UNIVERZA V LJUBLJANI
FAKULTETA ZA MATEMATIKO IN FIZIKO

Finančna matematika – 1. stopnja

Jakob Svetina, Enja Erker

Aproksimacija povprečne razdalje med točkami v dvodimenzionalnem evklidskem prostoru

Opis projekta pri predmetu Finančni praktikum

Mentorja: prof. dr. Sergio Cabello, asist. dr. Janoš Vidali

Ljubljana, 18. 12. 2020

1. Opis problema

Osrednji proučevani problem projekta pri predmetu Finančni praktikum je aproksimacija povprečne razdalje med točkami v dvodimenzionalnem evklidskem prostoru. Izhodišče predstavlja dvorazsežni evklidski prostor, torej ravnina ($\mathbb{R}^2$), v kateri leži množica n točk. V nadaljevanju bo ta množica označena s S, velja $S = \{P_1, P_2, ...P_n\} \subset \mathbb{R}^2$. Namen naloge je čim bolj ekonomično, s čim manjšo časovno zahtevnostjo, oceniti približek povprečne razdalje med vsemi točkami dane množice v ravnini.

Najpreprostejši pristop k problemu je postopno seštevanje vseh možnih razdalj med danimi točkami. Vendar je ta metoda časovno zelo potratna, njena časovna zahtevnost je $O(n^2)$. Z namenom poiskati metodo, katere časovna zahtevnost bi bila manjša od navedene, je bilo v preteklosti objavljenih kar nekaj študij na to temo. Ključno literaturo, na kateri temelji teoretično ozadje najine naloge, sestavljata dve deli. Gre za članek avtorjev K. Barhum, O. Goldreich in A. Shraibman z naslovom 'On Approximating the Average Distance Between Points' [1] in diplomsko delo avtorja S. Kolar Celarc 'Aproksimacijski algoritmi za računanje povprečne razdalje med točkami' [2]. Pristop, ki ga bova predstavila in implementirala v tem projektu je t. i. 'implementacija objekta' [1] oz. 'naključna projekcija' [2].

Ta pristop je osnovan na konceptu prevedbe problema iz dvodimenzionalnega na enodimenzionalnega. V primeru ene dimenzije je namreč mogoče problem rešiti zgolj v času $O(n \cdot log(n))$. Problem v dveh dimenzijah prevedemo na eno dimenzijo, tako da točke iz ravnine projiciramo na naključno premico z naključnim smernim vektorjem, ki gre skozi izhodišče. Nato seštejemo ustrezne dolžine med projekcijami točk, ki so sedaj kolinearne. Pričakovana vrednost povprečne razdalje med točkami, ki jih projiciramo na naključno premico, se razlikuje od dejanske povprečne razdalje med točkami za konstantni faktor in je neodvisna od izbire premice. Izkaže se, da lahko na premici vsoto vseh razdalj med točkami izračunamo bistveno hitreje kot v splošnem prostoru [2]. Podoben princip deluje tudi v treh in štirih dimenzijah, vendar z drugačnimi korekcijskimi faktorji oz. konstantami.

V projektu bova tako zasledovala dvojni cilj. Prvič, zasnovati natančen algoritem, ki oceni približek povprečne razdalje med danimi točkami v ravnini v čim krajšem času (skoraj linearnem času števila točk). Drugič, odgovoriti na ključno vprašanje glede napake in sicer, koliko eksperimentalnih ponovitev je treba narediti, da bo napaka čim manjša (npr. 5 %, 2 %).

2. Programsko okolje in implementacija

Za implementacijo zadanega problema bova uporabila program R. Ta program sva izbrala, ker je od nama znanih programov/programskih jezikov najbolj primeren za dobro implementacijo. V programu R bova napisala več algoritmov, vse v eno skripto. Tudi vizualizacija pridobljenih rezultatov bo izvedena v programu R.

Glavna ideja implementacije je možnost ponovitve z več različnimi parametri, saj je to zelo pomembno za drugi del projekta.

3. Načrt za delo

V program R bova napisala algoritem za aproksimacijo povprečne razdalje med točkami v dvodimenzionalnem evklidskem prostoru, ki bo slonel na algoritmu iz diplomskega dela 'Aproksimacijski algoritmi za računanje povprečne razdalje med točkami'. Poleg tega bova napisala še časovno bolj zahteven algoritem, ki bo računal točno razdaljo.

Nato bova v programu R izvedla več poskusov, s katerimi bova preverila natančnost aproksimacije. S poskušanjem bova ugotovila, koliko ponovitev algoritma za aproksimacijo je potrebnih, da je povprečna napaka enaka 5 % oziroma 2 %, glede na točno vrednost razdalje.

Literatura

[1] K. Barhum, O. Goldreich in A. Shraibman *On Approximating the Average Distance Between Points*, Lecture Notes in Computer Science (2007) 296–310.

[2] S. Kolar Celarc, *Aproksimacijski algoritmi za računanje povprečne razdalje med točkami*, Delo diplomskega seminarja (2019) 4–17.