Přednáška 8 – Quasi MC & náhodná čísla

Monte-Carlo integrace: $\int_\Omega f(x)\,dx \approx \frac{1}{N}\sum_1^N f(X_i)$

kde $X_1, \dots, X_N$ iid. Pak

$$\mathbb{P}\left(\left| \frac{1}{N}\sum_1^N f(X_i) - \int_\Omega f(x)\,dx \right| \leq \sqrt{\frac{var(f(X))}{\varepsilon \cdot N}} \right) \leq 1-\varepsilon$$

V praxi mám zafixované ε, ale (někdy) si můžu volit

- počet samplingů/pokusů N
- rozdělení, ze kterého vybírám X

$\rightsquigarrow$ chyba klesá (s pravděp. $1-\varepsilon$) $\begin{cases} \text{s rostoucím } N \\ \text{s klesající směr. odchylkou } var(f(X)) \end{cases}$

Na základě těchto pozorování máme několik odnoží Monte-Carlo:

- importance sampling MC
- quasi MC
- stratified / adaptive / antithetic / .... MC

Importance sampling

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

kde
- w je pdf (prob. density function: $w \geq 0$, $\int_{\Omega} w(x)dx = 1$, $w = 0$ pro nejvýše konečně bodů)
- W je cdf (cummulative distribution fun.: $\mathbb{P}(|X| \leq a) = \int_{\Omega} \chi_{|X|\leq a}(x)\, dW(x) \equiv \int_{\Omega} \chi_{|X|\leq a}(x) \cdot w(x)\, dx$)
- platí $\mathbb{E}(f(X)) = \mathbb{E}(\tilde{f}(X^w))$ $\left(:= \int_{\Omega} \tilde{f}(x)\, dW(x) \right)$

K čemu to je?

$$\int_{\Omega} f(x)\,dx = \int_{\Omega} \tilde{f}(x)\,dW(x) \approx \frac{1}{N}\sum_{1}^{N} \tilde{f}(X_i^w) \quad \ldots\ X_i \sim \text{rozdělení } W$$

Stejně jako u MC:

$$\mathbb{P}_W\left(\left| \frac{1}{N}\sum_{1}^{N} \tilde{f}(X_i^w) - \int_{\Omega} \tilde{f}(x)\,dW(x) \right| < \sqrt{\frac{var(\tilde{f}(X^w))}{\varepsilon \cdot N}} \right) = 1 - \varepsilon$$

$$var_w(\tilde{f}(X)) = \mathbb{E}(\tilde{f}^2(X^w)) - \underbrace{\mathbb{E}(\tilde{f}(X^w))^2}_{\mathbb{E}(f(X))^2} = \int_{\Omega} \left(\frac{f(x)}{w(x)}\right)^2 w(x)\,dx - \left(\mathbb{E}(f(X))\right)^2$$

pro porovnání: $var(f(X)) = \int_{\Omega} f^2(x)\,dx - \left(\mathbb{E}(f(X))\right)^2$

Importance sampling se používá, pokud mám $w(x)$ t.,že:
- $var(\tilde{f}(X^w)) \ll var(f(X))$
- umíme snadno vyhodnocovat/simulovat $\tilde{f}(X_i^w)$
- umíme snadno generovat $X_1, \ldots, X_N$ z distribuce W

$$\text{Importance sampling} \equiv \frac{1}{N}\sum_{1}^{N} \frac{f(X_i^w)}{w(X_i^w)}$$

oproti normálnímu MC má každá hodnota svojí váhu ≡ importance

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" $\rightsquigarrow$ tj. řešíme problém, který jsme měli s „interpolačními kvadraturami" pro vysoké d.
-> „uniformní" $\approx$ „maximalizujeme počet bodů pro daný rozptyl"

quasi Monte-Carlo $\equiv$ 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 ≡ rozdělíme Ω na podoblasti a aplikujeme MC nezávisle

Adaptivní MC ≡ uděláme stratifikované MC a na podoblastech s nejvyšší variancí přidáme další samples
⇝ řešíme problém pouze tam kde je

Antithetic MC ≡ 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
  - $\mathbb{E}(\text{antithetic MC}) = \int_0^1 f \, dx$
  - $\text{var}(\text{—''—}) = \frac{\text{var}(f(X))}{2n}(1+\rho)$

kde ρ je korelace X a Y

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