14. ORTOGONÁLNE PROJEKCIE A PODPRIESTORY

V tejto kapitole budeme pokračovať v štúdiu euklidovských priestorov s cieľom podať kvantitatívny popis vzájomnej polohy afinných podpriestorov v takomto priestore pomocou dvoch základných parametrov – ich *vzdialenosti* a *odchýlky (uhla)*. Našim hlavným nástrojom pri tom budú lineárne operátory *kolmého priemetu*, zvané tiež *ortogonálne projekcie*, vektorov do lineárnych podpriestorov. V závere kapitoly predvedieme dve aplikácie rozpracovaných pojmov a metód – jednak ich využijeme pri konštrukcii „najlepších približných riešení“ neriešiteľných sústav lineárnych rovníc a lineárnej regresii, jednak sa oboznámime s formuláciu základných pojmov teórie pravdepodobnosti v jazyku euklidovských priestorov.

14.1. Ortokomplement a ortogonálna projekcia

Relácia ortogonality (kolmosti) vo vektorovom priestore so skalárnym súčinom má niekoľko zrejmých vlastností, ktoré tu zaznamenáme bez dôkazu.

14.1.1. Tvrdenie. *Nech V je vektorový priestor so skalárnym súčinom. Potom pre ľubovoľné $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in V$, $c, d \in \mathbb{R}$ platí:*
(a) $\boldsymbol{x} \perp \mathbf{0}$;
(b) $\boldsymbol{x} \perp \boldsymbol{x} \;\Leftrightarrow\; \boldsymbol{x} = \mathbf{0}$;
(c) $\boldsymbol{x} \perp \boldsymbol{y} \;\Leftrightarrow\; \boldsymbol{y} \perp \boldsymbol{x}$;
(d) $\big(\boldsymbol{x} \perp \boldsymbol{y} \;\&\; \boldsymbol{x} \perp \boldsymbol{z}\big) \;\Rightarrow\; \boldsymbol{x} \perp (c\boldsymbol{y} + d\boldsymbol{z})$.

Ortogonálnym doplnkom alebo tiež *ortokomplementom* ľubovoľnej množiny $X \subseteq V$ vo vektorovom priestore so skalárnym súčinom nazveme množinu

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

všetkých vektorov $\boldsymbol{y} \in V$ kolmých na každý vektor $\boldsymbol{x} \in X$.

14.1.2. Tvrdenie. *Nech V je vektorový priestor so skalárnym súčinom. Potom pre ľubovoľné 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) $X \cap X^\perp = \{\mathbf{0}\}$, *ak* $\mathbf{0} \in X$, *a* $X \cap X^\perp = \emptyset$, *ak* $\mathbf{0} \notin X$;
(g) $(X \cup Y)^\perp = (X + Y)^\perp = X^\perp \cap Y^\perp$.

Dôkaz. Jednotlivé podmienky sú priamymi dôsledkami predošlého tvrdenia, ich jednoduché dôkazy preto prenechávame ako cvičenie čitateľovi. Na ukážku predvedieme, ako vyplýva (e) z podmienok (c) a (d). Podľa (d) platí $X \subseteq X^{\perp\perp}$, z čoho podľa (c) vyplýva $X^{\perp\perp\perp} \subseteq X^\perp$. Obrátená inklúzia $X^\perp \subseteq X^{\perp\perp\perp}$ je opäť dôsledkom (d).

Z podmienky (b) okrem iného vyplýva, že $X^\perp$ je lineárnym podpriestorom vo V pre *každú podmnožinu* $X \subseteq V$.

Nech $S \subseteq V$ je lineárny podpriestor priestoru so skalárnym súčinom V a $\boldsymbol{x} \in V$. Hovoríme, že vektor $\boldsymbol{z} \in S$ je *kolmým priemetom* alebo tiež *ortogonálnou projekciou* vektora $\boldsymbol{x}$ do podpriestoru S, ak $\boldsymbol{x} - \boldsymbol{z} \in S^\perp$. Tento vektor (ak existuje) budeme značiť $\boldsymbol{z} = \mathrm{pr}_S(\boldsymbol{x}) = \boldsymbol{x}_S$.

14.1.3. Veta. *Nech V je vektorový priestor so skalárnym súčinom, $S \subseteq V$ je jeho konečnorozmerný lineárny podpriestor a $\boldsymbol{x} \in V$. Potom*
(a) *kolmý priemet vektora $\boldsymbol{x}$ do podpriestoru S existuje a je jednoznačne určený rovnosťou*

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

kde $(\boldsymbol{u}_1, \dots, \boldsymbol{u}_k)$ je ľubovoľná ortonormálna báza podpriestoru S;
(b) *pre ľubovoľný vektor $\boldsymbol{y} \in S$ platí*

$$\|\boldsymbol{x} - \boldsymbol{x}_S\| \le \|\boldsymbol{x} - \boldsymbol{y}\|$$

pričom rovnosť nastane práve vtedy, keď $\boldsymbol{y} = \boldsymbol{x}_S$;
(c) *ak $\boldsymbol{x} \ne \mathbf{0}$ a $S \ne \{\mathbf{0}\}$, tak pre ľubovoľný vektor $\mathbf{0} \ne \boldsymbol{y} \in S$ platí*

$$\frac{\|\boldsymbol{x}_S\|}{\|\boldsymbol{x}\|} \ge \frac{|\langle \boldsymbol{x}, \boldsymbol{y} \rangle|}{\|\boldsymbol{x}\|\,\|\boldsymbol{y}\|},$$

pričom rovnosť nastane práve vtedy, keď vektory $\boldsymbol{x}_S$, $\boldsymbol{y}$ sú lineárne závislé.

Dôkaz. (a) Nech $(\boldsymbol{u}_1, \dots, \boldsymbol{u}_k)$ je ľubovoľná ortonormálna báza podpriestoru S. Ak má kolmý priemet vektora $\boldsymbol{x}$ do S existovať, musí mať tvar $\boldsymbol{x}_S = c_1\boldsymbol{u}_1 + \ldots + c_k\boldsymbol{u}_k$ pre nejaké $c_1, \dots, c_k \in \mathbb{R}$. Podmienka $\boldsymbol{x} - \boldsymbol{x}_S \in S^\perp$ je podľa predchádzajúceho tvrdenia ekvivalentná s konjunkciou podmienok $\boldsymbol{x} - \boldsymbol{x}_S \perp \boldsymbol{u}_i$ pre každé $i \le k$. Z toho vyplýva

$$0 = \langle \boldsymbol{x} - \boldsymbol{x}_S, \boldsymbol{u}_i \rangle = \langle \boldsymbol{x}, \boldsymbol{u}_i \rangle - \sum_{j=1}^{k} c_i \langle \boldsymbol{u}_j, \boldsymbol{u}_i \rangle = \langle \boldsymbol{x}, \boldsymbol{u}_i \rangle - c_i,$$

teda lineárna kombinácia $c_1\boldsymbol{u}_1 + \ldots + c_k\boldsymbol{u}_k$ je kolmým priemetom $\boldsymbol{x}$ do S práve vtedy, keď $c_i = \langle \boldsymbol{x}, \boldsymbol{u}_i \rangle$, pre každé i.

(b) Nech $\boldsymbol{y} \in S$. Potom tiež $\boldsymbol{y} - \boldsymbol{x}_S \in S$, teda $(\boldsymbol{x} - \boldsymbol{x}_S) \perp (\boldsymbol{y} - \boldsymbol{x}_S)$. Podľa Pytagorovej vety platí

$$\|\boldsymbol{x} - \boldsymbol{y}\|^2 = \|(\boldsymbol{x} - \boldsymbol{x}_S) + (\boldsymbol{y} - \boldsymbol{x}_S)\|^2 = \|\boldsymbol{x} - \boldsymbol{x}_S\|^2 + \|\boldsymbol{y} - \boldsymbol{x}_S\|^2 \ge \|\boldsymbol{x} - \boldsymbol{x}_S\|^2,$$

pričom rovnosť nastane práve vtedy, keď $\boldsymbol{y} = \boldsymbol{x}_S$.

(c) Pre $\boldsymbol{y} \in S$ máme $(\boldsymbol{x} - \boldsymbol{x}_S) \perp \boldsymbol{y}$, preto $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \langle \boldsymbol{x}_S, \boldsymbol{y} \rangle + \langle \boldsymbol{x} - \boldsymbol{x}_S, \boldsymbol{y} \rangle = \langle \boldsymbol{x}_S, \boldsymbol{y} \rangle$. Ak $\boldsymbol{y} \ne \mathbf{0}$, tak s použitím Cauchyho-Schwartzovej nerovnosti z toho dostávame

$$\frac{|\langle \boldsymbol{x}, \boldsymbol{y} \rangle|}{\|\boldsymbol{x}\|\,\|\boldsymbol{y}\|} = \frac{|\langle \boldsymbol{x}_S, \boldsymbol{y} \rangle|}{\|\boldsymbol{x}\|\,\|\boldsymbol{y}\|} \le \frac{\|\boldsymbol{x}_S\|\,\|\boldsymbol{y}\|}{\|\boldsymbol{x}\|\,\|\boldsymbol{y}\|} = \frac{\|\boldsymbol{x}_S\|}{\|\boldsymbol{x}\|}.$$

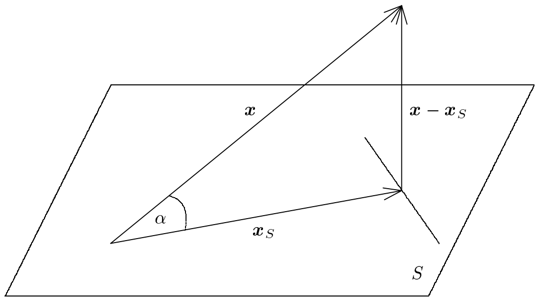
Rovnosť zrejme nastane práve vtedy, keď vektory $\boldsymbol{x}_S$, $\boldsymbol{y}$ sú lineárne závislé.

Podmienka (a) práve dokázanej vety má nasledujúci dôsledok.

14.1.4. Dôsledok. *Nech V je vektorový priestor so skalárnym súčinom a S, $T \subseteq V$ sú jeho konečnorozmerné lineárne podpriestory. 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 lineárny operátor;*
(c) $(\forall\, \boldsymbol{x} \in V)(\boldsymbol{x} \in S \ \Leftrightarrow\ \mathrm{pr}_S(\boldsymbol{x}) = \boldsymbol{x})$;
(d) $\mathrm{Im}\,\mathrm{pr}_S = S$ *a* $\mathrm{Ker}\,\mathrm{pr}_S = S^\perp$;
(e) $\boldsymbol{x} - \boldsymbol{x}_S$ *je kolmým priemetom vektora* $\boldsymbol{x}$ *do podpriestoru* $S^\perp$.

Lineárny operátor pr_S nazývame *ortogonálnou projekciou* na podpriestor S.

Poznámka. Ak V je euklidovský priestor, tak veta 14.1.3 a dôsledok 14.1.4 samozrejme platia pre *ľubovoľný* podpriestor $S \subseteq V$. Za istých okolností, ktorých rozbor však presahuje rámec lineárnej algebry (tzv. úplnosť priestoru V a uzavretosť podpriestoru S), spomínané výsledky platia aj pre nekonečnorozmerné podpriestory. V cvičeniach uvedieme príklad vektorového priestoru so skalárnym súčinom V a jeho nekonečnorozmerného podpriestoru $S \neq V$ takého, že $S^\perp = \{\boldsymbol{0}\}$. Potom pre žiaden vektor $\boldsymbol{x} \in V \smallsetminus S$ nemôže existovať jeho kolmý priemet do S; rovnako $S \oplus S^\perp = S \neq V$.



Predpokladajme, že kolmý priemet vektora $\boldsymbol{x}$ do lineárneho podpriestoru S existuje. Vysvetlíme si, ako možno za tohto predpokladu definovať vzdialenosť aj odchýlku vektora $\boldsymbol{x}$ od každého z podpriestorov S, $S^\perp$. Situácia je znázornená na obrázku. Vektor $\boldsymbol{x} - \boldsymbol{x}_S$ je kolmý na každú priamku v podpriestore S, špeciálne trojuholník tvorený vektormi $\boldsymbol{x}$, $\boldsymbol{x}_S$, $\boldsymbol{x} - \boldsymbol{x}_S$ je pravouhlý, s pravým uhlom pri „konci“ vektora $\boldsymbol{x}_S$.

Podmienka (b) vety 14.1.3 nás oprávňuje nazvať dĺžku vektora $\boldsymbol{x} - \boldsymbol{x}_S$ *vzdialenosťou* vektora $\boldsymbol{x}$ od podpriestoru S. Budeme ju značiť

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

Vzhľadom na podmienku (e) dôsledku 14.1.4 je vzdialenosť (anglicky *distance*) vektora $\boldsymbol{x}$ od podpriestoru $S^\perp$ daná vzťahom

$$\mathrm{dist}\big(\boldsymbol{x}, S^\perp\big) = \|\boldsymbol{x}_S\|.$$

Podobne, keďže kosínus je na intervale $\langle 0, \pi\rangle$ klesajúca funkcia, podmienka (c) vety 14.1.3 nás oprávňuje nazvať výraz

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

odchýlkou vektora $\boldsymbol{x} \neq \boldsymbol{0}$ od podpriestoru $S \neq \{\boldsymbol{0}\}$, prípadne *uhlom* vektora $\boldsymbol{x}$ a podpriestoru S. Odchýlka $\sphericalangle(\boldsymbol{x}, S)$ je teda jednoznačne určená ako také reálne číslo $\alpha \in \langle 0, \pi/2\rangle$, pre ktoré platí

$$\cos\alpha = \frac{\|\boldsymbol{x}_S\|}{\|\boldsymbol{x}\|}, \quad \text{alebo ekvivalentne,} \quad \sin\alpha = \frac{\|\boldsymbol{x} - \boldsymbol{x}_S\|}{\|\boldsymbol{x}\|}.$$

Zrejme opäť ide o *neorientovaný uhol.* Ak $\boldsymbol{x}_S \neq \boldsymbol{0}$, tak $\sphericalangle(\boldsymbol{x}, S) = \sphericalangle(\boldsymbol{x}, \boldsymbol{x}_S)$; ak $\boldsymbol{x}_S = \boldsymbol{0}$, t. j. ak $\boldsymbol{x} \in S^\perp$, tak samozrejme $\sphericalangle(\boldsymbol{x}, S) = \pi/2$. Ešte si všimnite, že zatiaľ čo uhol dvoch vektorov nadobúda hodnoty z intervalu $\langle 0, \pi\rangle$, hodnoty, ktoré nadobúda uhol vektora a podpriestoru, sú obmedzené na interval $\langle 0, \pi/2\rangle$. Vzhľadom na podmienku 14.1.3 (e), ak $S^\perp \neq \{\boldsymbol{0}\}$, tak odchýlka vektora $\boldsymbol{x} \neq \boldsymbol{0}$ od podpriestoru $S^\perp$ je daná vzťahom

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

Konečne časť (a) vety 14.1.3 nám dáva priamy návod, ako nájsť kolmý priemet vektora $\boldsymbol{x}$ do *konečnorozmerného* podpriestoru $S \subseteq V$, a tým aj vzdialenosti $\mathrm{dist}(\boldsymbol{x}, S)$, $\mathrm{dist}\big(\boldsymbol{x}, S^\perp\big)$ a odchýlky $\sphericalangle(\boldsymbol{x}, S)$, $\sphericalangle\big(\boldsymbol{x}, S^\perp\big)$ – potrebujeme však poznať aspoň jednu ortonormálnu bázu v S. Ak máme k dispozícii len nejakú „obyčajnú“ bázu podpriestoru S, je potrebné ju najprv ortonormalizovať, až potom možno použiť spomínaný vzorec. Diagonalizáciu Gramovej matice si však môžeme odpustiť, rovnaký výsledok sa totiž z nej dá získať aj priamo.

14.1.5. Tvrdenie. *Nech V je vektorový priestor so skalárnym súčinom, S je jeho konečnorozmerný lineárny podpriestor s bázou $\boldsymbol{\alpha} = (\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k)$ a $\boldsymbol{x} \in V$. Potom pre $\boldsymbol{c} = (c_1, \ldots, c_k)^T \in \mathbb{R}^k$ platí $\boldsymbol{x}_S = c_1\boldsymbol{u}_1 + \ldots + c_k\boldsymbol{u}_k$ práve vtedy, keď $\boldsymbol{c}$ je riešením sústavy lineárnych rovníc*

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

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

Dôkaz. Pre $\boldsymbol{c} = (c_1, \ldots, c_k)^T \in \mathbb{R}^k$ platí $\boldsymbol{x}_S = c_1\boldsymbol{u}_1 + \ldots + c_k\boldsymbol{u}_k$ práve vtedy, keď pre každé $i \leq k$ máme

$$0 = \langle \boldsymbol{x} - c_1\boldsymbol{u}_1 + \ldots + c_k\boldsymbol{u}_k,\, \boldsymbol{u}_i\rangle = \langle \boldsymbol{x}, \boldsymbol{u}_i\rangle - c_1\langle \boldsymbol{u}_1, \boldsymbol{u}_i\rangle - \ldots - c_k\langle \boldsymbol{u}_k, \boldsymbol{u}_i\rangle.$$

Inými slovami, $\boldsymbol{c}$ musí vyhovovať sústave

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

Ale $\boldsymbol{G}(\boldsymbol{\alpha})^T = \boldsymbol{G}(\boldsymbol{\alpha})$ vzhľadom na symetriu Gramovej matice. Vďaka regularite matice $\boldsymbol{G}(\boldsymbol{\alpha})$ ($\boldsymbol{\alpha}$ je báza S) má táto sústava jediné riešenie.

Ešte si všimnime, že rozšírená matica $\left(\boldsymbol{G}(\boldsymbol{\alpha}) \mid \langle \boldsymbol{x}, \boldsymbol{\alpha}\rangle^T\right)$ uvedenej sústavy je vlastne Gramovou maticou $\boldsymbol{G}(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k, \boldsymbol{x})$ rádu $k+1$, z ktorej sme vynechali posledný riadok. Ak $\boldsymbol{\alpha}$ je ortonormálna báza, tak $\boldsymbol{G}(\boldsymbol{\alpha}) = \boldsymbol{I}_k$, t. j. príslušná sústava je už vo vyriešenom tvare $\boldsymbol{c} = \langle \boldsymbol{x}, \boldsymbol{\alpha}\rangle^T$, presne v zhode s podmienkou (a) vety 14.1.3.

Poznámka. Tvrdenie zostáva bezo zmeny v platnosti, aj keď $\boldsymbol{\alpha}$ je ľubovoľný konečný (teda nie nutne lineárne nezávislý) systém generátorov v S. Jediné, čo sa zmení, je nejednoznačnosť *vyjadrenia* kolmého priemetu $\boldsymbol{x}_S = c_1\boldsymbol{u}_1 + \ldots + c_k\boldsymbol{u}_k$ ako lineárnej kombinácie vektorov systému $\boldsymbol{\alpha}$. Samotný kolmý priemet $\boldsymbol{x}_S$ je, samozrejme, určený jednoznačne. Rozmyslite si, prečo je to tak.

14.1.6. Príklad. V $\mathbb{R}^4$ so štandardným skalárnym súčinom je daný vektor $\boldsymbol{x} = (1,1,1,1)^T$ a rovina $S = [\boldsymbol{u}, \boldsymbol{v}]$, kde $\boldsymbol{u} = (0,-1,0,1)^T$, $\boldsymbol{v} = (1,-2,1,-3)^T$. Nájdeme kolmý priemet vektora $\boldsymbol{x}$ do roviny S a vypočítame vzdialenosť $\operatorname{dist}(\boldsymbol{x}, S)$ a odchýlku $\sphericalangle(\boldsymbol{x}, S)$. Kolmý priemet budeme hľadať v tvare $\boldsymbol{x}_S = c\boldsymbol{u} + d\boldsymbol{v}$, kde $(c,d)^T \in \mathbb{R}^2$ vyhovuje sústave s rozšírenou maticou

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

Jej riešením dostaneme $c = -3/29$, $d = -6/29$, teda kolmý priemet vektora $\boldsymbol{x}$ do roviny $[\boldsymbol{u}, \boldsymbol{v}]$ je

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

Ako skúšku správnosti si overte rovnosti $\langle \boldsymbol{x}-\boldsymbol{x}_S, \boldsymbol{u}\rangle = \langle \boldsymbol{x}-\boldsymbol{x}_S, \boldsymbol{v}\rangle = 0$. Pre vzdialenosť resp. odchýlku $\boldsymbol{x}$ od S potom dostávame

$$\operatorname{dist}(\boldsymbol{x}, S) = \|\boldsymbol{x} - \boldsymbol{x}_S\| = \left\| \frac{7}{29}(5,2,5,2)^T \right\| = \frac{7}{29}\sqrt{58},$$
$$\sin \sphericalangle(\boldsymbol{x}, S) = \frac{\|\boldsymbol{x} - \boldsymbol{x}_S\|}{\|\boldsymbol{x}\|} = \frac{7}{2\cdot 29}\sqrt{58} = \frac{7}{\sqrt{58}}.$$

S požitím kalkulačky možno zistiť

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

14.1.7. Príklad. Nech $\boldsymbol{A} \in R^{m\times n}$, pričom $m \geq n$ a $h(\boldsymbol{A}) = n$, t. j. stĺpce matice $\boldsymbol{A}$ sú lineárne nezávislé vektory v euklidovskom priestore $\mathbb{R}^m$ so štandardným skalárnym súčinom. Označme $S \subseteq \mathbb{R}^m$ lineárny podpriestor generovaný stĺpcami matice $\boldsymbol{A}$. Potom ortogonálna projekcia na podpriestor S je lineárny operátor $\operatorname{pr}_S\colon \mathbb{R}^m \to \mathbb{R}^m$. Nájdeme jeho maticu $\boldsymbol{B} = \left(\operatorname{pr}_S\right)_{\boldsymbol{\varepsilon},\boldsymbol{\varepsilon}} \in \mathbb{R}^{m\times m}$ vzhľadom na kanonickú ortonormálnu bázu $\boldsymbol{\varepsilon}$ priestoru $\mathbb{R}^m$.

Ak stotožníme maticu $\boldsymbol{A}$ s usporiadanou n-ticou jej stĺpcov, tak $\boldsymbol{A}$ je bázou S. Podľa tvrdenia 14.1.5 obraz $\boldsymbol{y} = \operatorname{pr}_S(\boldsymbol{x})$ vektora $\boldsymbol{x} \in \mathbb{R}^m$ dostaneme v tvare

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

kde $\boldsymbol{c} \in \mathbb{R}^n$ je (jediné) riešenie sústavy

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

Uvedomme si, že $\boldsymbol{G}(\boldsymbol{A}) = \boldsymbol{A}^T \cdot \boldsymbol{A}$ je regulárna matica. Ďalej platí $\langle \boldsymbol{x}, \boldsymbol{A}\rangle = \boldsymbol{x}^T \cdot \boldsymbol{A}$, čiže $\langle \boldsymbol{x}, \boldsymbol{A}\rangle^T = \boldsymbol{A}^T \cdot \boldsymbol{x}$. Z toho dostávame

$$\boldsymbol{c} = \boldsymbol{G}(\boldsymbol{A})^{-1} \cdot \langle \boldsymbol{x}, \boldsymbol{A}\rangle^T = \left(\boldsymbol{A}^T \cdot \boldsymbol{A}\right)^{-1} \cdot \boldsymbol{A}^T \cdot \boldsymbol{x},$$
$$\boldsymbol{y} = \boldsymbol{A} \cdot \boldsymbol{c} = \boldsymbol{A} \cdot \left(\boldsymbol{A}^T \cdot \boldsymbol{A}\right)^{-1} \cdot \boldsymbol{A}^T \cdot \boldsymbol{x}.$$

Teda hľadaná matica ortogonálnej projekcie pr_S je

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

14.2. Vzdialenosť dvoch afinných podpriestorov

Nech V je vektorový priestor so skalárnym súčinom a X, Y sú jeho dve neprázdne podmnožiny. *Vzdialenosťou množín* X, Y vo V nazývame číslo

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

Problematikou vzdialeností množín v plnom rozsahu sa tu zaoberať nebudeme. Obmedzíme sa len na vzdialenosti konečnorozmerných afinných podpriestorov. Úlohu prevedieme na určenie vzdialenosti vektora od lineárneho podpriestoru.

14.2.1. Lema. *Nech V je vektorový priestor so skalárnym súčinom a M, N sú jeho afinné podpriestory. Potom pre ľubovoľné body $\boldsymbol{p} \in M$, $\boldsymbol{q} \in M$ platí:*

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

Dôkaz. Označme $S = \operatorname{Dir} M$, $T = \operatorname{Dir} N$ zamerania podpriestrov M, N. Potom $M = \boldsymbol{p} + S$, $N = \boldsymbol{q} + T$. Podľa definície vzdialenosti platí

$$\operatorname{dist}(M, N) = \inf\{\|(\boldsymbol{p} + \boldsymbol{u}) - (\boldsymbol{q} + \boldsymbol{v})\|\,;\ \boldsymbol{u} \in S \ \&\ \boldsymbol{v} \in T\},$$
$$\operatorname{dist}(\boldsymbol{p} - \boldsymbol{q},\ S + T) = \inf\{\|(\boldsymbol{p} - \boldsymbol{q}) - (\boldsymbol{u} + \boldsymbol{v})\|\,;\ \boldsymbol{u} \in S \ \&\ \boldsymbol{v} \in T\}.$$

Stačí teda overiť rovnosť množín na pravých stranách. Túto jednoduchú úlohu prenechávame ako cvičenie čitateľovi.

Hovoríme, že body $\boldsymbol{p} \in M$, $\boldsymbol{q} \in N$ tvoria *priečku* afinných podpriestorov M, N, ak

$$\operatorname{dist}(M, N) = \|\boldsymbol{p} - \boldsymbol{q}\|,$$

t. j. ak sa vzdialenosť podpriestorov M, N realizuje ako dĺžka vektora $\boldsymbol{p} - \boldsymbol{q}$.

14.2.2. Tvrdenie. *Nech M, N sú konečnorozmerné afinné podpriestory vektorového priestoru so skalárnym súčinom V. Potom*

(a) *body $\boldsymbol{p} \in M$, $\boldsymbol{q} \in N$ tvoria priečku podpriestorov M, N práve vtedy, keď $\boldsymbol{p} - \boldsymbol{q} \in (\operatorname{Dir} M + \operatorname{Dir} N)^{\perp}$;*

(b) *pre ľubovoľné body $\boldsymbol{p} \in M$, $\boldsymbol{q} \in N$ a vektory $\boldsymbol{u} \in \operatorname{Dir} M$, $\boldsymbol{v} \in \operatorname{Dir} N$ platí: body $\boldsymbol{p} + \boldsymbol{u}$, $\boldsymbol{q} + \boldsymbol{v}$ tvoria priečku podpriestorov M, N práve vtedy, keď vektor $\boldsymbol{v} - \boldsymbol{u}$ je kolmým priemetom vektora $\boldsymbol{p} - \boldsymbol{q}$ do lineárneho podpriestoru $\operatorname{Dir} M + \operatorname{Dir} N$;*

(c) *existujú body $\boldsymbol{p} \in M$, $\boldsymbol{q} \in N$ tvoriace priečku podpriestorov M, N.*

Dôkaz. (a) je bezprostredným dôsledkom vety 14.1.3 (b) a lemy 14.2.1; (b) priamo vyplýva z (a). Konečne (c) dostaneme z (b) a vety 14.1.3 (a).

14.2.3. Dôsledok. *Pre konečnorozmerné afinné podpriestory M, $N \subseteq V$ vektorového priestoru so skalárnym súčinom platí $\operatorname{dist}(M, N) = 0$ práve vtedy, keď $M \cap N \neq \emptyset$.*

Podmienky 14.1.3 (a) a 14.2.2 (b) nám spolu s tvrdením 14.1.5 poskytujú priamy návod ako nájsť priečku a vzdialenosť ľubovoľných konečnorozmerných afinných podpriestorov. Ak $M = \boldsymbol{p} + [\boldsymbol{u}_1, \ldots, \boldsymbol{u}_m]$, $N = \boldsymbol{q} + [\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n]$ sú zadané parametricky, stačí nájsť jedno riešenie $\boldsymbol{c} = (c_1, \ldots, c_m, c_{m+1}, \ldots, c_{m+n})^T \in \mathbb{R}^{m+n}$ sústavy

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

kde $\boldsymbol{\gamma} = (\boldsymbol{u}_1, \ldots, \boldsymbol{u}_m, \boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$, a položiť

$$\boldsymbol{u} = c_1 \boldsymbol{u}_1 + \ldots c_m \boldsymbol{u}_m, \qquad \boldsymbol{v} = c_{m+1} \boldsymbol{v}_1 + \ldots + c_{m+n} \boldsymbol{v}_n.$$

Potom vektor $\boldsymbol{w} = \boldsymbol{u} + \boldsymbol{v} = \boldsymbol{\gamma} \cdot \boldsymbol{c}$ je kolmým priemetom vektora $\boldsymbol{p} - \boldsymbol{q}$ do lineárneho podpriestoru $\operatorname{Dir} M + \operatorname{Dir} N = [\boldsymbol{u}_1, \ldots \boldsymbol{u}_m, \boldsymbol{v}_1, \ldots, \boldsymbol{v}_n]$ a priečka podpriestorov M, N je tvorená bodmi $\boldsymbol{p} - \boldsymbol{u}$, $\boldsymbol{q} + \boldsymbol{v}$. V dôsledku toho

$$\operatorname{dist}(M, N) = \|(\boldsymbol{p} - \boldsymbol{u}) - (\boldsymbol{q} + \boldsymbol{v})\| = \|\boldsymbol{p} - \boldsymbol{q} - \boldsymbol{w}\|.$$

Rozmyslite si, prečo – napriek formulácii tvrdenia 14.1.5 – si nemusíme robiť starosti s lineárnou nezávislosťou vektorov $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_m, \boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$.

14.2.4. Príklad. V euklidovskom priestore $\mathbb{R}^4$ so štandardným skalárnym súčinom nájdeme vzdialenosť rovín

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

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

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

So zreteľom na poslednú otázku si riešenie sústavy napíšeme vo všeobecnom tvare $\boldsymbol{c}_t = (13/3 + t, -3 - t, -2/3 - t, t)^T$ s parametrom $t \in \mathbb{R}$. Položme

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

Potom pre každé $t \in \mathbb{R}$ dvojica bodov

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

tvorí priečku podpriestorov M, N. Vektory

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

však od parametra t nezávisia (a nie je to náhoda), rovnako ako vzdialenosť

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

Pokiaľ by nás teda zaujímala len vzdialenosť podpriestorov M, N, prípadne by sme chceli nájsť len akúkoľvek ich jednu priečku, skutočne by stačilo použiť iba jedno riešenie $\boldsymbol{c}$ uvažovanej sústavy. Ešte si rozmyslite, ako zo získaných výsledkov vyplýva čiastočná rovnobežnosť podpriestorov M, N (pozri paragraf 8.4).

14.3. Odchýlka dvoch afinných podpriestorov

Odchýlku alebo *uhol* dvoch *netriviálnych konečnorozmerných afinných podpriestorov* ve vektorovon priestrore so skalárnym súčinom V značíme $\sphericalangle(M, N)$ a definujeme ju ako odchýlku $\sphericalangle(\operatorname{Dir} M, \operatorname{Dir} N)$ ich zameraní. Stačí teda povedať, čo rozumieme pod *odchýlkou* alebo *uhlom* $\sphericalangle(S, T)$ dvoch *netriviálnych konečnorozmerných lineárnych podpriestorov* S, $T \subseteq V$. Pre $S \subseteq T$ alebo $T \subseteq S$ položíme

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

Ak $S \cap T = \{\boldsymbol{0}\}$, kladieme

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

Ak by sme takýmto spôsobom definovali odchýlku $\sphericalangle(S, T)$, aj keď $S \cap T \neq \{\boldsymbol{0}\}$, ľubovoľný spoločný nenulový vektor $\boldsymbol{x} \in S \cap T$ by sa postaral o to, aby platilo $\sphericalangle(S, T) = \sphericalangle(\boldsymbol{x}, \boldsymbol{x}) = 0$, čo nevyzerá príliš rozumne. Preto pokiaľ $S \cap T \neq \{\boldsymbol{0}\}$, $S \nsubseteq T$ ani $T \nsubseteq S$, položíme

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

Zrejme S_1, $T_1 \subseteq V$ sú netriválne lineárne podpriestory a $S_1 \cap T_1 = \{\boldsymbol{0}\}$ (za predpokladu $S \cap T = \{\boldsymbol{0}\}$ dokonca platí $S_1 = S$, $T_1 = T$). Preto môžeme konečne definovať

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

Takto definovaný uhol podpriestorov S, T je číslo z intervalu $\langle 0, \pi/2 \rangle$ a platí preň $\sphericalangle(S, T) = \sphericalangle(T, S)$, teda je to *neorientovaný uhol.*

Je užitočné si uvedomiť, že na výpočet odchýlky dvoch podpriestorov stačí minimalizovať odchýlku vhodných vektorov z jedného podpriestoru od druhého z nich.

Tvrdenie 14.3.1. *Nech V je vektorový priestor so skalárnym súčinom a S, T sú jeho konečnorozmerné lineárne podpriestory, pričom $S \nsubseteq T$ ani $T \nsubseteq S$. Potom*

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

Dôkaz. Pre ľubovoľné $\boldsymbol{x} \neq \mathbf{0}$ je $\boldsymbol{x}_T \in T$ a $\boldsymbol{x} - \boldsymbol{x}_T \in T^\perp$. Pre $\boldsymbol{x} \in S \cap (S\cap T)^\perp$ však platí

$$\boldsymbol{x}_T = \boldsymbol{x} - (\boldsymbol{x} - \boldsymbol{x}_T) \in (S\cap T)^\perp + T^\perp = (S\cap T)^\perp,$$

lebo z inklúzie $S\cap T \subseteq T$ podľa podmienky 14.1.2 (c) vyplýva $T^\perp \subseteq (S\cap T)^\perp$. Teda $\boldsymbol{x}_T \in T \cap (S\cap T)^\perp$, preto $\sphericalangle(S,T)$ je menšia alebo rovná, ako výraz na pravej strane. Keďže arccos je klesajúca funkcia, opačná nerovnosť je dôsledkom rovnosti $\sphericalangle(\boldsymbol{x},T) = \arccos\left(\|\boldsymbol{x}_T\|\,\|\boldsymbol{x}\|^{-1}\right)$ a vety 14.1.3 (c).

Výpočet odchýlky dvoch všeobecných konečnorozmerných podpriestorov si teda vyžaduje minimalizovať hodnotu istého výrazu. To naznačuje možnosť jej výpočtu s použitím diferenciálneho počtu ako extrému funkcie viac premenných. Nebudeme sa však púšťať touto cestou, lebo neskôr hodláme predviesť elegantnejší spôsob vyjadrenia odchýlky. Prostriedkami, ktoré máme zatiaľ k dispozícii, dokážeme zrátať len odchýlku ľubovoľného netriviálneho konečnorozmerného podpriestoru od priamky alebo nadroviny, prípadne odchýlky podpriestorov, ktoré možno na tieto prípady previesť.

Ako vyplýva z vety 14.1.3 (c), odchýlka priamky $[\boldsymbol{x}]$, kde $\boldsymbol{x} \neq \mathbf{0}$, a konečnorozmerného lineárneho podpriestoru $S \neq \{\mathbf{0}\}$ je daná vzťahom

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

Túto úlohu už vieme riešiť (pozri tvrdenie 14.1.5 a príklad 14.1.6).

14.3.2. Príklad. Vypočítame odchýlku rovín M, $N \subseteq \mathbb{R}^4$ z príkladu 14.2.4. Podľa definície $\sphericalangle(M,N) = \sphericalangle(S,T)$, kde $S = [\boldsymbol{e}_1+\boldsymbol{e}_2, \boldsymbol{e}_1+\boldsymbol{e}_2+\boldsymbol{e}_3]$, $T = [\boldsymbol{e}_2+\boldsymbol{e}_4, \boldsymbol{e}_2+\boldsymbol{e}_3+\boldsymbol{e}_4]$. Ľahko nahliadneme, že $S\cap T = [\boldsymbol{e}_3]$, teda $(S\cap T)^\perp = [\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_4]$. V dôsledku toho

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

Keďže $\langle \boldsymbol{e}_1+\boldsymbol{e}_2, \boldsymbol{e}_2+\boldsymbol{e}_4\rangle = 1 \geq 0$,

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

Podľa dôsledku 14.1.4 (a) má každý $(n-1)$-rozmerný lineárny podpriestor S v n-rozmernom euklidovskom priestore V tvar $S = [\boldsymbol{a}]^\perp$ pre vhodný nenulový vektor $\boldsymbol{a} \in V$; potom každá nadrovina $N \subseteq V$ so zameraním S má tvar $N = \boldsymbol{p} + [\boldsymbol{a}]^\perp$ pre nejaké $\boldsymbol{p} \in N$. Vektor $\boldsymbol{a}$ sa nazýva *normála* alebo *normálový vektor* nadroviny N. Normála nadroviny je zrejme určená jednoznačne až na skalárny násobok. Výpočet odchýlky nadroviny a netriviálneho vlastného afinného podpriestoru možno previesť na výpočet odchýlky normály nadroviny a tohto podpriestoru. Tento prevod sa zakladá na nasledujúcom tvrdení.

14.3.3. Tvrdenie. *Nech S je netriviálny, vlastný lineárny podpriestor euklidovského priestoru V a $\mathbf{0} \neq \boldsymbol{a} \in V$. Potom*

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

Dôkaz. Keďže druhá rovnosť je zrejmá, stačí dokázať prvú. Podmienky $\boldsymbol{a}_S = \mathbf{0}$, $\boldsymbol{a} \in S^\perp$ a $S \subseteq [\boldsymbol{a}]^\perp$ sú ekvivalentné. Podobne, $[\boldsymbol{a}]^\perp \subseteq S$ je ekvivalentné s podmienkou $S^\perp \subseteq [\boldsymbol{a}]$. V oboch prípadoch platí $\sphericalangle\left([\boldsymbol{a}]^\perp, S\right) = 0 = \sphericalangle\left(\boldsymbol{a}, S^\perp\right)$, $\sphericalangle(\boldsymbol{a},S) = \pi/2$.

Nech teda $\boldsymbol{a}_S \neq \mathbf{0}$ a $[\boldsymbol{a}]$, $S^\perp$ ani $[\boldsymbol{a}]^\perp$, S nie sú vo vzťahu inklúzie. Označme $\boldsymbol{a}_{ST}$ kolmý priemet vektora $\boldsymbol{a}_S$ do podpriestoru $T = [\boldsymbol{a}]^\perp$. Postupne dokážeme rovnosti

$$\sphericalangle\left([\boldsymbol{a}]^\perp, S\right) = \arccos\frac{\|\boldsymbol{a}_{ST}\|}{\|\boldsymbol{a}_S\|} = \frac{\pi}{2} - \sphericalangle(\boldsymbol{a},S).$$

Prvú z nich dokážme v dvoch krokoch. Najprv si uvedomme, že z dôsledku 14.1.4 (a) vyplýva $\left([\boldsymbol{a}]^\perp \cap S\right)^\perp = [\boldsymbol{a}] + S^\perp$, teda

$$\boldsymbol{a}_S = \boldsymbol{a} - (\boldsymbol{a} - \boldsymbol{a}_S) \in [\boldsymbol{a}] + S^\perp = \left([\boldsymbol{a}]^\perp \cap S\right)^\perp.$$

Takže $\sphericalangle\left([\boldsymbol{a}]^\perp, S\right) \leq \arccos\left(\|\boldsymbol{a}_{ST}\|\,\|\boldsymbol{a}_S\|^{-1}\right)$ na základe tvrdenia 14.3.1. Na dôkaz opačnej nerovnosti stačí podľa toho istého tvrdenia overiť

$$\frac{\|\boldsymbol{a}_{ST}\|}{\|\boldsymbol{a}_S\|} \geq \frac{\|\boldsymbol{x}_S\|}{\|\boldsymbol{x}\|}$$

pre všetky $\mathbf{0} \neq \boldsymbol{x} \in [\boldsymbol{a}]^\perp \cap \left([\boldsymbol{a}]^\perp \cap S\right)^\perp$. Zrejme $\langle \boldsymbol{x}, \boldsymbol{a}_S\rangle = \langle \boldsymbol{x}_S, \boldsymbol{a}_S\rangle$, lebo $\boldsymbol{x} - \boldsymbol{x}_S \in S^\perp$. Keďže $\boldsymbol{x} \in [\boldsymbol{a}] + S^\perp$, vektory $\boldsymbol{x}_S$, $\boldsymbol{a}_S$ sú lineárne závislé (rozmyslite si prečo); preto $|\langle \boldsymbol{x}_S, \boldsymbol{a}_S\rangle| = \|\boldsymbol{x}_S\|\,\|\boldsymbol{a}_S\|$. Nakoľko $\boldsymbol{x} \in [\boldsymbol{a}]^\perp = T$, s použitím vety 14.1.3 (c) z toho vyplýva

$$\frac{\|\boldsymbol{x}_S\|}{\|\boldsymbol{x}\|} = \frac{|\langle \boldsymbol{x}_S, \boldsymbol{a}_S\rangle|}{\|\boldsymbol{x}\|\,\|\boldsymbol{a}_S\|} = \frac{|\langle \boldsymbol{x}, \boldsymbol{a}_S\rangle|}{\|\boldsymbol{x}\|\,\|\boldsymbol{a}_S\|} \leq \frac{\|\boldsymbol{a}_{ST}\|}{\|\boldsymbol{a}_S\|}.$$

Druhú rovnosť overíme priamym výpočtom

$$\begin{aligned}\frac{\pi}{2} - \sphericalangle(\boldsymbol{a},S) &= \frac{\pi}{2} - \sphericalangle(\boldsymbol{a}_S,\boldsymbol{a}) = \frac{\pi}{2} - \sphericalangle\left(\boldsymbol{a}_S,[\boldsymbol{a}]\right)\\ &= \sphericalangle\left(\boldsymbol{a}_S,[\boldsymbol{a}]^\perp\right) = \arccos\frac{\|\boldsymbol{a}_{ST}\|}{\|\boldsymbol{a}_S\|}.\end{aligned}$$

Ako zvláštny prípad pre odchýlku dvoch nadrovín dostávame

14.3.4. Dôsledok. *Nech M, N sú dve nadroviny v euklidovskom priestore V s normálami $\boldsymbol{a}$, resp. $\boldsymbol{b}$. Potom*

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

V euklidovskom priestore $\mathbb{R}^n$ so štandardným skalárnym súčinom vystupuje normálový vektor danej nadroviny priamo v jej (všeobecnej) rovnici. Ak je totiž nadrovina M daná rovnicou

$$a_1x_1 + \ldots + a_nx_n = b,$$

tak $\boldsymbol{a} = (a_1,\ldots,a_n)^T \neq \mathbf{0}$ je jej normála a uvedenú rovnicu možno skrátene zapísať v tvare $\langle \boldsymbol{x}, \boldsymbol{a}\rangle = b$.

14.3.5. Príklad. V euklidovskom priestore V vypočítame odchýlku roviny $S = [\boldsymbol{u}, \boldsymbol{v}]$ a nadroviny $T = [\boldsymbol{a}]^\perp$. Podľa tvrdenia 14.3.5 platí

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

Súradnice c, d kolmého priemetu $\boldsymbol{a}_S = c\boldsymbol{u} + d\boldsymbol{v}$ vzhľadom na bázu $(\boldsymbol{u}, \boldsymbol{v})$ podpriestoru S získame riešením sústavy

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

podľa Cramerovho pravidla v tvare

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

Spätné dosadenie do vzorca pre odchýlku $\sphericalangle(S,T)$ si odpustíme. Tieto vzorce možno zrejmým spôsobom zovšeobecniť aj pre odchýlku k-rozmerného podpriestoru S a priamky, resp. nadroviny. vo V. Pre $k \geq 3$ však výpočet pomocou Gramových determinantov už nie je výhodný.

14.3.6. Príklad. Podaktorý čitateľ si možno kladie otázku, prečo sme miesto tvrdenia 14.3.3 nedokázali silnejší výsledok, totiž rovnosť

$$\sphericalangle(S,T) + \sphericalangle\left(S^\perp, T\right) = \frac{\pi}{2},$$

pre ***každý netriválny vlastný*** (teda nie len jednorozmerný) lineárny podpriestor S a každý netriviálny lineárny podpriestor T euklidovského priestoru V. Uvedené tvrdenie, ako aj náš geometrický názor nám totiž napovedajú, že také niečo by malo platiť. O to bizarnejší sa nám preto bude zdať nasledujúci veľmi jednoduchý príklad, ktorý v ľubovoľnej dimenzii $n > 3$ našu domnienku vyvracia.

V euklidovskom priestore $\mathbb{R}^n$ so štandardným skalárnym súčinom sú dané lineárne podpriestory $S = [\boldsymbol{e}_1, \boldsymbol{e}_2]$, $T = [\boldsymbol{e}_1, \boldsymbol{e}_3]$. Zrejme $S \cap T = [\boldsymbol{e}_1]$, teda $S \cap [\boldsymbol{e}_1]^\perp = [\boldsymbol{e}_2]$ a $T \cap [\boldsymbol{e}_1]^\perp = [\boldsymbol{e}_3]$. V dôsledku toho

$$\sphericalangle(S,T) = \sphericalangle\left([\boldsymbol{e}_2], [\boldsymbol{e}_3]\right) = \frac{\pi}{2}.$$

Na druhej strane $S^\perp = [\boldsymbol{e}_1, \boldsymbol{e}_2]^\perp = [\boldsymbol{e}_3, \ldots, \boldsymbol{e}_n]$. Takže $S^\perp \cap T = [\boldsymbol{e}_3]$, z čoho dostávame $S^\perp \cap [\boldsymbol{e}_3]^\perp = [\boldsymbol{e}_4, \ldots, \boldsymbol{e}_n]$ a $T \cap [\boldsymbol{e}_3]^\perp = [\boldsymbol{e}_1]$. Konečne

$$\sphericalangle\left(S^\perp, T\right) = \sphericalangle\left([\boldsymbol{e}_4, \ldots, \boldsymbol{e}_n], [\boldsymbol{e}_1]\right) = \frac{\pi}{2}.$$

Teda hodnota súčtu $\sphericalangle(S,T) + \sphericalangle\left(S^\perp, T\right)$ je tentokrát π, a nie $\pi/2$, ako sme očakávali. Voľne povedané, podpriestor T je kolmý tak na podpriestor S, ako aj na jeho ortokomplement $S^\perp$. Samozrejme, platí tiež $\sphericalangle\left(S, S^\perp\right) = \pi/2$.

14.4. Riešenie neriešiteľných sústav a lineárna regresia

V celom tomto paragrafe označujú m, n pevné kladné celé čísla. Stĺpcové vektorové priestory $\mathbb{R}^m$ a $\mathbb{R}^n$ sú vybavené štandardným skalárnym súčinom, takže to sú euklidovské priestory.

Nech $\boldsymbol{A} \in \mathbb{R}^{m\times n}$, $\boldsymbol{b} \in \mathbb{R}^m$. Uvažujme sústavu lineárnych rovníc

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

a označme $S = [\boldsymbol{s}_1(\boldsymbol{A}), \ldots, \boldsymbol{s}_n(\boldsymbol{A})]$ lineárny podpriestor v $\mathbb{R}^m$ generovaný stĺpcami matice $\boldsymbol{A}$. Podľa Frobeniovej vety naša sústava má nejaké riešenie $\boldsymbol{x} \in \mathbb{R}^n$ práve vtedy, keď $\boldsymbol{b} \in S$. Zložky riešenia $\boldsymbol{x} = (x_1, \ldots, x_n)^T \in \mathbb{R}^n$ sú potom koeficienty lineárnej kombinácie

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

No i v prípade, keď $\boldsymbol{b} \notin S$, t. j. riešenie sústavy neexistuje, sa môžeme pokúsiť nahradiť jej pravú stranu $\boldsymbol{b}$ čo najbližším vektorom podpriestoru S. Takto získaná nová sústava už má riešenie, ktoré môžeme právom považovať za najlepšie možné približné riešenie pôvodnej sústavy. Podľa vety 14.1.3 (b) je najbližší vektor podpriestoru S k vektoru $\boldsymbol{b}$ určený jednoznačne, a je to jeho kolmý priemet $\boldsymbol{b}_S$ do tohto podpriestoru. *Pseudoriešenie* neriešiteľnej (hovorí sa tiež *preurčenej*) sústavy $\boldsymbol{A} \cdot \boldsymbol{x} = \boldsymbol{b}$ teda definujeme ako riešenie (tento raz už istotne riešiteľnej) sústavy

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

Ak je pôvodná sústava riešiteľná, t. j. ak $\boldsymbol{b} \in S$, tak $\boldsymbol{b}_S = \boldsymbol{b}$ a obe sústavy splývajú, takže každé jej pseudoriešenie je priamo riešením pôvodnej sústavy.

14.4.1. Tvrdenie. *Nech $\boldsymbol{A} \in \mathbb{R}^{m\times n}$, $\boldsymbol{b} \in \mathbb{R}^m$. Potom $\boldsymbol{x} \in \mathbb{R}^n$ je pseudoriešením sústavy $\boldsymbol{A} \cdot \boldsymbol{x} = \boldsymbol{b}$ práve vtedy, keď $\boldsymbol{x}$ je riešením sústavy*

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

so štvorcovou maticou $\boldsymbol{A}^T \cdot \boldsymbol{A} \in \mathbb{R}^{n\times n}$ a ľavou stranou $\boldsymbol{A}^T \cdot \boldsymbol{b} \in \mathbb{R}^n$.

Dôkaz. Najprv si uvedomme, že $\boldsymbol{A}^T \cdot \boldsymbol{A} = \boldsymbol{G}(\boldsymbol{s}_1(\boldsymbol{A}), \ldots, \boldsymbol{s}_n(\boldsymbol{A}))$ je Gramovou maticou prislúchajúcou stĺpcom matice $\boldsymbol{A}$. Podľa tvrdenia 14.1.5 a za ním nasledujúcej poznámky je lineárna kombinácia $x_1\boldsymbol{s}_1(\boldsymbol{A}) + \ldots + x_n\boldsymbol{s}_n(\boldsymbol{A}) = \boldsymbol{A} \cdot \boldsymbol{x}$ kolmým priemetom vektora $\boldsymbol{b}$ do podpriestoru $S = [\boldsymbol{s}_1(\boldsymbol{A}), \ldots, \boldsymbol{s}_n(\boldsymbol{A})]$, t. j. platí $\boldsymbol{A} \cdot \boldsymbol{x} = \boldsymbol{b}_S$, práve vtedy, keď $\boldsymbol{A}^T \cdot \boldsymbol{A} \cdot \boldsymbol{x} = \boldsymbol{A}^T \cdot \boldsymbol{b}$.

Pseudoriešenie preurčenej sústavy $\boldsymbol{A} \cdot \boldsymbol{x} = \boldsymbol{b}$ teda hľadáme ako riešenie zaručene riešiteľnej sústavy $\boldsymbol{A}^T \cdot \boldsymbol{A} \cdot \boldsymbol{x} = \boldsymbol{A}^T \cdot \boldsymbol{b}$. V typickom prípade má pôvodná sústava viac rovníc než neznámych, čiže $m > n$ a $\boldsymbol{A}$ je obdĺžniková matica, „vyššia ako širšia“. Potom je veľmi pravdepodobné, že štvorcová matica $\boldsymbol{A}^T \cdot \boldsymbol{A}$ rádu n (ako Gramova matica „malého“ počtu stĺpcových vektorov v euklidovskom priestore „veľkej“ dimenzie) je regulárna, teda k nej existuje inverzná matica $\left(\boldsymbol{A}^T \cdot \boldsymbol{A}\right)^{-1}$. V takom prípade je pseudoriešenie pôvodnej sústavy určené jednoznačne:

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

Samozrejme, ak $m = n$ a už samotná matica $\boldsymbol{A}$ je regulárna, dostávame

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

a $\boldsymbol{x} = \boldsymbol{A}^{-1} \cdot \boldsymbol{b}$ je priamo jediným riešením pôvodnej sústavy.

V úlohách *lineárnej regresie* máme zadané hodnoty $y_1, \dots, y_m$ neznámej funkcie f v bodoch $x_1, \dots, x_m$ jej definičného oboru, získané väčšinou meraním. Funkciu f chceme aproximovať lineárnou kombináciou funkcií $f_1, \dots, f_n$, ktoré poznáme, či aspoň sú nám známe ich hodnoty $a_{ij} = f_j(x_i)$ v bodoch $x_1, \dots, x_m$. Zvyčajne je m podstatne väčšie ako n. V optimálnom prípade sa nám môže podariť zostrojiť funkciu $f = c_1 f_1 + \ldots + c_n f_n$ priamo ako lineárnu kombináciu funkcií f_j tak, aby f v bodoch x_i nadobúdala vopred predpísané hodnoty y_i, t.j.

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

Ak označíme $\boldsymbol{A} = (a_{ij}) \in \mathbb{R}^{m\times n}$, $\boldsymbol{y} = (y_1, \dots, y_m)^T \in \mathbb{R}^m$, $\boldsymbol{c} = (c_1, \dots, c_n)^T \in \mathbb{R}^n$, vidíme, že vlastne hľadáme riešenie $\boldsymbol{c}$ sústavy

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

Táto sústava je v typickom prípade preurčená, teda jej riešenie neexistuje. Úloha lineárnej regresie potom splýva s *metódou najmenších štvorcov* a spočíva v nájdení takých koeficientov c_j, ktoré minimalizujú výraz

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

Toto minimum sa nadobúda pre $\boldsymbol{c}$ také, že $\boldsymbol{A}\cdot\boldsymbol{c} = \boldsymbol{y}_S$, kde S je podpriestor v $\mathbb{R}^m$ generovaný stĺpcami matice $\boldsymbol{A}$. Inak povedané, hľadanú lineárnu kombináciu dostaneme pre pseudoriešenie $\boldsymbol{c}$ sústavy $\boldsymbol{A} \cdot \boldsymbol{c} = \boldsymbol{y}$. Pre hodnoty pochádzajúce z rozumne postavených praktických úloh je takmer isté, že matica $\boldsymbol{A}^T \cdot \boldsymbol{A}$ je regulárna. V takom prípade

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

čiže hľadaná lineárna kombinácia $f = c_1 f_1 + \ldots + c_n f_n = (f_1, \dots, f_n) \cdot \boldsymbol{c}$ je určená jednoznačne.

Metódu lineárnej regresie, ako aj to, čo približne rozumieme pod „rozumne postavenou úlohou", si ilustrujeme na jednom typickom a dôležitom príklade.

14.4.2. Príklad. V rovine $\mathbb{R}^2$ je daných $m \geq 2$ bodov $(x_1, y_1), \dots, (x_m, y_m)$, získaných meraním hodnôt nejakej neznámej funkcie f vo vybraných bodoch x_i jej definičného oboru. Túto funkciu hodláme aproximovať priamkou s rovnicou $y = a + bx$ tak, aby výraz $\sum_{i=1}^{m} (y_i - a - bx_i)^2$ bol minimálny. Ak si uvedomíme, že funkcia $y = a + bx$ je lineárnou kombináciou konštantnej funkcie $y = 1$ a identickej funkcie $y = x$, hneď vidíme, že ide o úlohu lineárnej regresie. Sústava

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

je okrem triviálneho prípadu, keď všetky body (x_i, y_i) ležia na jednej priamke, preurčená. Koeficienty a, b teda nájdeme ako pseudoriešenie tejto sústavy. Jej maticu si označíme $\boldsymbol{A}$. Jednoduchý výpočet dáva

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

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

Teda $\det\left(\boldsymbol{A}^T \cdot \boldsymbol{A}\right) = 0$ práve vtedy, keď $x_1 = x_2 = \ldots = x_m$. Asi sa zhodneme na tom, že v takomto prípade by šlo, mierne povedané, o veľmi nerozumne postavenú úlohu. Naopak, na základe jej formulácie je prirodzené očakávať, že všetky hodnoty x_i budú navzájom rôzne. V čo len trochu rozumnom prípade teda dostávame jednoznačne určené koeficienty a, b v tvare

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

V krátkosti preskúmame aj duálnu situáciu. Ak má sústava $\boldsymbol{A} \cdot \boldsymbol{x} = \boldsymbol{b}$ viac, t.j. nekonečne mnoho riešení (hovoríme tiež, že je *podurčená*), môže nám veľkosť riešenia (v zmysle euklidovskej normy) poslúžiť ako dodatočné kritérium výberu toho najvhodnejšieho (napr. najlacnejšieho) z nich. Hovoríme, že $\boldsymbol{x} \in \mathbb{R}^n$ je *minimálne riešenie* uvedenej sústavy, ak je jej riešením a pre každé riešenie $\boldsymbol{y}$ tejto sústavy platí $\|\boldsymbol{x}\| \leq \|\boldsymbol{y}\|$.

Samozrejme, ak má sústava jediné riešenie, tak je to zároveň aj minimálne riešenie.

14.4.3 Tvrdenie. *Nech $\boldsymbol{A} \in \mathbb{R}^{m\times n}$, $\boldsymbol{b} \in \mathbb{R}^m$. Potom $\boldsymbol{x} \in \mathbb{R}^n$ je minimálnym riešením sústavy $\boldsymbol{A}\cdot\boldsymbol{x} = \boldsymbol{b}$ práve vtedy, keď má tvar $\boldsymbol{x} = \boldsymbol{A}^T \cdot \boldsymbol{z}$, kde $\boldsymbol{z} \in \mathbb{R}^m$ je riešením sústavy*

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

so štvorcovou maticou $\boldsymbol{A} \cdot \boldsymbol{A}^T \in \mathbb{R}^{m\times m}$ a ľavou stranou $\boldsymbol{b} \in \mathbb{R}^m$.

Dôkaz. Pripomeňme si, že množina riešení $\mathcal{R}(\boldsymbol{A} \,|\, \boldsymbol{b})$ sústavy $\boldsymbol{A} \cdot \boldsymbol{x} = \boldsymbol{b}$ tvorí afinný podpriestor v $\mathbb{R}^n$ a jeho zameraním je lineárny podpriestor $\mathcal{R}(\boldsymbol{A})$ všetkých riešení homogénnej sústavy $\boldsymbol{A}\cdot\boldsymbol{x}=\boldsymbol{0}$. Podľa definície vektor $\boldsymbol{x}\in\mathbb{R}^n$ je minimálnym riešením pôvodnej sústavy práve vtedy, keď body $\boldsymbol{0}$ a $\boldsymbol{x}$ tvoria priečku afinných podpriestorov $\{\boldsymbol{0}\}$ a $\mathcal{R}(\boldsymbol{A}\,|\,\boldsymbol{b})$, čo je podľa tvrdenia 14.2.2 (a) ekvivalentné s podmienkou $\boldsymbol{x}\in\mathcal{R}(\boldsymbol{A})^\perp$. Ak si uvedomíme, že $\mathcal{R}(\boldsymbol{A})=R^\perp$ pre lineárny podpriestor $R=\left[\boldsymbol{r}_1(\boldsymbol{A})^T,\dots,\boldsymbol{r}_m(\boldsymbol{A})^T\right]$ v $\mathbb{R}^n$ generovaný stĺpcami $\boldsymbol{s}_i(\boldsymbol{A}^T)=\boldsymbol{r}_i(\boldsymbol{A})^T$ matice $\boldsymbol{A}^T$, a podľa dôsledku 14.1.4 (a) platí $\mathcal{R}(\boldsymbol{A})^\perp=R^{\perp\perp}=R$, vidíme, že $\boldsymbol{x}$ musí byť tvaru

$$\boldsymbol{x}=z_1\boldsymbol{r}_1(\boldsymbol{A})^T+\ldots+z_m\boldsymbol{r}_m(\boldsymbol{A})^T=\boldsymbol{A}^T\cdot\boldsymbol{z}$$

pre nejaké $\boldsymbol{z}\in\mathbb{R}^m$. Takýto vektor $\boldsymbol{x}$ je riešením pôvodnej sústavy práve vtedy, keď $\boldsymbol{A}\cdot\boldsymbol{A}^T\cdot\boldsymbol{z}=\boldsymbol{b}$.

Minimálne riešenie $\boldsymbol{x}$ podurčenej sústavy $\boldsymbol{A}\cdot\boldsymbol{x}=\boldsymbol{b}$ teda hľadáme tak, že najprv nájdeme nejaké (ľubovoľné) riešenie $\boldsymbol{z}$ sústavy $\boldsymbol{A}\cdot\boldsymbol{A}^T\cdot\boldsymbol{z}=\boldsymbol{b}$ a položíme $\boldsymbol{x}=\boldsymbol{A}^T\cdot\boldsymbol{z}$. V typickom prípade má pôvodná sústava menej rovníc než neznámych, čiže $m<n$ a $\boldsymbol{A}$ je obdĺžniková matica, „širšia ako vyššia“. Potom je veľmi pravdepodobné, že štvorcová matica $\boldsymbol{A}\cdot\boldsymbol{A}^T$ rádu m (ako Gramova matica „malého“ počtu riadkových vektorov v (riadkovom) euklidovskom priestore „veľkej“ dimenzie) je regulárna, teda k nej existuje inverzná matica $\left(\boldsymbol{A}\cdot\boldsymbol{A}^T\right)^{-1}$. V tom prípade je tak

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

ako aj minimálne riešenie pôvodnej sústavy

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

určené jednoznačne. Samozrejme, ak $m=n$ a už samotná $\boldsymbol{A}$ je regulárna, dostávame

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

a $\boldsymbol{x}=\boldsymbol{A}^{-1}\cdot\boldsymbol{b}$ je priamo jediným riešením pôvodnej sústavy.

Oba prístupy možno kombinovať. Ak totiž sústava $\boldsymbol{A}\cdot\boldsymbol{x}=\boldsymbol{b}$ nemá riešenie, avšak má viac pseudoriešení, jej *minimálne pseudoriešenie* dostaneme v tvare

$$\boldsymbol{x}=\boldsymbol{A}^T\cdot\boldsymbol{A}\cdot\boldsymbol{z},$$

kde $\boldsymbol{z}$ je riešenie sústavy

$$\left(\boldsymbol{A}^T\cdot\boldsymbol{A}\right)^2\cdot\boldsymbol{z}=\boldsymbol{A}^T\cdot\boldsymbol{b}.$$

Detaily si premyslite samostatne.

Na záver ešte poznamenajme, že pomocou podmienok tvrdení 14.4.1 a 14.4.3 možno definovať pojmy pseudoriešenia a minimálneho riešenia sústav lineárnych rovníc nad ľubovoľným poľom. Operácia transponovania matice totiž nie je nijako viazaná na štruktúru euklidovského priestoru. Vo všeobecnom prípade však uvedené pojmy strácajú svoj názorný geometrický význam. Neskôr uvidíme, že tento význam možno zachovať v prípade poľa komplexných čísel, ale toto zovšeobecnenie už bude trochu menej priamočiare, než sme naznačili.

14.5. Geometria pravdepodobnosti

V tomto paragrafe si predvedieme, ako možno základné pojmy teórie pravdepodobnosti formulovať v jazyku skalárneho súčinu a euklidovských priestorov. Kvôli jednoduchosti sa obmedzíme len na konečné pravdepodobnostné priestory. Bude to zároveň krátka a jednoduchá ukážka toho, čo sa zvykne nazývať „vnútornou jednotou matematiky“, prejavujúcou sa mnohorakými súvislosťami a hlbokou spätosťou jej jednotlivých disciplín. V tomto konkrétnom prípade dodáva geometrická interpretácia teórii pravdepodobnosti nový názorný rozmer.

Pravdepodobnostný priestor bude pre nás konečná množina X, vybavená funkciou $p\colon X\to\mathbb{R}$ takou, že

$$(\forall\, x\in X)(p(x)\geq 0)\qquad\text{a}\qquad\sum_{x\in X}p(x)=1,$$

nazývanou *rozdelením pravdepodobnosti* na X. Ako typický príklad si predstavte *rovnomerné pravdepodobnostné rozdelenie* $p(x)=1/n$ pre každé $x\in X$, kde $n=\#\,X$. Podmnožiny množiny X nazývame *javmi* a jej prvky (ktoré stotožňujeme s jednoprvkovými podmnožinami) *elementárnymi javmi*. *Pravdepodobnosťou javu* $A\subseteq X$ rozumieme číslo

$$P(A)=\sum_{x\in A}p(x).$$

Javy $A\subseteq X$ také, že $P(A)=0$ nazývame *nemožnými*. Keďže nemožné javy sú už zo svojej definície tie, ktoré nemôžu nastať, lebo majú nulovú pravdepodobnosť, bez ujmy na všeobecnosti môžeme prijať zjednodušujúci predpoklad, že medzi prvkami množiny X sa nevyskytujú nemožné javy, t. j. platí $p(x)>0$ pre každé $x\in X$ (v opačnom prípade môžeme nemožné elementárne javy z množiny X jednoducho vylúčiť).[1] Inak povedané, $A=\emptyset$ je jediný nemožný jav $A\subseteq X$.

Náhodnou premennou na pravdepodobnostnom priestore X rozumieme ľubovoľnú funkciu $f\colon X\to\mathbb{R}$. Množina $\mathbb{R}^X$ všetkých náhodných premenných tvorí vektorový priestor nad poľom $\mathbb{R}$ (pozri odstavec 1.6.5). Zrejme $\dim\mathbb{R}^X=\#\,X$. Vďaka nášmu predpokladu, podľa ktorého $p(x)>0$ pre každé $x\in X$, je rovnosťou

$$\langle f,g\rangle=\sum_{x\in X}f(x)g(x)p(x),$$

kde f, g sú náhodné premenné, definovaný skalárny súčin na $\mathbb{R}^X$. Priestor $\mathbb{R}^X$ všetkých náhodných premenných je tak euklidovským priestorom. Množina

$$C=\left\{f\in\mathbb{R}^X;\ (\forall\, x,y\in X)(f(x)=f(y))\right\}$$

všetkých konštantných náhodných premenných tvorí jednorozmerný lineárny podpriestor priestoru $\mathbb{R}^X$. Keďže zobrazenie, ktoré reálnemu číslu a priradí konštantnú náhodnú premennú $f\in\mathbb{R}^n$ takú, že $f(x)=a$ pre každé $x\in X$, je lineárny izomorfizmus $\mathbb{R}\cong C$, môžeme si dovoliť stotožniť toto číslo s príslušnou konštantnou náhodnou premennou.

[1] Treba poznamenať, že v nekonečnom pravdepodobnostnom priestore X by už takýto predpoklad spôsobil značnú ujmy na všeobecnosti.

Nakoľko konštantná náhodná premenná 1 tvorí zrejme ortonormálnu bázu podpriestoru C, lineárny operátor pr_C ortogonálnej projekcie na podpriestor C má tvar

$$\mathsf{E}(f) = \mathrm{pr}_C(f) = \langle f, 1\rangle = \sum_{x\in X} f(x)p(x)$$

pre $f \in \mathbb{R}^X$. Výraz $\mathsf{E}(f)$ nazývame *strednou* alebo aj *očakávanou hodnotou* náhodnej premennej f (označenie operátora $\mathsf{E}\colon \mathbb{R}^X \to \mathbb{R}^X$ pochádza z anglického slova *expectation*).

Ortokoplementom podpriestoru C je podľa dôsledku 14.1.4 (d) lineárny podpriestor

$$N = C^\perp = \operatorname{Ker}\mathsf{E} = \{f \in \mathbb{R}^X;\ \mathsf{E}(f) = 0\},$$

tvorený všetkými náhodnými premennými s nulovou strednou hodnotou. Potom náhodná premenná $f - \mathsf{E}(f)$ je kolmým priemetom $f \in \mathbb{R}^X$ do nadroviny N.

Disperziou alebo tiež *rozptylom* náhodnej premennej $f \in \mathbb{R}^X$ nazývame výraz $\mathsf{D}(f) = \|f - \mathsf{E}(f)\|^2$. Zrejme platí

$$\sqrt{\mathsf{D}(f)} = \|f - \mathsf{E}(f)\| = \operatorname{dist}(f, C).$$

To znamená, že toto číslo, nazývané *stredná kvadratická odchýlka* prípadne *smerodajná odchýlka* náhodnej premennej f, udáva akúsi strednú mieru nekonštantnosti f, presnejšie, vzdialenosť f od jej očakávanej hodnoty.

Nech $f, g \in \mathbb{R}^X$ sú náhodné premenné s nenulovými disperziami. Ich *korelačný koeficient* definujeme vzťahom

$$\mathsf{R}(f,g) = \cos\sphericalangle(f - \mathsf{E}(f),\ g - \mathsf{E}(g)) = \frac{\langle f - \mathsf{E}(f),\ g - \mathsf{E}(g)\rangle}{\|f - \mathsf{E}(f)\|\,\|g - \mathsf{E}(g)\|}.$$

To znamená, že $\arccos \mathsf{R}(f,g)$ je uhol, ktorý zvierajú kolmé priemety náhodných premenných f, g do podpriestoru N náhodných prememenných s nulovou strednou hodnotou. Špeciálne pre $f, g \in N$ je

$$\mathsf{R}(f,g) = \cos\sphericalangle(f,g) = \frac{\langle f, g\rangle}{\|f\|\,\|g\|}.$$

Niekoľko doplňujúcich informácií o „geometrii pravdepodobnosti“ môže čitateľ nájsť v cvičeniach.