalarny (25) generuje normę i metrykę w $L_2(D, w)$.

$$\|f\| = \sqrt{\langle f|f\rangle} = \sqrt{\int\limits_D w(x)|f(x)|^2\, dx} \tag{26}$$

$$d(f,g) = \|f-g\| = \sqrt{\langle f-g|f-g\rangle} = \sqrt{\int\limits_D w(x)|f(x)-g(x)|^2\, dx} \tag{27}$$

Przykład 9

Wielomiany są całkowalne z kwadratem na przedziale $[-1, 1]$ z wagą 1. Zatem, dla przykładu,

$$d(x^2, ax^3) = \left\| x^2 - ax^3 \right\| = \sqrt{\int\limits_{-1}^{1} (x^2 - ax^3)^2 \, dx} = \sqrt{\frac{2a^2}{7} + \frac{2}{5}} \quad (28)$$

Jest to, zdefiniowana w przestrzeni $L_2([-1, 1], 1)$, odległość *pomiędzy funkcjami* x^2, ax^3.

Zupełność

Posiadając pojęcie metryki, możemy zdefiniować granicę ciągów elementów $L_2(D,w)$ (ciągów funkcji całkowalnych z kwadratem), zastępując w definicjach wartości bezwzględne przez odległość $d(f,g) = \|f-g\|$. Możemy także zdefiniować ciągi Cauchy'ego: Niech $\{f_n \in L_2(D,w)\}_{n=1}^{\infty}$ będzie pewnym ciągiem funkcyjnym. Mówimy, że jest to ciąg Cauchy'ego, jeżeli

$$\forall\,\varepsilon > 0\ \exists\, n_0\ \forall\, n,m > n_0 : d(f_n - f_m) < \varepsilon \tag{29}$$

Twierdzenie: W przestrzeni $L_2(D,w)$ wszystkie ciągi Cauchy'ego są zbieżne, to znaczy wszystkie ciągi Cauchy'ego posiadają granice, które same należą do $L_2(D,w)$.

Przestrzeń o tej własności nazywam *przestrzenią zupełną*. (Np zbiór liczb wymiernych **nie** jest zupełny, bo ciągi wymierne mogą być zbieżne do liczb niewymiernych.)

Zupełną przestrzeń liniową z zadanym iloczynem skalarnym nazywam przestrzenią Hilberta.