Gerrymandering - krojenje izbornih jedinica i Metropolis-Hastings algoritam

Pleić, Nikolina

SVEUČILIŠTE U ZAGREBU

PRIRODOSLOVNO–MATEMATIČKI FAKULTET

MATEMATIČKI ODSJEK

Nikolina Pleić

GERRYMANDERING - KROJENJE IZBORNIH JEDINICA I METROPOLIS-HASTINGS ALGORITAM

Diplomski rad

Voditelj rada:
izv.prof.dr.sc. Matija Kazalicki

Zagreb, veljača, 2019.

Ovaj diplomski rad obranjen je dana ____________________ pred ispitnim povjerenstvom u sastavu:

1. ____________________________, predsjednik
2. ____________________________, član
3. ____________________________, član

Povjerenstvo je rad ocijenilo ocjenom ________________.

Potpisi članova povjerenstva:

1. ____________________________
2. ____________________________
3. ____________________________

Sadržaj

Uvod

Gerrymandering je politički pristrana podjela izbornih jedinica i kao takva predstavlja vrstu izbornog inženjeringa. Riječ je o obliku manipulacije kojim se granice izbornih jedinica svjesno prekrajaju prema određenim političkim interesima. Time se iskorištava zemljopisna raštrkanost birača različitih političkih stranaka kako bi se manipuliralo izbornim rezultatima. Sama riječ nastala je još 1812. godine u poznatom Boston Gazzette časopisu kada je opisan pokušaj senatora Elbridgea Gerryja da konstituira izborne jedinice tako da sebi i svojoj političkoj opciji osigura još jedan mandat. Tako je i nastala složenica dviju riječi – "salamander" mitsko čudovište i imena dotičnog senatora, a označava različite oblike zloupotrebe političke moći s promjenama izbornih jedinica u svrhu ostvarivanja političke prednosti.

Postoje dvije osnovne strategije gerrymanderinga. Prvo je "nagomilavanje" (eng. packing)– koncentriranje političke moći opozicije u jednu izbornu jedinicu kako bi im se smanjila politička moć u drugim izbornim jedinicama. Drugo je "razbijanje" (eng. cracking)- podjela političke moći oponenta na što više izbornih jedinica u cilju da se "razvodni" podrška.

Cilj ovog rada je dati pregled MCMC algoritma s pomičnom granicom i pokazati na koji se način može primjeniti u krojenju novih izbornih jedinica.
Općenito, MCMC algoritmi koriste se za simuliranje uzorka iz komplicirane distribucije koju znamo samo do na konstantu. Jedan od prvih MCMC algoritama razvio je 1953. Nicholas Metropolis te je algoritam po njemu i dobio ime. Generalizaciju Metropolis algoritma dao je W.K. Hastings 1970. godine, a algoritam je dobio ime Metropolis-Hastings. To je glavni algoritam koji ćemo obraditi u ovom radu i pokazati njegovu primjenu na problem izbornih raspodjela. U pozadini MCMC algoritama je teorija Markovljevih lanaca na općenitom skupu stanja. Stoga u prvom poglavlju uvodimo pojmove i rezultate iz te teorije koje ćemo u drugom poglavlju koristiti kako bismo dokazali svojstva Metropolis-Hastings algoritma. U trećem poglavlju matematički formaliziramo problem izbornih raspodjela te koristimo teorijske rezulate iz prva dva poglavlja kako bismo dokazali da su simulacije izbornih raspodjela generirane spomenutim algoritmom valjane.

Poglavlje 1

Markovljevi lanci na općenitom skupu stanja

U ovom poglavlju uvodimo osnovne pojmove teorije Markovljevih lanaca te navodimo rezultate koji će nam biti potrebni kako bismo uspostavili konvergenciju Metropolis-Hastings algoritma. Odmah na početku treba napomenuti kako ne promatramo Markovljeve lance s neprekidnim vremenom, obzirom da nas sama priroda simuliranja navodi na promatranje samo Markovljevih lanaca u diskretnom vremenu $(X_n)_{n\in\mathbb{N}}$.

Skup stanja označimo s $\mathcal{X}$, a s $\mathcal{B}(\mathcal{X})$ označimo σ-algebru na $\mathcal{X}$. Većina rezultata u ovom poglavlju vrijedi općenito, no mi ćemo pretpostavljati da je $\mathcal{X} = \mathbb{R}^m$ za neki $m \in \mathbf{N}$, a $\mathcal{B}(\mathcal{X})$ je Borelova σ-algebra $\mathcal{B}(\mathbb{R}^m)$ tj. σ-algebra generirana svim otvorenim skupovima u $\mathbb{R}^m$.

Markovljev lanac možemo zamišljati kao niz slučajnih skokova iz jednog stanja u drugo, pri čemu skok u novo stanje ovisi samo o trenutnom stanju u kojem se nalazimo. Stoga, za svako stanje u kojem se nalazimo, trebamo znati vjerojatnosnu distribuciju idućeg skoka, koja je dana sljedećom definicijom.

Definicija 1.0.1. *Prijelazna jezgra (eng. transition kernel) K je funkcija definirana na $\mathcal{X} \times \mathcal{B}(\mathcal{X})$ takva da vrijedi*
(i) za svaki $x \in \mathcal{X}$, $K(x, \cdot)$ je vjerojatnosna mjera;
(ii) za svaki $A \in \mathcal{B}(\mathcal{X})$, $K(\cdot, A)$ je izmjeriva funkcija na $\mathcal{X}$.

Kada je skup $\mathcal{X}$ diskretan tj. konačan ili prebrojiv, prijelazna jezgra je jednostavno matrica prijelaza

$$P_{xy} = P(X_n = y \mid X_{n-1} = x), \quad x, y \in \mathcal{X}$$

Ako je za svaki $x \in \mathcal{X}$ mjera $K(x, \cdot)$ apsolutno neprekidna u odnosu na Lebesgueovu mjeru λ^m, onda ćemo s $k(x, \cdot) : \mathcal{X} \to [0, \infty\rangle$ označavati uvjetnu gustoću s obzirom na mjeru

$K(X, \cdot)$, tj.

$$K(x, A) = \int_A k(x, y)\lambda^{(m)}(dy).$$

Takva funkcija postoji prema Radon-Nikodymovom teoremu za mjere. Ovaj slučaj zovemo neprekidan slučaj. Sada definiramo Markovljev lanac na skupu stanja $\mathcal{X}$.

Definicija 1.0.2. *Neka je $(\Omega, \mathcal{F}, \mathbb{P})$ vjerojatnosni prostor. Slučajni proces $X = (X_n : n \geq 0) : \Omega \to \prod_{i=0}^{\infty} X_i$ izmjeriv u odnosu na $(\mathcal{F}, \prod_{i=0}^{\infty} \mathcal{B}(\mathcal{X}_i))$ zovemo Markovljev lanac s prijelaznom jezgrom K ako, za svaki $n \in \mathbb{N}_0$ vrijedi*

$$P(X_{n+1} \in A \mid x_0, x_1, ..., x_n) = P(X_{n+1} \in A \mid x_n) = \int_A K(x_n, dx).$$

Vremenska homogenost znači da je distribucija od $(X_{n_1}, X_{n_2}, ..., X_{n_k})$ uz dano x_{n_0} ista kao i distribucija od $(X_{n_1-n_0}, X_{n_2-n_0}, ..., X_{n_k-n_0})$ uz dano x_0 za svaki $k \in \mathbb{N}$ i $n_0 \leq n_1 \leq ... \leq n_k$ prirodne brojeve. Formalno, to zapisujemo na način:

Definicija 1.0.3. *Neka je $(\Omega, \mathcal{F}, \mathbb{P})$ vjerojatnosni prostor. Slučajni proces $X = (X_n : n \geq 0) : \Omega \to \prod_{i=0}^{\infty} X_i$ izmjeriv u odnosu na $(\mathcal{F}, \prod_{i=0}^{\infty} \mathcal{B}(\mathcal{X}_i))$ zovemo vremenski homogen Markovljev lanac s prijelaznom jezgrom K i početnom distribucijom μ ako za svaki $n \in \mathbb{N}_0$ i za sve $A_0, A_1, ..., A_n \subset \mathbb{X}$ vrijedi*

$$\mathrm{P}(\mathrm{X}_0 \in A_0, X_1 \in A_1, ..., X_n \in A_n) = \int_{A_0} ... \int_{A_{n-1}} K(y_{n-1}, A_n)K(y_{n-2}, dy_{n-1})...K(y_0, dy_1)\mu(dy_0), \tag{1.1}$$

Kako ćemo u ovom poglavlju razmatrati samo vremenski homogene Markovljeve lance, izraz vremenski homogen zasada ćemo izostavljati. Može se pokazati da za svaku prijelaznu jezgru K i početnu distribuciju μ postoji vjerojatnosni prostor i na njemu slučajan proces tako da vrijedi (1).

Izraz $\prod_{i=0}^{\infty} \mathcal{B}(\mathcal{X}_i)$ zovemo produkt σ-algebri $\mathcal{B}(\mathcal{X}_i)$, pri čemu je $\mathcal{X}_i = \mathcal{X}$ za $i \in \mathbb{N}_0$. Ukoliko je $X : \Omega \to \mathcal{X}$ izmjeriva funkcija u odnosu na $(\mathcal{F}, \mathcal{B}(\mathcal{X}))$ te μ vjerojatnosna mjera na $(\mathcal{X}, \mathcal{B}(\mathcal{X}))$, tada ćemo s $X \sim \mu$ označavati da je zakon razdiobe od X jednak μ tj. da vrijedi

$$P(X \in A) = P_X(A) = \mu(A),$$

za svaki $A \in \mathcal{F}$. Važnu ulogu pri konstruiranju Markovljevog lanca osim prijelazne jezgre igra i početna mjera (početna distribucija). Ako je $X \sim \mu$, gdje je μ vjerojatnosna mjera na $(\mathcal{X}, \mathcal{B}(\mathcal{X}))$, tada kažemo da je μ početna distribucija lanca, a vjerojatnost P iz definicije

1.0.3. označavat ćemo s P_μ. Ukoliko lanac kreće iz stanja $x_0 \in \mathcal{X}$, što odgovara mjeri koncentriranoj u x_0 (Diracova mjera δ_{x_0}), onda ćemo pisati P_{x_0} umjesto $P_{\delta_{x_0}}$.

Primjer 1.0.1. *AR(1) proces*
Jednostavan primjer Markovljevog lanca na $\mathbb{R}$ možemo definirati na sljedeći način:

$$X_n = \theta X_{n-1} + \epsilon_n,\ \theta \in \mathbb{R},$$

pri čemu su $\epsilon_n \sim N(0, \sigma^2)$ međusobno nezavisne za svaki $n \in \mathbb{N}$. Za početno stanje uzmimo na primjer $X_0 = 0$. Početna distribucija je $\mu = \delta_0$. Jasno je da ako se nalazimo u stanju x, distribucija idućeg koraka je $N(\theta x, \sigma^2)$, čime smo dobili prijelaznu jezgru K.

Sada promatramo jednadžbu (1.1) koja nam govori kako računamo konačno dimenzionalne distribucije Markovljevog lanca. U specijalnom slučaju za neki $x \in \mathcal{X}$ i $A \in \mathcal{B}(\mathcal{X})$ izraze

$$\begin{aligned}
\mathbb{P}_x(X_1 \in A) &= K(x, A),\\
\mathbb{P}_x(X_2 \in A) &= \int_{\mathcal{X}} K(y, A) K(x, dy),\\
\mathbb{P}_x(X_n \in A) &= \underbrace{\int_{\mathcal{X}} \dots \int_{\mathcal{X}}}_{n-1 puta} K(y_{n-1}, A) K(y_{n-2}, dy_{n-1}) \dots K(y_1, dy_2) K(x, dy_1)
\end{aligned}$$

Sada možemo definirati prijelaznu jezgru u n koraka. Prvo označimo $K^1(x, A) = K(x, A)$. Definiramo za $n > 1$:

$$K^n(x, A) = \int_{\mathcal{X}} K^{n-1}(y, A) K(x, dy).$$

Vidimo da je $\mathbb{P}(X_n \in A) = K^n(x, A)$. Sada možemo definirati ireducibilnost Markovljevog lanca.

Definicija 1.0.4. *Neka je dana mjera ϕ na $\mathcal{B}(\mathcal{X})$. Markovljev lanac (X_n) s prijelaznom jezgrom K je **ϕ-ireducibilan** ako*

$$(\forall x \in \mathcal{X})\ (\forall A \in \mathcal{B}(\mathcal{X}))\ t.d.\ \phi(A) > 0 \implies (\exists n \in \mathbb{N})\ t.d.\ K^n(x, A) > 0.$$

*Lanac je **jako ϕ-ireducibilan** ukoliko je $n = 1$ za svaki $x \in \mathcal{X}$ i svaki $A \in \mathcal{B}(\mathcal{X})$.*

Riječima, lanac je ireducibilan u odnosu na mjeru ϕ ako za svaki skup koji nije mjere 0 postoji pozitivna vjerojatnost dolaska u taj skup u konačno mnogo koraka. Ireducibilnost

je bitna jer govori da lanac nije osjetljiv na početne uvjete (x_0 ili μ) jer ima pozitivnu vjerojatnost da posjeti sve skupove koji nisu mjere nula, u odnosu na ϕ, bez obzira iz kojeg stanja lanac krenuo. Ireducibilnost je krucijalna u postavljanju MCMC algoritama jer nas vodi k garanciji konvergencije (ka stacionarnoj distribuciji) i ujedno nas riješava brige o pogodnim početnim pozicijama iz kojih će naš lanac krenuti.
U diskretnom slučaju, Markovljev lanac je ireducibilan ako sva stanja komuniciraju, tj. ako

$$\mathbb{P}_x(\tau_y < \infty) > 0, \quad \forall x, y \in \mathcal{X}. \tag{1.2}$$

gdje je τ_y definirano s: $\tau_y = inf\{n \geq 1; X_n \in \{y\}\}$ i zovemo ga *vrijeme zaustavljanja* u stanju y. τ_y je dakle zapravo najmanji $n \geq 1$ za koji dostižemo stanje y. U mnogim slučajevima je $\mathbb{P}_x(\tau_y < \infty)$ uniformno jednako 0 stoga je nužno, kao što smo već napravili, uvesti pomoćnu mjeru ϕ na $\mathcal{B}(\mathcal{X})$ kako bi se točno definirao pojam ireducibilnosti.
Nadalje promatramo i broj dolazaka u skup A, definiran na način:

$$\eta_A = \sum_{n=1}^{\infty} \mathbb{1}_A(X_n).$$

Znamo da ireducibilni lanci imaju pozitivnu vjerojatnost dolaska u izmjerive skupove ne-nul mjere. Sljedeća definicija nam govori o očekivanom broju dolazaka u te skupove.

Definicija 1.0.5. *Markovljev lanac (X_n) je* ***povratan*** *ako*
(i) postoji mjera ψ takva da je (X_n) ψ-ireducibilan i
(ii) za svaki skup $A \in \mathcal{B}(\mathcal{X})$ takav da je $\psi(A) > 0$ vrijedi $\mathbb{E}_x[\eta_A] = \infty$ za svaki $x \in A$.

Za povratne lance vrijedi da ukoliko kreću iz izmjerivog skupa A ne-nul mjere, tada je očekivani broj povrataka u to stanje beskonačan. Povratnost lanca može se i ojačati. Naime, ukoliko još zahtijevamo i to da za svaku putanju lanca beskonačno puta posjetimo skup A dobivamo novu vrstu povratnosti koju je uveo Harris 1956. godine.

Definicija 1.0.6. *Markovljev lanac (X_n) je* ***Harris povratan*** *ako*
(i) postoji mjera ψ takva da je (X_n) ψ-ireducibilan i
(ii) za svaki izmjerivi skup A takav da je $\psi(A) > 0$ vrijedi $\mathbb{P}_x(\eta_A = \infty) = 1$ za sve $x \in A$.

U diskretnom slučaju tj. kada je skup stanja D konačan ili prebrojiv definiramo i pojam aperiodičnosti. Za stanje $i \in D$ period stanja i definira se kao najveći zajednički djelitelj skupa

$$\{n \geq 1 : \mathbb{P}_i(X_n = i) > 0\}. \tag{1.3}$$

Pokazuje se da je kod ireducibilnih lanaca period jednak za svako stanje. Ukoliko on iznosi 1, tada lanac zovemo **aperiodičnim**. Očito je da je dovoljan uvjet da ireducibilan lanac bude aperiodičan taj da je $\mathbb{P}_i(X_1 = i) > 0$ tj. da lanac može ostati u istom stanju.

Za lance na općenitom skupu stanja $\mathcal{X}$ definicija aperiodičnosti je dosta kompleksna pa je nećemo navoditi. Umjesto definicije dat ćemo dovoljan uvjet koji vrijedi i u diskretnom slučaju.

Napomena 1.0.2. *Ukoliko je Markovljev lanac ϕ-ireducibilan i vjerojatnost da ostane u istom stanju strogo veća od 0, onda je lanac* ***aperiodičan****.*

Marginalna distribucija lanca (X_n) je distribucija od X_n za $n \in \mathbb{N}$. Ako pretpostavimo da znamo $X_n \sim \pi$, pri čemu je π vjerojatnosna mjera, onda je:

$$\mathbb{P}_\mu(X_{n+1} \in A) = \mathbb{P}_\mu(X_{n+1} \in A, X_n \in \mathcal{X}) = \int_{\mathcal{X}} K(x, A)\pi(dx).$$

Viši stupanj stabilnosti Markovljevog lanca (X_n) dobivamo ukoliko marginalne distribucije lanca ne ovise o vremenu n. Formalnije, zahtijevamo postojanje vjerojatnosne mjere π t.d. $X_{n+1} \sim \pi$ ako $X_n \sim \pi$. Na ovom zahtjevu temelje se MCMC metode, a on ujedno definira i posebnu vrstu povratnosti koju nazivamo pozitivna povratnost. Markovljevi lanci konstruirani MCMC metodama gotovo uvijek imaju ovo svojstvo stabilnosti, osim u vrlo patološkim slučajevima. Sada uvodimo formalnu definiciju.

Definicija 1.0.7. *Za σ-konačnu mjeru π kažemo da je* ***invarijantna mjera*** *za prijelaznu jezgru K (i pridruženi lanac) ako vrijedi*

$$\pi(B) = \int_{\mathcal{X}} K(x, B)\pi(dx), \quad \forall B \in \mathcal{B}(\mathcal{X}).$$

Ukoliko je π vjerojatnosna mjera, invarijantnu mjeru često nazivamo ***stacionarna mjera*** *ili distribucija.*

Primjetimo da, ukoliko je π stacionarna distribucija lanca i uspijemo postići da je u nekom trenutku $X_n \sim \pi$, tada svaka iduća vrijednost lanca dolazi iz π distribucije što nam omogućuje da uzorkujemo iz te distribucije, iako uzorak neće nužno biti nezavisan. Za Markovljeve lance konstruirane MCMC metodom želimo provjeriti ovo svojstvo, no to će ponekad biti teško direktno provjeriti. Zato u nastavku definiramo dovoljan uvjet da bi lanac imao invarijantnu mjeru.

Definicija 1.0.8. *Neka je K prijelazna jezgra u neprekidnom slučaju, s uvjetnom gustoćom k. Markovljev lanac s prijelaznom jezgrom K zadovoljava* ***uvjet detaljne ravnoteže*** *ukoliko postoji funkcija $f : \mathcal{X} \to \mathbb{R}$ takva da vrijedi*

$$k(y, x)f(y) = k(x, y)f(x), \quad \text{za sve } (x, y) \in \mathcal{X} \times \mathcal{X}.$$

Uvjet detaljne ravnoteže nije nužan da f bude stacionarna distribucija lanca s prijelaznom jezgrom K, ali daje dovoljne uvjete koje nije teško provjeriti. Koristeći uvjet detaljne ravnoteže dokazat ćemo da lanac konstruiran MCMC algoritmima ima stacionarnu distribuciju f. U idućem teoremu dokazujemo da je uvjet detaljne ravnoteže zaista dovoljan za postojanje stacionarne distribucije.

Teorem 1.0.3. *Pretpostavimo da Markovljev lanac s prijelaznom jezgrom K zadovoljava uvjet detaljne ravnoteže s vjerojatnosnom funkcijom gustoće π. Tada je gustoća π invarijantna gustoća lanca.*

Dokaz. Napomenimo da, kada kažemo da je gustoća invarijantna za Markovljev lanac, mislimo na mjeru ν generiranu tom gustoćom, tj.

$$\nu(B) = \int_B \pi(x)\lambda^{(m)}(dx), \quad B \in \mathcal{B}(\mathcal{X}).$$

Za izmjerivi skup B računamo

$$\int_{\mathcal{X}} K(y,B)\nu(dy) = \int_{\mathcal{X}} K(y,B)\pi(y)\lambda^{(m)}(dy) =$$

$$\int_{\mathcal{X}}\int_B k(y,x)\pi(y)\lambda^{(m)}(dx)\lambda^{(m)}(dy) = \int_{\mathcal{X}}\int_B k(x,y)\pi(x)\lambda^{(m)}(dx)\lambda^{(m)}(dy) =$$

$$\int_B \pi(x)\,\underbrace{(\int_{\mathcal{X}} k(x,y)^{(m)}(dy))}_{K(x,\mathcal{X})=1}\,\lambda^{(m)}(dx) = \int_B \pi(x)\lambda^{(m)}(dx) = \nu(B),$$

pri čemu smo iskoristili uvjet detaljne ravnoteže u trećoj jednakosti te Fubinijev teorem u četvrtoj jednakosti, što je opravdano jer su podintegralne funkcije nenegativne. □

Promatrajući Markovljev lanac (X_n) s vremenske perspektive, prirodno je (i važno) odrediti prema čemu lanac konvergira. Postojanje (i jedinstvenost) invarijantne distribucije π čini tu distribuciju prirodnim kandidatom za graničnu distribuciju lanca. Sljedeći teorem daje dovoljne uvjete na lanac (X_n) kako bi njegova granična distribucija bila π, čime dobivamo da za velike n realizacije lanca možemo smatrati dobivenima iz distribucije π.
Norma u kojoj ćemo gledati konvergenciju je totalna varijacijska norma definirana izrazom:

$$\| \mu_1 - \mu_2 \|_{TV} = \sup_{A\in\mathcal{B}(\mathcal{X})} | \mu_1(A) - \mu_2(A) | .$$

Teorem 1.0.4. *Ako je Markovljev lanac* (X_n) *Harris povratan i aperiodičan sa stacionarnom mjerom* π*, onda vrijedi*

$$\lim_{n\to\infty} \left\| \int_{\mathcal{X}} K^n(x,\cdot)\mu(dx) - \pi \right\|_{TV} = 0,$$

za svaku početnu distribuciju μ*.*

Sada raspisujemo tvrdnju teorema. Iz definicije totalne varijacijske norme imamo:

$$\sup_{A\in\mathcal{B}(\mathcal{X})} \left| \int_{\mathcal{X}} K^n(x,A)\mu(dx) - \pi(A) \right| \longrightarrow 0, \quad n\to\infty,$$

iz čega slijedi

$$\forall A\in\mathcal{B}(\mathcal{X}), \quad \int_{\mathcal{X}} K^n(x,A)\mu(dx) \longrightarrow \pi(A), \quad n\to\infty,$$

tj.

$$\forall A\in\mathcal{B}(\mathcal{X}), \ \ \mathbb{P}_\mu(X_n\in A) \longrightarrow \pi(A), \quad n\to\infty.$$

Dakle, ovaj teorem nam govori da marginalne distribucije lanca, neovisno o početnoj distribuciji ili točki iz koje naš lanac kreće, konvergiraju prema stacionarnoj distribuciji. Za velike n marginalna distribucija od X_n aproksimira stacionarnu distribuciju π.

Dosada navedene tvrdnje važne su nam za definiranje MCMC metoda, no tvrdnje koje slijede su krucijalne za rad ovih algoritama. Naime, tvrdnje o konvergenciji koje smo do sada naveli odnose se samo na vjerojatnosnu mjeru $\mathbb{P}_x^n$ koja je zapravo "slika" lanca (X_n) u vremenu n. Dakle, ona određuje vjerojatnosna svojstva *prosječnog* ponašanja lanca u fiksiranom trenutku. Takva svojstva, iako nam pružaju opravdanja za metode simulacije, su od manje važnosti za kontrolu konvergencije dane simulacije, gdje su zapravo svojstva *realizacije* (x_n) lanca jedine karakteristike koje su nam zapravo bitne.
Stoga, moramo uočiti razliku između *vjerojatnosne analize*, koja opisuje prosječno ponašanje uzoraka, i *statističke analize*, koja svoje zaključke izvodi induktivno iz danog uzorka. Statistička analiza mora biti napravljena uvjetno na dani uzorak. Ovakvo razmatranje vodi k Bayesovskom pristupu. U teoriji Markovljevih lanaca, takva uvjetna analiza može iskoristiti svojstva konvergencije mjere $\mathbb{P}_x^n$ ka π samo kako bismo provjerili konvergenciju funkcija definiranih na promatranom dijelu lanca.
Problemi kod direktne primjene klasičnih teorema o konvergenciji (Zakon velikih brojeva, Centralni granični teorem, itd.) na uzorku $(X_1, ..., X_n)$ nastaju zbog Markovljeve ovisnosti između opažanja X_i i zbog ne-stacionarnosti niza (jedino u slučaju $X_0 \sim \pi$, gdje je π stacionarna distribucija, će lanac biti stacionaran). Stoga pretpostavljamo da lanac kreće iz

početnog stanja X_0, čija distribucija nije stacionarna distribucija lanca i promatramo nestacionaran lanac. Sljedeće tvrdnje su rezulati o konvergenciji ekvivalentni Zakonu velikih brojeva, a često se nazivaju i ergodski teoremi.

Kako bismo realizaciju lanca mogli koristiti kao uzorak dobiven iz distribucije π, potreban nam je *Ergodski teorem.* Uz dane opservacije $X_1, X_2, ..., X_n$ Markovljevog lanca, zanimat će nas ponašanje parcijalnih suma

$$S_n(h) = \frac{1}{n}\sum_{i=0}^{n} h(X_i),$$

pri čemu je h integrabilna funkcija s obzirom na odgovarajući prostor mjere.

Teorem 1.0.5. *(Ergodski teorem) Ako Markovljev lanac* (X_n) *ima* σ*-konačnu invarijantnu mjeru* π*, ekvivalentno je:*
(i) ako su $f, g \in L^1(\pi)$ *takve da* $\int_X g(x)\pi(dx) \neq 0$*, tada*

$$\lim_{n\to\infty} \frac{S_n(f)}{S_n(g)} = \frac{\int_X f(x)\pi(dx)}{\int_X g(x)\pi(dx)},$$

(ii) Markovljev lanac (X_n) *je Harris povratan.*

Ergodski teorem možemo koristiti tako da najprije dokažemo pretpostavke i tvrdnju *(ii)* teorema pa će tada vrijediti tvrdnja *(i)*. Markovljev lanac koji je aperiodičan i Harris povratan zovemo **ergodski** Markovljev lanac. Ako ergodski Markovljev lanac ima stacionarnu distribuciju π, onda na njega možemo primjeniti ergodski teorem (1.0.4) i (1.0.5), a to će nam biti ključno kod dokazivanja valjanosti MCMC algoritama.
Lako se vidi da, ukoliko je π vjerojatnosna mjera, onda je ona σ-konačna. U posebnom slučaju, kada je $g = 1$, imamo $\int_X g(x)\pi(dx) = 1$ pa slijedi

$$\lim_{n\to\infty} S_n(f) = \int_X f(x)\pi(dx) = \mathbb{E}_\pi[f(X)].$$

U izrazu $\mathbb{E}_\pi[f(X)]$ oznaka π znači da je varijabla X distribuirana prema mjeri π, tj. zakon razdiobe od X je $\mathbb{P}_X = \pi$. Ergodski teorem nam omogućuje da realizaciju lanca koristimo za izračunavanje raznih statistika vezanih uz distribuciju π, iako realizacija lanca ne daje nezavisan uzorak.
Teorem (1.0.4) nam daje rezultat asimptotske konvergencije marginalnih distribucija lanca prema stacionarnoj distribuciji. Precizniji zaključak o toj konvergenciji možemo dobiti ukoliko promatramo brzinu konvergencije K^n ka π. Razmatranje ove brzine bitno je za MCMC algoritme jer je brzina povezana s pravilima zaustavljanja ovih algoritama.

Definicija 1.0.9. *Ergodski Markovljev lanac sa stacionarnom distribucijom* π *je* ***uniformno ergodski*** *Markovljev lanac ako postoje* $\rho < 1$ *i* $M < \infty$ *takvi da za svaki* $x \in \mathcal{X}$ *vrijedi*

$$\|K^n(x, \cdot) - \pi(\cdot)\|_{TV} \leq M\rho^n,$$

za svaki $n \in \mathbb{N}$.

Iz definicije slijedi

$$\lim_{n\to\infty} \sup_{x\in\mathcal{X}} \|K^n(x, \cdot) - \pi(\cdot)\|_{TV} \leq \lim_{n\to\infty} M\rho^n = 0.$$

Slabije svojstvo od uniformne ergodičnosti lanca je geometrijska ergodičnost lanca.

Definicija 1.0.10. *Ergodski Markovljev lanac sa stacionarnom distribucijom* π *je* ***geometrijski ergodski*** *Markovljev lanac ako postoji* $\rho < 1$ *takav da*

$$\|K^n(x, \cdot) - \pi(\cdot)\|_{TV} \leq M(x)\rho^n,$$

za svaki $n \in \mathbb{N}$*, pri čemu je* $M(x) < \infty$ *za* π*-gotovo sve* $x \in \mathcal{X}$.

Poglavlje 2

Metropolis-Hastings algoritam

2.1 Općenite Monte Carlo metode

Prije no što detaljnije uđemo u proučavanje Metropolis-Hastings algoritma, ukratko navodimo motivaciju i ideje koje stoje iza Monte Carlo simulacijskih metoda.
Prisjetimo se, problem je aproksimirati integrale oblika:

$$\mathbb{E}_f[h(X)] = \int_X h(x)f(x)\lambda^m(dy) \tag{2.1}$$

Prirodno je promatrati uzorak $(X_1, ..., X_m)$ iz distribucije zadane gustoćom f te aproksimirati (2.1) uzoračkom sredinom:

$$\bar{h}_m = \frac{1}{m}\sum_{j=1}^{m} h(x_j)$$

s obzirom da $\bar{h}_m$ konvergira gotovo sigurno k $\mathbb{E}_f[h(X)]$ prema Jakom zakonu velikih brojeva. Nadalje, ako h^2 ima konačan drugi moment uz danu gustoću f, tada se brzina konvergencije od $\bar{h}_m$ može procijeniti obzirom da se varijanca:

$$var(\bar{h}_m) = \frac{1}{m}\int_X (h(x) - \mathbb{E}_f[h(X)])^2 f(x)\lambda^m(dy)$$

također može procijeniti iz uzorka $(X_1, ..., X_m)$ sa:

$$v_m = \frac{1}{m}\sum_{j=1}^{m}[h(x_j - \bar{h}_m]^2.$$

Za velike m je dakle

$$\frac{\bar{h}_m - \mathbb{E}_f[h(X)]}{\sqrt{v_m}}$$

približno distribuiran kao jedinična normalna varijabla $\mathcal{N}(0, 1)$.

2.2 MCMC metode

Markov Chain Monte Carlo tj. MCMC metode nam omogućuju aproksimaciju integrala

$$\mathbb{E}_f[h(X)] = \int_X h(x)f(x)\lambda^m(dy)$$

bez direktnog uzorkovanja iz distrbucije f. Naime, koristeći pomalo drugačiju strategiju, možemo dobiti uzorak $X_1, ..., X_n$ koji je približno iz distribucije f, bez da direktno uzorkujemo iz f. Osnovni princip u pozadini ove strategije je korištenje *ergodskog Markovljevog lanca s stacionarnom distribucijom f.*

Definicija 2.2.1. *MCMC algoritam za simuliranje uzoraka iz distribucije dane gustoćom f je svaka metoda koja generira ergodski Markovljev lanac (X_n) čija je stacionarna distribucija f.*

2.3 Metropolis-Hastings algoritam

U ovom poglavlju objasnit ćemo najopćenitiji algoritam za simuliranje uzorka iz distribucije zadane vjerojatnosnom funkcijom gustoće f koristeći Markovljeve lance - Metropolis-Hastings algoritam. Razvio ga je Nicholas Metropolis 1953. godine, a njegovu generalizaciju je 1970. godine dao W.K. Hastings te je algoritam dobio ime Metropolis-Hastings.
Želimo simulirati uzorak iz distribucije zadane gustoćom f koju znamo do na normalizirajuću konstantu kako bismo mogli aproksimirati integrale oblika

$$\mathbb{E}_f[h(X)] = \int_X h(x)f(x)\lambda^m(dy). \tag{2.2}$$

Konstruirat ćemo ergodski Markovljev lanac čija je stacionarna distribucija f. Da bismo to napravili, potrebno je konstruirati prijelaznu jezgru kojoj je f stacionarna distribucija i osigurati konvergenciju marginalnih distribucija lanca (X_n) prema f, te osigurati uvjete Ergodskog teorema 1.0.5. Na kraju ćemo dobiti aproksimaciju integrala (2.2) oblika

$$\hat{I}_N = \frac{1}{N}\sum_{n=1}^{N} h(X_n). \tag{2.3}$$

2.4 Algoritam

Metropolis-Hastings algoritam započinje s ciljnom (eng. *target*) distribucijom f tj. distribucijom iz koje želimo simulirati uzorak. Nadalje, za svako stanje $x \in \mathcal{X}$ u kojem se lanac može naći, moramo definirati vjerojatnosnu mjeru $K(x,\cdot)$. Za svaki x zadajemo uvjetnu funkciju gustoće $q(y|x)$ koju nazivamo prijedložna distribucija ili gustoća (engl. *proposal*). Tada vrijedi

$$K(x,A) = \int_A q(y|x)\lambda^m(dy), \quad A \in \mathcal{B}(\mathcal{X}).$$

Kako bi algoritam funkcionirao, nužno je da znamo simulirati iz prijedložne distribucije te da znamo izračunati omjer

$$\frac{f(y)}{q(y|x)}$$

do na konstantu koja ne ovisi o x. Algoritam generira Markovljev lanac $(X^{(t)})$ kroz sljedeće korake:

Metropolis-Hastings algoritam

1. Uz dano $X_n = x_n$
2. Generiraj $Y_n \sim q(\cdot|x_n)$
3. Generiraj $U \sim U([0,1])$
4. Izračunaj $\rho(x_n, Y_n)$, pri čemu je

$$\rho(x,y) = min\left\{\frac{f(y)}{f(x)}\frac{q(x|y)}{q(y|x)}, 1\right\}$$

5. Definiraj

$$X_{n+1} = \begin{cases} Y_n & ako\ U \leq \rho(x_n, Y_n) \\ x_n & inače \end{cases}$$

6. Ponavljaj za $n = n + 1$

Ovaj algoritam uvijek prihvaća vrijednosti Y_n za koje je omjer $f(Y_n)/q(Y_n|x_n)$ veći nego za prethodnu vrijednost $f(x_n)/q(x_n|Y_n)$.
Poseban slučaj imamo ako je *proposal* distribucija $q(\cdot|x)$ simetrična tj. $q(x|y) = q(y|x)$, tada prihvaćanje novog stanja ovisi samo o omjeru $f(Y_n)/f(x_n)$ i u tom slučaju radi se o Metropolis algoritmu.

Metropolis-Hastings algoritam ovisi samo o omjerima:

$$\frac{f(Y_n)}{f(x_n)} \quad i \quad \frac{q(x_n|Y_n)}{q(Y_n|x_n)}$$

pa stoga ne ovisi o normalizirajućim konstantama, tj. gustoću f je dovoljno znati samo do na konstantu. Također je i *prijedložnu* distribuciju q dovoljno znati samo do na konstantnu koja ne ovisi o x. Ipak, uvjetne gustoće sami zadajemo pa ih najčešće u potpunosti znamo, no to što nam u algoritmu treba samo njihov omjer može biti prednost pri implementaciji.

Algoritam započinje iz stanja x_0 koje biramo proizvoljno, ali tako da vrijedi $f(x_0) > 0$. Zatim generiramo vrijednost Y_n iz *prijedložne* distribucije. Izraz $\rho(x, y)$ je vjerojatnost da ćemo prihvatiti vrijednost iz *prijedložne* distribucije. Ukoliko ne prihvatimo novo stanje, ostajemo u trenutnom stanju i u idućem koraku. Algoritam zaustavljamo kada smatramo da je lanac napravio dovoljno koraka za naše potrebe.

Očito je vjerojatnost $\rho(x_n, Y_n)$, koju nazivamo i *Metropolis-Hastings vjerojatnost prihvaćanja* definirana samo za x_n za koje vrijedi $f(x_n) > 0$. Štoviše, ako lanac kreće iz stanja x_0 takvog da je $f(x_0) > 0$, slijedi da je $f(x_n) > 0$ za svaki $n \in \mathbb{N}$ obzirom da vrijednosti Y_n takve da je $f(Y_n) = 0$ vode k $\rho(x_n, Y_n) = 0$ i stoga ih algoritam odbacuje. Dogovorno uzimamo da je $\rho(x, y) = 0$ kada su i $f(x)$ i $f(y)$ jednake nula, kako bismo izbjegli teoretske poteškoće.

2.5 Svojstva Metropolis-Hastings algoritma

Iako je ovaj algoritam općenit, tj. definiran za proizvoljne f i q, ipak je potrebno ispuniti neke minimalne uvjete kako bi f bila stacionarna distribucija konstruiranog lanca (X_n).
Ukoliko nosač gustoće f, $supp f = \{x \in X : f(x) > 0\}$, nije povezan, važno je da q omogućuje povezivanje svih komponenti, tj. omogućuje lancu skok iz jedne komponente u drugu. Ako postoji skup $A \subset supp f$ takav da

$$\int_A f(x)\lambda^{(m)}(dx) > 0 \quad i \int_A q(y|x)\lambda^{(m)}(dy) = 0, \quad \textit{za svaki} \;\; x \in supp f$$

onda f nije stacionarna distribucija lanca dobivenog algoritmom, obzirom da za $x_0 \notin A$ lanac (X_n) nikada neće posjetiti skup A, a uzorak iz f može sadržavati vrijednost iz skupa A.
Dovoljan je uvjet da nosač od q sadrži nosač od f tj.

$$\bigcup_{x \in supp\, f} supp\; q(\cdot|x) \supset supp\; f.$$

Teorem 2.5.1. *Neka je (X_n) lanac dobiven Metropolis-Hastings algoritmom. Ako nosač od q sadrži nosač od f, tada vrijedi:*

(i) prijelazna jezgra K lanca (X_n) zadovoljava uvjet detaljne ravnoteže sa f,

(ii) f je stacionarna distribucija lanca.

Dokaz. Za dokaz tvrdnje *(i)* trebamo pokazati

$$k(y,x)f(y) = k(x,y)f(x), \quad za\ sve\ \ x,y \in \mathcal{X}.$$

Prisjetimo se da je $k(x,y)$ uvjetna gustoća prijelaza $K(x,\cdot)$ ukoliko se radi o neprekidnom slučaju. Za proizvoljan $A \in \mathcal{B}(\mathcal{X})$ i $x \in \mathcal{X}$ računamo

$$\begin{aligned}
&K(x,A) = \mathbb{P}_x(X_1 \in A) = \\
&\mathbb{P}_x(X_1 \in A,\ prihvatimo\ Y) + \mathbb{P}_x(X_1 \in A,\ odbijemo\ Y) = \\
&\mathbb{E}[\rho(x,Y)\mathbb{1}_{Y\in A}] + \mathbb{1}_A(x)\mathbb{P}_x(odbijemo\ Y) = \\
&\int_A \rho(x,y)q(y|x)\lambda^{(m)}(dy) + \underbrace{\delta_x(A)}_{\int_A \delta_x}\Big(1 - \underbrace{\int_A \rho(x,y)q(y|x)\lambda^{(m)}(dy)}_{r(x)}\Big),
\end{aligned}$$

pri čemu je $Y \sim q(\cdot|x)$ predložena vrijednost, a δ_x Diracova mjera koncentrirana u točki x. Dakle, imamo

$$\begin{aligned}
K(x,A) = \int_A \rho(x,y)q(y|x)\lambda^{(m)}(dy) + \int_A (1 - r(x))\delta_x(dy) = \\
\int_A \rho(x,y)q(y|x)\lambda^{(m)}(dy) + \int_A (1 - r(x))\delta_x(y)\lambda^{(m)}(dy),
\end{aligned}$$

pri čemu je sada δ_x Diracova delta "funkcija" koju shvaćamo kao:

$$\delta_x(y) = \begin{cases} +\infty & ako\ y = x \\ 0 & ako\ y \neq x \end{cases}$$

i vrijedi

$$\int_{\mathcal{X}} \delta_x(y)\lambda^{(m)}(dy) = 1.$$

Ovo valja shvatiti samo kao notaciju jer Diracova mjera nije apsolutno neprekidna u odnosu na Lebesgueovu mjeru pa ne postoji funkcija koja bi zadovoljavala gornju jednakost kod integrala.

Na kraju dobivamo uvjetnu gustoću prijelaza:

$$k(x, y) = \rho(x, y)q(y|x) + (1 - r(x))\delta_x(y).$$

Sada provjeravamo uvjet detaljne ravnoteže:

$$k(x, y)f(x) = \rho(x, y)q(y|x)f(x) + (1 - r(x))\delta_x(y)f(x).$$

Vrijedi $\rho(x, y) < 1 \implies \rho(y, x) = 1$ pa u tom slučaju imamo

$$\frac{f(y)}{f(x)}\frac{q(x|y)}{q(y|x)}q(y|x)f(x) = f(y)q(x|y) = f(y)q(x|y)\rho(y, x)$$

Ako je $\rho(x, y) = 1$, onda je $\rho(y, x) = \frac{f(x)}{f(y)}\frac{q(y|x)}{q(x|y)}$ pa imamo

$$\rho(x, y)q(y|x)f(x) = q(y|x)f(x) = q(y|x)f(x)\frac{f(y)}{f(y)}\frac{q(x|y)}{q(x|y)} = \rho(y, x)q(x|y)f(y).$$

Jasno je da vrijedi:

$$(1 - r(x))\delta_x(y)f(x) = (1 - r(y))\delta_y(x)f(x),$$

jer je $\delta_x(y) = 0$ ako je $x \neq y$. Dakle, vrijedi uvjet detaljne ravnoteže čime smo dokazali *(i)*. Dio *(ii)* slijedi iz teorema (1.0.3). □

Činjenica da je f stacionarna distribucija lanca povlači da ukoliko je $X_n \sim f$, tada je $X_{n+1} \sim f$, no zanima nas hoće li ikada doći do toga da u nekom koraku bude $X_n \sim f$? Idući korak nam je pokazati da Markovljev lanac generiran Metropolis-Hastings algoritmom uistinu konvergira ka stacionarnoj distribuciji f. Najprije moramo osigurati to da je lanac f-ireducibilan tj. da vrijedi

$$\forall x \in \mathcal{X}, \quad \forall A \in \mathcal{B}(\mathcal{X}), \quad \int_A f(x)\lambda^{(m)}(dx) > 0 \implies \exists n \in \mathbb{N}, \;\; K^n(x, A) > 0.$$

Dakle, za svaki skup mjere veće od 0, u odnosu na mjeru definiranu preko gustoće f, vjerojatnost dolaska u taj skup u konačno mnogo koraka mora biti veća od 0. Ako izaberemo *prijedložnu* distribuciju q tako da vrijedi

$$q(y|x) > 0, \quad za\ sve\ (x, y) \in supp\ f \times supp\ f,$$

tada u svaki skup ne-nula mjere možemo doći u jednom koraku, tj. naš lanac je f-ireducibilan.

Lema 2.5.2. *Ako je lanac dobiven Metropolis-Hastings algoritmom f-ireducibilan, onda je i Harris povratan.*

Obzirom da Metropolis-Hastings algoritam omogućuje lancu ostanak u istom stanju, lanac je i aperiodičan. Sada smo zadovoljili sve uvjete potrebne za teoreme iz prethodnog poglavlja. Rezultate ćemo objediniti u sljedećem teoremu koji navodimo bez dokaza.

Teorem 2.5.3. *Ukoliko je lanac (X_n) dobiven Metropolis-Hastings algoritmom f-ireducibilan, tada vrijedi:*

(i) f je stacionarna distribucija lanca (X_n),

(ii) za svaki $A \in \mathcal{B}(\mathcal{X})$ i svaku početnu distribuciju μ

$$\mathbb{P}_\mu(X_n \in A) \longrightarrow \int_A f(x)\lambda^{(m)}(dx), \ \ n \to \infty,$$

(iii) ako je $h \in L^1(\pi)$, pri čemu je π mjera generirana s f, tada

$$\hat{I}_N = \frac{1}{N}\sum_{n=1}^{N} h(X_n) \longrightarrow \int_{\mathcal{X}} h(x)f(x)\lambda^{(m)}(dy), \quad n \to \infty,$$

tj. (2.3) *je aproksimacija integrala* (2.2).

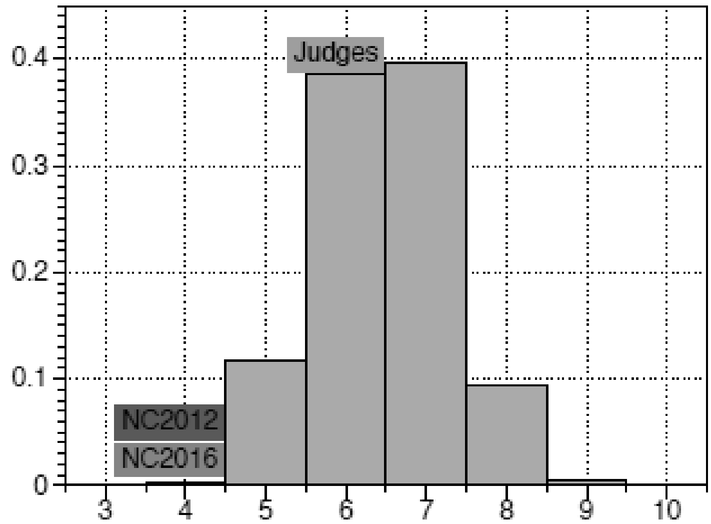
Poglavlje 3

Gerrymandering

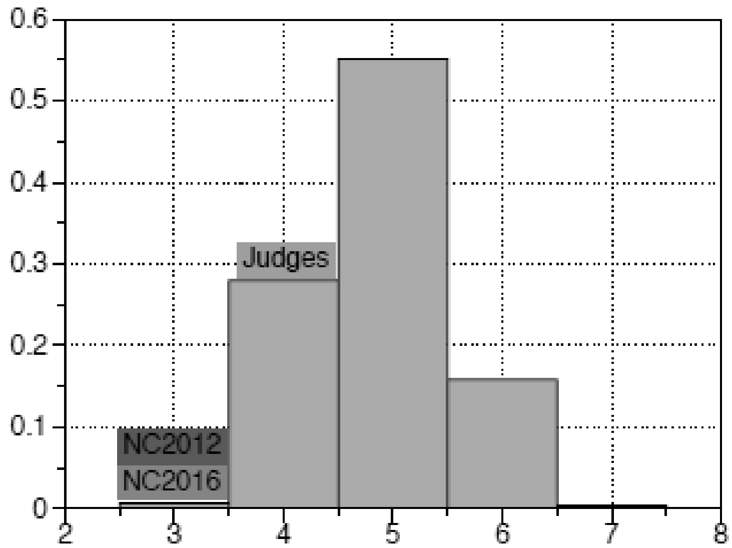
Gerrymandering je oblik manipulacije kojim se granice izbornih jedinica svjesno prekrajaju prema određenim političkim interesima. U ovom poglavlju navodimo jedan poznat primjer političkog gerrymanderinga kako bismo ilustrirali srž problema. Zatim predlažemo način na koji možemo stvarati nove izborne raspodjele koje bi poštovale određene zakonske odredbe. To postižemo koristeći Metropolis-Hastings algoritam objašnjen u prethodnom poglavlju.

3.1 Vizualizacija gerrymanderinga

Danas možda i najpoznatiji primjer gerrymanderinga tj. izbornog inženjeringa je onaj američke države North Caroline. Na izborima za američki kongres 2012. godine, koji su se temeljili na izbornim distriktima iz 2010. godine, 4 od 13 mjesta u Kongresu osvojili su Demokrati. No, s druge strane, na razini cijele države, većina glasova dana je upravo Demokratima. Upravo ovo odstupanje ponukalo je autore članka [1] da se zapitaju u kojoj mjeri rezultati izbora ovise o izbornim distriktima. Također, postavlja se pitanje koji su rezulati prosječnog izbora distrikta i kada bi se izborni plan trebao smatrati van norme? Autori stoga koriste stvarne glasove iz 2012. i 2016. i promatraju kako mijenjanje izbornih distrikta utječe na rezulate izbora. Kako bi mijenjali granice distrikta, iskoristili su upravo MCMC algoritam koji će biti objašnjen u ovom poglavlju. Generirali su 24000 nasumičnih, ali smislenih izbornih planova. Zatim su novim izbornim planovima pridružili stvarne glasove iz 2012. i 2016. kako bi dobili po 24000 izbornih rezultata za svaku izbornu godinu. Označit s ćemo NC2012 i NC2016 izborne planove North Caroline 2012. i 2016., redom. Izraz ”Judges” ovdje predstavlja rezultate izbora za distrikte kreirane od strane nepristranih umirovljenih sudaca okupljenih pod ”Beyond Gerrymandering” projektom.



Slika 3.1: Broj izabranih Demokrata (glasovi iz 2012.)



Slika 3.2: Broj izabranih Demokrata (glasovi iz 2016.)

Za stvarne glasove iz 2012. godine, i NC2012 i NC2016 rezultiraju s 4 demokratska mjesta u Kongresu, rezultat koji se javlja u manje od 0.3% od 24000 generiranih izbornih planova. Judges izborni plan rezultira izborom 6 demokrata, što se javlja u više od 39%

izbornih planova. Za glasove iz 2016., NC2012 i NC2016 rezultiraju s 3 demokratska mjesta u Kongresu, rezultat koji se javlja u manje od 0.7% generiranih izbornih planova. Judges izborni plan rezultira izborom 6 demokrata, što se javlja u 28% izbornih planova.

Jasno je da su izborni planovi koji proizvode rezultate koji su značajno različiti od prosječnih rezultata dobivenih iz nasumično uzorkovanih izbornih planova nedvojbeno u kontradikciji s ”voljom naroda” izraženom kroz stvarne glasove.

3.2 Krojenje izbornih jedinica

Obzirom da smo ograničeni zakonskim odredbama pri simuliranju novih raspodjela izbornih jedinica (dalje: izbornih planova), definiramo familiju vjerojatnosnih distribucija na prostoru svih izbornih planova prvo definirajući evaluacijsku funkciju na proizvoljnom izbornom planu. Ova funkcija vraća manji rezultat za one planove koji bolje poštuju zadane kriterije. Koristit ćemo ovu evaluacijsku funkciju kako bismo definirali vjerojatnosnu mjeru na prostoru svih izbornih planova.
Ideja generiranja izbornih jedinica u SAD-u je prisutna od 1960. Postoje tri glavne klasifikacije algoritama: konstruktivni randomizirajući algoritmi, optimizacijski algoritmi i MCMC algoritmi s pomičnom granicom.
Konstruktivni randomizirajući algoritmi započinju svaku novu raspodjelu s početnim nasumičnim ’seed’-om te tada ili generiraju fiksiran broj jedinica ili spajaju male podjedinice dok se ne postigne željeni broj izbornih jedinica.
Optimizacijski algoritmi ne uzorkuju iz prostora svih izbornih raspodjela, već se prvenstveno fokusiraju na generiranje jedne ili kolekcije ”elitnih” izbornih jedinica.
MCMC algoritmi s pomičnom granicom traže nove raspodjele tako da mijenjaju granice postojećih jedinica kako bi uzorkovali iz određene distribucije izbornih planova. MCMC algoritmi su, u načelu, bolji u uzorkovanju od konstruktivnih algoritama. MCMC algoritmi teoretski mogu simulirati prostor izbornih planova s točnom vjerojatnosnom distribucijom, dok konstruktivni algoritmi mogu konstruirati puno sličnih izbornih planova, a pri tom ne simulirati druge vrste jednako valjanih izbornih planova, što vodi k pomaknutoj (skewed) distribuciji. Kod konstruktivnih algoritama, nismo sigurni koju distribuciju algoritam generira. Jedna prednost MCMC pristupa naprema svim drugima jest da MCMC pristup simulira uzorak iz eksplicitno određene i konstruirane vjerojatnosne distribucije na izbornim planovima.
Korištenje MCMC metode zahtijeva da definiramo familiju vjerojatnosnih distribucija na prostoru svih izbornih planova.

Definiranje evaluacijske funkcije

Kako bismo definirali evaluacijsku funkciju, uvodimo nekoliko matematičkih formalizama. Inicijalno stanje ξ bi u ovom slučaju bila podjela RH na izborne jedinice iz 2016. godine. Prvo, reprezentirat ćemo RH kao graf G s bridovima E i vrhovima V. Svaki vrh $v \in V$ predstavlja jednu jedinicu lokalne samouprave (grad ili općinu), brid $e \in E$ između dva vrha iz V postoji ako te dvije jedinice lokalne samouprave dijele granice s ne-nula duljinom. Ovakvo definiranje grafa $G(V, E)$ omogućava nam da formalno definiramo raspodjelu izbornih jedinica (dalje: izborni plan).

Pretpostavljamo da svaka jedinica lokalne samouprave pripada samo jednoj izbornoj jedinici. Ovdje se javlja problem grada Zagreba koji je u aktualnom izbornom planu podijeljen među 4 izborne jedinice, stoga ćemo ta 4 dijela grada Zagreba nadalje promatrati kao 4 zasebne jedinice lokalne samouprave. Izborni plan je definiran kao fukcija sa skupa vrhova V u skup svih mogućih izbornih jedinica, koje su ovdje predstavljene kao niz cijelih brojeva od 1 do 10 obzirom da na teritorijalnom području cijele Hrvatske postoji 10 izbornih jedinica isključujući dijasporu i manjine. Stoga je izborni plan defniran kao fukcija $\xi \colon V \to \{1, ..., 10\}$. Funkciju izbornog plana interpretiramo na sljedeći način: Ako je jedinica lokalne samouprave označena kao vrh $v \in V$, onda $\xi(v) = i$ znači da ta jedinica lokalne samouprave pripada i izbornoj jedinici; slično za proizvoljan $i \in \{1, 2, ..., 10\}$ i plan ξ, i-ta izborna jedinica, oznaka $D_i(\xi)$ je zapravo skup $\{v \in V \colon \xi(v) = i\}$. Ograničavamo skup svih izbornih planova tako da za svaku izbornu jedinicu $D_i(\xi)$ vrijedi da je ona jedna povezana komponenta grafa; ovaj uvjet, zajedno s uvjetom na bridovima osigurava da su zakonski zahtjevi uvijek zadovoljeni.

Kolekciju svih izbornih planova s povezanim izbornim jedinicama označimo s R. Izborni plan ξ vrednujemo s našom evaluacijskom funkcijom, oznaka J. J svakom izbornom planu $\xi \in R$ pridružuje nenegativan broj. Manja vrijednost evaluacijske funkcije ukazuje na to da je izborni plan dobro prilagođen zakonskim zahtjevima. Kako bismo definirali evaluacijsku fukncijU J, uvodimo nekoliko pod-funkcija koje će mjeriti koliko dobro izborni plan zadovoljava pojedinačne zakonske zahtjeve.

Označimo ove pod-funkcije s J_p, J_I, J_c; *populacijski parametar* $J_p(\xi)$ mjeri koliko dobro izborni plan ξ dijeli populaciju RH na 10 jednakih populacijskih grupa; *izoperimetrijski parametar* $_I(\xi)$ mjeri kompaktnost izbornih jedinica; *županijski parametar* $J_c(\xi)$ mjeri broj županija koje su podijeljene između više izbornih jedinica.

Jednom kad su ove pod-funkcije određene, naša evaluacijska funkcija J je definirana kao težinska suma J_p, J_I i J_c. Kako nam sve pod-funkcije nisu na istoj skali, koristimo njihovu težinsku linearnu kombinaciju kako bismo uravnotežili efekt svakog od kriterija. Dakle, definiramo:

$$J(\xi) = w_p J_p(\xi) + w_I J_I + w_c J_c, \tag{3.1}$$

gdje w_p, w_c i w_I predstavljaju pozitivne koeficijente.

Kako bismo opisali pojedine pod-funkcije, podatke o jedinicama lokalne samouprave treba povezati s grafom G što nam omogućava dohvaćanje pojedinih svojstava svake jedinice lokalne samouprave tj. vrha $v \in V$. Pozitivne funkcije $pop(v)$ i $area(v)$, definirane na vrhu $v \in V$, predstavljaju ukupnu populaciju i geografsku površinu, redom. Funkcije $pop(v)$ i $area(v)$ proširujemo na kolekciju vrhova $B \subset V$ na način:

$$pop(B) = \sum_{v \in B} pop(v)$$
$$area(B) = \sum_{v \in B} area(v)$$

Rub izborne jedinice $D_i(\xi)$, oznaka $\partial D_i(\xi)$, je podskup skupa bridova E koji povezuju vrhove unutar $D_i(\xi)$ s vrhovima izvan $D_i(\xi)$. Prema ovoj definiciji, jedinice lokalne samouprave koje graniče s morem ili drugom državom neće imati ovakav brid za tu relaciju. Kako bismo uvažili podatke o državnim granicama, uvodimo vrh o u skup V, koji predstavlja inozemstvo i bridom spojimo taj vrh sa svakim vrhom $v \in V$ koji predstavlja jedinicu lokalne samouprave koja granicu dijeli s državnom granicom. Pretpostavljamo da proizvoljan izborni plan raspodjele uvijek zadovoljava uvjet $\xi(v) = 0$ ako i samo ako $v = o$; kako ξ uvijek zadovoljava $\xi(o) = 0$, i $o \notin D_i(\xi)$, za $i \geq 1$, nema potrebe definirati $pop(o)$ i $area(o)$ obzirom da o nikad nije uključen u izborne jedinice.

Neka je dan brid $e \in E$ koji spaja dva vrha $v, \tilde{v} \in V$, tada $granica(e)$ predstavlja duljinu zajedničke granice dviju jedinica lokalne samouprave reprezentiranih s v i $\tilde{v}$.
Kao i prije, proširujemo definiciju na podskup bridova $B \subset E$ na način:

$$granica(B) = \sum_{e \in E} granica(e).$$

Sada kada smo uveli osnovne pojmove, možemo definirati evaluacijske pod-funkcije kako bismo procijenili valjanost izbornog plana.

Populacijski parametar

Populacijski parametar, koji mjeri koliko su izborne jednice podjednako naseljene, definiramo s:

$$J_p(\xi) = \sqrt{\sum_{i=1}^{10} \left(\frac{pop(D_i(\xi))}{pop_{ideal}} - 1 \right)^2}$$

,

$$pop_{ideal} = \frac{N_{pop}}{10}$$

gdje je N_{pop} ukupni broj stanovništva RH, a $pop(D_i(\xi))$ je broj stanovnika izborne jedinice $D_i(\xi)$ i pop_{ideal} je broj stanovnika koji bi svaka izborna jedinica trebala imati prema "jedna osoba jedan glas" standardu, tj. pop_{ideal} je jednak jednoj desetini ukupnog broja stanovnika RH.

Izoperimetrijski parametar

Izoperimetrijski parametar, koji mjeri ukupnu kompaktnost izbornog plana, definiran je s:

$$J_I(\xi) = \sum_{i=1}^{10} \frac{[granica(\partial D_i(\xi))]^2}{area(D_i(\xi))}$$

$J_i(\xi)$ je omjer kvadratnog opsega i ukupne površine svake od izbornih jedinica. Vrijednost izoperimetrijskog parametra je minimalna za krug, koji je najkompaktniji geometrijski lik. Ova mjera kompaktnosti je jedna od dvije najčešće korištene mjere u američkoj zakonskoj literaturi. Recipročna vrijednost ove mjere je proporcionalna *Polsby-Popper score*-u ili 'perimeter score-u'.

Županijski parametar

Županijski parametar mjeri koliko je, i do kojeg stupnja, županija rastavljeno među različitim izbornim jedinicama. Ako dvije jedinice lokalne samouprave pripadaju različitim izbornim jedinicama, ali istoj županiji, tada se ta županija naziva podijeljena županija. Ovaj uvjet osigurava minimalno razbijanje županija na više izbornih jedinica kako bi se sačuvala upravno-politička podjela RH. Parametar je definiran na način:

$$J_Z(\xi) = \{\text{županije podijeljene u 2 izborne jedinice}\} \times W_2(\xi)$$

$$+ \mathrm{M}_Z \times \{\text{županije podijeljene u } \geq 3 \text{ izborne jedinice}\} \times W_3(\xi)$$

Gdje je M_Z velika konstanta koja teško kažnjava podjelu županije između 3 izborne jedinice, a W_2 i W_3 su težinski faktori koji ublažavaju nagle tranzicije između podijeljenih/nepodijeljenih županija i definirani su s:

$$W_2(\xi) = \sum_{\substack{županije\ podijeljene \\ u\ 2\ izborne\ jedinice}} \begin{pmatrix} udio\ županijskih\ jedinica \\ lokalne\ samouprave \\ u\ drugom\ najvećem\ presjeku \\ izborne\ jedinice\ s\ županijom \end{pmatrix}^{1/2}$$

,

$$W_3(\xi) = \sum_{\substack{županije\ podijeljene \\ u\ \geq 3\ izborne\ jedinice}} \begin{pmatrix} udio\ županijskih\ jedinica \\ lokalne\ samouprave\ koje\ nisu \\ u\ prvom\ i\ drugom\ najvećem\ presjeku \\ izborne\ jedinice\ s\ županijom \end{pmatrix}^{1/2}$$

Ako je proizvoljna županija nejednako podijeljena između dvije izborne jedinice, recimo u omjeru 90:10, tada se u definiciji koeficijenta W_2 promatra manji presjek jedinice s županijom. Analogno postupamo u slučaju koeficijenta W_3 i promatramo najmanji presjek županije s izbornom jedinicom.

Dosada smo evaluacijsku funkciju J definirali na R tj. kolekciji svih izbornih planova s povezanim izbornim jedinicama. Sada tu definiciju proširujemo na sve moguće izborne planove s:

$$J(\xi) = \begin{cases} w_p J_p(\xi) + w_I J_I + w_c J_c,\ ako \quad \xi \in R \\ \infty,\ ako \quad \xi \notin R \end{cases}$$

Pokazat ćemo da, ako je $J(\xi) = \infty$, onda je vjerojatnost da će izborni plan ξ biti prihvaćen jednaka nuli.

Familija vjerojatnosnih distribucija na izbornim planovima

Sada koristimo evaluacijsku funkciju $J(\xi)$ kako bismo svakom izbornom planu $\xi \in R$ dodijelili vjerojatnost takvu da izborni planovi s manjom vrijednošću evaluacijske funkcije budu više vjerojatni. Fiksiramo $\beta > 0$ i definiramo vjerojatnost od ξ, oznaka $P_\beta(\xi)$ s:

$$P_\beta(\xi) = \frac{e^{-\beta J(\xi)}}{Z_\beta} \tag{3.2}$$

gdje je Z_β normalizacijska konstanta definirana tako da je $P_\beta(R) = 1$. Posebno,

$$Z_\beta = \sum_{\xi \in R} e^{-\beta J(\xi)}. \tag{3.3}$$

Primjetimo da, ako $\xi \notin R$, onda je $e^{-\beta J(\xi)} = e^{-\infty} = 0$ i vidimo da je $P_\beta(R) = 1$. Iz ovog slijedi da je sva vjerojatnost koncentrirana na izbornim planovima koji imaju povezane izborne jedinice.

Pozitivna konstanta β se često naziva 'inverzna temperatura' u analogiji s statističkom mehanikom i dinamikom plina. Kada je β jako mala (područje visoke temperature) različiti elementi od R imaju skoro pa jednaku vjerojatnost. Kako se β povećava ('temperatura se snižava'), mjera koncentrira vjerojatnost oko izbornih planova $\xi \in R$ koji minimiziraju $J(\xi)$.

Uzorkovanje iz distribucije

Ako zanemarimo činjenicu da izborne jedinice unutar izbornog plana moraju biti povezane komponente, tada postoji više od 10^{560} izbornih planova, što je veće i od trenutne procjene broja atoma u svemiru (između 10^{78} i 10^{82}) i procjene broja sekunda od Velikog praska (4.3×10^{17}). Iako u skupu R (skup svih izbornih planova s povezanim izbornim jedinicama) ima značajno manje izbornih planova, evaluiranje svih izbornih planova kako bi se našli oni s najmanjom vrijednošću funkcije J, nije praktično. Uobičajen i efektivan način kako bi se izbjego ovaj problem dimenzionalnosti jest korištenje MCMC algoritma kako bismo uzorkovali iz distribucije P_β. Koristit ćemo standardan Metropolis-Hastings algoritam. No, prije nego uvedemo konstrukciju algoritma, navodimo intuitivno objašnjenje. Prvo, što nam je zapravo cilj? Cilj nam je dobiti uzorke iz vjerojatnosne distribucije bez da znamo njenu točnu visinu na bilo kojoj točki (ne moramo znati normalizirajuću konstantu). Način na koji MCMC ovo uspijevaju je da "šetaju" po toj distribuciji na način da je količina vremena provedenog na pojedinoj lokaciji proporcionalna visini distribucije na toj lokaciji. Ako je taj proces "šetanja" točno postavljen, možemo dokazati da je navedena proporcionalnost (između vremena i visine distribucije) zaista postignuta. Intuitivno, ono što želimo raditi jest "šetati" po nekoj "brdovitoj" površini na način da je količina provedenog vremena (ili broj dobivenih uzoraka) na svakoj lokaciji proporcionalna visini površine te lokacije. To znači da npr. želimo provesti duplo više vremena na brdu visine 100m nego na susjednom brdu koje je visine 50m. Zgodna stvar je da ovo možemo napraviti čak i ako ne znamo apsolutne visine pojedinih lokacija, dovoljno je da znamo njihove relativne visine, npr. ako je jedno brdo A duplo više od brda B, onda na brdu A želimo provesti duplo više vremena nego na brdu B. Najjednostavnija varijanta Metropolis-Hastings algortima (Nezavisni Metropolis-Hastings) postiže ovo na sljedeći način: pretpostavimo da u svakom (diskretnom) vremenskom trenutku izaberemo nasumičnu novu "predloženu" lokaciju (koju biramo uniformno na cijeloj površini). Ako je predložena lokacija viša od naše trenutne lokacije, prelazimo na nju. Ako je predložena lokacija niža od naše trenutne, onda prelazimo na nju s vjerojatnošću p, gdje je p omjer visine predložene lokacije i visine naše trenutne lokacije. Ove prelaske možemo zamišljati i kao bacanje kovanice koja ima vjerojatnost p za stranu glave; ako dobijemo glavu, prelazimo na novu lokaciju, ako dobijemo pismo, ostajemo gdje jesmo. Vodimo listu svih lokacija na kojima smo bili u svakom vremenskom trenutku i ta lista će onda asimptotski imati točnu proporciju provedenog vremena za svaku lokaciju na našoj površini. I za brda A i B navedena gore, dobit ćemo da je vjerojatnost prelaska iz B u A duplo veća nego vjerojatnost prelaska iz A u B.

Kako smo već objasnili u prethodnom poglavlju, Metropolis-Hastings algoritam je dizajniran da koristi jednu prijelaznu jezgru Q (prijedložni lanac) kako bi uzorkovao iz druge prijelazne jezgre koja ima jedinstvenu stacionarnu distribuciju μ (ciljana distribucija). U našem slučaju, $Q(\xi, \xi')$ daje vjerojatnost prelaska iz izbornog plana ξ u izborni plan ξ'. Koristimo Q kako bismo dobili uzorak iz ciljane distribucije. Algoritam izgleda ovako:

1. Uz dano $X_n = \xi$
2. Generiraj $Y_n = \xi' \sim Q(\xi, \cdot)$
3. Generiraj $U \sim U([0, 1])$
4. Izračunaj $p(\xi, \xi')$, pri čemu je
$$p(\xi, \xi') = min\left(1, \frac{P_\beta(\xi')Q(\xi', \xi)}{P_\beta(\xi)Q(\xi, \xi')}\right)$$
5. Definiraj
$$X_{n+1} = \begin{cases} \xi' & ako\ U \le p(\xi, \xi') \\ \xi & inače \end{cases}$$

6. Ponavljaj za $n = n + 1$

Algoritam zaustavljamo kada smatramo da je lanac napravio dovoljno koraka za naše potrebe.
Primjetimo da, u slučaju da je predloženi izborni plan ξ' nepovezan, to je ekvivalentno postavljanju $J(\xi') = \infty$ iz čega slijedi $P_\beta(\xi') = 0$ pa će naš algoritam uvijek odbaciti nepovezan izborni plan ξ'.

Sada definiramo prijedložni lanac Q koji koristimo za predlaganje izbornih planova na način:

1. Proizvoljno i uniformno izaberemo brid 'u sukobu'. Brid $e = (u, v)$ je brid 'u sukobu' ako $\xi(u) \neq \xi(v), \xi(u) \neq 0, \xi(v) \neq 0$.

2. Za izabrani brid $e = (u, v)$, s vjerojatnošću $1/2$ defniramo novi izborni plan ξ' na način:

$$\xi'(w) = \begin{cases} \xi(w), & w \neq u \\ \xi(v), & w = u \end{cases} \qquad \text{ili} \qquad \xi'(w) = \begin{cases} \xi(w), & w \neq v \\ \xi(u), & w = v \end{cases}$$

Neka je $sukobljeni(\xi)$ broj sukobljenih bridova za izborni plan ξ. Tada imamo $Q(\xi, \xi') = \frac{1}{2}\frac{1}{sukobljeni(\xi)}$. Vjerojatnost prihvaćanja sada je dana s:

$$p = min\left(1, \frac{sukobljeni(\xi)}{sukobljeni(\xi')} e^{-\beta(J(\xi') - J(\xi))}\right)$$

Primjetimo također da ne zahtijevamo da je naša ciljana distribucija P_β normalizirana obzirom da se normalizacijske konstante ponište u razlomku. Upravo ova činjenica, koja nam omogućava da uzorkujemo iz ne-normaliziranih distribucija je jedna od najvećih prednosti Metropolis-Hastings algoritma. Uzorkovanje iz ne-normaliziranih distribucija je česta pojava u Bayesovskim modelima gdje je izračunavanje normalizacijske konstante teško ili nepraktično, baš kao i u slučaju našeg problema.

Stacionarna distribucija i konvergencija k stacionarnoj distribucji

Sada pokazujemo da je P_β uistinu stacionarna distribucija dobivenog Markovljevog lanca. Koristeći Teorem 1.0.3 i Teorem 2.5.1 dovoljno je pokazati da P_β zadovoljava uvjet detaljne ravnoteže s prijelaznom jezgrom dobivenog lanca. Kako bismo to dokazali, potrebno je pronaći prijelaznu jezgru lanca. Obzirom da je skup svih izbornih raspodjela konačan skup, prijelazna jezgra će nam zapravo biti prijelazna matrica. Svako novo stanje ξ' predlažemo s vjerojatnošću $Q(\xi,\xi') = \frac{1}{2}\frac{1}{s(\xi)}$, gdje je $s(\xi)$ oznaka za $sukobljeni(\xi)$ koju uvodimo radi jednostavnosti zapisa. Vjerojatnost prihvaćanja dana je s $p = min\left(1, \frac{P_\beta(\xi')Q(\xi',\xi)}{P_\beta(\xi)Q(\xi,\xi')}\right)$ i ovdje ćemo je označiti s $\alpha_{\xi,\xi'}$.

Sada tražimo vrijednosti $p_{\xi,\xi'}$ prijelazne matrice dobivenog lanca. Neka $\{\mathbf{tp}\}$ označava događaj prihvaćanja predloženog prijelaza, a neka $\{\mathbf{tp}\}^c$ označava njegov komplement. Tada, za $\xi \neq \xi'$ imamo:

$$p_{\xi,\xi'} = P(X_{n+1} = \xi'|X_n = \xi) = P(Y_n = \xi', \mathbf{tp}|X_n = \xi)$$

prema definiciji prijedložne vjerojatnosti i vjerojatnosti prihvaćanja. Nadalje,

$$= P(Y_n = \xi'|\mathbf{tp}, X_n = \xi)P(\mathbf{tp}|X_n = \xi)$$

i, s obzirom da predloženu vrijednost generiramo neovisno o događaju prihvaćanja, imamo:

$$= P(Y_n = \xi'|X_n = \xi)\alpha_{\xi,\xi'}$$

Dakle, za $\xi \neq \xi'$ dobili smo:

$$p_{\xi,\xi'} = Q(\xi,\xi')min\left(1, \frac{P_\beta(\xi')Q(\xi',\xi)}{P_\beta(\xi)Q(\xi,\xi')}\right) = Q(\xi,\xi')\alpha_{\xi,\xi'}.$$

Na dijagonali tj. za $\xi = \xi$ imamo:

$$p_{\xi,\xi} = P(X_{n+1} = \xi|X_n = \xi) = P(Y_n = \xi, \mathbf{tp}|X_n = \xi) + P(Y_n \neq \xi, \mathbf{tp}^c|X_n = \xi)$$

pa analogno kao i gore imamo:

$$\begin{aligned}
&= P(Y_n = \xi|X_n = \xi)P(\mathbf{tp}|X_n = \xi) + P(Y_n \neq \xi|X_n = \xi)P(\mathbf{tp}^c|X_n = \xi) \\
&= P(Y_n = \xi|X_n = \xi)\alpha_{\xi,\xi} + \sum_{\xi' \neq \xi} P(Y_n = \xi'|X_n = \xi)(1 - \alpha_{\xi,\xi'}) \\
&= Q(\xi,\xi)\alpha_{\xi,\xi} + \sum_{\xi' \neq \xi} P(Y_n = \xi'|X_n = \xi)(1 - \alpha_{\xi,\xi'}) \\
&= Q(\xi,\xi) + \sum_{\xi' \neq \xi} Q(\xi,\xi')(1 - \alpha_{\xi,\xi'})
\end{aligned}$$

Čime smo dobili prijelaznu matricu Markovljevog lanca dobivenog navedenim algoritmom. Sada provjeravamo uvjet detaljne ravnoteže. Promatramo $\xi \neq \xi'$ i pretpostavljamo $\alpha_{\xi,\xi'} < 1$ z čega slijedi $\alpha_{\xi',\xi} = 1$ pa u tom slučaju imamo:

$$\alpha_{\xi,\xi'}Q(\xi,\xi')P_\beta(\xi) = \frac{P_\beta(\xi')Q(\xi',\xi)}{P_\beta(\xi)Q(\xi,\xi')}Q(\xi,\xi')P_\beta(\xi) = P_\beta(\xi')Q(\xi',\xi) = P_\beta(\xi')Q(\xi',\xi)\alpha_{\xi',\xi}$$

Sada prepostavimo $\alpha_{\xi,\xi'} = 1$, onda je $\alpha_{\xi',\xi} = \frac{P_\beta(\xi)}{P_\beta(\xi')}\frac{Q(\xi,\xi')}{Q(\xi',\xi)}$ pa imamo:

$$\alpha_{\xi,\xi'}Q(\xi,\xi')P_\beta(\xi) = Q(\xi,\xi')P_\beta(\xi) = Q(\xi,\xi')P_\beta(\xi)\frac{P_\beta(\xi')}{P_\beta(\xi')}\frac{Q(\xi',\xi)}{Q(\xi',\xi)} = \alpha_{\xi',\xi}Q(\xi',\xi)P_\beta(\xi').$$

čime je zadovoljen uvjet detaljne ravnoteže pa po Teoremu 1.0.3 i Teoremu 2.5.1 slijedi da je P_β stacionarna distrubucija dobivenog lanca.

Nadalje, tvrdimo da Markovljev lanac dan ovim algoritmom konvergira k željenoj distribuciji ako se algoritam izvodi dovoljno dugo. Koristeći tvrdnje navedene u prethodnom poglavlju, posebice Teorem 2.5.3 dovoljno je da pokažemo kako je lanac ireducibilan i aperiodičan. S obzirom da kroz korake algoritma iz proizvoljnog izbornog plana možemo dobiti neki drugi proizvoljan izborni plan, lanac je, po tvrdnji (1.2) prvog poglavlja koja se odnosi na diskretan slučaj, ireducibilan. Aperiodičnost slijedi iz činjenice da postoje izborni planovi koji su povezani sami sa sobom kroz petlju koja se sastoji od 2 koraka i kroz petlju koja se sastoji od 3 koraka. Kako su 2 i 3 prosti brojevi i stoga imaju 1 kao najveći zajednički djelitelj, koristeći definiciju perioda (1.3) lanac je aperiodičan. Dakle, nakon dovoljnog broja tranzicija lanca (eng. burnin period) stanja našeg Markovljevog lanca možemo promatrati kao uzorke iz željene ciljane distribucije.

Definiranje težinskih parametara evaluacijske funkcije

U definiciji evaluacijske funkcije (3.1) koristimo 3 težinska faktora (w_p, w_I i w_c) kako bismo ujednačili efekt vrijednosti različitih pod-funkcija na sveukupnu vrijednost evaluacijske funkcije $J(\xi)$. Uz ove faktore također imamo i niske i visoke temperature koje odgovaraju maksimalnom i minimalnom β, redom. Fiksiramo minimalnu vrijednost $\beta = 0$ što odgovara beskonačnoj temperaturi. U ovom području nijedna izborna jedinica nije preferirana naprema drugima što omogućava izbornom planu da slobodno pretražuje prostor svih mogućih izbornih planova. Sada nam preostaje još fiksirati visoku vrijednost parametra β. Obzirom da u definiciji target distribucije β množi ostale parametre, jedno od ovih ograničenja je redundantno pa ga možemo proizvoljno postaviti. Fiksiramo nisku temperaturu tj. visoku vrijednost od β na 1.

Kako bismo odredili faktore, koristimo sljedeću metodu:

1. Postavi sve faktore na 0
2. Simuliraj izborne planove za proizvoljne vrijednosti težinskog parametra w_p. Nađi najmanji w_p takav da dio rezultata zadovoljava određeni uvjet (npr. možemo zahtijevati da 25% izbornih planova bude ispod 0.5% populacijske devijacije)

3. Koristeći w_p iz prethodnog koraka, simuliraj izborne planove za navedeni w_p i proizvoljne w_I. Nađi najmanji w_I takav da dio rezultata ima sve izborne jedinice ispod dane vrijednosti izoperimetrijske funkcije (npr. možemo zahtijevati da 80% izbornih planova ima vrijednost izoperimetrijskog parametra manju od 60)

4. Ako populacijski uvjet više nije zadovoljen, ponovi korake 2 do 4 dok oba uvjeta ne budu zadovoljena

5. Koristeći w_p i w_I iz prethodnih koraka, simuliraj izborne planove za navedene w_p i w_I i proizvoljne w_c. Nađi najmanji w_c takav da gotovo uvijek imamo podjele županija samo na 2 izborne jedinice i da broj takvih 2-podjela, u prosjeku, bude manji od 25

6. Ako populacijski uvjet više nije zadovoljen, ponovi korake 2 do 6. Ako uvjet kompaktnosti više nije zadovoljen, ponovi korake 3 do 6. Inače, kraj.

Bibliografija

[1] Sachet Bangia, Christy Vaughn Graves, Gregory Herschlag, Han Sung Kang, Justin Luo, Jonathan C. Mattingly i Robert Ravier, *Redistricting: Drawing the Line*, arXiv e-prints (2017), arXiv:1704.03360.

[2] Timo Koski, *Sf2955 COMPUTER INTENSIVE METHODS MCMC On the Discrete Metropolis-Hastings Algorithm 2009*, (2019).

[3] Christian P. Robert i George Casella, *Monte Carlo Statistical Methods*, Springer Publishing Company, Incorporated, 2010, ISBN 1441919392, 9781441919397.

[4] Z. Vondraček, *Markovljevi lanci - predavanja, PMF-MO*, 2008.

[5] H. Šikić, *Mjera i integral, predavanja, PMF-MO*, 2014.

Sažetak

U prvom poglavlju definirali smo Markovljeve lance na općenitom skupu stanja. Definirali smo pojmove kao što su prijelazna jezgra Markovljevog lanca u Definiciji 1.0.1, invarijantna mjera u Definiciji 1.0.7 te detaljna ravnoteža. Pokazali smo da je uvjet detaljne ravnoteže dovoljan uvjet da bi Markovljev lanac imao invarijantnu mjeru. Nakon toga drugom poglavlju primjenili smo tu teoriju na Markovljev lanac dobiven Metropolis-Hastings algoritmom. Naveli smo dovoljne uvjete kako bi dobiveni lanac imao stacionarnu distribuciju. U Teoremu 2.5.3 objedinili smo rezulate o egzistenciji stacionarne distribucije i konvergenciji lanca k stacionarnoj distribuciji. U trećem poglavlju promatrali smo politički problem gerrymanderinga te smo ga vizualizirali na primjeru američke države North Caroline. Uveli smo evaluacijsku funkciju J kako bismo svakom izbornom planu dodijelili vjerojatnost takvu da izborni planovi s manjom vrijednošću evaluacijske funkcije budu više vjerojatni. Pomoću evaluacijske funkcije definirali smo ciljanu vjerojatnosnu distribuciju. Konačno, predložili smo način stvaranja novih izbornih raspodjela koristeći Metropolis-Hastings algoritam.

Summary

In the first chapter we discussed the theory of Markov chains on a general state space. We defined terms such as transition kernel in Definition 1.0.1, invariant measure in 1.0.7 and detailed balance condition. We have shown that the detailed balance condition is sufficient for a Markov chain to have an invariant measure. After that, in the second chapter, we have applied that theory to a Markov chain produced by Metropolis-Hastings algorithm. We list the sufficient conditions for the produced Markov chain to have a stationary distribution. In Theorem 2.5.3 we combine results of existence of stationary distribution and the convergence of a chain to the stationary distribution. In third chapter, we examined the political problem of gerrymandering. We visualized it on an example of an US state of North Carolina. We introduced a score function J in order to assign a value to each redistricting so that redistrictings with lower scores are more likely to occur. Using this score function we have defined a target distribution. Finally, we proposed a way of creating new redistrictings using a standard Metropolis-Hastings algorithm.

Životopis

Rođena sam 10. srpnja 1993. godine u Splitu. Nakon završene Osnovne škole "Sućidar" upisujem IV. gimnaziju "Marko Marulić" u Splitu. Akademske godine 2012./2013. upisujem preddiplomski studij Matematike na Prirodoslovno - matematičkom fakultetu Sveučilišta u Splitu, na kojem, 2016. godine, stječem akademski naziv sveučilišne prvostupnice matematike (bacc. univ. math.). Iste godine upisujem diplomski studij Matematička statistika na Matematičkom odsjeku Prirodoslovno - matematičkog fakulteta u Zagrebu.