

Mešanje z naključnimi transpozicijami

Matevž Mišič

Mentor: doc. dr. Urban Jezernik

1 Uvod

Predstavljajmo si, da imamo pred seboj kupček z n kartami, na katerih so napisane številke od 1 do n . Na začetku so karte razporejene po velikosti, zato jih bomo premešali. Na vsakem koraku mešanja naključno izberemo dve karti iz kupčka. Če sta izbrani karti različni, ju zamenjamo med seboj, sicer pa kart ne premikamo. Lahko se vprašamo, koliko takšnih korakov moramo narediti, da bodo karte dobro premešane.

Vrstni red kart si lahko predstavljamo kot permutacijo prvih n naravnih števil, torej kot element grupe S_n , zamenjavo dveh kart, pa potem predstavlja množenje te permutacije z ustrezno transpozicijo z desne strani. Na vsakem koraku torej permutacijo pomnožimo s transpozicijo, ki ustreza menjavi izbranih kart, če sta ti dve karti različni. Če sta karti enaki, pa vrstni red ostane enak, kar pomeni, da permutacijo pomnožimo z identiteto. S pomočjo tega lahko ugotovimo verjetnosti vseh permutacij po izbranem številu korakov. Verjetnost permutacije σ po k korakih označimo s $P^k(\sigma)$.

Karte so dobro premešane, če so vse permutacije približno enako verjetne. V idealnem primeru bi si želeli, da je verjetnost vsake permutacije σ enaka $U(\sigma) = \frac{1}{n!}$. Ker bi radi znali oceniti tudi, kako dobro so karte premešane, uvedimo razdaljo $\|P - Q\| = \frac{1}{2} \sum_{\sigma \in S_n} |P(\sigma) - Q(\sigma)|$ med preslikavama $P, Q : S_n \rightarrow \mathbb{R}$. Manjša kot je razdalja $\|P^k - U\|$, boljše so karte premešane.

Želimo torej določiti zadostno število korakov k , da bo razdalja $\|P^k - U\|$ blizu 0. Izkaže se, da obstaja tak $c > 0$, da je razdalja velika, če je k manjši od $\frac{1}{2}n \log n - cn$, in zelo majhna, če je k večji od $\frac{1}{2}n \log n + cn$. To pomeni, da so karte slabo premešane, če naredimo malo manj kot $\frac{1}{2}n \log n$ in dobro premešane, če naredimo malo več kot toliko korakov. Temu pojavu rečemo odrez. V diplomski nalogi bomo dokazali, da v primeru mešanja z naključnimi transpozicijami pride do tega pojava.

2 Naključni prehodi po grupah

Začnimo z vpeljavo nekaj osnovnih pojmov. *Verjetnostna porazdelitev na množici* X je preslikava $P : X \rightarrow \mathbb{R}$, za katero velja $P(x) \geq 0$ za vse $x \in X$ in $\sum_{x \in X} P(x) = 1$. Namesto verjetnostna porazdelitev bomo včasih rekli kar verjetnost na množici X . Za vsak $a \in X$ imamo na primer porazdelitev δ_a na X , za katero je $\delta_a(a) = 1$ in $\delta_a(x) = 0$ za $x \neq a$. Drug primer porazdelitve je enakomerna porazdelitev U na množici X , ki je dana z $U(x) = \frac{1}{|X|}$ za vsak $x \in X$. Za verjetnostno porazdelitev P na X in podmnožico $A \subset X$ naj $P(A)$ označuje verjetnost, da se nahajamo v A , torej $P(A) = \sum_{a \in A} P(a)$.

Markovska veriga je proces, pri katerem se naključno premikamo med stanji iz neke množice Ω . Če smo na nekem koraku v stanju $x \in \Omega$, se v naslednjem koraku premaknemo v stanje $y \in \Omega$ z verjetnostjo $P(x, y)$. Ta pojem formalizira naslednja definicija.

Definicija 1. *Končna markovska veriga z množico stanj Ω , začetno porazdelitvijo S na množici Ω in prehodno matriko $P : \Omega \times \Omega \rightarrow \mathbb{R}$ je zaporedje naključnih spremenljivk $(X_0, X_1, \dots)$ z vrednostmi v Ω , za katero velja:*

- $\mathbb{P}(X_0 = x) = S(x)$ za vse $x \in \Omega$.
- Za vse $k \in \mathbb{N}_0$ in vse $x_0, x_1, \dots, x_k, y \in \Omega$ velja $\mathbb{P}(X_{k+1} = y \mid X_0 = x_0, \dots, X_k = x_k) = P(x_k, y)$.

Druga lastnost pove, da je stanje, v katerem smo po $k + 1$ korakih, odvisno le od stanja, v katerem se nahajamo po k korakih, ne pa tudi od tega, kako smo v to stanje prišli.

Nas ne bodo zanimale poljubne markovske verige ampak le takšne, ki predstavljajo naključne prehode po grupah.

Definicija 2. Naj bo P verjetnost na končni grupi G . *Naključni sprehod na G z verjetnostjo P* je markovska veriga z množico stanj G , pri kateri je začetna porazdelitev enaka δ_1 in za prehodno matriko $Q : G \times G \rightarrow \mathbb{R}$ velja $Q(x, y) = P(yx^{-1})$.

Sprehod na grupi se torej vedno začne v enoti grupe, verjetnost, da gremo v nekem koraku z x na hx , pa je enaka $P(h)$ ne glede na x . Sprehod po grupi torej izgleda enako v vsakem elementu grupe.

Naj bosta G in P kot v zgornji definiciji in naj $Q(x)$ označuje verjetnost, da se nahajamo v elementu $x \in G$. Zanima nas, kakšna je verjetnost, da bomo v naslednjem koraku v elementu $y \in G$. Verjetnost, da se nahajamo v x , je enaka $Q(x)$, verjetnost, da gremo iz tega elementa v y , pa je $P(yx^{-1})$, torej je iskana verjetnost enaka

$$\sum_{x \in G} P(yx^{-1})Q(x).$$

Definicija 3. *Konvolucija* funkcij $f, g : G \rightarrow \mathbb{R}$, je funkcija $f * g : G \rightarrow \mathbb{R}$ s predpisom

$$(f * g)(y) = \sum_{x \in G} f(yx^{-1})g(x).$$

Torej je verjetnost, da se po k korakih sprehoda nahajamo v x , enaka $(P^{*k} * \delta_1)(x) = (P^{*k})(x)$, kjer P^{*k} označuje konvolucijo k kopij P .

Mešanje n kart, kot je opisano v uvodu, je tako sprehod po grupi S_n , kjer za verjetnost vzamemo P_n , za katero velja $P_n(id) = \frac{1}{n}$, $P_n(\tau) = \frac{2}{n^2}$ za vsako transpozicijo τ in $P_n(\sigma) = 0$ za ostale permutacije σ .

3 Variacijska razdalja

Radi bi znali oceniti razdaljo od neke verjetnostne porazdelitve na grupi G do enakomerne. Zdi se nam, da sta dve porazdelitvi blizu, če za vsako podmnožico $A \subset G$ velja, da sta verjetnosti, da se nahajamo v A , za obe porazdelitvi podobni. Torej lahko razdaljo definiramo na naslednji način.

Definicija 4. *Variacijska razdalja* med verjetnostma P in Q na grupi G je

$$\|P - Q\| = \max_{A \subset G} |P(A) - Q(A)|$$

Ta definicija ni najprimernejša za določanje variacijske razdalje, saj bi morali načeloma obravnavati vse podmnožice $A \subset G$. Velja pa naslednja trditev, ki nam to olajša.

Trditev 1. *Za verjetnosti P in Q na G velja*

$$\|P - Q\| = \frac{1}{2} \sum_{x \in G} |P(x) - Q(x)|.$$

4 Odrez

Naj bo G grupa in P verjetnostna porazdelitev na njej. Opazujmo graf $\|P^{*k} - U\|$ v odvisnosti od k . Pojavu pri katerem tak graf v nekem ozkem oknu preide od vrednosti blizu 1 k vrednostim blizu 0, rečemo odrez. V tem razdelku bomo natančno definirali ta pojav, nato pa bomo formulirali glavni izrek, ki bo v diplomski nalogi dokazan.

Paru (G, P) , kjer je G neka grupa in P verjetnostna porazdelitev na njej, bomo rekli *grupa z verjetnostjo*.

Definicija 5. Naj bo (G_n, P_n) zaporedje grup z verjetnostmi. Pravimo, da ima to zaporedje *odrez*, če obstaja tako zaporedje pozitivnih realnih števil (t_n) z $\lim_{n \rightarrow \infty} t_n = \infty$, da za vsak $0 < \epsilon < 1$ velja:

- Za $k_n = \lfloor (1 - \epsilon)t_n \rfloor$ velja $\lim_{n \rightarrow \infty} \|P_n^{*k_n} - U_n\| = 1$,
- Za $k_n = \lceil (1 + \epsilon)t_n \rceil$ velja $\lim_{n \rightarrow \infty} \|P_n^{*k_n} - U_n\| = 0$.

Zavedati se je potrebno, da odrez ni lastnost le ene grupe z verjetnostjo, ampak zaporedja grup z verjetnostmi. Vrednosti t_n predstavljajo približno število korakov potrebnih, da graf spremeni vrednost. Ta sprememba se zgodi v nekem intervalu s središčem v t_n , ki mora biti za velike n zanemarljivo kratek v primerjavi s t_n . Čeprav zaporedje (t_n) ni enolično določeno, pogosto pravimo, da ima (G_n, P_n) odrez ob času (t_n) .

Naslednji izrek bo glavni rezultat diplomske naloge.

Izrek 1. Za vsak $n \in \mathbb{N}$ naj bo P_n verjetnost na grupi S_n , s predpisom $P_n(id) = \frac{1}{n}$, $P_n(\tau) = \frac{2}{n^2}$ za vsako transpozicijo τ in $P_n(\sigma) = 0$ za ostale permutacije σ . Potem za zaporedje grup z verjetnostmi (S_n, P_n) velja odrez ob času $\frac{1}{2}n \log n$.

Za dokaz izreka bomo razvili nekomutativno Fourierovo transformacijo, ki konvolucijo transformira v produkt transformirank. Produkt je veliko bolj oprijemljiv kot konvolucija in to nam bo pomagalo pri dokazu izreka. Za obravnavo te transformacije, pa moramo najprej spoznati upodobitve grup.

5 Upodobitve grup

V tem razdelku bomo spoznali pojma upodobitve grupe in nekomutativne Fourierove transformacije.

Definicija 6. Upodobitev grupe G je homomorfizem grup $\rho : G \rightarrow GL(V)$, kjer je $GL(V)$ grupa avtomorfizmov končnorazsežnega kompleksnega vektorskega prostora V . Dimenziji $\dim V$ rečemo *stopnja upodobitve* ρ .

Ker se bomo ukvarjali predvsem s simetričnimi grupami, si oglejmo nekaj upodobitev simetrične grupe S_3 .

1. Trivialen homomorfizem $\rho : G \rightarrow GL(V)$ v enodimenzionalen kompleksen vektorski prostor V je upodobitev grupe S_3 . Tej upodobitvi rečemo *trivialna upodobitev*.
2. Homomorfizem $\rho : G \rightarrow GL(V)$ v enodimenzionalen kompleksen vektorski prostor V , za katerega velja $\rho(\sigma)v = \text{sgn}(\sigma)v$ za vse $\sigma \in S_3$ in $v \in V$, je *alternirajoča upodobitev* grupe S_3 .

3. Naj bo zdaj V vektorski prostor z bazo $B = \{e_1, e_2, e_3\}$. Za $\sigma \in S_3$ bomo endomorfizem $\rho(\sigma)$ prostora V definirali na bazi B : $\rho(\sigma)e_j = e_{\sigma(j)}$. Ker $\rho(\sigma)$ preslika B na B , je surjektiven in zato avtomorfizem V . Za vsaka $\sigma, \pi \in S_3$ velja tudi $\rho(\sigma) \circ \rho(\pi) = \rho(\sigma\pi)$, torej je $\rho : S_3 \rightarrow GL(V)$ upodobitev. Takšni upodobitvi rečemo *permutacijska upodobitev grupe S_3* .

Definicija 7. Naj bo $f : G \rightarrow \mathbb{C}$ funkcija. *Fourierova transformiranka funkcije f* je preslikava $\widehat{f}$, ki upodobitev ρ grupe G preslika v linearno preslikavo $\widehat{f}(\rho) = \sum_{x \in G} f(x)\rho(x)$.

Ena od najpomembnejših lastnosti Fourierove transformacije je to, da je transformiranka konvolucije enaka produkt transformirank.

Trditev 2. Naj bosta $f, g : G \rightarrow \mathbb{C}$ funkciji na G in ρ upodobitev grupe G . Velja $\widehat{(f * g)}(\rho) = \widehat{f}(\rho)\widehat{g}(\rho)$.

Verjetnostna porazdelitev, ki jo dobimo po k korakih našega sprehoda po S_n je enaka P_n^{*k} , njena Fourierova transformiranka pa je po trditvi enaka $\widehat{P_n^{*k}} = \widehat{P_n}^k$. Izkaže se, da je za nekatere upodobitve ρ , ki jim rečemo *nerazcepne upodobitve*, $\widehat{P_n}(\rho)$ kar skalarni večkratnik identitete. Ker je takšne endomorfizme enostavno potencirati, lahko izračunamo $\widehat{P_n^{*k}}(\rho) = (\widehat{P_n}(\rho))^k$. Hkrati pa je P_n^{*k} enolično določena z vrednostmi svoje Fourierove transformirank v nerazcepnih upodobitvah, zato bi lahko na tak način ugotovili, kakšna je porazdelitev P_n^{*k} in jo primerjali z U_n . S pomočjo tega bi nato lahko pokazali, da res pride do odreza. V primeru mešanja z naključnimi transpozicijami se dobro vidi uporabnost nekomutativne Fourierove transformacije.

Literatura

- [1] D. A. Levin in Y. Peres, *Markov chains and mixing times*, American Mathematical Society, Providence, 2017.
- [2] P. Diaconis, *Group representations in probability and statistics*, Lecture Notes–Monograph Series **11**, Institute of Mathematical Statistics, Hayward, 1988.