

UNIVERZA V LJUBLJANI
FAKULTETA ZA MATEMATIKO IN FIZIKO

Finančna matematika – 1. stopnja

Timotej Vesel

**Mrežna pravila in kvazi-Monte Carlo metode za integracijo
funkcij**

Delo diplomskega seminarja

Mentor: prof. dr./doc. dr. Ime Priimek

Ljubljana, 2019

KAZALO

1. Uvod	4
2. Neki	4
3. Metoda Monte Carlo	4
4. Kvazi-Monte Carlo metode	5

Mrežna pravila in kvazi-Monte Carlo metode za integracijo funkcij

POVZETEK

V povzetku na kratko opiši vsebinske rezultate dela. Sem ne sodi razlaga organizacije dela – v katerem poglavju/razdelku je kaj, pač pa le opis vsebine.

Angleški naslov dela

ABSTRACT

Prevod zgornjega povzetka v angleščino.

Math. Subj. Class. (2010): navedi vsaj eno klasifikacijsko oznako – dostopne so na www.ams.org/mathscinet/msc/msc2010.html

Ključne besede: navedi nekaj ključnih pojmov, ki nastopajo v delu

Keywords: angleški prevod ključnih besed

1. UVOD

2. NEKI

Naš cilj je izračunati integral poljubne integrabilne funkcije f , ki je definiran na s -dimenzionalni hiperkocki $[0, 1]^s$,

$$I(f) = \int_0^1 \cdots \int_0^1 f(x_1, \dots, x_s) dx_1 \cdots dx_s = \int_{[0,1]^s} f(\mathbf{x}) d\mathbf{x}.$$

V praksi takšnega integrala za večino funkcij ne znamo izračunati eksaktno zato želimo izračunati numerični približek. Uporabili bi lahko kakšno deterministično (sestavljeno) integracijsko oziroma kvadraturno pravilo, npr. trapezno pravilo, Simpsonovo pravilo, 3/8 pravilo itd. To bi šlo, če bi bila dimenzija majhna (npr. $s = 4$), pri večjem številu integracijskih spremenljivk pa se srečamo s t.i. *prekletstvom dimenzionalnosti*. Že pri $s = 40$ in samo dveh vozlih v vsaki smeri bi potrebovali 2^{40} , kar je več kot bilijon izračunov vrednosti funkcije. (S povprečnim računalnikom bi to trajalo ??? časa, zato za funkcije z veliko spremenljivkami potrebujemo druge metode.)

3. METODA MONTE CARLO

Metoda Monte Carlo (MC) se uporablja za izračun približka pričakovane vrednosti $\mathbb{E}(X)$ slučajne spremenljivke X . Vzemimo funkcijo $f : (0, 1)^s \rightarrow \mathbb{R}$ in naj bo $\mathbf{u} = (u_1, \dots, u_s) \in (0, 1)^s$ in $\mathbf{U} \sim \mathcal{U}(0, 1)^s$ slučajni vektor (enakomerno porazdeljen na s -razsežni hiperkocki). Potem lahko zapišemo

$$I(f) = \int_0^1 \cdots \int_0^1 f(u_1, \dots, u_s) du_1 \cdots du_s = \int_{[0,1]^s} f(\mathbf{u}) d\mathbf{u} = \mathbb{E}[f(\mathbf{U})]$$

Izrek 3.1 (Krepki zakon velikih števil Kolmogorova). *Naj bodo X_1, X_2, \dots, X_n neodvisne in enako porazdeljene slučajne spremenljivke na skupnem verjetnostnem prostoru, za katerega obstaja pričakovana vrednost $\mathbb{E}(X_i) = \mu$. Tedaj zaporedje vzorčnih povprečij $S_n = \frac{X_1 + \dots + X_n}{n}$ skoraj gotovo konvergira k μ .*

KZVŠ nam torej zagotavlja, da bo za večji n S_n boljši približek za matematično upanje in posledično je

$$\mathbb{E}[f(\mathbf{U})] \approx \frac{f(\mathbf{U}_1) + \dots + f(\mathbf{U}_n)}{n}$$

Klasična Monte Carlo Metoda zato $\mathbb{E}[f(\mathbf{U})]$ oziroma $I(f)$ aproksimira z

$$Q_n(f) = \frac{1}{n} \sum_{i=0}^{n-1} f(\mathbf{U}_i)$$

kjer so $\mathbf{U}_0, \dots, \mathbf{U}_{n-1}$ neodvisni $\mathcal{U}(0, 1)^s$ slučajni vektorji. V implementaciji slučajne vektorje \mathbf{U}_i zamenjamo z vektorji “(psevdo) naključnih števil”. To je zelo preprosta ter široko uporabljena metoda.

Poleg tega opazimo, da je $Q_n(f)$ nepristranska cenilka za $\mathbb{E}[f(\mathbf{U})] = I(f)$, t.j. $\mathbb{E}[Q_n(f)] = \mathbb{E}[f(\mathbf{U})]$:

$$\mathbb{E}[Q_n(f)] = \mathbb{E}\left(\frac{1}{n} \sum_{i=0}^{n-1} f(\mathbf{U}_i)\right) = \frac{1}{n} \sum_{i=0}^{n-1} \mathbb{E}[f(\mathbf{U}_i)] = \mathbb{E}[f(\mathbf{U})].$$

Izračunamo lahko še srednjo kvadratično napako, ki nam pove koliko se približek integrala (cenilka) razlikuje od točne vrednosti integrala:

$$\begin{aligned} SKN[Q_n(f)] &= \mathbb{E}[(Q_n(f) - \mathbb{E}[f(\mathbf{U})])^2] \\ &= \mathbb{E}[(Q_n(f) - \mathbb{E}[Q_n(f)])^2] = \\ &= \text{Var}[Q_n(f)] \\ &= \text{Var}\left(\frac{1}{n} \sum_{i=0}^{n-1} f(\mathbf{U}_i)\right) \\ &= \frac{1}{n^2} \sum_{i=0}^{n-1} \text{Var}[f(\mathbf{U}_i)] \\ &= \frac{\text{Var}[f(\mathbf{U})]}{n}, \end{aligned}$$

kjer je $\sigma_f^2 = \text{Var}[f(\mathbf{U})] = \int_{(0,1)^s} f^2(\mathbf{u}) d\mathbf{u} - I^2(f)$ in zadnja enakost velja, ker so \mathbf{U}_i neodvisni in enako porazdeljeni slučajni vektorji. Parameter σ_f^2 nam ponavadi ni eksplisitno poznan, vendar ga lahko statistično ocenimo z vzorčno disperzijo

$$S^2 = \frac{1}{n-1} \sum_{i=0}^{n-1} (f(\mathbf{U}_i) - Q_n(f))^2.$$

Iz tega sledi, da je koren iz srednje kvadratične napake enak S/\sqrt{n} . Pravimo, da je konvergenca Monte Carlo Metode reda " $1/\sqrt{n}$ ", kar označimo z $\mathcal{O}(1/\sqrt{n})$. Red konvergence je neodvisen od dimenzije problema. Takšen red konvergence v praksi pomeni, da če bi želeli napako prepoloviti, bi za to potrebovali štirikrat več vzorčnih točk (slučajnih spremenljivk).

Izkaže se, da so za probleme nizkih dimenzij deterministična integralska pravila učinkovitejša za gladke funkcije, v višjih dimenzijah pa so Monte Carlo metode veliko hitrejša.

Kljub temu, da so Monte Carlo metode večinoma hitreje od kvadraturnih pravil in so zelo široko uporabljene, pa je red konvergence $\mathcal{O}(1/\sqrt{n})$ velikokrat prepočasen za praktično uporabo. Za zmanjšanje variance obstajajo različne tehnike (npr. stratifikacija), vendar v praksi MC metode pogosto ostanejo prepočasne.

4. KVAZI-MONTE CARLO METODE

Kvazi-Monte Carlo (QMC) metode so deterministična različica Monte Carlo metod. Pri QMC metodah zamenjamo neodvisne slučajne točke \mathbf{U}_i z množico determinističnih točk, ki hiperkocko $[0, 1]^s$ pokrijejo bolj enakomerno.

Recimo, da je naš cilj izračunati:

$$I(f) = \int_{[0,1]^s} f(\mathbf{u}) d\mathbf{u}.$$

Opazimo, da je območje integracije nekoliko drugačno kot pri MC metodi, kjer je vseeno ali vključimo robove hiperkocke. Pri QMC pa so nekater metode takšne, da imajo deterministične točke pogosto koordinate v 0, prav tako to zahtevajo nekatere definicije, ki jih bomo potrebovali za QMC.

QMC ocena zgornjega problema pa je podobna oceni za standardno MC metodo.

Definicija 4.1. (Kvazi-Monte Carlo ocena). Naj bodo $\mathbf{u}_0, \mathbf{u}_1, \dots, \mathbf{u}_{n-1}$ (“deterministično”) izbrane točke v hiperkocki $[0, 1]^s$. QMC ocena integrala $I(f)$ je definirana kot

$$\bar{I}_{QMC} = \frac{1}{n} \sum_{i=0}^n f(\mathbf{u}_i).$$

Definicija 4.2. (Diskrepanca??). Naj bo \mathcal{B} družina vseh podmnožic $B \subset [0, 1]^s$. Diskrepanca množice točk $P = \{\mathbf{x}_1, \dots, \mathbf{x}_n\}$ je definirana kot

$$D_n(P) = \sup_{B \in \mathcal{B}} \left| \frac{\sum_{i=1}^n \mathbb{I}_B(\mathbf{x}_i)}{n} - \lambda(B) \right|, \quad (1)$$

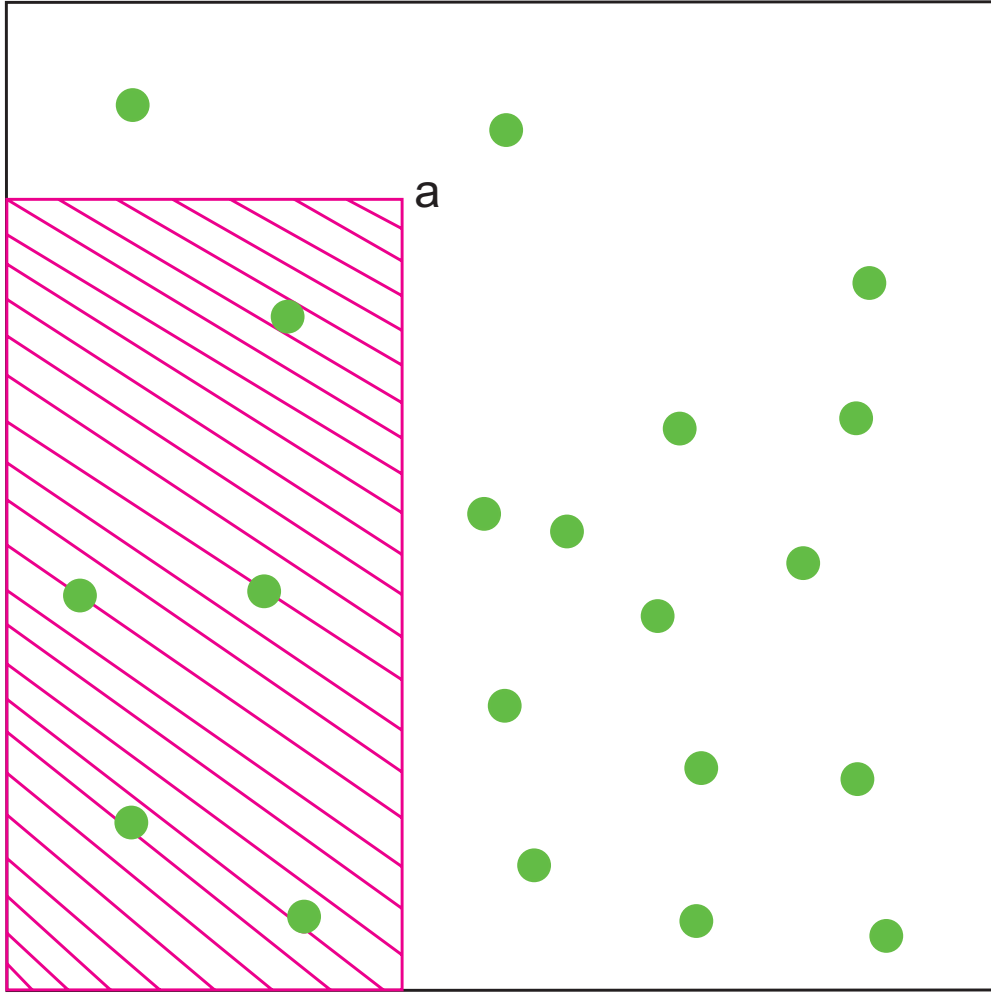
kjer je

$$\mathbb{I}_B(\mathbf{x}_i) = \begin{cases} 1; & \text{točka } \mathbf{x}_i \text{ je v množici } B, \\ 0; & \text{sicer} \end{cases}$$

in $\lambda(B)$ je Lebesgueova mera?? oziroma, ker je B n -dimenzionalna škatla (n-orthotope??), kar n -dimenzionalni volumen??? množice B .

Definicija 4.3. (Star Discrepancy). Naj bo \mathcal{J}^* družina vseh podintervalov $J^* \subset [0, 1]^s$ oblike $\prod_{i=1}^s [0, u_i)$, $0 \leq u_i \leq 1$ (poznani tudi kot ??zasidrane škatle?? anchored boxes), the star discrepancy?? množice točk $P = \{\mathbf{x}_1, \dots, \mathbf{x}_n\}$ v $[0, 1]^s$, $D_n^*(P)$, je definirana kot v enačbi (1) z $\mathcal{J}^* = \mathcal{B}$.

Slika 1 prikazuje to diskrepanco. Na sliki so zasidrana škatla $[0, a) = [0, 0.4) \times [0, 0.8) \in [0, 1]^2$ in množica $n = 20$ točk. Zasidrana škatla vsebuje 5 izmed 20 točk, zato je $\frac{1}{n} \sum_{i=1}^n \mathbb{I}_{0 \leq \mathbf{x}_i < a}(\mathbf{x}_i) = 0.20$. Ploščina (volumen) zasidrane škatle pa je 0.32, zato je $|0.2 - 0.32| = 0.12$. The star diskrepanco D_{20}^* najdemo z maksimiziranjem razlike po vseh zasidranih škatlah $[0, a)$.



Slika 1: V enotskem kvadratu je prikazanih 20 točk in zasidrana škatla (senčeno) od $(0, 0)$ do $a = (0.4, 0.8)$. zasidrana škatla ima ploščino 0.32 in vsebuje $5/20 = 0.2$ vseh točk.

Za $\mathbf{x}_i \sim \mathcal{U}[0, 1]^s$ je Chung pokazal, da je

$$\limsup_{n \rightarrow \infty} \frac{\sqrt{2n} D_n^*}{\sqrt{\log(\log(n))}} = 1, \quad (2)$$

torej je $D_n^* = \mathcal{O}((\frac{\log \log(n)}{n})^{\frac{1}{2}})$ z verjetnostjo 1. Ko se n povečuje, logaritem raste zelo počasi, zato bo D_n^* le malo boljša kot $\mathcal{O}(1/\sqrt{n})$ za velike n .

Znano je, da lahko z deterministično izbiro točk $\{\mathbf{x}_i\}$ dosežemo, da je D_n^* veliko manjši kot (2). Obstajajo neskončna zaporedja $\{\mathbf{x}_i\}$ v $[0, 1]^s$ z $D_n^*(\mathbf{x}_1, \dots, \mathbf{x}_n) = \mathcal{O}(\log(n)^s/n)$. Takšna zaporedja se imenujejo *zaporedja z nizko diskrepanco* (low-discrepancy sequences) in njih na temeljijo Kvazi-Monte Carlo metode.

Definicija 4.4. (Zaporedja z nizko diskrepanco). Zaporedje točk $P = \{\mathbf{x}_n\}_{n \in \mathbb{N}} \subset [0, 1]^s$ se imenuje zaporedje z nizko diskrepanco, če je star discrepancy $D_n^*(P)$ prvih n točk množice P enaka $\mathcal{O}(\frac{\log(n)^s}{n})$.

Obstaja povezava med manjšo diskrepanco in boljšim približkom za integral. Povezuje jo *Hlawka-Koksma neenakost*?

Izrek 4.5 (Neenakost Hlawka-Koksma). *Naj ima funkcija f variacijo? $V(f)$ v smislu Hardy-Krause. Potem za poljubno množico točk $P = \{\mathbf{x}_1, \dots, \mathbf{x}_n\}$ v $[0, 1]^s$ velja*

$$\left| \frac{\sum_{i=1}^n f(\mathbf{x}_i)}{n} - \int_{[0,1]^s} f(\mathbf{x}) d\mathbf{x} \right| \leq V(f) D_n^*(P). \quad (3)$$

Definicija za $V(f)$??? (podana je v [26] Harald Niederreiter. Random Number Generation and Quasi-Monte Carlo Methods. S.I.A.M., Philadelphia, PA, 1992.)