K oddílu VI.2 (první část) – regulární matice, lineární kombinace, lineární nezávislost

K definici regulární matice a inverzní matice

- Motivace: Máme-li dána (reálná či komplexní) čísla a, b a chceme vyřešit rovnici $ax = b$ (s neznámou x), mohou nastat následující možnosti:

  $a = b = 0$ **:** V tomto případě je každé číslo x řešením.

  $a = 0$, $b \neq 0$ **:** V tomto případě rovnice nemá žádné řešení.

  $a \neq 0$ **:** V tomto případě má rovnice právě jedno řešení, které získáme tak, že rovnici vydělíme číslem a, tedy $x = \frac{b}{a}$.
  Přitom „vydělit rovnici číslem a“ znamená vynásobit ji převrácenou hodnotou (číslem $\frac{1}{a} = a^{-1}$), která existuje pro každé nenulové číslo (viz oddíl I.3, vlastnosti reálných čísel, osmá vlastnost první skupiny).

- Jedním z důvodů, proč se zabýváme maticemi, je skutečnost, že některé důležité úlohy lze převést na řešení maticové rovnice $\mathbb{A}\mathbb{X} = \mathbb{B}$, kde $\mathbb{A}$ a $\mathbb{B}$ jsou dané matice a $\mathbb{X}$ je neznámá matice. (Budeme se tím zabývat například v oddílu VI.4). Proto bychom v tomto případě potřebovali kritéria řešitelnosti a postup řešení podobně, jako máme pro rovnice s čísly. Je to ovšem složitější.

  Uvažujme maticovou rovnici $\mathbb{A}\mathbb{X} = \mathbb{B}$ v rámci čtvercových matic řádu 2. Pak mohou nastat například následující situace:

  $\mathbb{A}$ i $\mathbb{B}$ jsou nulové matice: V tomto případě je každá matice $\mathbb{X}$ (čtvercová řádu 2) řešením.

  $\mathbb{A}$ je nulová matice a $\mathbb{B}$ není nulová matice: V tomto případě rovnice nemá žádné řešení, protože levá strana je vždy nulová matice.

  $\mathbb{A} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, $\mathbb{B} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$ **:** V tomto případě nemá rovnice žádné řešení, protože levá strana má tvar

  $$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix} = \begin{pmatrix} x_{11} & x_{12} \\ 0 & 0 \end{pmatrix},$$

  a tedy se nemůže rovnat matici $\mathbb{B}$.

$\mathbb{A} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, $\mathbb{B} = \begin{pmatrix} 1 & 2 \\ 0 & 0 \end{pmatrix}$ **:** V tomto případě má levá strana opět tvar

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix} = \begin{pmatrix} x_{11} & x_{12} \\ 0 & 0 \end{pmatrix},$$

a tedy se rovná matici $\mathbb{B}$, právě když $x_{11} = 1$ a $x_{12} = 2$. Tedy množina všech řešení je tvořena maticemi

$$\begin{pmatrix} 1 & 2 \\ x_{21} & x_{22} \end{pmatrix}, \quad x_{21}, x_{22} \in \mathbb{R}.$$

$\mathbb{A} = \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}$ **:** V tomto případě má levá strana tvar

$$\begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix} = \begin{pmatrix} x_{11} & x_{12} \\ 3x_{21} & 3x_{22} \end{pmatrix},$$

a tedy rovnice má pro každou $\mathbb{B}$ právě jedno řešení

$$\mathbb{X} = \begin{pmatrix} b_{11} & b_{12} \\ \frac{1}{3}b_{21} & \frac{1}{3}b_{22} \end{pmatrix}.$$

Vidíme tedy, že situace je opravdu složitější:

- Může se stát, že neexistuje řešení i v případě, že $\mathbb{A}$ není nulová.
- Může se stát, že rovnice má nekonečně mnoho řešení a přitom ne každá matice je řešením.

- Abychom uměli zkoumat řešitelnost maticových rovnic, zavádí se pojem regulární matice a inverzní matice. Začněme s definicí inverzní matice:

  Nechť $\mathbb{A}$ je čtvercová matice řádu n. Pak čtvercovou matici $\mathbb{B}$ řádu n nazýváme inverzní maticí k $\mathbb{A}$, pokud $\mathbb{AB} = \mathbb{BA} = \mathbb{I}$.

  Vysvětlení a poznámky:

  - Inverzní matice je jakousi náhradou za převrácenou hodnotou čísla. Maticí sice nemůžeme dělit, ale můžeme násobit inverzní maticí. Máme-li maticovou rovnici $\mathbb{AX} = \mathbb{B}$ jako výše a $\mathbb{C}$ je inverzní matice k $\mathbb{A}$, pak můžeme rovnici zleva vynásobit $\mathbb{C}$ a dostameme

$$\mathbb{C}(\mathbb{AX}) = \mathbb{CB},$$

přičemž levou stranu lze upravit

$$\mathbb{C}(\mathbb{A}\mathbb{X}) = (\mathbb{C}\mathbb{A})\mathbb{X} = \mathbb{I}\mathbb{X} = \mathbb{X}.$$

Dostáváme tedy řešení $\mathbb{X} = \mathbb{C}\mathbb{B}$.

– Inverzní matice může být jen jedna. Kdyby totiž $\mathbb{B}$ a $\mathbb{C}$ byly dvě matice splňující $\mathbb{A}\mathbb{B} = \mathbb{B}\mathbb{A} = \mathbb{I}$ a $\mathbb{A}\mathbb{C} = \mathbb{C}\mathbb{A} = \mathbb{I}$, pak platí

$$\mathbb{B} = \mathbb{B}\mathbb{I} = \mathbb{B}(\mathbb{A}\mathbb{C}) = (\mathbb{B}\mathbb{A})\mathbb{C} = \mathbb{I}\mathbb{C} = \mathbb{C}.$$

Proto dává smysl inverzní matici k $\mathbb{A}$ označit $\mathbb{A}^{-1}$ (pokud existuje).

– Protože maticové násobení není komutativní, požadujeme v definici inverzní matice, aby platilo zároveň $\mathbb{A}\mathbb{B} = \mathbb{I}$ i $\mathbb{B}\mathbb{A} = \mathbb{I}$. Nicméně, jak se přesvědčíme v oddílu VI.5, stačila by i jedna z těchto rovností, druhá pak platí automaticky (tj. pokud $\mathbb{A}, \mathbb{B}$ jsou čtvercové matice téhož řádu, pak z rovnosti $\mathbb{A}\mathbb{B} = \mathbb{I}$ plyne rovnost $\mathbb{B}\mathbb{A} = \mathbb{I}$).

- Nechť $\mathbb{A}$ je čtvercová matice řádu n. Pak inverzní matice může a nemusí existovat. Například

$$\begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{3} \end{pmatrix}, \qquad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}^{-1} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$

zatímco inverzní matice k $\mathbb{A} = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$ neexistuje, protože pro každou čtvercovou matici $\mathbb{B}$ řádu 2 má součin $\mathbb{A}\mathbb{B}$ nulový druhý řádek, a tedy se nemůže rovnat jednotkové matici.

- Čtvercové matice, k nimž existuje inverzní matice, se nazývají regulární.

K Větě VI.4:

- Chceme-li dokázat, že nějaká matice je regulární, znamená to dokázat, že k ní existuje inverzní matice. Například tak, že inverzní matici najdeme.

- Dokázat, že $\mathbb{A}^{-1} = \mathbb{B}$ znamená dokázat, že $\mathbb{A}\mathbb{B} = \mathbb{B}\mathbb{A} = \mathbb{I}$.

- Důkaz bodu (a): Chceme dokázat, že $(\mathbb{A}^{-1})^{-1} = \mathbb{A}$. To jest, že platí

  $$\mathbb{A}^{-1}\mathbb{A} = \mathbb{A}\mathbb{A}^{-1} = \mathbb{I}.$$

  To ovšem platí, protože $\mathbb{A}^{-1}$ je inverzní matice k $\mathbb{A}$.

  Protože jsme dokázali, že k $\mathbb{A}^{-1}$ existuje inverzní matice, je $\mathbb{A}^{-1}$ regulární.

- Důkaz bodu (b): Dokažme uvedenou rovnost. K tomu potřebujeme spočítat

  $$\begin{aligned} \mathbb{A}^T(\mathbb{A}^{-1})^T &= (\mathbb{A}^{-1}\mathbb{A})^T = \mathbb{I}^T = \mathbb{I}, \\ (\mathbb{A}^{-1})^T\mathbb{A}^T &= (\mathbb{A}\mathbb{A}^{-1})^T = \mathbb{I}^T = \mathbb{I}. \end{aligned}$$

  Výpočty v obou řádcích jsou podobné. První rovnost v obou případech plyne z bodu (iii) Věty VI.3, druhá z definice inverzní matice. Třetí rovnost plyne z pozorování, že jednotková matice je symetrická.

  Dokázali jsme tedy uvedenou rovnost, tedy, že $\mathbb{A}^T$ má inverzní matici, a proto je regulární.

- Důkaz bodu (c): Opět dokážeme uvedenou rovnost pomocí následujícící výpočtů:

  $$\begin{aligned} (\mathbb{A}\mathbb{B})(\mathbb{B}^{-1}\mathbb{A}^{-1}) &= ((\mathbb{A}\mathbb{B})\mathbb{B}^{-1})\mathbb{A}^{-1} = (\mathbb{A}(\mathbb{B}\mathbb{B}^{-1}))\mathbb{A}^{-1} = (\mathbb{A}\mathbb{I})\mathbb{A}^{-1} \\ &= \mathbb{A}\mathbb{A}^{-1} = \mathbb{I}, \\ (\mathbb{B}^{-1}\mathbb{A}^{-1})(\mathbb{A}\mathbb{B}) &= (\mathbb{B}^{-1}\mathbb{A}^{-1})\mathbb{A})\mathbb{B} = (\mathbb{B}^{-1}(\mathbb{A}^{-1}\mathbb{A}))\mathbb{B} = (\mathbb{B}^{-1}\mathbb{I})\mathbb{B} \\ &= \mathbb{B}^{-1}\mathbb{B} = \mathbb{I}. \end{aligned}$$

  Ve výpočtu jsme použili asociativitu maticového násobení (Věta VI.2(i)), definici inverzní matice a vlastnosti jednotkové matice.

  Opět můžeme konstatovat, že jsme dokázali uvedenou rovnost, tedy, že $\mathbb{A}\mathbb{B}$ má inverzní matici, a proto je regulární.

Poznámka: Nyní víme, co je regulární a inverzní matice, dokázali jsme nějaké základní vlastnosti. Ale zatím nemáme žádnou metodu, jak zjistit, zda matice je regulární, ani metodu výpočtu inverzní matice. Tomu bude věnován zbytek tohoto oddílu. Proto teď zdánlivě změníme téma. (Přitom následující pojmy a věty se budou hodit i k dalším věcem.)

K definici lineární kombinace, lineární závislosti a nezávislosti

- Lineární kombinací rozumíme uvedený výraz. Je dána vektory $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m$ a koeficienty $\lambda_1, \ldots, \lambda_m$. Onen výraz zároveň můžeme spočítat a můžeme uvažovat jeho výsledek. Tím je opět nějaký vektor. Je potřebné rozlišovat mezi výrazem a jeho výsledkem, mezi lineární kombinací a vektorem, kterému se rovná. Například, máme-li vektory

$$(0\ \ 1),\ (1\ \ 0),\ (1\ \ 1),$$

můžeme uvažovat jejich různé lineární kombinace

$$\begin{aligned}
&1 \cdot (0\ \ 1) + 2 \cdot (1\ \ 0) + (-1) \cdot (1\ \ 1),\\
&0 \cdot (0\ \ 1) + 1 \cdot (1\ \ 0) + 0 \cdot (1\ \ 1),\\
&1 \cdot (0\ \ 1) + 1 \cdot (1\ \ 0) + (-1) \cdot (1\ \ 1),\\
&0 \cdot (0\ \ 1) + 0 \cdot (1\ \ 0) + 0 \cdot (1\ \ 1).
\end{aligned}$$

To jsou čtyři různé lineární kombinace, protože koeficienty jsou různé. Ale první dvě mají stejný výsledek, obě se rovnají vektoru $(1\ \ 0)$. Stejně tak poslední dvě se rovnají témuž vektoru, a to $(0\ \ 0)$.

- Triviální lineární kombinace je ta, jejíž všechny koeficienty jsou nulové. Jejím výsledkem je samozřejmě nulový vektor. Ve výše uvedeném příkladu je čtvrtá lineární kombinace triviální.

- Netriviální lineární kombinace je taková, že alespoň jeden z jejích koeficientů není nulový. Ve výše uvedeném příkladu první tři lineární kombinace jsou netriviální. Výsledkem netriviální lineární kombinace může být vektor nenulový (jako v prvních dvou případech) nebo nulový (jako ve třetím případě).

- Vektory $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m$ jsou lineárně závislé, pokud nějaká jejich netriviální lineární kombinace je rovna nulovému vektoru. Například výše uvedené tři vektory jsou lineárně závislé, protože třetí z lineárních kombinací je netriviální a je rovna nulovému vektoru.

- Pokud vektory $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m$ nejsou lineárně závislé, říkáme jim lineárně nezávislé. To znamená, že jediná jejich lineární kombinace, která je rovna nulovému vektoru, je triviální lineární kombinace. Například dvojice vektorů

$$(0\ \ 1),\ (1\ \ 0)$$

je lineárně nezávislá. Pokud máme totiž jejich lineární kombinaci s koeficienty λ_1, λ_2, pak

$$\lambda_1 \cdot (0 \;\; 1) + \lambda_2 \cdot (1 \;\; 0) = (\lambda_2 \;\; \lambda_1)$$

a jediný případ, kdy výsledkem je nulový vektor je $\lambda_1 = \lambda_2 = 0$. Podobě trojice vektorů

$$(0 \;\; 0 \;\; 1),\; (0 \;\; 1 \;\; 0),\; (1 \;\; 0 \;\; 0)$$

je lineárně nezávislá. Pokud máme totiž jejich lineární kombinaci s koeficienty $\lambda_1, \lambda_2, \lambda_3$, pak

$$\lambda_1 \cdot (0 \;\; 0 \;\; 1) + \lambda_2 \cdot (0 \;\; 1 \;\; 0) + \lambda_3 \cdot (1 \;\; 0 \;\; 0) = (\lambda_3 \;\; \lambda_2 \;\; \lambda_1)$$

a jediný případ, kdy výsledkem je nulový vektor je $\lambda_1 = \lambda_2 = \lambda_3 = 0$.

- Lineární závislost či nezávislost vektorů $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m$ je vlastnost této m-tice jako celku, nikoli jednotlivých vektorů nebo jejich dvojic. Například výše uvedená trojice vektorů

$$(0 \;\; 1),\; (1 \;\; 0),\; (1 \;\; 1)$$

je lineárně závislá (jak je vysvětleno výše). Ale každá dvojice těchto vektorů je lineárně nezávislá. Pro dvojici $(0 \;\; 1)$, $(1 \;\; 0)$ to bylo vysvětleno výše. Rozmysleme si to pro dvojici $(0 \;\; 1)$, $(1 \;\; 1)$. Uvažme lineární kombinaci s koeficienty λ_1, λ_2:

$$\lambda_1 \cdot (0 \;\; 1) + \lambda_2 \cdot (1 \;\; 1) = (\lambda_2 \;\; \lambda_1 + \lambda_2).$$

Aby vyšel nulový vektor, musí být $\lambda_2 = 0$ a $\lambda_1 + \lambda_2 = 0$. Je zřejmé, že jediná možnost je $\lambda_1 = \lambda_2 = 0$.

Cvičení:

1. *Ukažte, že jeden vektor je lineárně závislý, právě když je nulový.*

2. *Ukažte, že dvojice vektorů $\boldsymbol{v}_1, \boldsymbol{v}_2$ je lineárně závislá, právě když jeden z nich je násobkem druhého. (Tj. buď $\boldsymbol{v}_1$ je násobkem $\boldsymbol{v}_2$ nebo $\boldsymbol{v}_2$ je násobkem $\boldsymbol{v}_1$.)*

3. *Ukažte, že vektory $\boldsymbol{v}_1, \dots, \boldsymbol{v}_m$ jsou lineárně závislé, právě když jeden z nich lze vyjádřit jako lineární kombinaci ostatních, tj. existuje k a koeficienty $\lambda_1, \dots, \lambda_{k-1}, \lambda_{k+1}, \dots, \lambda_m$, že*

$$\boldsymbol{v}_k = \lambda_1 \boldsymbol{v}_1 + \dots + \lambda_{k-1}\boldsymbol{v}_{k-1} + \lambda_{k+1}\boldsymbol{v}_{k+1} + \dots \lambda_m \boldsymbol{v}_m.$$

4. *Pokud jeden z vektorů $\boldsymbol{v}_1, \dots, \boldsymbol{v}_m$ je nulový, ukažte, že jsou lineárně závislé.*

5. *Mějme vektory $\boldsymbol{v}_1, \dots, \boldsymbol{v}_m$. Pokud se dva z nich rovnají (tj. existují i, j, že $i \neq j$ a přitom $\boldsymbol{v}_i = \boldsymbol{v}_j$), ukažte, že tato m-tice je lineárně závislá.*

K definici hodnosti matice a schodovité matice

- To, že hodnost matice je rovna m znamená, že:
  - Mezi jejími řádky lze vybrat m-tici, která je lineárně nezávislá.
  - Pokud vybereme více než m řádků, už musí být lineárně závislé.
- Hodnost matice nemůže být větší než počet jejích řádků.
- Hodnost nulové matice je nula.
- Pokud $\mathbb{A}$ není nulová matice, je její hodnost aspoň jedna. (Existuje nenulový řádek, on sám je pak lineárně nezávislý).
- Příklady schodovitých matic:

$$\begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 7 \\ 0 & 0 & 0 & 5 \end{pmatrix}, \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 5 & 7 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 5 & 7 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

- Hodnost schodovité matice je rovna počtu nenulových řádků:

  $\leq$: Nechť $\mathbb{A}$ je (schodovitá) matice a m je počet jejích nenulových řádků. Pokud vybereme více než m řádků, jeden z nich musí být nulový, takže jsou lineárně závislé. Proto nemůže být více než m lineárně nezávislých řádků.

$\geq$ **:** Nechť $\mathbb{A}$ je schodovitá matice a m je počet jejích nenulových řádků. Musí to být prvních k řádků. Označme je $\boldsymbol{v}_1, \ldots \boldsymbol{v}_m$. Ukážeme, že jsou lineárně nezávislé. Mějme tedy lineární kombinaci

$$\lambda_1 \boldsymbol{v}_1 + \cdots + \lambda_m \boldsymbol{v}_m$$

a předpokládejme, že se rovná nulovému vektoru.

Podívejme se nyní, jak vypadá vektor $\boldsymbol{v}_j$. Víme, že to je j-tý řádek matice $\mathbb{A}$, tedy

$$\boldsymbol{v}_j = (a_{j1} \quad a_{j2} \quad \ldots \quad a_{jn}),$$

kde n je počet sloupců matice $\mathbb{A}$. Víme, že $\boldsymbol{v}_j$ není nulový vektor, a tedy má nějakou nenulovou složku. Vezměme první z nenulových složek, nechť je na místě k_j. To znamená, že

$$a_{j,k_j} \neq 0 \text{ a } a_{ji} = 0 \text{ pro } i < k_j.$$

Protože $\mathbb{A}$ je schodovitá, dostaneme

$$k_1 < k_2 < \cdots < k_m.$$

Připomeňme, že

$$\lambda_1 \boldsymbol{v}_1 + \cdots + \lambda_m \boldsymbol{v}_m = \boldsymbol{o}.$$

Dosaďme postupně složky $k_1, \ldots, k_m$. Po dosazení složky k_1 dostaneme

$$\lambda_1 a_{1,k_1} = 0,$$

protože další řádky mají ma místě k_1 nulu. Protože $a_{1,k_1} \neq 0$, dostaneme $\lambda_1 = 0$.

Dosadíme-li složku k_2, dostaneme

$$\lambda_1 a_{1,k_2} + \lambda_2 a_{2,k_2} = 0,$$

protože další řádky mají ma místě k_2 nulu. Protože $\lambda_1 = 0$ a $a_{2,k_2} \neq 0$, dostaneme $\lambda_2 = 0$.

Takto pokračujeme dále a postupně dostaneme, že všechny koeficienty jsou nulové. To znamená, že vektory jsou lineárně nezávislé, a tedy hodnost maticem $\mathbb{A}$ je alespoň m.

Ilustrace předchozího postupu na příkladu:
Uvažme matici

$$\mathbb{A} = \begin{pmatrix} 1 & 1 & 0 & 0 & 7 \\ 0 & 3 & -1 & 7 & 2 \\ 0 & 0 & 0 & 5 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

To je schodovitá matice a první tři řádky jsou lineárně nezávislé:
Předpokládejme, že

$$\lambda_1 \cdot (\text{řádek } 1) + \lambda_2 \cdot (\text{řádek } 2) + \lambda_3 \cdot (\text{řádek } 3) = \boldsymbol{o}.$$

Ve značení z předchozího důkazu máme $k_1 = 1$, $k_2 = 2$, $k_3 = 4$.
Dosazením postupně první, druhé a čtvrté složky dostaneme

$$\begin{aligned} \lambda_1 \cdot 1 &= 0, \\ \lambda_1 \cdot 1 + \lambda_2 \cdot 3 &= 0, \\ \lambda_1 \cdot 0 + \lambda_2 \cdot 7 + \lambda_3 \cdot 5 &= 0. \end{aligned}$$

Z toho snadno plyne, že $\lambda_1 = \lambda_2 = \lambda_3 = 0$.