K oddílu VI.5 – lineární zobrazení

K definici lineárního zobrazení

- Zobrazení $f : \mathbb{R}^n \to \mathbb{R}^m$ je lineární, pokud má dvě uvedené vlastnosti.

  Tyto vlastnosti znamenají, že f „zachovává operace a sčítání a násobení reálným číslem“. Tyto dvě operace určují „lineární strukturu“ prostoru $\mathbb{R}^n$. Proto zobrazení je lineární, pokud „zachovává lineární strukturu“.

  Tato intuitivní vyjádření vystihují důležitost lineárních zobrazení. Tu ale více pochopíme až v obecnějším kontextu v Matematice III. V tomto oddílu se zabýváme pouze speciálním případem, který má ilustrovat použití maticového násobení.

- Lineární zobrazení $\mathbb{R}$ do $\mathbb{R}$:

  Pokud $a \in \mathbb{R}$, pak zobrazení $f : \mathbb{R} \to \mathbb{R}$ definované vzorcem

  $$f(x) = ax, \quad x \in \mathbb{R},$$

  je lineární. To plyne z vlastností operací na $\mathbb{R}$, protože pro $x, y \in \mathbb{R}$ platí

  $$f(x + y) = a(x + y) = ax + ay = f(x) + f(y)$$

  a pro $x \in \mathbb{R}$ a $\lambda \in \mathbb{R}$ platí

  $$f(\lambda x) = a(\lambda x) = \lambda(ax) = \lambda f(x).$$

  Navíc, každé lineární zobrazení $\mathbb{R}$ do $\mathbb{R}$ má uvedený tvar. Pokud totiž $f : \mathbb{R} \to \mathbb{R}$ je lineární a označíme $a = f(1)$, pak pro každé $x \in \mathbb{R}$ platí

  $$f(x) = f(x \cdot 1) = x \cdot f(1) = x \cdot a = ax.$$

  Shrneme-li to, pak zobrazení $f : \mathbb{R} \to \mathbb{R}$ je lineární, právě když je dáno vzorcem $f(x) = ax$ pro nějaké $a \in \mathbb{R}$, tedy právě když grafem f je nějaká přímka procházející počátkem.

- Lineární zobrazení $\mathbb{R}$ do $\mathbb{R}^m$:

  Situace je analogická předchozímu případu. Pokud $\boldsymbol{a} \in \mathbb{R}^m$, pak zobrazení $f : \mathbb{R} \to \mathbb{R}^m$ definované vzorcem

  $$f(x) = x\boldsymbol{a}, \quad x \in \mathbb{R}$$

je lineární. To plyne z vlastností operací na $\mathbb{R}$ a $\mathbb{R}^m$, protože pro $x, y \in \mathbb{R}$ platí

$$f(x+y) = (x+y)\boldsymbol{a} = x\boldsymbol{a} + y\boldsymbol{a} = f(x) + f(y)$$

a pro $x \in \mathbb{R}$ a $\lambda \in \mathbb{R}$ platí

$$f(\lambda x) = (\lambda x)\boldsymbol{a} = \lambda(x\boldsymbol{a}) = \lambda f(x).$$

Navíc opět platí, že každé lineární zobrazení $\mathbb{R}$ do $\mathbb{R}^m$ má uvedený tvar. Pokud totiž $f : \mathbb{R} \to \mathbb{R}^m$ je lineární a označíme $\boldsymbol{a} = f(1)$, pak pro každé $x \in \mathbb{R}$ platí

$$f(x) = f(x \cdot 1) = x \cdot f(1) = x\boldsymbol{a}.$$

- Lineární zobrazení $\mathbb{R}^2$ do $\mathbb{R}$:

  Situace je opět analogická, i když o něco složitější. Pokud $a, b \in \mathbb{R}$, pak zobrazení $f : \mathbb{R}^2 \to \mathbb{R}$ definované vzorcem

$$f(x, y) = ax + by, \quad [x, y] \in \mathbb{R}^2$$

je lineární. To plyne z vlastností operací na $\mathbb{R}$ a $\mathbb{R}^2$, protože pro $[x, y], [u, v] \in \mathbb{R}^2$ platí

$$\begin{aligned} f([x, y] + [u, v]) &= f(x+u, y+v) = a(x+u) + b(y+v) \\ &= ax + by + au + bv = f(x, y) + f(u, v) \end{aligned}$$

a pro $[x, y] \in \mathbb{R}^2$ a $\lambda \in \mathbb{R}$ platí

$$f(\lambda[x, y]) = f(\lambda x, \lambda y) = a \cdot \lambda x + b \cdot \lambda y = \lambda(ax + by) = \lambda f(x, y).$$

I v tomto případě platí, že každé lineární zobrazení $\mathbb{R}^2$ do $\mathbb{R}$ má uvedený tvar. Mějme totiž $f : \mathbb{R}^2 \to \mathbb{R}$ lineární zobrazení. Označme

$$a = f(1, 0), \qquad b = f(0, 1).$$

Pak pro každé $[x, y] \in \mathbb{R}^2$ platí

$$\begin{aligned} f(x, y) &= f([x, 0] + [0, y]) = f(x \cdot [1, 0] + y \cdot [0, 1]) \\ &= f(x \cdot [1, 0]) + f(y \cdot [0, 1]) = x \cdot f(1, 0) + y \cdot f(0, 1) \\ &= x \cdot a + y \cdot b = ax + by. \end{aligned}$$

Tedy, zobrazení $f : \mathbb{R}^2 \to \mathbb{R}$ je lineární, právě když jeho grafem je nějaká rovina procházející počátkem.

- Obecný případ – lineární zobrazení $\mathbb{R}^n$ do $\mathbb{R}^m$.

  Pokud $\mathbb{A}$ je matice typu $m \times n$, pak zobrazení $f : \mathbb{R}^n \to \mathbb{R}^m$ definované vzorcem
  $$f(\boldsymbol{u}) = \mathbb{A}\boldsymbol{u}, \qquad \boldsymbol{u} \in \mathbb{R}^n,$$
  kde prvky $\mathbb{R}^n$ a $\mathbb{R}^m$ reprezentujeme jako sloupcové vektory (v souladu s poznámkou na začátku oddílu), je lineární.

  To plyne z vlastností maticového násobení: Pro $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^n$ platí
  $$f(\boldsymbol{u} + \boldsymbol{v}) = \mathbb{A}(\boldsymbol{u} + \boldsymbol{v}) = \mathbb{A}\boldsymbol{u} + \mathbb{A}\boldsymbol{v} = f(\boldsymbol{u}) + f(\boldsymbol{v})$$
  a pro $\boldsymbol{u} \in \mathbb{R}^n$ a $\lambda \in \mathbb{R}$ platí
  $$f(\lambda\boldsymbol{u}) = \mathbb{A}(\lambda\boldsymbol{u}) = \lambda\mathbb{A}\boldsymbol{u} = \lambda f(\boldsymbol{u}).$$

  V tomto případě říkáme, že matice $\mathbb{A}$ reprezentuje zobrazení f.

  Opět platí, že každé lineární zobrazení má tento tvar. To je obsahem Věty VI.18.

K Větě VI.18:

- Důkaz věty:

  Implikace $\Leftarrow$: Pokud je f reprezentováno nějakou maticí, je lineární. To bylo dokázáno již výše.

  Implikace $\Rightarrow$: Nechť $f : \mathbb{R}^n \to \mathbb{R}^m$ je lineární zobrazení. Ukážeme, že je reprezentováno nějakou maticí. Budeme postupovat podobně jako ve speciálních případech výše.

  Připomeňme, že $\boldsymbol{e}_j$ je prvek $\mathbb{R}^n$, který má na j-tém místě číslo 1 a jinde má nuly. Označme $\boldsymbol{a}_j = f(\boldsymbol{e}_j)$. Pak $\boldsymbol{a}_j \in \mathbb{R}^m$.

  Nechť $\mathbb{A}$ je matice, která má sloupce $\boldsymbol{a}_1, \dots, \boldsymbol{a}_n$ (připomeňme, že prvky $\mathbb{R}^n$ reprezentujeme jako sloupcové vektory). Tj.,
  $$\mathbb{A} = (\boldsymbol{a}_1 \;\; \boldsymbol{a}_2 \;\; \dots \;\; \boldsymbol{a}_n).$$
  Pak $\mathbb{A}$ je matice typu $m \times n$.

Vezměme nyní

$$\boldsymbol{u} = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} \in \mathbb{R}^n.$$

Pak

$$\boldsymbol{u} = u_1\boldsymbol{e}_1 + u_2\boldsymbol{e}_2 + \cdots + u_n\boldsymbol{e}_n,$$

a tedy

$$\begin{aligned} f(\boldsymbol{u}) &= f(u_1\boldsymbol{e}_1 + u_2\boldsymbol{e}_2 + \cdots + u_n\boldsymbol{e}_n) \\ &= f(u_1\boldsymbol{e}_1) + f(u_2\boldsymbol{e}_2) + \cdots + f(u_n\boldsymbol{e}_n) \\ &= u_1 f(\boldsymbol{e}_1) + u_2 f(\boldsymbol{e}_2) + \cdots + u_n f(\boldsymbol{e}_n) \\ &= u_1\boldsymbol{a}_1 + u_2\boldsymbol{a}_2 + \cdots + u_n\boldsymbol{a}_n = \mathbb{A}\boldsymbol{u}, \end{aligned}$$

kde poslední rovnost plyne z definice maticového násobení (viz důkaz ekvivalence (1) $\Leftrightarrow$ (3) z Věty VI.16).

Tedy $\mathbb{A}$ reprezentuje f.

Jednoznačnost reprezentující matice: Matice, která reprezentuje zobrazení f je jednoznačně určena:

Nechť matice $\mathbb{A}$ reprezentuje zobrazení f. Pro každé $j \in \{1, \ldots, n\}$ je součin $\mathbb{A}\boldsymbol{e_j}$ roven j-tému sloupci matice $\mathbb{A}$ (to plyne snadno z definice maticového násobení).

Protože $\mathbb{A}\boldsymbol{e}_j = f(\boldsymbol{e}_j)$, plyne z toho, že v j-tém sloupci matice $\mathbb{A}$ je vektor $f(\boldsymbol{e}_j)$.

Tedy matice $\mathbb{A}$ je jednoznačně určena. (Přesně takto jsme ji konstruovali výše.)

- Význam této věty:

Tato věta je příkladem tzv. „reprezentačních vět“ v matematice. Nějaký objekt (v našem případě lineární zobrazení $\mathbb{R}^n$ do $\mathbb{R}^m$) máme definovaný abstraktně, pomocí nějakých vlastností (v našem případě jsou to dvě vlastnosti z definice). A tato věta říká, že každý z těchto abstraktních objektů (každé zobrazení $\mathbb{R}^n$ do $\mathbb{R}^m$ splňující ony dvě vlastnosti) má velmi konkrétní formu (takové zobrazení je reprezentované maticí, tj. dáno jednoduchým vzorcem).

K Větě VI.19:

- V této větě je podstatné, že jde o zobrazení $\mathbb{R}^n$ do $\mathbb{R}^n$, tj. výchozí a cílový prostor jsou stejné.
- Triviální implikace:

  Implikace $(i) \Rightarrow (ii)$ a $(i) \Rightarrow (iii)$ jsou triviální. Důležité jsou ty obrácené.
- Přeformulování pomocí reprezentující matice:

  Podle Věty VI.18 víme, že zobrazení f je reprezentováno nějakou čtvercovou maticí $\mathbb{A}$ řádu n. S využitím tohoto faktu a toho, co znamená, že $\mathbb{A}$ reprezentuje f, můžeme tvrzení (i)–(iii) přeformulovat následovně:

  (i') Pro každé $\boldsymbol{b} \in \mathbb{R}^n$ existuje právě jedno $\boldsymbol{x} \in \mathbb{R}^n$, pro které $\mathbb{A}\boldsymbol{x} = \boldsymbol{b}$.

  (ii') Pro každé $\boldsymbol{b} \in \mathbb{R}^n$ existuje nejvýše jedno $\boldsymbol{x} \in \mathbb{R}^n$, pro které $\mathbb{A}\boldsymbol{x} = \boldsymbol{b}$.

  (iii') Pro každé $\boldsymbol{b} \in \mathbb{R}^n$ existuje alespoň jedno $\boldsymbol{x} \in \mathbb{R}^n$, pro které $\mathbb{A}\boldsymbol{x} = \boldsymbol{b}$.

  Nyní si všimneme souvislosti s Větou VI.15: Obsahem Věty VI.15 je platnost ekvivalencí

  $$(i') \Leftrightarrow (iii') \Leftrightarrow \mathbb{A} \text{ je regulární.}$$

  Protože implikace $(i') \Rightarrow (ii')$ je triviální, zbývá dokázat implikaci $(ii') \Rightarrow (i')$.
- Důkaz implikace $(ii') \Rightarrow (i')$:

  Můžeme dokázat obměnu, tj. $non(i') \Rightarrow non(ii')$. Jak je vysvětleno výše, díky Větě VI.15 stačí dokázat implikaci

  $\mathbb{A}$ není regulární $\Rightarrow non(ii')$.

  Nechť $\mathbb{A}$ není regulární. Pak ani $\mathbb{A}^T$ není regulární (Věta VI.4), tedy $h(\mathbb{A}^T) < n$ (Věta VI.7).

  Tedy řádky matice $\mathbb{A}^T$ jsou lineárně závislé.

  Proto sloupce matice $\mathbb{A}$ jsou též lineárně závislé (sloupce matice $\mathbb{A}$ jsou vlastně řádky matice $\mathbb{A}^T$).

Označme $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$ sloupce matice $\mathbb{A}$. To, že jsou lineárně závislé, znamená, že existuje jejich netriviální lineární kombinace, která je rovna nulovému vektoru. Tj., existují čísla $y_1, \ldots, y_n$, z nichž alespoň jedno je nenulové, která splňují

$$y_1\boldsymbol{a}_1 + y_2\boldsymbol{a}_2 + \cdots + y_n\boldsymbol{a}_n = \boldsymbol{o}.$$

Tato rovnost znamená

$$\mathbb{A}\boldsymbol{y} = \boldsymbol{o},$$

kde $\boldsymbol{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$. Přitom vektor $\boldsymbol{y}$ není nulový (aspoň jedna ze souřadnic není nulová).

Protože zřejmě $\mathbb{A}\boldsymbol{o} = \boldsymbol{o}$, máme dva různé vektory $\boldsymbol{x}$ splňující $\mathbb{A}\boldsymbol{x} = \boldsymbol{o}$, a to $\boldsymbol{y}$ a $\boldsymbol{o}$.

Proto tvrzení (ii') neplatí (pro $\boldsymbol{b} = \boldsymbol{o}$).

Tím je důkaz hotov.

- Aplikace na rozšíření Věty VI.15:

  Důsledkem Věty VI.19 je, že k Větě VI.15 lze přidat ještě jednu ekvivalentní podmínku:

  (iv) Pro každý vektor pravých stran $\boldsymbol{b}$ má soustava $\mathbb{A}\boldsymbol{x} = \boldsymbol{b}$ nejvýše jedno řešení.

  To je totiž přesně podmínka (ii') výše.

  Dokázali jsme navíc, že není-li $\mathbb{A}$ regulární, pak soustava $\mathbb{A}\boldsymbol{x} = \boldsymbol{o}$ (s nulovým vektorem pravých stran) má více různých řešení (a to $\boldsymbol{o}$ a pak ještě nějaký nenulový vektor).

K Větě VI.20:

- Důkaz:

  Důkaz je vpodstatě triviální. Předpoklady říkají, že

  $$f(\boldsymbol{x}) = \mathbb{A}\boldsymbol{x} \text{ pro } \boldsymbol{x} \in \mathbb{R}^n \quad \text{a} \quad g(\boldsymbol{u}) = \mathbb{B}\boldsymbol{u} \text{ pro } \boldsymbol{u} \in \mathbb{R}^m.$$

  Z těchto předpokladů a definice složeného zobrazení plyne, že pro $\boldsymbol{x} \in \mathbb{R}^n$ platí

  $$(g \circ f)(\boldsymbol{x}) = g(f(\boldsymbol{x})) = \mathbb{B}f(\boldsymbol{x}) = \mathbb{B}(\mathbb{A}\boldsymbol{x}) = (\mathbb{B}\mathbb{A})\boldsymbol{x}.$$

  Tedy matice $\mathbb{B}\mathbb{A}$ reprezentuje zobrazení $g \circ f$, a tedy toto zobrazení je lineární.

- Význam: Tato věta mimo jiné vysvětluje, proč je maticové násobení přirozenou operací – odpovídá skládání lineárních zobrazení.

- Aplikace Vět VI.20 a VI.19 na důkaz implikace

  $$\mathbb{B}\mathbb{A} = \mathbb{I} \Rightarrow \mathbb{A}\mathbb{B} = \mathbb{I}$$

  pro čtvercové matice řádu n.

  Nechť $\mathbb{A}$ a $\mathbb{B}$ jsou matice řádu n splňující $\mathbb{B}\mathbb{A} = \mathbb{I}$.

  Nechť f je zobrazení reprezentované maticí $\mathbb{A}$ a g je zobrazení reprezentované maticí $\mathbb{B}$.

  Podle Věty VI.20 je zobrazení $g \circ f$ reprezentované maticí $\mathbb{B}\mathbb{A} = \mathbb{I}$.

  Tedy $(g \circ f)(\boldsymbol{x}) = \mathbb{I}\boldsymbol{x} = \boldsymbol{x}$ pro $\boldsymbol{x} \in \mathbb{R}^n$. Tedy složené zobrazení $g \circ f$ je identické zobrazení, speciálně je prosté a na.

  Proto f je prosté a g je na. *(Rozmystete si to.)*

  Tedy podle Věty VI.19 jsou f a g prosté i na.

  Nyní ukážeme, že $f(g(\boldsymbol{x})) = \boldsymbol{x}$ pro každé $\boldsymbol{x} \in \mathbb{R}^n$:

  Nechť $\boldsymbol{x} \in \mathbb{R}^n$ je libovolné. Protože f je na, existuje $\boldsymbol{y} \in \mathbb{R}^n$, pro které $f(\boldsymbol{y}) = \boldsymbol{x}$. Pak

  $$f(g(\boldsymbol{x})) = f(g(f(\boldsymbol{y}))) = f(\boldsymbol{y}) = \boldsymbol{x},$$

kde v druhé rovnosti jsme použili skutečnost, že

$$g(f(\boldsymbol{y})) = \boldsymbol{y}.$$

Tedy $f \circ g$ je reprezentováno jednotkovou maticí $\mathbb{I}$. Podle Věty VI.20 je $f \circ g$ reprezentováno maticí $\mathbb{A}\mathbb{B}$. Protože reprezentující matice je jednoznačně určena, dostáváme $\mathbb{A}\mathbb{B} = \mathbb{I}$.

Tím je důkaz hotov.