Metody Monte Carlo

Stojíme před následujícím problémem

- potřebujeme spočítat $\int_{a_1}^{b_1}\int_{a_2}^{b_2}\cdots\int_{a_{50}}^{b_{50}} f(x_1,\dots,x_{50})\,dx_1\dots dx_{50}$
- nemůžeme si dovolit klasickou kvadraturu, protože bychom potřebovali 3-4 body „v každé dimenzi" $\to$ odpovídá 3^{50} vyhodnocení f $\leadsto$ příliš časově náročné

Jsme tedy omezeni počtem bodů $\to$ dejme tomu, že chceme aproximovat

$$\left.\begin{array}{ll} \displaystyle\int_\Omega f(\vec{x})\,d\vec{x} & f:\Omega\to\mathbb{R} \\ & \Omega\in\mathbb{R}^d,\ d\in\mathbb{N} \text{ vysoké} \end{array}\right\}\ (\mathrm{I})$$

a máme praktické omezení $\to$ smíme f vyhodnotit nejvýše v N bodech pro nějaké pevně dané $N\in\mathbb{N}$ (např. dáno výkonností počítače). Otázky:

Jak aproximovat (I) na základě $f(\vec{x}_1),\dots,f(\vec{x}_N)$?

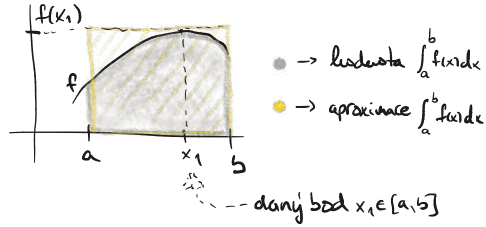
Jak volit $\vec{x}_1,\dots,\vec{x}_N\in\Omega$?

Odvození v $\mathbb{R}$ pro dané $x_1, \dots, x_N$

$\Omega \equiv [a,b]$ & máme dané body $x_1, \dots, x_N \in [a,b]$

• $N=1$ → jak aproximujeme $\int_a^b f(x)\,dx$ na základě $x_1 \in [a,b]$ & $f(x_1) \in \mathbb{R}$?

↝ obdélníkové pravidlo :



$$\Rightarrow \int_a^b f(x)\,dx \approx (b-a)\cdot f(x_1)$$

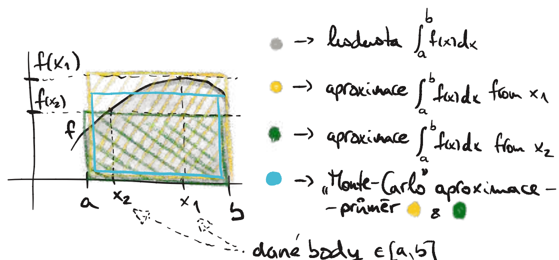
• $N>1$ → nechceme zobecnit „$N=1$" jako u kvadratur:

• problém pro zobecnění do vyšších dimenzí

• nemáme kontrolu nad body $x_1, \dots, x_N$

→ místo toho „$N>1$" ≡ „N-krát $N=1$"

↝ obdélníkové pravidlo :



⇒ když máme $N=2$ aproximace

$$\int_a^b f(x)dx \approx (b-a)\,f(x_1) \quad \& \quad \int_a^b f(x)\,dx \approx (b-a)\,f(x_2) \rightarrow \text{co s nimi ?}$$

„Monte-Carlo": vezmeme jejich průměr → $\boxed{\int_a^b f(x)dx \approx (b-a)\,\frac{f(x_1)+f(x_2)}{2}}$

$$\Rightarrow \int_a^b f(x)\,dx \approx \frac{b-a}{N}\sum_{i=1}^{N} f(x_i)$$ ..... „Monte-Carlo" pro $\Omega = [a,b]$

Zobecnění pro $\mathcal{R} \subseteq \mathbb{R}^d$:

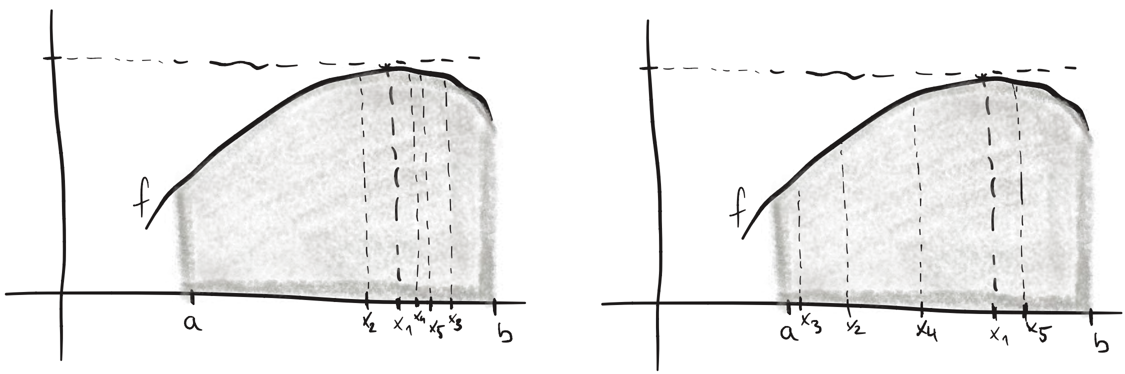
$\mathrm{vol}(\mathcal{R}) \equiv$ „Lebesgueova míra $\mathcal{R}$"
→ přirozené zobecnění délky, obsahu, objemu do $\mathbb{R}^d$

$$\int_{\mathcal{R}} f(\vec{x})\, d\vec{x} \overset{(*)}{\approx} \boxed{\mathrm{vol}(\mathcal{R})} \frac{1}{N} \sum_{i=1}^{N} f(\vec{x}_i) \quad \text{pro } \vec{x}_1, \dots, \vec{x}_N \in \mathcal{R} \subseteq \mathbb{R}^d$$

⇒ Pokud máme $\vec{x}_1, \dots, \vec{x}_N \in \mathcal{R} \subseteq \mathbb{R}^d$ dané, definovali jsme jak na jejich základě aproximovat $\int_{\mathcal{R}} f(\vec{x})\, d\vec{x}$.

Jak volit $\vec{x}_1, \dots, \vec{x}_N$ aby aproximace (*) fungovala dobře?

Která z těchto variant nám dá lepší aproximaci?



⇝ zjevně varianta vpravo, protože se body neakumulují v žádné konkrétní podoblasti $\mathcal{R} = [a,b]$

V našem odhadu $\frac{1}{N} \sum_{i=1}^{N} f(x_i)$ jsou si všechny body „rovnocenné", dává smysl „rovnoceně" pokrývali $\mathcal{R}$. Víme, že brát $\{x_i\}_1^N \subset \mathcal{R}$ „rovnoměrně rozdělené" nechceme

→ <u>vezmeme $\{x_i\}_1^N \subset \mathcal{R}$ jako náhodné body z rovnoměrného rozdělení</u>

Intuitivně: Pokud N není „příliš malé", pak $x_1, \dots, x_N$ pokrývají Ω „rovnoměrně". Ale nemusíme pokrývat „rovnoměrně" každou dimenzi $\rightarrow$ pracujeme automaticky v $\Omega \subseteq \mathbb{R}^d$.

Kombinací myšlenek výše dostáváme Monte-Carlo aproximaci (I).
Další krok je vše formálně zapsat a ukázat, že to opravdu funguje
$\equiv$ chyba $\rightarrow 0$ pro $N \rightarrow +\infty$

Opáčko z teorie pravděpodobnosti

Náhodná veličina $\vec{X}$ na Ω odpovídá rozdělení

- s distribuční funkcí W $\Rightarrow$ $\mathbb{P}(\vec{X} \in \Omega_0) = \int_{\Omega_0} dW(\vec{x}) \quad \forall \Omega_0 \subseteq \Omega$

  $\equiv$ 1. způsob jak zadat rozdělení

  $\equiv$ pravděpodobnost, že realizace náh. vel. X leží v $\Omega_0 \subseteq \Omega$

  integrál vzhledem k pravděpodobnostní míře dané distr. funkcí $W(x)$
  $\equiv \int_{\Omega_0} 1 \, dW(\vec{x}) \equiv \int_{\Omega} \chi_{\Omega_0}(\vec{x}) \, dW(\vec{x})$

- s hustotou $w(\vec{x})$ $\Rightarrow$ $\mathbb{P}(\vec{X} \in \Omega_0) = \int_{\Omega_0} w(\vec{x}) \, d\vec{x} \quad \forall \Omega_0 \subseteq \Omega$

  $\equiv$ 2. způsob jak zadat rozdělení

  $\equiv$ pravděpodobnost, že realizace náh. vel. X leží v $\Omega_0 \subseteq \Omega$

  integrál vzhledem k pravděpodobnostní míře dané distr. funkcí $W(x)$
  $\equiv \int_{\Omega} \chi_{\Omega_0}(\vec{x}) \cdot w(\vec{x}) \, d\vec{x}$

Náhodná veličina $\vec{X}$ na Ω má

- střední hodnotu $\mathbb{E}_W(\vec{X}) = \int_{\Omega} \vec{x} \, dW(\vec{x}) \equiv \int_{\Omega} \vec{x} \cdot w(\vec{x}) \, d\vec{x}$

  $\equiv$ expected value / mean

- rozptyl $\mathrm{var}_W(\vec{X}) = \mathbb{E}_W\left((\vec{X} - \mathbb{E}_W(\vec{X}))^2 \right) \equiv \int_{\Omega} |\vec{x} - \mathbb{E}_W(\vec{x})|^2 \, w(\vec{x}) \, d\vec{x}$

  $\equiv$ variance

Spojité funkce náhodných veličin jsou náhodné veličiny.
(ale typicky s jiným rozdělením)

$\vec{X} \equiv \begin{bmatrix} x_1 \\ \vdots \\ x_N \end{bmatrix} \in \mathbb{R}^N$ a $x_i \in \mathbb{R}$ jsou nezávislé náhodné veličiny z rozdělení s hustotami $w_i(x)$ $\forall i$

pak je náhodná veličina $\vec{X}$ z rozdělení s hustotou $w(\vec{x}) \equiv \prod_{i=1}^{N} w_i(x_i)$

Teorie Monte-Carlo

Pro dané $\Omega \subseteq \mathbb{R}^d$ omezenou & $f: \Omega \to \mathbb{R}$ integrovatelnou definujeme odhad integrálu $\int_\Omega f(\vec{x})\,d\vec{x}$ klasickou metodou Monte-Carlo s N body jako

$$I_{MC}(N) := \frac{vol(\Omega)}{N} \sum_{i=1}^{N} f(\vec{X}_i) \quad \text{pro libovolné } N \in \mathbb{N},$$

kde $\{\vec{X}_i\}_{i=1}^N$ jsou nezávislé a odpovídají náhodné veličině na Ω s rovnoměrným rozdělením.

Chyba klasické metody Monte-Carlo odpovídá $\left| \int_\Omega f(\vec{x})\,d\vec{x} - I_{MC}(N) \right|$

Krok 1: $\vec{X}_1, \dots, \vec{X}_N$ jsou nezávislé náhodné veličiny ze stejného rozdělení $\Rightarrow$
tzv. iid (independent, identically distributed)

$\Rightarrow I_{MC}(N)$ je také náhodná veličina & střední hodnota $I_{MC}(N)$ odpovídá

$E(\alpha_1 Y_1 + \alpha_2 Y_2) = \alpha_1 E(Y_1) + \alpha_2 E(Y_2)$

$$\mathbb{E}(I_{MC}(N)) = \mathbb{E}\left(\frac{vol(\Omega)}{N} \sum_{i=1}^{N} f(\vec{X}_i) \right) = \frac{vol(\Omega)}{N} \sum_{i=1}^{N} \mathbb{E}(f(\vec{X}_i)) =$$

$$= \frac{vol(\Omega)}{N} \sum_{i=1}^{N} \int_\Omega f(\vec{x}_i) \cdot w(\vec{x}_i)\,d\vec{x} = vol(\Omega) \cdot \int_\Omega f(\vec{x})\, w(\vec{x})\,d\vec{x} = \int_\Omega f(\vec{x})\,d\vec{x}$$

$w(\vec{x})$ je hustota rovnoměrného rozdělení

$\vec{x}_i$ jsou iid

hustota rovnoměrného rozdělení na Ω odpovídá $1/vol(\Omega)$

$\Rightarrow$ „$I_{MC}(N) \approx \int_\Omega f(\vec{x})\,d\vec{x}$" ve smyslu pravděpodobnosti/střední hodnoty.

Krok 2: Chebyshevovo lemma: $\mathbb{P}(|Z - \mathbb{E}(Z)| \geq \delta) \leq \dfrac{var(Z)}{\delta^2}$

Důkaz: $\mathbb{P}(|Z - \mathbb{E}(Z)| \geq \delta) \equiv \int_\Omega \chi_{\{|z - E(z)| \geq \delta\}}(z)\,dW(z) = \int_\Omega \chi^2_{\{|z - E(z)| \geq \delta\}}(z)\,dW(z)$

$$\leq \int_\Omega \frac{|z - E(z)|^2}{\delta^2}\,dW(z) = \frac{var(Z)}{\delta^2}.$$

Krok 3: Díky Chebyshevovu lemma dostáváme

$$\mathbb{P}\Big(\big|I_{MC}(N) - \underbrace{\int_\Omega f(\vec{x})d\vec{x}}_{\equiv \mathbb{E}(I_{MC}(N))}\big| \geq \delta\Big) \leq \frac{var(I_{MC}(N))}{\delta^2}$$

$\Updownarrow$ $\mathbb{P}(|\alpha| \geq \delta) = p \iff \mathbb{P}(|\alpha| < \delta) = 1-p$

$$\mathbb{P}\Big(\big|I_{MC}(N) - \int_\Omega f(\vec{x})d\vec{x}\big| < \delta\Big) \geq 1 - \frac{var(I_{MC}(N))}{\delta^2}$$

$\Updownarrow$ $\delta \equiv \sqrt{\frac{var(I_{MC}(N))}{\varepsilon}}$

$$\mathbb{P}\left(\big|I_{MC}(N) - \int_\Omega f(\vec{x})d\vec{x}\big| < \sqrt{\frac{var(I_{MC}(N))}{\varepsilon}}\right) \geq 1 - \varepsilon$$

$\equiv$ pravděpodobnost, že chyba odhadu I_{MC} bude menší než $\sqrt{\frac{var(I_{MC}(N))}{\varepsilon}}$ je alespoň $1-\varepsilon$ $\rightsquigarrow$ tj. pro $\varepsilon \ll 1$ je s velkou pravděpodobností chyba přinejhorším $\sqrt{\frac{var(I_{MC}(N))}{\varepsilon}}$.

Krok 4: opáčko pravděpodobnosti $\to$ $var(I_{MC}(N)) \equiv$

$$\equiv var\left(\frac{vol(\Omega)}{N}\sum_{i=1}^N f(X_i)\right) = \frac{vol(\Omega)^2}{N^2} \cdot var\left(\sum_{i=1}^N f(X_i)\right) \overset{*}{=}$$

$*$ X_i jsou iid $\Rightarrow$ $\Rightarrow f(X_i)$ jsou iid & a mají stejný rozptyl

$$= \frac{vol(\Omega)^2}{N^2}\sum_{i=1}^N var(f(X_i)) = \frac{vol(\Omega)^2}{N} var(f(X_1))$$ a tedy dostáváme

$$\mathbb{P}\left(\big|I_{MC}(N) - \int_\Omega f(\vec{x})d\vec{x}\big| < \sqrt{\frac{1}{N}} \cdot \frac{vol(\Omega)\cdot\sqrt{var(f(X))}}{\sqrt{\varepsilon}}\right) \geq 1 - \varepsilon$$

$\equiv$ chyba odhadu $I_{MC}(N)$ klesá alespoň jako $(\varepsilon \cdot N)^{-1/2}$ s pravděpodobností $1-\varepsilon$.