Lineární algebra II (pro fyziky)

Skalární součin

Dalibor Šmíd

MFF UK

Skalární součin máme zatím definovaný na $\mathbb{R}^n$ předpisem

$$\mathbf{x}.\mathbf{y} = \sum_{i=1}^{n} x_i y_i \equiv \mathbf{x}^T\mathbf{y}$$

Pomocí něj jsme pak zavedli velikost vektoru, vzdálenost vektorů, ortogonální doplněk, ortogonální projekci, ortogonální matici a další pojmy.

Budeme chtít tuto definici zobecnit na libovolný vektorový prostor V. Uvažujme jeho báze B, B' a vektory $u, v \in V$ a označme $\mathbf{x} := [u]^B$, $\mathbf{x}' := [u]^{B'}$, $\mathbf{y} := [v]^B$, $\mathbf{y}' := [v]^{B'}$. Platí $\mathbf{x} = R\mathbf{x}'$, kde R je matice přechodu od B k B'. Pak

$$\mathbf{x}^T\mathbf{y} = \mathbf{x}'^T R^T R \mathbf{y}'$$

Protože $R^T R$ není obecně jednotková matice, výrazy $\mathbf{x}^T\mathbf{y}$ a $\mathbf{x}'^T\mathbf{y}'$ se liší. Nelze tedy skalární součin vektorů $u, v \in V$ zadefinovat jednoznačně pomocí jejich reprezentací.

Připusťme tedy, že skalárních součinů na V může být víc. Jaké vlastnosti mají splňovat?

Definice

Nechť V je vektorový prostor nad $\mathbb{R}$ a $g : V \times V \to \mathbb{R}$ je zobrazení splňující

1. $\forall u, v, w \in V$, $r, s \in \mathbb{R}$:

$$g(ru + sv, w) = rg(u, w) + sg(v, w)$$
$$g(u, rv + sw) = rg(u, v) + sg(u, w)$$

2. $\forall u, v \in V : g(u, v) = g(v, u)$
3. $\forall u \in V, u \neq 0 : g(u, u) > 0$

Pak řekneme, že g je *skalární součin* na V.

Zobrazení $g : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ definované $g(\mathbf{x}, \mathbf{y}) := \mathbf{x}^T \mathbf{y}$ tyto vlastnosti splňuje ♣ a nazývá se *standardní skalární součin* na vektorovém prostoru $\mathbb{R}^n$.

Každé zobrazení $g : V \times V \to \mathbb{R}$ splňující vlastnost (1), se nazývá *bilineární forma*. Pokud $B = (u_1, \ldots, u_n)$, $u, v \in V$, pak

$$g(u, v) = g\left(\sum_{i=1}^{n} x_i u_i, \sum_{j=1}^{n} y_j u_j\right) = \sum_{i,j=1}^{n} x_i y_j g(u_i, u_j)$$

Matice $G \in \mathbb{R}^{n \times n}$, jejíž ij-tý element je $g_{ij} := g(u_i, u_j)$, se značí $[g]_B$ a nazývá *maticí*, případně *reprezentací bilineární formy* vzhledem k bázi B. Pokud $B' = (u'_1, \ldots, u'_n)$, kde $u'_p = \sum_{i=1}^{n} r_{ip} u_i$ (čili $R = [\mathrm{Id}]^B_{B'}$), pak

$$g(u'_p, u'_q) = g\left(\sum_{i=1}^{n} r_{ip} u_i, \sum_{j=1}^{n} r_{jq} u_j\right) = \sum_{i,j=1}^{n} r^T_{pi} g_{ij} r_{jq},$$

neboli

$$[g]_{B'} \equiv G' = R^T G R = \left([\mathrm{Id}]^B_{B'}\right)^T [g]_B [\mathrm{Id}]^B_{B'}$$

Matice bilineární formy se tedy transformuje jinak než matice endomorfismu. Vlastnost (2) je ekvivalentní s tím, že G je symetrická matice, mluvíme o *symetrické bilineární formě*. Vlastnost (3) znamená, že $\mathbf{x}^T G\mathbf{x} > 0$ pro libovolný nenulový $\mathbf{x} \in \mathbb{R}^n$, taková matice (i forma) se nazývá *pozitivně definitní*. Kdyby $G = \operatorname{diag}(g_{11}, \ldots, g_{nn})$, pak

$$\mathbf{x}^T G\mathbf{x} = g_{11}x_1^2 + \ldots + g_{nn}x_n^2 > 0 \Leftrightarrow \forall i \in \{1, \ldots, n\} : g_{ii} > 0$$

U nediagonální G stačí tedy najít R tak, aby $G' = R^T GR$ diagonální byla, neboli reprezentovat g vzhledem k tzv. *polární bázi*. To jde vždy ♣:

Věta

Každá symetrická bilineární forma má polární bázi.

Pokud g je symetrická pozitivně definitní bilineární forma (a tedy skalární součin), pak se její polární báze označují jako báze *ortogonální*. Je-li dokonce $[g]_B = E$, je B *ortonormální*.

Uvažujme třeba $g : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ s předpisem

$$g(\mathbf{x}, \mathbf{y}) := \mathbf{x}^T \begin{pmatrix} 1 & 1 \\ 1 & 5 \end{pmatrix} \mathbf{y} \equiv x_1 y_1 + x_1 y_2 + x_2 y_1 + 5 x_2 y_2$$

Použijeme-li jako R^T matici vhodné řádkové úpravy, získáme

$$R^T G R = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 5 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 4 \end{pmatrix} =: [g]_B,$$

kde polární bázi $B = ((1,0), (-1,1))$ čteme jako sloupce matice přechodu $R = [\mathrm{Id}]^K_B$. Násobení maticí R zprava se chová jako sloupcová úprava odpovídající řádkové úpravě R^T. Vzhledem k asociativitě násobení matic je jedno, která z nich se provede dřív. Úpravám typu $G \to R^T G R$ se říká *symetrické úpravy*.
Tedy g je pozitivně definitní, B je ortogonální báze,
$\tilde{B} = ((1,0), \frac{1}{2}(-1,1))$ je báze ortonormální $\clubsuit$.

Pro V nad $\mathbb{C}$ uvedená definice nefunguje, už standardní skalární součin by nebyl pozitivně definitní:

$$(1 \quad i)\begin{pmatrix}1\\ i\end{pmatrix} = 0,\ (0 \quad i)\begin{pmatrix}0\\ i\end{pmatrix} = -1,\ (0 \quad 1+i)\begin{pmatrix}0\\ 1+i\end{pmatrix} = 2i$$

Definice

Nechť V je vektorový prostor nad $\mathbb{C}$. Zobrazení $g : V \times V \to \mathbb{C}$ je *skalární součin* na V, pakliže

1. $\forall u, v, w \in V$, $r, s \in \mathbb{C}$:

$$g(ru + sv, w) = \bar{r}g(u, w) + \bar{s}g(v, w)$$
$$g(u, rv + sw) = rg(u, v) + sg(u, w)$$

2. $\forall u, v \in V : g(u, v) = \overline{g(v, u)}$
3. $\forall u \in V, u \neq 0 : g(u, u) > 0$

- *Standardní skalární součin* na $\mathbb{C}^n$ je $g(\mathbf{x}, \mathbf{y}) = \mathbf{x}^+\mathbf{y}$.
- První část vlastnosti (1) se nazývá *antilinearita* v prvním argumentu, celá vlastnost pak *seskvilinearita*. Často se používá i konvence s antilinearitou v druhém a linearitou v prvním argumentu.
- Matice seskvilineární formy $[g]_B \equiv G := (g(u_i, u_j))$ splňující vlastnost (2) je hermitovská, mluvíme pak o *hermitovské seskvilineární formě*.
- Díky vlastnosti (2) je $g(u, u)$ vždy v $\mathbb{R}$, tedy vlastnost (3) dává smysl. Komplexní skalární součin je tedy hermitovská pozitivně definitní seskvilineární forma.
- Při změně báze $G \to R^+GR$, polární báze se tedy dá hledat *hermitovskými úpravami* (nejprve řádková, pak komplexně sdružená sloupcová) a vždy to jde $\clubsuit$.
- Ortogonální a ortonormální báze jsou definovány stejně. Je praktické vztáhnout definici skalárního součinu nad $\mathbb{C}$ i na $\mathbb{R}$ a všude pak jen ignorovat symboly komplexního sdružení.

Dvojice (V, g), kde V je vektorový prostor a g skalární součin na něm, se nazývá *unitární prostor*. Skalární součin budeme odteď značit místo $g(u, v)$ pouze $\langle u, v\rangle$. Definujeme pak

- *Kolmost:* $u \perp v \Leftrightarrow \langle u, v\rangle = 0$
- *Ortogonální doplněk:* $M^{\perp} = \{u \in V | \forall v \in M : u \perp v\}$
- *Normu vektoru:* $\|u\| := \sqrt{\langle u, u\rangle}$
- *Vzdálenost vektorů:* $\rho(u, v) := \|u - v\|$
- *Úhel vektorů:* $\arccos \frac{|\langle u,v\rangle|}{\|u\|\|v\|}$ nad $\mathbb{C}$ nebo $\arccos \frac{\langle u,v\rangle}{\|u\|\|v\|}$ nad $\mathbb{R}$
- *Ortogonální projekci na $v \in V$:* $P_v : V \to V$, $P_v(u) = \frac{\langle v,u\rangle}{\|v\|^2} v$
- *Ortogonální projekci na og. doplněk:* $P_{v^{\perp}} = \mathrm{Id} - P_v$

Z rozpisu $\|u + v\|^2 = \langle u, u\rangle + \langle u, v\rangle + \langle v, u\rangle + \langle v, v\rangle$ plyne tzv. polarizační identita

$$\mathrm{Re}\langle u, v\rangle = \frac{1}{2}(\|u + v\|^2 - \|u\|^2 - \|v\|^2)$$

a pro $u \perp v$ též Pythagorova věta $\|u + v\|^2 = \|u\|^2 + \|v\|^2$.

Tvrzení

Nechť V je unitární prostor nad $\mathbb{F}$, $u, v \in V$, $v \neq o$. Pak

1. $P_v, P_{v^\perp}$ *jsou lineární zobrazení*
2. $P_v(v) = v$, $\operatorname{Ker} P_v = v^\perp$, $\operatorname{Im} P_v = \langle v \rangle$,
3. $\operatorname{Ker} P_{v^\perp} = \langle v \rangle$, $\operatorname{Im} P_{v^\perp} = v^\perp$
4. $P_v \circ P_v = P_v$, $P_{v^\perp} \circ P_{v^\perp} = P_{v^\perp}$
5. *Vektor u lze jednoznačně zapsat jako $u_\parallel + u_\perp$, kde $u_\parallel \in \langle v \rangle$, $u_\perp \perp v$. Navíc pak $u_\parallel = P_v(u)$ a $u_\perp = P_{v^\perp}(u)$.*
6. $|\langle u, v \rangle| \leq \|u\|\|v\|$ *a rovnost platí pouze pro $u \in \langle v \rangle$.*

Důkazy jsou analogické jako v první kapitole ♣. Ze Schwarzovy nerovnosti pak plyne ♣ trojúhelníková nerovnost, tj. $\forall u, v \in V$:

$$\|u + v\| \leq \|u\| + \|v\|$$

Uvažujme posloupnost $(w_1, \ldots, w_k)$ ve V, pro kterou platí $\langle w_i, w_j \rangle = \delta_{ij}$ (*Kroneckerovo delta*, tj. 1 pro $i = j$ a 0 jindy), taková posloupnost se nazývá *ortonormální*. Označme $W \leq V$ její lineární obal, a $P_W : V \to V$ zobrazení

$$P_W(u) := \langle w_1, u \rangle w_1 + \ldots + \langle w_n, u \rangle w_n \equiv \sum_{i=1}^{k} P_{w_i}(u)$$

Pro zobrazení P_W platí ♣:

1. $u \in W$ právě když $P_W(u) = u$; $\operatorname{Ker} P_W = W^\perp$, $\operatorname{Im} P_W = W$.
2. $\forall u \in V : \|P_W(u)\| \leq \|u\|$, přičemž rovnost nastává právě když $u \in W$
3. $\forall v \in W : \|u - P_W(u)\| \leq \|u - v\|$, přičemž rovnost nastává právě když $v = P_W(u)$.

Třetí bod říká, že $P_W(u)$ je vektor W, který je nejblíže vektoru $u \in V$, a že je tento vektor určen jednoznačně. Tedy zobrazení P_W je ortogonální projekce na podprostor W a nezáleží, pomocí které ortonormální báze W jej zadefinujeme.