

Splošna definicija diferencirane zasebnosti

Metod Jazbec

Mentor: doc. dr. Aljoša Peperko

Fakulteta za matematiko in fiziko
Univerza v Ljubljani

Diplomski seminar, 2018

Uvod

- Diferencirana zasebnost - matematična definicija zasebnosti pri javni objavi ter rudarjenju podatkov.

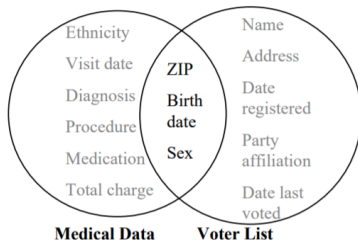


Figure 1: Metoda anonimizacije in napad s pomožno podatkovno bazo.

Podatkovna baza

- (U, ρ) metrični prostor, $D \subseteq U$. Podatkovno bazo predstavimo z vektorjem $\mathbf{d} = (d_1, \dots, d_n) \in D^n$.
- Dve podatkovni bazi, $\mathbf{d} = (a_1, \dots, a_n)$ in $\mathbf{d}' = (b_1, \dots, b_n)$, sta *sosednji*, če se razlikujeta v natanko enem vnosu. Označimo $\mathbf{d} \sim \mathbf{d}'$.
- Množico U opremimo z Borelovo σ -algebro, označimo jo z \mathcal{A}_U .

Poizvedba

- Poizvedba (angl. query) je način pridobitve željenih informacij iz podatkovne baze (SQL ...).
- Metrični prostor vseh možnih odgovor na posamezno poizvedbo $(E_Q, \rho_Q, \mathcal{A}_Q)$.
- Poizvedba je merljiva funkcija $Q : U^n \rightarrow E_Q$.

Odzivni mehanizem

Naj bo $(\Omega, \mathcal{F}, \mathbb{P})$ verjetnostni prostor. *Odzivni mehanizem* (za izbran nabor poizvedb $\mathcal{Q}(n)$) je definiran kot družina slučajnih spremenljivk

$$\{X_{Q,\mathbf{d}} : \Omega \rightarrow E_Q | Q \in \mathcal{Q}(n), \mathbf{d} \in D^n\}. \quad (1)$$

- Perturbacija podatkovne baze:

$$X_{Q,\mathbf{d}} = Q \circ Y_{\mathbf{d}} \quad (2)$$

- Perturbacija odgovorov na poizvedbo:

$$X_{Q,\mathbf{d}} = Z_{Q(\mathbf{d})} \quad (3)$$

Definicija diferencirane zasebnosti

Definicija (Diferencirana zasebnost za posamezno poizvedbo)

Naj bo $\epsilon > 0$ in $0 \leq \delta \leq 1$. Odzivni mehanizem je (ϵ, δ) -diferencirano zaseben za poizvedbo Q , če za vse $\mathbf{d} \sim \mathbf{d}' \in D^n$ in za vse $A \in \mathcal{A}_Q$ velja

$$\mathbb{P}(X_{Q,\mathbf{d}} \in A) \leq e^\epsilon \mathbb{P}(X_{Q,\mathbf{d}'} \in A) + \delta \quad (4)$$

- Simetričnost definicije!

Zadostne testne množice

Izrek 1

Naj bosta podana odzivni mehanizem (1) in poizvedba (E_Q, \mathcal{A}_Q, Q) . Naj bo $S \subset \mathcal{A}_Q$ algebra in naj velja $\sigma(S) = \mathcal{A}_Q$. Če (4) velja za vse $A \in S$, potem velja za vse $A \in \mathcal{A}_Q$.

Identična poizvedba

- $I_n : D^n \rightarrow D^n$, $I_n(\mathbf{d}) = \mathbf{d}$ (javna objava celotne podatkovne baze).

Izrek 2

Naj bo odzivni mehanizem s perturbacijo podatkovne baze (ϵ, δ) -diferencirano zaseben glede na identično poizvedbo $(U^n, \mathcal{A}_{U^n}, I_n)$. Potem sledi, da je tak mehanizem (ϵ, δ) -diferencirano zaseben glede na katerokoli poizvedbo (E_Q, \mathcal{A}_Q, Q) .

- Posledica: mehanizmi oblike (2) so robustni na to, kolikokrat ponovimo določeno poizvedbo.

Poenostavitev na 1D baze

Izrek 3

Naj bo podana družina 1-dimenzionalnih diferencirano zasebnih mehanizmov $\{Y_d : \Omega \rightarrow U \mid d \in D\}$. Velja torej

$$\mathbb{P}(Y_d \in A) \leq e^\epsilon \mathbb{P}(Y_{d'} \in A) + \delta$$

za vse $d, d' \in D, A \in \mathcal{A}_D$. Če definiramo n -dimenzionalni odzivni mehanizem kot $Y_{\mathbf{d}}(\omega) = (Y_{d_1}(\omega), \dots, Y_{d_n}(\omega))$, potem sledi, da je tudi ta diferencirano zaseben:

$$\mathbb{P}(Y_{\mathbf{d}} \in A) \leq e^\epsilon \mathbb{P}(Y_{\mathbf{d}'} \in A) + \delta$$

za vse $\mathbf{d} \sim \mathbf{d}' \in D^n, A \in \mathcal{A}_{D^n}$.

- Uporaba na diskretnem metričnem prostoru.

Laplaceov mehanizem na numerične podatke

- $D \subset \mathbb{R}$, kompakten. $\text{diam}(D) := \max_{d, d' \in D} \rho(d, d')$.
- $L : \Omega \rightarrow \mathbb{R}$ Laplaceova slučajna spremenljivka s parametroma $(0, b)$, $b > 0$, $f(x) = \frac{1}{2b} e^{-\frac{|x|}{b}}$.
- Za vsak $d \in D$ definiramo 1-dimenzionalni mehanizem kot $Y_d(\omega) = d + L(\omega)$.
- Parameter b izberimo tako, da velja

$$b \geq \frac{\text{diam}(D)}{\epsilon - \log(1 - \delta)}.$$

Potem sledi, da je vsak n -dimenzionalen mehanizem oblike $Y_{\mathbf{d}}(\omega) = (Y_{d_1}(\omega), \dots, Y_{d_n}(\omega))$ (ϵ, δ) -diferencirano zaseben za vsako poizvedbo Q .

Mehanizem za diskretne podatke

- D diskreten prostor, $|D| = m + 1$.
- Y_d mehanizem (slučajna spremenljivka):
 $\mathbb{P}(Y_d = d) = 1 - pm, \quad \mathbb{P}(Y_d = d') = p, \quad d, d' \in D$
- (ϵ, δ) –diferencirana zasebnost bo dosežena natanko tedaj, ko

$$p \geq \frac{1 - \delta}{m + e^\epsilon}.$$

Natančnost odzivnih mehanizmov

- $\gamma := \max_{d \in D} \mathbb{E}[\rho(Y_d, d)]$ – maksimalna pričakovana napaka danega mehanizma Y_d . D kompakten metričen prostor.

Izrek 4

Naj bo podana družina 1-dimenzionalnih diferencirano zasebnih mehanizmov $\{Y_d : \Omega \rightarrow U \mid d \in D\}$. Potem velja

$$\gamma \geq (1 - \delta) \left(\frac{\text{diam}(D)}{2(1 + e^\epsilon)} \right).$$

Izrek 5

Naj bo D diskreten metrični prostor z $|D| = m + 1$ in $\kappa = \min_{d, d' \in D} \rho(d, d')$. Naj bo podana družina 1-dimenzionalnih diferencirano zasebnih mehanizmov $\{Y_d : \Omega \rightarrow U \mid d \in D\}$. Potem velja

$$\gamma \geq (1 - \delta) \left(\frac{\kappa m}{(m + e^\epsilon)} \right).$$

Implementacija Laplaceovega mehanizma

ϵ	δ	b
0.1	0.1	14589
2	0.5	1112
11	0.7	245

Table 1: b predstavlja parameter Laplaceove porazdelitve

	povprečje	min	max
prvotni podatki	2995	1504	4500
(0.1 , 0.1)	3402	-111499	109729
(2, 0.5)	2999	-6110	11411
(11, 0.7)	2990	765	5360

Table 2: Vrednosti nekaterih osnovnih poizvedb pri različnih vrednostih parametrov (ϵ, δ).

Implementacija Laplaceovega mehanizma

(ϵ, δ)	(0.1, 0.1)	(2, 0.5)	(11, 0.7)
spodnja meja za γ	640	89	0.0075
dejanska največja napaka	110600	7203	1803

Table 3: Spodnja meja največje napake γ in dejansko opažena največja napaka pri različnih vrednostih parametrov (ϵ, δ) .

- $\Delta Q = \max_{\mathbf{d} \sim \mathbf{d}'} \|Q(\mathbf{d}) - Q(\mathbf{d}')\|_1$ (velika občutljivost identične poizvedbe).
- Uber (sistem za SQL poizvedbe).
- Skalirana metrika, problem?

Implementacija mehanizma za diskretne podatke

(ϵ, δ)	verjetnost, s katero podamo pravi odgovor
(0.1, 0.1)	0.12
(2, 0.5)	0.57
(7, 0.6)	0.98

Table 4: Rezultati mehanizma za diskretne podatke.

(ϵ, δ)			
prvotni podatki	TX (114)	CA (102)	NY (63)
(0.1, 0.1)	TX (35)	OH (31)	CA (30)
(2, 0.5)	TX (68)	CA (62)	FL (45)
(7, 0.6)	TX (113)	CA (102)	NY (62)

Table 5: Najpogostejše tri zvezne države v prvotnih podatkih in pri različnih vrednostih parametrov (ϵ, δ) .

Implementacija mehanizma za diskretne podatke

(ϵ, δ)	(0.1, 0.1)	(2, 0.5)	(7, 0.6)
spodnja meja za γ	0.879	0.432	0.016
dejanska največja napaka	0.880	0.437	0.018

Table 6: Podatki o spodnjih mejah pri diskretnih podatkih. Dejanska največja napaka je tu izračunana kot razmerje med številom nepravilnih odgovorov ter številom posameznikov v bazi.

Dodatno o funkcijskih podatkih

- $\mathbf{d} = (d_1, \dots, d_n)$, $d_i \in \mathbb{R}^m$ (d_i dobljeni iz porazdelitve z gostoto f)
- Jedrna cenilka za gostoto:

$$\hat{f}(x) = \frac{1}{n} \sum_1^n w\left(\frac{\|x - d_i\|}{h}\right), x \in \mathbb{R}^m$$

(W jedrna funkcija, h parameter dosega).

- $D = \mathbb{R}^m$, $E_Q \subseteq \mathbb{R}^D$, $Q(\mathbf{d}) = f_{\mathbf{d}}$, $X_{Q(\mathbf{d})} = \tilde{f}_{\mathbf{d}}$

Hilbertovi prostori z reprodukcijskim jedrom

Definicija

Naj bo \mathcal{H} Hilbertov prostor, katerega elementi so funkcije oblike $f : X \rightarrow \mathbb{R}$ (pri tem je X poljubna množica). Označimo z L_x linearen funkcional, ki vsako funkcijo $f \in \mathcal{H}$ izvrednoti v x , torej

$$L_x : f \rightarrow f(x).$$

Če je L_x zvezen nad \mathcal{H} za vsak $x \in X$, je \mathcal{H} Hilbertov prostor z reprodukcijskim jedrom.

- $f(x) = L_x(f) = \langle f, K_x \rangle_{\mathcal{H}}, \quad \forall f \in \mathcal{H}$ (Riesz)
- $K_x(y) = L_y(K_x) = \langle K_x, K_y \rangle_{\mathcal{H}}$
- Reprodukcijsko jedro $K : X \times X \rightarrow \mathbb{R}, \quad K(x, y) := \langle K_x, K_y \rangle_{\mathcal{H}}$

Gaussov proces

Definicija

Gaussov proces, parametriziran z indeksno množico T , je slučajni proces $\{X_t : t \in T\}$, za katerega velja, da je za vsak končni nabor točk $t_1, \dots, t_n \in T$ slučajni vektor

$$(X_{t_1}, \dots, X_{t_n})$$

porazdeljen večrazsežno normalno.

Gaussov proces je popolnoma določen s funkcijama povprečja in kovariance:

$$m(t) := \mathbb{E}X_t, \quad K(s, t) := \text{Cov}(X_s, X_t).$$

Izrek 6

Naj bo G trajektorija Gaussovega procesa s povprečjem 0 in kovariančno funkcijo K . Naj bodo $x_1, \dots, x_n \in D$. Naj bo matrika

$$M(x_1, \dots, x_n) = \begin{pmatrix} K(x_1, x_1) & \cdots & K(x_1, x_n) \\ \vdots & \ddots & \vdots \\ K(x_n, x_1) & \cdots & K(x_n, x_n) \end{pmatrix}$$

pozitivno definitna. Potem bo odzivni mehanizem

$$\tilde{f}_{\mathbf{d}} = f_{\mathbf{d}} + \sqrt{2 \log \frac{2}{\delta} \frac{\Delta}{\epsilon}} G$$

(ϵ, δ) -diferencirano zaseben, ko bo veljalo

$$\sup_{\mathbf{d} \sim \mathbf{d}'} \sup_{n < \infty} \sup_{(x_1, \dots, x_n) \in D^n} \left\| M^{-1/2}(x_1, \dots, x_n) \begin{pmatrix} f_{\mathbf{d}}(x_1) - f_{\mathbf{d}'}(x_1) \\ \vdots \\ f_{\mathbf{d}}(x_n) - f_{\mathbf{d}'}(x_n) \end{pmatrix} \right\|_2 \leq \Delta.$$

Izrek 7

Naj bo $f \in \mathcal{H}$, kjer je \mathcal{H} Hilbertov prostor z reprodukcijskim jedrom K . Za vsako x_1, \dots, x_n končno zaporedje različnih točk v \mathbb{R}^m , za katero je matrika

$$M(x_1, \dots, x_n) = \begin{pmatrix} K(x_1, x_1) & \cdots & K(x_1, x_n) \\ \vdots & \ddots & \vdots \\ K(x_n, x_1) & \cdots & K(x_n, x_n) \end{pmatrix}$$

pozitivno definitna, velja

$$\left\| M^{-1/2}(x_1, \dots, x_n) \begin{pmatrix} f(x_1) \\ \vdots \\ f(x_n) \end{pmatrix} \right\|_2 \leq \|f\|_{\mathcal{H}}.$$

Posledica

Naj E_Q podmnožica Hilbertovega prostora \mathcal{H} z reprodukcijskim jedrom K , ki je enak kovariančni funkciji Gaussovega procesa (povprečje procesa naj bo enako 0). Če z G označimo trajektorijo tega Gaussovega procesa, potem bo mehanizem

$$\tilde{f}_{\mathbf{d}} = f_{\mathbf{d}} + \sqrt{2 \log \frac{2}{\delta} \frac{\Delta}{\epsilon}} G$$

(ϵ, δ) –diferencirano zaseben, ko bo veljajo

$$\sup_{\mathbf{d} \sim \mathbf{d}'} \|f_{\mathbf{d}} - f_{\mathbf{d}'}\|_{\mathcal{H}} \leq \Delta.$$

Uporaba na primeru Gaussovega jedra

- $\mathbf{d} = (d_1, \dots, d_n)$, $d_i \in \mathbb{R}^m$ (d_i dobljeni iz porazdelitve z gostoto f)

- Cenilka za gostoto z Gaussovim jedrom:

$$f_{\mathbf{d}}(x) = \frac{1}{n(2\pi h^2)^{d/2}} \sum_{i=1}^n \exp\left\{\frac{-\|x-d_i\|_2^2}{2h^2}\right\}, \quad x \in \mathbb{R}^m$$

- $\mathbf{d} \sim \mathbf{d}'$:

$$(f_{\mathbf{d}} - f_{\mathbf{d}'})(x) = \frac{1}{n(2\pi h^2)^{d/2}} \left(\exp\left\{\frac{-\|x-d_n\|_2^2}{2h^2}\right\} - \exp\left\{\frac{-\|x-d'_n\|_2^2}{2h^2}\right\} \right)$$

- Za kovariančno funkcijo Gaussovega procesa $\{X_t : t \in T\}$ vzemimo Gaussovo jedro $K(x, y) = \exp\left\{\frac{-\|x-y\|_2^2}{2h^2}\right\}$

- $K(x, y) = K_x(y) = \exp\left\{\frac{-\|x-y\|_2^2}{2h^2}\right\}$, $f_{\mathbf{d}} - f_{\mathbf{d}'} = \frac{1}{n(2\pi h^2)^{d/2}} (K_{d_n} - K_{d'_n})$

- $\|f_{\mathbf{d}} - f_{\mathbf{d}'}\|_{\mathcal{H}}^2 \leq 2 \left(\frac{1}{n(2\pi h^2)^{d/2}} \right)^2$

Uporaba na primeru Gaussovega jedra

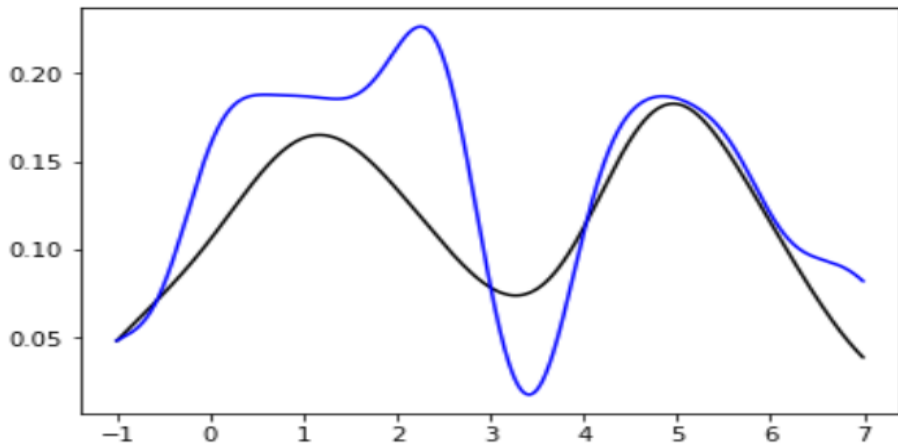


Figure 2: $\epsilon = 0.1, \delta = 0.1$

Uporaba na primeru Gaussovega jedra

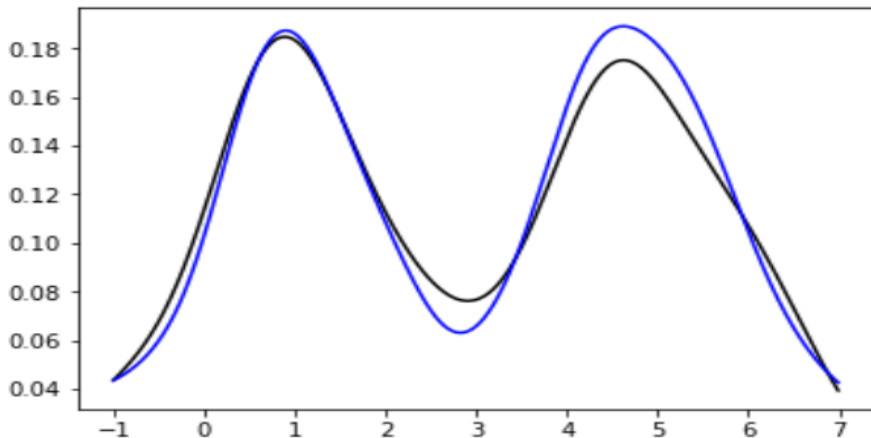


Figure 3: $\epsilon = 1, \delta = 0.1$