

Metoda Monte Carlo a Markovské řetězce

Ondřej Matyšek, Václav Hudský, Matouš Richter,
Miroslav Holeček, Kryštof Basista

Gymnázium ALTIS, Gymnázium dr. V. Šmejkal,

Gymnázium Tanvald, Masarykovo Gymnázium Plzeň, Gymnázium Josefa Kainara
ondramat007@gmail.com, vasek.hudsky@gmail.com, 2018-richtermatous@gymtan.cz

miroslav.holecek.188b@mgplzen.cz, basista@student.ghlucin.cz

18. 6. 2024, MFF UK, Karlín, Sokolovská 49/83

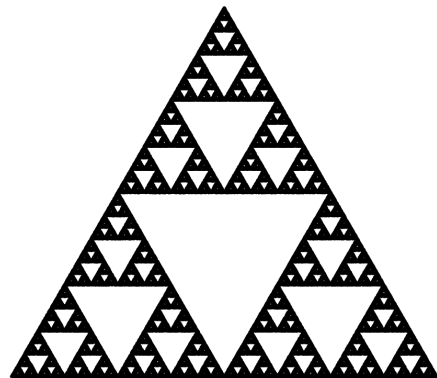
Abstrakt

Tato práce se zabývá řešením úlohy týkající se vzorkováním z předem stanovené Boltzmannovy distribuce (Isingův model) a její analýzy k zodpovězení zadaných otázek za pomoci metody Monte Carlo a Markovských řetězců (zkráceně MCMC). Prvně si představíme základní pojmy a znalosti společně s příklady nutnými k pochopení naší úlohy. V druhé části se podíváme na průběh, metody řešení a výsledky, které jsme získali a jejich interpretaci. Nakonec v závěru shrneme úspěšnost naší práce a její implikace.

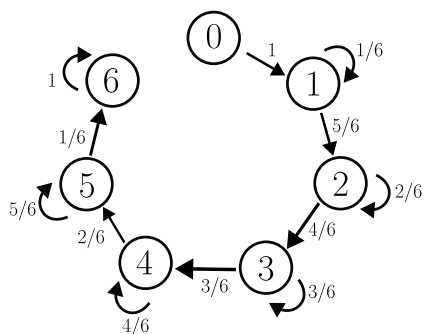
1 Teoretický úvod

Začněme představením jednotlivých termínů a pojmů, které budeme v naší práci používat. Markovský řetězec je množina stavů S popisující náhodný diskretní proces, kde platí, že pravděpodobnost přechodu do následujícího stavu závisí pouze na současném stavu a ne na předchozích stavech, tj. $\mathbb{P}[X_{k+1} = y \mid X_k = x] = p(y|x)$ (funkce závislá pouze na x, y).

Zajímavý příklad výskytu Markovských řetězců je třeba tato úloha, při které nám rozdělení relaxuje na Sierpenského trojúhelník. Úloha je konstruována tak, že si v rovnostranném trojúhelníku ABC zvolíme náhodný bod X_1 , a následně do náhodného vrcholu A, B, C vedeme přímku tímto bodem. Ve středu nově vzniklé úsečky X_1A, X_1B nebo X_1C bude ležet nový bod X_2 . Tento proces následně mnohokrát opakujeme a krásně na něm vidíme, že každý bod je určen pouze tím bodem bezprostředně předešlým a ne celou řadou předchozích. Po velkém množství opakování nám vyjde obrazec totožný Sierpenskému trojúhelníku (obrázek ??).



Obrázek 1: Sierpenského trojúhelník vygenerovaný náhodnými body

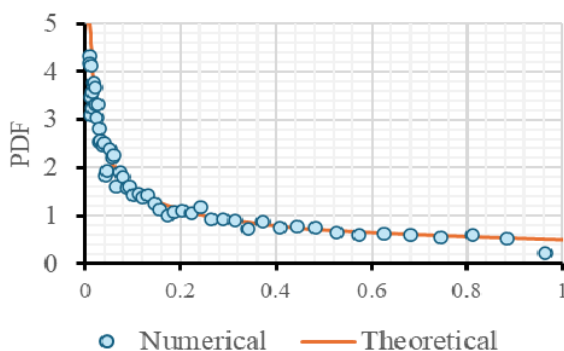


Obrázek 2: Markovův řetězec hodů kostkou

Další zajímavé využití Markovských řetězců je například u úloh s hodem kostkou. Budeme mít takovou úlohu: Kolikrát musí člověk hodit kostkou, aby potkal všechna čísla od 1 – 6? Jednotlivé body znázorňují počet čísel, která jsme už hodili. Šipky nám ukazují, jak velká je pravděpodobnost přechodu do dalšího stavu. Například: po prvním hozeném čísle vidíme, že je šance 1/6, že hodíme to stejné číslo a počet čísel se nezmění a šance 5/6, že se počet čísel již hozených zvýší. Z tohoto diagramu lze už vypočítat průměrný počet hodů, potřebný k potkání všech čísel 1 – 6 (Obrázek ??).

Monte Carlo je numerická metoda využívající vzorkování dat a jejich extrahování za účelem zjištění charakteristik dat. Tohoto se využívá například u počítání kumulativních distribučních funkcí (CDF z ang. „Cumulative distribution function“) $P_X(x) \stackrel{\text{def}}{=} \mathbb{P}[X \leq x]$ za pomoci mnoha funkčních hodnot neznámé hustoty (PDF z ang. „Probability distribution function“) $p_X(x) \stackrel{\text{def}}{=} \frac{d}{dx}(P_X(x))$ a počítání jejich relativních četností a jejich zařazování do uměle vytvořených intervalů.

Příklad použití této metody je například na zjištění $P_Y(y)$, jestliže víme: $Y = X^2$ a $X \sim \text{Unif}(0, 1)$. Po mnoho iteracích a procesu popsaném výše dostaneme naši výslednou $P_Y(y) \approx \sqrt{y}$. Dále po numerické derivaci dostaneme distribuční funkci $p_Y(y) \approx \frac{1}{2\sqrt{y}}$, kterou vidíte na Obrázku ?? . Rovněž pomocí této metody můžeme vzorkovat data a zjistit jejich očekávanou hodnotu $\mathbb{E}X \stackrel{\text{def}}{=} \lim_{n \rightarrow \infty} \bar{x}_n$, kde \bar{x}_n je aritmetický průměr. Při zpracování dat je také důležité řešit standardní odchylku $s_n = \sqrt{\frac{1}{n-1} \sum_i (x_i - \bar{x})^2}$. Ta udává variabilitu dat od daného průměru.



Obrázek 3: distribuční funkce

Kombinací těchto dvou metod je výše zmiňovaná MCMC (Monte Carlo Markov chain) metoda. V této metodě známe počáteční funkci $f(x)$, kdy $f(x) \propto p(x)$ (cílové rozdělení). Uvažujme, že naše rozdělení $p(x)$ je stacionární rozdělení nějakého Markovova řetězce, z něhož budeme vzorkovat. Zvolíme si nějakou počáteční hodnotu x_0 libovolně, a dále budeme generovat řetězec (x_0, x_1, x_2, \dots) . Pro danou hodnotu $x = x_k$ navrhuje přeskok do y z rozdělení $q(\bullet|x)$, které lze volit libovolně. Následující hodnota řetězce x_{k+1} je volena jako y s pravděpodobností $\alpha(y|x)$ (akceptace), jinak volíme znovu $x_{k+1} = x$ (rejekce). Akceptační poměr spočteme pomocí volby Metropolis-Hastings [?]

$$\alpha(y|x) = \min \left\{ 1, \frac{f(y)q(x|y)}{f(x)q(y|x)} \right\}. \quad (1)$$

Jestliže máme spočítaný akceptační poměr, stačí nám si vygenerovat náhodné číslo rovnoměrně rozdělené ($U \sim \text{Unif}(0, 1)$) na intervalu $(0, 1)$ a jestliže toto číslo bude menší, tak y akceptujeme a jestliže ne, tak ho odmítneme.

2 Zjednodušený Isingův model

2.1 Zadání

Pomocí MCMC jsme řešili otázky v následujícím modelu: Je dán prostor bodů $\Lambda = \{0, 1, 2, \dots, 255\}^2$ a každý bod $\lambda \in \Lambda$ může být ve dvou stavech: $x(\lambda) \in \{0, 1\}$. Za pomoci těchto stavů je definována energie daného rozdělení stavů x :

$$E(x) = \sum_{\{\lambda, \lambda'\}: \|\lambda - \lambda'\|=1} |x(\lambda) - x(\lambda')|, \quad (2)$$

kde součet probíhá přes sousední políčka (nahoru/dolu, doleva/doprava). Dále známe, do jaké distribuce bude naše náhodné rozdělení relaxovat

$$\pi(x) = \frac{1}{Z} \exp\left(-\frac{E(x)}{k_B T}\right). \quad (3)$$

V našem modelu volíme $k_B = 1$ a Z volit nemusíme, protože nám stačí úměrnost k rozdělení. Nyní musíme za pomoci těchto znalostí a metody MCMC splnit následující úkoly:

- Vykreslete závislost průměrné energie \bar{E} na teplotě.
- Nechme C být počet černých regionů. Vykreslete závislost očekávané hodnoty C a teploty.
- Nechme L být pravděpodobnost výskytu perkolací (shlukování do uzlů a jejich následná možnost vytvoření řetězce, spojujícího vrchol a spodek viz Obrázek ??). Tuto hodnotu porovnáme s teplotou.

2.2 Řešení

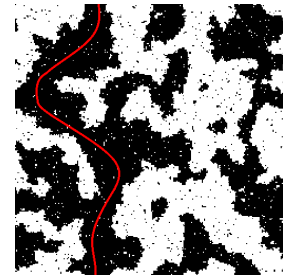
Nyní si do našeho problému zaimplementujeme metodu MCMC. Náhodně si vygenerujeme počáteční stav x_0 (padesátiprocentní šance, že políčko je černé) a spočítáme jeho celkovou energii $E(x_0)$. Následně tvoříme řetěz stavů mřížky (x_0, x_1, x_2, \dots) . Pro daný stav $x = x_k$ v řetězci navrhneme nový stav x' následovně: Náhodně zvolíme políčko $\lambda_0 \in \Lambda$ a změníme jeho barvu, ostatní políčka zůstanou nezměněna. Není třeba počítat celkovou energii $E(x')$, protože se liší od $E(x)$ jen v počtu různých sousedů λ_0 . Pro rozdíl energií platí

$$\Delta(x'|x) = E(x') - E(x) = \sum_{\lambda: \|\lambda - \lambda_0\|=1} (|x'(\lambda) - x'(\lambda_0)| - |x(\lambda) - x(\lambda_0)|). \quad (4)$$

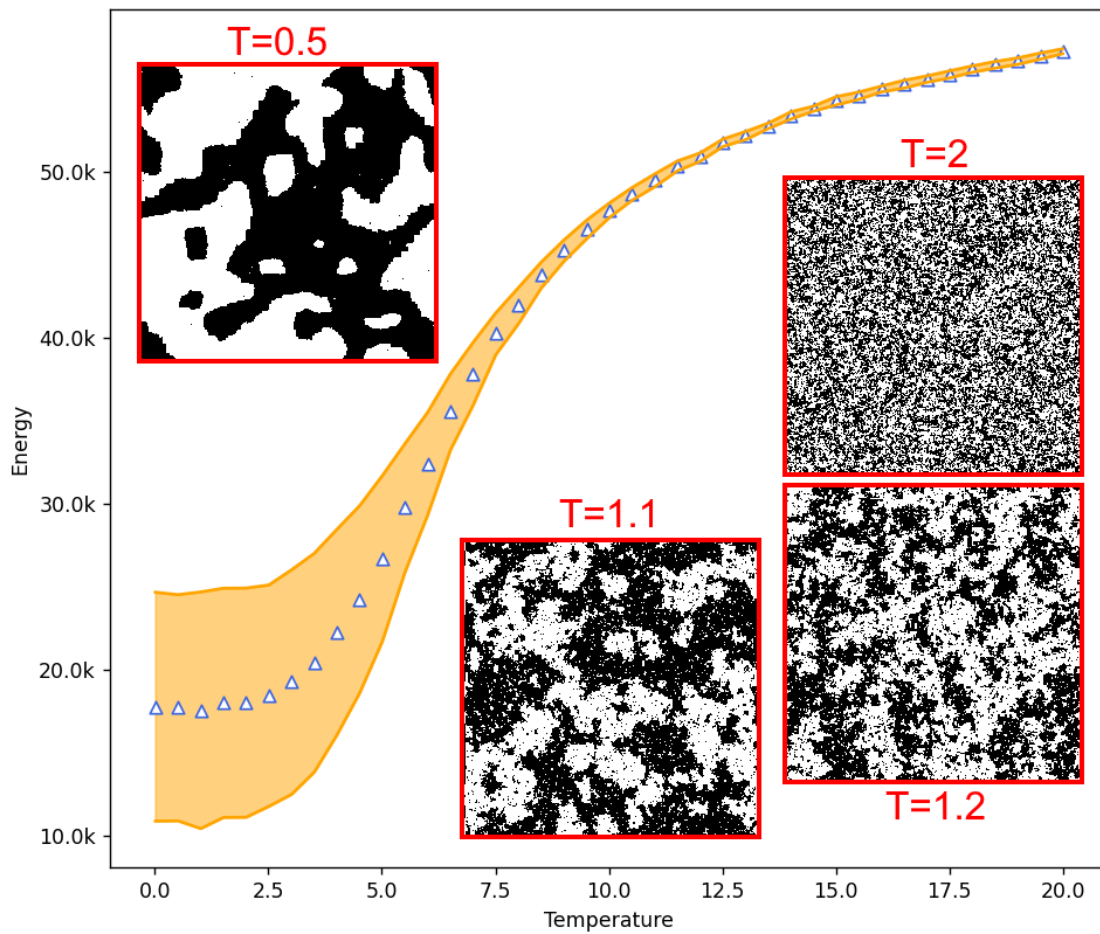
Pomocí tohoto výpočtu dokážeme spočítat akceptační poměr, který se spočte vzorcem:

$$\alpha(x'|x) = \exp\left(\min\left\{0, -\frac{\Delta(x'|x)}{k_B T}\right\}\right) \quad (5)$$

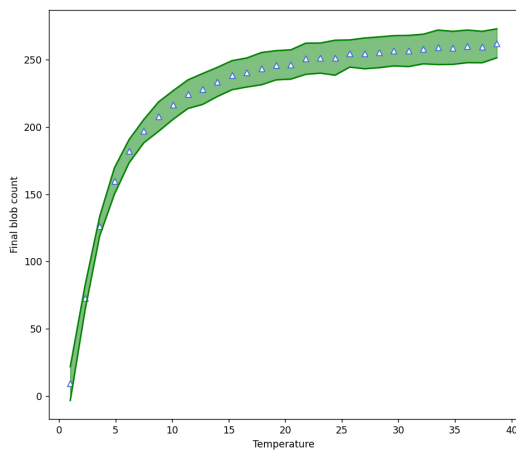
Za pomoci procesu popsaného výše stačí vždy akceptovat či odmítnout x' a toto rozdělení bude relaxovat do rozdělení popsaného výše.



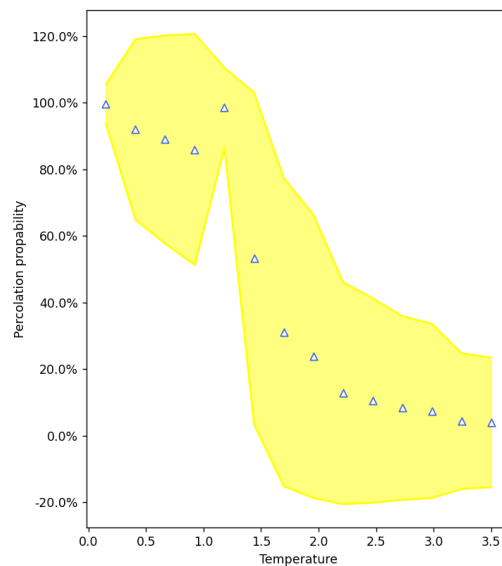
Obrázek 4: Perkolace



Obrázek 5: Klidová rozdělení a střední energie s teplotou



Obrázek 6: Počet blobů s teplotou



Obrázek 7: Perkolace s teplotou

2.3 Výsledky

Výsledky jsme zpracovali tak, že jsme vzali z 10 miliónů iterací poslední milión. Jednotlivé průměry jsou v grafech naznačeny body, vybarvená plocha značí standardní odchylku (viz Obrázky ??, ??, ??)

3 Závěr

V naší práci jsme řešili tři hlavní problémy v Isingově modelu, který se podobá modelu reálného plynu s nekonstatním počtem částic. První úloha se zabývala vykreslováním funkce průměrné energie s teplotou, kdy jsme získali funkci prvně rostoucí strmě a následně pomalu rostoucí. V Isingově modelu si totiž můžeme všimnout, že energie je analogická s entropií, neboť závisí na počtu různých sousedů a tento počet se s rostoucí teplotou bude zmenšovat, neboť částice nemají tolik volných stavů energie resp. nejsou pravděpodobné. To se dá jednoduše vysvětlit, aby částice měla vysokou energii tj. měla hodně různých sousedů, tak při malých teplotách musí vzít hodně energie ostatním bodům a tedy budou částice, které mají hodně energie a ty, co mají hodně málo, ovšem z venku tento stav vypadá jinak než stavy, kde energie je rovnoměrně rozložená, tedy je to jiný makrostav a tento makrostav má mnohem méně způsobu, jak by existoval než když každá částice má přibližně stejné množství energie, neboť v tom případě můžete prakticky zaměnit každou částici s každou. Stejně jako entropie, tak růst energie bude s rostoucím množstvím možných stavů energie zmenšovat a při přechodu do vyšších teplot rapidně poroste. Druhá úloha se zabývá očekávaným počtem blobů na teplotě a lze vidět, jak bylo popsáno výše, že pro malé teploty bude blobů méně, neboť body nemají energii, při dané teplotě se rozdělí, protože to není nejpravděpodobnější. Nakonec, úloha 3 se zabývá pravděpodobností výskytu perkolací. Toto rovněž souvisí s entropií, neboť když se body více kupí, tak je větší šance, že povedou od shora dolů (existuje i kritická Curiova teplota). Tedy jejich počet bude klesat stejně jako množství blobů, což jsme již zjistili.

Poděkování

Děkujeme za odborné vedení miniprojektu našemu garantovi Mgr. et Mgr. Dominiku Beckovi a za jeho trpělivost při řešení všech problémů, na které jsme v průběhu narazili. Dále děkujeme všem organizátorům a sponzorům Týdne Vědy.

Reference

- [1] Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., and Teller, E. *Equation of State Calculations by Fast Computing Machines*. Journal of Chemical Physics, 21, 1087-1092. (1953)
- [2] Lague, S. *[unity] procedural cave generation (E01. cellular automata)*. YouTube, <https://www.youtube.com/watch?v=v7yyZZjF1z4>. (2015)
- [3] Ising, E. *Beitrag zur Theorie des Ferromagnetismus*. Zeitschrift für Physik, 31, 253-258. (1925)