**math.e**

Hrvatski matematički elektronički časopis

Kvazislučajni brojevi i numerička integracija

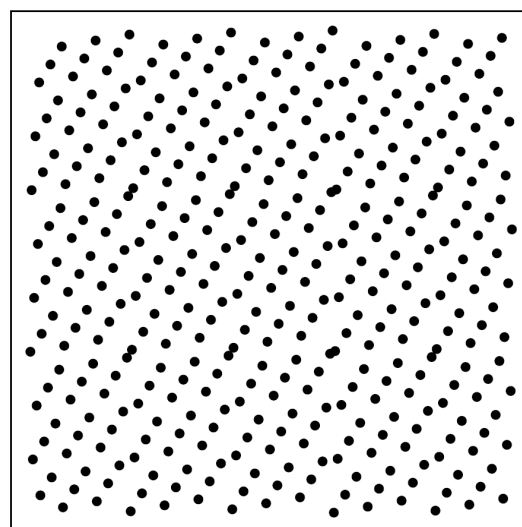
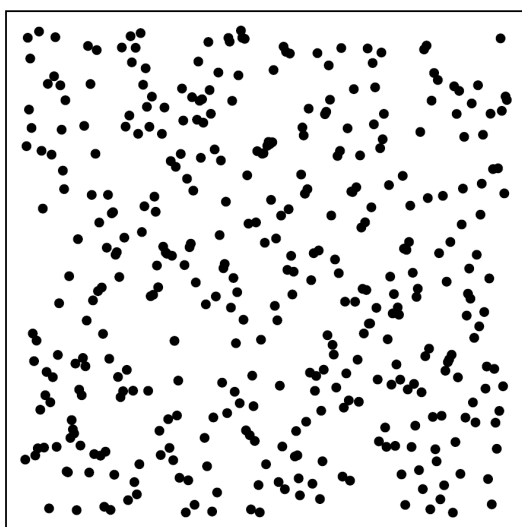
Vjekoslav Kovač¹ Luka Turčić²

Sažetak. U ovom članku izlažu se osnove Kvazi Monte Carlo metode, kojom se generiraju tzv. kvazislučajni brojevi. Ona je sasvim determinističke prirode, ali ime, motivaciju i korisnost dijeli sa svojom poznatijom, stohastičkom varijantom — Monte Carlo metodom. Upoznat ćemo se s teorijskim aspektima teme, koji se uglavnom tiču pojmova ekvidistribuiranosti i diskrepancije, kao i s najvažnijom primjenom metode, za numeričko integriranje funkcija. (Ovaj članak je izrađen na temelju diplomskog rada jednog od autora [9].)

1 Slučajni, pseudoslučajni i kvazislučajni brojevi

Stavimo pred čitatelja poznato trik-pitanje koje provjerava naše poimanje "slučajnosti".

- *Koji od kvadrata na slici 1 je ispunjen konačnim nizom slučajno generiranih točaka?*



Slika 1: Pseudoslučajni niz točaka (lijevo) i kvazislučajni niz točaka (desno).

Lako može zavesti činjenica da lijevi kvadrat sadrži i nakupine točaka i praznine, a to je nešto što neiskusni čitatelj možda ne očekuje od

slučajnog niza. Zavarati može i činjenica da točke u desnom kvadratu ravnomjerno ispunjavaju njegovu površinu, što bismo mogli dovesti u vezu s činjenicom da u svaki djelić kvadrata slučajno odabrana točka pada s jednakom vjerojatnosti. Oba ta argumenta, koji navode na odgovor "desni kvadrat", su pogrešni! Kod slučajno generiranog niza točaka s velikom vjerojatnosti se formiraju razne praznine i nakupine, dok s malom vjerojatnosti on baš ravnomjerno počinje ispunjavati kvadrat. Radi toga je "lijevi kvadrat" *točniji* odgovor, u smislu da nam simulacija slučajnog niza točaka vrlo vjerojatno izgleda baš kao na lijevoj polovici slike.

Ako smo pak sasvim rigorozni, onda je prethodno pitanje besmisleno! S teorijskog stanovišta oba scenarija na slici 1 su moguća i jednako (ne)vjerojatna. Nema razloga zašto bismo neki konkretni niz brojeva ili točaka zvali slučajnim ili ne. Slučajni brojevi ne postoje — možemo samo govoriti o nizu slučajnih varijabli; za definiciju vidjeti udžbenike iz elementarne teorije vjerojatnosti [7] i [8]. Ljevi kvadrat je ispunjen nizom tzv. *pseudoslučajnih brojeva*, koji se trudi izgledati slučajno, tj. što vjerodostojnije idealnoj pravoj simulaciji niza slučajnih varijabli. Desni kvadrat je ispunjen nizom tzv. *kvazislučajnih brojeva*, koji se naprosto trudi što ravnomjernije popunjavati kvadrat; rigorozna definicija će uslijediti u jednom od kasnijih odjeljaka. Obje konstrukcije svojju pravu vrijednost realiziraju u primjenama, a u ovom ih radu povežemo s naizgled nevezanim matematičkim konceptom: integriranjem funkcija.

2 Klasična numerička integracija

čitatelj je zasigurno upoznat s pojmom integrala funkcije, makar na intuitivnom nivou. *Integral* (tj. *određeni integral*) dovoljno "lijepo" (recimo neprekidne) funkcije $f: [a, b] \rightarrow \mathbb{R}$ je broj pisan

$$\int_a^b f(x) dx, \quad \text{ili} \quad \int_{[a,b]} f(x) dx, \quad \text{ili samo} \quad \int_{[a,b]} f,$$

a kojeg se interpretira kao površinu ispod grafa funkcije, a iznad apscise. Pritom se površina uzima s negativnim predznakom tamo gdje je funkcija negativna. Osnove diferencijalnog i integralnog računa te, posebno, rigoroznu definiciju tzv. *Riemannovog integrala* čitatelj može naći u knjigama [3] i [4]. Ipak, to znanje nije ključno za razumijevanje narednog materijala. Komponiranjem funkcije f s pogodno odabranom afinom funkcijom općeniti se slučaj lako svodi na interval $[a, b] = [0, 1]$ pa ćemo to već pretpostavljati u daljnjem tekstu.

U srži numeričkog računanja integrala leži njegova aproksimacija diskretnim sumama, a najjednostavniji slučaj je

$$\int_{[0,1]} f(x) dx \approx \frac{1}{N} \sum_{n=1}^N f(x_n), \quad (1)$$

pri čemu za točke $x_n \in [0, 1]$, $n = 1, 2, \dots, N$, možemo uzeti recimo $x_n = (n-1)/N$ ili $x_n = n/N$ pa dobivamo tzv. *lijevu i desnu pravokutnu formulu*. Za svaki od ta dva izbora greška aproksimacije (1) iznosi najviše $C(f')N^{-1}$, pri čemu je $C(f')$ neka konstanta proporcionalna maksimumu apsolutne vrijednosti derivacije f' na intervalu $[0, 1]$. Pokazuje se da je zapravo pametnije gledati prosjek dviju pravokutnih formula, čime dobivamo tzv. *trapeznu formulu*,

$$\int_{[0,1]} f(x) dx \approx \frac{1}{2N} \left(f(0) + 2f\left(\frac{1}{N}\right) + 2f\left(\frac{2}{N}\right) + \dots + 2f\left(\frac{N-1}{N}\right) + f(1) \right).$$

Greška aproksimacije (2) je najviše $C(f'')N^{-2}$, pri čemu konstanta

$C(f''')$ ovisi o drugoj derivaciji od f ; dakle trebamo nešto veću glatkoću funkcije. Ista teorijska ocjena greške vrijedi i ako u (1) naprosto uzmemo $x_n = (2n - 1)/2N$; to je tzv. *formula srednjih točaka*. Dokazi se mogu naći u skripti [1].

Kod višestrukog integrala, $\int_{[0,1]^d} f$, tj. integrala d -dimenzionalne funkcije $f: [0, 1]^d \rightarrow \mathbb{R}$, postupamo slično. U svakoj koordinati možemo promatrati točke $0, 1/M, 2/M, \dots, (M - 1)/M$ pa onda za čvorove x_n u d -dimenzionalnom poopćenju lijeve pravokutne formule uzimamo sve različite d -torke tih brojeva; ima ih $N = M^d$. Usrednjivanjem se dobiva i d -dimenzionalna trapezna formula čija greška je reda veličine $M^{-2} = N^{-2/d}$, čim funkcija ima neprekidne sve druge parcijalne derivacije.

Postoje i mnogo naprednije metode numeričke integracije od navedenih; pogledajte skriptu [1]. Ipak, sve one imaju sljedeće nedostatke, koji ih u nekim posebnim situacijama čine neupotrebljivima.

Nedostatak 1. Sve klasične formule se u velikom broju dimenzija susreću s tzv. *kletvom dimenzionalnosti*. Npr. ako je teorijska greška aproksimacije $CN^{-2/d}$, a želimo garantirati preciznost ε , tada treba uzeti $N \geq (C/\varepsilon)^{d/2}$, tj. broj čvorova u kojima se funkcija treba izračunati raste eksponencijalno s povećanjem dimenzije d . Poteškoću možemo vidjeti već i zdravo-razumski: ako u $d = 100$ dimenzija želimo imati bilo kakvu razumnu aproksimaciju, onda kod gornjih aproksimacija u svakoj koordinati moramo uzeti barem po dvije točke (a i to je vrlo skromno), što nas vodi na ukupno $2^{100} \approx 1.27 \cdot 10^{30}$ čvorova — definitivno previše za bilo koje računalo!

Nedostatak 2. Klasične metode pretpostavljaju izvjesnu glatkoću funkcije koja se integrira. Sofisticiranije formule daju bolju grešku aproksimacije, ali i pretpostavljaju povećanu glatkoću; npr. poznata *Simpsonova formula* u jednoj dimenziji ima teorijsku grešku reda veličine $C(f^{(4)})N^{-4}$ — ona je reda veličine N^{-4} , ali uz pretpostavku ograničenosti četvrte derivacije funkcije f .

Nedostatak 3. Numerička integracija može iziskivati računanje funkcije f u velikom broju čvorova x_n , a to može biti vrlo spor proces ukoliko je funkcija složena. Ponekad nas zahtjev za greškom manjom od ε vodi na veći broj računskih operacija nego ih računalo može podnijeti u danom vremenu. Zato bismo voljeli imati aproksimacijski algoritam kojeg možemo zaustaviti nakon bilo kojeg broja obrađenih čvorova. Uočimo da gore spomenute formule za numeričku integraciju nemaju to svojstvo; ako zaustavimo računanje trapezne formule (2) nakon obrađene polovine čvorova, nećemo imati izračunatu nijednu vrijednost od f na podintervalu $\langle 1/2, 1 \rangle$ pa zato ni aproksimacija integrala $\int_{[0,1]} f$ neće uopće biti smisljena.

3 Monte Carlo metoda

Postoji tehnika numeričke integracije koja zaobilazi sva tri spomenuta nedostatka i zove se *Monte Carlo metoda*. Kod nje su čvorovi x_n slučajni brojevi iz intervala $[0, 1]$ ili pak slučajne točke iz d -dimenzionalne kocke $[0, 1]^d$. U teoriji je x_1, x_2, x_3, \dots niz nezavisnih slučajnih varijabli iz intervala $[0, 1]$. Kvadratni prosjek greške aproksimacije integrala je tada oblika $C(f)N^{-1/2}$, pri čemu je konstanta $C(f)$ konačna već kada $|f|^2$ samo ima konačan integral; vidjeti [5]. Ova metoda zaobilazi sva tri ranije spomenuta nedostatka: dimenzija d ne igra nikakvu ulogu u ocjeni greške, od funkcije f se ne traži nikakva glatkoća, a računanje vrijednosti $f(x_n)$,

$n = 1, 2, \dots, N$, možemo zaustaviti kod bilo kojeg broja N . Mi ne možemo doista raditi s nizom nezavisnih slučajnih varijabli, već ga *simuliramo*, a tzv. *jaki zakon velikih brojeva* garantira da za gotovo svaku realizaciju tog niza prosjeci na desnoj strani od (1) doista konvergiraju prema integralu na lijevoj strani; vidjeti [7] i [8].

Jedno od prvih elektroničkih računala, imena ENIAC, dovršeno je 1945. godine. Već je iduće godine poznati poljsko-američki matematičar Stanisław Ulam došao na ideju da iskoristi njegovu brzinu za Monte Carlo simulacije vezane uz razne fizikalne probleme, a posebno one vezane uz razvoj nuklearnog oružja u laboratoriju u Los Alamosu. Ništa manje slavni, mađarsko-američki matematičar John von Neumann je, nadahnut idejom svojeg kolege, već 1947. razradio plan implementacije te metode na računalo ENIAC; detalji te priče mogu se pročitati u popularnom članku [2]. Navodno je naziv "Monte Carlo" predložio Ulamov prijatelj, u skladu s čestom interpretacijom slučajnog uzorkovanja preko igara na sreću, a prema gradu gdje je Ulamov ujak običavao kockati.

Monte Carlo metoda pak ima neke svoje nedostatke.

Nedostatak 4. Za implementaciju te tehnike treba generirati niz slučajnih brojeva, tj. savršeno simulirati niz nezavisnih slučajnih varijabli. U praksi to nikada nije moguće! Von Neumann je prilikom svoje implementacije Monte Carlo metode za ENIAC opisao neke moguće konstrukcije *pseudoslučajnih brojeva*. Oni su deterministički generirani, ali po mnogim statističkim svojstvima imitiraju slučajnost. Ipak, tada postoji opasnost od nesrazmjera teorijske pozadine i praktične realizacije. Konkretnim problemom simulacije slučajnih varijabli se ovdje nećemo baviti, jer je on težak već sam za sebe.

Nedostatak 5. Monte Carlo metoda u praksi dosta sporo konvergira. To vidimo već i iz ocjene greške oblika $C(f)N^{-1/2}$; ona je inferiorna naspram greškama ranije spomenutih determinističkih metoda, koje su redom reda veličina N^{-1} , N^{-2} , odnosno čak N^{-4} . To nas ne treba čuditi jer ona ne uspijeva iskoristiti informaciju o mogućoj dodatnoj glatkoći zadane funkcije f — algoritam uvijek postupa jednako, što je s jedne strane njegova prednost, a s druge strane njegova mana.

Nedostatak 6. Spomenuta ocjena greške je vjerojatnosna, tj. ne možemo baš ništa zaključiti o greški aproksimacije dobivene pomoću neke konkretne simulacije niza slučajnih varijabli. Kako možemo znati da naša simulacija neće započeti baš s tisuću nula? U tom slučaju izraz s desne strane od (1) aproksimira integral naprosto brojem $f(0)$, što je vrlo loša procjena. Teoretski je moguće da se to dogodi, ali je vrlo malo vjerojatno.

4 Kvazi Monte Carlo metoda

Tehnika koju ćemo sada opisati zove se *kvazi Monte Carlo metoda* i ona stoji na pola puta između pristupa iz prethodna dva odjeljka: deterministička je, ali po mnogo čemu imitira Monte Carlo metodu. Ona nipošto nije univerzalno rješenje problema numeričke integracije, već naprosto vrlo dobar kompromis između nedostataka 1–6 prethodnih metoda.

Glavna ideja kvazi Monte Carlo metode je uzeti čvorove x_n koji će naizgled dobro oponašati slučajne, na način da će ravnomjerno popunjavati prostor. Pokazat će se da će takvi čvorovi davati povoljnu numeričku aproksimaciju integrala, uz dodatno dobro svojstvo determiniranosti. Tu ćemo ideju sada i formalizirati. Uglavnom ćemo se zadržati u dimenziji $d = 1$.

Neka je $\mathbf{x} = (x_n)_{n=1}^{\infty}$ proizvoljan niz brojeva iz intervala $[0, 1]$. Za skup $S \subseteq [0, 1]$ i broj $N \in \mathbb{N}$ definiramo tzv. *brojeću funkciju* kao

$\text{Br}_N(\mathbf{x}, S) :=$ broj indeksa $n \in \{1, 2, \dots, N\}$ takvih da je $x_n \in S$.

Riječima, između prvih N članova niza \mathbf{x} njih točno $\text{Br}_N(\mathbf{x}, S)$ se nalazi u skupu S . Reći ćemo da je niz \mathbf{x} *ekvidistribuiran*³ ako za svaki interval $I \subseteq [0, 1]$ vrijedi

$$\lim_{N \rightarrow \infty} \frac{\text{Br}_N(\mathbf{x}, I)}{N} = \text{duljina}(I).$$

Kakve veze ima pojam ekvidistribuiranosti s računanjem integrala funkcije? Pokazuje se da za ekvidistribuirani niz \mathbf{x} brojeva iz $[0, 1]$ i za Riemann-integrabilnu funkciju $f: [0, 1] \rightarrow \mathbb{R}$ uvijek vrijedi

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N f(x_n) = \int_{[0,1]} f(x) dx. \quad (3)$$

Dakle, integral funkcije je prikazan kao limes niza konačnih suma, prosjeka funkcije u prvih N članova niza \mathbf{x} . Formula (3) ne govori ništa o brzini konvergencije niza na njezinoj lijevoj strani; ne znamo koliko članova niza treba uzeti da bi aproksimacija bila zadovoljavajuća. Radi toga moramo uvesti kvantitativnu mjeru ekvidistribuiranosti niza.

Za niz $\mathbf{x} = (x_n)_{n=1}^{\infty}$ brojeva iz intervala $[0, 1]$ i prirodni broj N definiramo *diskrepanciju* od \mathbf{x} kao broj

$$d_N(\mathbf{x}) := \sup_I \left| \frac{\text{Br}_N(\mathbf{x}, I)}{N} - \text{duljina}(I) \right|,$$

pri čemu se supremum uzima po svim intervalima $I \subseteq [0, 1]$. što je diskrepancija manja, to su točke ravnomjernije raspoređene.

Kako se sada povezuje diskrepancija niza \mathbf{x} s aproksimacijom integrala zasnovanom na formuli (3)? Usredotočimo se na tzv. *funkcije ograničene varijacije*. Za funkciju $f: [0, 1] \rightarrow \mathbb{R}$ definiramo (*totalnu varijaciju*) kao broj

$$V(f) := \sup_{0 \leq a_0 < \dots < a_m \leq 1} \sum_{j=1}^m |f(a_j) - f(a_{j-1})|,$$

pri čemu se supremum uzima po svim $m \in \mathbb{N}$ i svim mogućim izborima brojeva $a_0, a_1, \dots, a_{m-1}, a_m$ takvih da je $0 \leq a_0 < a_1 < \dots < a_m \leq 1$. Zanimaju nas funkcije za koje je $V(f) < \infty$; npr. sve neprekidno diferencijabilne funkcije su takve. Koksma je prvi dokazao nejednakost

$$\left| \frac{1}{N} \sum_{n=1}^N f(x_n) - \int_{[0,1]} f(x) dx \right| \leq V(f) d_N(\mathbf{x}), \quad (4)$$

koja eksplicitno ocjenjuje grešku prethodno spomenute aproksimacije integrala. Iz (4) vidimo da je najpametnije birati niz \mathbf{x} za koji je diskrepancija $d_N(\mathbf{x})$ najmanja moguća, makar što se tiče reda veličine u terminima broja N kada $N \rightarrow \infty$. Schmidt je pokazao da postoji konstanta $c > 0$ takva da za svaki niz brojeva \mathbf{x} iz $[0, 1]$ za beskonačno mnogo prirodnih brojeva N vrijedi

$$d_N(\mathbf{x}) \geq c \frac{\log N}{N}.$$

Sa druge strane, u idućem ćemo odjeljku navesti neke primjere nizova \mathbf{x} takvih da za svaki $N \in \mathbb{N}$ vrijedi

$$d_N(\mathbf{x}) \leq C \frac{\log N}{N}, \quad (5)$$

uz neku konstantu C ovisnu o nizu \mathbf{x} .

Vrlo slično definiraju se ekvidistribuiranost i diskrepancija niza točaka u d -dimenzionalnoj kocki $[0, 1]^d$ za $d \geq 2$. Poznata otvorena slutnja glasi da postoji konstanta $c > 0$, ovisna samo o dimenziji d , takva da za svaki niz točaka \mathbf{x} iz $[0, 1]^d$ možemo naći beskonačno mnogo $N \in \mathbb{N}$ za koje vrijedi

$$d_N(\mathbf{x}) \geq c \frac{(\log N)^d}{N}.$$

Osim za $d = 1$ ta slutnja je još jedino provjerena u $d = 2$ dimenzije. Svi navedeni jednodimenzionalni rezultati također imaju i svoje višedimenzionalne analogone. Detalje čitatelj može naći u knjizi [5], a mi ih izostavljamo radi jednostavnosti izlaganja.

Kvazislučajan niz brojeva iz $[0, 1]$ ili točaka iz $[0, 1]^d$ nam od sada naprosto znači niz s niskom diskrepancijom. Na osnovu rečenog pod time se obično misli niz \mathbf{x} takav da za svaki $N \in \mathbb{N}$ vrijedi

$$d_N(\mathbf{x}) \leq C \frac{(\log N)^d}{N}, \quad (6)$$

uz neku konstantu C . To se doista i može postići i to raznolikim konstrukcijama, od kojih ćemo neke navesti u nastavku.

čitatelj sada već pogađa: *Kvazi Monte Carlo metoda* aproksimira integral tako da na desnoj strani formule (1) uzima upravo neki kvazislučajni niz x_1, x_2, x_3, \dots . Lako je argumentirati prednosti ove metode po svakom od 6 spomenutih nedostataka prethodnih tehnika. Npr., u usporedbi s klasičnom numeričkom integracijom, zaobiđen je nedostatak br. 3: definicija diskrepancije traži da za *svaki* $N \in \mathbb{N}$ prvih N članova niza ravnomjerno ispunjava interval, kvadrat, ili kocku. Nadalje, u usporedbi s Monte Carlo metodom, nedostatak br. 6 zaobiđen je tako što biramo konkretne nizove čija diskrepancija se može ocijeniti pa je onda i ocjena greške aproksimacije (4) deterministička i sasvim konkretna. Obzirom da niz točaka biramo pažljivo, tako da je njegova diskrepancija jednaka naslućenom teorijskom minimumu (6), dobivamo da je ta greška oblika $C(f)N^{-1}(\log N)^d$, što je samo malo lošije od $C(f)N^{-1}$ kada $N \rightarrow \infty$, a svakako superiorno nad greškom Monte Carlo metode, koja je oblika $C(f)N^{-1/2}$.

5 Generiranje kvazislučajnih brojeva

Nakon što smo uvidjeli korisnost nizova s niskom diskrepancijom, prirodno se pitamo kako generirati takve nizove. Opišimo dvije klasične konstrukcije.

5.1 Van der Corputov i Haltonov niz

Proizvoljan prirodni broj n ima jedinstveni raspis u bazi $b \geq 2$ oblika

$$n = \sum_{j=0}^{\infty} a_j(n)b^j,$$

gdje su $a_j(n) \in \{0, 1, \dots, b-1\}$ njegove znamenke i $a_j(n) = 0$ za sve dovoljno velike indekse j ; odnosno gornja suma je zapravo

konačna. Sada, tzv. obrtanjem znamenaka, za svaki n možemo definirati broj

$$\phi_b(n) := \sum_{j=0}^{\infty} a_j(n) b^{-j-1},$$

koji pripada jediničnom intervalu $[0, 1]$. Za danu bazu $b \geq 2$ van der Corputov niz je upravo niz brojeva x_0, x_1, \dots definiran sa $x_n = \phi_b(n)$ za sve $n \in \mathbb{N}$ i on je primjer niza čija diskrepancija zadovoljava (5), tj. on je kvazislučajan.

Računalno praktičan način dobivanja višedimenzionalnih nizova niske diskrepancije je generalizacija van der Corputovog niza. Koristeći netom uvedenu funkciju ϕ_b Haltonov niz u bazama $b_1, \dots, b_d \geq 2$ definiramo kao niz točaka x_0, x_1, \dots iz $[0, 1]^d$ takav da vrijedi

$$x_n = (\phi_{b_1}(n), \dots, \phi_{b_d}(n)) \quad \text{za sve } n \in \mathbb{N}.$$

Ukoliko su brojevi b_1, \dots, b_d relativno prosti, tada gornji niz zadovoljava (6), tj. to je kvazislučajni niz točaka.

Ukoliko samo želimo kvadrat $[0, 1]^2$ ravnomjerno ispuniti unaprijed zadanim brojem točaka N , možemo koristiti i tzv. Hammersleyev skup za bazu $b \geq 2$:

$$\left\{ \left(\frac{n}{N+1}, \phi_b(n) \right) : n = 1, 2, \dots, N \right\}.$$

Naprimjer, desna polovica slike 1 ilustrira upravo taj skup točaka za $b = 3$ i $N = 385$. Ipak, treba napomenuti da na ovaj način nije definiran kvazislučajni niz u smislu naše definicije, jer je veličina skupa unaprijed određena.

5.2 Nizovi dobiveni od iracionalnih brojeva

Za proizvoljan broj $u \in \mathbb{R}$ definiramo decimalni dio broja u kao $\{u\} := u - \lfloor u \rfloor$. Drugim riječima, za nenegativan u , imamo da je $\{u\}$ dio zapisa od u koji se nalazi iza decimalne točke. Niz niske diskrepancije $\mathbf{x} = (x_n)_{n=1}^{\infty}$ tada dobijemo promatranjem decimalnih dijelova višekratnika fiksniranog iracionalnog broja z , odnosno možemo definirati $x_n = \{nz\}$ za sve $n \in \mathbb{N}$. Svaki tako definiran niz brojeva \mathbf{x} je ekvidistribuiran. Ipak, ako inzistiramo na optimalnoj ocjeni za diskrepanciju (5), tada moramo pažljivo birati iracionalni broj z ; recimo tzv. zlatni omjer $z = \frac{\sqrt{5}-1}{2} = 0.618\dots$ je jedan dobar odabir.

Slično možemo dobiti višedimenzionalne nizove niske diskrepancije. Neka je $z = (z_1, \dots, z_d) \in \mathbb{R}^d$ takav da su brojevi $1, z_1, \dots, z_d$ linearno nezavisni nad racionalnim brojevima, tj.

$$\alpha_0 + \sum_{i=1}^d \alpha_i z_i = 0, \quad \alpha_i \in \mathbb{Q} \implies \alpha_i = 0, \quad i = 0, 1, \dots, d,$$

i neka je niz točaka x_0, x_1, \dots definiran sa

$$x_n = (\{nz_1\}, \dots, \{nz_d\}) \quad \text{za sve } n \in \mathbb{N}.$$

Takav niz točaka je uvijek makar ekvidistribuiran, ali mu ocjena diskrepancije vodi na ozbiljne probleme u području teorije brojeva.

6 Numerički primjer

Pokušajmo numerički izračunati određeni integral

$$\int_3^4 x e^x dx. \quad (7)$$

Radi usporedbe s pravom vrijednosti odabrali smo integral kojeg znamo i egzaktno izračunati: parcijalnom integracijom lako dobivamo

$$3e^4 - 2e^3 = 123.623376 \dots$$

U svrhu primjene gornjih metoda najprije integral (7) supstitucijom $x = t + 3$ svodimo na integral po intervalu $[0, 1]$:

$$\int_0^1 (t + 3)e^{t+3} dt. \quad (8)$$

Za funkciju f zadanu formulom $f(t) = (t + 3)e^{t+3}$ koristili smo Monte Carlo metodu i kvazi Monte Carlo metodu uz van der Corputov niz za bazu $b = 3$. Usporedni numerički rezultati dani su tablicama na slici 3. Možemo vidjeti kako je relativna greška Monte Carlo metode prilično nepredvidiva, dok kvazi Monte Carlo metoda u praksi daje sve bolju aproksimaciju s rastućim brojem članova niza. (Pseudoslučajni niz točaka (lijevo) i kvazislučajni niz točaka (desno).)

N	10	100	1000	100000
num. rezultat	120.5889	124.4008	122.4759	123.4507
rel. odstupanje	2.45%	0.63%	0.93%	0.14%

Slika 2: Monte Carlo metoda

N	10	100	1000	100000
num. rezultat	113.6413	121.7685	123.3847	123.6190
rel. odstupanje	8.07%	1.50%	0.19%	0.004%

Slika 3: Kvazi Monte Carlo metoda (van der Corput, $b = 3$)

7 Zaključak

Integrali su jedan od osnovnih matematičkih pojmova s brojnim primjenama, kako unutar, tako i izvan same matematike. Česte su situacije kada je integral nepraktično, a ponekad i nemoguće egzaktno izračunati. U takvim je slučajevima korisno poslužiti se metodama numeričke integracije, odnosno numeričkom aproksimacijom integrala u pitanju, za što postoje brojne klasične metode (trapezna formula, Simpsonova formula, itd.). Takve metode, uz brojne prednosti, posjeduju i neke spomenute mane.

Bitno različit postupak, koji predstavlja poboljšanje u odnosu na neke klasične nedostatke, je Monte Carlo metoda, odnosno korištenje slučajno generiranih točaka pri procjeni integrala. Monte Carlo metoda, uz broj čvorova N , daje grešku aproksimacije reda veličine $N^{-\frac{1}{2}}$, što je lošije od klasičnih metoda, ali je neovisno od broja dimenzija d . Kvazi Monte Carlo metoda je determinističko poboljšanje Monte Carlo metode koje uspijeva zadržati neke njezine prednosti. U njoj se koristimo kvazislučajnim nizovima, odnosno nizovima niske diskrepancije, za koje smo neke od klasičnih konstrukcija dali u ovome radu.

Već smo bili komentirali da kvazi Monte Carlo metoda ima grešku oblika $CN^{-1}(\log N)^d$, koja je asimptotski (tj. kada $N \rightarrow \infty$) bolja od $CN^{-1/2}$, što je pak greška Monte Carlo metode. Ipak, za konkretni neveliki N i za visoki broj dimenzija d na temelju ocjena grešaka očekujemo da će kvazi Monte Carlo metoda sporije konvergirati od Monte Carlo metode. U praksi se pokazalo da nije nužno tako. Jedna od tipičnih primjena kvazi Monte Carlo metode je u financijama, gdje su integrali u $d = 100$ ili $d = 1000$ varijabli sasvim uobičajeni. Godine 1995. Paskov i Traub su u radu [6] pokazali da kvazi Monte Carlo metoda može davati povoljne rezultate unatoč velikom broju dimenzija; konkretno kod njih je $d = 360$. To je bio jedan od prvih primjera korištenja kvazi Monte Carlo metode u financijama. Oni su procjenjivali vrijednost kolateraliziranih hipotekarnih obveza (collateralized mortgage obligation ili CMO), što je tip financijskog instrumenta u kojem isplata proizlazi iz skupa hipoteka kao kolaterala, u ovom slučaju kroz 30 godina (odnosno 30 puta po 12 mjeseci). Na [6] su nastavili brojni istraživački radovi, a kvazi Monte Carlo metoda i danas ostaje često korišteni alat u financijskoj matematici.

Bibliografija

- [1] Z. Drmač, V. Hari, M. Marušić, M. Rogina, S. Singer, S. Singer, *Numerička analiza: Predavanja i vježbe*, skripta, Sveučilište u Zagrebu, 2003.
- [2] R. Eckhardt, *Stan Ulam, John von Neumann, and the Monte Carlo method*, Los Alamos Science **15**, 1987.
- [3] S. Kurepa, *Matematička analiza 1*, školska knjiga, Zagreb, 1997.
- [4] S. Kurepa, *Matematička analiza 2*, školska knjiga, Zagreb, 1997.
- [5] H. Niederreiter, *Random Number Generation and Quasi-Monte Carlo Methods*, CBMS-NSF Regional Conference Series in Applied Mathematics **63**, SIAM, Philadelphia, 1992.
- [6] S. H. Paskov i J. F. Traub, *Faster valuation of financial derivatives*, J. Portfolio Management **22** (1995), 113–120.
- [7] N. Sarapa, *Vjerojatnost i statistika: I. dio: osnove vjerojatnosti, kombinatorika*, školska knjiga, Zagreb, 1993.
- [8] N. Sarapa, *Vjerojatnost i statistika: II. dio: osnove statistike, slučajne varijable*, školska knjiga, Zagreb, 1996.
- [9] L. Turčić, *Kvazi Monte Carlo metoda za numeričku integraciju*, diplomski rad, Sveučilište u Zagrebu, 2018.

¹Autor je izvanredni profesor na Matematičkom odsjeku PMF-a Sveučilišta u Zagrebu.
E-mail: vjekovac@math.hr.

²Autor je kvantitativni analitičar u SKDD-CCP Smart Clear d.d. E-mail:
ltzs1403@gmail.com.

³Ponekad se, kao sinonim, koristi pojam *uniformno distribuiran*, ali mi želimo naglasiti da je to svojstvo determinističkog niza, a ne niza slučajnih varijabli.

