Lineární algebra I (pro fyziky)

Vektory a zobrazení v $\mathbb{R}^n$

Dalibor Šmíd

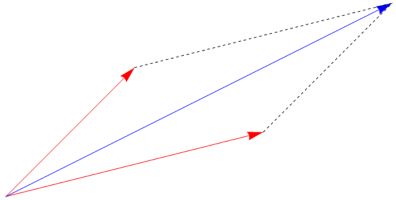
MFF UK

Vektor v rovině můžeme zapsat jako uspořádanou dvojici reálných čísel $\mathbf{x} := \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$. Součet dvou vektorů a násobek vektoru skalárem (číslem $r \in \mathbb{R}$) můžeme definovat dvěma ekvivalentními způsoby:

algebraicky

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} := \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \end{pmatrix}$$
$$r \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} := \begin{pmatrix} r x_1 \\ r x_2 \end{pmatrix}$$

geometricky



Algebraickou definici operací snadno zobecníme na vektory v prostoru (uspořádané trojice) nebo rovnou na uspořádané n-tice, tedy vektory z $\mathbb{R}^n \equiv \underbrace{\mathbb{R} \times \ldots \times \mathbb{R}}_{n}$:

$$\mathbf{x} + \mathbf{y} \equiv \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} := \begin{pmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{pmatrix}$$

$$r\mathbf{x} \equiv r \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} := \begin{pmatrix} rx_1 \\ \vdots \\ rx_n \end{pmatrix}$$

Čísla x_i jsou *složky* vektoru $\mathbf{x}$. Vektor, jehož všechny složky jsou nulové, označujeme $\mathbf{o}$ (*nulový vektor*), vektor $-\mathbf{x} := (-1)\mathbf{x}$ je *opačný vektor*. Množině $\mathbb{R}^n$ budeme říkat *n-rozměrný prostor*.

Skalární součin dvojice vektorů v rovině je definován předpisem

$$\mathbf{x}.\mathbf{y} := x_1 y_1 + x_2 y_2$$

Pomocí tohoto předpisu lze zapsat *velikost (normu)* vektoru

$$\|\mathbf{x}\| := \sqrt{\mathbf{x}.\mathbf{x}} \equiv \sqrt{x_1^2 + x_2^2}$$

Tvrzení

Je-li $\phi \in \langle 0, \pi \rangle$ *úhel mezi vektory* $\mathbf{x}, \mathbf{y} \in \mathbb{R}^2$, *pak*

$$\mathbf{x}.\mathbf{y} = \|\mathbf{x}\| \|\mathbf{y}\| \cos \phi$$

Návod k důkazu tohoto tvrzení naleznete ve cvičeních, což zde i v dalších podobných případech označíme symbolem $\clubsuit$.

Zavedeme skalární součin a normu vektoru analogicky i na $\mathbb{R}^n$:

$$\mathbf{x}.\mathbf{y} := \sum_{i=1}^{n} x_i y_i, \quad \|\mathbf{x}\| := \sqrt{\mathbf{x}.\mathbf{x}}$$

V třírozměrném prostoru platí stejné tvrzení o souvislosti skalárního součinu a úhlu mezi vektory. V obecném $\mathbb{R}^n$ nevíme, co to úhel mezi vektory je, ale můžeme jej (pro nenulové vektory $\mathbf{x}$, $\mathbf{y}$) zadefinovat jako $\phi := \arccos \frac{\mathbf{x}.\mathbf{y}}{\|\mathbf{x}\|\|\mathbf{y}\|}$. Aby tato definice byla korektní, potřebujeme vědět, že

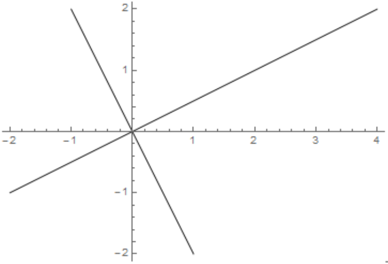
$$-1 \leq \frac{\mathbf{x}.\mathbf{y}}{\|\mathbf{x}\|\|\mathbf{y}\|} \leq 1,$$

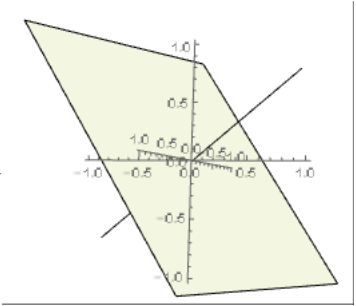
neboli $|\mathbf{x}.\mathbf{y}| \leq \|\mathbf{x}\|\|\mathbf{y}\|$ (Schwarzova nerovnost ♣).

Dva vektory prohlásíme za *kolmé* (značíme $\mathbf{x} \perp \mathbf{y}$), právě když $\mathbf{x}.\mathbf{y} = 0$, tedy když je jeden z nich nulový nebo když svírají úhel $\phi = \frac{\pi}{2}$. Množina

$$\mathbf{x}^{\perp} := \{\mathbf{y} \in \mathbb{R}^n | \mathbf{x}.\mathbf{y} = 0\}$$

všech vektorů $\mathbf{y}$ kolmých na vektor $\mathbf{x}$ se nazývá *ortogonální doplněk vektoru* $\mathbf{x}$. V $\mathbb{R}^2$ je to přímka, v $\mathbb{R}^3$ rovina, obě procházejí počátkem:





Množinám tvaru $\{\mathbf{y} \in \mathbb{R}^n | \mathbf{x}.\mathbf{y} = c\}$ říkáme obecně *nadroviny*. Nadrovina obsahuje vektor $\mathbf{o}$, právě když $c = 0$.

Vektorový součin $\mathbf{x}, \mathbf{y} \in \mathbb{R}^3$ je vektor definovaný po složkách jako

$$(\mathbf{x} \times \mathbf{y})_k = \sum_{i=1}^{3} \sum_{j=1}^{3} \varepsilon_{ijk} x_i y_j,$$

kde $\varepsilon_{123} = \varepsilon_{231} = \varepsilon_{312} = 1$, $\varepsilon_{213} = \varepsilon_{321} = \varepsilon_{132} = -1$ a $\varepsilon_{ijk} = 0$, jsou-li dva indexy stejné.

Tvrzení

Nechť $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^3$, $r, s \in \mathbb{R}$, $\mathbf{x}$ *a* $\mathbf{y}$ *svírají úhel* ϕ. *Pak*

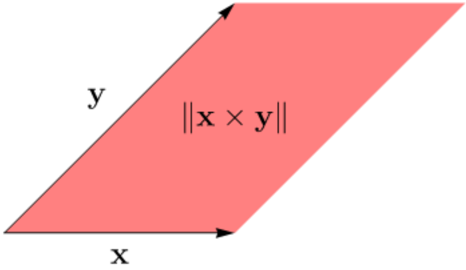
1. $\mathbf{x} \times \mathbf{y} = -\mathbf{y} \times \mathbf{x}$, *speciálně* $\mathbf{x} \times \mathbf{x} = 0$
2. $\mathbf{x} \times (r\mathbf{y} + s\mathbf{z}) = r(\mathbf{x} \times \mathbf{y}) + s(\mathbf{x} \times \mathbf{z})$
3. $\|\mathbf{x} \times \mathbf{y}\| = \|\mathbf{x}\|\|\mathbf{y}\||\sin\phi|$
4. $|(\mathbf{x} \times \mathbf{y}).\mathbf{z}|$ *je objem rovnoběžnostěnu definovaného vektory* $\mathbf{x}, \mathbf{y}, \mathbf{z}$, *speciálně* 0 *pro* $z = r\mathbf{x} + s\mathbf{y}$.

Důkaz ♣. Z druhé části posledního bodu plyne, že $\mathbf{x} \times \mathbf{y}$ je kolmý na rovinu definovanou $\mathbf{x}$ a $\mathbf{y}$.

Oproti skalárnímu součinu se nenabízí žádné přímočaré zobecnění vektorového součinu do libovolné dimenze n. V $\mathbb{R}^2$ lze vektoru $\mathbf{x}$ přiřadit vektor

$$\tilde{\mathbf{x}} := \begin{pmatrix} \mathbf{x} \\ 0 \end{pmatrix} \in \mathbb{R}^3$$

a definovat $\mathbf{x} \times \mathbf{y} := (\tilde{\mathbf{x}} \times \tilde{\mathbf{y}})_3 \in \mathbb{R}$. Pak je $\|\mathbf{x} \times \mathbf{y}\|$ obsah rovnoběžníka definovaného $\mathbf{x}$ a $\mathbf{y}$ (♣):



Při zobecnění do dimenze $n > 3$ chceme, aby vektorový součin dál počítal objemy rovnoběžnostěnů, což lze splnit definicí

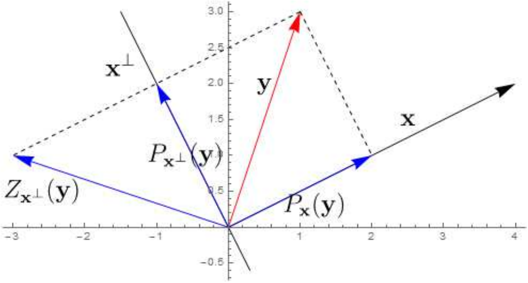
$$(\mathbf{x}_1 \times \mathbf{x}_2 \times \ldots \times \mathbf{x}_{n-1})_{k_n} := \sum_{k_1=1}^{n} \sum_{k_2=1}^{n} \cdots \sum_{k_{n-1}=1}^{n} \varepsilon_{k_1 k_2 \ldots k_n} \mathbf{x}_{1k_1} \mathbf{x}_{2k_2} \ldots \mathbf{x}_{n-1,k_{n-1}},$$

kde symbol ε je opět definován tak, aby měnil znaménko při výměně libovolné dvojice indexů a aby $\varepsilon_{12\ldots n} = 1$.

Zobecněný vektorový součin tedy již není funkce dvou, ale $n-1$ vektorových argumentů, výsledkem je opět vektor. Jsou ale splněny analogie vlastností $1, 2$ a 4 z tvrzení ($\clubsuit$). Například objem rovnoběžnostěnu určeného vektory $\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d} \in \mathbb{R}^4$ je roven

$$|(\mathbf{a} \times \mathbf{b} \times \mathbf{c}).\mathbf{d}| = \left| \sum_{i=1}^{4} \sum_{j=1}^{4} \sum_{k=1}^{4} \sum_{l=1}^{4} \varepsilon_{ijkl} a_i b_j c_k d_l \right|$$

Pomocí skalárního součinu můžeme vyjádřit kolmý průmět (*ortogonální projekci*) vektoru $\mathbf{y}$ do směru jiného (nenulového) vektoru $\mathbf{x}$, kolmý průmět do směru jeho ortogonálního doplňku a vektor zrcadlený podle nadroviny $\mathbf{x}^\perp$.



$$P_{\mathbf{x}}(\mathbf{y}) = \|\mathbf{y}\| \cos\phi \frac{\mathbf{x}}{\|\mathbf{x}\|} = \frac{\mathbf{x}.\mathbf{y}}{\|\mathbf{x}\|^2}\mathbf{x}$$

$$P_{\mathbf{x}^\perp}(\mathbf{y}) = \mathbf{y} - P_{\mathbf{x}}(\mathbf{y}) = \mathbf{y} - \frac{\mathbf{x}.\mathbf{y}}{\|\mathbf{x}\|^2}\mathbf{x}$$

$$Z_{\mathbf{x}^\perp}(\mathbf{y}) = \mathbf{y} - 2P_{\mathbf{x}}(\mathbf{y}) = \mathbf{y} - 2\frac{\mathbf{x}.\mathbf{y}}{\|\mathbf{x}\|^2}\mathbf{x}$$

Stejné vzorce fungují i v obecné dimenzi n. Platí totiž, že $P_{\mathbf{x}}(\mathbf{y}) = \mathbf{o}$, právě když $\mathbf{x} \perp \mathbf{y}$, a

$$\mathbf{x}.(\mathbf{y} - P_{\mathbf{x}}(\mathbf{y})) = \mathbf{x}.\mathbf{y} - \frac{\mathbf{x}.\mathbf{y}}{\|\mathbf{x}\|^2}\mathbf{x}.\mathbf{x} = 0,$$

tedy rozdíl $\mathbf{y}$ a $P_{\mathbf{x}}(\mathbf{y})$ je vždy kolmý na vektor $\mathbf{x}$.

$P_{\mathbf{x}}$ lze chápat jako *zobrazení*, tedy „mašinku“, do které se vloží libovolný prvek $\mathbf{y} \in \mathbb{R}^n$ a ona mu přiřadí jednoznačně určený prvek $P_{\mathbf{x}}(\mathbf{y}) \in \mathbb{R}^n$. Navíc $\forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, $\forall r, s \in \mathbb{R}$ platí

$$P_{\mathbf{x}}(r\mathbf{y} + s\mathbf{z}) = rP_{\mathbf{x}}(\mathbf{y}) + sP_{\mathbf{x}}(\mathbf{z}),$$

neboli zobrazení $P_{\mathbf{x}}$ je *lineární*. Podobné úvahy vedou k tomu, že i $P_{\mathbf{x}^\perp}, Z_{\mathbf{x}^\perp} : \mathbb{R}^n \to \mathbb{R}^n$ jsou lineární zobrazení, realizující projekci, resp. zrcadlení podle nadroviny $\mathbf{x}^\perp$.

Otočení (*rotaci*) vektoru $\mathbf{y} \in \mathbb{R}^2$ okolo počátku o úhel ϕ proti směru hodinových ručiček lze zapsat jako zobrazení

$$\mathbf{y} := \begin{pmatrix} \|y\| \cos(\alpha) \\ \|y\| \sin(\alpha) \end{pmatrix} \mapsto R_\phi(\mathbf{y}) := \begin{pmatrix} \|y\| \cos(\alpha + \phi) \\ \|y\| \sin(\alpha + \phi) \end{pmatrix}$$

Po aplikaci součtových vzorců máme vyjádření ve složkách:

$$R_\phi(\mathbf{y}) := \begin{pmatrix} y_1 \cos\phi - y_2 \sin\phi \\ y_1 \sin\phi + y_2 \cos\phi \end{pmatrix} = \cos\phi \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} + \sin\phi \begin{pmatrix} -y_2 \\ y_1 \end{pmatrix}$$

Několik pozorování:

- R_ϕ je lineární zobrazení.
- $R_0(\mathbf{y}) = \mathbf{y}$, tedy $R_0 = \mathrm{Id}$ (*identita*).
- Složení dvou rotací $R_\phi \circ R_\psi$ je rotace $R_{\phi+\psi}$. Na pořadí složení zde nezáleží, říkáme, že R_ϕ a R_ψ *komutují*.
- $R_{-\phi} \circ R_\phi = R_\phi \circ R_{-\phi} = R_0 = \mathrm{Id}$, tedy zobrazení $R_{-\phi}$ je *inverzní* k R_ϕ, $R_{-\phi} = (R_\phi)^{-1}$.

Rotaci vektoru $\mathbf{y} \in \mathbb{R}^3$ okolo osy dané vektorem $\mathbf{v}$ ($\|\mathbf{v}\| = 1$) můžeme zapsat například pomocí Rodriguesovy formule $\clubsuit$:

$$R_{\mathbf{v},\phi}(\mathbf{y}) = \cos\phi\ \mathbf{y} + (1 - \cos\phi)(\mathbf{v}.\mathbf{y})\mathbf{v} + \sin\phi\ \mathbf{v} \times \mathbf{y}$$

Rotace v $\mathbb{R}^3$ komutují, pokud mají společnou osu, jindy komutovat nemusí. Vektory

$$\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \mathbf{e}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

a jejich přímočará zobecnění do $\mathbb{R}^n$ se nazývají *prvky kanonické báze*. Nekomutativita rotací se projeví třeba na příkladu

$$R_{\mathbf{e}_2,\frac{\pi}{2}} \circ R_{\mathbf{e}_3,\frac{\pi}{2}}(\mathbf{e}_3) = R_{\mathbf{e}_2,\frac{\pi}{2}}(\mathbf{e}_3) = \mathbf{e}_1$$
$$R_{\mathbf{e}_3,\frac{\pi}{2}} \circ R_{\mathbf{e}_2,\frac{\pi}{2}}(\mathbf{e}_3) = R_{\mathbf{e}_3,\frac{\pi}{2}}(\mathbf{e}_1) = \mathbf{e}_2$$

Dalším příkladem lineárního zobrazení z $\mathbb{R}^n$ do $\mathbb{R}^n$ je roztažení(*dilatace*) faktorem $\lambda \in \mathbb{R}$:

$$D_\lambda(\mathbf{y}) := \lambda \mathbf{y}$$

Naopak posunutí (*translace*) o nenulový vektor $\mathbf{b} \in \mathbb{R}^n$

$$T_{\mathbf{b}}(\mathbf{y}) := \mathbf{y} + \mathbf{b}$$

lineárním zobrazením není. Lineární zobrazení F totiž musí splňovat $F(\mathbf{o}) = F(0.\mathbf{x} + 0.\mathbf{y}) = 0F(\mathbf{x}) + 0F(\mathbf{y}) = \mathbf{o}$, pro translaci ale $T_{\mathbf{b}}(\mathbf{o}) = \mathbf{b}$.

Jak tedy poznáme, že je nějaké zobrazení $F : \mathbb{R}^n \to \mathbb{R}^n$ lineární? Každý vektor $\mathbf{x} \in \mathbb{R}^n$ lze zapsat jako

$$\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \ldots + x_n\mathbf{e}_n,$$

neboli jako *lineární kombinaci* prvků kanonické báze. Z linearity zobrazení F plyne, že

$$\begin{aligned} F(\mathbf{x}) &= F(x_1\mathbf{e}_1 + \ldots + x_n\mathbf{e}_n) = x_1F(\mathbf{e}_1) + \ldots + x_nF(\mathbf{e}_n) \\ &= x_1\mathbf{a}_1 + \ldots + x_n\mathbf{a}_n, \end{aligned}$$

kde jsme jako $\mathbf{a}_j \in \mathbb{R}^n$ označili obraz $F(\mathbf{e}_j)$ vektoru $\mathbf{e}_i$. Označíme-li i-tou složku vektoru $\mathbf{a}_j$ jako a_{ij}, je to dohromady n^2 reálných čísel, která společně zobrazení F jednoznačně určují. Tato čísla souhrnně označujeme jako *matici lineárního zobrazení* a značíme $A \equiv (a_{ij})$.

Zobrazení určené maticí A označujeme symbolem F_A. Pak tedy můžeme zavést zápis lineárního zobrazení pomocí jeho *matice*:

$$\begin{aligned} F_A(\mathbf{x}) &:= x_1 \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{pmatrix} + x_2 \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{n2} \end{pmatrix} + \ldots + x_n \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{nn} \end{pmatrix} \\ &=: \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \equiv A\mathbf{x} \end{aligned}$$

Druhou rovností jsme jednak zavedli standardní zápis matice jako tabulky čísel, v níž první index čísluje řádky a druhý sloupce. Zároveň jsme také zadefinovali *součin matice a vektoru*.

Součin matice A a vektoru $\mathbf{x}$ je vlastně lineární kombinací sloupců matice A s koeficienty rovnými složkám vektoru $\mathbf{x}$. Užívá se proto také *sloupcový zápis matice* $A = (\mathbf{a}_1|\mathbf{a}_2|\ldots|\mathbf{a}_n)$. Zobrazení s maticí

$$E := (\mathbf{e}_1|\ldots|\mathbf{e}_n) \equiv \begin{pmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 1 \end{pmatrix}$$

přiřazuje každému vektoru $\mathbf{x}$ vektor

$$F_E(\mathbf{x}) = E\mathbf{x} = x_1\mathbf{e}_1 + \ldots + x_n\mathbf{e}_n = \mathbf{x},$$

je to tedy identita $F_E = \mathrm{Id}$. Matici E se říká *jednotková matice*.

Označíme-li symbolem

$$\tilde{\mathbf{a}}_i := \begin{pmatrix} a_{i1} \\ a_{i2} \\ \vdots \\ a_{in} \end{pmatrix}$$

i-tý řádek matice A, zapsaný ovšem jako sloupcový vektor, pak i-tá složka obrazu vektoru $\mathbf{x}$ v zobrazení F_A je

$$F_A(\mathbf{x})_i \equiv (A\mathbf{x})_i = a_{i1}x_1 + a_{i2}x_2 + \ldots a_{in}x_n \equiv \sum_{j=1}^{n} a_{ij}x_j,$$

což lze interpretovat jako skalární součin vektorů $\tilde{\mathbf{a}}_i$ a $\mathbf{x}$. Pokud tedy $F_A(\mathbf{x})_i = 0$, je vektor $\mathbf{x}$ kolmý na $\tilde{\mathbf{a}}_i$. Podobně pokud $F_A(\mathbf{x})_i = c$, leží $\mathbf{x}$ v nadrovině s rovnicí $\tilde{\mathbf{a}}_i.\mathbf{x} = c$.