UNIVERZA V LJUBLJANI
FAKULTETA ZA MATEMATIKO IN FIZIKO

Matematika – 1. stopnja

Luka Lodrant

O geometriji diferencirane zasebnosti

Delo diplomskega seminarja

Mentor: doc. dr. Aljoša Peperko

Ljubljana, 2019

Kazalo

O geometriji diferencirane zasebnosti

Povzetek

V delu najprej predstavimo pojem diferencirane zasebnosti, kot strogo matematično definicijo zasebnosti podatkov, ki pride do izraza pri njihovi javni objavi. Definiramo splošno okolje za numerične podatke, nato pa ocenimo spodnjo mejo napake, ki jo zaseben odzivni mehanizem mora vnesti v podatke. Predstavimo Laplaceov mehanizem, podrobneje pa še K-normni in rekurzivni NIM mehanizem. Za vse izpeljemo tudi zgornjo mejo napake in tako za K-normni ter NIM mehanizem ocenimo, da sta na določenih razredih poizvedb asimptotsko optimalna. Mehanizme implementiramo in obravnavamo težave, ki pri tem nastanejo.

On the Geometry of Differential Privacy

Abstract

In this work we present the concept of differential privacy as a rigorous mathematical definition of privacy, which is required for publishing private data. We define a general setting for numerical data and derive a lower bound for the required error of private mechanisms. Laplace mechanism, K-norm mechanism and recursive NIM mechanism are presented, each with an upper bound on its error. We conclude that NIM and K-norm mechanism are asimptotically optimal for specific classes of queries. Mechanisms are implemented and problems which arise during the implementation are addressed.

Math. Subj. Class. (2010): 62-07, 52A22

Ključne besede: diferencirana zasebnost, odzivni mehanizem, K-normni mehanizem, izotropski položaj

Keywords: differential privacy, response mechanism, K-norm mechanism, isotropic position

1. Uvod

Skupaj z razvojem tehnologij obdelave podatkov, strojnega učenja in umetne inteligence se vedno pogosteje pojavljajo tudi problemi z zasebnostjo naših podatkov. Podjetja in institucije jih zbirajo čedalje več, vprašanje pa je, kako te podatke obdelati, da bodo čim bolje opisali splošno populacijo, hkrati pa ohranili zasebnost vsakega posameznika. *Diferencirana zasebnost* nam ponudi dosledno definicijo zasebnosti, s pomočjo katere lahko statistično analiziramo podatke z zagotovilom, da napadalec, tudi če poseduje dodatne informacije, ne bo mogel iz njih izolirati podatkov o posameznem subjektu. Ideja diferencirano zasebnih mehanizmov je, da prisotnost posameznika v podatkovni bazi ne vpliva na rezultat določene poizvedbe. To se zagotovi z uvedbo naključnih napak v odgovor oz. v samo podatkovno bazo.

Do težav z zasebnostjo pride predvsem pri objavi podatkov, zato obstajajo razne metode anonimizacije podatkov. Najbolj naiven sistem je recimo preprost izbris podatkov kot so ime, priimek in EMŠO, izkaže pa se, da je že s poznavanjem malega števila parametrov (npr. prebivališče in rojstni datum) mogoče določeno osebo identificirati. Diferencirano zasebni mehanizmi so dokazljivo varni pred razkritjem zasebnih podatkov, za kar pa morajo podatke na neki način pokvariti. V tem prispevku se bomo najbolj posvetili ravno kompromisu med količino vnešene napake in zagotovljeno mero zasebnosti (večja napaka, večja zasebnost). Več informacij o diferencirani zasebnosti nasploh in njenih osnovnih lastnostih je na voljo v delu C. Dwork in A. Rotha [4].

Podlaga diplomskega dela bo članek M. Hardta in K. Talwara [5]. V njem bomo najprej pripravili matematično podlago in splošno okolje za predstavitev našega problema, nato bomo izpeljali spodnjo mejo za napako, ki jo vsak diferencirano zaseben algoritem mora uvesti. Ta meja bo pogojena glede na volumen slike krogle. Izpeljali bomo tudi mehanizem, ki pri določenih predpostavkah to mejo doseže. Vse meje bodo asimptotske v odvisnost od števila poizvedb in elementov v podatkovni bazi, kar pa pomeni, da jih v našem primeru praktično ne bo mogoče predstaviti. Na koncu bomo ta mehanizem tudi implementirali in analizirali njegove odgovore.

2. Matematična podlaga

Z B_p^d bomo označevali zaprto enotsko kroglo p-norme v $\mathbb{R}^d$. Za označevanje ℓ_p-norm bomo uporabljali $\|\cdot\|_p$, kjer bo $\|\cdot\|$ krajši zapis za klasično evklidsko normo $\|\cdot\|_2$. Če je K konveksno glede na izhodišče simetrično telo v $\mathbb{R}^d$, bomo z $\|\cdot\|_K$ označevali seminormo Minkovskega definirano z $\|x\|_K = \inf\{r > 0 : x \in rK\}$. Z $\log(n)$ bomo v delu, če ni dodatno določeno, označevali dvojiški logaritem $\log_2(n)$. Če je K omejeno konveksno telo, z $a \sim K$ povemo, da je a slučajna spremenljivka enakomerno porazdeljena po telesu K, torej s porazdelitveno gostoto $f(r) = \frac{1}{\mathrm{Vol}(K)}$ za $r \in K$.

Izrek 2.1 (neenakost Markova)**.** *Naj bo X nenegativna slučajna spremenljivka in $a > 0$, potem velja:*

$$\mathbb{P}(X \geq a) \leq \frac{\mathbb{E}(X)}{a},$$

kjer $\mathbb{E}(X)$ označuje matematično upanje slučajne spremenljivke X.

Dokaz. Naj bo $f(x)$ gostota slučajne spremenljivke X. Potem ocenimo

$$\begin{aligned}\mathbb{E}(X) &= \int_{-\infty}^{\infty} xf(x)dx \\ &= \int_{0}^{\infty} xf(x)dx && X \text{ je nenegativna} \\ &\geq \int_{a}^{\infty} xf(x)dx \\ &\geq a\int_{a}^{\infty} f(x)dx = a\,\mathbb{P}(X \geq A).\end{aligned}$$

□

Izrek 2.2 (Jensenova neenakost, [8, Izrek 7.11])**.** *Naj bo X slučajna spremenljivka v $\mathbb{R}$ in $f : \mathbb{R} \to \mathbb{R}$ konveksna funkcija. Tedaj velja*

$$f(\mathbb{E}[X]) \leq \mathbb{E}[f(x)].$$

Lema 2.3. *Naj bo $F : \mathbb{R}^n \to \mathbb{R}^d$ linearna preslikava in naj bodo $v_i \in \mathbb{R}^d$ stolpci njej pridružene matrike v standardni bazi. Potem je $K = FB_1^n$ simetrična konveksna ogrinjača točk $V = \{v_1, \ldots, v_n\}$, torej konveksna ogrinjača točk $\{\pm v_1, \ldots, \pm v_n\}$.*

Dokaz. Če razpišemo x na komponente $x = (\alpha_1, \ldots, \alpha_n)$, dobimo

$$\begin{aligned}FB_1^n &= \{Fx : x \in B_1^n\} \\ &= \{Fx : x \in \mathbb{R}^n \wedge \|x\|_1 \leq 1\} \\ &= \left\{\sum_{i=1}^{n} \alpha_i v_i : \alpha_i \in \mathbb{R} \wedge \sum_{i=1}^{n} |\alpha_i| \leq 1\right\} \\ &= \left\{\sum_{i=1}^{n} (\alpha_i - \beta_i) v_i : \alpha_i, \beta_i > 0 \wedge \sum_{i=1}^{n} \alpha_i + \beta_i \leq 1\right\} \\ &= \text{symconv}(V).\end{aligned}$$

S tem je enakost pokazana. □

Definicija 2.4 (asimptotska notacija)**.** Naj bosta $f(n)$ in $g(n)$ nenegativni funkciji. Za njiju lahko definiramo asimptotsko zgornjo, spodnjo in pa točno mejo.

- *Zgornja meja*: $f(n) \leq O(g(n))$ natanko tedaj, ko obstajata taki konstanti $C > 0$ ter $N > 0$, da za vse $n > N$ velja $f(n) \leq Cg(n)$.
- *Spodnja meja*: $f(n) \geq \Omega(g(n))$ natanko tedaj, ko obstajata taki konstanti $C > 0$ ter $N > 0$, da za vse $n > N$ velja $f(n) \geq Cg(n)$.
- *Točna meja*: $f(n) = \Theta(g(n))$ natanko tedaj, ko je $f(n) \leq O(g(n))$ in $f(n) \geq \Omega(g(n))$.

Opomba 2.5. To asimptotsko notacijo bomo uporabljali tekom celotnega dela, zato si oglejmo nekaj primerov, kako deluje. Pogosto bomo videli, da je $f = O(1)$. Če si pogledamo definicijo, hitro vidimo, da to pomeni, da je f navzgor omejena s konstanto. Najpomembnejša lastnost tega zapisa bo, da lahko ohranimo le dominantni člen. Recimo funkcija $n^2 + 10n + \log n$ je v tem zapisu kar $\Theta(n^2)$, saj ta člen dominira ostale. Uporabno je tudi dejstvo, da sta O in Ω inverzna v smislu, da je $f(n) = O(g(n)) \iff g(n) = \Omega(f(n))$. To inverzno razmerje bomo ponavadi uporabljali v neenakostih, kjer bo veljalo

$$A(n) \geq \Omega(f(n)) \iff A(n)^{-1} \leq O(f^{-1}(n)).$$

Lema 2.6. *Naj bo $U \subseteq \mathbb{R}^d$ vektorski podprostor prostora $\mathbb{R}^d$ in $P : \mathbb{R}^d \to U$ pravokotna projekcija v ta podprostor. Potem velja za vsak $x \in \mathbb{R}^d$*

$$\|Px\|_2 \leq \|x\|_2.$$

Dokaz. Ker je P pravokotna projekcija na U, za vsak $x \in \mathbb{R}^d$ velja $\langle Px, x - Px\rangle = 0$, saj je $Px \in U$ ter $Px - x \in U^\perp$. Torej zaradi aditivnosti skalarnega produkta ter Cauchy-Schwarzeve neenakosti velja

$$\|Px\|_2^2 = \langle Px, Px\rangle = \langle Px, x\rangle - \langle Px, x - Px\rangle = \langle Px, x\rangle \leq \|Px\|_2\|x\|_2. \qquad \square$$

Definicija 2.7. *Gama* funkcijo označimo z Γ in je za vse $x > 0$ definirana z

$$\Gamma(x) = \int_0^\infty t^{x-1}e^{-t}dt.$$

Znano je, da za $n \in \mathbb{N}$ velja $\Gamma(n) = (n-1)!$.

Omenimo še dve verjetnostni porazdelitvi, ki se bosta uporabljali v kasneje opisanih mehanizmih.

Definicija 2.8. *Laplaceovo porazdelitev* s parametrom lokacije $\mu > 0$ in merilom $b \in \mathbb{R}$ označujemo z Laplace(μ, b) in je za $r \in \mathbb{R}$ porazdeljena z verjetnostno gostoto

$$f(r) = \frac{1}{2b}\exp\left(-\frac{|r-\mu|}{b}\right).$$

Za oceni zgornje meje napake bomo potrebovali momente te porazdelitve, za $\mu = 0$ nam jih poda sledeča trditev.

Trditev 2.9. *Naj bo $r \sim$ Laplace$(0, b)$. Potem za $m \in 2\mathbb{N}$ velja*

$$\mathbb{E}\left[r^m\right] = b^m\,\Gamma(m+1).$$

Dokaz.

$$\begin{aligned}
\mathbb{E}\left[r^m\right] &= \frac{1}{2b}\int_{-\infty}^{\infty} r^m \exp\left(-\frac{|r|}{b}\right)dr \\
&= \frac{2}{2b}\int_0^\infty r^m \exp\left(-\frac{r}{b}\right)dr \\
&= \frac{1}{b}\int_0^\infty (bz)^m \exp\left(-z\right)b\,dz \qquad\qquad \text{substitucija } r = bz \\
&= b^m\int_0^\infty z^m \exp(-z)dz \\
&= b^m\,\Gamma(m+1) \qquad\qquad \square
\end{aligned}$$

Definicija 2.10. *Gama porazdelitev* s parametrom oblike $k > 0$ in merila $\theta > 0$ označujemo z Gamma(k, θ) in je za $r > 0$ porazdeljena z verjetnostno gostoto

$$f(r) = \frac{1}{\Gamma(k)\theta^k}\; r^{k-1}\exp(-r/\theta).$$

Trditev 2.11. *Naj bo $r \sim$ Gamma(k, θ). Potem za $m \in \mathbb{N}$ velja*

$$\mathbb{E}\left[r^m\right] = \frac{\theta^m\,\Gamma(k+m)}{\Gamma(k)}.$$

Dokaz.

$$\begin{aligned}
\mathbb{E}\left[r^m\right] &= \frac{1}{\Gamma(k)\theta^k}\int_0^\infty r^{k-1+m}\exp(-r/\theta)dr \\
&= \frac{1}{\Gamma(k)\theta^k}\int_0^\infty (\theta z)^{k+m-1}\exp(-z)\theta dz \qquad \text{substitucija } r=\theta z \\
&= \frac{\theta^{k+m}}{\Gamma(k)\theta^k}\int_0^\infty z^{k+m-1}\exp(-z)dz \\
&= \frac{\Gamma(k+m)\theta^{k+m}}{\Gamma(k)\theta^k} = \frac{\Gamma(k+m)\theta^m}{\Gamma(k)}.
\end{aligned}$$

□

3. Priprava splošnega okolja

Če želimo dosledno analizirati zasebnost podatkov, moramo najprej postaviti okolje, v katerem bomo lahko to počeli. Kljub temu, da je v splošnem diferencirano zasebnost mogoče definirati na kakršni koli vrsti podatkov (glej [6]), se bomo v našem delu omejili na numerične podatke iz množice realnih števil.

Podatkovno bazo bomo predstavili kot vektor $x \in \mathbb{R}^n$, *poizvedbo* na taki podatkovni bazi pa z linearno kombinacijo členov x. Natančneje lahko d poizvedb združimo v linearno preslikavo $F : \mathbb{R}^n \to \mathbb{R}^d$. V tem delu se bomo omejili na take matrike, ki imajo vse elemente v intervalu $[-1, 1]$, torej $F \in \mathbb{R}^{d\times n}$. Poleg tega bomo zahtevali še $d \leq n$, kar je v dejanski uporabi naravna predpostavka.

Primer 3.1. Osnovni primer take podatkovne baze je *histogram*. Recimo, da štejemo, koliko ljudi se ukvarja z določenim športom. Taka podatkovna baza bo oblike $x = \{\text{nogomet} : 30, \text{odbojka} : 15, \text{plezanje} : 20, \text{tek} : 3\}$ oz. bolj formalno $x = (30, 15, 20, 3)^T$. V tem primeru bomo s poizvedbo $F = [1, 1, 0, 0]$ ugotovili, koliko ljudi prisega na športe z žogo. ♢

Odzivni mehanizem bo v tem primeru slučajni algoritem, ki kot vhod vzame podatkovno bazo $x \in \mathbb{R}^n$ ter poizvedbo $F : \mathbb{R}^n \to \mathbb{R}^d$, vrne pa rezultat v obliki $a \in \mathbb{R}^d$. Tak mehanizem lahko analitik uporablja za izvajanje analiz na podatkovni bazi x. Neformalno bi tak mehanizem bil diferencirano zaseben, če bi se za dve dovolj podobni podatkovni bazi porazdelitev odgovorov razlikovala za multiplikativen faktor največ $\exp(\varepsilon)$. Tu je ε parameter, ki pove, kako močno zaseben je obravnavani mehanizem. Manjši ε pomeni višjo zasebnost. *Napaka* takega algoritma je pričakovana evklidska razdalja med pravilnim odgovorom Fx in dejanskim odgovorom mehanizma.

Omenjeni slučajen algoritem je tak algoritem, ki v svojem delovanju uporabi stopnjo naključnosti, ponavadi tako, da odgovore vzorči iz neke znane porazdelitve. Rezultat z istimi vhodnimi podatki je zato načeloma različen vsakič, ko ta algoritem izvedemo.

3.1. Diferencirana zasebnost.

Prej smo odzivni mehanizem definirali kot naključen algoritem, kar je res v dejanski uporabi. Za teoretično analizo pa ga lahko opišemo samo kot porazdelitev odgovorov za vsako možno podatkovno bazo. S pomočjo tega lahko potem tudi diferencirano zasebnost opišemo kot bližino porazdelitev odgovorov mehanizma.

Definicija 3.2. *Odzivni mehanizem* M je družina slučajnih spremenljivk $M = \{M_x : x \in \mathbb{R}^n\}$, kjer je vsak M_x porazdeljen po $\mathbb{R}^d$ in opisuje porazdelitev odgovorov mehanizma za pripadajočo podatkovno bazo.

Opomba 3.3. Slučajne spremenljivke M_x so odvisne tudi od poizvedbe F, ampak ker redko primerjamo delovanje mehanizma za dve različni poizvedbi, te ne dodamo k zapisu.

Primer 3.4. Najenostavnejši tak mehanizem je kar tisti, ki odgovori z resničnim rezultatom na poizvedbo. Če je $Fx \in \mathbb{R}^d$ ta odgovor za podatkovno bazo $x \in \mathbb{R}^n$, bo slučajna spremenljivka M_x izrojena s vso verjetnostjo koncentrirano v Fx. Torej $\mathbb{P}(M_x = Fx) = 1$. ♢

Za diferencirano zasebnost bo morala biti porazdelitev odgovorov dovolj podobna na vsaki merljivi podmnožici $\mathbb{R}^d$. Neformalno so merljive podmnožice v $\mathbb{R}^d$ takšne, ki jih lahko dobimo kot štev no unijo oz. presek odprtih in zaprtih intervalov.

Definicija 3.5. Odzivni mehanizem je *ε-diferencirano zaseben*, če za vse $x, y \in \mathbb{R}^n$, za katere je $\|x - y\|_1 \leq 1$, in vsako merljivo podmnožico $S \subseteq \mathbb{R}^n$ velja

$$\mathbb{P}(M_x \in S) \leq \exp(\varepsilon)\, \mathbb{P}(M_y \in S).$$

Opomba 3.6. V literaturi se diferencirana zasebnost pogosteje definira tudi v Hammingovi razdalji, namesto v 1-normi, kot smo jo definirali tukaj. V [5] je pokazano, da je mogoče spodnjo mejo, ki jo bomo izpeljali, dokazati tudi za to obliko diferencirane zasebnosti.

Obstaja tudi šibkejša oblika ε-diferencirane zasebnosti, ki se zaradi manjše občutljivosti na računske napake in približke pogosteje uporablja v praksi, a je tukaj ne bomo natančneje obravnavali. Razlikuje se v konstanti δ, ki pove, koliko lahko mehanizem odstopa od stroge diferencirane zasebnosti.

Definicija 3.7. Odzivni mehanizem je *δ-približno ε-diferencirano zaseben*, če za vse $x, y \in \mathbb{R}^n$, za katere je $\|x - y\|_1 \leq 1$, in vsako merljivo podmnožico $S \subseteq \mathbb{R}^n$ velja

$$\mathbb{P}(M_x \in S) \leq \exp(\varepsilon)\, \mathbb{P}(M_y \in S) + \delta.$$

Trditev 3.8. *Naj bo $\lambda \in \mathbb{N}$ in naj bo odzivni mehanizem M ε-diferencirano zaseben. Potem za vse $x, y \in \mathbb{R}^n$, za katere je $\|x - y\|_1 \leq \lambda$, in vsako merljivo podmnožico $S \subseteq \mathbb{R}^n$ velja*

$$\mathbb{P}(M_x \in S) \leq \exp(\varepsilon\lambda)\, \mathbb{P}(M_y \in S).$$

Dokaz. Z x_i za $1 \leq i \leq \lambda$ označimo točko $x + \frac{i(x-y)}{\lambda}$. Za vsaki x_i in x_{i+1} zato velja

$$\begin{aligned} \|x_i - x_{i+1}\|_1 &= \left\| x + \frac{i(x-y)}{\lambda} - x - \frac{(i+1)(x-y)}{\lambda} \right\|_1 \\ &= \left\| \frac{x-y}{\lambda} \right\|_1 = \frac{\|x-y\|_1}{\lambda} \leq 1. \end{aligned}$$

Točke x_i sedaj predstavljajo pot med x in y s koraki velikosti manj od 1. Zato zahtevana enakost sledi iz zaporedne uporabe definicije diferencirane zasebnosti mehanizma M. □

Za obravnavo diferencirano zasebnih mehanizmov sta pomembna tudi pojma napake in občutljivosti.

Trditev 3.9. *V tem delu bomo obravnavali samo poizvedbe v obliki linearne preslikave $F : \mathbb{R}^n \to \mathbb{R}^d$, katerih pridružena matrika v standardni bazi bo imele vse*

elemente v $[-1, 1]$. *Za njih bo veljala Lipschitzeva lastnost glede na normo* $\|\cdot\|_1$, *torej*

$$\sup_{x \in B_1^n} \|Fx\|_1 \leq d.$$

d bomo imenovali tudi občutljivost matrike.

Dokaz. Pišimo $x = (\alpha_1, \ldots, \alpha_n) \in B_1^n$ in $F = (f_{ij})$. Potem za vsako komponento $Fx = (\beta_1, \ldots, \beta_d)$ velja

$$|\beta_j| = \left| \sum_{i=1}^{n} \alpha_i f_{ij} \right| \leq \sum_i |\alpha_i f_{ij}| \leq \sum_i |\alpha_i| = \|x\|_1 \leq 1.$$

Torej velja

$$\|Fx\|_1 = \sum_{i=1}^{d} |\beta_j| \leq d. \qquad \square$$

Definicija 3.10. *Napako* mehanizma M za poizvedbo F po normi ℓ definiramo kot $\mathrm{err}_\ell(M, F) = \sup_{x \in \mathbb{R}^n} \mathbb{E}\|M_x - Fx\|_\ell$. Tu je kot prej $F : \mathbb{R}^n \to \mathbb{R}^d$ linearna preslikava, ki opisuje poizvedbo. Če ni drugače označeno, se za normo $\|\cdot\|_\ell$ vzame evklidska norma $\|\cdot\|_2$.

Za diferencirano zasebnost velja tudi sledeča lastnost, ki nam omogoča podatke po uporabi zasebnega mehanizma poljubno obdelovati, ne da bi zmanjšali zahtevano zasebnost.

Izrek 3.11 (neobčutljivost na obdelavo)**.** *Naj bo* $M : \mathbb{R}^n \to \mathbb{R}^d$ ε*-diferencirano zaseben mehanizem za* $\varepsilon > 0$ *in naj bo* $f : \mathbb{R}^d \to \mathbb{R}^s$ *poljubna merljiva preslikava. Potem je* $f \circ M$ *prav tako* ε*-diferencirano zaseben.*

Dokaz. Naj bosta x, y podatkovni bazi, za kateri velja $\|x - y\|_1 \leq 1$ in naj bo $S \subseteq \mathbb{R}^d$ poljubna merljiva podmnožica $\mathbb{R}^d$. Naj bo T praslika množice S, torej $T = \{r \in \mathbb{R}^d : f(r) \in S\}$. Potem velja

$$\begin{aligned} \mathbb{P}[f(M_x) \in S] &= \mathbb{P}[M_x \in T] \\ &\leq \exp(\varepsilon)\, \mathbb{P}[M_y \in T] \\ &= \exp(\varepsilon)\, \mathbb{P}[f(M_y) \in S]. \end{aligned} \qquad \square$$

Eden izmed ciljev našega dela bo, da ocenimo napako mehanizmov glede na zahtevano diferencirano zasebnost. Za to bomo izpeljali spodnjo mejo napake in za vse opisane mehanizme poskusili izpeljati tudi zgornjo mejo za kar se da splošne poizvedbe. V tabeli 1 jih predstavimo, v prihajajočih poglavjih pa jih bomo tudi dokazali.

Mehanizem	ℓ_2**-napaka**
Laplaceov	$O(\varepsilon^{-1} d\sqrt{d})$
K-normni	$O(\varepsilon^{-1} d\sqrt{\log(n/d)})$
spodnja meja	$\Omega(\varepsilon^{-1} d) \min\{\sqrt{\log(n/d)}, \sqrt{d}\}$

Tabela 1. Meje za napake mehanizmov.

4. Spodnja meja

4.1. **Spodnja meja prek ocene volumna.** Naša prva naloga bo, da navzdol ocenimo napako, ki jo mora v odgovor vnesti vsak diferencirano zaseben mehanizem. Ta bo seveda odvisna tudi od poizvedbe F, bolj natančno, odvisna bo od volumna konveksnega telesa $K = FB_1^n$.

Za dokaz izreka o spodnji meji bomo potrebovali spodnjo geometrijsko lemo.

Definicija 4.1. Množica točk $Y \subseteq \mathbb{R}^d$ se imenuje *r-pakiranje* , če je $\|y - y'\|_2 \geq r$ za vsak $y, y' \in Y, y \neq y'$.

Lema 4.2 ([5, Trditev 3.2])**.** *Naj bo $K \subseteq \mathbb{R}^d$ konveksno telo in $R = \mathrm{Vol}(K)^{1/d}$. Potem K vsebuje podmnožico Y z vsaj $\exp(d)$ elementi, ki je $\Omega(R\sqrt{d})$-pakiranje.*

Izrek 4.3. *Naj bo $\varepsilon > 0$, $F : \mathbb{R}^n \to \mathbb{R}^d$ linearna preslikava in $K = FB_1^n$. Potem ima vsak ε-zaseben mehanizem M napako vsaj $\mathrm{err}(M, F) \geq \Omega(\varepsilon^{-1} d\sqrt{d} \cdot \mathrm{Vol}(K)^{1/d})$.*

Dokaz. Naj bo $\lambda \geq 1$ in $R = \mathrm{Vol}(K)^{1/d}$. Iz leme 4.2 sledi, da $\lambda K = \lambda FB_1^n$ vsebuje $\Omega(\lambda R\sqrt{d})$-pakiranje Y z vsaj $\exp(d)$ elementi. Naj bo $X \subseteq \mathbb{R}^n$ množica poljubnih praslik $y \in Y$, tako da je $FX = Y$ in imata množici X in Y enako število elementov. Iz linearnosti F sledi $\lambda FB_1^n = F(\lambda B_1^n)$, torej jih lahko izberemo tako, da za vse praslike $x \in X$ velja $||x||_1 \leq \lambda$.

Od tod naprej bomo dokazovali z uporabo protislovja. Naj bo $M = \{M_x : x \in \mathbb{R}^n\}$ ε-diferencirano zaseben mehanizem z napako $\mathrm{err}(M, F) = cd\sqrt{d}R/\varepsilon$, protislovje bomo izpeljali za neki dovolj majhen $c > 0$. Naj bo sedaj $\lambda = d/2\varepsilon$, z B_x pa označimo 2-kroglo s polmerom $2cd\sqrt{d}R/\varepsilon = 4c\lambda R\sqrt{d}$ in središčem v Fx. Z uporabo neenakosti Markova na slučajni spremenljivki $\|M_x - Fx\|_2$ (napaki mehanizma) za vsak $x \in X$ vidimo

$$\begin{aligned}\mathbb{P}(M_x \in B_x) &= \mathbb{P}(\|M_x - Fx\|_2 < 4c\lambda R\sqrt{d}) \\ &\geq 1 - \frac{\mathbb{E}\|M_x - Fx\|_2}{2cd\sqrt{d}R/\varepsilon} \\ &\geq 1 - \frac{cd\sqrt{d}R/\varepsilon}{2cd\sqrt{d}R/\varepsilon} = \frac{1}{2}.\end{aligned}$$

Ker je $\|x\|_1 \leq \lambda$, iz trditve 3.8 sledi

$$\mathbb{P}(M_0 \in B_x) \geq \exp(-\varepsilon\lambda)\,\mathbb{P}(M_x \in B_x) \geq \frac{1}{2}\exp(-d/2).$$

Za dovolj majhen c bodo krogle $\{B_x : x \in X\}$ disjunktne po definiciji $\Omega(\lambda R\sqrt{d})$-pakiranja, saj je $Y = FX$. Vse sestavimo skupaj in dobimo

$$1 \geq \mathbb{P}\left(M_0 \in \bigcup_{x \in X} B_x\right) = \sum_{x \in X} \mathbb{P}(M_0 \in B_x) \geq \frac{1}{2}\exp(-d/2)\exp(d),$$

kar pa je za $d > 2$ protislovje, saj je tedaj $\frac{1}{2}\exp(d/2) > 1$. Z $d \leq 2$ se nam ni potrebno ukvarjati, saj nas zanima le asimptotska meja. Iz tega protislovja sedaj sledi, da mehanizem s pričakovano napako $\mathrm{err}(M, F) = cd\sqrt{d}R/\varepsilon$ ne more obstajati. □

Od tod naprej bomo z $\mathrm{VolLB}(F, \varepsilon)$ označevali v tem izreku dokazano spodnjo mejo za napako mehanizma. Natančneje

$$\mathrm{VolLB}(F, \varepsilon) = \varepsilon^{-1} d\sqrt{d}\,\mathrm{Vol}(FB_1^n)^{1/d}.$$

Dokazali smo, da vsak ε-diferencirano zaseben mehanizem mora dodati šum velikosti vsaj $\Omega(\mathrm{VolLB}(F, \epsilon))$. Proti koncu dela bomo potrebovali tudi sledečo posledico tega izreka, ki poda spodnjo mejo v primeru, ko je K blizu nižje-dimenzionalnega podprostora in nam lahko volumen projekcije v ta prostor poda boljšo spodnjo mejo. To se zgodi v primeru, da F ni polnega ranga oz. je blizu taki matriki (očitno, saj je v tem primeru $\mathrm{Vol}(FB_1^n) = 0$).

Posledica 4.4. *Naj bo $\varepsilon > 0$, $F : \mathbb{R}^n \to \mathbb{R}^d$ linearna preslikava in $K = FB_1^n$. Poleg tega naj P označuje pravokotno projekcijo na k-dimenzionalen podprostor E prostora $\mathbb{R}^d$ za neki $1 \leq k \leq d$. Potem mora vsak ε-diferencirano zaseben mehanizem M imeti napako vsaj*

$$\mathrm{err}(M, F) \geq \Omega(\varepsilon^{-1} k\sqrt{k}\,\mathrm{Vol}_k(PK)^{1/k}).$$

Dokaz. Opazimo, da lahko odgovor diferencirano zasebnega mehanizma M projiciramo v odgovor Pa na poizvedbo PF in zaradi izreka 3.11 dobimo nov ε-diferencirano zaseben mehanizem, za katerega mora prav tako veljati izrek 4.3. Ker je P ortogonalna, zanjo velja $\|x\|_2 \geq \|Px\|_2$, zato lahko napako osnovnega mehanizma ocenimo z napako nižje-dimenzionalnega in dobimo zgornjo oceno. □

Podobno kot prej z $\mathrm{GVolLB}(F, \varepsilon)$ definirajmo najboljšo spodnjo mejo, ki jo lahko dobimo iz te posledice, torej

$$\mathrm{GVolLB}(F, \varepsilon) = \sup_{k,P} \varepsilon^{-1} k\sqrt{k}\,\mathrm{Vol}_k(PFB_1^n)^{1/k},$$

kjer je $1 \leq k \leq d$, P pa teče čez vse projekcije v k-dimenzionalne podprostore.

4.2. **Ocena volumna telesa K.** V prejšnjem delu smo analizo napake mehanizma prevedli na analizo volumna telesa $K = FB_1^n$, za boljše razumevanje pa lahko tudi tega navzgor ocenimo. Izpeljali bomo dve oceni, prva je enostavnejša in jo bomo dokazali, druga pa je delo I. Bárány in Z. Füredi ter prestavljena v [1].

Izrek 4.5. *Če je $F \in [-1, 1]^{d\times n}$, velja ocena volumna $\mathrm{Vol}(FB_1^n)^{1/d} \leq O(1)$.*

Dokaz. Lema 2.3 pravi, da je FB_1^n simetrična konveksna ogrinjača stolpcev matrike F. Za vsak stolpec v_i očitno velja $v_i \in B_\infty^d$. Ker je B_∞^d simetrično konveksno telo, velja tudi $FB_1^n \subseteq B_\infty^d$. Volumen lahko zato ocenimo z $\mathrm{Vol}(FB_1^n) \leq \mathrm{Vol}(B_\infty^d) = 2^d$. □

Izrek 4.6 ([1, poglavje 2])**.** *Če s K označimo konveksno ogrinjačo točk $v_1, \ldots, v_d$, velja ocena volumna*

$$\mathrm{Vol}(K)^{1/d} \leq O\left(\sqrt{\frac{\log(n/d)}{d}}\right).$$

Posledica 4.7. *Če je $F \in [-1, 1]^{d\times n}$ in $d \leq n/2$, velja*

$$\mathrm{Vol}(FB_1^n)^{1/d} \leq O(1) \min\left\{1, \sqrt{\frac{\log(n/d)}{d}}\right\}.$$

Dokaz. Združimo oba prejšnja izreka. Prvega uporabimo direktno, pri drugem pa na telo K gledamo kot na konveksno ogrinjačo $2d$ točk in dobimo

$$\begin{aligned}\mathrm{Vol}(FB_1^n)^{1/d} &\leq O\left(\sqrt{\frac{\log(2n/d)}{d}}\right)\\ &\leq O\left(\sqrt{\frac{\log(n/d)+\log(2)}{d}}\right)\\ &= O\left(\sqrt{\frac{\log(n/d)}{d}}\right). \qquad \text{ker je } n/d > 2 \qquad \square\end{aligned}$$

5. Mehanizmi diferencirane zasebnosti

5.1. **Eksponentni mehanizmi.** Začeli bomo tako, da opišemo širšo družino diferencirano zasebnih mehanizmov predstavljenih v [11], v katero bodo spadali vsi v preostanku dela obravnavani mehanizmi. To nam bo koristilo, saj za posamezne mehanizme ne bo potrebno dokazovati njihove zasebnosti, temveč se lahko samo skličemo na to, da so primeri eksponentnih mehanizmov.

Definicija 5.1. (eksponentni mehanizem) Za funkcijo $q : \mathbb{R}^n \times \mathbb{R}^d \to \mathbb{R}$ in podatkovno bazo $x \in \mathbb{R}^n$ definiramo odzivni mehanizem $\mathrm{EM}(\varepsilon, q) = \{M_x : x \in \mathbb{R}^n\}$ tako, da je vsaka slučajna spremenljivka M_x podana z gostoto:

$$f(a) = C^{-1}\exp(\varepsilon q(x, a)),$$

kjer je C normalizacijska konstanta.

Opomba 5.2. Da bo ta mehanizem dobro definiran, mora seveda biti integral $\int_{\mathbb{R}^d}\exp(\varepsilon q(x,a))da$ omejen za vsak $x \in \mathbb{R}^n$.

Naj bo Δq največja možna razlika funkcije $q(x, r)$ za dve sosednji podatkovni bazi, za kateri velja $\|x - y\|_1 \leq 1$, pri katerem koli r. Torej

$$\Delta q = \sup_{\substack{x,y\in\mathbb{R}^n\\ r\in\mathbb{R}^d}} |q(x, r) - q(y, r)|.$$

Izrek 5.3. *Eksponentni mehanizem* $\mathrm{EM}(\varepsilon, q)$ *je* $(2\varepsilon\Delta q)$*-diferencirano zaseben.*

Dokaz. Z uporabo porazdelitve mehanizma $\mathrm{EM}(\varepsilon, q)$ dobimo za bazi $x, y \in \mathbb{R}^n$ in vsako merljivo $S \subseteq \mathbb{R}^d$

$$\begin{aligned}\mathbb{P}(M_x \in S) &= \frac{\int_S \exp(\varepsilon q(x, r))dr}{\int_{\mathbb{R}^d}\exp(\varepsilon q(x, r))dr}\\ &\leq \frac{\int_S \exp(\varepsilon(q(y, r) + \Delta q))dr}{\int_{\mathbb{R}^d}\exp(\varepsilon(q(y, r) - \Delta q))dr}\\ &= \exp(2\varepsilon\Delta q)\ \mathbb{P}(M_y \in S). \qquad \square\end{aligned}$$

Opomba 5.4. Ta rezultat je smiseln le v primeru, ko je Δq omejen. V večini primerov je to lahko naravna predpostavka.

Opomba 5.5. Tehnično gledano lahko vsak zvezno porazdeljen diferencirano zaseben mehanizem M opišemo z nekim EM, tako da za funkcijo $q(x, r)$ vzamemo logaritem gostote slučajne spremenljivke M_x.

5.2. **Laplaceov mehanizem.** Za kasnejšo primerjavo si najprej oglejmo Laplaceov odzivni mehanizem, ki je trenutno *de facto* standard za numerične podatke. Kot prej naj bo $x \in \mathbb{R}^n$ podatkovna baza. Laplaceov mehanizem deluje tako, da resničnemu odgovoru na poizvedbo prištejemo Laplaceovo porazdeljen šum. Odgovor mehanizma je torej podan z Laplaceovo porazdelitvijo, s parametrom b odvisnim od zahtevane zasebnosti in parametrom μ enakim Fx.

Definicija 5.6. (Laplaceov mehanizem). Naj bo $F : \mathbb{R}^n \to \mathbb{R}^d$ linearna preslikava in $\varepsilon > 0$. Odzivni mehanizem $\mathrm{LM}(F, d, \varepsilon) = \{M_x : x \in \mathbb{R}^n\}$ definiramo tako, da je vsaka slučajna spremenljivka M_x definirana na $\mathbb{R}^d$ in porazdeljena z gostoto

$$f(a) = \frac{\varepsilon^d}{(2d)^d} \exp\left(-\frac{\varepsilon \|Fx - a\|_1}{d}\right).$$

Drugače gledano je to le porazdelitev, ki jo dobimo, če dejanskemu odgovoru Fx prištejemo slučajni vektor iz d neodvisnih Laplace$(0, d\varepsilon^{-1})$ porazdeljenih slučajnih spremenljivk.

Trditev 5.7. *Laplaceov mehanizem* $\mathrm{LM}(F, d, \varepsilon)$ *je* 2ε*-diferencirano zaseben.*

Dokaz. Za ta mehanizem hitro opazimo, da je primer eksponentnega mehanizma za funkcijo $q(x, r) = -\frac{\|Fx-a\|_1}{d}$, kjer je Fx dejanski odgovor na poizvedbo. Ocenimo še parameter Δq za $\|x - y\|_1 \leq 1$:

$$\Delta q = \frac{1}{d} \sup_{\substack{x,y \in \mathbb{R}^n \\ a \in \mathbb{R}}} (\|Fx - a\|_1 - \|Fy - a\|_1) \leq \frac{1}{d} \sup_{x,y \in \mathbb{R}^n} (\|Fx - Fy\|_1) \leq \frac{d}{d} = 1.$$

Najprej smo uporabili trikotniško neenakost za 1-normo, nato pa smo upoštevali, da smo za naše poizvedbe v definiciji 3.9 privzeli, da je njihova občutljivost manjša od d. Sedaj iz izreka 5.3 sledi, da je ta mehanizem 2ε-diferencirano zaseben. □

Trditev 5.8. *Naj bo* $F : \mathbb{R}^n \to \mathbb{R}^d$ *linearna preslikava in* $\varepsilon > 0$*. Potem za pričakovano napako* ε*-diferencirano zasebnega Laplaceovega mehanizma velja*

$$\mathrm{err}(LM, F) \leq O(\varepsilon^{-1} d\sqrt{d}).$$

Dokaz. Pričakovana vrednost 2-norme napake tega mehanizma je pričakovana vrednost 2-norme slučajnega vektorja $Y = (Y_1, \ldots, Y_d)$, sestavljenega iz d neodvisnih Laplaceovo porazdeljenih slučajnih spremenljivk. Z uporabo Jensenove neenakosti in trditve 2.9 dobimo

$$\mathrm{err}(\mathrm{LM}, F) = \mathbb{E}\|Y\| \leq \sqrt{\mathbb{E}\|Y\|^2} = \sqrt{\sum_{i=1}^{d} \mathbb{E}\, Y_i^2} = \sqrt{\sum_{i=1}^{d} 2(d\varepsilon^{-1})^2} \leq O(\varepsilon^{-1} d\sqrt{d}). \quad \square$$

5.3. K**-normni mehanizem.** V tem podpoglavju bomo opisali novejši diferencirano zaseben mehanizem, ki ga bomo imenovali K-normni mehanizem. Poleg Laplaceovega je tudi ta le primer eksponentnega mehanizma. Njegova bistvena prednost je, da je porazdelitev dodanega šuma direktno odvisna od uporabljene poizvedbe in lahko zato le tega doda manj.

Definicija 5.9. (K-normni mehanizem). Za linearno poizvedbo $F : \mathbb{R}^n \to \mathbb{R}^d$ in $\varepsilon > 0$ naj bo $K = FB_1^n$. Mehanizem $\mathrm{KM}(F, d, \varepsilon) = \{M_x : x \in \mathbb{R}^n\}$ definiramo tako, da je vsaka slučajna spremenljivka M_x definirana na $\mathbb{R}^d$ in porazdeljena z gostoto

$$f(a) = C^{-1} \exp(-\varepsilon \|Fx - a\|_K),$$

kjer je C normalizacijska konstanta, torej

$$C = \int_{\mathbb{R}^d} \exp(-\varepsilon \|Fx - a\|_K) da = \Gamma(d+1)\operatorname{Vol}(\varepsilon^{-1}K)$$

Do tega mehanizma lahko pridemo tudi na malo bolj praktičen način, s čimer bomo tudi dokazali zadnjo enakost v normalizacijski konstanti. Vzorec iz porazdelitve M_x lahko dobimo z naslednjim postopkom:

(1) Vzorčimo r iz Gamma$(d+1, \varepsilon^{-1})$ porazdelitve, torej iz porazdelitve z gostoto

$$f_r(z) = \frac{1}{\varepsilon^{-d}\,\Gamma(d+1)} \exp(-\varepsilon z) z^d.$$

(2) Izberemo a enakomerno iz množice $Fx + rK$.

Pogojno na r je a porazdeljena z gostoto $f_{a|r}(z|t) = \frac{1}{\operatorname{Vol}(tK)} = \frac{1}{t^d \operatorname{Vol}(K)}$ za vse $z \in (Fx + tK) \iff t \geq \|z - Fx\|_K$. Iz pogojne porazdelitve lahko izračunamo brezpogojno

$$\begin{aligned} f_a(z) &= \int_{-\infty}^{\infty} f_{a|r}(z|t) f_r(t) dt \\ &= \frac{1}{\varepsilon^{-d}\,\Gamma(d+1)} \int_{\|z-Fx\|_K}^{\infty} \frac{e^{-\varepsilon t} t^d}{t^d \operatorname{Vol}(K)} dt \\ &= \frac{e^{-\varepsilon \|z-Fx\|_K}}{\Gamma(d+1)\operatorname{Vol}(\varepsilon^{-1}K)}, \end{aligned}$$

kar se ujema z začetno definicijo. Od tod tudi razberemo lepšo obliko normalizacijske konstante. V sledečem izreku bomo dokazali, da je ta mehanizem res diferencirano zaseben, hkrati pa bomo izrazili pričakovano vrednost njegove napake.

Izrek 5.10. *Naj bo $\varepsilon > 0$, $F : \mathbb{R}^n \to \mathbb{R}^d$ linearna preslikava in $K = FB_1^n$. Potem je mehanizem $\mathrm{KM}(F, d, \varepsilon)$ (2ϵ)-diferencirano zaseben in je za vsak $p > 0$ njegova napaka omejena z $E\|Fx - M_x\|_2^p \leq \frac{\Gamma(d+1+p)}{\varepsilon^p\,\Gamma(d+1)} \mathbb{E}_{z\in K}\|z\|_2^p$. Napaka mehanizma kot definirana v 3.10 je torej največ $\frac{d+1}{\varepsilon} \mathbb{E}_{z\in K}\|z\|_2$.*

Dokaz. Sledili bomo alternativni izpeljavi K-normnega algoritma. Če je kot v prej opisanih korakih $r \sim \text{Gamma}(d+1, \varepsilon^{-1})$, lahko za vsak $x \in \mathbb{R}^n$ napako ocenimo takole:

$$\begin{aligned} \mathbb{E}\|Fx - M_x\|_2^p = \mathbb{E}\|M_0\|_2^p &= \mathop{\mathbb{E}}_{r} \mathop{\mathbb{E}}_{a\sim rK} \|a\|_2^p \\ &= \mathop{\mathbb{E}}_{r} \mathop{\mathbb{E}}_{z\sim K} r^d \|z\|_2^p \\ &= \left[\mathop{\mathbb{E}}_{r} r^p\right] \cdot \mathop{\mathbb{E}}_{z\sim K} \|z\|_2^p \\ &= \frac{\Gamma(d+1+p)}{\varepsilon^p\,\Gamma(d+1)} \mathop{\mathbb{E}}_{z\sim K} \|z\|_2^p. \end{aligned}$$

Na koncu smo r-ti moment Gama porazdelitve izrazili v skladu s trditvijo 2.11. V posebnem primeru, ko je $p = 1$, sledi $\frac{\Gamma(d+1+p)}{\Gamma(d+1)} = d + 1$.

Diferencirane zasebnosti ni potrebno dodatno dokazovati, saj je ta mehanizem poseben primer eksponentnega algoritma, kjer za našo funkcijo q izberemo $q(x, a) = -\|Fx - a\|_K$. Ocenimo še Δq z

$$\Delta q = \|Fx - a\|_K - \|Fy - a\|_K \leq \|Fx - Fy\|_K \leq 1.$$

Tu smo najprej uporabili trikotniško enakost za normo Minkowskega, nato pa opazili, da je $\|x-y\|_1 \leq 1$, torej $F(x-y) \in K = FB_1^n$ ter zato po definiciji norme Minkovskega $\|F(x-y)\|_K \leq 1$. Iz izreka 5.3 sledi, da je KM 2ε-diferencirano zaseben. □

Opomba 5.11. Kljub temu, da smo tukaj dokazali, da je ta mehanizem 2ε-diferencirano zaseben, bomo ponavadi privzeli, da je kar ε-diferencirano zaseben, saj ga lahko uporabimo z zmanjšanim ε.

6. Ujemanje mej za naključne poizvedbe

V prejšnjih delih smo dokazali spodnjo mejo za napako ε-diferencirano zasebnega mehanizma, za naš K-normni mehanizem pa tudi zgornjo mejo. V tem delu bomo pokazali, da se ti dve meji ujemata za naključne poizvedbe F v obliki Bernoullijevih matrik. Ključnega pomena bo, da na telo $K = FB_1^n$ gledamo kot na simetrično konveksno ogrinjačo n točk $\{v_1, \ldots, v_n\} \subseteq \mathbb{R}^d$, kjer je v_i i-ti stolpec matrike F.

Ogrinjače te vrste so že bile množično raziskovane v teoriji slučajnih politopov. V [9] so Litvak, Pajor, Rudelson in Tomczak-Jaegermann predstavili sledeč izrek, s pomočjo katerega bomo za to kategorijo poizvedb lepše izrazili spodnjo mejo napake.

Izrek 6.1 ([9, Izrek 4.8])**.** *Naj bodo $2d \leq n \leq 2^d$ in naj bo F slučajna $d \times n$ Bernoullijeva matrika (vsi elementi matrike so Bernoullijevo porazdeljene neodvisne slučajne spremenljivke). Potem za vsak $\beta \in (0, \frac{1}{2})$ z verjetnostjo večjo od $1-\exp(-\Omega(d^\beta n^{1-\beta}))$ velja*

$$\mathrm{Vol}(FB_1^n)^{1/d} \geq \Omega(1)\sqrt{\log(n/d)/d}.$$

To lahko združimo z oceno spodnje meje iz izreka 4.3 in dobimo sledeči izrek.

Izrek 6.2. *Naj bo $\varepsilon > 0$ in $0 < d \leq n/2$. Potem mora, za skoraj vse matrike $F \in \{-1,1\}^{d\times n}$, imeti vsak ε-diferencirano zaseben mehanizem M napako vsaj*

$$\mathrm{err}(M,F) \geq \Omega(d/\varepsilon) \cdot \min\{\sqrt{d}, \sqrt{\log(n/d)}\}.$$

6.1. Zgornja meja K-normnega mehanizma za naključne poizvedbe.

Naš cilj je pokazati, da je K-normni mehanizem skoraj optimalen za naključne poizvedbe. Spodnjo mejo splošnega mehanizma sedaj poznamo, kar pomeni, da moramo omejiti še količino $\mathbb{E}_{z\sim K}\|z\|_p$. Uporabili bomo sledeči izrek, delo Klartaga in Kozme iz [7].

Izrek 6.3 ([7, Posledica 3.1])**.** *Naj bo F slučajna Bernoullijeva matrika in $K = FB_1^n$. Potem obstaja konstanta $C > 0$, tako da z verjetnostjo večjo od $1 - Ce^{-O(n)}$ velja*

$$\frac{1}{\mathrm{Vol}(K)} \int_{z\in K} \|z\|^2 dz \leq C\log(n/d).$$

Ta izrek lahko sedaj uporabimo na zgornji meji KM mehanizma in tako za skoraj vse Bernoullijeve poizvedbe dobimo ujemajočo spodnjo in zgornjo mejo.

Posledica 6.4. *Naj bo $\varepsilon > 0$ in $0 < d \leq n/2$. Potem je za skoraj vse matrike $F \in \{-1,1\}^{d\times n}$ mehanizem $\mathrm{KM}(F,d,\varepsilon)$ ε-diferencirano zaseben z napako največ*

$$O(d/\varepsilon) \cdot \min\{\sqrt{d}, \sqrt{\log(n/d)}\}.$$

Dokaz. Izraz $\mathbb{E}_{z\sim B_\infty^d}\|z\|$ lahko ocenimo na dva načina. Najprej kot v izreku 4.2 opazimo, da je $K = FB_1^n \subseteq B_\infty^d$, zato lahko ocenimo:

$$\mathrm{err}(M,F) \leq O(d/\varepsilon) \mathop{\mathbb{E}}_{z\sim K}\|z\| \leq O(d/\varepsilon) \mathop{\mathbb{E}}_{z\sim B_\infty^d}\|z\| \leq O(d/\varepsilon)\sqrt{d}.$$

Za drugi način pa uporabimo zgoraj predstavljen izrek in Jensenovo neenakost ter tako dobimo

$$\begin{aligned}\operatorname{err}(M,F) &\leq O(d/\varepsilon) \mathop{\mathbb{E}}_{z\sim K} \|z\| \\ &\leq O(d/\varepsilon)\sqrt{\mathop{\mathbb{E}}_{z\sim K} \|z\|^2} \\ &\leq O(d/\varepsilon)\sqrt{\log(n/d)}.\end{aligned}$$

Ti dve meji lahko združimo in dobimo željeno omejitev. □

S tem smo sedaj dokazali, da je K-normni mehanizem asimptotsko gledano optimalen za skoraj vse poizvedbe $F \in \{-1,1\}^{d\times n}$, v prihodnjih poglavjih pa bomo poskusili podobno dokazati tudi za bolj splošne oblike poizvedb.

7. Meje za približno izotropska telesa

Za K-normni mehanizem se da pokazati, da je asimptotsko optimalen za vsako poizvedbo, za katero bo telo $K = FB_1^n$ v tako imenovanem *izotropskem položaju*. Podrobnejša analiza takšnih teles skupaj z dokazi uporabljenih trditev je na voljo v delu Milmana in Pajorja [12], ali pa v bolj razširjeni raziskavi Giannopoulosa [3].

Definicija 7.1 (izotropski položaj)**.** Za konveksno telo $K \subseteq \mathbb{R}^d$ rečemo, da je v *izotropskem položaju* z izotropsko konstanto L_k, če za vsak enotski vektor $v \in \mathbb{R}^d$ velja

$$\frac{1}{\operatorname{Vol}(K)} \int_K \langle z, v\rangle^2 dz = L_K^2 \operatorname{Vol}(K)^{2/d}.$$

Trditev 7.2 ([12, poglavje 1.6])**.** *Za vsako konveksno telo $K \subseteq \mathbb{R}^d$ obstaja taka do ortogonalne transformacije natančno določena linearna preslikava T, ki ohranja volumen, da je TK v izotropskem položaju.*

S pomočjo te trditve lahko za katerokoli konveksno telo K torej definiramo izotropsko konstanto L_K kot L_{TK}, kjer je T preslikava, ki prenese K v izotropski položaj. Ta konstanta je dobro definirana, saj je T enolična do ortogonalne transformacije natančno (ortogonalna preslikava ohranja skalarni produkt).

Definicija 7.3 (približno izotropski položaj)**.** Za konveksno telo $K \subseteq \mathbb{R}^d$ rečemo, da je v *c-približno izotropskem položaju*, če za vsak enotski vektor $v \in \mathbb{R}^d$ velja

$$\frac{1}{\operatorname{Vol}(K)} \int_K \langle z, v\rangle^2 dz \leq c^2 L_K^2 \operatorname{Vol}(K)^{2/d}.$$

L_K je tu definiran za splošno telo kot prej opisano.

Sedaj lahko pokažemo, da je K-normni mehanizem asimptotsko optimalen, če je K c-približno izotropski.

Izrek 7.4. *Naj bo $\varepsilon > 0$ in $F : \mathbb{R}^n \to \mathbb{R}^d$ linearna preslikava. Če je $K = FB_1^n$ v c-približno izotropskem položaju, bo K-normni mehanizem ε-diferencirano zaseben s pričakovano napako največ $O(cL_K) \cdot \operatorname{VolLB}(F,\varepsilon)$.*

Dokaz. Po izreku 5.10 je K-normni mehanizem ε-diferencirano zaseben in ima napako največ $O(d/\varepsilon)\,\mathbb{E}_{z\sim K}\|z\|$. Naj bo e_i i-ti bazni vektor standardne ortogonalne

baze $\mathbb{R}^d$ in ocenimo

$$\begin{aligned}\mathop{\mathbb{E}}_{z\sim K}\|z\|^2 &= \int_K \frac{1}{\mathrm{Vol}(K)}\|z\|^2 dz \\ &= \frac{1}{\mathrm{Vol}(K)}\int_K (z_1^2+\cdots+z_d^2)dz \\ &= \sum_{i=1}^d \frac{1}{\mathrm{Vol}(K)}\int_K \langle z, e_i\rangle^2 dz \\ &\le dc^2 L_K^2 \,\mathrm{Vol}(K)^{2/d}.\end{aligned}$$

Sedaj združimo to z napako KM in po uporabi Jensenove neenakosti dobimo

$$\begin{aligned}O(d/\varepsilon)\mathop{\mathbb{E}}_{z\sim K}\|z\| &\le O(d/\varepsilon)\sqrt{\mathop{\mathbb{E}}_{z\sim K}\|z\|^2} \\ &\le O(\varepsilon^{-1}d\sqrt{d}\,\mathrm{Vol}(K)^{1/d}cL_K) \\ &= O(cL_K)\,\mathrm{VolLB}(F,\varepsilon).\end{aligned}$$ □

Iz tega izreka vidimo, da se splošna spodnja meja in pa zgornja meja K-normnega mehanizma ujemata do faktorja cL_K natančno. Ocena L_K za splošno konveksno telo je dobro znan odprt problem v konveksni geometriji. Trenutno najboljša dokazana zgornja meja za splošno konveksno telo je $L_K \le O(d^{1/4})$, medtem ko obstaja domneva, da je $L_K = O(1)$. Ta je bolj podrobno obravnavana v delu [3].

Domneva 7.5 (hiperravninska domneva)**.** *Obstaja tak $C > 0$, da je za vsak d in vsako konveksno telo $K \subseteq \mathbb{R}^d$, $L_K < C$.*

Če privzamemo to domnevo, dobimo ujemanje mej za približno izotropska konveksna telesa. S pomočjo posledice 4.7 pa lahko dobljeno zgornjo mejo še poenostavimo in dobimo sledeč izrek.

Izrek 7.6. *Naj bo $\varepsilon > 0$ in privzemimo hiperravninsko domnevo. Potem je za vsako matriko $F \in [-1,1]^{d\times n}$, za katero je $K = FB_1^n$ v c-približno izotropskem položaju, K-normni mehanizem ε-diferencirano zaseben s pričakovano napako največ*

$$\mathrm{err}(M,F) \le O(cd/\varepsilon)\min\{\sqrt{d}, \sqrt{\log(n/d)}\}.$$

8. Neizotropska telesa

Dosedaj opisan K-normni mehanizem je torej skoraj optimalen za izotropska telesa, za neizotropska telesa pa je lahko daleč od optimalnega.

V tem razdelku bomo prilagodili K-normni mehanizem, tako da bo bolje deloval tudi na neizotropskih telesih, za to pa bomo potrebovali še nekaj teorije iz konveskne geometrije.

Definicija 8.1. Konveksno telo $K \subseteq \mathbb{R}^d$ ima center mase v 0, če zanj velja $\int_K x dx = 0$. Kovariančna matrika takega telesa K se označi z $M_K \in \mathbb{R}^{d\times d}$, njeni elementi pa so:

$$M_{ij} = \frac{1}{\mathrm{Vol}(K)}\int_K x_i x_j dx.$$

To je tudi natanko kovariančna matrika enakomerne porazdelitve na množici K.

8.1. **Rekurzivni mehanizem.** Ideja novega mehanizma je, da se na različnih lastnih podprostorih, ki pripadajo lastnim vrednostim kovariančne matrike, obnaša različno. Natančneje, mehanizem bo na tistih podprostorih, ki pripradajo večjim lastnim vrednostim, uporabil nižje-dimenzionalno obliko K-normnega mehanizma, na koncu pa vse rezultate združil.

Mehanizem bomo označevali z $\mathrm{NIM}(F, d, \varepsilon)$, kjer je $F : \mathbb{R}^n \to \mathbb{R}^d$ linearna preslikava, $d \in \mathbb{N}$ in $\varepsilon > 0$. V vsakem rekurzivnem koraku bomo konveksno telo F razdelili na dva dela glede na velikost lastnih vrednosti matrike M_K, nato pa na enem delu uporabili KM mehanizem, drugega pa se bomo lotili rekurzivno.

Algoritem 1: Neizotropski mehanizem – NIM

Vhod: linearna preslikava $F : \mathbb{R}^n \to \mathbb{R}^d, x \in \mathbb{R}^n, d \in \mathbb{N}, \varepsilon > 0$
Izhod: $(\varepsilon \log n)$-diferencirano zaseben odgovor na poizvedbo F

1. Naj bo $K = FB_1^n$, $\omega_1 \geq \omega_2 \geq \cdots \geq \omega_d$ pa so lastne vrednosti kovariančne matrike M_K. Poiščemo še pripadajočo ortonormirano bazo iz lastnih vektorjev $u_1, \ldots, u_d$.
2. Naj bo $d' = \lfloor d/2 \rfloor$. Prostor U naj razpenjajo lastni vektorji, ki pripadajo večji polovici lastnih vrednosti, prostor V pa manjši. Torej $U = \mathrm{span}\{u_1, \ldots, u_{d'}\}$ in $V = \mathrm{span}\{u_{d'+1}, \ldots, u_d\}$. S P_U in P_V označimo ortogonalna projekcijska operatorja na prostora U in V.
3. Uporabimo K-normni mehanizem, da dobimo $a \sim \mathrm{KM}(F, d, \varepsilon)$
4. Če je $d = 1$, vrnemo $P_V a$, drugače vrnemo $\mathrm{NIM}(P_U F, d', \varepsilon) + P_V a$.

Opomba 8.2. Rekurzivni klic v tem mehanizmu uporabimo na sliki preslikave $P_U F$. To storimo tako, da uporabimo d' dimenzionalni K-normni algoritem v bazi prostora U in rezultat na koncu preslikamo nazaj v $\mathbb{R}^d$.

Očitno je, da bo v mehanizmu uporabljenih največ $\log d$ rekurzivni klicev. Za vsak korak $m \in \{0, \ldots, \log d\}$ z a_m označimo porazdelitev odgovora K-normnega mehanizma iz tretjega koraka.

Lema 8.3. *Mehanizem* $\mathrm{NIM}(F, d, \varepsilon)$ *je* $(\varepsilon \log d)$*-diferencirano zaseben.*

Dokaz. Končni odgovor mehanizma je porazdeljen kot funkcija slučajnih spremenljivk a_m. Vsaka izmed njih je zaradi izreka 5.10 ε-diferencirano zasebna. Naj bo $l = \log d$ in si oglejmo skupno porazdelitev vektorja $a = (a_1, \ldots, a_l)$ ter upoštevajmo neodvisnost različnih izvedb K-normnega mehanizma.

$$f_a(x_1, \ldots, x_l) = f_{a_1}(x_1) \cdot \cdots \cdot f_{a_l}(x_l)$$

Tu so f_{a_m} porazdelitvene funkcije posameznih členov, ki so vse ε-diferencirano zasebne, kar pa torej pomeni, da je skupna porazdelitev $(\varepsilon \log(d))$-diferencirano zasebna. Do konca nas sedaj pripelje izrek 3.11, ki zagotavlja zasebnost po kasnejši obdelavi. □

Ta del je bil enostavnejši, analiza napake tega algoritma pa bo zahtevala več truda. Najti bomo morali povezavo med volumnom telesa $P_U K$ in normo vektorja $P_V a$. Naprej se lotimo telesa $P_U K$.

8.2. **Volumen lastnih prostorov kovariančne matrike.** Cilj tega razdelka bo izraziti volumen izotropskega telesa K s pomočjo lastnih vrednosti njegove kovariančne matrike. To bomo kasneje potrebovali za oceno napake algoritma za neizotropska telesa. V [12] najdemo sledečo formulo za k-dimenzionalen volumen preseka izotropskega telesa s k-dimenzionalnim podprostorom.

Izrek 8.4 ([12, Izrek 3.11])**.** *Naj bo $K \subseteq \mathbb{R}^d$ izotropsko telo z $\mathrm{Vol}(K) = 1$, E pa k-dimenzionalen podprostor $\mathbb{R}^d$, kjer je $1 \leq k \leq d$. Potem velja*

$$\mathrm{Vol}_k(E \cap K)^{1/(d-k)} = \Theta\left(\frac{L_{B_K}}{L_K}\right).$$

B_K je tukaj neko eksplicitno definirano konveksno telo neodvisno od podprostora E.

Za lažje označevanje pišimo $\alpha_K = L_{B_K}/L_K$, torej spodnjo mejo volumna iz zgornjega izreka. Za neizotropske K lahko podobno kot za L_K definiramo tudi α_K kot α_{TK}, kjer je T linearna preslikava, ki prenese K v izotropsko pozicijo. Če predpostavimo hiperravninsko domnevo, očitno velja tudi $\alpha_K = \Omega(1)$.

Posledica 8.5. *Naj bo $K \subseteq \mathbb{R}^d$ izotropsko telo z $\mathrm{Vol}(K) = 1$, E pa k-dimenzionalen podprostor v $\mathbb{R}^d$, kjer je $1 \leq k \leq d$. Če s P označimo ortogonalno projekcijo na podprostor E, velja*

$$\mathrm{Vol}_k(PK)^{1/(d-k)} \geq \Omega(\alpha_K).$$

Dokaz. Ker je P identiteta na prostoru E velja $E \cap K \subseteq PK$. Torej s pomočjo prejšnjega izreka dobimo

$$\mathrm{Vol}_k(PK)^{1/(d-k)} \geq \mathrm{Vol}_k(E \cap K)^{1/(d-k)} \geq \Omega(\alpha_K). \qquad \square$$

Teh ocen ne moremo uporabiti direktno, saj veljajo le za izotropske K, v tem poglavju pa se ukvarjamo z neizotropskimi telesi. V nadaljevanju bomo poskusili telo K pretvoriti v izotropsko telo, v vsakem koraku pa paziti na to, koliko se mu je spremenil volumen. Za to bomo potrebovali še sledečo lemo.

Lema 8.6. *Naj bo $K \subseteq \mathbb{R}^d$ simetrično konveksno telo ($K = -K$) in T simetrična matrika z lastnimi vrednostmi $\lambda_1, \ldots, \lambda_1$ ter pripadajočimi enotskimi lastnimi vektorji $u_1, \ldots, u_d$. Naj bo $1 \leq k \leq d$ in $E = \mathrm{span}\{u_1, \ldots, u_k\}$, s P pa označimo projekcijo na podprostor E. Potem velja*

$$\mathrm{Vol}_k(PK) \geq \mathrm{Vol}_k(PTK) \prod_{i=1}^{k} \lambda_i^{-1}.$$

Dokaz. Brez škode za splošnost lahko predpostavimo, da so lastni vektorji T kar standardni bazni vektorji $e_1, \ldots, e_d$, saj so zaradi simetričnosti T ortogonalni in lahko to enostavno dosežemo z rotacijo prostora, ki ne spremeni nobenega volumna. Torej je matrika T oblike $\mathrm{diag}(\lambda_1, \ldots, \lambda_d)$, projekcija P pa je kar I_k, torej diagonalna matrika z enkami na prvih k diagonalnih elementih in ničlami drugod. Sedaj velja za $S = \mathrm{diag}(\lambda_1^{-1}, \ldots, \lambda_k^{-1}, 0, \ldots, 0)$ formula $P = SPT$, iz česar pa dobimo

$$\mathrm{Vol}_k(PK) = \mathrm{Vol}_k(SPTK) = \det(S_{|E}) \,\mathrm{Vol}_k(PTK) = \prod_{i=1}^{k} \lambda_i^{-1} \,\mathrm{Vol}_k(PTK). \qquad \square$$

Preden dokončamo izpeljavo meje za volumen PK, bomo potrebovali še sledečo trditev, ki izrazi izotropsko konstanto telesa K z determinanto njegove kovariančne matrike M_K.

Trditev 8.7 ([12, poglavje 1.6]). *Naj bo $K \subseteq \mathbb{R}^d$ konveksno telo. Potem za njegovo izotropsko konstanto velja:*

$$L_K^2 \operatorname{Vol}(K)^{\frac{2}{d}} = \det(M_K)^{\frac{1}{d}}.$$

Še več, K je v izotropskem položaju natanko tedaj, ko velja $M_K = L_K^2 \operatorname{Vol}(K)^{\frac{2}{d}} I$.

Vse to nas sedaj pripelje do ključnega izreka tega podpoglavja, v katerem bomo volumen telesa PK ocenili z lastnimi vrednostmi njegove kovariančne matrike.

Izrek 8.8. *Naj bo $K \subseteq \mathbb{R}^d$ simetrično konveksno telo in M_K njegova kovariančna matrika. Z $(u_1, \ldots, u_d)$ označimo ortonormirano bazo iz lastnih vektorjev te matrike, z $(\sigma_1, \ldots, \sigma_d)$ pa pripadajoče lastne vrednosti. Naj bo $1 \leq k \leq \lceil \frac{d}{2} \rceil$ in $E = span\{u_1, \ldots, u_k\}$. Če je P projekcijski operator na podprostor E, bo veljalo*

$$\operatorname{Vol}_k(PK)^{1/(d-k)} \geq \Omega(\alpha_K L_K^{-2}) \cdot \left(\prod_{i=1}^{k} \sigma_i^{1/2} \right)^{1/(d-k)}.$$

Uporabljen α_K je razreda $\Omega(d^{-\frac{1}{4}})$, L_k^{-2} pa razreda $\Omega(d^{-\frac{1}{2}})$. Če privzamemo še hiperravninsko domnevo, pa sta oba kar razreda $\Omega(1)$.

Dokaz. Naj bo slučajni vektor $X \sim K$ enakomerno porazdeljena po simetričnem konveksnem telesu K, potem je $M_K = \mathbb{E}[XX^T]$. Če je T obrnljiva linearna preslikava $T : \mathbb{R}^d \to \mathbb{R}^d$, je slučajni vektor TX porazdeljen enakomerno po TK, za njegovo kovariančno matriko pa velja $M_{TK} = \mathbb{E}[TXX^TT^T] = TM_KT^T$.

Kot kovariančna matrika je M_K pozitivno semidefinitna. Iz trditve 8.7 lahko sklepamo tudi, da je $\det(M_K) > 0$, saj je $L_K^2 \operatorname{Vol}(K)^{2/d} > 0$. Zato je M_K pozitivno definitna in je dobro definirana matrika $T = M_K^{-1/2}$. Zanjo torej velja $M_{TK} = M_K^{-1/2} M_K M_K^{-1/2} = I$, Iz trditve 8.7 zato sledi, da je telo TK v izotropskem položaju. Ker je $\det(M_{TK}) = 1$, velja tudi $\operatorname{Vol}(TK)^{1/d} = 1/L_{TK} = 1/L_K$. Če raztegnemo telo TK z faktorjem $\lambda = L_K$, dobimo

$$\operatorname{Vol}(\lambda TK) = \lambda^d \operatorname{Vol}(TK) = L_K^d \operatorname{Vol}(TK) = 1.$$

Preslikava λT ima lastne vrednosti $\lambda\sigma_1^{-\frac{1}{2}}, \ldots, \lambda\sigma_1^{-\frac{1}{2}}$, na njej pa lahko uporabimo lemo 8.6, iz česar dobimo

$$\operatorname{Vol}_k(PK) \geq \operatorname{Vol}_k(P\lambda TK) \prod_{i=1}^{k} \frac{\sqrt{\sigma_i}}{\lambda}.$$

Ker je $\operatorname{Vol}(\lambda TK) = 1$ in je telo v izotropskem položaju, lahko na njem uporabimo tudi posledico 8.5, iz česar zaključimo:

$$\begin{aligned} \operatorname{Vol}_k(PK)^{1/(d-k)} &\geq \operatorname{Vol}_k(P\lambda TK)^{1/(d-k)} \left(\prod_{i=1}^{k} \frac{\sqrt{\sigma_i}}{\lambda} \right)^{1/(d-k)} \\ &\geq \Omega(\alpha_K) \prod_{i=1}^{k} \left(\frac{\sqrt{\sigma_i}}{\lambda} \right)^{1/(d-k)}. \end{aligned}$$

Do končnega izreka nam torej manjka le še ocena faktorja $\lambda^{-\frac{k}{d-k}}$. Ker je $\frac{k}{d-k} \leq 2$, velja

$$\lambda^{-\frac{k}{d-k}} = L_K^{-\frac{k}{d-k}} \geq L_K^{-2}.$$

To pa je zagotovo $\Omega(d^{-1/2})$, oziroma, če predpostavimo hiperravninsko domnevo, tudi $\Omega(1)$. □

8.3. **Optimalnost** NIM **mehanizma.** Zdaj imamo pripravljeno že skoraj vse, kar bomo potrebovali, da končno ocenimo volumne projekcij iz našega mehanizma. Sledeča trditev nam bo pomagala prevesti problem ocenjevanja pričakovane vrednosti napake na ocenjevanje lastnih vrednosti matrike, ki smo ga obdelali v prejšnjem poglavju.

Trditev 8.9. *Naj bo $K \subseteq \mathbb{R}^d$ simetrično konveksno telo, $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_d$ lastne vrednosti kovariančne matrike M_K in $u_1, \ldots, u_d$ pripadajoči ortonormalni lastni vektorji. Potem za vsak $1 \leq i \leq d$ velja*

$$\sigma_i = \mathop{\mathbb{E}}_{z \sim K} \langle z, u_i \rangle^2.$$

Dokaz. Iz $M_K u_i = \sigma_i u_i$ dobimo $u_i^T M_k u_i = u_i^T \sigma_i u_i = \sigma_i$, torej, če razpišemo skalarni produkt

$$\begin{aligned}
\mathop{\mathbb{E}}_{z \in K} \langle z, u_i \rangle^2 &= \frac{1}{\mathrm{Vol}(K)} \int_K \langle z, u_i \rangle^2 dz \\
&= \frac{1}{\mathrm{Vol}(K)} \int_K \left(\sum_{j=1}^{d} z_j u_{ij} \right)^2 dz \\
&= u_i^T M_K u_i = \sigma_i.
\end{aligned}$$

□

Lema 8.10. *Naj bo a slučajna spremenljivka, ki jo vrne K-normni mehanizem v tretjem koraku* $\mathrm{NIM}(F, d, \varepsilon)$ *mehanizma za podatkovno bazo $x = 0$. Za njo velja*

$$\mathrm{GVolLB}(F, \varepsilon)^2 \geq \Omega(\alpha_K^2 L_K^{-4}) \, \mathbb{E} \|P_V a\|_2^2.$$

Dokaz. Predpostavimo, da je d sod, torej je $d - d' = d'$. Izpeljava za lih d je simetrična. Izrek 5.10 nam pove, da pri $p = 2$ in podatkovni bazi $x = 0$ slučajna spremenljivka a zadošča sledeči enačbi.

$$\begin{aligned}
\mathbb{E} \|P_V a\|_2^2 &= \frac{\Gamma(d+3)}{\varepsilon^2 \, \Gamma(d+1)} \mathop{\mathbb{E}}_{z \in K} \|P_V z\|_2^2 \\
&= \frac{(d+2)(d+1)}{\varepsilon^2} \mathop{\mathbb{E}}_{z \in K} \|P_V z\|_2^2 \\
&= O\left(\frac{d^2}{\varepsilon^2}\right) \mathop{\mathbb{E}}_{z \in K} \left\| P_V \left(\sum_{i=1}^{d} \langle z, u_i \rangle u_i \right) \right\|_2^2 && \text{(zapišemo v bazi } u_i\text{)} \\
&= O\left(\frac{d^2}{\varepsilon^2}\right) \sum_{i=d'+1}^{d} \mathop{\mathbb{E}}_{z \in K} \langle z, u_i \rangle^2 && \text{(projekcija ohrani } u_i \text{ za } i \geq d+1\text{)} \\
&= O\left(\frac{d^2}{\varepsilon^2}\right) \sum_{i=d'+1}^{d} \sigma_i && \text{(trditev 8.9)} \\
&\leq O\left(\frac{d^3}{\varepsilon^2}\right) \sigma_{d'+1}.
\end{aligned}$$

Po drugi strani pa iz definicije GVolLB in leme dobimo

$$\mathrm{GVolLB}(F,\varepsilon)^2 \geq \Omega\left(\frac{d^3}{\varepsilon^2}\right) Vol_{d'}(P_U K)^{2/d'} \geq \Omega\left(\frac{d^3}{\varepsilon^2}\right) \Omega(\alpha_K^2 L_K^{-4}) \left(\prod_{i=1}^{d'} \sigma_i\right)^{1/d'}.$$

Ker je $\sigma_{d'} \geq \sigma_{d'+1}$, lahko ti dve enačbi združimo in dobimo

$$\mathrm{GVolLB}(F,\varepsilon)^2 \geq \Omega(\alpha_K^2 L_K^{-4})\, \mathbb{E}\|P_V a\|^2. \qquad \square$$

Lema 8.11. *Če privzamemo hiperravninsko domnevo, bo za l_2-napako mehanizma* $\mathrm{NIM}(F,d,\varepsilon)$ *veljalo*

$$\mathrm{err}(\mathrm{NIM},F) \leq O(\sqrt{\log(d)})\, \mathrm{GVolLB}(F,\varepsilon).$$

Dokaz. Sešteti moramo napake vseh rekurzivnih korakov v našem mehanizmu. S $P_{V_m} a_m$ označimo projekcijo K-normnega mehanizma iz tretjega koraka v pripadajoči lastni podprostor V_m, $a \in \mathbb{R}^d$ pa naj označuje končni rezultat mehanizma. Podobno kot v oceni zgornje meje za K-normni mehanizem lahko opazimo, da je pričakovana napaka neodvisna od podatkovne baze x, saj je v končnem rezultatu mehanizma vedno vsebovana $\sum_{m=1}^{d} P_{V_m} F x = Fx$. Zato lahko ocenimo napako za podatkovno bazo $x = 0$ tako

$$\begin{aligned}
\mathbb{E}\|a\|_2 &\leq \sqrt{\mathbb{E}\|a\|_2^2} && \text{(Jensenova neenakost)}\\
&= \sqrt{\sum_{i=1}^{\log d}\sum_{k=1}^{\log d} \mathbb{E}\langle P_{V_i} a_i, P_{V_j} a_j\rangle}\\
&= \sqrt{\sum_{m=1}^{\log d} \mathbb{E}\|P_{V_m} a_m\|_2^2}\\
&\leq \sqrt{\sum_{m=1}^{\log d} O(\alpha_{K_m}^{-2} L_K^4)\cdot \mathrm{GVolLB}(P_{U_m} F,\varepsilon)^2} && \text{(po lemi 8.10)}\\
&\leq O\left(\sqrt{\log d}\max_m \left[\alpha_{K_m}^{-1} L_{K_m}^2\right]\right) \mathrm{GVolLB}(F,\varepsilon).
\end{aligned}$$

Tukaj smo uporabili tudi $\mathrm{GVolLB}(F,\varepsilon) \geq \mathrm{GVolLB}(P_U F,\varepsilon)$, ki sledi kar iz definicije GVolLB. Po predpostavki hiperravninske hipoteze dobimo še $max_m \alpha_{K_m}^{-1} = O(1)$ kar zaključi dokaz. $\square$

Posledica 8.12. *Naj bo $\varepsilon > 0$ in $F : \mathbb{R}^n \to \mathbb{R}^d$ linearna poizvedba. Če predpostavimo hiperravninsko hipotezo, obstaja ε-diferencirano zaseben mehanizem M z napako največ*

$$\mathrm{err}(M,F) \leq O(\log(d)^{3/2} \cdot \mathrm{GVolLB}(F,\varepsilon)).$$

Dokaz. Lema 8.3 nam pove, da je mehanizem $\mathrm{NIM}(F,d,\varepsilon/\log(d))$ ε-diferencirano zaseben, napaka pa je direktna posledica leme 8.11. $\square$

S tem smo pokazali, da tako spodnja meja GVolLB in zgornja meja mehanizma NIM od optimuma odstopata za največ faktor $O(\log(d)^{3/2})$.

9. Implementacija mehanizmov

V tem zadnjem poglavju se bomo lotili problema še malo praktično, tako da bomo implementirali Laplaceov in K-normni mehanizem ter ju preizkusili na neki podatkovni bazi. To bomo storili v programskem jeziku Python, veliko pa se bomo opirali na knjižnici *numpy* in *scipy*. Implementacija Laplaceovega mehanizma je precej preprosta, medtem ko bomo za implementacijo K-normnega mehanizma potrebovali malo več dela. Težava bo nastala pri enakomernem vzorčenju iz telesa $K = FB_1^d$ o katerem vemo zelo malo. Implementacije NIM mehanizma pa se bomo dotaknili le teoretično, saj je v praksi brez dodatnih optimizacij zelo počasen.

9.1. **Laplaceov mehanizem.** Za primerjavo bomo najprej implementirali Laplaceov mehanizem. Vse, kar moramo storiti, je, da vsaki komponenti dejanskega odgovora prištejemo Laplaceovo porazdeljen šum s konstanto $b = 2\varepsilon^{-1}$. To storimo kar z uporabo statističnih metod knjižnice *numpy*.

9.2. **K-Normni mehanizem.** Kot že povedano, je ta mehanizem bolj zanimiv z vidika implementacije. Za začetek si še enkrat oglejmo, kako ta mehanizem deluje in analizirajmo vsak korak posebej.

Prvi korak ni problematičen, tukaj le vzorčimo r iz porazdelitve $\mathrm{Gamma}(d+1, \varepsilon)$ s pomočjo knjižnice *numpy*, v drugem koraku pa moramo vzorčiti še iz enakomerne porazdelitve po množici $K = FB_1^n$. Ker je K konveksno telo, lahko to storimo s pomočjo naključnih sprehodov po prostoru, katerih porazdelitev se bliža enakomerni. Ti so predstavljeni v raziskavi L. Lovasza [14]. Dva najpreprostejša sta sprehod po mreži in pa sprehod po kroglah. V prvem začnemo v neki točki in se nato v vsakem koraku prestavimo na naključno izbrano sosednjo točko na mreži $\gamma\mathbb{Z}^d$, če je le-ta znotraj našega telesa. V sprehodu po kroglah, ki ga bomo uporabljali mi, pa najprej točko, v katero se bomo premaknili, izberemo enakomerno iz 2-krogle z radijem γ okrog trenutne točke. Po določenem številu korakov (znano je, da polinomsko mnogo) se bo porazdelitev trenutne točke dovolj približala enakomerni, da bo naš mehanizem deloval.

Za uporabo tega sprehoda mora za telo K veljati sledeče:

(1) K mora biti omejen iz obeh strani, torej obstajati morata konstanti r in R, da je $rB_2^d \subseteq K \subseteq RB_2^d$
(2) Potrebujemo učinkovit način preverjanja ali je točka x znotraj telesa K

Za $K = FB_1^n$ pri $F \in [-1, 1]^{d\times n}$ je prvi zahtevi naravno zadoščeno, saj velja $K \subseteq B_\infty^d \subseteq \sqrt{d}B_2^d$. Poleg tega lahko namesto, da v mehanizmu uporabljamo direktno K, uporabimo $K' = K + B_2^d$. V primeru, da je končni rezultat element B_2^d, bo za napako mehanizma KM' v 2-normi očitno veljalo

$$\mathrm{err}(\mathrm{KM}', F) = \mathrm{err}(\mathrm{KM}, F) + O(1).$$

9.2.1. *Preverjanje vsebovanih točk.* Druga zahteva je implementacija algoritma, ki bo za vsak $a \in \mathbb{R}^d$ ugotovil, ali obstaja tak $x \in B_1^n$, da zanj velja $Fx = a$. Ta problem lahko zastavimo tudi kot problem konveksne optimizacije, bolj natačno problem najmanjših kvadratov z omejitvijo v 1-normi.

Zanima nas torej:

$$\min \frac{1}{2}\|Fx - a\|_2$$
$$\text{p.p. } \|x\|_1 \leq 1.$$

Optimizacijsko funkcijo lahko na klasičen način prevedemo na problem kvadratičnega programiranja, kasneje pa bomo še omejitve prevedli v linearno obliko.

$$\begin{aligned}\frac{1}{2}\|Fx - a\|_2 &= \frac{1}{2}(Fx - a)^T(Fx - a) \\ &= \frac{1}{2}(x^T F^T Fx - a^T Fx - x^T F^T a + a^T a) \\ &\sim \frac{1}{2}x^T F^T Fx - a^T x\end{aligned}$$

Člena $a^T Fx$ in $x^T F^T a$ sta oba skalarna, zato tudi enaka, člen $a^T a$ pa je konstanten in zato ne spremeni x pri katerem pride do minimuma. Ostane nam še, da pogoj $\|x\|_1 \leq 1$ spravimo v linearno obliko, torej v $Cx \geq d$ za neko matriko C in vektor d. Najbolj očiten način kako to storiti je, da z 1 omejimo vsako možno kombinacijo predznakov komponent x_i. To nam bi prineslo 2^d omejitev, kar pa za učinkovito reševanje ni uporabno. V [13] je opisan lepši način kjer namesto z matriko F, delamo z matriko $Q = [F, -F]$. Vsako spremenljivko uvodnega problema x_i lahko zapišemo kot razliko dveh nenegativnih spremenljivk, kjer x_i^+ predstavlja pozitivni x_i^- pa negativni del spremenljivke:

$$x_i = x_i^+ - x_i^-.$$

Omejitve tega prilagojenega problema so

$$\begin{aligned} x_i^+ &\geq 0 \\ x_i^- &\geq 0 \\ \sum_{i=1}^{d}(x_i^+ + x_i^-) &\leq 1.\end{aligned}$$

Hitro lahko preverimo, iz teh omejitev sledi začetna:

$$\|x\|_1 = \sum_{i=1}^{d}|x_i| = \sum_{i=1}^{d}|x_i^+ - x_i^-| \leq \sum_{i=1}^{d} x_i^+ + x_i^- \leq 1.$$

S tem naš problem prevedemo na sledeč problem kvadratičnega programiranja z linearnimi omejitvami, ki pa ga znamo rešiti (v sami implementaciji uporabimo nabor orodij *cvxopt*). Če pišemo $Q = [F, -F]$ dobimo

$$\min_x \frac{1}{2}x^T Q^T Qx - a^T Qx$$
$$\text{p.p.} \begin{bmatrix} I_{2d} \\ \mathbb{1}_{2d}^T \end{bmatrix} x \geq \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}$$

Na koncu nas zanima le, ali je pri x, pri katerem je dosežen ta minimum, $\|Fx - a\|_2 = 0$, kar bi pomenilo, da je začetna točka a vsebovana v FB_1^n.

Opomba 9.1. Zaradi računanja s plavajočo vejico dejanski minimum ni skoraj nikoli natančno 0, zato zahtevamo le, da je manjši od npr. 10^{-5}.

9.2.2. *Enakomeren izbor iz B_2-krogle.* Zdaj, ko smo zadostili zahtevam za uporabo slučajnega sprehoda po kroglah, bomo očitno potrebovali tudi način, kako iz B_2 krogle enakomerno izberemo točke. To bomo storili z uporabo sledečega izreka, delo Bartheja.

Izrek 9.2 ([2, Izrek 1])**.** *Naj bodo $X_1, X_2, \ldots, X_n$ neodvisne porazdeljene slučajne spremenljivke z gostoto*

$$f(x) = \frac{p}{2\,\Gamma(p)} e^{|x|^p}$$

E naj bo od njih neodvisna eksponentno porazdeljena slučajna spremenljivka z gostoto $f(x) = e^{-x}$. Potem je vektor

$$(U_1, \ldots, U_n) = \frac{(X_1, \ldots, X_n)}{(\sum_{i=1}^{n} X_i^p + E)^{1/p}}$$

porazdeljen enakomerno na krogli B_2^p.

To je vse kar smo potrebovali za implementacijo K-normnega mehanizma, izvorna koda se nahaja v Dodatku 9.4.

Opomba 9.3. Tukaj nismo dokazali, da je mehanizem diferencirano zaseben tudi ob uporabi približno enakomerne porazdelitve. Klasično se oceni, da je približek porazdelitve, ki ga dobimo s sprehodom po kroglah aditivno blizu prave enakomerne porazdelitve, kar bi nam zagotovilo le δ-približno ε-zasebnost. Z modifikacijo standardnih argumentov pa je mogoče dokazati, da je lahko s polinomskim številom korakov zagotovljena tudi ε-zasebnost. Več v [5].

Opomba 9.4. Vsi koraki našega algoritma se zgodijo v polinomskem času, zato bi lahko rekli, da je učinkovit. A vendar to ne pomeni, da je v realni uporabi hiter, saj že samo preverjanje ali je $x \in K$ zahteva reševanje kvadratičnega problema z omejitvami dimenzij $2n \times 2n$, ki lahko za večje podatkovne baze vzame kar precej časa, poleg tega pa ga moramo za dosego dobrega približka enakomerne porazdelitve uporabiti mnogokrat. Za praktično uporabo na večjih podatkovnih bazah, ga bi bilo potrebno še bolj optimizirati.

9.3. **NIM mehanizem.** Dodatno delo nastane, saj moramo, če telo K ni izotropsko, poleg samega K-normnega mehanizma izračunati tudi lastna podprostora U in V kovariančne matrike M_K. Ker te matrike ne znamo natančno izračunati, jo moramo nekako dovolj dobro oceniti. Tu nam pomaga ravno dejstvo, da je M_K kovariančna matrika enakomerne porazdelitve na K. V prejšnjem poglavju smo predstavili algoritem za vzorčenje iz te porazdelitve, torej lahko z dovolj vzorci iz te porazdelitve ocenimo matriko M_K s kovariančno matriko vzorca $\tilde{M}_K$.

Od tod naprej je sama implementacija dokaj neposredna, vendar pa je treba ravno tako kot v prejšnjem poglavju, tudi sedaj paziti, da ε-diferencirana zasebnost ni očitno zagotovljena zaradi uporabe približka matrike M_K. Tudi tukaj gre vse dokaze izrekov iz poglavja 8 prilagoditi tako, da delujejo z rahlo pokvarjeno matriko. Več v [5].

Igralec	**Starost**
Ja'Wuan James	27
Lane Johnson	29
Ricky Wagner	30
Rob Havenstein	26

Tabela 2. Primer podatkov.

9.4. **Primerjava rezultatov.** Naše mehanizme bomo preizkusili na podatkih o starosti profesionalnih nogometašev. V podatkovni bazi bo 200 nogometašev, katerih starosti se gibljejo med 22 in 38 let. V tabeli 2 je prikazan primer teh podatkov.

Poizvedbe, ki jih bomo uporabljali bodo naključne, kot v poglavju 6.1 Bernoullijevo porazdeljene matrike velikosti $d \times n$, eno z $d = n/5$ in eno z $d = n/10$ poizvedbami. Za vsako bomo nato izračunali 100 vzorcev K-normnega in Laplaceovega mehanizma za vsak $\varepsilon \in \{0.1, 1, 2\}$. Zanimala nas bo predvsem povprečna napaka mehanizma v 2-normi in pa čas izvajanja. Rezultati so predstavljeni v tabeli 3.

Opomba 9.5. Tukaj bomo predstavili podatke dveh naključnih poizvedb, mehanizme pa smo preizkusili še na drugih, a so rezultati večinoma enaki.

ε	**Laplace** err	**Laplace čas**	K**-normni** err	K**-normni čas**
0.1	90.1	0.01 s	326.2	528 s
1	8.74	0.01 s	32.3	560 s
2	4.52	0.01 s	16.3	537 s

0.1	60.1	0.01 s	186.3	241 s
1	6.0	0.01 s	18.3	229 s
2	3.0	0.01 s	9.1	228 s

Tabela 3. Rezultati mehanizmov za naključni poizvedbi.

Opazimo, da so napake K-normnega mehanizma v poprečju večje od napak Laplaceovega mehanizma pri isti zahtevani zasebnosti, medtem ko pa je čas izvajanja bistveno daljši. To seveda ni v nasprotju z v tem delu obravnavano optimalnostjo K-norma mehanizma, saj je le-ta optimalen v asimptotskem smislu, ko gresta n in d proti neskončno. Prav tako izpeljane spodnje in zgornje meje napake, tudi z veliko vzorci, ki bi dobro ocenili pričakovano napako, zaradi njene asimptotske narave ni mogoče preveriti. Je pa že na majhnem število uporabljenih ε za oba mehanizma opaziti, da je napaka sorazmerna faktorju ε^{-1}.

Dodatek A: Programska koda

Tu je na voljo v zadnjem poglavju uporabljena programska koda za Laplaceov in K-normni mehanizem. Napisana je v programskem jeziku *Python 3*, za njeno uporabo pa so zahtevane sledeče knjižnice:

- *numpy* - numerično računanje
- *pandas* - uvoz in obdelava podatkov
- *cvxopt* - konveksna optimizacija

- *scipy* - verjetnostne porazdelitve

Vsa koda je dosegljiva na Githubu, glej [10].

```
# za poizvedbo F in podatkovno bazo x odgovori z
# ε-diferencirano zasebnim odgovorom z uporabo Laplacevega mehanizma
def laplace(F, x, eps):
    result = F @ np.transpose(x)
    b = 2 / eps
    return result + np.random.laplace(0,b, size = result.shape)

# reši problem min‖Cx − d‖_2
# pri pogojih Ax >= b
def least_squares(C, d, A, b):
    C = C.astype('float64')
    P = C.T @ C
    q = - C.T @ d
    G = -A
    h = -b
    sol = qpsolvers.solve_qp(P, q, G, h, solver='cvxopt')
    return np.linalg.norm(C @ sol - d, ord=2)

# vrne ali je a ∈ FB_1^n + B_1^d
def oracle(F, a):
    d, n = F.shape
    if np.linalg.norm(a, ord=1) <= 1:
        return True

    F_double = np.hstack((F, -F))
    A = np.vstack((np.identity(2*n),-np.ones(2*n)))
    b = np.zeros(2*n+1)
    b[2*n] = -1
    res = least_squares(F_double, a, A, b)

    return res < 1e-5

# vrne vzorec enakomerne porazdelitve na B_p^n krogli
def uniform_ball(d, delta, p=2):
    normals = stats.gennorm(beta=p).rvs(size=d)
    exp = np.random.exponential()
    denom = (np.sum(abs(normals)**p)+exp)**(1/p)
    return (normals / denom) * delta

# naredi en korak sprehoda po kroglah velikosti δ iz točke a po telesu FB_1^n + B_1^d
def ballwalk(F, a, delta):
    d, n = F.shape
    move = uniform_ball(d, delta)
    while not oracle(F, a + move):
        print('fail')
        move = uniform_ball(d, delta)
    print(end - start)
    return move + a

# za poizvedbo F in podatkovno bazo x odgovori z
# ε-diferencirano zasebnim odgovorom z uporabo K-normnega mehanizma
def knorm(F, x, eps):
    d, n = F.shape

    r = np.random.gamma(d+1, scale=(1/eps))
    a = uniform_ball(d, 1, p=1)

    for i in range(n^2):
```

```
        a = ballwalk(F, a, 1/np.sqrt(d))

    return F @ x + r * a
```

Slovar strokovnih izrazov

approximate differential privacy približna diferencirana zasebnost
database podatkovna baza
differential privacy diferencirana zasebnost
isotropic position izotropski položaj
random algorithm naključen algoritem
response mechanism odzivni mehanizem
query poizvedba
sensitivity občutljivost

Literatura

[1] I. Bárány in Z. Füredi, *Approximation of the sphere by polytopes having few vertices*, Proc. Amer. Math. Soc. **102** (1988) 651–659

[2] F. Barthe, O. Guédon, S. Mendelson in A. Naor, *A probabilistic approach to the geometry of the ℓ_p^N ball*, Ann. Probab. **33** (2005) 480–513.

[3] S. Brazitikos, A. Giannopoulos, P. Valettas in B. H. Vritsiou, *Geometry of Isotropic Convex Bodies*, Mathematical surveys and monographs **196**, AMS, Providence, 2014.

[4] C. Dwork in A. Roth, *The algorithmic foundations of differential privacy*, Found. Trends Theor. Comput. Sci. **9** (2014) 211-407

[5] M. Hardt in K. Talwar, *On the geometry of differential privacy*, Conf. Proc. Theory Comput. **42** (2010) 705–714

[6] M. Jazbec, *Splošna definicija diferencirane zasebnosti*, diplomsko delo, Fakulteta za matematiko in fiziko, Univerza v Ljubljani, 2018.

[7] B. Klartag in G. Kozma, *On the hyperplane conjecture for random convex sets*, Israel J. Math. **170** (2009) 253–268.

[8] A. Klenke, *Probability theory: A comprehensive course*, Universitext, Springer, London, 2008.

[9] A. E. Litvak, A. Pajor, M. Rudelson in N. Tomczak-Jaegermann, *Smallest singular value of random matrices and geometry of random polytopes*, Adv. Math **195** (2005).

[10] L. Lodrant, *Programska koda diferencirano zasebnih mehanizmov*, 2019, dostopno na: `github.com/lodrantl/diplomski-seminar`.

[11] F. McSherry in K. Talwar, *Mechanism design via differential privacy* , Proceedings of the 48th annual IEEE symposium on foundations of computer science (2007) 94–103.

[12] V. D. Milman in A. Pajor, *Isotropic position and inertia ellipsoids and zonoids of unit ball of a normed n-dimensional space*, v: Geometric aspects of functional analysis (ur. J. Lindenstrauss in V. .D. Milman), Lecture notes in mathematics **1376**, Springer, Berlin, 1989, str. 64–104.

[13] M. Schmidt, *Least squares optimization with L1-norm regularization*, 2005, dostopno na: `www.cs.ubc.ca/~schmidtm/Documents/2005_Notes_Lasso.pdf`.

[14] S. Vempala, *Geometric random walks: a survey*, Combinatorial and computational geometry **22** (2005) 573–612.