MATEMATICKO-FYZIKÁLNÍ FAKULTA
Univerzita Karlova

BAKALÁŘSKÁ PRÁCE

Jakub Fedor

Kvantily pre smerové dáta

Katedra pravděpodobnosti a matematické statistiky

Vedoucí bakalářské práce: Mgr. Stanislav Nagy, Ph.D.
Studijní program: Finanční matematika

Praha 2023

Prohlašuji, že jsem tuto bakalářskou práci vypracoval(a) samostatně a výhradně s použitím citovaných pramenů, literatury a dalších odborných zdrojů. Tato práce nebyla využita k získání jiného nebo stejného titulu.

Beru na vědomí, že se na moji práci vztahují práva a povinnosti vyplývající ze zákona č. 121/2000 Sb., autorského zákona v platném znění, zejména skutečnost, že Univerzita Karlova má právo na uzavření licenční smlouvy o užití této práce jako školního díla podle §60 odst. 1 autorského zákona.

V ..... dne ....

Podpis autora

Chcel by som sa primárne poďakovať vedúcemu mojej bakalárskej práce Mgr. Stanislavovi Nagyovi, Ph.D. za cenné pripomienky, skvelú komunikáciu a čas, ktorý mi venoval. Rovnako by som sa chcel poďakovať mojim rodičom a priateľom za podporu počas celého môjho vysokoškolského štúdia.

Název práce: Kvantily pre smerové dáta

Autor: Jakub Fedor

Katedra: Katedra pravděpodobnosti a matematické statistiky

Vedoucí bakalářské práce: Mgr. Stanislav Nagy, Ph.D., Katedra pravděpodobnosti a matematické statistiky

Abstrakt: V tejto práci budeme skúmať špeciálny typ viacrozmerných dát – smerové dáta. Hlavnou časťou práce je predstavenie a porovnanie rôznych spôsobov usporadúvania dát a zavedenie zovšeobecnených definícií pre medián smerových dát. Najdôležitejším prezentovaným prístupom pre usporadúvanie smerových dát sú uhlové hĺbky. Popíšeme dôležité vlastnosti uhlových hĺbok a overíme, či nami definované uhlové hĺbky tieto vlastnosti spĺňajú. Pomocou definovaných uhlových hĺbok a mediánov zavedieme spôsoby pre vytvorenie krabicového diagramu pre smerové dáta.

Klíčová slova: kvantil, medián, směrová data, cirkulární data, uhlová hloubka, krabicový diagram

Title: Quantiles for directional data

Author: Jakub Fedor

Department: Department of Probability and Mathematical Statistics

Supervisor: Mgr. Stanislav Nagy, Ph.D., Department of Probability and Mathematical Statistics

Abstract: In this thesis we study a special type of multidimentional data – directional data. The main part of this thesis consists of defining and comparing different ways of ordering directional data. The most important functions used for ordering directional data presented in this thesis are angular depths. We will describe importnant properties of angular depths and we will discuss wheter each angular depth satisfies the formulated desirable properties. Using previously defined angular depths and median we show two ways of drawing the circular version of a boxplot.

Keywords: quantile, median, directional data, circular data, angular depth, boxplot

Obsah

Úvod **2**

Použité značenie **3**

1 Základné definície **4**
1.1 Smerové dáta 4
1.2 Intuitívne prístupy k usporiadaniu smerových dát 5
1.3 Uhlové hĺbky 6

2 Definície mediánu a kvantilov **12**

3 Vlastnosti uhlových hĺbok **16**

4 Definície empirických ekvivalentov **23**

5 Krabicový diagram pre smerové dáta **24**
5.1 Cirkulárny krabicový diagram 24
5.2 Analógia krabicového diagramu pomocou empirických projekčných kvantilov 26
5.3 Porovnanie navrhnutých prístupov 29

Záver **32**

Zoznam použitej literatúry **33**

Úvod

Pozorovania, ktorých výsledky sú smermi v dvoch alebo troch dimenziách, sa prirodzene vyskytujú v rôznych prírodných vedách. Biológovia, takzvaným trasovaním zvierat, skúmajú migráciu zvierat, výsledkom trasovania bývajú smery v dvoch dimenziách. Uskutočnenie rôznych javov (extrémy teploty, tlaku a pod.) v cirkadiánnych rytmoch vieme reprezentovať ako bod na jednotkovej kružnici, pričom kružnica ako celok bude predstavovať jeden cyklus. V meteorológii sú sledované smery vetra, ktoré vieme zobraziť na jednotkovej kružnici, podobne v geológii je skúmaná orientácia magnetického poľa Zeme, ktorú predstavuje smer v trojrozmernom priestore. V seizmológii sú pozorované epicentrá zemetrasení, ktoré reprezentujeme ich geografickými súradnicami. Tento typ dát sa často nazýva smerovými dátami.

Cieľom tejto bakalárskej práce je porovnanie zaužívaných definícií hĺbkových funkcií a iných prístupov k definíciám charakteristík predstavujúcich obdobu kvantilov pre smerové dáta a posúdenie ich vhodnosti pre využitie pri analýze smerových dát v rôznych oblastiach.

V prvej kapitole formálne definujeme smerové dáta, pozrieme sa na intuitívne prístupy usporadúvania smerových dát a definujeme niekoľko uhlových hĺbok. V druhej kapitole uvedieme rôzne definície mediánu a pomocou nich odvodených charakteristík. V tretej kapitole budeme skúmať spĺňanie istých užitočných vlastností pri nami definovaných uhlových hĺbkach. V štvrtej kapitole sa budeme venovať empirickým verziám nami definovaných funkcií. V poslednej – piatej kapitole sa pozrieme na dva návrhy istých ekvivalentov krabicového diagramu pre smerové dáta.

Použité značenie

$\mathbf{0}_n$ vektor $(0,\ldots,0)^\top \in \mathbb{R}^n$.

$\|\cdot\|$ euklidovská norma.

$\langle\cdot,\cdot\rangle$ štandardný skalárny súčin v $\mathbb{R}^k$.

$\mathbb{N}$ množina prirodzených čísel.

$\mathbb{R}^k$ k-rozmerný euklidovský priestor.

$\mathbb{S}^{k-1}$ množina $\{\boldsymbol{x} \in \mathbb{R}^k \colon \|\boldsymbol{x}\| = 1\}$.

matica rotácie v $\mathbb{R}^n$ ortogonálna matica $O \in \mathbb{R}^{n\times n}$.

1. Základné definície

1.1 Smerové dáta

Cieľom tejto podkapitoly je zavedenie formálnej definície smerových dát, ktorých výskyt v praxi sme načrtli v úvode práce.

Definícia 1.1 (Náhodná veličina)
Majme pravdepodobnostný priestor $(\Omega, \mathcal{A}, P)$*, merateľné zobrazenie* $\boldsymbol{X}\colon (\Omega, \mathcal{A}) \to (\mathcal{X}, \mathcal{B})$ *nazývame* ***náhodnou veličinou****. Ak navyše* $\mathcal{X} = \mathbb{S}^{k-1}$ *a* $\mathcal{B}$ *je borelovská* σ*-algebra na* $\mathbb{S}^{k-1}$*, budeme* $\boldsymbol{X}$ *nazývať* ***náhodnou veličinou na k-sfére****.*

Definícia 1.2 (Pravdepodobnostné rozdelenie na k-sfére)
Majme pravdepodobnostný priestor $(\Omega, \mathcal{A}, P)$ *a nech* $\boldsymbol{X}$ *je náhodná veličina na* k*-sfére, potom mieru* R *takú, že platí*

$$R(B) = P(\{\omega \in \Omega\colon \boldsymbol{X}(\omega) \in B\}); \quad B \in \mathcal{B}$$

*budeme nazývať (****pravdepodobnostným****)* ***rozdelením na k-sfére****. Rozdelenie na* k*-sfére budeme nazývať* ***diskrétnym****, ak platí* $\sum_{i=1}^{\infty} R(\{\boldsymbol{y}_i\})$ *pre nejakú postupnosť bodov* $\{\boldsymbol{y}_i\}_{i=1}^{\infty}$*;* $\boldsymbol{y}_i \in \mathbb{S}^{k-1}$*,* ***spojitým****, ak platí* $R(T) = 0$ *pre každú nadrovinu* T *priestoru* $\mathbb{R}^k$*.*

Poznámka 1.1. Budeme hovoriť, že náhodná veličina $\boldsymbol{X}$ má rozdelenie R, ak medzi nimi platí vzťah popísaný v definícii 1.2. Túto skutočnosť budeme značiť $\boldsymbol{X} \sim R$.

Príklad 1.1. Majme pravdepodobnostné rozdelenie dané hustotou

$$f(\boldsymbol{x}, \boldsymbol{\mu}, \kappa) = \begin{cases} c_k(\kappa) e^{\kappa \boldsymbol{\mu}^\top \boldsymbol{x}}, & \text{pre } \boldsymbol{x} \in \mathbb{S}^{k-1}, \\ 0, & \text{pre } \boldsymbol{x} \notin \mathbb{S}^{k-1}, \end{cases}$$

kde $c_k(\kappa)$ je normalizačná konštanta daná predpisom

$$c_k(\kappa) = \frac{\kappa^{k/2-1}}{(2\pi)^{k/2} I_{k/2-1}(\kappa)},$$

kde I_j je modifikovaná Besselova funkcia prvého druhu j-tého rádu. Takto definované rozdelenie spĺňa definíciu spojitého rozdelenia na k-sfére a je známe pod názvom von Mises-Fisher rozdelenie. Toto rozdelenie je považované za istý ekvivalent normálneho rozdelenia na k-sfére. Parameter $\boldsymbol{\mu} \in \mathbb{S}^{k-1}$ predstavuje bod okolo, ktorého je toto rozdelenie sústredené. Parameter $\kappa \geq 0$ je parametrom koncentrácie, čím vyššia je hodnota κ, tým viac je rozdelenie koncentrované okolo bodu $\boldsymbol{\mu}$. Pre $k = 2$ je toto rozdelenie známe aj pod názvom von Mises rozdelenie, v tomto prípade sa hustota zjednoduší na tvar

$$f(\theta, \mu, \kappa) = \begin{cases} \frac{1}{2\pi I_0(\kappa)} e^{\kappa \cos(\theta - \mu)}, & \text{pre } \theta \in [0, 2\pi), \\ 0, & \text{pre } \theta \notin [0, 2\pi). \end{cases}$$

Čitatelia sa viac o von Mises-Fisher rozdelení a odhadoch jeho parametrov $\boldsymbol{\mu}$ a κ môžu dozvedieť v článku Dhillon a Sra (2003).

Definícia 1.3 (Smerové dáta)
Nech $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n$ je náhodný výber z rozdelenia na k-sfére. Potom tento náhodný výber nazývame ***k-rozmernými smerovými dátami****. Špeciálne 2-rozmerné smerové dáta budeme nazývať* ***cirkulárnymi dátami****.*

Vďaka prirodzenému výskytu smerových dát je rozumné ich skúmať pomocou štatistických metód. Na analýzu týchto dát, ale nemôžeme použiť štandardné štatistické nástroje definované pre dáta v euklidovských priestoroch. Štandardná definícia strednej hodnoty aplikovaná v priestore $\mathbb{R}^k$ umožňuje výsledky mimo obor pozorovaných dát. Ukáže sa, že ani štandardné definície charakteristiky variability nie sú vhodné. Všetky tieto charakteristiky je preto potrebné pre smerové dáta definovať inými spôsobmi. Rovnako musíme pristupovať aj k štandardným metódam definovania kvantilov pre jednorozmerné euklidovské dáta. Tie v tomto prípade nebudú vhodné, keďže určenie usporiadania na kruhu je podriadené zvoleniu nulového smeru a určeniu, ktorý smer otáčania budeme považovať za kladný. Na jednotkových sférach vo vyšších dimenziách sa vyskytnú podobné problémy. V tejto kapitole sa budeme venovať možným definíciám charakteristík pre smerové dáta, ktoré budú predstavovať obdobu kvantilov.

Pred predstavením rôznych postupov ešte uvedieme vlastnosť rotačnej invariancie, ktorá bude ekvivalentom translačnej invariancie pre smerové dáta. Rotačná invariancia bude jedna z vlastností, ktoré budeme skúmať pri rôznych prístupoch k usporiadaniu smerových dát.

Definícia 1.4 (Rotačne invariantné zobrazenie)
Majme zobrazenie $f : \mathbb{R}^k \to \mathbb{R}^k$. Zobrazenie f nazývame ***rotačne invariantné****, keď platí $f(\boldsymbol{x}) = f(O\boldsymbol{x})$ pre všetky $\boldsymbol{x} \in \mathbb{R}^k$ a všetky O matice rotácií v $\mathbb{R}^k$.*

1.2 Intuitívne prístupy k usporiadaniu smerových dát

Priamočiarym prístupom je transformácia jednotkovej k-sféry $\mathbb{S}^{k-1}$ na $\mathbb{R}^{k-1}$ a aplikovanie definícií kvantilov, respektíve ich alternatív, na transformované dáta. Tento postup má však viaceré problémy. Už zvolenie transformácie nie je jednoznačné, navyše rôzne voľby transformácie dávajú rôzne výsledky. Body, ktorých vzdialenosť na sfére bola malá, môžu byť v transformovanom priestore naopak veľmi vzdialené. Výsledok transformácie je podmienený výberu nulového smeru, teda v závislosti na výbere nulového smeru je možné dostať rôzne výsledky. Takýto postup určite nemôže byť rotačne invariantný, od definícií v ďalších odstavcoch budeme túto vlastnosť požadovať.

V prípade dát, ktoré sú vysoko koncentrované je možné časť sféry s najväčšou koncentráciou aproximovať dotykovou nadrovinou a na tejto nadrovine využívať postupy štandardne platné v euklidovských priestoroch. Na rozdiel od prvého spomenutého postupu je tento prístup použiteľný, ale dáta, pre ktoré by sa dal považovať za jeden z vhodných, sú v praxi skôr vzácne. V ďalšej podkapitole budeme definovať funkcie, ktoré budú použiteľné ako obdoba kvantilov bez ohľadu na koncentráciu dát.

1.3 Uhlové hĺbky

Takzvané hĺbkové funkcie sú bežne používané pre viacrozmerné dáta. Tieto funkcie slúžia ako ukazovateľ centrality a k usporiadaniu dát od centra von, ktoré môže slúžiť ako obdoba kvantilov pre viacrozmerné dáta. V tejto podkapitole budeme skúmať analógie hĺbkových funkcií pre smerové dáta, tie sa v literatúre zvyknú označovať ako **uhlové hĺbky**. Všetky takéto funkcie budú spĺňať našu požiadavku rotačnej invariancie. Ako prvé uvedieme dve uhlové hĺbky, ktoré sú ekvivalentom zaužívaných hĺbok pre viacrozmerné dáta v euklidovských priestoroch: uhlovú polpriestorovú hĺbku a uhlovú simplexovú hĺbku, ktoré boli predstavené v článkoch Small (1987) a Liu a Singh (1992). Ako ďalšie uvedieme uhlové hĺbky odvodené od rôznych definícií vzdialeností na jednotkových sférach, tieto hĺbky boli zavedené v článkoch Liu a Singh (1992) a Pandolfo a kol. (2018).

Definícia 1.5 (Polsféra)
Majme jednotkovú k-sféru $\mathbb{S}^{k-1}$, ďalej nech $\boldsymbol{u} \in \mathbb{S}^{k-1}$.
*Množinu $A = \{\langle \boldsymbol{x},\boldsymbol{u}\rangle \geq 0 \colon \boldsymbol{x} \in \mathbb{S}^{k-1}\}$ budeme nazývať **uzavretou polsférou**.*
*A množinu $B = \{\langle \boldsymbol{x},\boldsymbol{u}\rangle > 0 \colon \boldsymbol{x} \in \mathbb{S}^{k-1}\}$ budeme nazývať **otvorenou polsférou**.*

Definícia 1.6 (Uhlová polpriestorová hĺbka)
*Nech R je pravdepodobnostné rozdelenie na k-sfére, potom definujeme **uhlovú polpriestorovú hĺbku** bodu $\boldsymbol{x}$ vzhľadom k R, ako*

$$uPH(\boldsymbol{x},R) = \inf_{A\colon \boldsymbol{x} \in A} R(A),$$

kde infimum berieme cez všetky uzavreté polsféry A obsahujúce $\boldsymbol{x}$.

Ďalej uvedieme tri tvrdenia o uhlovej polriestorovej hĺbke. Prvé z tvrdení je zovšeobecnenie tvrdenia formulovaného v článku Liu a Singh (1992) a hovorí o obmedzenosti zhora, druhé tvrdenie hovorí o obmedzení zdola a nadobúdaní minima. Ako posledné uvedieme tvrdenie o rotačnej invariancii tejto uhlovej hĺbky.

Tvrdenie 1.1. *Nech R je spojité pravdepodobnostné rozdelenie na k-sfére. Potom uhlová polpriestorová hĺbka je zhora obmedzená $1/2$ a nadobúda maxima $1/2$ v bode $\boldsymbol{x}$ práve, keď pre každú polsféru A obsahujúcu $\boldsymbol{x}$ platí $R(A) \geq 1/2$.*

Tvrdenie 1.2. *Nech R je spojité pravdepodobnostné rozdelenie na k-sfére $\mathbb{S}^{k-1}$. Potom uhlová polpriestorová hĺbka nadobúda svojho minima na celej uzavretej polsfére A^*, pre ktorú platí*

$$A^* = \operatorname*{arg\,min}_{A \subset \mathbb{S}^{k-1}} R(A),$$

kde argument minima hľadáme medzi všetkými uzavretými polsférami A.

Dôkaz. Vyplýva z tvrdenia „Theorem 7“ z článku Nagy a Laketa (2022) a z definície spojitého rozdelenia na k-sfére.

□

Tvrdenie 1.2 nám hovorí, že uhlová polpriestorová hĺbka je vždy konštantná aspoň na jednej polsfére. To znamená, že táto hĺbka nám umožňuje vytvoriť

usporiadanie len na jednej polsfére. Navyše, ak toto rozdelenie bude rovnomerné, uhlová polpriestorová hĺbka je rovná 1/2 na celej sfére. Uvidíme, že podobné problémy sa pre rovnomerné rozdelenie na sfére vyskytnú aj pre iné definície uhlových hĺbok.

Tvrdenie 1.3. *Nech R je pravdepodobnostné rozdelenie na k-sfére, potom uhlová polpriestorová hĺbka je rotačne invariantná.*

Dôkaz. Nech $\boldsymbol{Y} \sim R$. Označme R_O rozdelenie transformovaného vektoru $O\boldsymbol{Y}$. Majme zobrazenie $h : \boldsymbol{x} \mapsto O\boldsymbol{x}$. Dostávame teda

$$P(h(\boldsymbol{Y}) \in h(B)) = P(h^{-1}(h(\boldsymbol{Y})) \in h^{-1}(h(B))) = P(\boldsymbol{Y} \in B), \tag{1.1}$$
$$\boldsymbol{Y} \in B \Leftrightarrow h(\boldsymbol{Y}) \in h(B)$$

pre všetky borelovské množiny $B \subset \mathbb{R}^k$. Takže platia aj nasledovné rovnosti

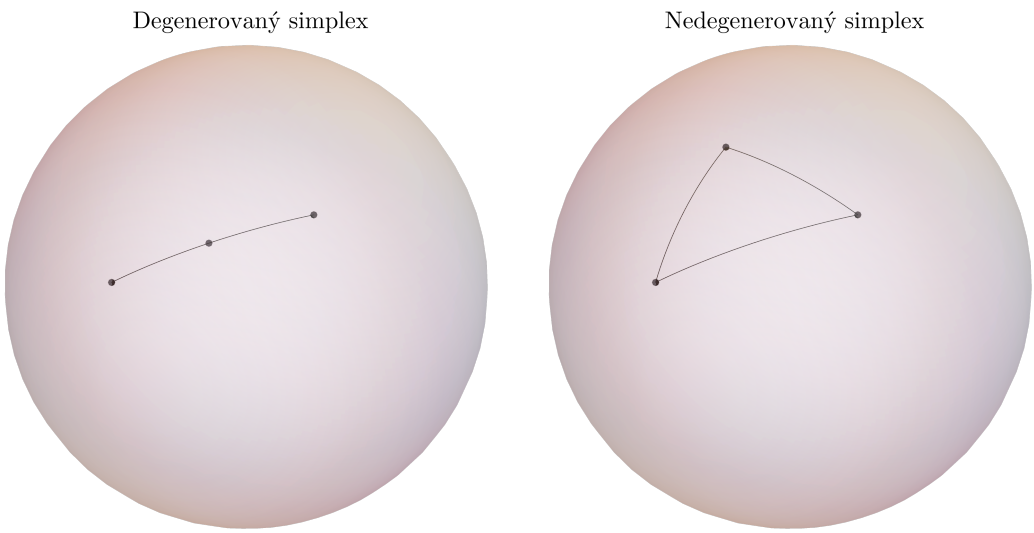
$$\begin{aligned} uPH(\boldsymbol{x},R) &= \inf_{A:\, \boldsymbol{x} \in A} R(A) = \inf_{A:\, \boldsymbol{x} \in A} R_O(h(A)) \\ &= \inf_{h(A):\, O\boldsymbol{x} \in h(A)} R_O(h(A)) = uPH(O\boldsymbol{x},R_O). \end{aligned}$$

□

Simplexom v $\mathbb{R}^k$ rozumieme konvexný obal $k+1$ affínne nezávislých bodov. Podobne vieme definovať simplex na jednotkovej k-sfére $\mathbb{S}^{k-1}$. Hrany takéhoto simplexu budú tvorené kratšími kruhovými oblúkmi, ktoré spájajú jednotlivé vrcholy.

Definícia 1.7 (Simplex na k-sfére)
Majme jednotkovú k-sféru $\mathbb{S}^{k-1}$ a jej podmnožiny M, $L \subset \mathbb{S}^{k-1}$. Hovoríme, že M je ***sféricky konvexná****, keď kužeľ $K = \{\boldsymbol{a} \in \mathbb{R}^k : \boldsymbol{a} = t\boldsymbol{x}, t \in [0,\infty), \boldsymbol{x} \in M\}$ je konvexný. Ďalej definujeme* ***sférický konvexný obal*** *L ako najmenšiu sféricky konvexnú množinu obsahujúcu L.*
Simplex na k-sfére *definujeme ako sférický konvexný obal k bodov.*

V definícii 1.7 dovoľujeme aj degenerované prípady, keď simplex tvoria vrcholy, ktoré nie sú affínne nezávislé. V takomto prípade výsledný útvar nie je $(k-1)$-dimenzionálny útvar, ale útvar menšej dimenzie ako si môžeme všimnúť na obrázku 1.1, na ktorom je porovnaný degenerovaný a nedegenerovaný simplex na 3-sfére. Podobne dostávame degenerované simplexy v prípade ak sú súčasťou množiny bodov, ktorá simplex určuje zároveň bod $\boldsymbol{x} \in \mathbb{S}^{k-1}$ a bod $-\boldsymbol{x} \in \mathbb{S}^{k-1}$. V prípade simplexu na 2-sfére daného bodmi $-\boldsymbol{x}$, $\boldsymbol{x} \in \mathbb{S}^1$ tvoria tento simplex len body $\boldsymbol{x}$ a $-\boldsymbol{x}$. V prípade simplexu na 3-sfére daného bodmi $-\boldsymbol{x}$, $\boldsymbol{x}$, $\boldsymbol{y} \in \mathbb{S}^2$ je výsledný útvar polkružnica l spájajúca body $-\boldsymbol{x}$, $\boldsymbol{x}$, pre ktorú platí $\boldsymbol{y} \in l$.



Obr. 1.1: Degenerovaný a nedegenerovaný simplex na 3-sfére.

Definícia 1.8 (Uhlová simplexová hĺbka)
Nech R je pravdepodobnostné rozdelenie na k-sfére a $\boldsymbol{X}_1,\ldots,\boldsymbol{X}_k$ je náhodným výberom z tohto rozdelenia, potom definujeme ***uhlovú simplexovú hĺbku*** *bodu $\boldsymbol{x}$ vzhľadom k R, ako*

$$uSH(\boldsymbol{x},R) = R(\boldsymbol{x} \in \Delta^k(\boldsymbol{X}_1,\ldots,\boldsymbol{X}_k)),$$

kde $\Delta^k(\boldsymbol{X}_1,\ldots,\boldsymbol{X}_k)$ značí simplex na jednotkovej k-sfére $\mathbb{S}^{k-1}$, ktorý je určený bodmi $\boldsymbol{X}_1,\ldots,\boldsymbol{X}_k$.

Ak je R spojité rozdelenie, simplex $\Delta^k(\boldsymbol{X}_1,\ldots,\boldsymbol{X}_k)$ bude degenerovaný práve s pravdepodobnosťou nula. Avšak, ak R nebude spojité, simplex $\Delta^k(\boldsymbol{X}_1,\ldots,\boldsymbol{X}_k)$ bude degenerovaný s pravdepodobnosťou väčšou ako nula. Táto situácia sa môže vyskytnúť aj v prípade empirických dát, teda je žiaduce skúmať aj tieto prípady.

Tvrdenie 1.4. *Nech R je pravdepodobnostné rozdelenie na k-sfére, potom uhlová simplexová hĺbka je rotačne invariantná.*

Dôkaz. Nech $\boldsymbol{Y} \sim R$. Označme R_O rozdelenie transformovaného vektoru $O\boldsymbol{Y}$. Majme zobrazenie $h : \boldsymbol{x} \mapsto O\boldsymbol{x}$. Podobne ako v dôkaze tvrdenia 1.3 využijeme vlastnosť (1.1). Takže platí aj nasledovné:

$$\begin{aligned} uSH(\boldsymbol{x},R) &= R(\boldsymbol{x} \in \Delta^k(\boldsymbol{X}_1,\ldots,\boldsymbol{X}_k)) = R(h(\boldsymbol{x}) \in h(\Delta^k(\boldsymbol{X}_1,\ldots,\boldsymbol{X}_k))) \\ &= R(O\boldsymbol{x} \in \Delta^k(h(\boldsymbol{X}_1),\ldots,h(\boldsymbol{X}_k))) = R_O(O\boldsymbol{x} \in \Delta^k(\boldsymbol{X}_1,\ldots,\boldsymbol{X}_k)) \\ &= uSH(O\boldsymbol{x},R_O). \end{aligned}$$

□

V článku Liu a Singh (1992) autori uviedli ekvivalentnú podmienku pre vzťah $\boldsymbol{x} \in \Delta^k(\boldsymbol{X}_1,\ldots,\boldsymbol{X}_k)$, ktorá tento vzťah prevedie na sústavu lineárnych rovníc.

Tvrdenie 1.5. *Vzťah* $\boldsymbol{x} \in \Delta^k(\boldsymbol{X}_1,\ldots,\boldsymbol{X}_k)$ *je ekvivalentný tomu, že prienik úsečky* $\overline{\boldsymbol{0}_k\boldsymbol{x}} = \{t\boldsymbol{x}; t \in (0,1]\}$ *a simplexu v* $\mathbb{R}^k$ *daného bodmi* $\boldsymbol{X}_1,\ldots,\boldsymbol{X}_k$ *je neprázdny. Teda je ekvivalentný sústave rovností* $\alpha_0\boldsymbol{x} = \alpha_1\boldsymbol{X}_1 + \cdots + \alpha_n\boldsymbol{X}_k$ *za podmienky* $\alpha_0 \in [0,1]$, $\sum_{i=1}^n \alpha_i = 1$, $\alpha_i \in [0,1]$ *pre* $i = 1,\ldots,n$.

Dôkaz. Sférický simplex $\Delta^k(\boldsymbol{X}_1,\ldots,\boldsymbol{X}_k)$ jednoznačne určuje konvexný kužeľ v $\mathbb{R}^k$. Označme tento kužeľ K. Kužeľu K náležia všetky body tvaru $\alpha\boldsymbol{X}_1 + (1-\alpha)\boldsymbol{X}_2$; $\alpha \in [0,1]$. Indukciou teda vieme ukázať, kužeľu K náležia aj body tvaru $\alpha_1\boldsymbol{X}_1 + \cdots + \alpha_n\boldsymbol{X}_n$; $\sum_{i=1}^k \alpha_i = 1$ a $\alpha_i \in [0,1]$ pre $i = 1,\ldots,n$. Teda $\Delta_R^k(\boldsymbol{X}_1,\ldots,\boldsymbol{X}_k) \subset K$, kde $\Delta_R^k(\boldsymbol{X}_1,\ldots,\boldsymbol{X}_k)$ značí euklidovský simplex v $\mathbb{R}^k$ daný bodmi $\boldsymbol{X}_1,\ldots,\boldsymbol{X}_k$. Tento simplex tvorí rez kužeľom K. Ak teda platí $\boldsymbol{x} \in K$ a $\boldsymbol{x} \in \mathbb{S}^{k-1}$, musí byť prienik úsečky $\overline{\boldsymbol{0}_k\boldsymbol{x}}$ a $\Delta_R^k(\boldsymbol{X}_1,\ldots,\boldsymbol{X}_k)$ neprázdny. Navyše platí $(\boldsymbol{x} \in K$ a $\boldsymbol{x} \in \mathbb{S}^{k-1}) \Leftrightarrow \boldsymbol{x} \in \Delta^k(\boldsymbol{X}_1,\ldots,\boldsymbol{X}_k)$.

□

Uhlové hĺbky odvodené od metrík

V nasledujúcich definíciách postupne zavedieme metriky na k-sfére a triedu uhlových hĺbok odvodených od rôznych obmedzených metrík. Podrobnejšie budeme skúmať tri uhlové hĺbky z tejto triedy: oblúkovú, kosínusovú a tetivovú.

Definícia 1.9 (Metrika na k-sfére)
Funkciu d budeme nazývať ***metrikou na k-sfére*** $\mathbb{S}^{k-1}$*, ak je tvaru:*

$$d : \mathbb{S}^{k-1} \times \mathbb{S}^{k-1} \to [0,\infty)$$

a pre všetky $\boldsymbol{p},\boldsymbol{q},\boldsymbol{r} \in \mathbb{S}^{k-1}$ *platí nasledujúce:*

1. $d(\boldsymbol{p},\boldsymbol{q}) \geq 0$,
2. $d(\boldsymbol{p},\boldsymbol{q}) = d(\boldsymbol{q},\boldsymbol{p})$,
3. $d(\boldsymbol{p},\boldsymbol{q}) = 0 \Leftrightarrow \boldsymbol{p} = \boldsymbol{q}$,
4. $d(\boldsymbol{p},\boldsymbol{r}) \leq d(\boldsymbol{p},\boldsymbol{q}) + d(\boldsymbol{q},\boldsymbol{r})$.

Definícia 1.10 (Uhlová hĺbka odvodená od metriky)
Nech d je obmedzená metrika na jednotkovej k-sfére $\mathbb{S}^{k-1}$ *a R je pravdepodobnostné rozdelenie na k-sfére. Potom definujeme* ***uhlovú hĺbku odvodenú od metriky d*** *vzhľadom k R, ako*

$$uH_d(\boldsymbol{x},R) = \sup_{\boldsymbol{p},\,\boldsymbol{q} \in \mathbb{S}^{k-1}} d(\boldsymbol{p},\boldsymbol{q}) - \mathsf{E}_R[d(\boldsymbol{x},\boldsymbol{Y})], \tag{1.2}$$

kde E_R *je stredná hodnota za predpokladu, že* $\boldsymbol{Y}$ *má rozdelenie R.*

Nás budú zaujímať uhlové hĺbky odvodené od rotačne invariantných metrík, keďže odvodenie od rotačne invariantnej metriky zaistí, že aj uhlová hĺbka bude rotačne invariantná. V nasledujúcom tvrdení ukážeme vlastnosť rotačne invariantných metrík, ktorá nám umožní jednoduchšie skúmanie uhlových hĺbok odvodených od takýchto metrík. Toto tvrdenie a dôkaz boli formulované v článku Pandolfo a kol. (2018). V tejto práci uvedieme rozšírený a sprehľadnený dôkaz tohto tvrdenia.

Tvrdenie 1.6. *Nech d je metrika na k-sfére $\mathbb{S}^{k-1}$. Potom je d rotačne invariantná práve vtedy keď existuje funkcia f taká, že platí*

$$\begin{aligned} &f : [-1,1] \to [0,\infty), \\ &d(\boldsymbol{p},\boldsymbol{q}) = f(\boldsymbol{p}^\top \boldsymbol{q}) = f(\langle \boldsymbol{p},\boldsymbol{q}\rangle), \end{aligned} \tag{1.3}$$

pre všetky $\boldsymbol{p}, \boldsymbol{q} \in \mathbb{S}^{k-1}$.

Dôkaz. Nech je d rotačne invariantná. Majme $\boldsymbol{p},\boldsymbol{q} \in \mathbb{S}^{k-1}$ také, že $\boldsymbol{p} \neq \boldsymbol{q}$, položme $\boldsymbol{s} = (\boldsymbol{q} - (\boldsymbol{p}^\top \boldsymbol{q})\boldsymbol{p})/\|\boldsymbol{q} - (\boldsymbol{p}^\top \boldsymbol{q})\boldsymbol{p}\|$. Označme Q maticu tvaru $k \times (k-2)$ takú, že $O = (\boldsymbol{p} \mid \boldsymbol{s} \mid Q)$ je ortogonálna. Keďže d je rotačne invariantná, platí $d(\boldsymbol{p},\boldsymbol{q}) = d(O^\top \boldsymbol{p},O^\top \boldsymbol{q})$, $\boldsymbol{p}^\top \boldsymbol{p} = \boldsymbol{q}^\top \boldsymbol{q} = 1$, $\boldsymbol{p}^\top \boldsymbol{s} = 0$ a $Q^\top \boldsymbol{p} = \boldsymbol{0}_{k-2}$, navyše

$$\begin{aligned} \boldsymbol{s}^\top \boldsymbol{q} &= \frac{(\boldsymbol{q} - (\boldsymbol{p}^\top \boldsymbol{q})\boldsymbol{p})^\top \boldsymbol{q}}{\|\boldsymbol{q} - (\boldsymbol{p}^\top \boldsymbol{q})\boldsymbol{p}\|} = \frac{1 - (\boldsymbol{p}^\top \boldsymbol{q})^2}{\sqrt{(\boldsymbol{q} - (\boldsymbol{p}^\top \boldsymbol{q})\boldsymbol{p})^\top (\boldsymbol{q} - (\boldsymbol{p}^\top \boldsymbol{q})\boldsymbol{p})}} \\ &= \frac{1 - (\boldsymbol{p}^\top \boldsymbol{q})^2}{\sqrt{(\boldsymbol{q}^\top (\boldsymbol{q} - (\boldsymbol{p}^\top \boldsymbol{q})\boldsymbol{p}) - 0}} = \frac{1 - (\boldsymbol{p}^\top \boldsymbol{q})^2}{\sqrt{1 - (\boldsymbol{p}^\top \boldsymbol{q})^2}} = \sqrt{1 - (\boldsymbol{p}^\top \boldsymbol{q})^2}, \\ Q^\top \boldsymbol{s} &= \frac{Q^\top (\boldsymbol{q} - (\boldsymbol{p}^\top \boldsymbol{q})\boldsymbol{p})}{\|\boldsymbol{q} - (\boldsymbol{p}^\top \boldsymbol{q})\boldsymbol{p}\|} = \boldsymbol{0}_{k-2} \Rightarrow Q^\top \boldsymbol{q} = (\boldsymbol{p}^\top \boldsymbol{q}) Q^\top \boldsymbol{p} = \boldsymbol{0}_{k-2}. \end{aligned}$$

Dostávame teda

$$\begin{aligned} O^\top \boldsymbol{p} &= (1,0,\ldots,0)^\top = \boldsymbol{e}_1, \\ O^\top \boldsymbol{q} &= (\boldsymbol{p}^\top \boldsymbol{q}, \sqrt{1 - (\boldsymbol{p}^\top \boldsymbol{q})^2}, 0, \ldots, 0)^\top =: \boldsymbol{l}. \end{aligned}$$

Teda $d(\boldsymbol{p},\boldsymbol{q}) = d(\boldsymbol{e}_1,\boldsymbol{l})$ závisí na $\boldsymbol{p},\boldsymbol{q}$ len skrz $\boldsymbol{p}^\top \boldsymbol{q}$, takže platí $d(\boldsymbol{p},\boldsymbol{q}) = f(\boldsymbol{p}^\top \boldsymbol{q})$ pre nejakú funkciu f. Pre $\boldsymbol{p} = \boldsymbol{q}$ máme $d(\boldsymbol{p},\boldsymbol{q}) = 0$, položme $f(\boldsymbol{p}^\top \boldsymbol{p}) = 0$. Takto dostávame funkciu f, pre ktorú platí $d(\boldsymbol{p},\boldsymbol{q}) = f(\boldsymbol{p}^\top \boldsymbol{q})$ pre všetky $\boldsymbol{p},\boldsymbol{q} \in \mathbb{S}^{k-1}$.

Nech je teraz naopak d definovaná vzťahom (1.3). Jednoduchým odvodením dostávame:

$$d(\boldsymbol{p},\boldsymbol{q}) = f(\langle \boldsymbol{p},\boldsymbol{q}\rangle) = f(\boldsymbol{p}^\top \boldsymbol{q}) = f(\boldsymbol{p}^\top O^\top O \boldsymbol{q}) = f(\langle O\boldsymbol{p}, O\boldsymbol{q}\rangle) = d(O\boldsymbol{p},O\boldsymbol{q}\rangle),$$

pre všetky O matice rotácií v $\mathbb{R}^k$.

□

Definícia 1.11 (Oblúková, kosínusová, tetivová metrika)
Postupne definujeme ***oblúkovú, kosínusovú a tetivovú metriku*** *ako:*

$$\begin{aligned} d_o(\boldsymbol{p},\boldsymbol{q}) &= g_1(\langle \boldsymbol{p},\boldsymbol{q}\rangle); \quad g_1(t) = \arccos t, \\ d_k(\boldsymbol{p},\boldsymbol{q}) &= g_2(\langle \boldsymbol{p},\boldsymbol{q}\rangle); \quad g_2(t) = 1 - t, \\ d_t(\boldsymbol{p},\boldsymbol{q}) &= g_3(\langle \boldsymbol{p},\boldsymbol{q}\rangle); \quad g_3(t) = \sqrt{2(1-t)}. \end{aligned}$$

Poznámka 1.2. Keďže oblúkovú, kosínusovú a tetivovú hĺbku definujeme pomocou funkcií g_1, g_2, g_3, ktoré spĺňajú podmienky pre funkciu f z tvrdenia 1.6, teda z tohto tvrdenia vyplýva rotačná invariancia týchto metrík.

Dosadením vyššie definovaných metrík do rovnosti (1.2) dostávame postupne **uhlovú oblúkovú**, **kosínusovú** a **tetivovú hĺbku** vzhľadom k R, ktoré značíme $uH_{d_o}(\boldsymbol{x},R), uH_{d_c}(\boldsymbol{x},R), uH_{d_t}(\boldsymbol{x},R)$.

Poznámka 1.3. Všetky vyššie definované metriky nadobúdajú svojho maxima, keď platí $\boldsymbol{p}^\top \boldsymbol{q} = -1$, dostávame teda

$$uH_d(\boldsymbol{x},R) = g(-1) - \mathsf{E}_R(d(\boldsymbol{x},\boldsymbol{Y})).$$

Postupne máme

$$\begin{aligned} uH_{d_o}(\boldsymbol{x},R) &= \pi - \mathsf{E}_R(\arccos(\boldsymbol{x}^\top \boldsymbol{Y})), \\ uH_{d_k}(\boldsymbol{x},R) &= 2 - \mathsf{E}_R(1 - (\boldsymbol{x}^\top \boldsymbol{Y})) = 1 + E_R(\boldsymbol{x}^\top \boldsymbol{Y}), \\ uH_{d_t}(\boldsymbol{x},R) &= 2 - \mathsf{E}_R(\sqrt{2(1 - \boldsymbol{x}^\top \boldsymbol{Y})}). \end{aligned}$$

Ďalej ukážeme rotačnú invarianciu uhlových hĺbok odvodených od oblúkovej, kosínusovej a tetivovej metriky. Táto vlastnosť spomínaných uhlových hĺbok je priamym dôsledkom rotačnej invariancie metrík, od ktorých sú dané uhlové hĺbky odvodené.

Tvrdenie 1.7. *Uhlová oblúková, uhlová kosínusová a uhlová tetivová hĺbka sú rotačne invariantné.*

Dôkaz. Metriky, od ktorých sú tieto hĺbky odvodené, sú podľa tvrdenia 1.6 rotačne invariantné, máme teda

$$d(\boldsymbol{p},\boldsymbol{q}) = d(O\boldsymbol{p},O\boldsymbol{q}); \quad \text{pre všetky } O \text{ matice rotácií v } \mathbb{R}^k.$$

Navyše pre všetky O matice rotácií v $\mathbb{R}^k$ platí

$$(O\boldsymbol{x})^\top O\boldsymbol{y} = \boldsymbol{x}^\top O^\top O\boldsymbol{y} = \boldsymbol{x}^\top O^{-1} O\boldsymbol{y} = \boldsymbol{x}^\top \boldsymbol{y}.$$

Dostávame

$$\begin{aligned} d(O\boldsymbol{p},O\boldsymbol{q}) - \mathsf{E}_R(d(O\boldsymbol{x},O\boldsymbol{Y})) &= d(O\boldsymbol{p},O\boldsymbol{q}) - E_R(g((O\boldsymbol{x})^\top O\boldsymbol{Y})) \\ &= d(\boldsymbol{p},\boldsymbol{q}) - \mathsf{E}_R(g(\boldsymbol{x}^\top \boldsymbol{Y})) \\ &= d(\boldsymbol{p},\boldsymbol{q}) - \mathsf{E}_R(d(\boldsymbol{x},\boldsymbol{Y})) \end{aligned}$$

pre všetky O matice rotácií v $\mathbb{R}^k$.

□

Poznámka 1.4. Pre funkcie $g_1(t)$ a $g_2(t)$ z definície 1.11 platí

$$\begin{aligned} g_1(t) + g_1(-t) &= \arccos t + \arccos(-t) = \pi, \text{ pre všetky } t \in [-1,1] \\ g_2(t) + g_2(-t) &= 1 + t + 1 - t = 2, \text{ pre všetky } t \in [-1,1]. \end{aligned}$$

Z toho pre všetky $\boldsymbol{p},\boldsymbol{q} \in \mathbb{S}^{k-1}$ dostávame

$$\begin{aligned} d_o(\boldsymbol{p},\boldsymbol{q}) + d_o(-\boldsymbol{p},\boldsymbol{q}) = \pi &\Rightarrow \mathsf{E}d_o(\boldsymbol{p},\boldsymbol{q}) + \mathsf{E}d_o(-\boldsymbol{p},\boldsymbol{q}) = \pi \\ &\Rightarrow uH_{d_o}(\boldsymbol{x},R) + uH_{d_o}(-\boldsymbol{x},R) = \pi, \\ d_k(\boldsymbol{p},\boldsymbol{q}) + d_k(-\boldsymbol{p},\boldsymbol{q}) = 2 &\Rightarrow \mathsf{E}d_k(\boldsymbol{p},\boldsymbol{q}) + \mathsf{E}d_k(-\boldsymbol{p},\boldsymbol{q}) = 2 \\ &\Rightarrow uH_{d_k}(\boldsymbol{x},R) + uH_{d_k}(-\boldsymbol{x},R) = 2. \end{aligned}$$

Pre uhlovú tetivovú hĺbku žiadne podobné vyjadrenie neexistuje.

2. Definície mediánu a kvantilov

V tejto kapitole uvedieme rôzne definície pre medián smerových dát a predstavíme koncept projekčných kvantilov prvýkrát definovaných v článku Ley a kol. (2014).

Definícia 2.1 (Fisherov medián)
*Nech R je pravdepodobnostné rozdelenie na k-sfére a $\boldsymbol{Y} \sim R$, potom **Fisherov medián** náhodného vektoru $\boldsymbol{Y}$ definujeme ako*

$$m_F(\boldsymbol{Y}) = \underset{\boldsymbol{x} \in \mathbb{S}^{k-1}}{\arg\min}\, \mathsf{E}(d_o(\boldsymbol{x},\boldsymbol{Y})),$$

*kde d_o značí oblúkovú metriku na $\mathbb{S}^{k-1}$. Hovoríme, že Fisherov medián je **unikátny**, ak $|m_F(\boldsymbol{Y})| = 1$, kde $|M|$ značí mohutnosť množiny M. Ak $|m_F(\boldsymbol{Y})| > 1$, hovoríme, že Fisherov medián je **viacnásobný**.*

Obdoba tejto definície pre trojrozmerné smerové dáta sa prvýkrát objavila v článku Fisher (1985). Jej myšlienka je rovnaká ako v prípade jednorozmerného alebo priestorového mediánu – minimalizácia vzdialenosti. Na rozdiel od priestorového mediánu, ale zohľadňuje zakrivenie sféry, keďže ako metriku berie dĺžku kratšieho oblúku hlavnej kružnice spájajúcej dva body. Rotačná invariancia tohto mediánu vyplýva z rotačnej invariancie oblúkovej metriky.

K zavedeniu mediánu môžeme použiť aj uhlové hĺbky, ktoré sme definovali v predchádzajúcej kapitole. Keďže sme dokázali rotačnú invarianciu týchto funkcií, bude aj medián definovaný pomocou nich rotačne invariantný.

Definícia 2.2 (Medián daný uhlovou hĺbkou)
*Nech R je pravdepodobnostné rozdelenie na k-sfére a $\boldsymbol{Y} \sim R$ a nech uH je uhlová hĺbka, potom **medián** náhodného vektoru $\boldsymbol{Y}$ **daný uhlovou hĺbkou uH** definujeme ako*

$$m_{uH}(\boldsymbol{Y}) = \underset{\boldsymbol{x} \in \mathbb{S}^{k-1}}{\arg\max}\, uH(\boldsymbol{x},R).$$

*Hovoríme, že medián daný uhlovou hĺbkou uH je **unikátny**, ak $|m_{uH}(\boldsymbol{Y})| = 1$. Ak $|m_{uH}(\boldsymbol{Y})| > 1$, hovoríme, že medián daný uhlovou hĺbkou uH je **viacnásobný**.*

Definícia 2.3 (Antimedián)
*Nech R je pravdepodobnostné rozdelenie na k-sfére a $\boldsymbol{Y} \sim R$ a nech $\boldsymbol{m}$ je unikátny medián tohto rozdelenia, potom **antimedián** náhodného vektoru $\boldsymbol{Y}$ definujeme ako*

$$\boldsymbol{m}_a = -\boldsymbol{m}.$$

Ďalej definujeme triedu rozdelení na k-sférach, ktorá bude charakteristická istou symetriou. Táto trieda rozdelení je typická viacnásobným mediánom. Do tejto triedy mimo iných spadá aj spojité rovnomerné rozdelenie na k-sfére. Pre rozdelenia z tejto triedy bude problematické vytvoriť rozumné usporiadanie. Najlepšie čo môžeme urobiť, je vytvoriť usporiadanie pre nejakú polsféru A a pre jej opačnú polsféru $-A$.

Definícia 2.4 (Antipodálne rozdelenia)
Nech R je pravdepodobnostné rozdelenie na k-sfére. R nazývame ***antipodálnym*** *rozdelením, ak platí*

$$R(B) = R(-B) \text{ pre všetky borelovské množiny } B \subset \mathbb{S}^{k-1},$$

kde $-B = \{-\boldsymbol{y}; \boldsymbol{y} \in B\}$.

Ekvivalentne R je antipodálne, ak platí $\boldsymbol{Y} \sim R \Leftrightarrow -\boldsymbol{Y} \sim R$.

Tvrdenie 2.1. *Antipodálne rozdelenie má viacnásobný medián bez ohľadu na výber definície mediánu.*

Dôkaz. Zobrazenie $\boldsymbol{x} \to -\boldsymbol{x}$ je dané maticou $-I_k$, kde I_k je jednotková matica. Maticu $-I_k$ môžeme chápať ako maticu rotácie. Nech $\boldsymbol{Y} \sim R$. Označme R_{-I_k} rozdelenie transformovaného vektoru $-I_k\boldsymbol{Y}$. V tvrdeniach 1.3, 1.4, 1.7 sme dokázali rotačnú invarianciu uhlových hĺbok, platí teda

$$uH(\boldsymbol{x},R) = uH(-I_k\boldsymbol{x},R_{-I_k}) = uH(-\boldsymbol{x},R).$$

V poslednej rovnosti využívame to, že platí $\boldsymbol{Y} \sim R \Leftrightarrow -\boldsymbol{Y} \sim R$, teda R a R_{-I_k} sú rovnaké rozdelenia. Z toho vyplýva

$$\boldsymbol{y} \in \underset{\boldsymbol{x} \in \mathbb{S}^{k-1}}{\arg\max}\, uH(\boldsymbol{x},R) \Leftrightarrow -\boldsymbol{y} \in \underset{\boldsymbol{x} \in \mathbb{S}^{k-1}}{\arg\max}\, uH(\boldsymbol{x},R).$$

Dokázali sme viacnásobnosť mediánu daného uhlovými hĺbkami definovanými v prvej kapitole. Dokážeme ešte viacnásobnosť Fisherovho mediánu. Majme R antipodálne rozdelenie a $\boldsymbol{Y} \sim R$. Keďže oblúková hĺbka je definovaná pomocou skalárneho súčinu, skoro iste platí

$$d_o(-\boldsymbol{x},\boldsymbol{Y}) = d_o(\boldsymbol{x}, -\boldsymbol{Y}).$$

Teda

$$\boldsymbol{y} \in \underset{\boldsymbol{x} \in \mathbb{S}^{k-1}}{\arg\min}\, \mathsf{E}(d_o(\boldsymbol{x},\boldsymbol{Y})) \Leftrightarrow -\boldsymbol{y} \in \underset{\boldsymbol{x} \in \mathbb{S}^{k-1}}{\arg\min}\, \mathsf{E}(d_o(\boldsymbol{x},\boldsymbol{Y})).$$

□

Pre uhlovú polpriestorovú a uhlovú simplexovú hĺbku platia silnejšie tvrdenia, ktoré boli formulované v článku Liu a Singh (1992) pre $k = 2, 3$. My uvedieme zovšeobecnenú verziu tvrdenia pre uhlovú polriestorovú hĺbku a tvrdenie pre uhlovú simplexovú hĺbku, ktoré bolo uvedené v spomenutom článku.

Tvrdenie 2.2. *Nech R je spojité antipodálne rozdelenie. Potom:*

$$uPH(\boldsymbol{x},R) = \frac{1}{2}, \text{ pre všetky } \boldsymbol{x} \in \mathbb{S}^{k-1}.$$

Dôkaz. Majme antipodálne spojité rozdelenie R. Teda platí $R(B) = R(-B)$ pre všetky merateľné množiny B. Z tejto vlastnosti vieme odvodiť, že každá uzavretá polsféra $A \subset \mathbb{S}^{k-1}$ nesie rovnakú pravdepodobnosť. Navyše musí platiť $P(A) + P(-A) = 1$, teda $P(A) = 1/2$ pre všetky polsféry $A \subset \mathbb{S}^{k-1}$.

□

Tvrdenie 2.3. *Nech R je absolútne spojité rozdelenie na 2-sfére a nech F je absolútne spojité rozdelenie na 3-sfére. Potom:*
R je antipodálne práve, keď platí $uSH(\boldsymbol{x},R) = 1/4$ pre všetky $\boldsymbol{x} \in \mathbb{S}^1$.
F je antipodálne práve, keď platí $uSH(\boldsymbol{x},F) = 1/8$ pre všetky $\boldsymbol{x} \in \mathbb{S}^2$.

Dôkaz. Dôkaz pre rozdelenie na 2-sfére nájdeme v článku Liu a Singh (1992) ako dôkaz tvrdenia „Proposition 3.3“, dôkaz pre rozdelenie na 3-sfére nájdeme v rovnakom článku ako dôkaz tvrdenia „Proposition 3.6“.

□

Ďalej definujeme projekčné kvantily, ktoré budú odvodené od mediánov. Projekčné kvantily budeme chápať ako kvantily rozdelenia daného projekciou $\boldsymbol{Y} \sim R$ na medián $\boldsymbol{m}$, teda kvantily náhodnej veličiny $\boldsymbol{Y}^\top \boldsymbol{m}$. Vidíme, že takýto postup bude dávať zmysel len pre rozdelenia s unikátnym Fisherovým mediánom. Po definovaní projekčných kvantilov zavedieme spodný a horný α-kvantilový uzáver, ktoré predstavujú isté rozdelenie jednotkovej k-sféry odvodené od projekčných kvantilov. Ako posledné definujeme hladinové α-kvantily, ktorých definícia bude taktiež založená na projekčných kvantiloch.

Definícia 2.5 (Projekčné kvantily)
*Nech R je spojité pravdepodobnostné rozdelenie na k-sfére, $\boldsymbol{Y} \sim R$ a nech $\boldsymbol{m}$ je unikátny Fisherov medián náhodného vektoru $\boldsymbol{Y}$. Potom definujeme **projekčný $\boldsymbol{\alpha}$-kvantil** náhodného vektoru $\boldsymbol{Y}$ ako*

$$c_\alpha(R) = \underset{c \in [-1,1]}{\arg\min}\, \mathsf{E}[\rho_\alpha(\boldsymbol{Y}^\top \boldsymbol{m} - c)]; \quad \textit{pre } \alpha \in (0,1),$$

kde $\rho_\alpha(u) = u(\alpha - \mathbb{1}\,(u \leq 0))$.

Predpis $\arg\min_{c \in [-1,1]} \mathsf{E}[\rho_\alpha(X - c)]$; pre $\alpha \in (0,1)$ je možné použiť ako definíciu kvantilov náhodnej veličiny X. „Lemma 3.4“ v skriptách Omelka (2022) tvrdí, že takto môžeme definovať empirické kvantily, dá sa však ukázať, že takýmto spôsobom je možné definovať kvantily aj na teoretickej úrovni.

Definícia 2.6 (Spodný a horný α-kvantilový uzáver)
*Nech R je pravdepodobnostné rozdelenie na k-sfére, c_α je projekčný α-kvantil rozdelenia R a $\boldsymbol{m}$ je jeho unikátny medián. Potom definujeme **spodný** a **horný $\boldsymbol{\alpha}$-kvantilový uzáver** rozdelenia R ako*

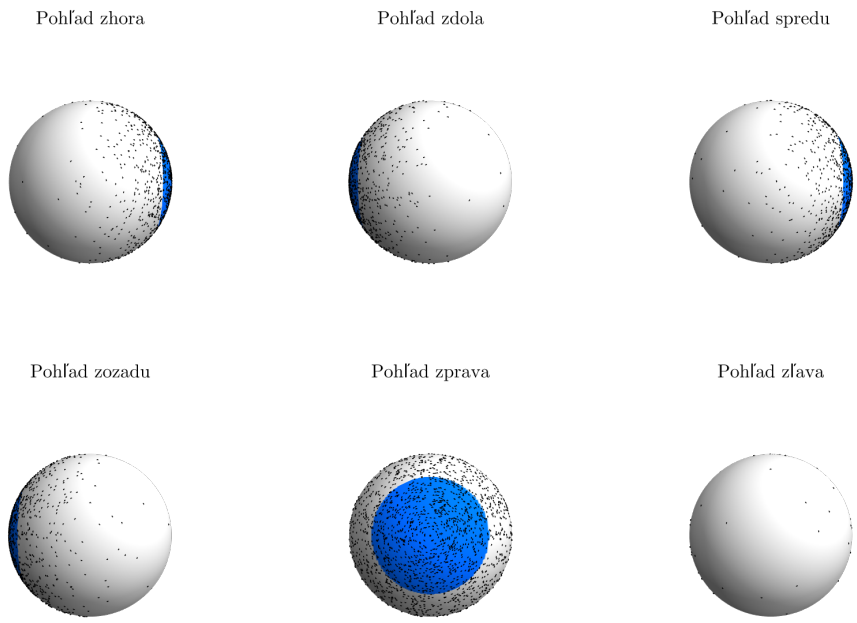
$$c_\alpha^-(R) = \{\boldsymbol{x} \in \mathbb{S}^{k-1} \colon \boldsymbol{x}^\top \boldsymbol{m} \leq c_\alpha\},$$
$$c_\alpha^+(R) = \{\boldsymbol{x} \in \mathbb{S}^{k-1} \colon \boldsymbol{x}^\top \boldsymbol{m} \geq c_\alpha\},$$

pre $\alpha \in (0,1)$.

Definícia 2.7 (Hladinový α-kvantil)
*Nech R je pravdepodobnostné rozdelenie na k-sfére, c_α je projekčný α-kvantil rozdelenia R a $\boldsymbol{m}$ je jeho unikátny medián. Potom definujeme **hladinový $\boldsymbol{\alpha}$-kvantil rozdelenia** R ako*

$$k_\alpha(R) = c_\alpha\, \boldsymbol{m}; \quad \textit{pre } \alpha \in (0,1).$$

Takto definované kvantily ležia na úsečke danej predpisom $c\,\boldsymbol{m}$; $c \in [-1,1]$. Tieto kvantily netvoria usporiadanie smerových dát, každý kvantil k_α ale jednoznačne určuje nadrovinu kolmú na $\boldsymbol{m}$. Označme ju T_α. Táto nadrovina rozdelí jednotkovú k-sféru $\mathbb{S}^{k-1}$ na množiny c_α^- a c_α^+, teda „nad“ touto rovinou je časť jednotkovej sféry $\mathbb{S}^{k-1}$, ktorá nesie pravdepodobnosť α a „pod“ touto rovinou je časť tejto sféry, ktorá nesie pravdepodobnosť $1-\alpha$. Označme prienik $T_\alpha \cap \mathbb{S}^{k-1}$ ako C_α. Túto množinu budeme nazývať **kvantilová $\boldsymbol{\alpha}$-hladina**. Tieto kvantilové hladiny tvoria vrstvovité usporiadanie smerových dát od mediánu $\boldsymbol{m}$ až k takzvanému antimediánu $-\boldsymbol{m}$. Každú takúto hladinu tvorí $(k-1)$-sféra (respektíve dva body pre dáta na jednotkovej kružnici).



Obr. 2.1: Empirické verzie spodného (znázorneného modrou) a horného (znázorneného šedou) 0.5-kvantilového uzáveru ilustrované na realizácii náhodného výberu o rozsahu 2000 z von Mises–Fisher rozdelenia s parametrami $\mu = (1,0,0)^\top$, $\kappa = 3.5$.

Ďalej definujeme ďalšiu uhlovú hĺbku – uhlovú Mahalanobisovu hĺbku, ktorá bola, rovnako ako projekčné kvantily, prvýkrát predstavená v článku Ley a kol. (2014).

Definícia 2.8 (Uhlová Mahalanobisova hĺbka)
Nech R je spojité pravdepodobnostné rozdelenie na k-sfére a $\boldsymbol{m}$ je unikátny Fisherov medián tohto rozdelenia. Označme

$$M_{\boldsymbol{x}} = \arg\min\{\alpha \in [0,1] \colon c_\alpha(R) \geq \boldsymbol{x}^\top \boldsymbol{m}\}.$$

*Potom **uhlovú Mahalanobisovu hĺbku** bodu $\boldsymbol{x}$ definujeme ako*

$$uMH(\boldsymbol{x},R) = \frac{1}{1 + M_{\boldsymbol{x}}^{-1}}.$$

3. Vlastnosti uhlových hĺbok

Po spomínaných hĺbkových funkciách definovaných pre viacrozmerné euklidovské dáta sú často požadované nasledujúce vlastnosti:

E1 **Affínna invariancia**
Majme hĺbku H, pravdepodobnostné rozdelenie R, regulárnu maticu $B \in \mathbb{R}^{k\times k}$ a vektor $\boldsymbol{b} \in \mathbb{R}^k$. Ďalej nech $\boldsymbol{X} \sim R$ a označme $R_{B\boldsymbol{X}+\boldsymbol{b}}$ rozdelenie transformovaného vektoru $A\boldsymbol{X} + \boldsymbol{b}$, potom platí

$$H(\boldsymbol{x},R) = H(B\boldsymbol{x} + \boldsymbol{b},R_{B\boldsymbol{X}+\boldsymbol{b}}), \text{ pre všetky } \boldsymbol{x} \in \mathbb{R}^k.$$

E2 **Maximalita v centre**
Majme pravdepodobnostné rozdelenie R s „centrom“ v bode $\boldsymbol{\mu} \in \mathbb{R}^k$. Potom platí

$$H(\boldsymbol{\mu},R) = \sup_{\boldsymbol{x} \in \mathbb{R}^k} H(\boldsymbol{x},R). \tag{3.1}$$

E3 **Monotónnosť pozdĺž lúčov**
Majme pravdepodobnostné rozdelenie R a $\boldsymbol{\mu} \in \mathbb{R}^k$ bod, ktorý spĺňa (3.1). Potom platí

$$H(\boldsymbol{x},R) \leq H(\boldsymbol{\mu} + \alpha(\boldsymbol{x} - \boldsymbol{\mu}),R), \text{ pre všetky } \boldsymbol{x} \in \mathbb{R}^k \text{ a všetky } \alpha \in (0,1).$$

E4 **Nulovosť v nekonečne**
Majme pravdepodobnostné rozdelenie R. Potom platí

$$\lim_{\|\boldsymbol{x}\|\to\infty} H(\boldsymbol{x},R) = 0.$$

V článku Nagy a kol. (2022) sú formulované ekvivalenty týchto vlastností pre uhlové hĺbky. Vlastnosť E2 ostáva takmer nezmenená, vlastnosť E1 nahradí rotačná invariancia, vlastnosť E3 bude nahradená monotónnosťou pozdĺž hlavných kružníc a vlastnosť E4 bude nahradená minimalitou v antimediáne, ktorý predstavuje akýsi ekvivalent nekonečna na jednotkovej k-sfére $\mathbb{S}^{k-1}$, keďže je „najvzdialenejším“ bodom od mediánu. Formulujme takto popísané vlastnosti formálne:

S1 **Rotačná invariancia**
Majme uhlovú hĺbku uH, pravdepodobnostné rozdelenie na k-sfére R, O maticu rotácie v $\mathbb{R}^k$. Ďalej nech $\boldsymbol{X} \sim R$ a označme R_O rozdelenie transformovaného vektoru $O\boldsymbol{X}$. Potom

$$uH(\boldsymbol{x},R) = uH(O\boldsymbol{x},R_O), \text{ pre všetky } \boldsymbol{x} \in \mathbb{S}^{k-1}.$$

S2 **Maximalita v centre**
Majme pravdepodobnostné rozdelenie na k-sfére R s „centrom“ v $\boldsymbol{\mu} \in \mathbb{S}^{k-1}$. Potom platí

$$uH(\boldsymbol{\mu},R) = \sup_{\boldsymbol{x} \in \mathbb{S}^{k-1}} uH(\boldsymbol{x},R). \tag{3.2}$$

S3 **Monotónnosť pozdĺž hlavných kružníc**
Majme pravdepodobnostné rozdelenie na k-sfére R a $\boldsymbol{\mu} \in \mathbb{S}^{k-1}$ bod, pre ktorý platí (3.2). Potom

$$uH(\boldsymbol{x},R) \leq uH\left(\frac{\boldsymbol{\mu} + \alpha(\boldsymbol{x} - \boldsymbol{\mu})}{\|\boldsymbol{\mu} + \alpha(\boldsymbol{x} - \boldsymbol{\mu})\|},R\right),$$

pre všetky $\boldsymbol{x} \in \mathbb{S}^{k-1} \setminus \{-\boldsymbol{\mu}\}, \alpha \in (0,1)$.

S4 **Minimalita v anticentre**
Majme pravdepodobnostné rozdelenie na k-sfére R a $\boldsymbol{\mu} \in \mathbb{S}^{k-1}$, bod pre ktorý platí (3.2). Potom

$$uH(-\boldsymbol{\mu},R) = \inf_{\boldsymbol{x} \in \mathbb{S}^{k-1}} uH(\boldsymbol{x},R).$$

Uvažujme teraz rozdelenia, pre ktoré platí $\boldsymbol{X} \sim R \Leftrightarrow O\boldsymbol{X} \sim R$, pre všetky O matice rotácií v $\mathbb{R}^k$ také, že platí $O\boldsymbol{\mu} = \boldsymbol{\mu}$. Takéto rozdelenia budeme nazývať **rotačne symetrickými okolo $\boldsymbol{\mu}$**. Vágny pojem „centrum“ vo vlastnosti S2 budeme chápať ako bod $\boldsymbol{\mu}$, okolo ktoréhe je nejaké rozdelenie R rotačne symetrické, takéto body sú však vždy minimálne dva konkrétne $\boldsymbol{\mu}$ a $-\boldsymbol{\mu}$. Jeden z týchto bodov teda zaujme úlohu „centra“ z vlastnosti S2, druhý bude hrať úlohu „anicentra“ v zmysle vlastnosti S4. Bod S3 popisuje monotóniu na kratších oblúkoch hlavných kružníc spájajúcich body $\boldsymbol{\mu}$ a $\boldsymbol{x}$, keďže body $\boldsymbol{\mu}$ a $-\boldsymbol{\mu}$ spája nekonečne veľa rovnako dlhých oblúkov (resp. 2 v prípade $k = 2$), musíme predpokladať $\boldsymbol{x} \neq -\boldsymbol{\mu}$.

Rotačnú invarianciu uhlových hĺbok definovaných v kapitole 1 sme v spomenutej kapitole aj dokázali. Overíme teda, či uhlové hĺbky definované v kapitole 1 splňujú aj ostatné požadované vlastnosti.

Pozrime sa najprv na uhlovú simplexovú hĺbku. Na príkladoch ukážeme, že táto hĺbka nespĺňa vlastnosti S2, S3 a S4. Uvažujme bod $\boldsymbol{t}(\theta) = (\cos(\theta), \sin(\theta))^\top \in \mathbb{S}^1$ ako funkciu uhla $\theta \in (-\pi,\pi]$. Majme diskrétne pravdepodobnostné rozdelenie R_1 na jednotkovej kružnici, teda na $\mathbb{S}^1$ sústredené v štyroch bodoch

$$\begin{aligned}
R_1\left(\{\boldsymbol{t}(0)\}\right) &= \frac{1}{4}, \\
R_1\left(\{\boldsymbol{t}\left(\frac{\pi}{3}\right)\}\right) &= \frac{1}{4}, \\
R_1\left(\{\boldsymbol{t}\left(\frac{7\pi}{6}\right)\}\right) &= \frac{1}{2}, \\
R_1\left(\{\boldsymbol{t}\left(\frac{3\pi}{2}\right)\}\right) &= \frac{1}{8}.
\end{aligned}$$

Teda máme

$$uSH(\boldsymbol{t}(\theta),R_1) = \begin{cases} \frac{8}{64}, & \text{pre } \theta \in (-\pi, -\frac{5\pi}{6}), \\ \frac{48}{64}, & \text{pre } \theta = -\frac{5\pi}{6}, \\ \frac{24}{64}, & \text{pre } \theta \in (-\frac{5\pi}{6}, -\frac{\pi}{2}), \\ \frac{39}{64}, & \text{pre } \theta = -\frac{\pi}{2}, \\ \frac{22}{64}, & \text{pre } \theta \in (-\frac{\pi}{2},0), \\ \frac{30}{64}, & \text{pre } \theta = 0, \\ \frac{6}{64}, & \text{pre } \theta \in (0,\frac{\pi}{3}), \\ \frac{15}{64}, & \text{pre } \theta = \frac{\pi}{3}, \\ \frac{8}{64}, & \text{pre } \theta \in (\frac{\pi}{3},\pi)). \end{cases}$$

Jediný bod, pre ktorý platí (3.2), je bod $\boldsymbol{t}(-\frac{5\pi}{6})$ teda, ak by uhlová simplexová hĺbka spĺňala S3, musela by byť uhlová simplexová hĺbka chápaná ako funkcia θ nerastúca na intervale $(-\frac{5\pi}{6},0)$, to však neplatí. Ďalej majme diskrétne pravdepodobnostné rozdelenie R_2 na kružnici, sústredené v dvoch bodoch $R_2[\{F\}] = \frac{1}{2}$, $R_2[\{G\}] = \frac{1}{2}$, kde $F,\ G$ sú body na jednotkovej kružnici prislúchajúce postupne uhlom 0, π. Toto rozdelenie je symetrické okolo osi danej bodmi $(0,1)^\top$, $(0,-1)^\top$. Oba tieto body však majú menšiu uhlovú simplexovú hĺbku ako body F, G, teda neplatí S2. Rovnosť (3.2) platí pre body F, G navyše platí $F = -G$, teda simplexová hĺbka nemôže všeobecne spĺňať ani S4.

Skúmajme teraz uhlové hĺbky odvodené od metrík. Ukážeme, že uhlová oblúková a uhlová kosínusová hĺbka spĺňajú S3. Z vlastnosti odvodenej v poznámke 1.4 dostávame $\boldsymbol{\mu} = \arg\max uH_{d_o} \Rightarrow -\boldsymbol{\mu} = \arg\min uH_{d_o}$. Podobne pre uhlovú kosínusovú hĺbku. Navyše, ak poznáme uhlovú hĺbku všetkých bodov ľubovoľnej polsféry $A \subset \mathbb{S}^{k-1}$, vieme dopočítať vďaka spomínanému vzťahu uhlovú hĺbku všetkých bodov bodov $\boldsymbol{x} \in \mathbb{S}^{k-1} \setminus A$. Uvažujme opäť funkciu $\boldsymbol{t}(\theta)$ definovanú vyššie ďalej uvažujme diskrétne rozdelenie R_3 na $\mathbb{S}^1$, sústredené v štyroch bodoch:

$$\begin{aligned} R_3(\{\boldsymbol{t}(a)\}) &= \frac{p}{2}, \\ R_3(\{\boldsymbol{t}(\pi - a)\}) &= \frac{p}{2}, \\ R_3(\{\boldsymbol{t}(-b)\}) &= \frac{1-p}{2}, \\ R_3(\{\boldsymbol{t}(b - \pi)\}) &= \frac{1-p}{2}, \end{aligned}$$

pre nejaké $a,b \in \left[-\frac{\pi}{2},\frac{\pi}{2}\right]$ a $p \in (0,1)$. Toto rozdelenie bolo navrhnuté v článku Nagy a kol. (2022). Takto skonštruované rozdelenie je vždy symetrické okolo osi danej bodmi $(0,1)^\top$, $(0,-1)^\top$. Pozrime sa na uhlovú hĺbku bodu $\boldsymbol{t} = (t_1, t_2)^\top \in \mathbb{S}^1$ vzhľadom k rozdeleniu R_3. Z definície dostávame

$$\begin{aligned} uH_{d_o}(\boldsymbol{t},R_3) = \pi - \mathsf{E}(\arccos(\boldsymbol{t}^\top Y)) = \pi - \frac{p}{2}&(\arccos(t_1\cos(a) + t_2\sin(a)) \\ &+ \arccos(t_1\cos(\pi - a) + t_2\sin(\pi - a))) \\ &- \frac{1-p}{2}(\arccos(t_1\cos(-b) + t_2\sin(-b)) \\ &+ \arccos(t_1\cos(b - \pi) + t_2\sin(b - \pi))). \end{aligned}$$

Ďalej uvažujme opäť bod $\boldsymbol{t}(\theta) = (\cos(\theta),\ \sin(\theta))^\top \in \mathbb{S}^1$ ako funkciu uhla $\theta \in (-\pi,\pi]$. Máme teda

$$\begin{aligned} uH_{d_o}(\boldsymbol{t}(\theta),R_3) = \pi &- \frac{p}{2}(\arccos(\cos(\theta)\cos(a) + \sin(\theta)\sin(a)) \\ &+ \arccos(\cos(\theta)\cos(\pi - a) + \sin(\theta)\sin(\pi - a))) \\ &- \frac{1-p}{2}(\arccos(\cos(\theta)\cos(-b) + \sin(\theta)\sin(-b)) \\ &+ \arccos(\cos(\theta)\cos(b - \pi) + \sin(\theta)\sin(b - \pi))). \end{aligned}$$

Pomocou goniometrických identít $\cos(y)\cos(z) = \frac{\cos(y-z)+\cos(y+z)}{2}$, $\sin(y)\sin(z) = \frac{\cos(y-z)-\cos(y+z)}{2}$ a vlastností funkcie $\arccos(\cos(y))$, pre $y \in (-\pi,\ \pi]$ dostávame postupne

$$\begin{aligned} uH_{d_o}(\boldsymbol{t}(\theta),R_3) &= \pi - \frac{p}{2}\left(\arccos(\cos(-\theta + a)) + \arccos(\cos(-\theta + \pi + a))\right. \\ &\quad - \frac{1-p}{2}\left(\arccos(\cos(\theta + b) + \arccos(\cos(\theta + \pi - b))\right. \\ &= \pi - \frac{p}{2}\left(m(-\theta + a) + m(-\theta + \pi - a)\right) \\ &\quad - \frac{1-p}{2}\left(m(\theta + b) + m(\theta + \pi - b)\right), \end{aligned}$$

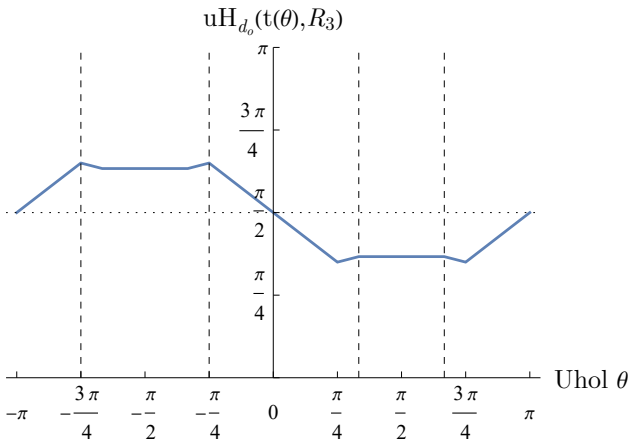
kde $m(x) = \min(|x|, 2\pi - |x|)$. Majme teraz $a = \frac{\pi}{3}$, $b = \frac{\pi}{4}$, $p = 0.2$. Pre tieto hodnoty máme

$$uH_{d_o}(\boldsymbol{t}(\theta),R_3) = \begin{cases} \frac{19}{30}\pi, & \text{pre } \theta \in [-\frac{\pi}{2}, -\frac{\pi}{3}), \\ \frac{1}{10}(7\pi + 2\theta), & \text{pre } \theta \in [-\frac{\pi}{3}, -\frac{\pi}{4}), \\ \frac{1}{10}(5\pi - 6\theta), & \text{pre } \theta \in [-\frac{\pi}{4}, -\frac{\pi}{4}), \\ \frac{1}{10}(3\pi + 2\theta), & \text{pre } \theta \in [\frac{\pi}{4}, \frac{\pi}{3}), \\ \frac{11}{30}\pi, & \text{pre } \theta \in [\frac{\pi}{3}, \frac{\pi}{2}]. \end{cases}$$

Hodnoty pre druhú polkružnicu vieme dopočítať zo vzťahu odvodeného v poznámke 1.4. Uhlová oblúková hĺbka nadobúda maxima v bodoch daných uhlami $-\frac{\pi}{4}, \frac{3\pi}{4}$, teda ani v jednom z bodov $(0,1)^\top$, $(0, -1)^\top$ maxima nenadobúda, to znamená, že uhlová oblúková hĺbka nespĺňa bod S2. Vieme, že platí

$$(\cos(a + \pi), \sin(a + \pi))^\top = -(\cos(a), \sin(a))^\top, \text{ pre } a \in \mathbb{R},$$

teda aby uhlová oblúková hĺbka spĺňala S3, musela by byť nerastúca na intervale $\left[-\frac{\pi}{4}, \frac{3\pi}{4}\right]$, to však, ako vidíme na obrázku 3.1, neplatí.

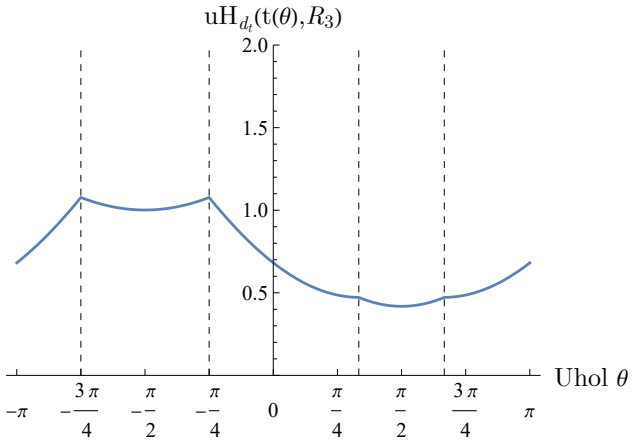


Obr. 3.1: Graf uhlovej oblúkovej hĺbky vzhľadom k rozdeleniu R_3, pre bod $\boldsymbol{t}(\theta)$ ako funkcia uhla $\theta \in (-\pi,\pi]$.

Pozrime sa ďalej na uhlovú tetivovú hĺbku. Uvažujme opäť rozdelenie R_3 s hodnotami $a = \frac{\pi}{3}$, $b = \frac{\pi}{4}$, $p = 0.2$ a bod $\boldsymbol{t}(\theta) = (\cos(\theta),\ \sin(\theta))^\top \in \mathbb{S}^1$ ako funkciu uhla $\theta \in (-\pi,\pi]$. Podobnými úpravami ako v prípade uhlovej oblúkovej hĺbky dostávame

$$\begin{aligned} uH_{d_t}(\boldsymbol{t}(\theta),R_3) = \pi &- \frac{p}{2}\left(\sqrt{2(1-\cos(\theta-a))} + \sqrt{2(1-\cos(-\theta+\pi+a))}\right) \\ &- \frac{1-p}{2}\left(\sqrt{2(1-\cos(\theta+b))} + \sqrt{2(1-\cos(\theta+\pi-b))}\right). \end{aligned}$$

Táto funkcia má opäť maximum v prípadoch $\theta = -\frac{\pi}{4}$ a $\theta = \frac{3\pi}{4}$. Teda ani uhlová tetivová hĺbka nespĺňa S2. Podobne ako v predchádzajúcom prípade, ak by pre uhlovú tetivovú hĺbku platilo S3, musela by byť nerastúca na intervale $\left[-\frac{\pi}{4},\frac{3\pi}{4}\right]$, to opäť neplatí ako vidíme na obrázku 3.2. Uhlová tetivová hĺbka navyše nespĺňa ani S4, keďže v bode $\boldsymbol{x}(\frac{3\pi}{4}) = -\boldsymbol{x}(-\frac{\pi}{4})$ táto funkcia nemá minimum.



Obr. 3.2: Graf uhlovej tetivovej hĺbky vzhľadom k rozdeleniu R_3, pre bod $\boldsymbol{t}(\theta)$ ako funkcia uhla $\theta \in (-\pi,\pi]$.

Ďalej skúmajme uhlovú kosínusovú hĺbku, pre tú už máme dokázanú platnosť vlastností S1 a S4, pozrime sa teraz na vlastnosť S3. Majme bod $\boldsymbol{\mu}$, pre

ktorý platí (3.2) potom s použitým vlastností strednej hodnoty a transponovania dostávame

$$\begin{aligned} uH_{d_k}\left(\frac{\boldsymbol{\mu}+\alpha(\boldsymbol{x}-\boldsymbol{\mu})}{\|\boldsymbol{\mu}+\alpha(\boldsymbol{x}-\boldsymbol{\mu})\|},R\right) &= 1+\mathsf{E}_R\left(\left(\frac{\boldsymbol{\mu}+\alpha(\boldsymbol{x}-\boldsymbol{\mu})}{\|\boldsymbol{\mu}+\alpha(\boldsymbol{x}-\boldsymbol{\mu})\|}\right)^\top \boldsymbol{Y}\right) \\ &= 1+k\ \mathsf{E}_R\left((\boldsymbol{\mu}+\alpha(\boldsymbol{x}-\boldsymbol{\mu}))^\top \boldsymbol{Y}\right) \\ &= 1+k\left(\mathsf{E}_R\left(\boldsymbol{\mu}^\top\boldsymbol{Y}\right)+\alpha\mathsf{E}_R\left(\boldsymbol{x}^\top\boldsymbol{Y}-\boldsymbol{\mu}^\top\boldsymbol{Y}\right)\right) \\ &= 1+k\left((1-\alpha)\mathsf{E}_R\left(\boldsymbol{\mu}^\top\boldsymbol{Y}\right)+\alpha\mathsf{E}_R\left(\boldsymbol{x}^\top\boldsymbol{Y}\right)\right), \end{aligned}$$

kde $k=\frac{1}{\|\boldsymbol{\mu}+\alpha(\boldsymbol{x}-\boldsymbol{\mu})\|}$. Navyše pomocou trojuholníkovej nerovnosti a vlastností normy vieme odvodiť nasledujúci vzťah

$$\|\boldsymbol{\mu}+\alpha(\boldsymbol{x}-\boldsymbol{\mu})\| \le \|\boldsymbol{\mu}\|+\|\alpha(\boldsymbol{x}-\boldsymbol{\mu})\| \le \|\boldsymbol{\mu}\|+\alpha\|\boldsymbol{x}\|+\alpha\|-\boldsymbol{\mu}\|.$$

Keďže máme $-\boldsymbol{\mu}$, $\boldsymbol{\mu}$, $\boldsymbol{x}\in\mathbb{S}^{k-1}$ a $\alpha\in(0,1)$ dostávame

$$\|\boldsymbol{\mu}\|+\alpha\|\boldsymbol{x}\|+\alpha\|-\boldsymbol{\mu}\| = 1+\alpha+\alpha \Rightarrow k \ge \frac{1}{1+2\alpha}.$$

Ďalej za využitia (3.2) máme

$$\begin{aligned} & uH_{d_k}\left(\frac{\boldsymbol{\mu}+\alpha(\boldsymbol{x}-\boldsymbol{\mu})}{\|\boldsymbol{\mu}+\alpha(\boldsymbol{x}-\boldsymbol{\mu})\|},R\right) - uH_{d_k}(\boldsymbol{x},R) \\ & \ge 1+\frac{1}{1+2\alpha}\left((1-\alpha)\mathsf{E}_R\left(\boldsymbol{\mu}^\top\boldsymbol{Y}\right)+\alpha\mathsf{E}_R\left(\boldsymbol{x}^\top\boldsymbol{Y}\right)\right) - uH_{d_k}(\boldsymbol{x},R) \\ & = \frac{1}{1+2\alpha}\left((1-\alpha)\mathsf{E}_R\left(\boldsymbol{\mu}^\top\boldsymbol{Y}\right)-(1-\alpha)\mathsf{E}_R\left(\boldsymbol{x}^\top\boldsymbol{Y}\right)\right) \ge 0. \end{aligned}$$

Teda uhlová kosínusová hĺbka spĺňa S3. Majme teraz rozdelenie R_4 rozdelenie rotačne symetrické okolo nejakého bodu μ. Teda platí $\boldsymbol{Y}\sim R_4 \Leftrightarrow O\boldsymbol{Y}\sim R_4$, pre také O matice rotácií v $\mathbb{R}^k$, pre ktoré platí $O\boldsymbol{\mu}=\boldsymbol{\mu}$. Navyše uhlová kosínusová hĺbka bodu $\boldsymbol{x}$ závisí na R_4 len skrz $\mathsf{E}_{R_4}(\langle\boldsymbol{x},\boldsymbol{Y}\rangle)=\langle\boldsymbol{x},\mathsf{E}_{R_4}(\boldsymbol{Y})\rangle$ a platí

$$O\mathsf{E}_{R_4}(\boldsymbol{Y})=\mathsf{E}_{R_4}(O\boldsymbol{Y})=\mathsf{E}_{R_4}(\boldsymbol{Y}),\ \text{pre všetky } O\text{, pre ktoré platí } O\boldsymbol{\mu}=\boldsymbol{\mu}.$$

Teda $\mathsf{E}_{R_4}(\boldsymbol{Y})$ leží na úsečke spájajúcej body $-\boldsymbol{\mu}$ a $\boldsymbol{\mu}$, to znamená, že $\mathsf{E}_{R_4}(\boldsymbol{Y})=c\ \boldsymbol{\mu}$, pre nejaké $c\in[-1,1]$. Z toho vyplýva, že uhlová kosínusová hĺbka musí nadobúdať maxima práve v jednom z bodov $-\boldsymbol{\mu}$, $\boldsymbol{\mu}$. Teda uhlová kosínusová hĺbka spĺňa aj S2. To, že uhlová Mahalanobisova hĺbka spĺňa všetky z bodov S1-S4 dokázali už autori článku Ley a kol. (2014), v článku Pandolfo a kol. (2018) je spomenuté, že táto hĺbka je vo vzťahu jedna k jednej s uhlovou kosínusovou hĺbkou, teda ak sme dokázali spĺňanie vlastností S1-S4 pre uhlovú kosínusovú hĺbku vieme, že aj uhlová Mahalanobisova hĺbka musí spĺňať tieto vlastnosti. Uhlová polpriestorová hĺbka taktiež spĺňa všetky nami požadované vlastnosti ako bolo dokázané v článku Nagy a Laketa (2022).

Výsledky nakoniec zhrnieme v tabuľke, kde symbol ✓ značí, že daná uhlová hĺbka spĺňa požadovanú vlastnosť a symbol ✗ znamená naopak, že danú vlastnosť uhlová hĺbka nespĺňa.

Uhlová hĺbka		S1	S2	S3	S4
Uhlová polpriestorová hĺbka	uPH	✓	✓	✓	✓
Uhlová simplexová hĺba	uSH	✓	✗	✗	✗
Uhlová oblúková hĺbka	uH_{d_o}	✓	✗	✗	✓
Uhlová tetivová hĺbka	uH_{d_t}	✓	✗	✗	✗
Uhlová kosínusová hĺbka	uH_{d_k}	✓	✓	✓	✓
Uhlová Mahalanobisova hĺbka	uMH	✓	✓	✓	✓

Tabuľka 3.1: Zhrnutie spĺňania vlastností S1-S4.

4. Definície empirických ekvivalentov

V tejto kapitole postupne zavedieme definície empirických verzií uhlových hĺbok, mediánov a projekčných kvantilov. Ako prvé definujeme empirické ekvivalenty uhlových hĺbok zavedených v prvej kapitole, tie, ako aj ich teoretické ekvivalenty, budú rotačne invariantné. Pre túto kapitolu predpokladajme nasledovné: Nech R je pravdepodobnostné rozdelenie na k-sfére a $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n$ je náhodný výber z tohto rozdelenia. Označme $\mathbb{X} = (\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n)^\top$ a $\widehat{P}_{\mathbb{X}}$ empirickú pravdepodobnostnú mieru príslušnú náhodnému výberu $\mathbb{X}$ definovanú nasledovne:

$$\widehat{P}_{\mathbb{X}}(B) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}\left(\boldsymbol{X}_i \in B\right), \text{ pre borelovskú množinu } B.$$

V definíciách z kapitol 1 a 2 sa vyskytujú len stredné hodnoty transformácií náhodných veličín na k-sfére, ktorých výsledkom sú jednorozmerné náhodné veličiny. Za ich empirickú ekvivalentu teda môžeme považovať výberový priemer rovnakej transformácie realizácií náhodného výberu. Ak v tejto kapitole nedefinujeme empirickú verziu inak, považujeme za empirickú verziu funkcií definovaných v kapitolách 1 a 2 funkcie, v ktorých pravdepodobnostnú mieru, strednú hodnotu, respektíve už definovanú funkciu nahradíme jej empirickou verziou.

Definícia 4.1 (Empirická uhlová simplexová hĺbka)
Nech R je pravdepodobnostné rozdelenie na k-sfére a $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n$; $n \geq k$ je náhodným výberom z tohto rozdelenia, ďalej označme $\mathbb{X} = (\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n)^\top$. Potom definujeme ***empirickú uhlovú simplexovú hĺbku*** *vzhľadom k $\mathbb{X}$ ako*

$$uSH_n(\boldsymbol{x}, \mathbb{X}) = \binom{n}{k}^{-1} \sum \mathbb{1}\left(\boldsymbol{x} \in \Delta^k(\boldsymbol{X}_{i_1}, \ldots, \boldsymbol{X}_{i_k})\right),$$

kde $\Delta^k(\boldsymbol{X}_{i_1}, \ldots, \boldsymbol{X}_{i_k})$ značí simplex na jednotkovej k-sfére $\mathbb{S}^{k-1}$ určený bodmi $\boldsymbol{X}_{i_1}, \ldots, \boldsymbol{X}_{i_k}$ a sumu berieme cez všetky možné k-tice $\boldsymbol{X}_{i_1}, \ldots, \boldsymbol{X}_{i_k}$, $i_j \neq i_l$, pre $l \neq j$ tvoriace simplexy $\Delta^k(\boldsymbol{X}_{i_1}, \ldots, \boldsymbol{X}_{i_k})$.

5. Krabicový diagram pre smerové dáta

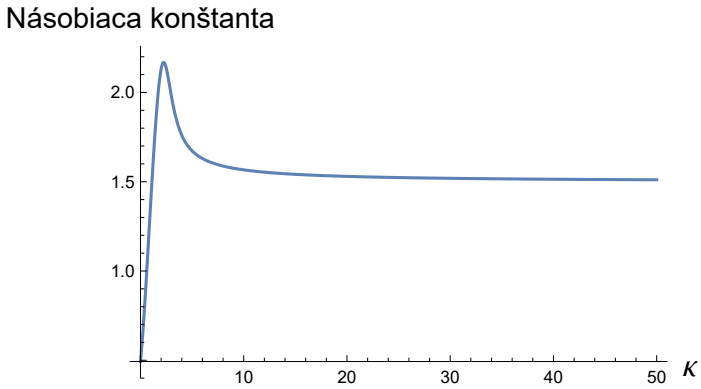
Krabicový diagram je jedným z najpoužívanejších nástrojov na grafickú analýzu, je teda rozumné snažiť sa nájsť jeho ekvivalent pre smerové dáta. Jeden z návrhov priniesli autori v článku Buttarazzi a kol. (2018). V spomenutom článku je popísaný postup pre vytvorenie analógie krabicového diagramu pre cirkulárne dáta. V tejto kapitole tento postup popíšeme a porovnáme ho s prístupom založeným na projekčných kvantiloch, ktorého výsledkom bude istá analógia krabicového diagramu.

5.1 Cirkulárny krabicový diagram

V tejto podkapitole popíšeme postup zavedený v článku Buttarazzi a kol. (2018). Prvým krokom je nájdenie Fisherovho mediánu a k nemu prislúchajúceho antimediánu. Následne očíslujeme body tak, ako na nich narazíme pri pohybe po smere hodinových ručičiek začínajúc v antimediáne. Zopakujeme predchádzajúci krok, ale pohyb po smere nahradíme pohybom proti smeru hodinových ručičiek. V ďalšom kroku priradíme každému bodu minimum z predchádzajúcich dvoch očíslovaní. Takto priradené číslo nazveme **hĺbkou** bodu. Budeme uvažovať aj polovičnú hĺbku a to nasledovným spôsobom. Bod, ktorý bude mať hĺbku $m + 0.5$, $m \in \mathbb{N}$ definujeme ako „priemer“ bodov s hĺbkou m a $m + 1$. Na kruhu budeme ako „priemer“ bodov K, L, chápať bod, ktorý sa nachádza presne v polovici kratšieho z oblúkov spájajúcich body K a L. Majme n pozorovaní, potom stred boxu bude tvoriť bod s hĺbkou $\lfloor (1+n)/2 \rfloor$. Okraje boxu budú tvorené bodmi s hĺbkou $\frac{\lfloor (1+n)/2 \rfloor}{2}$. „Fúzy“ krabicového diagramu v jednorozmernom prípade získavame vynásobením šírky boxu istou násobiacou konštantou. Štandardná voľba násobiacej konštanty je 1.5, za predpokladu normality pri tejto voľbe konštanty môžeme očakávať, že približne 0.7 % pozorovaní spadne mimo „fúzov“. Podobným spôsobom by sme chceli dostať aj „fúzy“ cirkulárneho boxplotu, teda chceme, aby za istého referenčného rozdelenia spadlo mimo „fúzov“ 0.7 % pozorovaní. Kvantily normálneho rozdelenia sú ekvivariantné voči posunutiu a násobeniu kladnou konštantou. Oproti tomu, referenčné rozdelenie na kruhu, teda von Mises rozdelenie túto vlastnosť nemá, toto rozdelenie je len rotačne ekvivariantné. Z toho plynie, že pri výpočte násobiacej konštanty musíme istým spôsobom zohľadniť parameter koncentrácie von Mises rozdelenia κ. Násobiaca konštanta NK bude, za predpokladu, že γ % pozorovaní spadne mimo „fúzov“, funkciou parametru κ, ktorá vyzerá nasledovne

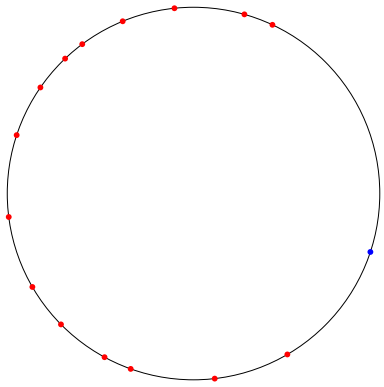
$$NK_\gamma(\kappa) = \frac{Q_{1-\gamma/200}(\kappa) - Q_{0.75}(\kappa)}{Q_{0.75}(\kappa) - Q_{0.25}(\kappa)} = \frac{Q_{0.25}(0,\kappa) - Q_{\gamma/200}(\kappa)}{Q_{0.75}(\kappa) - Q_{0.25}(0,\kappa)},$$

kde $Q_\alpha(\kappa)$ značí α-kvantil von Mises rozdelenia s parametrami $\mu = 0$ a κ. Graf funkcie $Nk_\gamma(\kappa)$ pre $\gamma = 0.7$ je na obrázku 5.1.

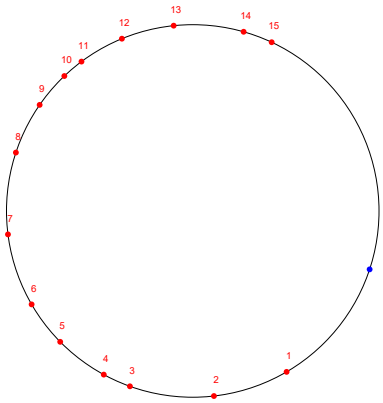


Obr. 5.1: Graf závislosti násobiacej konštanty na parametre koncentrácie von Mises rozdelenia κ, za predpokladu, že požadujeme, aby 0.7 % pozorovaní spadlo mimo „fúzov“.

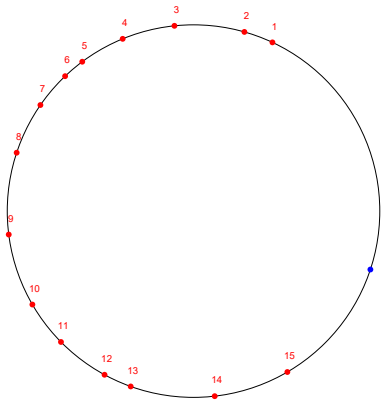
Autori článku, v ktorom bol tento postup prvýkrát predstavený, používajú pre výpočet násobiacej konštanty odhad parametru κ, ktorý získali netriviálnym spôsobom popísaným v prílohe „Web Appendix A“ článku Buttarazzi a kol. (2018). My budeme používať maximálne vierohodný odhad tohto parametru. Posledným krokom je vykreslenie odľahlých pozorovaní, teda pozorovaní, ktoré spadnú mimo „fúzov“. Popísaný postup je ilustrovaný na obrázku 5.2. Na ilustráciu použijeme realizáciu náhodného výberu z von Mises rozdelenia s parametrami $\mu = \pi,\ \kappa = \frac{1}{2}$ o rozsahu 15.



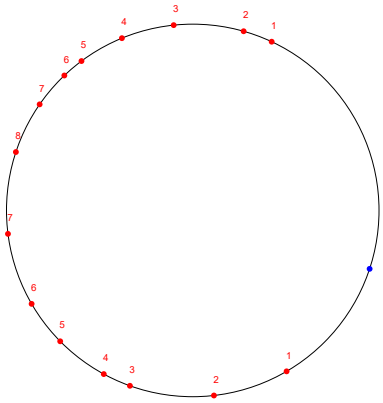
(a) Vykreslenie pozorovaní na kružnicu



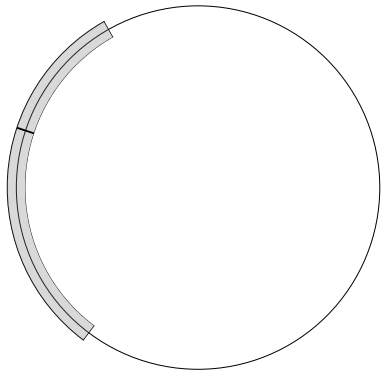
(b) Očíslovanie po smere hodinových ručičiek



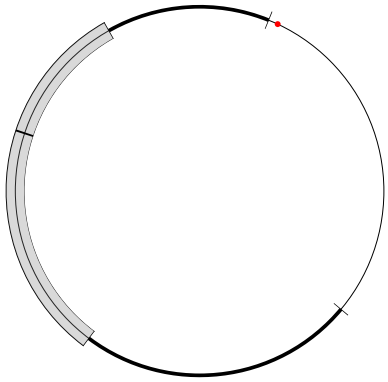
(c) Očíslovanie proti smeru hodinových ručičiek



(d) Priradenie hĺbky



(e) Vykreslenie boxu



(f) Vykreslenie „fúzov“ a odľahlých pozorovaní

Obr. 5.2: Zobrazenie krokov zostavenia krabicového diagramu pre smerové dáta, realizácie náhodného výberu zobrazené červenou, antimedián modrou.

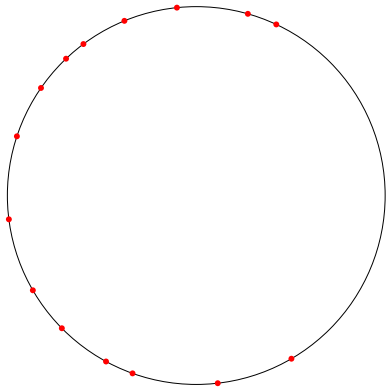
5.2 Analógia krabicového diagramu pomocou empirických projekčných kvantilov

V kapitole 2 sme zaviedli projekčné kvantily a pomocou nich definovali horný a dolný kvantilový uzáver. Pomocou horných kvantilových uzáverov vieme jedno-

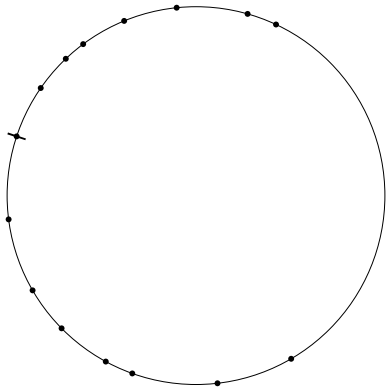
ducho vytvoriť postup, ktorého výsledkom bude istá analógia cirkulárneho krabicového diagramu predstaveného v predchádzajúcej podkapitole. Rovnako ako v predchádzajúcom prípade je prvým krokom nájdenie Fisherovho mediánu. Pomocou mediánu získame projekčné kvantily. Po vypočítaní projekčných kvantilov vieme prejsť priamo k vykresleniu boxu. Box bude v tomto prípade tvorený množinou $c^+_{0.5}(\mathbb{X})$, teda empirickým horným 0.5-kvantilovým uzáverom náhodného výberu $\mathbb{X}$, ktorého teoretická verzia bola definované v definícii 2.6. „Fúzy“ krabicového diagramu bude potom tvoriť množina $c^+_{0.07}(\mathbb{X}) \setminus c^+_{0.5}(\mathbb{X})$. Postup skrátene zopakujeme v bodoch:

1. Nájdenie mediánu.
2. Výpočet množiny $c^+_{0.5}(\mathbb{X})$ a vykreslenie boxu.
3. Výpočet množiny $c^+_{0.07}(\mathbb{X}) \setminus c^+_{0.5}(\mathbb{X})$ a vykreslenie fúzov.

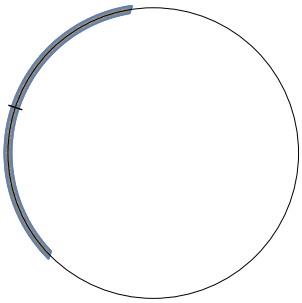
Na obrázku 5.3 ilustrujeme tento postup, opäť na ilustráciu použijeme náhodný výber z von Mises rozdelenia, ktorý bol použitý aj v predchádzajúcej podkapitole. Na obrázku 5.4 porovnáme výsledky postupu popísaného v tejto podkapitole a postupu popísaného v predchádzajúcej podkapitole. Ako prílohu k bakalárskej práci prikladáme funkciu použitú na vytváranie analógie krabicového diagramu popísanej v tejto podkapitole s komentárom pre používateľov a niekoľkými príkladmi.



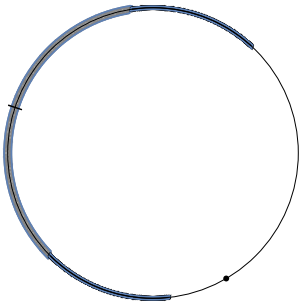
(a) Vykreslenie pozorovaní na kružnicu



(b) Označenie mediánu

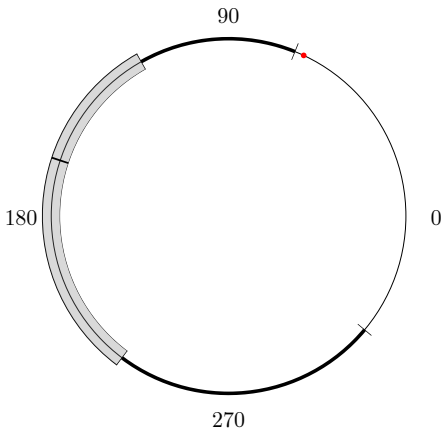


(c) Vykreslenie boxu

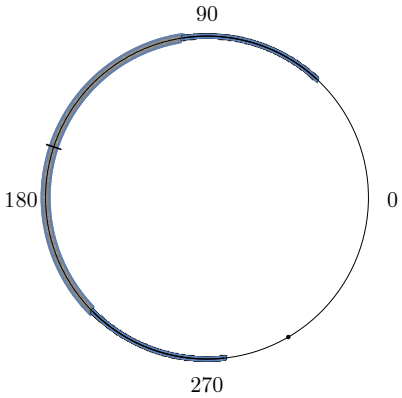


(d) Vykreslenie „fúzov“ a odľahlých pozorovaní

Obr. 5.3: Zobrazenie krokov zostavenia „krabicového diagramu“ pomocou empirických horných kvantilových uzáverov.



(a) Krabicový diagram zavedený v podkapitole 5.1

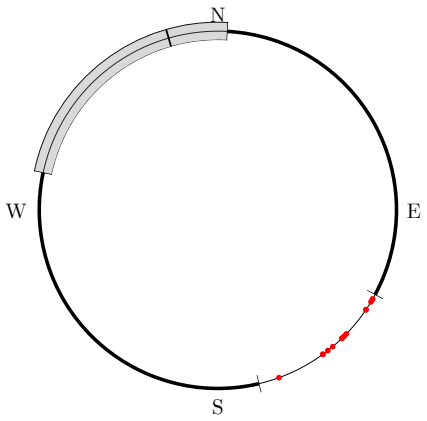


(b) Analógia krabicového diagramu predstavená v podkapitole 5.2

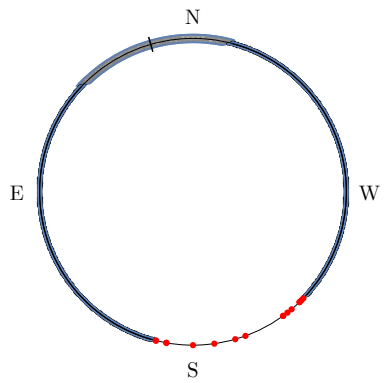
Obr. 5.4: Porovnanie krabicového diagramu zavedeného v podkapitole 5.1 a postupu zavedeného v podkapitole 5.2, oba boli aplikované na náhodný výber z von Mises rozdelenia s parametrami $\mu = \pi$ a $\kappa = 1/2$ o rozsahu 15.

5.3 Porovnanie navrhnutých prístupov

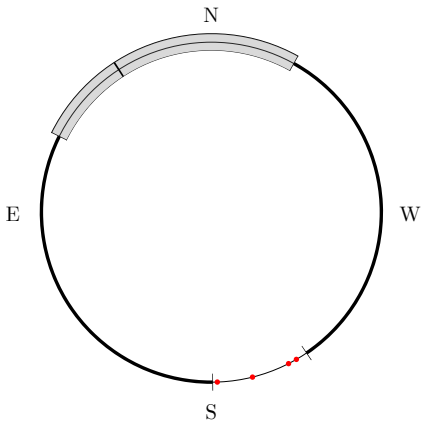
Postupy popísané v podkapitolách 5.1 a 5.2 dávajú vizuálne podobné výsledky. V oboch prípadoch je stredom krabicového diagramu medián, box znázorňuje časť kruhu, do ktorej by mala spadnúť polovica pozorovaní a mimo fúzov by malo spadnúť približne 0.7 % pozorovaní. Krabicový diagram zostavený postupom z podkapitoly 5.2 je vždy symetrický okolo mediánu, teda žiadnym spôsobom nezohľadňuje šikmosť dát (emprická šikmosť smerových dát je definovaná v Fisher (1993)). Cirkulárny krabicový diagram naopak nie je nutne symetrický okolo mediánu, teda je možné zistiť prípadné zošikmenie dát. Prístup predstavený v podkapitole 5.2 je jednoducho aplikovateľný aj na dáta vyšších dimenzií. Cirkulárny krabicový diagram, ako už naznačuje aj názov, je použiteľný len pre cirkulárne dáta a jeho zovšeobecnenie pre viacrozmerné smerové dáta by bolo netriviálne. Na obrázku 5.5 vidíme oba spôsoby aplikované na reálne dáta získané zo stránky `https://www.visualcrossing.com/`. Tieto dáta predstavujú merania smerov vetra v prvých desiatich dňoch šiesteho mesiaca roku 2023. Merania boli uskutočňované každú hodinu a zaznamenávané v stupňoch. Teda každá dátová sada predstavuje zoznam 240 pozorovaní. Môžeme si všimnúť, že v strednej Európe v tomto období prevládal vietor zo severo-východu, napriek tomu sú však smery vetra v tejto oblasti Európy rozložené relatívne rovnomerne. V Istanbule rovnako prevažoval vietor zo severo-východu, v tomto meste však vietor z iných smerov takmer nefúkal. V Madride prevládal prevažne vietor z východu, ale podobne ako v strednej Európe boli aj ostatné smery reprezentované relatívne rovnomerne.



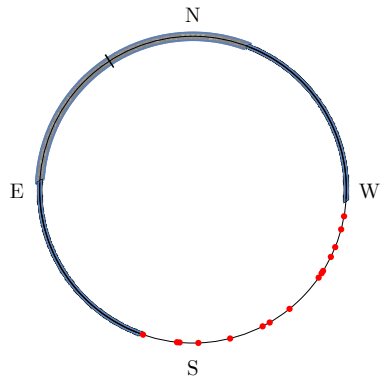
(a) Krabicový diagram pre dáta o smeroch vetra v Košiciach z prvých 10 dní 6. mesiaca roku 2023



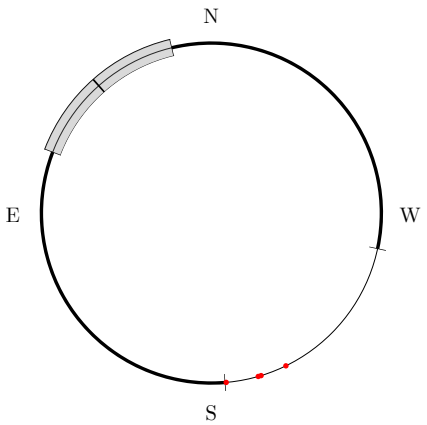
(b) Analógia krabicového diagramu pre dáta o smeroch vetra v Košiciach z prvých 10 dní 6. mesiaca roku 2023



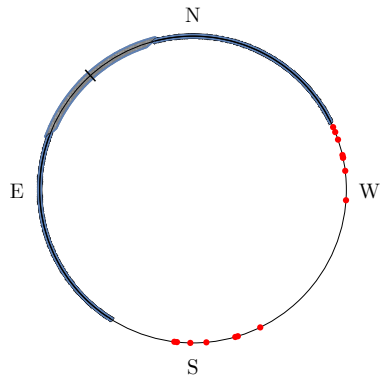
(c) Krabicový diagram pre dáta o smeroch vetra v Poprade z prvých 10 dní 6. mesiaca roku 2023



(d) Analógia krabicového diagramu pre dáta o smeroch vetra v Poprade z prvých 10 dní 6. mesiaca roku 2023

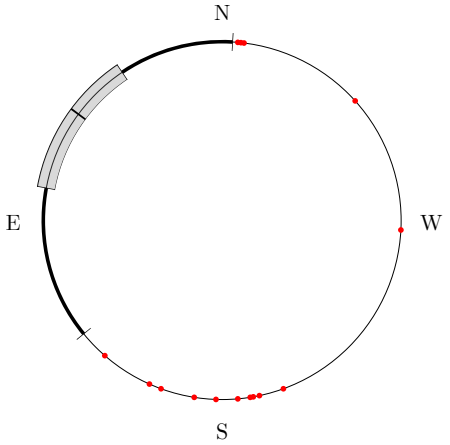


(e) Krabicový diagram pre dáta o smeroch vetra v Berlíne z prvých 10 dní 6. mesiaca roku 2023

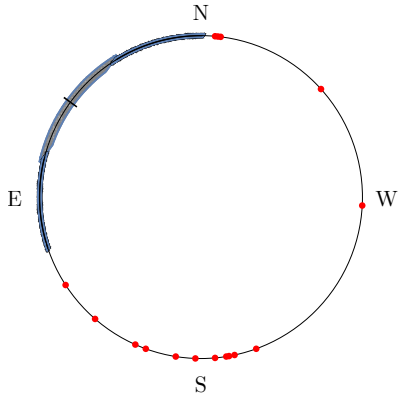


(f) Analógia krabicového diagramu pre dáta o smeroch vetra v Berlíne z prvých 10 dní 6. mesiaca roku 2023

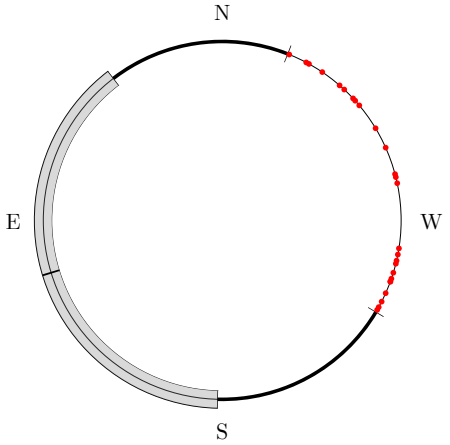
Obr. 5.5: Porovnanie prístupov popísaných v podkapitole 5.1 (naľavo) a v podkapitole 5.2 (napravo) aplikovaných na dáta o smere vetra poskytnutých spoločnosťou Visual Crossing Corporation. Visual Crossing Weather.



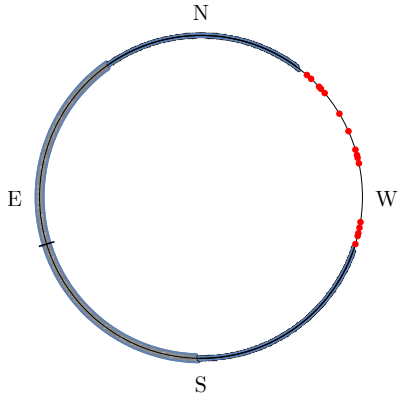
(a) Krabicový diagram pre dáta o smeroch vetra v Istanbule z prvých 10 dní 6. mesiaca roku 2023



(b) Analógia krabicového diagramu pre dáta o smeroch vetra v Istanbule z prvých 10 dní 6. mesiaca roku 2023



(c) Krabicový diagram pre dáta o smeroch vetra v Madride z prvých 10 dní 6. mesiaca roku 2023



(d) Analógia krabicového diagramu pre dáta o smeroch vetra v Madride z prvých 10 dní 6. mesiaca roku 2023

Obr. 5.6: Porovnanie prístupov popísaných v podkapitole 5.1 (naľavo) a v podkapitole 5.2 (napravo) aplikovaných na dáta o smere vetra poskytnutých spoločnosťou Visual Crossing Corporation. Visual Crossing Weather.

Záver

V práci sme zaviedli formálnu definíciu smerových dát a skúmali sme možnosti usporiadania týchto dát.

V úvode prvej kapitoly sme formálne definovali pravdepodobnostné rozdelenia na k-sfére a smerové dáta. Ďalej sme sa venovali intuitívnym prístupom k usporiadaniu smerových dát. Ďalšiu časť prvej kapitoly tvorili definície uhlových hĺbok, konkrétne to boli definície uhlovej simplexovej, polpriestorovej, oblúkovej kosínusovej a tetivovej hĺbky.

V druhej kapitole sme definovali medián pre smerové dáta. Pomocou definície mediánu sme potom definovali projekčné kvantily, uhlové kvantilové uzávery a hladinové kvantily a uhlovú Mahalanobisovu hĺbku.

V tretej kapitole sme popísali isté žiaduce vlastnosti pre uhlové hĺbky. Dokázali sme, respektíve vyvrátili platnosť vlastností pre všetky nami definované uhlové hĺbky.

V štvrtej kapitole sme popísali, ako dostaneme empirické ekvivalenty funkcií definovaných v prvej a druhej kapitole.

V poslednej kapitole sme popísali dva prístupy k vytvoreniu krabicového diagramu pre cirkulárne dáta a porovnali sme ich na reálnych dátach.

Zoznam použitej literatúry

Buttarazzi, D., Pandolfo, G. a Porzio, G. C. (2018). A boxplot for circular data. *Biometrics*, **74**(4), 1492–1501. doi:10.1111/biom.12889.

Dhillon, I. S. a Sra, S. (2003). Modeling data using directional distributions. Technical report, Citeseer.

Fisher, N. I. (1985). Spherical medians. *Journal of the Royal Statistical Society: Series B (Methodological)*, **47**(2), 342–348. doi:10.1111/j.2517-6161.1985.tb01362.x.

Fisher, N. I. (1993). *Statistical Analysis of Circular Data*. Cambridge University Press. doi:10.1017/CBO9780511564345.

Ley, C., Sabbah, C. a Verdebout, T. (2014). A new concept of quantiles for directional data and the angular mahalanobis depth. *Electronic Journal of Statistics*, **8**(1). doi:10.1214/14-ejs904.

Liu, R. Y. a Singh, K. (1992). Ordering directional data: Concepts of data depth on circles and spheres. *The Annals of Statistics*, **20**(3). doi:10.1214/aos/1176348779.

Nagy, S. a Laketa, P. (2022). Theoretical properties of angular halfspace depth. Technická správa.

Nagy, S., Demni, H., Buttarazzi, D. a Porzio, G. C. (2022). Theory of angular depth for classification of directional data. Technická správa.

Omelka, M. (2022). NMSA331 Matematická statistika 1. Poznámky k přednášce. Univerzita Karlova. `https://www2.karlin.mff.cuni.cz/~omelka/Soubory/nmsa331/ms1.pdf`. Accessed: 2023-07-18.

Pandolfo, G., Paindaveine, D. a Porzio, G. C. (2018). Distance-based depths for directional data. *Canadian Journal of Statistics*, **46**(4), 593–609. doi:10.1002/cjs.11479.

Small, C. G. (1987). Measures of centrality for multivariate and directional distributions. *Canadian Journal of Statistics*, **15**(1), 31–39. doi:10.2307/3314859.

Visual Crossing Corporation. Visual Crossing Weather. (2023). `https://www.visualcrossing.com/`.