Afinní podprostory, ortogonalita: úvod

Petr Olšák

petr@olsak.net

http://petr.olsak.net/

Afinní kombinace a odvozené pojmy

- Afinní kombinace je lineární kombinace prvků z $\mathbb{R}^n$, jejíž koeficienty mají součet 1, tedy:

$$\mathrm{AK}(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k) = \alpha_1 \mathbf{x}_1 + \alpha_2 \mathbf{x}_2 + \cdots + \alpha_k \mathbf{x}_k, \quad \text{kde } \sum_{i=1}^{k} \alpha_i = 1$$

- V kontextu afinních kombinací je užitečné si při geometrické interpreatači prvky z $\mathbb{R}^n$ (vektory) představovat jako body v prostoru, ne jako orientované úsečky. AK totiž nezávisí na poloze počátku.
- Afinní obal zadaných bodů $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k$ je množina všech afinních kombinací těchto bodů.
- Afinní podprostor je podmnožina $M \subseteq \mathbb{R}^n$ uzavřená na afinní kombinace. Tedy je-li $\mathbf{x}_i \in M$, pak také všechny $\mathrm{AK}(\mathbf{x}_i) \in M$.
- Afinní zobrazení je zobrazení $f\colon \mathbb{R}^n \to \mathbb{R}^m$ zachovávající afinní kombinace, tedy pro $\sum_{i=1}^{k} \alpha_i = 1$ platí

$$\mathbf{f}(\alpha_1 \mathbf{x}_1 + \alpha_2 \mathbf{x}_2 + \cdots + \alpha_k \mathbf{x}_k) = \alpha_1 \mathbf{f}(\mathbf{x}_1) + \alpha_2 \mathbf{f}(\mathbf{x}_2) + \cdots + \alpha_k \mathbf{f}(\mathbf{x}_k)$$

- Analogie: lineární kombinace, lineární obal, podprostor, zobrazení.

Afinní kombinace a podprostory: vlastnosti

- Afinní kombinace afinních kombinací je afinní kombinace.
- Afinní obal je afinní podprostor. Každý afinní podprostor lze zapsat jako afinní obal nějakých bodů.
- Množina řešení soustavy $\mathbf{Ax} = \mathbf{b}$ je prázdná nebo to je afinní podprostor.
- Každý afinní podprostor lze zapsat jako součet jednoho vektoru s vektory nějakého lineárního prostoru, přesněji: je-li $A \subseteq \mathbb{R}^n$ afinní prostor, pak existuje $\mathbf{x}_0 \in \mathbb{R}^n$ a lineární podprostor $X \subseteq \mathbb{R}^n$ tak, že

$$A = \{\mathbf{x}_0 + \mathbf{x};\ \mathbf{x} \in X\} = \mathbf{x}_0 + X.$$

- Ke každému afinnímu podprostoru A existuje soustava $\mathbf{Ax} = \mathbf{b}$, jejíž množina řešení je rovna A.
- Afinní podprostor je z geometrického pohledu lineární podprostor plus posun.
- Dimenze afinního prostoru se definuje jako dimenze příslušného posunutého lineárního podprostoru.

Afinní zobrazení: vlastnosti

- Zobrazení $\mathbf{f}$ definované vztahem $\mathbf{f}(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{b}$ je afinní.
- Ke každému afinnímu zobrazení $\mathbf{f}\colon \mathbb{R}^n \to \mathbb{R}^m$ existuje jednoznačně lineární zobrazení $\mathbf{g}\colon \mathbb{R}^n \to \mathbb{R}^m$ a vektor $\mathbf{b} \in \mathbb{R}^m$ tak, že $\mathbf{f}(\mathbf{x}) = \mathbf{g}(\mathbf{x}) + \mathbf{b}$ pro všechna $\mathbf{x} \in \mathbb{R}^n$.
- Ke každému afinnímu zobrazení $\mathbf{f}$ existuje jednoznačně matice $\mathbf{A}$ a vektor $\mathbf{b}$ tak, že $\mathbf{f}(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{b}$.
- Afinní zobrazení je z geometrického pohledu lineární zobrazení plus konstantní posun, například rotace a posun, zkosení a posun atd.
- Hodnota posunu je rovna $\mathbf{f}(\mathbf{0})$.
- Afinní zobrazení zobrazuje afinní podprostory na afinní podporostory.

Příklad

Zkusíme si geometricky znázornit, jak „pracuje“ zobrazení $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ definované předpisem:

$$\mathbf{f}(x_1, x_2) = \begin{bmatrix} 2 & 1/2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 3 \\ 2 \end{bmatrix}$$

Různá vyjádření afinních prostorů

Od rovnicového popisu k parametrickému

- Vyjádříme afinní podprostor $A = \{\mathbf{x};\ \mathbf{A}\mathbf{x} = \mathbf{b}\}$ parametricky, tedy jako $A = \mathbf{x}_0 + \mathrm{Span}\{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_k\}$.

  Řešení: najdeme bázi $\mathbf{b}_1, \dots, \mathbf{b}_k$ řešení přidružené homogenní soustavy $\mathbf{A}\mathbf{x} = \mathbf{0}$ a jedno partikulární řešení $\mathbf{x}_0$.

Od parametrického popisu k rovnicovému

- Pro afinní prostor zadaný parametricky $A = \mathbf{x}_0 + \mathrm{Span}\{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_k\}$ najdeme matici $\mathbf{A}$ a vektor $\mathbf{b}$ tak, že $A = \{\mathbf{x};\ \mathbf{A}\mathbf{x} = \mathbf{b}\}$.

  Řešení: zapíšeme vektory $\mathbf{b}_i$ do řádků matice $\mathbf{B}$ a najdeme bázi prostoru řešení soustavy $\mathbf{B}\mathbf{x} = \mathbf{0}$. Tyto bázové vektory zapíšeme do řádků matice $\mathbf{A}$. Pak je $\mathrm{Span}\{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_k\} = \{\mathbf{x};\ \mathbf{A}\mathbf{x} = \mathbf{0}\}$. Dále volíme $\mathbf{b} = \mathbf{A}\mathbf{x}_0$.

Skalární součin: úhly, velikosti

- Skalární součin $\mathbf{x}^T\mathbf{y}$ umožňuje měřit vzdálenosti a úhly:

$$\begin{aligned} \text{Euklidovská norma:} \quad & \|\mathbf{x}\| = \sqrt{\mathbf{x}^T\mathbf{x}} \\ \text{Euklidovská metrika (vzdálenost):} \quad & d(x, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\| \\ \text{Úhel:} \quad & \cos\varphi = \frac{\mathbf{x}^T\mathbf{y}}{\|\mathbf{x}\|\,\|\mathbf{y}\|} \\ \text{Pravý úhel:} \quad & \mathbf{x}^T\mathbf{y} = 0 \quad (\text{pro } \mathbf{x}, \mathbf{y} \neq \mathbf{0}) \end{aligned}$$

- Skalární součin geometricky: číslo $\|x\| \cos\varphi$ je orientovaná velikost průmětu vektoru $\mathbf{x}$ na přímku generovanou vektorem $\mathbf{y}$. Vynásobíme-li toto číslo velikostí $\|\mathbf{y}\|$, máme skalární součin:

$$\mathbf{x}^T\mathbf{y} = \|\mathbf{x}\|\,\|\mathbf{y}\|\,\cos\varphi.$$

- Zakreslíme-li vektory korespondující s $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n \in \mathbb{R}^n$ jako šipky na sebe kolmé a s jednotkovou velikostí, pak výše uvedené pojmy (úhel, velikost vektoru, vzdálenost) zavedené v $\mathbb{R}^n$ „numericky“ mají přesně odpovídající geometrický význam.

Připomenutí

- Schwartzova nerovnost: $|\mathbf{x}^T\mathbf{y}| \leq \|\mathbf{x}\|\,\|\mathbf{y}\|$, takže úhel je definován pro každé dva nenulové vektory. Rovnost nastává jen pro LZ vektory.
- Trojúhelníková nerovnost: $\|\mathbf{x}+\mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$.
- Příklad odvození Pythagorovy věty (předpokládáme $\mathbf{x} \perp \mathbf{y}$):

$$\|\mathbf{x}-\mathbf{y}\|^2 = (\mathbf{x}-\mathbf{y})^T(\mathbf{x}-\mathbf{y}) = \mathbf{x}^T\mathbf{x} - 2\mathbf{x}^T\mathbf{y} + \mathbf{y}^T\mathbf{y} = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2.$$

Terminologie, základní vlastnosti

- Vektory jsou ortogonální (píšeme $\mathbf{x} \perp \mathbf{y}$), když $\mathbf{x}^T\mathbf{y} = 0$.
- Vektory jsou ortonormální, jsou-li ortogonální a mají jednotkovou velikost.
- Matice je s ortonormálními sloupci, jsou-li každé dva různé sloupce této matice ortonormální.
- Matice je ortogonální, je-li čtvercová a s ortonormálními sloupci.
- **Tvrzení:** Skupina nenulových vektorů, kde je každý s každým ortogonální, je lineárně nezávislá.
- **Důsledek:** Matice s ortonormálními sloupci je úzká nebo čtvercová a vždy s plnou hodností. Ortogonální matice je regulární.
- **Tvrzení:** Nechť $\mathbf{U}$ je s ortonormálními sloupci, $\mathbf{x} \in \operatorname{rng} \mathbf{U}$. Tedy sloupce $\mathbf{U}$ (označíme je $\mathbf{u}_1, \ldots, \mathbf{u}_k$) tvoří ortonormální bázi podprostoru $\operatorname{rng} \mathbf{U}$. Pak souřadnice vektoru $\mathbf{x}$ vzhledem k této bázi jsou $\mathbf{u}_i^T\mathbf{x}$, tedy

$$\mathbf{x} = (\mathbf{u}_1^T\mathbf{x})\,\mathbf{u}_1 + (\mathbf{u}_2^T\mathbf{x})\,\mathbf{u}_2 + \cdots + (\mathbf{u}_k^T\mathbf{x})\,\mathbf{u}_k.$$

- **Tvrzení:** Matice $\mathbf{U}$ je s ortonormálními sloupci, právě když $\mathbf{U}^T\mathbf{U} = \mathbf{I}$.

Vlastnosti ortogonální matice

- Následující vlastnosti jsou ekvivalentní.
  - $\mathbf{U}$ je ortogonální matice.
  - $\mathbf{U}^{-1} = \mathbf{U}^T$
  - $\mathbf{U}^T$ je ortogonální matice.

  Vidíme tedy, že jakmile je matice ortogonální, má ortonormální nejen sloupce, ale i řádky.
- **Tvrzení:** Součin ortogonálních matic je ortogonální matice.

Příklady ortogonálních matic

- Rotační matice v $\mathbb{R}^2$.
- Matice $\mathbf{P}$ obsahuje ve sloupcích bázové vektory $\mathbf{e}_i$ v libovolném pořadí: permutační matice. Pak vektor $\mathbf{Px}$ má stejné složky jako vektor $\mathbf{x}$, jen v permutovaném pořadí.
- Householderova matice je definována pro libovolný vektor $\mathbf{u}$ jednotkové velikosti takto: $\mathbf{H_u} = \mathbf{I} - 2\mathbf{uu}^T$. Je ortogonální.

Lineární isometrie

Nechť $\mathbf{U}$ je matice s ortonormálními sloupci. Definujme lineární zobrazení $\mathbf{f}\colon \mathbb{R}^n \to \mathbb{R}^m$ předpisem $\mathbf{f}(\mathbf{x}) = \mathbf{U}\mathbf{x}$. Pak platí:

- $\mathbf{f}$ zachovává skalární součin, přesněji: $\mathbf{x}^T\mathbf{y} = \mathbf{f}(\mathbf{x})^T\mathbf{f}(\mathbf{y})$.
- $\mathbf{f}$ zachovává velikosti vektorů (velikost vzoru je rovna velikosti obrazu).
- $\mathbf{f}$ zachovává úhly (úhel mezi vzory je stejný jako mezi jejich obrazy).

Takové zobrazení $\mathbf{f}$ se nazývá isometrie.

Isometrie do stejného prostoru

- Je-li $\mathbf{f}\colon \mathbb{R}^n \to \mathbb{R}^n$ (tj. je to transformace), pak z geometrického pohledu je zřejmé, že $\mathbf{f}$ může být pouze rotace nebo rotace složená se zrcadlením.
- Nechť $\mathbf{U}$ je ortogonální matice, $\mathbf{f}(\mathbf{x}) = \mathbf{U}\mathbf{x}$. Pak $\det \mathbf{U} = \pm 1$. Podle znaménka determinantu poznáme, zda součástí této transformace je zrcadlení.

Ortogonalita lineárních podprostorů

- Podprostory X a Y jsou na sebe ortogonální (píšeme $X \perp Y$), když každý vektor $\mathbf{x} \in X$ je kolmý na každý vektor $\mathbf{y} \in Y$, tedy pro ně platí $\mathbf{x}^T\mathbf{y} = 0$.
- Díky linearitě skalárního součinu stačí ověřit, že všechny bázové vektory z X jsou kolmé na všechny bázové vektory z Y.
- Pro $X \perp Y$ zřejmě platí $X \cap Y = \{\mathbf{0}\}$.
- Jestiže $X \perp Y$ a navíc X a Y generují celý prostor, pak říkáme, že X je ortogonálním doplňkem Y nebo též Y je ortogonálním doplňkem X a značíme $X = Y^\perp$ nebo $Y = X^\perp$.
- Je-li dáno X svou bází, jak najdeme bázi $X^\perp$?
- Je-li dáno X jako množina řešení soustavy $\mathbf{A}\mathbf{x} = \mathbf{0}$, tedy $X = \{\mathbf{x};\ \mathbf{A}\mathbf{x} = \mathbf{0}\}$. Jak najdeme bázi $X^\perp$?
- Povšimneme si, že $(\operatorname{rng}\mathbf{A})^\perp = \operatorname{Null}\mathbf{A}^T$ a $(\operatorname{Null}\mathbf{A})^\perp = \operatorname{rng}\mathbf{A}^T$.
- Důsledek: Soustava $\mathbf{A}\mathbf{x} = \mathbf{b}$ nemá řešení, právě když existuje řešení soustavy $\mathbf{A}^T\mathbf{y} = \mathbf{0}$, pro které neplatí $\mathbf{y} \perp \mathbf{b}$.