Přednáška 8 – Quasi MC & náhodná čísla

Opáčko: viděli jsme metodu Monte-Carlo pro integraci:

- pro $X_1, \dots, X_N$ iid odpovídající rovnoměrnému rozdělení na Ω máme

$$\boxed{\frac{vol(\Omega)}{N} \cdot \sum_{i=1}^{N} f(X_i) \approx \int_{\Omega} f(x)\,dx} \;\&\; \boxed{\mathbb{P}\left(\left| \frac{vol(\Omega)}{N} \sum_{1}^{N} f(X_i) - \int_{\Omega} f(x)\,dx \right| \leq \sqrt{\frac{vol(\Omega)^2 \cdot var(f(x))}{\varepsilon \cdot N}} \right) > 1 - \varepsilon}$$

⇝ aproximace & odhad chyby jsou nezávislé na dimenzi integrálu (tj. na d, kde $\Omega \subseteq \mathbb{R}^d$)

:)

Otevřené otázky/problémy:

- chyba klesá pouze jako $N^{-1/2}$ ... N je # samplovaných bodů ... jak zrychlit?
- jak vlastně generujeme náhodné veličiny & jak to ovlivňuje Monte-Carlo?
- co dělat v případě $vol(\Omega) = +\infty$?
- co když chci samplovat z jiného rozdělení?

Ty otevřené otázky vedou k adaptacím/odnožím MC metod:

- Importance sampling MC
- Quasi-MC
- stratified/adaptive/antithetic MC

Importance sampling

- $w \geq 0$ & $w = 0$ nejvýše pro konečně mnoho bodů
- $\int_{\Omega} w(x)\,dx = 1$

Krok 1: pro libovolné funkce $f(x)$ a $w(x)$, kde $\boxed{w(x) \text{ je pdf}}$ platí

$$\int_{\Omega} f(x)\,dx = \int_{\Omega} \frac{f(x)}{w(x)} \cdot w(x)\,dx \equiv \int_{\Omega} \frac{f(\tilde{x})}{w(\tilde{x})}\, dW_{\tilde{x}}$$

W je pravděpodobnostní míra odpovídající $w(x)$ → → klasicky ztotožňovaná s tzv. cumulative distribution function (cdf)

→ vlastně analogické odvozování na minulé přednášce.

Krok 2: $$\int_{\Omega} f(x)\,dx = \int_{\Omega} \frac{f(x)}{w(x)}\,dW_x = \mathbb{E}_W\left(\frac{f(X)}{w(X)}\right) \approx \frac{1}{N} \sum_{i=1}^{N} \frac{f(X_i)}{w(X_i)}$$

pro iid náhodné veličiny $X_1, \dots, X_n$ z rozdělení daného cdf $W(X)$.

$\rightarrow$ opět postup jako minule:

$X_1, \dots, X_N \sim$ iid náhodné veličiny z rozdělení odpovídajícímu $W(x)$

$Y_N := \frac{1}{N} \sum_{1}^{N} \frac{f(X_i)}{w(X_i)}$ & máme:

- $\mathbb{E}_W(Y_N) = \frac{1}{N} \sum_{1}^{N} \int_{\mathbb{R}} \frac{f(x)}{w(x)} dW(x) = \int_{\mathbb{R}} f(x)\, dx$
- $var_W(Y_N) = \frac{var_W\left(\frac{f(X)}{w(X)}\right)}{N} = \mathbb{E}_W\left(\left(\frac{f(X)}{w(X)}\right)^2\right) - \underbrace{\mathbb{E}_W\left(\frac{f(X)}{w(X)}\right)^2}_{= \mathbb{E}(f(X))^2}$

Pak nám Chebyshevovo lemma dává:

$$\mathbb{P}_W\left(\left|\frac{1}{N}\sum_{1}^{N}\frac{f(X_i)}{w(X_i)} - \int_{\mathbb{R}} f(x)dx\right| < \sqrt{\frac{var_W\left(\frac{f(X)}{w(X)}\right)}{\varepsilon \cdot N}}\right) > 1 - \varepsilon$$

$\leadsto$ přímé & úplné zobecnění MC pro libovolně rozdělené náhodné body.

<u>Potřebujeme:</u>
- umět samplovat body X_i z daného rozdělení
- umět vyhodnocovat pdf daného rozdělení ($= w(X_i)$)

<u>Pozorování:</u>
- vyřešili jsme ty poslední 2 otázky/problémy
- může se stát, že $var_W\left(\frac{f(X)}{w(X)}\right) << var(f(X)) \rightarrow$

  $\rightarrow$ záleží na porovnání $\mathbb{E}_W\left(\left(\frac{f(X)}{w(X)}\right)^2\right)$ vs. $\mathbb{E}(f(X)^2)$.

  V takovém případě jsme i zlepšili přesnost.
- lze psát $\int_{\mathbb{R}} f(x)dx \approx \sum_{i=1}^{N} w_i f(x_i)$ kde $w_i = \frac{1}{N \cdot w(x_i)}$

  $\leadsto$ zjevně tedy stále nejde o klasickou kvadraturu, ale vidíme, že je to stále lineární kombinace funkčních hodnot v různých bodech & ty koeficienty nám zohledňují jak moc je výskyt daného bodu pravděpodobný $\rightarrow$ „je-li x_i 2× více pravděpodobný pak má poloviční váhu" $\rightarrow$

  $\rightarrow$ tudíž název „importance sampling".

Quasi-Monte-Carlo

-> jak v počítači generujeme/simulujeme náhodná čísla?

tzv. pseudo-random generators

-> numpy/scipy/MATLAB/R/.... standard
-> Krok 1 -> vygenerovat sekvenci náhodných čísel v $(0,1)$ (... uniformní)
   Krok 2 -> transf. na normální/exponenciální/....
-> vygenerovaná čísla nejsou opravdu náhodná, ale zkonstruovaná tak, aby šlo dokázat, že splňují klasické testy náhodnosti

Ale pokud si nechám vygenerovat 10^4 bodů v $(0,1)^2$ tak nedostanu (s velkou pravděp.) „rovnoměrné pokrytí" toho $(0,1)^2$.
Pozorujeme něco jako „lokální clusterování".
To lze spravit tím, že nepoužijeme „pseudo-random generators", ale tzv. „quasi-random generators".

-> ty jsou zkonstruovány tak, aby uniformně pokrývali $(0,1)^d$
-> v principu negenerujeme náhodné body, ale deterministické body
-> výhoda je, že uniformní pokrytí $(0,1)^d$ jich potřebujeme **daleko** méně než „uniformní pokrytí $|(0,1)|^d$" ~> tj. řešíme problém, který jsme měli s „interpolačními kvadraturami" pro vysoké d.
-> „uniformní" ≈ „maximalizujeme počet bodů pro daný rozptyl"

quasi Monte-Carlo ≡ X_i nejsou generovány za použití pseudo-random generátorů, ale za použití quasi-random generátorů.

- místo chyby $\mathcal{O}(N^{-1/2})$ lze dokázat $\mathcal{O}\left(\log^d(N) \cdot N^{-1}\right)$

Stratifikované MC $\equiv$ rozdělíme Ω na podoblasti a aplikujeme MC nezávisle

Adaptivní MC $\equiv$ uděláme stratifikované MC a na podoblastech s nejvyšší variancí přidáme další samples

$\rightsquigarrow$ řešíme problém pouze tam kde je

Antithetic MC $\equiv$ využijeme symetrii Ω, pokud nějaká je

- když $X \in (0,1)$ pak pro $Y := 1-X$ také $Y \in (0,1)$
- antithetic MC: $\frac{1}{2N} \sum_{1}^{N} f(X_i) + f(Y_i)$
- platí
  - $\mathbb{E}(\text{antithetic MC}) = \int_0^1 f \, dx$
  - $\text{var}(\text{—}''\text{—}) = \frac{\text{var}(f(X))}{2n} (1+\rho)$

kde ρ je korelace X a Y

$\Rightarrow$ pokud umíme najít jednoduchou transformaci $X \to Y$ která bude mít malou korelaci, pak „zadarmo" získáme „dvojnásobek sample bodů bez zvětšení variance".