Optimalizace

2. Vybraná témata z lineární algebry

Tomáš Kroupa

FEL ČVUT

Linearita

Základní koncepty a fakta

- Lineární prostor $\mathbb{R}^n$ a lineární podprostory $X \subseteq \mathbb{R}^n$
- Vektory $\mathbf{x} \in \mathbb{R}^n$ (sloupcové!) a matice $\mathbf{A} \in \mathbb{R}^{m \times n}$
- Elementární vektorové/maticové operace
- Pojem báze a dimenze
- Hodnost matice $\operatorname{rank} \mathbf{A}$
- Lineární zobrazení $\mathbf{f}\colon \mathbb{R}^n \to \mathbb{R}^m$ je vyjádřitelné jako

  $$\mathbf{f}(\mathbf{x}) = \mathbf{A}\mathbf{x}$$

  pro nějakou matici $\mathbf{A} \in \mathbb{R}^{m \times n}$ (v závislosti na volbě báze!)
- Struktura řešení soustavy lineárních rovnic $\mathbf{A}\mathbf{x} = \mathbf{b}$

Sloupcový a řádkový pohled na součin Ax

Vyjádříme $\mathbf{A} \in \mathbb{R}^{m \times n}$ prostřednictvím sloupců $\mathbf{a}_1, \ldots, \mathbf{a}_n$:

$$\mathbf{Ax} = \begin{bmatrix} \mathbf{a}_1 \ldots \mathbf{a}_n \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n$$

Vyjádříme $\mathbf{A} \in \mathbb{R}^{m \times n}$ prostřednictvím řádků $\mathbf{a}_1^\mathsf{T}, \ldots, \mathbf{a}_m^\mathsf{T}$:

$$\mathbf{Ax} = \begin{bmatrix} \mathbf{a}_1^\mathsf{T} \\ \vdots \\ \mathbf{a}_m^\mathsf{T} \end{bmatrix} \mathbf{x} = \begin{bmatrix} \mathbf{a}_1^\mathsf{T}\mathbf{x} \\ \vdots \\ \mathbf{a}_m^\mathsf{T}\mathbf{x} \end{bmatrix}$$

Řešení soustavy lineárních rovnic $\mathbf{A}\mathbf{x} = \mathbf{b}$ geometricky

Sloupcově

Hledáme vektor $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$ splňující

$$x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n = \mathbf{b}$$

Řádkově

Hledáme vektor $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$ ležící v průniku

$$\{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{a}_1^\mathsf{T}\mathbf{x} = b_1\} \cap \cdots \cap \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{a}_m^\mathsf{T}\mathbf{x} = b_m\}$$

Prostor obrazů matice $\mathbf{A} \in \mathbb{R}^{m \times n}$

$$\operatorname{rng} \mathbf{A} := \{\mathbf{A}\mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n\} \quad \subseteq \mathbb{R}^m$$

Interpretace

- Obor hodnot (range, image) lineárního zobrazení $\mathbf{f}(\mathbf{x}) = \mathbf{A}\mathbf{x}$
- Množina všech vektorů $\mathbf{y}$ takových, že $\mathbf{A}\mathbf{x} = \mathbf{y}$ má řešení
- Podle sloupcového pohledu na součin $\mathbf{A}\mathbf{x}$ je to lineární obal sloupců matice $\mathbf{A}$

Hodnost matice $\mathbf{A}$ je číslo

$$\operatorname{rank} \mathbf{A} := \dim \operatorname{rng} \mathbf{A}.$$

Platí $\operatorname{rank} \mathbf{A} = \operatorname{rank} \mathbf{A}^\mathsf{T}$.

Nulový prostor matice $\mathbf{A} \in \mathbb{R}^{m \times n}$

$$\operatorname{null} \mathbf{A} := \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{A}\mathbf{x} = \mathbf{0}\} \quad \subseteq \mathbb{R}^n$$

Interpretace

- Množina vektorů, které se zobrazí na nulový vektor
- Množina řešení soustavy lineárních rovnic $\mathbf{A}\mathbf{x} = \mathbf{0}$
- Podle řádkového pohledu na součin $\mathbf{A}\mathbf{x}$ je $\operatorname{null} \mathbf{A}$ množina všech vektorů kolmých na každý řádek matice $\mathbf{A}$

Jaký je vztah mezi $\operatorname{rng} \mathbf{A}$ a $\operatorname{null} \mathbf{A}$?

Existence/jednoznačnost řešení soustav lineárních rovnic

Pro matici $\mathbf{A} \in \mathbb{R}^{m \times n}$ jsou tvrzení pod sebou ekvivalentní:

Prostor obrazů

1. $\operatorname{rng} \mathbf{A} = \mathbb{R}^m$
2. $\mathbf{A}\mathbf{x} = \mathbf{y}$ má řešení $\forall \mathbf{y}$
3. $\operatorname{rank} \mathbf{A} = m$
4. $\mathbf{A}$ má lin. nezávislé řádky
5. $\mathbf{A}$ má pravou inverzi
6. $\mathbf{A}\mathbf{A}^\mathsf{T}$ je regulární

Nulový prostor

1. $\operatorname{null} \mathbf{A} = \{\mathbf{0}\}$
2. $\mathbf{A}\mathbf{x} = \mathbf{0}$ řeší jen $\mathbf{x} = \mathbf{0}$
3. $\operatorname{rank} \mathbf{A} = n$
4. $\mathbf{A}$ má lin. nezávislé sloupce
5. $\mathbf{A}$ má levou inverzi
6. $\mathbf{A}^\mathsf{T}\mathbf{A}$ je regulární

Dimenze fundamentálních maticových podprostorů

Věta

Pro každou matici $\mathbf{A} \in \mathbb{R}^{m \times n}$ platí

$$\dim \operatorname{rng} \mathbf{A} + \dim \operatorname{null} \mathbf{A} = n.$$

- Protože $\operatorname{rank} \mathbf{A} = \operatorname{rank} \mathbf{A}^\mathsf{T}$, platí také

$$\dim \operatorname{rng} \mathbf{A}^\mathsf{T} + \dim \operatorname{null} \mathbf{A} = n$$

- Dimenze nulového prostoru popisuje míru degenerace lineárního zobrazení $\mathbf{f}(\mathbf{x}) = \mathbf{A}\mathbf{x}$
- Počet lineárně nezávislých řešení soustavy lineárních rovnic $\mathbf{A}\mathbf{x} = \mathbf{0}$ je $n - \operatorname{rank} \mathbf{A}$

Rank factorization

Rozklad matice podle hodnosti

Pro každou matici $\mathbf{A} \in \mathbb{R}^{m \times n}$ hodnosti r existují matice $\mathbf{B} \in \mathbb{R}^{m \times r}$ a $\mathbf{C} \in \mathbb{R}^{r \times n}$ splňující

$$\mathbf{A} = \mathbf{B}\mathbf{C}.$$

- Úspora paměti pro $r \ll \min\{m, n\}$
- Rozklad není jednoznačný
- Užitím tohoto vztahu lze dokázat, že $\operatorname{rank} \mathbf{A} = \operatorname{rank} \mathbf{A}^{\mathsf{T}}$

Afinní podprostor

Afinní kombinace vektorů $\mathbf{x}_1, \ldots, \mathbf{x}_k \in \mathbb{R}^n$ je

$$\alpha_1 \mathbf{x}_1 + \cdots + \alpha_k \mathbf{x}_k, \qquad \text{kde } \alpha_1 + \cdots + \alpha_k = 1.$$

Afinní obal vektorů $\mathbf{x}_1, \ldots, \mathbf{x}_k \in \mathbb{R}^n$ je množina

$$\{\alpha_1 \mathbf{x}_1 + \cdots + \alpha_k \mathbf{x}_k \mid \alpha_1 + \cdots + \alpha_k = 1,\ \ \alpha_i \in \mathbb{R}, i = 1, \ldots, k\}.$$

Definice

Afinní podprostor je množina $Y \subseteq \mathbb{R}^n$ splňující

$$\left[\alpha_1, \ldots, \alpha_k \in \mathbb{R},\ \mathbf{x}_1, \ldots, \mathbf{x}_k \in Y,\ \sum_{i=1}^{k} \alpha_i = 1\right] \ \Rightarrow\ \sum_{i=1}^{k} \alpha_i \mathbf{x}_i \in Y.$$

Příklady afinních podprostorů

- Bod, přímka, rovina, nadrovina v $\mathbb{R}^n$
- Je-li X lineární podprostor $\mathbb{R}^n$ a $\mathbf{x}_0 \in \mathbb{R}^n$, pak je množina

$$X + \mathbf{x}_0 := \{\mathbf{x} + \mathbf{x}_0 \mid \mathbf{x} \in X\}$$

  afinní podprostor $\mathbb{R}^n$

Nejobecnější příklad afinního podprostoru

Pro každý afinní podprostor $Y \neq \emptyset$ existuje jediný lineární podprostor X a nějaký vektor $\mathbf{x}_0 \in Y$ splňující

$$Y = X + \mathbf{x}_0$$

Tvrzení

Nechť $Y \subseteq \mathbb{R}^n$. Následující tvrzení jsou ekvivalentní.

1. Y je afinní podprostor.
2. Y je množinou řešení soustavy lineárních rovnic,

$$Y = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{A}\mathbf{x} = \mathbf{b}\},$$

   pro nějakou matici $\mathbf{A}$ a vektor $\mathbf{b}$.

Afinní zobrazení

Zobrazení $\mathbf{f}\colon \mathbb{R}^n \to \mathbb{R}^m$ je afinní, pokud

$$\mathbf{f}(\alpha_1\mathbf{x}_1 + \cdots + \alpha_k\mathbf{x}_k) = \alpha_1\mathbf{f}(\mathbf{x}_1) + \cdots + \alpha_k\mathbf{f}(\mathbf{x}_k)$$

platí pro všechna $\mathbf{x}_1, \ldots, \mathbf{x}_k \in \mathbb{R}^n$, $\alpha_1, \ldots, \alpha_k \in \mathbb{R}$, $\sum_{i=1}^k \alpha_i = 1$.

Tvrzení

Nechť $\mathbf{f}\colon \mathbb{R}^n \to \mathbb{R}^m$. Následující výroky jsou ekvivalentní.

1. Zobrazení $\mathbf{f}$ je afinní.
2. Existuje matice $\mathbf{A} \in \mathbb{R}^{m\times n}$ a vektor $\mathbf{b} \in \mathbb{R}^m$ tak, že

$$\mathbf{f}(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{b}.$$

Analogie

Lineární	**Afinní**
lineární kombinace	afinní kombinace
lineární obal	afinní obal
lineární podprostor X	afinní podprostor $X + \mathbf{x}_0$
řešení soustavy $\mathbf{Ax} = \mathbf{0}$	řešení soustavy $\mathbf{Ax} = \mathbf{b}$
$\operatorname{null} \mathbf{A}$	$\mathbf{x}_0 + \operatorname{null} \mathbf{A}$
lineární zobrazení $\mathbf{Ax}$	afinní zobrazení $\mathbf{Ax} + \mathbf{b}$

Ortogonalita

Eukleidovský prostor $\mathbb{R}^n$

- Standardní skalární součin $\mathbf{x}^\mathsf{T}\mathbf{y} = x_1y_1 + \cdots + x_ny_n$
- Eukleidovská norma $\|\mathbf{x}\| = \sqrt{\mathbf{x}^T\mathbf{x}} = \sqrt{x_1^2 + \cdots + x_n^2}$
- Eukleidovská metrika $d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|$

Definice

Ortogonální vektory $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ splňují $\mathbf{x}^\mathsf{T}\mathbf{y} = 0$. Píšeme

$$\mathbf{x} \perp \mathbf{y}.$$

Pro ortogonální vektory $\mathbf{x}$ a $\mathbf{y}$ platí Pythagorova věta:

$$\|\mathbf{x} + \mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2$$

Motivační úloha

Pro zadaný lineární podprostor $X \subseteq \mathbb{R}^n$ a vektor $\mathbf{z} \in \mathbb{R}^n$ minimalizuj funkci $d(\mathbf{z}, \mathbf{x}) = \|\mathbf{z} - \mathbf{x}\|$ za podmínky $\mathbf{x} \in X$.

Řešení

- Existuje jediný vektor $\mathbf{x}$ minimalizující vzdálenost
- Pro libovolné $\mathbf{z} \in \mathbb{R}^n$ tak lze psát $\mathbf{x} := \mathbf{f}(\mathbf{z})$, kde $\mathbf{f}\colon \mathbb{R}^n \to \mathbb{R}^n$
- Zobrazení $\mathbf{f}$ je lineární a jeho matice $\mathbf{P}$ má tyto vlastnosti:
  - $\mathbf{P} = \mathbf{P}^2$
  - $\mathbf{P} = \mathbf{P}^\mathsf{T}$

Ortogonalita podprostorů

- Vektor $\mathbf{y} \in R^n$ je ortogonální na podprostor $X \subseteq \mathbb{R}^n$, pokud $\mathbf{y} \perp \mathbf{x}$ pro všechna $\mathbf{x} \in X$. Píšeme

$$\mathbf{y} \perp X$$

- Ortogonální podprostory $X, Y \subseteq \mathbb{R}^n$ splňují $\mathbf{y} \perp X$ pro všechna $\mathbf{y} \in Y$. Píšeme

$$X \perp Y$$

- Ortogonální doplněk podprostoru $X \subseteq \mathbb{R}^n$ je množina

$$X^\perp := \{\mathbf{y} \in \mathbb{R}^n \mid \mathbf{y} \perp X\}$$

Ortogonální doplňky $\operatorname{null} \mathbf{A}$ a $\operatorname{rng} \mathbf{A}$

Povšimněme si:

$$\operatorname{null} \mathbf{A} = (\operatorname{rng} \mathbf{A}^\mathsf{T})^\perp$$

Tvrzení

Pro každou matici $\mathbf{A}$ platí:

$$(\operatorname{null} \mathbf{A})^\perp = \operatorname{rng} \mathbf{A}^\mathsf{T}$$
$$(\operatorname{rng} \mathbf{A})^\perp = \operatorname{null} \mathbf{A}^\mathsf{T}$$

Ortonormální množina vektorů

Ortonormální množina vektorů $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ splňuje

$$\mathbf{u}_i^\top \mathbf{u}_j = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases}$$

Důležité postřehy

- Ortonormální množina je lineárně nezávislá
- i-tá souřadnice vektoru $\mathbf{x} \in \operatorname{span}\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ je $\mathbf{u}_i^\top \mathbf{x}$, tedy

$$\mathbf{x} = (\mathbf{u}_1^\top \mathbf{x})\mathbf{u}_1 + \cdots + (\mathbf{u}_n^\top \mathbf{x})\mathbf{u}_n$$

Matice s ortonormálními sloupci

- Matice $\mathbf{U} \in \mathbb{R}^{m \times n}$ s ortonormálními sloupci splňuje $\mathbf{U}^\mathsf{T}\mathbf{U} = \mathbf{I}$
- Nutně $m \geq n$
- Lineární zobrazení $\mathbf{f}(\mathbf{x}) := \mathbf{U}\mathbf{x}$ zachovává skalární součin

Definice

Ortogonální matice je matice $\mathbf{U} \in \mathbb{R}^{n \times n}$ s ortonormálními sloupci.

Ekvivalentní podmínky:

- $\mathbf{U}^\mathsf{T} = \mathbf{U}^{-1}$
- $\mathbf{U}\mathbf{U}^\mathsf{T} = \mathbf{I}$

Ortogonální matice - příklady

Rotační matice v $\mathbb{R}^2$

Rotace vektoru kolem počátku o úhel φ proti směru hodinových ručiček je vyjádřeno maticí

$$\begin{bmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{bmatrix}$$

Householderova matice

Zrcadlení podle nadroviny procházející počátkem s normálovým vektorem $\mathbf{u} \in \mathbb{R}^n$, kde $\|\mathbf{u}\| = 1$, reprezentuje matice

$$\mathbf{H}_\mathbf{u} = \mathbf{I} - 2\mathbf{u}\mathbf{u}^\mathsf{T}$$

Ortogonální projekce

Definice

Ortogonální projekce vektoru $\mathbf{z} \in \mathbb{R}^m$ na lineární podprostor $X \subseteq \mathbb{R}^m$ je vektor $\mathbf{x} \in X$ takový, že $(\mathbf{z} - \mathbf{x}) \perp X$.

Každý lineární podprostor $X \subseteq \mathbb{R}^m$ lze vyjádřit ve tvaru $X = \operatorname{rng} \mathbf{U}$ pro nějakou matici $\mathbf{U}$ s ortonormálními sloupci.

Věta

Nechť má matice $\mathbf{U} \in \mathbb{R}^{m \times n}$ ortonormální sloupce $\mathbf{u}_1, \ldots, \mathbf{u}_n$. Ortogonální projekce $\mathbf{z} \in \mathbb{R}^m$ na $\operatorname{rng} \mathbf{U}$ je vektor

$$\mathbf{x} = \mathbf{U}\mathbf{U}^\top \mathbf{z} = (\mathbf{u}_1^\top \mathbf{z})\mathbf{u}_1 + \cdots + (\mathbf{u}_n^\top \mathbf{z})\mathbf{u}_n$$

Ortogonální projekce minimalizuje vzdálenost

Věta o kolmici

Ortogonální projekce $\mathbf{x} \in X$ vektoru $\mathbf{z} \in \mathbb{R}^m$ na lineární podprostor $X \subseteq \mathbb{R}^m$ splňuje nerovnost

$$\|\mathbf{z} - \mathbf{x}\| < \|\mathbf{z} - \mathbf{x}'\| \qquad \forall \mathbf{x}' \in X, \quad \mathbf{x}' \neq \mathbf{x}$$

Věta ukazuje, že jediné řešení původní motivační úlohy je ve tvaru ortogonální projekce. Hodnota v bodě optima $\mathbf{x} \in X$ je vzdálenost bodu $\mathbf{z}$ od podprostoru X:

$$\|\mathbf{z} - \mathbf{x}\| = \|\mathbf{z} - \mathbf{U}\mathbf{U}^\mathsf{T}\mathbf{z}\|$$

Vlastnosti ortogonálního projektoru $\mathbf{P} := \mathbf{U}\mathbf{U}^\mathsf{T}$ na X

Charakterizující vlastnosti

$$\mathbf{P} = \mathbf{P}^2 = \mathbf{P}^\mathsf{T}$$

Další vlastnosti

- $\operatorname{rng} \mathbf{P} = X$ a $\operatorname{null} \mathbf{P} = X^\perp$
- Matice $\mathbf{I} - \mathbf{P}$ je ortogonální projektor na $X^\perp$
- Vzdálenost bodu $\mathbf{z}$ od podprostoru $X^\perp$ je

$$\|\mathbf{P}\mathbf{z}\| = \|\mathbf{U}^T\mathbf{z}\|$$

QR rozklad

Věta

Pro každou matici $\mathbf{A} \in \mathbb{R}^{m \times n}$ existuje ortogonální matice $\mathbf{Q} \in \mathbb{R}^{m \times m}$ a horní trojúhelníková matice $\mathbf{R} \in \mathbb{R}^{m \times n}$ splňující

$$\mathbf{A} = \mathbf{QR}.$$

Redukovaný QR rozklad matice $\mathbf{A}$ definujeme pro $m > n$ takto:

- Čtvercová matice $\mathbf{R} \in \mathbb{R}^{n \times n}$ je horní trojúhelníková
- Matice $\mathbf{Q} \in \mathbb{R}^{m \times n}$ má ortonormální sloupce