14. ORTOGONÁLNÍ PROJEKCE A PODPROSTORY

Jan Paseka

Masarykova univerzita Brno

2. dubna 2020

Abstrakt

Budeme pokračovat ve studiu euklidovských prostorů s cílem podat kvantitativní popis vzájemné polohy afinních podprostorů v takovémto prostoru pomocí dvou základních parametrů – jejich ***vzdálenosti*** a ***odchylky (úhlu)***.

Abstrakt

Budeme pokračovat ve studiu euklidovských prostorů s cílem podat kvantitativní popis vzájemné polohy afinních podprostorů v takovémto prostoru pomocí dvou základních parametrů – jejich ***vzdálenosti*** a ***odchylky (úhlu)***.

Naším hlavním nástrojem při tom budou lineární operátory ***kolmého průmětu***, zvané též ***ortogonální projekce***, vektorů do vektorových podprostorů.

Abstrakt

Budeme pokračovat ve studiu euklidovských prostorů s cílem podat kvantitativní popis vzájemné polohy afinních podprostorů v takovémto prostoru pomocí dvou základních parametrů – jejich ***vzdálenosti*** a ***odchylky (úhlu)***.

Naším hlavním nástrojem při tom budou lineární operátory ***kolmého průmětu***, zvané též ***ortogonální projekce***, vektorů do vektorových podprostorů.

V závěru kapitoly předvedeme aplikace rozpracovaných pojmů a metod.

Obsah přednášky I

- Ortokomplement a ortogonální projekce
  - Kolmý průmět vektoru.
  - Vzdálenost vektoru od podprostoru.
  - Odchylka vektoru od podprostoru.
  - Matice ortogonální projekce

Obsah přednášky I

- Ortokomplement a ortogonální projekce
  - Kolmý průmět vektoru.
  - Vzdálenost vektoru od podprostoru.
  - Odchylka vektoru od podprostoru.
  - Matice ortogonální projekce
- Vzdálenost a odchylka dvou afinních podprostorů
  - Vzdálenost dvou afinních podprostorů.
  - Odchylka dvou afinních podprostorů

Ortokomplement a ortogonální projekce I

Relace ortogonality (kolmosti) má několik následujících zřejmých vlastností.

Tvrzení

Nechť V je vektorový prostor se skalárním součinem. Potom pro všechna $\mathbf{x}, \mathbf{y}, \mathbf{z} \in V$, $c, d \in \mathbb{R}$ (resp. $c, d \in \mathbb{C}$) platí:

(a) $\mathbf{x} \perp \mathbf{0}$;

(b) $\mathbf{x} \perp \mathbf{x} \;\Leftrightarrow\; \mathbf{x} = \mathbf{0}$;

(c) $\mathbf{x} \perp \mathbf{y} \;\Leftrightarrow\; \mathbf{y} \perp \mathbf{x}$;

(d) $\big(\mathbf{x} \perp \mathbf{y} \;\&\; \mathbf{x} \perp \mathbf{z}\big) \;\Rightarrow\; \mathbf{x} \perp (c\mathbf{y} + d\mathbf{z})$.

Ortokomplement a ortogonální projekce I

Relace ortogonality (kolmosti) má několik následujících zřejmých vlastností.

Tvrzení

Nechť V je vektorový prostor se skalárním součinem. Potom pro všechna $\mathbf{x}, \mathbf{y}, \mathbf{z} \in V$, $c, d \in \mathbb{R}$ (resp. $c, d \in \mathbb{C}$) platí:

(a) $\mathbf{x} \perp \mathbf{0}$;

(b) $\mathbf{x} \perp \mathbf{x} \;\Leftrightarrow\; \mathbf{x} = \mathbf{0}$;

(c) $\mathbf{x} \perp \mathbf{y} \;\Leftrightarrow\; \mathbf{y} \perp \mathbf{x}$;

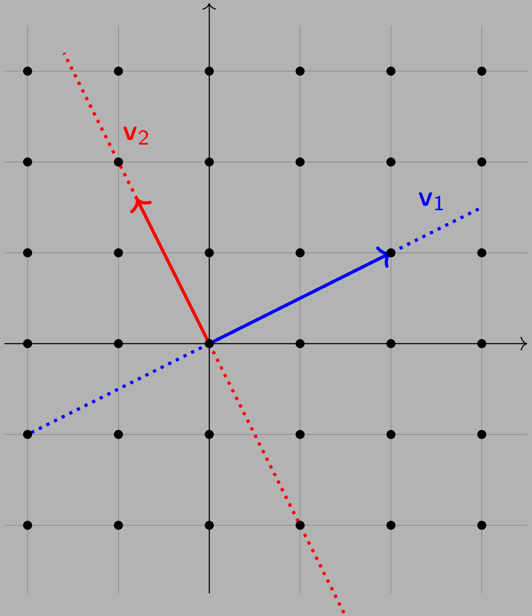
(d) $\big(\mathbf{x} \perp \mathbf{y} \;\&\; \mathbf{x} \perp \mathbf{z}\big) \;\Rightarrow\; \mathbf{x} \perp (c\mathbf{y} + d\mathbf{z})$.

Ortogonálním doplňkem nebo též ***ortokomplementem*** libovolné množiny $X \subseteq V$ ve vektorovém prostoru se skalárním součinem nazveme množinu

$$X^{\perp} = \{\mathbf{y} \in V;\, (\forall\, \mathbf{x} \in X)(\mathbf{x} \perp \mathbf{y})\}$$

všech vektorů $\mathbf{y} \in V$ kolmých na každý vektor $\mathbf{x} \in X$.

Ortokomplement a ortogonální projekce II



Ortokomplement a ortogonální projekce III

Tvrzení

Nechť V je vektorový prostor so skalárním součinem. Potom pro všechny množiny $X, Y \subseteq V$ platí:

(a) $\emptyset^\perp = \{\mathbf{0}\}^\perp = V, \;\; V^\perp = \{\mathbf{0}\}$;

(b) $X^\perp = [X]^\perp = [X^\perp]$;

(c) $X \subseteq Y \;\; \Rightarrow \;\; Y^\perp \subseteq X^\perp$;

(d) $X \subseteq X^{\perp\perp}$;

(e) $X^{\perp\perp\perp} = X^\perp$;

(f) $X \cap X^\perp = \{\mathbf{0}\}$, *pokud* $\mathbf{0} \in X$, *a* $X \cap X^\perp = \emptyset$, *pokud* $\mathbf{0} \notin X$;

(g) $(X \cup Y)^\perp = (X + Y)^\perp = X^\perp \cap Y^\perp$.

Ortokomplement a ortogonální projekce III

Tvrzení

Nechť V je vektorový prostor so skalárním součinem. Potom pro všechny množiny $X, Y \subseteq V$ platí:

(a) $\emptyset^\perp = \{\mathbf{0}\}^\perp = V, \;\; V^\perp = \{\mathbf{0}\}$;

(b) $X^\perp = [X]^\perp = [X^\perp]$;

(c) $X \subseteq Y \;\Rightarrow\; Y^\perp \subseteq X^\perp$;

(d) $X \subseteq X^{\perp\perp}$;

(e) $X^{\perp\perp\perp} = X^\perp$;

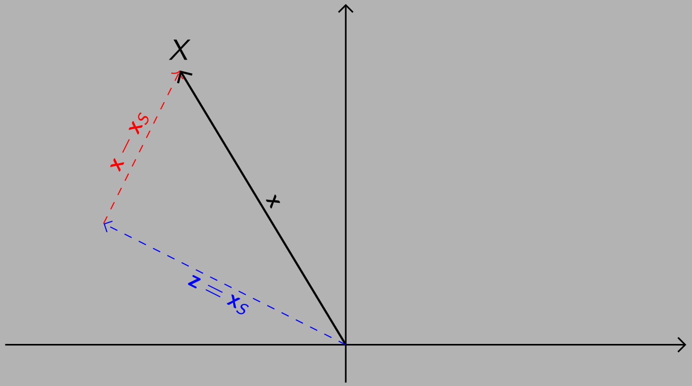
(f) $X \cap X^\perp = \{\mathbf{0}\}$, *pokud* $\mathbf{0} \in X$, *a* $X \cap X^\perp = \emptyset$, *pokud* $\mathbf{0} \notin X$;

(g) $(X \cup Y)^\perp = (X + Y)^\perp = X^\perp \cap Y^\perp$.

Z podmínky (b) mimo jiné plyne, že $X^\perp$ je vektorový podprostor ve V pro *každou podmnožinu* $X \subseteq V$.

Ortokomplement a ortogonální projekce IV

Nechť $S \subseteq V$ je lineární podprostor prostoru so skalárním součinem V a $\mathbf{x} \in V$. Říkáme, že vektor $\mathbf{z} \in S$ je ***kolmý průmět*** nebo též ***ortogonální projekce*** vektoru $\mathbf{x}$ do podprostoru S, pokud $\mathbf{x} - \mathbf{z} \in S^{\perp}$. Tento vektor (pokud existuje) budeme značit $\mathbf{z} = \mathrm{pr}_S(\mathbf{x}) = \mathbf{x}_S$.



Ortokomplement a ortogonální projekce V

Věta

Nechť V je vektorový prostor so skalárním součinem, $S \subseteq V$ je jeho konečně rozměrný lineární podprostor a $\mathbf{x} \in V$. Potom

(a) *kolmý průmět vektoru $\mathbf{x}$ do podprostoru S existuje a je jednoznačně určený rovností*

$$\mathrm{pr}_S(\mathbf{x}) = \mathbf{x}_S = \sum_{i=1}^{k} \langle \mathbf{x}, \mathbf{u}_i \rangle \mathbf{u}_i,$$

kde $(\mathbf{u}_1, \ldots, \mathbf{u}_k)$ je libovolná ortonormální báze podprostoru S;

(b) *pro libovolný vektor $\mathbf{y} \in S$ platí*

$$\|\mathbf{x} - \mathbf{x}_S\| \leq \|\mathbf{x} - \mathbf{y}\|$$

přičemž rovnost nastane právě tehdy, když $\mathbf{y} = \mathbf{x}_S$;

Ortokomplement a ortogonální projekce VI

(c) *pokud* $\mathbf{x} \neq \mathbf{0}$ *a* $S \neq \{\mathbf{0}\}$, *tak pro libovolný vektor* $\mathbf{0} \neq \mathbf{y} \in S$ *platí*

$$\frac{\|\mathbf{x}_S\|}{\|\mathbf{x}\|} \geq \frac{|\langle \mathbf{x}, \mathbf{y} \rangle|}{\|\mathbf{x}\| \, \|\mathbf{y}\|},$$

přičemž rovnost nastane právě tehdy, když vektory $\mathbf{x}_S$, $\mathbf{y}$ *jsou lineárně závislé.*

Ortokomplement a ortogonální projekce VI

(c) *pokud* $\mathbf{x} \neq \mathbf{0}$ *a* $S \neq \{\mathbf{0}\}$, *tak pro libovolný vektor* $\mathbf{0} \neq \mathbf{y} \in S$ *platí*

$$\frac{\|\mathbf{x}_S\|}{\|\mathbf{x}\|} \geq \frac{|\langle \mathbf{x}, \mathbf{y} \rangle|}{\|\mathbf{x}\| \, \|\mathbf{y}\|},$$

přičemž rovnost nastane právě tehdy, když vektory $\mathbf{x}_S$, $\mathbf{y}$ *jsou lineárně závislé.*

Vektor $\mathbf{x} - \mathbf{x}_S$ je kolmý na každou přímku v podprostoru S, speciálně trojúhelník tvořený vektory $\mathbf{x}$, $\mathbf{x}_S$, $\mathbf{x} - \mathbf{x}_S$ je pravoúhlý, s pravým úhlem při "konci" vektoru $\mathbf{x}_S$.

Ortokomplement a ortogonální projekce VI

(c) *pokud* $\mathbf{x} \neq \mathbf{0}$ *a* $S \neq \{\mathbf{0}\}$, *tak pro libovolný vektor* $\mathbf{0} \neq \mathbf{y} \in S$ *platí*

$$\frac{\|\mathbf{x}_S\|}{\|\mathbf{x}\|} \geq \frac{|\langle \mathbf{x}, \mathbf{y} \rangle|}{\|\mathbf{x}\| \, \|\mathbf{y}\|},$$

přičemž rovnost nastane právě tehdy, když vektory $\mathbf{x}_S$, $\mathbf{y}$ *jsou lineárně závislé.*

Vektor $\mathbf{x} - \mathbf{x}_S$ je kolmý na každou přímku v podprostoru S, speciálně trojúhelník tvořený vektory $\mathbf{x}$, $\mathbf{x}_S$, $\mathbf{x} - \mathbf{x}_S$ je pravoúhlý, s pravým úhlem při "konci" vektoru $\mathbf{x}_S$.

Podmínka (b) předcházející věty nás oprávňuje nazvat délku vektoru $\mathbf{x} - \mathbf{x}_S$ ***vzdáleností*** vektoru $\mathbf{x}$ od podprostoru S. Budeme ji značit

$$\mathrm{dist}(\mathbf{x}, S) = \|\mathbf{x} - \mathbf{x}_S\| = \min\{\|\mathbf{x} - \mathbf{y}\|; \, \mathbf{y} \in S\}.$$

Ortokomplement a ortogonální projekce VII

Důsledek

Nechť V je vektorový prostor se skalárním součinem a $S, T \subseteq V$ jsou jeho konečně rozměrné lineární podprostory. Potom

(a) $S = S^{\perp\perp}$, $(S \cap T)^{\perp} = S^{\perp} + T^{\perp}$ *a* $V = S \oplus S^{\perp}$;

(b) $\mathrm{pr}_S : V \to V$ *je lineární operátor;*

(c) $(\forall \mathbf{x} \in V)(\mathbf{x} \in S \ \Leftrightarrow \ \mathrm{pr}_S(\mathbf{x}) = \mathbf{x})$;

(d) $\mathrm{Im}\, \mathrm{pr}_S = S$ *a* $\mathrm{Ker}\, \mathrm{pr}_S = S^{\perp}$;

(e) $\mathbf{x} - \mathbf{x}_S$ *je kolmý průmět vektoru* $\mathbf{x}$ *do podprostoru* $S^{\perp}$.

Z podmínky (e) výše uvedeného důsledku je vzdálenost vektoru $\mathbf{x}$ od podprostoru $S^{\perp}$ daná vztahem

$$\mathrm{dist}\left(\mathbf{x}, S^{\perp}\right) = \|\mathbf{x}_S\|.$$

Ortokomplement a ortogonální projekce VIII

Podobně, protože kosinus je na intervalu $\langle 0, \pi \rangle$ klesající funkce, podmínka (c) předcházející věty nás oprávňuje nazvat výraz

$$\sphericalangle(\mathbf{x}, S) = \arccos \frac{\|\mathbf{x}_S\|}{\|\mathbf{x}\|} = \min\{\sphericalangle(\mathbf{x}, \mathbf{y}); \mathbf{0} \neq \mathbf{y} \in S\}$$

odchylkou vektoru $\mathbf{x} \neq \mathbf{0}$ od podprostoru $S \neq \{\mathbf{0}\}$, případně ***úhlem*** vektoru $\mathbf{x}$ a podprostoru S.

Ortokomplement a ortogonální projekce VIII

Podobně, protože kosinus je na intervalu $\langle 0, \pi \rangle$ klesající funkce, podmínka (c) předcházející věty nás opravňuje nazvat výraz

$$\sphericalangle(\mathbf{x}, S) = \arccos \frac{\|\mathbf{x}_S\|}{\|\mathbf{x}\|} = \min\{\sphericalangle(\mathbf{x}, \mathbf{y}); \mathbf{0} \neq \mathbf{y} \in S\}$$

odchylkou vektoru $\mathbf{x} \neq \mathbf{0}$ od podprostoru $S \neq \{\mathbf{0}\}$, případně ***úhlem*** vektoru $\mathbf{x}$ a podprostoru S.

Odchylka $\sphericalangle(\mathbf{x}, S)$ je tedy jednoznačně určená jako takové reálné číslo $\alpha \in \langle 0, \pi/2 \rangle$, pro které platí

$$\cos\alpha = \frac{\|\mathbf{x}_S\|}{\|\mathbf{x}\|} \quad \text{t. j.} \quad \sin\alpha = \frac{\|\mathbf{x} - \mathbf{x}_S\|}{\|\mathbf{x}\|}.$$

Ortokomplement a ortogonální projekce VIII

Podobně, protože kosinus je na intervalu $\langle 0, \pi \rangle$ klesající funkce, podmínka (c) předcházející věty nás oprávňuje nazvat výraz

$$\sphericalangle(\mathbf{x}, S) = \arccos \frac{\|\mathbf{x}_S\|}{\|\mathbf{x}\|} = \min\{\sphericalangle(\mathbf{x}, \mathbf{y}); \mathbf{0} \neq \mathbf{y} \in S\}$$

odchylkou vektoru $\mathbf{x} \neq \mathbf{0}$ od podprostoru $S \neq \{\mathbf{0}\}$, případně ***úhlem*** vektoru $\mathbf{x}$ a podprostoru S.

Odchylka $\sphericalangle(\mathbf{x}, S)$ je tedy jednoznačně určená jako takové reálné číslo $\alpha \in \langle 0, \pi/2 \rangle$, pro které platí

$$\cos \alpha = \frac{\|\mathbf{x}_S\|}{\|\mathbf{x}\|} \quad \text{t. j.} \quad \sin \alpha = \frac{\|\mathbf{x} - \mathbf{x}_S\|}{\|\mathbf{x}\|}.$$

Zřejmě opět půjde o ***neorientovaný uhel***. Pokud $\mathbf{x}_S \neq \mathbf{0}$, tak $\sphericalangle(\mathbf{x}, S) = \sphericalangle(\mathbf{x}, \mathbf{x}_S)$; pokud $\mathbf{x}_S = \mathbf{0}$, t. j. pokud $\mathbf{x} \in S^\perp$, tak samozřejmě $\sphericalangle(\mathbf{x}, S) = \pi/2$.

Ortokomplement a ortogonální projekce IX

Úhel dvou vektorů nabývá hodnoty z intervalu $\langle 0, \pi \rangle$, hodnoty, které nabývá úhel vektoru a podprostoru, jsou omezené na interval $\langle 0, \pi/2 \rangle$. Z podmínky (e) předchozí věty, pokud $S^\perp \neq \{\mathbf{0}\}$, tak odchylka vektoru $\mathbf{x} \neq \mathbf{0}$ od podprostoru $S^\perp$ je daná vztahem

$$\sphericalangle(\mathbf{x}, S^\perp) = \arccos \frac{\|\mathbf{x} - \mathbf{x}_S\|}{\|\mathbf{x}\|} = \arcsin \frac{\|\mathbf{x}_S\|}{\|\mathbf{x}\|}.$$

Ortokomplement a ortogonální projekce IX

Úhel dvou vektorů nabývá hodnoty z intervalu $\langle 0, \pi \rangle$, hodnoty, které nabývá úhel vektoru a podprostoru, jsou omezené na interval $\langle 0, \pi/2 \rangle$. Z podmínky (e) předchozí věty, pokud $S^\perp \neq \{\mathbf{0}\}$, tak odchylka vektoru $\mathbf{x} \neq \mathbf{0}$ od podprostoru $S^\perp$ je daná vztahem

$$\sphericalangle\left(\mathbf{x}, S^\perp\right) = \arccos \frac{\|\mathbf{x} - \mathbf{x}_S\|}{\|\mathbf{x}\|} = \arcsin \frac{\|\mathbf{x}_S\|}{\|\mathbf{x}\|}.$$

Z předcházející věty, část (a) máme přímý návod, jak najít kolmý průmět vektoru $\mathbf{x}$ do ***konečně rozměrného*** podprostoru $S \subseteq V$, a tím i vzdálenosti $\operatorname{dist}(\mathbf{x}, S)$, $\operatorname{dist}\left(\mathbf{x}, S^\perp\right)$ a odchylky $\sphericalangle(\mathbf{x}, S)$, $\sphericalangle\left(\mathbf{x}, S^\perp\right)$.Potřebujeme však mít k dispozici ***alespoň jednu ortonormální bázi*** v S.

Ortokomplement a ortogonální projekce X

Tvrzení

Nechť V je reálný vektorový prostor so skalárním součinem, S je jeho konečně rozměrný lineární podprostor s bazí $\boldsymbol{\alpha} = (\mathbf{u}_1, \ldots, \mathbf{u}_k)$ a $\mathbf{x} \in V$. Potom pro $\mathbf{c} = (c_1, \ldots, c_k)^T \in \mathbb{R}^k$ platí $\mathbf{x}_S = c_1\mathbf{u}_1 + \ldots + c_k\mathbf{u}_k$ právě tehdy, když $\mathbf{c}$ je řešením soustavy lineárních rovnic

$$\mathbf{G}(\boldsymbol{\alpha}) \cdot \mathbf{c} = \langle \mathbf{x}, \boldsymbol{\alpha} \rangle^T,$$

kde $\langle \mathbf{x}, \boldsymbol{\alpha} \rangle$ označuje řádkový vektor $\big(\langle \mathbf{x}, \mathbf{u}_1 \rangle, \ldots, \langle \mathbf{x}, \mathbf{u}_k \rangle\big) \in \mathbb{R}^k$.

Ortokomplement a ortogonální projekce X

Tvrzení

Nechť V je reálný vektorový prostor so skalárním součinem, S je jeho konečně rozměrný lineární podprostor s bazí $\boldsymbol{\alpha} = (\mathbf{u}_1, \ldots, \mathbf{u}_k)$ a $\mathbf{x} \in V$. Potom pro $\mathbf{c} = (c_1, \ldots, c_k)^T \in \mathbb{R}^k$ platí $\mathbf{x}_S = c_1\mathbf{u}_1 + \ldots + c_k\mathbf{u}_k$ právě tehdy, když $\mathbf{c}$ je řešením soustavy lineárních rovnic

$$\mathbf{G}(\boldsymbol{\alpha}) \cdot \mathbf{c} = \langle \mathbf{x}, \boldsymbol{\alpha} \rangle^T,$$

kde $\langle \mathbf{x}, \boldsymbol{\alpha} \rangle$ označuje řádkový vektor $\big(\langle \mathbf{x}, \mathbf{u}_1 \rangle, \ldots, \langle \mathbf{x}, \mathbf{u}_k \rangle\big) \in \mathbb{R}^k$.

Rozšířená matice $\left(\mathbf{G}(\boldsymbol{\alpha}) \,|\, \langle \mathbf{x}, \boldsymbol{\alpha} \rangle^T\right)$ uvedené soustavy je Gramovou maticí $\mathbf{G}(\mathbf{u}_1, \ldots, \mathbf{u}_k, \mathbf{x})$ řádu $k+1$, ze které jsme vynechali poslední řádek.

Ortokomplement a ortogonální projekce XI

Je-li $\boldsymbol{\alpha}$ ortonormální báze, tak $\mathbf{G}(\boldsymbol{\alpha}) = \mathbf{I}_k$, t. j. příslušná soustava je už ve vyřešeném tvaru $\mathbf{c} = \langle \mathbf{x}, \boldsymbol{\alpha} \rangle^T$, t.j. ve shodě s podmínkou (a) předcházející věty.

Ortokomplement a ortogonální projekce XI

Je-li $\boldsymbol{\alpha}$ ortonormální báze, tak $\mathbf{G}(\boldsymbol{\alpha}) = \mathbf{I}_k$, t. j. příslušná soustava je už ve vyřešeném tvaru $\mathbf{c} = \langle \mathbf{x}, \boldsymbol{\alpha} \rangle^T$, t.j. ve shodě s podmínkou (a) předcházející věty.

Totiž

$$\mathrm{pr}_S(\mathbf{x}) = \mathbf{x}_S = \langle \mathbf{x}, \boldsymbol{\alpha} \rangle^T \cdot \boldsymbol{\alpha} = \mathbf{c} \cdot \boldsymbol{\alpha}.$$

Zároveň předcházející tvrzení platí i pro případ, kdy $\boldsymbol{\alpha}$ je posloupnost generátorů podprostoru S.

Příklad

V $\mathbb{R}^4$ se standardním skalárním součinem je daný vektor $\mathbf{x} = (1, 1, 1, 1)^T$ a rovina $S = [\mathbf{u}, \mathbf{v}]$, kde $\mathbf{u} = (0, -1, 0, 1)^T$, $\mathbf{v} = (1, -2, 1, -3)^T$.
Najdeme kolmý průmět vektoru $\mathbf{x}$ do roviny S a vypočítáme vzdálenost $\mathrm{dist}(\mathbf{x}, S)$ a odchylku $\sphericalangle(\mathbf{x}, S)$.

Ortokomplement a ortogonální projekce XII

Kolmý průmět budeme hledat ve tvaru $\mathbf{x}_S = c\mathbf{u} + d\mathbf{v}$, kde $(c, d)^T \in \mathbb{R}^2$ vyhovuje soustavě s rozšírenou maticí

$$\left(\begin{array}{cc|c} \langle \mathbf{u}, \mathbf{u} \rangle & \langle \mathbf{v}, \mathbf{u} \rangle & \langle \mathbf{x}, \mathbf{u} \rangle \\ \langle \mathbf{u}, \mathbf{v} \rangle & \langle \mathbf{v}, \mathbf{v} \rangle & \langle \mathbf{x}, \mathbf{v} \rangle \end{array}\right) = \left(\begin{array}{cc|c} 2 & -1 & 0 \\ -1 & 15 & -3 \end{array}\right).$$

Jejím řešením dostaneme $c = -3/29$, $d = -6/29$, tedy kolmý průmět vektoru $\mathbf{x}$ do roviny $[\mathbf{u}, \mathbf{v}]$ je

$$\mathbf{x}_S = (\mathbf{u}, \mathbf{v}) \cdot \begin{pmatrix} c \\ d \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & -2 \\ 0 & 1 \\ 1 & -3 \end{pmatrix} \cdot \begin{pmatrix} -3/29 \\ -6/29 \end{pmatrix} = \tfrac{3}{29} \begin{pmatrix} -2 \\ 5 \\ -2 \\ 5 \end{pmatrix}.$$

Ortokomplement a ortogonální projekce XIII

Pro vzdálenost $\mathbf{x}$ od S potom dostáváme

$$\mathrm{dist}(\mathbf{x}, S) = \|\mathbf{x} - \mathbf{x}_S\| = \left\| \frac{7}{29}(5,2,5,2)^T \right\| = \frac{7}{29}\sqrt{58}.$$

Pro odchylku $\mathbf{x}$ od S dostaneme

$$\sin \sphericalangle(\mathbf{x}, S) = \frac{\|\mathbf{x} - \mathbf{x}_S\|}{\|\mathbf{x}\|} = \frac{7}{2 \cdot 29}\sqrt{58} = \frac{7}{\sqrt{58}}.$$

S použitím kalkulačky či tabulek můžeme zjistit, že

$$\sphericalangle(\mathbf{x}, S) = \arcsin \frac{7}{\sqrt{58}} \approx 1,1659\,\mathrm{rad} \approx 66^\circ\, 48'\, 5''.$$

Ortokomplement a ortogonální projekce XIV

Příklad

Nechť $\mathbf{A} \in R^{m \times n}$, přičemž $m \geq n$ a $h(\mathbf{A}) = n$, t. j. sloupce matice $\mathbf{A}$ jsou lineárně nezávislé vektory v euklidovském prostoru $\mathbb{R}^m$ se standardním skalárním součinem.

Označme $S \subseteq \mathbb{R}^m$ lineární podprostor generovaný sloupci matice $\mathbf{A}$. Potom ortogonální projekce na podprostor S je lineární operátor $\mathrm{pr}_S : \mathbb{R}^m \to \mathbb{R}^m$.

Ortokomplement a ortogonální projekce XIV

Příklad

Nechť $\mathbf{A} \in R^{m\times n}$, přičemž $m \geq n$ a $h(\mathbf{A}) = n$, t. j. sloupce matice $\mathbf{A}$ jsou lineárně nezávislé vektory v euklidovském prostoru $\mathbb{R}^m$ se standardním skalárním součinem.

Označme $S \subseteq \mathbb{R}^m$ lineární podprostor generovaný sloupci matice $\mathbf{A}$. Potom ortogonální projekce na podprostor S je lineární operátor $\mathrm{pr}_S : \mathbb{R}^m \to \mathbb{R}^m$.

Najděme jeho matici $\mathbf{B} = (\mathrm{pr}_S)_{\varepsilon,\varepsilon} \in \mathbb{R}^{m\times m}$ vzhledem ke kanonické ortonormální bázi ε prostoru $\mathbb{R}^m$.

Pokud ztotožníme matici $\mathbf{A}$ s uspořádanou n-ticí jejich sloupců, tak $\mathbf{A}$ je bazí S.

Ortokomplement a ortogonální projekce XV

Podle předcházejícího tvrzení obraz $\mathbf{y} = \mathrm{pr}_S(\mathbf{x})$ vektoru $\mathbf{x} \in \mathbb{R}^m$ dostaneme ve tvaru

$$\mathbf{y} = \mathbf{A} \cdot \mathbf{c},$$

kde $\mathbf{c} \in \mathbb{R}^n$ je jediné řešení soustavy

$$\mathbf{G}(\mathbf{A}) \cdot \mathbf{c} = \langle \mathbf{x}, \mathbf{A} \rangle^T.$$

Z nezávislosti sloupců matice $\mathbf{A}$ víme, že $\mathbf{G}(\mathbf{A}) = \mathbf{A}^T \cdot \mathbf{A}$ je regulární matice.

Dále platí $\langle \mathbf{x}, \mathbf{A} \rangle = \mathbf{x}^T \cdot \mathbf{A}$, tedy $\langle \mathbf{x}, \mathbf{A} \rangle^T = \mathbf{A}^T \cdot \mathbf{x}$.

Ortokomplement a ortogonální projekce XVI

Po dosazení

$$\begin{aligned} \mathbf{c} &= \mathbf{G}(\mathbf{A})^{-1} \cdot \langle \mathbf{x}, \mathbf{A} \rangle^{T} = \left(\mathbf{A}^{T} \cdot \mathbf{A}\right)^{-1} \cdot \mathbf{A}^{T} \cdot \mathbf{x}, \\ \mathbf{y} &= \mathbf{A} \cdot \mathbf{c} = \mathbf{A} \cdot \left(\mathbf{A}^{T} \cdot \mathbf{A}\right)^{-1} \cdot \mathbf{A}^{T} \cdot \mathbf{x}. \end{aligned}$$

Tedy hledaná matice ortogonální projekce pr_{S} je

$$\mathbf{B} = \left(\mathrm{pr}_{S}\right)_{\boldsymbol{\varepsilon},\boldsymbol{\varepsilon}} = \mathbf{A} \cdot \left(\mathbf{A}^{T} \cdot \mathbf{A}\right)^{-1} \cdot \mathbf{A}^{T}.$$

Vzdálenost dvou afinních podprostorů I

Nechť V je vektorový prostor se skalárním součinem a X, Y jsou jeho dvě neprázdné podmnožiny. ***Vzdáleností množin*** X, Y v prostoru V nazýváme číslo

$$\operatorname{dist}(X, Y) = \inf\{\|\mathbf{x} - \mathbf{y}\| \,;\, \mathbf{x} \in X \;\&\; \mathbf{y} \in Y\}.$$

Vzdálenost dvou afinních podprostorů I

Nechť V je vektorový prostor se skalárním součinem a X, Y jsou jeho dvě neprázdné podmnožiny. ***Vzdáleností množin*** X, Y v prostoru V nazýváme číslo

$$\operatorname{dist}(X, Y) = \inf\{\|\mathbf{x} - \mathbf{y}\| \,;\, \mathbf{x} \in X \;\&\; \mathbf{y} \in Y\}.$$

Lemma

Nechť V je vektorový prostor se skalárním součinem a M, N jsou jeho afinní podprostory. Potom pro libovolné body $\mathbf{p} \in M$, $\mathbf{q} \in M$ platí:

$$\operatorname{dist}(M, N) = \operatorname{dist}(\mathbf{p} - \mathbf{q}, \operatorname{Dir} M + \operatorname{Dir} N).$$

Vzdálenost dvou afinních podprostorů I

Nechť V je vektorový prostor se skalárním součinem a X, Y jsou jeho dvě neprázdné podmnožiny. ***Vzdáleností množin*** X, Y v prostoru V nazýváme číslo

$$\mathrm{dist}(X, Y) = \inf\{\|\mathbf{x} - \mathbf{y}\| \,;\, \mathbf{x} \in X \;\&\; \mathbf{y} \in Y\}.$$

Lemma

Nechť V je vektorový prostor se skalárním součinem a M, N jsou jeho afinní podprostory. Potom pro libovolné body $\mathbf{p} \in M$, $\mathbf{q} \in M$ platí:

$$\mathrm{dist}(M, N) = \mathrm{dist}(\mathbf{p} - \mathbf{q}, \, \mathrm{Dir}M + \mathrm{Dir}N).$$

Říkáme, že body $\mathbf{p} \in M$, $\mathbf{q} \in N$ tvoří ***příčku*** afinních podprostorů M, N, pokud

$$\mathrm{dist}(M, N) = \|\mathbf{p} - \mathbf{q}\|,$$

t. j. pokud se vzdálenost podprostorů M, N realizuje jako délka vektoru $\mathbf{p} - \mathbf{q}$.

Vzdálenost dvou afinních podprostorů II

$H = \{\mathbf{x} \in \mathbb{R}^3 | a^T x = b\}, \|a\| = 1, b \neq 0$

$b = \mathrm{dist}(\{\mathbf{0}\}, H), \mathbf{q}_{[\mathbf{a}]} = b \cdot \mathbf{a}$

$\mathbf{a}^T\mathbf{p} - b = \mathrm{dist}(\{\mathbf{p}\}, H)$

$\mathbf{p}_{[\mathbf{a}]} = (\mathbf{a}^T\mathbf{p}) \cdot \mathbf{a}$

$\mathbf{p}' = \overline{\mathbf{0p}} \cap H = \lambda\mathbf{p}$

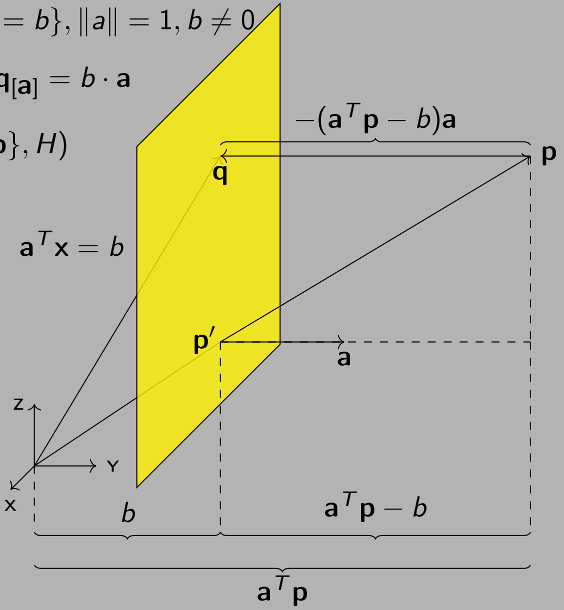
$\mathbf{a}^T\mathbf{p}' = b = \mathbf{a}^T\lambda\mathbf{p}$

$\mathbf{p}' = \frac{b}{\mathbf{a}^T\mathbf{p}}\mathbf{p}$

$\mathbf{a}^T(\mathbf{p} - \mathbf{p}') =$

$\mathbf{a}^T\mathbf{p} - b$

$\mathbf{p}'_{[\mathbf{a}]} = b \cdot \mathbf{a}$



Vzdálenost dvou afinních podprostorů III

Tvrzení

Nechť M, N jsou konečně rozměrné afinní podprostory vektorového prostoru se skalárním součinem V. Potom

(a) *body $\mathbf{p} \in M$, $\mathbf{q} \in N$ tvoří příčku podprostorů M, N právě tehdy, když $\mathbf{p} - \mathbf{q} \in (\mathrm{Dir} M + \mathrm{Dir} N)^{\perp}$;*

(b) *pro libovolné body $\mathbf{p} \in M$, $\mathbf{q} \in N$ a vektory $\mathbf{u} \in \mathrm{Dir} M$, $\mathbf{v} \in \mathrm{Dir} N$ platí: body $\mathbf{p} + \mathbf{u}$, $\mathbf{q} + \mathbf{v}$ tvoří příčku podprostorů M, N právě tehdy, když vektor $\mathbf{v} - \mathbf{u}$ je kolmým průmětem vektoru $\mathbf{p} - \mathbf{q}$ do lineárního podprostoru $\mathrm{Dir} M + \mathrm{Dir} N$;*

(c) *existují body $\mathbf{p} \in M$, $\mathbf{q} \in N$ tvořící příčku podprostorů M, N.*

Vzdálenost dvou afinních podprostorů III

Tvrzení

Nechť M, N jsou konečně rozměrné afinní podprostory vektorového prostoru se skalárním součinem V. Potom

(a) *body $\mathbf{p} \in M$, $\mathbf{q} \in N$ tvoří příčku podprostorů M, N právě tehdy, když $\mathbf{p} - \mathbf{q} \in (\mathrm{Dir} M + \mathrm{Dir} N)^{\perp}$;*

(b) *pro libovolné body $\mathbf{p} \in M$, $\mathbf{q} \in N$ a vektory $\mathbf{u} \in \mathrm{Dir} M$, $\mathbf{v} \in \mathrm{Dir} N$ platí: body $\mathbf{p} + \mathbf{u}$, $\mathbf{q} + \mathbf{v}$ tvoří příčku podprostorů M, N právě tehdy, když vektor $\mathbf{v} - \mathbf{u}$ je kolmým průmětem vektoru $\mathbf{p} - \mathbf{q}$ do lineárního podprostoru $\mathrm{Dir} M + \mathrm{Dir} N$;*

(c) *existují body $\mathbf{p} \in M$, $\mathbf{q} \in N$ tvořící příčku podprostorů M, N.*

Důsledek

Pro konečně rozměrné afinní podprostory $M, N \subseteq V$ vektorového prostoru se skalárním součinem platí $\mathrm{dist}(M, N) = 0$ právě tehdy, když $M \cap N \neq \emptyset$.

Vzdálenost dvou afinních podprostorů IV

Přímý návod jak najít příčku a vzdálenost libovolných konečně rozměrných afinních podprostorů

Jsou-li $M = \mathbf{p} + [\mathbf{u}_1, \dots, \mathbf{u}_m]$, $N = \mathbf{q} + [\mathbf{v}_1, \dots, \mathbf{v}_n]$ zadané parametricky, stačí najít jedno řešení $\mathbf{c} = (c_1, \dots, c_m, c_{m+1}, \dots, c_{m+n})^T \in \mathbb{R}^{m+n}$ soustavy

$$\mathbf{G}(\gamma) \cdot \mathbf{c} = \langle \mathbf{p} - \mathbf{q}, \gamma \rangle^T,$$

kde $\boldsymbol{\gamma} = (\mathbf{u}_1, \dots, \mathbf{u}_m, \mathbf{v}_1, \dots, \mathbf{v}_n)$, a položit

$$\mathbf{u} = c_1 \mathbf{u}_1 + \dots c_m \mathbf{u}_m, \;\; \mathbf{v} = c_{m+1} \mathbf{v}_1 + \dots + c_{m+n} \mathbf{v}_n.$$

Vzdálenost dvou afinních podprostorů IV

Přímý návod jak najít příčku a vzdálenost libovolných konečně rozměrných afinních podprostorů

Jsou-li $M = \mathbf{p} + [\mathbf{u}_1, \dots, \mathbf{u}_m]$, $N = \mathbf{q} + [\mathbf{v}_1, \dots, \mathbf{v}_n]$ zadané parametricky, stačí najít jedno řešení $\mathbf{c} = (c_1, \dots, c_m, c_{m+1}, \dots, c_{m+n})^T \in \mathbb{R}^{m+n}$ soustavy

$$\mathbf{G}(\gamma) \cdot \mathbf{c} = \langle \mathbf{p} - \mathbf{q}, \gamma \rangle^T,$$

kde $\boldsymbol{\gamma} = (\mathbf{u}_1, \dots, \mathbf{u}_m, \mathbf{v}_1, \dots, \mathbf{v}_n)$, a položit

$$\mathbf{u} = c_1 \mathbf{u}_1 + \dots c_m \mathbf{u}_m, \;\; \mathbf{v} = c_{m+1} \mathbf{v}_1 + \dots + c_{m+n} \mathbf{v}_n.$$

Potom vektor $\mathbf{w} = \mathbf{u} + \mathbf{v} = \gamma \cdot \mathbf{c}$ je kolmým průmětem vektoru $\mathbf{p} - \mathbf{q}$ do lineárního podprostoru

$$\mathrm{Dir} M + \mathrm{Dir} N = [\mathbf{u}_1, \dots \mathbf{u}_m, \mathbf{v}_1, \dots, \mathbf{v}_n]$$

a příčka podprostorů M, N je tvořená body $\mathbf{p} - \mathbf{u}$, $\mathbf{q} + \mathbf{v}$.

Vzdálenost dvou afinních podprostorů V

Tedy

$$\begin{aligned}\operatorname{dist}(M, N) &= \|(\mathbf{p} - \mathbf{u}) - (\mathbf{q} + \mathbf{v})\| \\ &= \|\mathbf{p} - \mathbf{q} - \mathbf{w}\| \,.\end{aligned}$$

Vzdálenost dvou afinních podprostorů V

Tedy

$$\begin{aligned}\mathrm{dist}(M,N) &= \|(\mathbf{p}-\mathbf{u})-(\mathbf{q}+\mathbf{v})\| \\ &= \|\mathbf{p}-\mathbf{q}-\mathbf{w}\|\,.\end{aligned}$$

Příklad

V euklidovském prostoru $\mathbb{R}^4$ se standardním skalárním součinem máme najít vzdálenost rovin

$$\begin{aligned}M &= (1,1,2,-2)^T + [\mathbf{e}_1+\mathbf{e}_2, \mathbf{e}_1+\mathbf{e}_2+\mathbf{e}_3],\\ N &= (0,0,5,-1)^T + [\mathbf{e}_2+\mathbf{e}_4, \mathbf{e}_2+\mathbf{e}_3+\mathbf{e}_4].\end{aligned}$$

Vzdálenost dvou afinních podprostorů VI

Z príslušných skalárnych súčinov zostavíme (takmer Gramovu) rozšírenú maticu sústavy $\mathbf{G}(\gamma) \cdot \mathbf{c} = \langle \mathbf{p} - \mathbf{q}, \gamma \rangle^T$ a upravíme ju na redukovaný stupňovitý tvar

$$\left(\begin{array}{cccc|c} 2 & 2 & 1 & 1 & 2 \\ 2 & 3 & 1 & 2 & -1 \\ 1 & 1 & 2 & 2 & 0 \\ 1 & 2 & 2 & 3 & -3 \end{array}\right) \sim \left(\begin{array}{cccc|c} 1 & 0 & 0 & -1 & 13/3 \\ 0 & 1 & 0 & 1 & -3 \\ 0 & 0 & 1 & 1 & -2/3 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

Vzdálenost dvou afinních podprostorů VI

Z príslušných skalárnych súčinov zostavíme (takmer Gramovu) rozšírenú maticu sústavy $\mathbf{G}(\gamma)\cdot\mathbf{c} = \langle \mathbf{p}-\mathbf{q}, \gamma\rangle^T$ a upravíme ju na redukovaný stupňovitý tvar

$$\left(\begin{array}{cccc|c} 2 & 2 & 1 & 1 & 2 \\ 2 & 3 & 1 & 2 & -1 \\ 1 & 1 & 2 & 2 & 0 \\ 1 & 2 & 2 & 3 & -3 \end{array}\right) \sim \left(\begin{array}{cccc|c} 1 & 0 & 0 & -1 & 13/3 \\ 0 & 1 & 0 & 1 & -3 \\ 0 & 0 & 1 & 1 & -2/3 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

Řešení soustavy zapíšeme vo všeobecném tvaru $\mathbf{c}_t = (13/3+t, -3-t, -2/3-t, t)^T$ s parametrem $t\in\mathbb{R}$.
Položme

$$\begin{aligned} \mathbf{u}_t &= c_1(\mathbf{e}_1+\mathbf{e}_2) + c_2(\mathbf{e}_1+\mathbf{e}_2+\mathbf{e}_3) = (4/3, 4/3, -3-t, 0)^T, \\ \mathbf{v}_t &= c_3(\mathbf{e}_2+\mathbf{e}_4) + c_3(\mathbf{e}_2+\mathbf{e}_3+\mathbf{e}_4) = (0, -2/3, t, -2/3)^T. \end{aligned}$$

Vzdálenost dvou afinních podprostorů VII

Potom pro každé $t \in \mathbb{R}$ dvojice bodů

$$\begin{aligned}\mathbf{p}_t &= (1,1,2,-2)^T - \mathbf{u}_t = (-1/3,-1/3,5+t,-2)^T\\ \mathbf{q}_t &= (0,0,5,-1)^T + \mathbf{v}_t = (0,-2/3,5+t,-5/3)^T\end{aligned}$$

tvoří příčku podprostorů M, N.

Vektory

$$\mathbf{u}_t + \mathbf{v}_t = (4/3,2/3,-3,-2/3)^T, \qquad \mathbf{p}_t - \mathbf{q}_t = \frac{1}{3}(-1,1,0,-1)^T$$

ale od parametru t nezávisí stejně jako vzdálenost

$$\operatorname{dist}(M,N) = \|\mathbf{p}_t - \mathbf{q}_t\| = \frac{1}{\sqrt{3}}.$$

Odchylka dvou afinních podprostorů I

Odchylku neboli ***úhel*** dvou ***netriviálních konečně rozměrných afinních podprostorů*** ve vektorovém prostoru so skalárním součinem V značíme $\sphericalangle(M, N)$ a definujeme ji jako odchylku $\sphericalangle(\mathrm{Dir} M, \mathrm{Dir} N)$ jejich zaměření.

Odchylka dvou afinních podprostorů I

Odchylku neboli ***úhel*** dvou ***netriviálních konečně rozměrných afinních podprostorů*** ve vektorovém prostoru so skalárním součinem V značíme $\sphericalangle(M, N)$ a definujeme ji jako odchylku $\sphericalangle(\mathrm{Dir} M, \mathrm{Dir} N)$ jejich zaměření.

Odchylku neboli ***úhel*** $\sphericalangle(S, T)$ dvou ***netriviálních konečně rozměrných lineárních podprostorů*** $S,\ T \subseteq V$ definujeme následovně:

Pro $S \subseteq T$ nebo $T \subseteq S$ položíme

$$\sphericalangle(S, T) = 0.$$

Pokud $S \cap T = \{\mathbf{0}\}$, klademe

$$\sphericalangle(S, T) = \inf\{\sphericalangle(\mathbf{x}, \mathbf{y});\ \mathbf{0} \neq \mathbf{x} \in S\ \&\ \mathbf{0} \neq \mathbf{y} \in T\}.$$

Odchylka dvou afinních podprostorů II

Pokud bychom takovýmto způsobem definovali odchylku $\sphericalangle(S, T)$, i když $S \cap T \neq \{\mathbf{0}\}$, libovolný společný nenulový vektor $\mathbf{x} \in S \cap T$ by se postaral o to, aby platilo $\sphericalangle(S, T) = \sphericalangle(\mathbf{x}, \mathbf{x}) = 0$, což nevypadá příliš rozumně.

Tedy pro $S \cap T \neq \{\mathbf{0}\}$, $S \not\subseteq T$, $T \not\subseteq S$, položíme

$$S_1 = S \cap (S \cap T)^{\perp}, \qquad T_1 = T \cap (S \cap T)^{\perp}.$$

Odchylka dvou afinních podprostorů II

Pokud bychom takovýmto způsobem definovali odchylku $\sphericalangle(S, T)$, i když $S \cap T \neq \{\mathbf{0}\}$, libovolný společný nenulový vektor $\mathbf{x} \in S \cap T$ by se postaral o to, aby platilo $\sphericalangle(S, T) = \sphericalangle(\mathbf{x}, \mathbf{x}) = 0$, což nevypadá příliš rozumně.

Tedy pro $S \cap T \neq \{\mathbf{0}\}$, $S \not\subseteq T$, $T \not\subseteq S$, položíme

$$S_1 = S \cap (S \cap T)^{\perp}, \qquad T_1 = T \cap (S \cap T)^{\perp}.$$

Zřejmě $S_1,\ T_1 \subseteq V$ jsou netriválni lineární podprostory a $S_1 \cap T_1 = \{\mathbf{0}\}$ (za předpokladu $S \cap T = \{\mathbf{0}\}$ dokonce platí $S_1 = S$, $T_1 = T$).

Proto můžeme konečně definovat

$$\sphericalangle(S, T) = \sphericalangle(S_1, T_1).$$

Odchylka dvou afinních podprostorů II

Pokud bychom takovýmto způsobem definovali odchylku $\sphericalangle(S,T)$, i když $S \cap T \neq \{\mathbf{0}\}$, libovolný společný nenulový vektor $\mathbf{x} \in S \cap T$ by se postaral o to, aby platilo $\sphericalangle(S,T) = \sphericalangle(\mathbf{x},\mathbf{x}) = 0$, což nevypadá příliš rozumně.

Tedy pro $S \cap T \neq \{\mathbf{0}\}$, $S \not\subseteq T$, $T \not\subseteq S$, položíme

$$S_1 = S \cap (S \cap T)^{\perp}, \qquad T_1 = T \cap (S \cap T)^{\perp}.$$

Zřejmě $S_1,\ T_1 \subseteq V$ jsou netriválnì lineární podprostory a $S_1 \cap T_1 = \{\mathbf{0}\}$ (za předpokladu $S \cap T = \{\mathbf{0}\}$ dokonce platí $S_1 = S$, $T_1 = T$).

Proto můžeme konečně definovat

$$\sphericalangle(S,T) = \sphericalangle(S_1,T_1).$$

Takto definovaný úhel podprostorů S, T je číslo z intervalu $\langle 0, \pi/2 \rangle$ a platí pro něj $\sphericalangle(S,T) = \sphericalangle(T,S)$, tedy je to ***neorientovaný úhel***.

Odchylka dvou afinních podprostorů III

Tvrzení

Nechť V je vektorový prostor so skalárním součinem a S, T jsou jeho konečně rozměrné lineární podprostory, přičemž $S \not\subseteq T$ ani $T \not\subseteq S$. Potom

$$\sphericalangle(S, T) = \inf\left\{\sphericalangle(\mathbf{x}, T);\ \mathbf{0} \neq \mathbf{x} \in S \cap (S \cap T)^{\perp}\right\}.$$

Odchylka dvou afinních podprostorů III

Tvrzení

Nechť V je vektorový prostor so skalárním součinem a S, T jsou jeho konečně rozměrné lineární podprostory, přičemž $S \not\subseteq T$ ani $T \not\subseteq S$. Potom

$$\sphericalangle(S, T) = \inf\left\{\sphericalangle(\mathbf{x}, T);\ \mathbf{0} \neq \mathbf{x} \in S \cap (S \cap T)^{\perp}\right\}.$$

Odchylka přímky $[\mathbf{x}]$, kde $\mathbf{x} \neq \mathbf{0}$, a konečně rozměrného lineárního podprostoru $S \neq \{\mathbf{0}\}$ je daná vztahem

$$\begin{aligned} \sphericalangle\big([\mathbf{x}], S\big) &= \sphericalangle(\mathbf{x}, S) \\ &= \arccos \frac{\|\mathbf{x}_S\|}{\|\mathbf{x}\|} = \begin{cases} \sphericalangle(\mathbf{x}, \mathbf{x}_S), & \text{pokud } \mathbf{x}_S \neq \mathbf{0}, \\ & \text{t. j. } \mathbf{x} \notin S^{\perp}, \\ \pi/2, & \text{pokud } \mathbf{x}_S = \mathbf{0}, \\ & \text{t. j. } \mathbf{x} \in S^{\perp}. \end{cases} \end{aligned}$$

Odchylka dvou afinních podprostorů IV

Příklad

V euklidovském prostoru $\mathbb{R}^4$ se standardním skalárním součinem máme najít odchylku rovin

$$\begin{aligned} M &= (1,1,2,-2)^T + [\mathbf{e}_1+\mathbf{e}_2, \mathbf{e}_1+\mathbf{e}_2+\mathbf{e}_3], \\ N &= (0,0,5,-1)^T + [\mathbf{e}_2+\mathbf{e}_4, \mathbf{e}_2+\mathbf{e}_3+\mathbf{e}_4]. \end{aligned}$$

Podle definice $\sphericalangle(M,N) = \sphericalangle(S,T)$, kde $S = [\mathbf{e}_1+\mathbf{e}_2, \mathbf{e}_1+\mathbf{e}_2+\mathbf{e}_3]$ a $T = [\mathbf{e}_2+\mathbf{e}_4, \mathbf{e}_2+\mathbf{e}_3+\mathbf{e}_4]$.

Vidíme, že $S \cap T = [\mathbf{e}_3]$, tedy $(S \cap T)^\perp = [\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_4]$. Nutně pak

$$S_1 = S \cap (S \cap T)^\perp = [\mathbf{e}_1+\mathbf{e}_2], \qquad T_1 = T \cap (S \cap T)^\perp = [\mathbf{e}_2+\mathbf{e}_4].$$

Protože $\langle \mathbf{e}_1+\mathbf{e}_2, \mathbf{e}_2+\mathbf{e}_4 \rangle = 1 \geq 0$ a $\|\mathbf{e}_1+\mathbf{e}_2\| = \|\mathbf{e}_2+\mathbf{e}_4\| = \sqrt{2}$,

$$\sphericalangle(M,N) = \sphericalangle(\mathbf{e}_1+\mathbf{e}_2,\ \mathbf{e}_2+\mathbf{e}_4) = \arccos\frac{1}{2} = \frac{\pi}{3} = 60^\circ.$$

Odchylka dvou afinních podprostorů V

Každý $(n-1)$-rozměrný lineární podprostor S v n-rozměrném euklidovském prostoru V má tvar $S = [\mathbf{a}]^{\perp}$ pro vhodný nenulový vektor $\mathbf{a} \in V$.

Odchylka dvou afinních podprostorů V

Každý $(n-1)$-rozměrný lineární podprostor S v n-rozměrném euklidovském prostoru V má tvar $S = [\mathbf{a}]^\perp$ pro vhodný nenulový vektor $\mathbf{a} \in V$.

Každá nadrovina $N \subseteq V$ se zaměřením S má tvar $N = \mathbf{p} + [\mathbf{a}]^\perp$ pro nějaké $\mathbf{p} \in N$.

Odchylka dvou afinních podprostorů V

Každý $(n-1)$-rozměrný lineární podprostor S v n-rozměrném euklidovském prostoru V má tvar $S = [\mathbf{a}]^{\perp}$ pro vhodný nenulový vektor $\mathbf{a} \in V$.

Každá nadrovina $N \subseteq V$ se zaměřením S má tvar $N = \mathbf{p} + [\mathbf{a}]^{\perp}$ pro nějaké $\mathbf{p} \in N$.

Vektor **a** se nazývá ***normála*** neboli ***normálový vektor*** nadroviny N. Normála nadroviny je určená jednoznačně až na skalární násobek.

Odchylka dvou afinních podprostorů V

Každý $(n-1)$-rozměrný lineární podprostor S v n-rozměrném euklidovském prostoru V má tvar $S = [\mathbf{a}]^{\perp}$ pro vhodný nenulový vektor $\mathbf{a} \in V$.

Každá nadrovina $N \subseteq V$ se zaměřením S má tvar $N = \mathbf{p} + [\mathbf{a}]^{\perp}$ pro nějaké $\mathbf{p} \in N$.

Vektor $\mathbf{a}$ se nazývá ***normála*** neboli ***normálový vektor*** nadroviny N. Normála nadroviny je určená jednoznačně až na skalární násobek.

V euklidovském prostoru $\mathbb{R}^n$ se standardním skalárním součinem vystupuje normálový vektor dané nadroviny přímo v její (obecné) rovnici. Pokud je totiž nadrovina N daná rovnicí

$$a_1 x_1 + \ldots + a_n x_n = b,$$

tak $\mathbf{a} = (a_1, \ldots, a_n)^T \neq \mathbf{0}$ je její normála a uvedenou rovnici můžeme zapsat ve tvaru $\langle \mathbf{x}, \mathbf{a} \rangle = b$.

Odchylka dvou afinních podprostorů VI

Tvrzení

Nechť S je netriviální, vlastní lineární podprostor euklidovského prostoru V a $\mathbf{0} \neq \mathbf{a} \in V$. Potom

$$\sphericalangle\left([\mathbf{a}]^{\perp}, S\right) = \frac{\pi}{2} - \sphericalangle(\mathbf{a}, S) = \sphericalangle\left(\mathbf{a}, S^{\perp}\right).$$

Odchylka dvou afinních podprostorů VI

Tvrzení

Nechť S je netriviální, vlastní lineární podprostor euklidovského prostoru V a $\mathbf{0} \neq \mathbf{a} \in V$. Potom

$$\sphericalangle\big([\mathbf{a}]^{\perp}, S\big) = \frac{\pi}{2} - \sphericalangle(\mathbf{a}, S) = \sphericalangle\big(\mathbf{a}, S^{\perp}\big).$$

Důsledek

Nechť M, N jsou dvě nadroviny v euklidovském prostoru V s normálami $\mathbf{a}$, resp. $\mathbf{b}$. Potom

$$\sphericalangle(M, N) = \sphericalangle(\mathbf{a}, [\mathbf{b}]) = \min\big\{\sphericalangle(\mathbf{a}, \mathbf{b}),\ \sphericalangle(\mathbf{a}, -\mathbf{b})\big\}.$$

Odchylka dvou afinních podprostorů VII

Příklad

V euklidovském prostoru V vypočteme odchylku roviny $S = [\mathbf{u}, \mathbf{v}]$ a nadroviny $T = [\mathbf{a}]^{\perp}$.
Podle předcházejícího tvrzení platí

$$\sphericalangle(S, T) = \frac{\pi}{2} - \sphericalangle(\mathbf{a}, S) = \arcsin \frac{\|\mathbf{a}_S\|}{\|\mathbf{a}\|}.$$

Odchylka dvou afinních podprostorů VII

Příklad

V euklidovském prostoru V vypočteme odchylku roviny $S = [\mathbf{u}, \mathbf{v}]$ a nadroviny $T = [\mathbf{a}]^\perp$.
Podle předcházejícího tvrzení platí

$$\sphericalangle(S, T) = \frac{\pi}{2} - \sphericalangle(\mathbf{a}, S) = \arcsin \frac{\|\mathbf{a}_S\|}{\|\mathbf{a}\|}.$$

Souřadnice $\mathbf{c}$, $\mathbf{d}$ kolmého průmětu $\mathbf{a}_S = c\mathbf{u} + d\mathbf{v}$ vzhledem k bázi $(\mathbf{u}, \mathbf{v})$ podprostoru S získáme řešením soustavy

$$\mathbf{G}(\mathbf{u}, \mathbf{v}) \cdot \begin{pmatrix} c \\ d \end{pmatrix} = \begin{pmatrix} \langle \mathbf{a}, \mathbf{u} \rangle \\ \langle \mathbf{a}, \mathbf{v} \rangle \end{pmatrix}$$

pomocí Cramerova pravidla.

Odchylka dvou afinních podprostorů VIII

Platí

$$c = \frac{\begin{vmatrix} \langle \mathbf{a}, \mathbf{u} \rangle & \langle \mathbf{v}, \mathbf{u} \rangle \\ \langle \mathbf{a}, \mathbf{v} \rangle & \langle \mathbf{v}, \mathbf{v} \rangle \end{vmatrix}}{|\mathbf{G}(\mathbf{u}, \mathbf{v})|}, \qquad d = \frac{\begin{vmatrix} \langle \mathbf{u}, \mathbf{u} \rangle & \langle \mathbf{a}, \mathbf{u} \rangle \\ \langle \mathbf{u}, \mathbf{v} \rangle & \langle \mathbf{a}, \mathbf{v} \rangle \end{vmatrix}}{|\mathbf{G}(\mathbf{u}, \mathbf{v})|}.$$

Řešení neřešitelných soustav a lineární regrese I

V celém tomto paragrafu označují m, n pevná kladná čísla. Sloupcové vektorové prostory $\mathbb{R}^m$ a $\mathbb{R}^n$ jsou opatřeny standardním skalárním součinem, takže tvoří euklidovský prostor.

Řešení neřešitelných soustav a lineární regrese I

V celém tomto paragrafu označují m, n pevná kladná čísla.
Sloupcové vektorové prostory $\mathbb{R}^m$ a $\mathbb{R}^n$ jsou opatřeny standardním skalárním součinem, takže tvoří euklidovský prostor.

Nechť $\mathbf{A} \in \mathbb{R}^{m \times n}$, $\mathbf{b} \in \mathbb{R}^m$. Uvažujme soustavu lineárních rovnic

$$\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$$

a označme $S = [\mathbf{s}_1(\mathbf{A}), \ldots, \mathbf{s}_n(\mathbf{A})]$ lineární podprostor v $\mathbb{R}^m$ generovaný sloupci matice $\mathbf{A}$.

Řešení neřešitelných soustav a lineární regrese I

V celém tomto paragrafu označují m, n pevná kladná čísla. Sloupcové vektorové prostory $\mathbb{R}^m$ a $\mathbb{R}^n$ jsou opatřeny standardním skalárním součinem, takže tvoří euklidovský prostor.

Nechť $\mathbf{A} \in \mathbb{R}^{m \times n}$, $\mathbf{b} \in \mathbb{R}^m$. Uvažujme soustavu lineárních rovnic

$$\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$$

a označme $S = [\mathbf{s}_1(\mathbf{A}), \ldots, \mathbf{s}_n(\mathbf{A})]$ lineární podprostor v $\mathbb{R}^m$ generovaný sloupci matice $\mathbf{A}$.

Podle Frobeniova kritéria má naše soustava nějaké řešení $\mathbf{x} \in \mathbb{R}^n$ právě tehdy, když $\mathbf{b} \in S$. Složky řešení $\mathbf{x} = (x_1, \ldots, x_n)^T \in \mathbb{R}^n$ jsou pak koeficienty lineární kombinace

$$x_1 \mathbf{s}_1(\mathbf{A}) + \ldots + x_n \mathbf{s}_n(\mathbf{A}) = \mathbf{A} \cdot \mathbf{x} = \mathbf{b}.$$

Ale i v případě, kdy $\mathbf{b} \notin S$, tj. řešení soustavy neexistuje, se můžeme pokusit nahradit její pravou stranu $\mathbf{b}$ co nejbližším vektorem z podprostoru S. Takto získaná nová soustava už má řešení, které můžeme právem považovat za nejlepší možné přibližné řešení původní soustavy.

Řešení neřešitelných soustav a lineární regrese II

Podle věty o ortogonální projekci je nejbližší vektor z podprostoru S k vektoru $\mathbf{b}$ určený jednoznačně, a je to jeho kolmý průmět $\mathbf{b}_S$ do tohoto podprostoru.

Řešení neřešitelných soustav a lineární regrese II

Podle věty o ortogonální projekci je nejbližší vektor z podprostoru S k vektoru $\mathbf{b}$ určený jednoznačně, a je to jeho kolmý průmět $\mathbf{b}_S$ do tohoto podprostoru.

Pseudořešení (i neřešitelné) soustavy $\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$ definujeme jako řešení (tentokrát již jistě řešitelné) soustavy

$$\mathbf{A} \cdot \mathbf{x} = \mathbf{b}_S.$$

Řešení neřešitelných soustav a lineární regrese II

Podle věty o ortogonální projekci je nejbližší vektor z podprostoru S k vektoru $\mathbf{b}$ určený jednoznačně, a je to jeho kolmý průmět $\mathbf{b}_S$ do tohoto podprostoru.

Pseudořešení (i neřešitelné) soustavy $\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$ definujeme jako řešení (tentokrát již jistě řešitelné) soustavy

$$\mathbf{A} \cdot \mathbf{x} = \mathbf{b}_S.$$

Pokud je původní soustava řešitelná, tj. pokud $\mathbf{b} \in S$, tak $\mathbf{b}_S = \mathbf{b}$ a obě soustavy splývají, takže každé její pseudořešení je přímo řešením původní soustavy.

Řešení neřešitelných soustav a lineární regrese III

Tvrzení

Nechť $\mathbf{A} \in \mathbb{R}^{m\ cxn}$, $\mathbf{b} \in \mathbb{R}^m$. *Potom* $\mathbf{x} \in \mathbb{R}^n$ *je pseudořešením soustavy* $\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$ *právě když* $\mathbf{x}$ *je řešením soustavy*

$$\mathbf{A}^T \cdot \mathbf{A} \cdot \mathbf{x} = \mathbf{A}^T \cdot \mathbf{b}$$

se čtvercovou maticí $\mathbf{A}^T \cdot \mathbf{A} \in \mathbb{R}^{n \times n}$ *a levou stranou* $\mathbf{A}^T \cdot \mathbf{b} \in \mathbb{R}^n$.

Řešení neřešitelných soustav a lineární regrese III

Tvrzení

Nechť $\mathbf{A} \in \mathbb{R}^{m\ cxn}$, $\mathbf{b} \in \mathbb{R}^m$. *Potom* $\mathbf{x} \in \mathbb{R}^n$ *je pseudořešením soustavy* $\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$ *právě když* $\mathbf{x}$ *je řešením soustavy*

$$\mathbf{A}^T \cdot \mathbf{A} \cdot \mathbf{x} = \mathbf{A}^T \cdot \mathbf{b}$$

se čtvercovou maticí $\mathbf{A}^T \cdot \mathbf{A} \in \mathbb{R}^{n \times n}$ *a levou stranou* $\mathbf{A}^T \cdot \mathbf{b} \in \mathbb{R}^n$.

Pseudořešení soustavy $\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$ tedy hledáme jako řešení zaručeně řešitelné soustavy $\mathbf{A}^T \cdot \mathbf{A} \cdot \mathbf{x} = \mathbf{A}^T \cdot \mathbf{b}$.

Řešení neřešitelných soustav a lineární regrese III

Tvrzení

Nechť $\mathbf{A} \in \mathbb{R}^{m\ cxn}$, $\mathbf{b} \in \mathbb{R}^m$. *Potom* $\mathbf{x} \in \mathbb{R}^n$ *je pseudořešením soustavy* $\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$ *právě když* $\mathbf{x}$ *je řešením soustavy*

$$\mathbf{A}^T \cdot \mathbf{A} \cdot \mathbf{x} = \mathbf{A}^T \cdot \mathbf{b}$$

se čtvercovou maticí $\mathbf{A}^T \cdot \mathbf{A} \in \mathbb{R}^{n \times n}$ *a levou stranou* $\mathbf{A}^T \cdot \mathbf{b} \in \mathbb{R}^n$.

Pseudořešení soustavy $\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$ tedy hledáme jako řešení zaručeně řešitelné soustavy $\mathbf{A}^T \cdot \mathbf{A} \cdot \mathbf{x} = \mathbf{A}^T \cdot \mathbf{b}$.

V typickém případě má původní soustava více rovnic než neznámých neboli $m > n$ a $\mathbf{A}$ je obdélníková matice, „vyšší než širší“.

Řešení neřešitelných soustav a lineární regrese III

Tvrzení

Nechť $\mathbf{A} \in \mathbb{R}^{m\ cxn}$, $\mathbf{b} \in \mathbb{R}^m$. *Potom* $\mathbf{x} \in \mathbb{R}^n$ *je pseudořešením soustavy* $\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$ *právě když* $\mathbf{x}$ *je řešením soustavy*

$$\mathbf{A}^T \cdot \mathbf{A} \cdot \mathbf{x} = \mathbf{A}^T \cdot \mathbf{b}$$

se čtvercovou maticí $\mathbf{A}^T \cdot \mathbf{A} \in \mathbb{R}^{n \times n}$ *a levou stranou* $\mathbf{A}^T \cdot \mathbf{b} \in \mathbb{R}^n$.

Pseudořešení soustavy $\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$ tedy hledáme jako řešení zaručeně řešitelné soustavy $\mathbf{A}^T \cdot \mathbf{A} \cdot \mathbf{x} = \mathbf{A}^T \cdot \mathbf{b}$.

V typickém případě má původní soustava více rovnic než neznámých neboli $m > n$ a $\mathbf{A}$ je obdélníková matice, „vyšší než širší“.

Pak je velmi pravděpodobné, že čtvercová matice $\mathbf{A}^T \cdot \mathbf{A}$ řádu large n (jako Gramova matice „malého“ počtu sloupcových vektorů v euklidovském prostoru „velké“ dimenze) je regulární, tedy k ní existuje regulární matice $\left(\mathbf{A}^T \cdot \mathbf{A}\right)^{-1}$.

Řešení neřešitelných soustav a lineární regrese IV

V takovémto případě je pseudořešení původní soustavy určené jednoznačně:

$$\mathbf{X} = \left(\mathbf{A}^T \cdot \mathbf{A}\right)^{-1} \cdot \mathbf{A}^T \cdot \mathbf{b}.$$

Samozřejmě, pokud $m = n$ a už samotná matice $\mathbf{A}$ je regulární, dostáváme

$$\left(\mathbf{A}^T \cdot \mathbf{A}\right)^{-1} \cdot \mathbf{A}^T = \mathbf{A}^{-1}$$

a $\mathbf{x} = \mathbf{A}^{-1} \cdot \mathbf{b}$ je přímo jediným řešením původní soustavy.

Řešení neřešitelných soustav a lineární regrese IV

V takovémto případě je pseudořešení původní soustavy určené jednoznačně:

$$\mathbf{X} = \left(\mathbf{A}^T \cdot \mathbf{A}\right)^{-1} \cdot \mathbf{A}^T \cdot \mathbf{b}.$$

Samozřejmě, pokud $m = n$ a už samotná matice $\mathbf{A}$ je regulární, dostáváme

$$\left(\mathbf{A}^T \cdot \mathbf{A}\right)^{-1} \cdot \mathbf{A}^T = \mathbf{A}^{-1}$$

a $\mathbf{x} = \mathbf{A}^{-1} \cdot \mathbf{b}$ je přímo jediným řešením původní soustavy.

V úlohách ***lineární regrese*** máme zadané hodnoty $y_1, \dots, y_m$ neznámé funkce f v bodech $x_1, \dots, x_m$ jejího definičního oboru, získané většinou měřením. Funkci f chceme aproximovat lineární kombinací funkcí $f_1, \dots, f_n$, které známe, či alespoň jsou nám známé jejich hodnoty $a_{ij} = f_j(x_i)$ v bodech $x_1, \dots, x_m$.

Řešení neřešitelných soustav a lineární regrese V

Obvykle je m podstatně větší než n. V optimálním případě se nám může podařit sestrojit funkci $f = c_1 f_1 + \ldots + c_n f_n$ přímo jako lineární kombinaci funkcí f_j tak, aby f v bodech x_i nabývala předem předepsané hodnoty y_i, tj.

$$y_i = f(x_i) = \sum_{j=1}^{n} c_j f_j(x_i) = \sum_{j=1}^{n} a_{ij} c_j.$$

Řešení neřešitelných soustav a lineární regrese V

Obvykle je m podstatně větší než n. V optimálním případě se nám může podařit sestrojit funkci $f = c_1 f_1 + \ldots + c_n f_n$ přímo jako lineární kombinaci funkcí f_j tak, aby f v bodech x_i nabývala předem předepsané hodnoty y_i, tj.

$$y_i = f(x_i) = \sum_{j=1}^{n} c_j f_j(x_i) = \sum_{j=1}^{n} a_{ij} c_j.$$

Pokud označíme $\mathbf{A} = (a_{ij}) \in \mathbb{R}^{m \times n}$, $\mathbf{y} = (y_1, \ldots, y_m)^T \in \mathbb{R}^m$, $\mathbf{c} = (c_1, \ldots, c_n)^T \in \mathbb{R}^n$, vidíme, že vlastně hledáme řešení $\mathbf{c}$ soustavy

$$\mathbf{A} \cdot \mathbf{c} = \mathbf{y}.$$

Řešení neřešitelných soustav a lineární regrese V

Obvykle je m podstatně větší než n. V optimálním případě se nám může podařit sestrojit funkci $f = c_1 f_1 + \ldots + c_n f_n$ přímo jako lineární kombinaci funkcí f_j tak, aby f v bodech x_i nabývala předem předepsané hodnoty y_i, tj.

$$y_i = f(x_i) = \sum_{j=1}^{n} c_j f_j(x_i) = \sum_{j=1}^{n} a_{ij} c_j.$$

Pokud označíme $\mathbf{A} = (a_{ij}) \in \mathbb{R}^{m \times n}$, $\mathbf{y} = (y_1, \ldots, y_m)^T \in \mathbb{R}^m$, $\mathbf{c} = (c_1, \ldots, c_n)^T \in \mathbb{R}^n$, vidíme, že vlastně hledáme řešení $\mathbf{c}$ soustavy

$$\mathbf{A} \cdot \mathbf{c} = \mathbf{y}.$$

Tato soustava je v typickém případě neřešitelná.

Řešení neřešitelných soustav a lineární regrese VI

Úloha lineární regrese pak splývá s ***metodou nejmenších čtverců*** a spočívá v nalezení takových koeficientů c_j, které minimalizují výraz

$$\sum_{i=1}^{m}\left(y_i - \sum_{j=1}^{n} a_{ij} c_j\right)^2 = \|\mathbf{y} - \mathbf{A} \cdot \mathbf{c}\|^2 .$$

Řešení neřešitelných soustav a lineární regrese VI

Úloha lineární regrese pak splývá s ***metodou nejmenších čtverců*** a spočívá v nalezení takových koeficientů c_j, které minimalizují výraz

$$\sum_{i=1}^{m}\Big(y_i - \sum_{j=1}^{n} a_{ij} c_j\Big)^2 = \|\mathbf{y} - \mathbf{A}\cdot\mathbf{c}\|^2 .$$

Toto minimum nastává pro $\mathbf{c}$ takové, že $\mathbf{A}\cdot\mathbf{c} = \mathbf{y}_S$, kde S je podprostor v $\mathbb{R}^m$ generovaný sloupci matice $\mathbf{A}$. Jinak řečeno, hledanou lineární kombinaci dostaneme jako pseudořešení $\mathbf{c}$ soustavy $\mathbf{A}\cdot\mathbf{c} = \mathbf{y}$.

Řešení neřešitelných soustav a lineární regrese VI

Úloha lineární regrese pak splývá s ***metodou nejmenších čtverců*** a spočívá v nalezení takových koeficientů c_j, které minimalizují výraz

$$\sum_{i=1}^{m}\Big(y_i - \sum_{j=1}^{n} a_{ij}c_j\Big)^2 = \|\mathbf{y} - \mathbf{A}\cdot\mathbf{c}\|^2 .$$

Toto minimum nastává pro $\mathbf{c}$ takové, že $\mathbf{A}\cdot\mathbf{c} = \mathbf{y}_S$, kde S je podprostor v $\mathbb{R}^m$ generovaný sloupci matice $\mathbf{A}$. Jinak řečeno, hledanou lineární kombinaci dostaneme jako pseudořešení $\mathbf{c}$ soustavy $\mathbf{A}\cdot\mathbf{c} = \mathbf{y}$.

Pro hodnoty pocházející z rozumných praktických úkolů je téměř jisté, že matice $\mathbf{A}^T\cdot\mathbf{A}$ je regulární. V takovémto případě

$$\mathbf{C} = \left(\mathbf{A}^T\cdot\mathbf{A}\right)^{-1}\cdot\mathbf{A}^T\cdot\mathbf{y},$$

čili hledaná lineární kombinace
$f = c_1f_1 + \ldots + c_nf_n = (f_1, \ldots, f_n)\cdot\mathbf{c}$ je určena jednoznačně.

Řešení neřešitelných soustav a lineární regrese VII

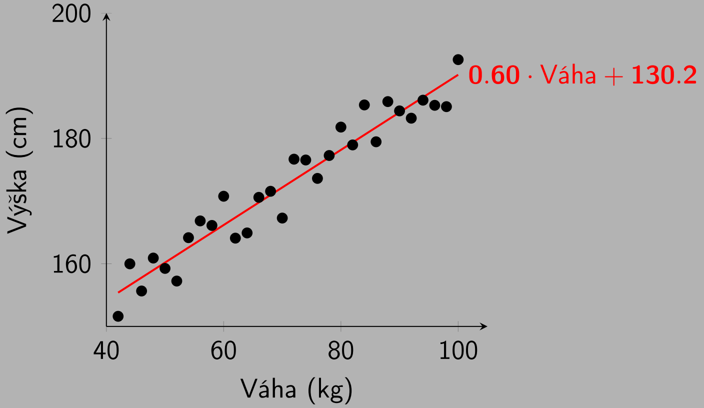
Příklad

V rovině $\mathbb{R}^2$ je daných $m \geq 2$ bodů $(x_1, y_1), \ldots, (x_m, y_m)$, získaných měřením hodnot nějaké neznámé funkce f ve vybraných bodech x_i jejího definičního oboru. Tuto funkci hodláme aproximovat přímkou s rovnicí $y = a + bx$ tak, aby výraz $\sum_{i=1}^{m}(y_i - a - bx_i)^2$ byl minimální.

Řešení neřešitelných soustav a lineární regrese VII

Příklad

V rovině $\mathbb{R}^2$ je daných $m \geq 2$ bodů $(x_1, y_1), \ldots, (x_m, y_m)$, získaných měřením hodnot nějaké neznámé funkce f ve vybraných bodech x_i jejího definičního oboru. Tuto funkci hodláme aproximovat přímkou s rovnicí $y = a + bx$ tak, aby výraz $\sum_{i=1}^{m}(y_i - a - bx_i)^2$ byl minimální.



Řešení neřešitelných soustav a lineární regrese VIII

Pokud si uvědomíme, že funkce $y = a + bx$ je lineární kombinací konstantní funkce $y = 1$ a identické funkce $y = x$, hned vidíme, že jde o úlohu lineární regrese.

Řešení neřešitelných soustav a lineární regrese VIII

Pokud si uvědomíme, že funkce $y = a + bx$ je lineární kombinací konstantní funkce $y = 1$ a identické funkce $y = x$, hned vidíme, že jde o úlohu lineární regrese.

Soustava

$$\begin{pmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_m \end{pmatrix} \cdot \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}$$

je kromě triviálního případu, kdy všechny body (x_i, y_i) leží na jedné přímce, neřešitelná. Koeficienty a, b tedy najdeme jako pseudořešení této soustavy. Její matici si označíme $\mathbf{A}$.

Řešení neřešitelných soustav a lineární regrese VIII

Pokud si uvědomíme, že funkce $y = a + bx$ je lineární kombinací konstantní funkce $y = 1$ a identické funkce $y = x$, hned vidíme, že jde o úlohu lineární regrese.

Soustava

$$\begin{pmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_m \end{pmatrix} \cdot \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}$$

je kromě triviálního případu, kdy všechny body (x_i, y_i) leží na jedné přímce, neřešitelná. Koeficienty a, b tedy najdeme jako pseudořešení této soustavy. Její matici si označíme $\mathbf{A}$.

Jednoduchý výpočet dává

$$\mathbf{A}^T \cdot \mathbf{A} = \begin{pmatrix} m & \sum x_i \\ \sum x_i & \sum x_i^2 \end{pmatrix}.$$

Řešení neřešitelných soustav a lineární regrese IX

Tedy platí

$$\det\left(\mathbf{A}^T \cdot \mathbf{A}\right) = m \sum_{i=1}^{m} x_i^2 - \left(\sum_{i=1}^{m} x_i \right)^2 = \sum_{i<j} (x_i - x_j)^2.$$

Odtud $\det\left(\mathbf{A}^T \cdot \mathbf{A}\right) = 0$ právě tehdy, když $x_1 = x_2 = \ldots = x_m$.

Asi se shodneme na tom, že v takovémto případě by šlo, mírně řečeno, o velmi nerozumně postavenou úlohu.

Naopak, na základě její formulace je přirozené očekávat, že všechny hodnoty x_i budou navzájem různé.
V jen trochu rozumném případě tedy dostáváme jednoznačně určené koeficienty a, b ve tvaru

$$\begin{pmatrix} a \\ b \end{pmatrix} = \left(\mathbf{A}^T \cdot \mathbf{A}\right)^{-1} \cdot \mathbf{A}^T \cdot \mathbf{y}.$$

Řešení neřešitelných soustav a lineární regrese X

Po úpravě

$$\begin{aligned}
\begin{pmatrix} a \\ b \end{pmatrix} &= \frac{1}{\sum_{i<j}(x_i - x_j)^2} \begin{pmatrix} \sum x_i^2 & -\sum x_i \\ -\sum x_i & m \end{pmatrix} \cdot \begin{pmatrix} \sum y_i \\ \sum x_i y_i \end{pmatrix} \\
&= \frac{1}{\sum_{i<j}(x_i - x_j)^2} \begin{pmatrix} \sum x_i^2 \sum y_i - \sum x_i \sum x_i y_i \\ m \sum x_i y_i - \sum x_i \sum y_i \end{pmatrix} \\
&= \frac{1}{\sum_{i<j}(x_i - x_j)^2} \begin{pmatrix} \sum_{i<j}(x_i - x_j)(x_i y_j - x_j y_i) \\ \sum_{i<j}(x_i - x_j)(y_i - y_j) \end{pmatrix} .
\end{aligned}$$

Minimální řešení soustavy I

Pokud má soustava $\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$ více, tj. nekonečně mnoho řešení, může nám velikost řešení (ve smyslu euklidovské normy) posloužit jako dodatečné kritérium výběru toho nejvhodnějšího řešení (např. nejlevnějšího) z možných řešení.

Minimální řešení soustavy I

Pokud má soustava $\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$ více, tj. nekonečně mnoho řešení, může nám velikost řešení (ve smyslu euklidovské normy) posloužit jako dodatečné kritérium výběru toho nejvhodnějšího řešení (např. nejlevnějšího) z možných řešení.

Říkáme, že $\mathbf{x} \in \mathbb{R}^n$ je ***minimální řešení*** uvedené soustavy, pokud je jejím řešením a pro každé řešení $\mathbf{y}$ této soustavy platí $\|\mathbf{x}\| \leq \|\mathbf{y}\|$.

Minimální řešení soustavy I

Pokud má soustava $\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$ více, tj. nekonečně mnoho řešení, může nám velikost řešení (ve smyslu euklidovské normy) posloužit jako dodatečné kritérium výběru toho nejvhodnějšího řešení (např. nejlevnějšího) z možných řešení.

Říkáme, že $\mathbf{x} \in \mathbb{R}^n$ je ***minimální řešení*** uvedené soustavy, pokud je jejím řešením a pro každé řešení $\mathbf{y}$ této soustavy platí $\|\mathbf{x}\| \leq \|\mathbf{y}\|$.

Samozřejmě, pokud má soustava jediné řešení, tak je to zároveň i minimální řešení.

Minimální řešení soustavy I

Pokud má soustava $\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$ více, tj. nekonečně mnoho řešení, může nám velikost řešení (ve smyslu euklidovské normy) posloužit jako dodatečné kritérium výběru toho nejvhodnějšího řešení (např. nejlevnějšího) z možných řešení.

Říkáme, že $\mathbf{x} \in \mathbb{R}^n$ je ***minimální řešení*** uvedené soustavy, pokud je jejím řešením a pro každé řešení $\mathbf{y}$ této soustavy platí $\|\mathbf{x}\| \leq \|\mathbf{y}\|$.

Samozřejmě, pokud má soustava jediné řešení, tak je to zároveň i minimální řešení.

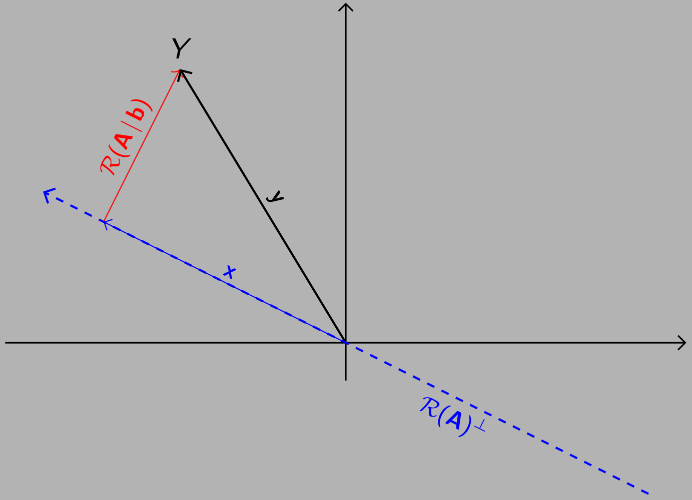
Tvrzení

Nechť $\mathbf{A} \in \mathbb{R}^{m \times n}$, $\mathbf{b} \in \mathbb{R}^m$. *Potom* $\mathbf{x} \in \mathbb{R}^n$ *je minimálním řešením soustavy* $\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$ *právě tehdy, když má tvar* $\mathbf{x} = \mathbf{A}^T \cdot \mathbf{z}$, *kde* $\mathbf{z} \in \mathbb{R}^m$ *je řešením soustavy*

$$\mathbf{A} \cdot \mathbf{A}^T \cdot \mathbf{z} = \mathbf{b}$$

se čtvercovou maticí $\mathbf{A} \cdot \mathbf{A}^T \in \mathbb{R}^{m \times m}$ *a levou stranou* $\mathbf{b} \in \mathbb{R}^m$.

Minimální řešení soustavy II



Minimální řešení soustavy III

Minimální řešení $\mathbf{x}$ řešitelné soustavy $\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$ tedy hledáme tak, že nejprve najdeme nějaké (libovolné) řešení $\mathbf{z}$ soustavy $\mathbf{A} \cdot \mathbf{A}^T \cdot \mathbf{z} = \mathbf{b}$ a položíme $\mathbf{x} = \mathbf{A}^T \cdot \mathbf{z}$. V typickém případě má původní soustava méně rovnic než neznámých neboli $m < n$ a $\mathbf{A}$ je obdélníková matice, „širší než vyšší“. Pak je velmi pravděpodobné, že čtvercová matice $\mathbf{A} \cdot \mathbf{A}^T$ řádu m (jako Gramova matice „malého“ počtu řádkových vektorů v (řádkovém) euklidovském prostoru „velké“ dimenze) je regulární, tedy k ní existuje regulární matice $\left(\mathbf{A} \cdot \mathbf{A}^T\right)^{-1}$. V tomto případě jsou tedy

$$\mathbf{z} = \left(\mathbf{A} \cdot \mathbf{A}^T\right)^{-1} \cdot \mathbf{b},$$

a minimální řešení původní soustavy

$$\mathbf{x} = \mathbf{A}^T \cdot \mathbf{z} = \mathbf{A}^T \cdot \left(\mathbf{A} \cdot \mathbf{A}^T\right)^{-1} \cdot \mathbf{b}$$

určené jednoznačně.