

# Pokroky matematiky, fyziky a astronomie

---

Václav Dupač  
Monte Carlo podle Markova

*Pokroky matematiky, fyziky a astronomie*, Vol. 46 (2001), No. 1, 7--17

Persistent URL: <http://dml.cz/dmlcz/141059>

## Terms of use:

© Jednota českých matematiků a fyziků, 2001

Institute of Mathematics of the Academy of Sciences of the Czech Republic provides access to digitized documents strictly for personal use. Each copy of any part of this document must contain these *Terms of use*.



This paper has been digitized, optimized for electronic delivery and stamped with digital signature within the project *DML-CZ: The Czech Digital Mathematics Library* <http://project.dml.cz>

vyřešit problémy s životností laserů hlavně odstraněním defektů z aktivní oblasti, ale také snížením proudového a teplotního zatížení součástek — znovu s přispěním heterostruktur.

Výběr laureátů nebyl jistě jednoduchý, neboť struktury dnes používané v popisovaných aplikacích jsou již mnohem složitější a o jejich vývoj se zasloužila řada dalších badatelů. Určit skutečné autory těch prvních prapůvodních nápadů také nebývá jednoduché. Jisté je, že letošní držitelé Nobelovy ceny byli v pravý čas na pravém místě, měli ty správné nápady a včas je publikovali.

## L i t e r a t u r a

- [1] HOŘEJŠÍ, J., HOŠEK, J.: *Nobelova cena za fyziku 1999*. Pokroky MFA 45 (2000), 1–6.
- [2] STŘEDA, P.: *Kvantové Hallovy jevy*. Pokroky MFA 44 (1999), 177–186.
- [3] KROEMER, H.: *Quasi-Electric and Quasi-Magnetic Fields in Non-Uniform Semiconductors*. RCA Review 18 (1957), 332.
- [4] KROEMER, H.: *A Proposed Class of Heterojunction Injection Lasers*. Proc. IEEE 51 (1963), 1782.
- [5] ALFEROV, ZH. I., KHAZARINOV, R. F.: *Semiconductor laser with electric pumping*. Patent USSR N181737, March 30<sup>th</sup> (1963).
- [6] ALFEROV, ZH. I. a další: *Effect of the heterostructure parameters on the laser threshold current and the realization of continuous generation at room temperature*. Fiz. Tekh. Polupr. 4 (1970), 1826 (Sov. Phys. Semicond. 4 (1971), 1573);  
HAYASHI, I., PANISH, M. B.: *Junction lasers which operate at room temperature*. Appl. Phys. Lett. 17 (1970), 109.

# Monte Carlo podle Markova

Václav Dupač, Praha

## 1. Simulace náhodné veličiny

Ve statistické fyzice, v matematické statistice i v technice výpočtů se setkáváme s touto úlohou:

$\mathcal{X}$  je konečná množina,  $X$  je náhodná veličina, která nabývá hodnot  $x \in \mathcal{X}$  s vesměs kladnými pravděpodobnostmi  $\pi_x$ . Rozdělení pravděpodobností náhodné veličiny  $X$  označíme stručně  $\pi$ . Máme realizovat (simulovat) hodnotu náhodné veličiny  $X$  nebo

---

Prof. RNDr. VÁCLAV DUPAČ, DrSc. (1929), katedra pravděpodobnosti a matematické statistiky MFF UK, Sokolovská 83, 186 75 Praha 8, e-mail: [dupac@karlin.mff.cuni.cz](mailto:dupac@karlin.mff.cuni.cz)  
Napsáno za podpory grantu GAČR 201/00/0770.

častěji hodnoty  $n$  nezávislých kopií náhodné veličiny  $X$ ; v tomto případě mluvíme o náhodném výběru rozsahu  $n$  z rozdělení  $\pi$ . K dispozici máme náhodnou veličinu  $U$  s rovnoměrným rozdělením na intervalu  $[0, 1]$ , v případě náhodného výběru máme k dispozici  $n$  nezávislých kopií náhodné veličiny  $U$ , říkáme jim náhodná čísla. Za jejich realizace lze s dostatečnou přesností považovat pseudonáhodná čísla, vytvářená počítačovými programy. Očíslujeme prvky množiny  $\mathcal{X}$  čísly  $1, 2, \dots, N$  a ztotožňujeme je na okamžik s těmito čísly.

Pro  $u \in [0, 1]$  definujeme funkci  $g$  s hodnotami v  $\mathcal{X}$ :

$$g(u) = \min\{v : \pi_1 + \dots + \pi_v \geq u\};$$

je pak

$$\begin{aligned} \mathbb{P}(g(U) = x) &= \mathbb{P}(\min\{v : \pi_1 + \dots + \pi_v \geq U\} = x) = \\ &= \mathbb{P}(\pi_1 + \dots + \pi_{x-1} < U \leq \pi_1 + \dots + \pi_x) = \pi_x, \quad x \in \mathcal{X}, \end{aligned}$$

kde  $\mathbb{P}(A)$  značí pravděpodobnost jevu  $A$ .

Transformace  $g$  náhodného čísla  $U$  tedy simuluje náhodnou veličinu  $X$ ; transformace  $g(U_1), \dots, g(U_n)$  simulují náhodný výběr  $X_1, \dots, X_n$  z rozdělení  $\pi$ .

Popsaná metoda je univerzální (pro některá rozdělení existují metody ad hoc, úspornější na počítačový čas). I ona však selhává, je-li počet prvků množiny  $\mathcal{X}$  enormně velký, zejména jsou-li navíc pravděpodobnosti  $\pi_x$  dány až na normující konstantu, kterou je nutno nalézt jejich sečtením přes všechna  $x \in \mathcal{X}$ .

Univerzální metoda je použitelná i pro enormně velká  $\mathcal{X}$  tehdy, je-li pravděpodobnostní míra  $\pi$  součinná, tedy je-li  $\mathcal{X} = \prod_{\alpha=1}^a \mathcal{X}^{(\alpha)}$ ,  $\pi = \prod_{\alpha=1}^a \pi^{(\alpha)}$ , a je-li počet prvků každého  $\mathcal{X}^{(\alpha)}$  nevelký. Náhodnou veličinu  $X = (X^{(1)}, \dots, X^{(a)})$  lze pak simulovat tak, že simulujeme nezávisle její souřadnice. K tomu ovšem potřebujeme náhodné číslo „ $a$ -rozměrné“,  $U = (U^{(1)}, \dots, U^{(a)})$ . Pro pozdější účely bude vhodné chápat náhodné číslo ještě obecněji, jako náhodnou veličinu  $U$  s hodnotami v nějaké množině  $\mathcal{U}$  takovou, že pro vhodnou funkci  $g : \mathcal{U} \rightarrow \mathcal{X}$  má  $g(U)$  stejné rozdělení jako  $X$ .

## 2. Metoda MCMC

Není-li pravděpodobnost  $\pi$  součinná, lze postupovat takto: Najdeme matici pravděpodobností přechodu  $\mathbf{P} = (p_{xy} : x, y \in \mathcal{X})$  tak, že Markovův řetězec  $(X_t, t \geq t_0)$  s množinou stavů  $\mathcal{X}$ , popsaný vztahy

$$\mathbb{P}(X_{t+1} = y \mid X_t = x, X_{t-1} = x_{t-1}, X_{t-2} = x_{t-2}, \dots) = p_{xy},$$

- (i) je nerozložitelný (z každého stavu je po konečně mnoha krocích dosažitelný každý stav),
- (ii) je neperiodický (neexistuje  $d > 1$  tak, že návraty do výchozího stavu jsou možné jen po násobcích času  $d$ ),

- (iii) má v každém řádku matice  $\mathbf{P}$  jen nevelký počet kladných prvků,
- (iv) má stacionární rozdělení  $\pi$ , tj. pro všechna  $x, y \in \mathcal{X}$  platí

$$P(X_t = y \mid X_{t_0} = x) \rightarrow \pi_y \quad \text{pro } t \rightarrow \infty.$$

Definujeme zobrazení  $\varphi(x, u)$  z  $\mathcal{X} \times \mathcal{U}$  do  $\mathcal{X}$  tak, že je-li  $U$  náhodné číslo, pak

$$P(\varphi(x, U) = y) = p_{xy}, \quad x, y \in \mathcal{X}.$$

Ještě definujeme složená zobrazení

$$\Phi_{t_1}^{t_2}(x; u_{t_1}, u_{t_1+1}, \dots, u_{t_2-1}) = \varphi(\dots \varphi(\varphi(x, u_{t_1}), u_{t_1+1}) \dots, u_{t_2-1});$$

stručně

$$\Phi_{t_1}^{t_2} = \varphi_{t_2-1} \circ \dots \circ \varphi_{t_1+1} \circ \varphi_{t_1},$$

kde  $\varphi_t$  značí funkci  $\varphi(\cdot, u_t)$ . (Jako identitu dodefinujeme  $\Phi_{t_1}^{t_1}$ .)

Jsou-li  $U_{t_1}, \dots, U_{t_2-1}$  náhodná čísla, pak

$$P(\Phi_{t_1}^{t_2}(x; U_{t_1}, \dots, U_{t_2-1}) = y) = P(X_{t_2} = y \mid X_{t_1} = x).$$

Dále budeme vždy předpokládat, že za druhý argument zobrazení  $\varphi$  a  $\Phi$  jsou dosazena náhodná čísla, ale nebudeme je v zápise vyznačovat.

Náhodná posloupnost  $(\Phi_{t_0}^t(x), t \geq t_0)$  tedy simuluje trajektorii Markovova řetězce  $(X_t, t \geq t_0)$  začínající v čase  $t_0$  ve stavu  $x$ .

Vzhledem k vlastnosti (iv) lze  $\Phi_{t_0}^T(x)$  pro velké  $T$  již považovat přibližně za simulaci náhodné veličiny  $X$  s rozdělením  $\pi$ . Naproti tomu nelze  $\Phi_{t_0}^T(x)$ ,  $\Phi_{t_0}^{T+1}(x)$ ,  $\dots$ ,  $\Phi_{t_0}^{T+n-1}(x)$  považovat ani přibližně za náhodný výběr z rozdělení  $\pi$ , neboť veličiny tvořící Markovův řetězec jsou vzájemně závislé. Tuto závislost bychom mohli zanedbat teprve pro  $\Phi_{t_0}^T(x)$ ,  $\Phi_{t_0}^{T_1}(x)$ ,  $\Phi_{t_0}^{T_2}(x)$ ,  $\dots$  s dostatečně „řidkou“ posloupností okamžiků  $T < T_1 < T_2 < \dots$ .

Celou posloupnost  $\Phi_{t_0}^T(x)$ ,  $\Phi_{t_0}^{T+1}(x)$ ,  $\dots$  simulující  $X_T, X_{T+1}, \dots$  lze však použít k přibližnému výpočtu výrazů typu  $I = \sum_{x \in \mathcal{X}} f(x)\pi_x$ , kde  $f$  je reálná funkce na  $\mathcal{X}$ . Náš řetězec má totiž také ergodickou vlastnost, podle níž časové průměry  $\frac{1}{n} \sum_{t=T}^{T+n-1} f(X_t)$  konvergují pro  $n \rightarrow \infty$  s pravděpodobností 1 k součtu  $I$ , a to přesto, že  $X_t$  nejsou nezávislé.

Popsané metodě simulace náhodného výběru pomocí Markovova řetězce je více než půl století. Atomovým fyzikům v USA byla známa již ve čtyřicátých letech; první veřejná publikace [2] je z roku 1953. V literatuře se metoda označuje zkratkou MCMC (Markov Chain Monte Carlo). Kniha [1] je téměř encyklopedií teorie i užití metody MCMC.

### 3. Gibbsův výběr; Metropolisův algoritmus

Podstatná je ovšem volba matice  $\mathbf{P}$ . Požadavkům na ni lze vyhovět zejména tehdy, je-li  $\mathcal{X} = \prod_{\alpha=1}^a \mathcal{X}^{(\alpha)}$ , kde počet prvků každé  $\mathcal{X}^{(\alpha)}$  je nevelký; přitom  $\pi$  obecně *není* součinnová pravděpodobnost. Matici  $\mathbf{P}$  volíme pak ve tvaru  $\mathbf{P} = (1/a) \sum_{\alpha=1}^a \mathbf{P}^{(\alpha)}$ , kde  $\mathbf{P}^{(\alpha)}$  jsou rovněž matice pravděpodobností přechodu.

Ty lze volit například tak, že

$$p_{xy}^{(\alpha)} = \frac{\pi_y}{\sum \pi_y}$$

pro stavy  $y$ , které se ve všech souřadnicích shodují se stavem  $x$ , nejvýš vyjímaje souřadnici  $\alpha$ -tou, a kde také součet ve jmenovateli se rozumí jen přes tato  $y$ ; pro všechny ostatní stavy  $y$  jsou  $p_{xy}^{(\alpha)} = 0$ . Takováto volba se nazývá *Gibbsův výběr*. Znamená, že nejprve volíme náhodně  $\alpha \in \{1, 2, \dots, a\}$ , kteroukoli hodnotu s pravděpodobností  $1/a$ ;  $\alpha$ -tou souřadnici pak opravíme (nebo ponecháme) ve shodě s podmíněným rozdělením  $\pi$  za podmínky, že všechny ostatní souřadnice stavu  $x$  jsou zachovány. Ty pak spolu s opravenou  $\alpha$ -tou souřadnicí tvoří stav  $y$ .

Nebo volíme  $p_{xy}^{(\alpha)}$  takto ( $N^{(\alpha)}$  je počet prvků množiny  $\mathcal{X}^{(\alpha)}$ ):

$$p_{xy}^{(\alpha)} = \frac{1}{N^{(\alpha)} - 1} \min\left(\frac{\pi_y}{\pi_x}, 1\right) \quad \text{pro výše popsaná } y \neq x,$$

$$p_{xx}^{(\alpha)} = 1 - \sum_{y \neq x} p_{xy}^{(\alpha)}; \quad p_{xy}^{(\alpha)} = 0 \quad \text{jinak.}$$

Toto je *Metropolisův algoritmus*. I zde volíme náhodně  $\alpha$ . Pak ještě volíme náhodně hodnotu, na kterou by se mohla  $\alpha$ -tá souřadnice změnit. Takto změněný stav (označme jej  $y$ ) přijmeme, je-li  $\pi_y \geq \pi_x$ ; nebo jej přijmeme jen s pravděpodobností  $\pi_y/\pi_x$ , je-li  $\pi_y < \pi_x$ , a s opačnou pravděpodobností zůstaneme ve stavu  $x$ .

V obou algoritmech se místo náhodné volby  $\alpha$ , tj. místo  $\mathbf{P} = \frac{1}{a} \sum_{\alpha=1}^a \mathbf{P}^{(\alpha)}$ , používá také, a to častěji, systematická volba  $\alpha = 1, 2, \dots, a$ , tj.  $\mathbf{P} = \mathbf{P}^{(1)}\mathbf{P}^{(2)} \dots \mathbf{P}^{(a)}$ .

Gibbsův i Metropolisův algoritmus skutečně definují dva Markovovy řetězce s týmž stacionárním rozdělením  $\pi$ . K tomu stačí ověřit tzv. podmínku vyváženosti pro jednotlivá  $\mathbf{P}^{(\alpha)}$ :

$$\pi_x p_{xy}^{(\alpha)} = \pi_y p_{yx}^{(\alpha)}, \quad x \neq y;$$

po dosazení za  $p_{xy}^{(\alpha)}$ ,  $p_{yx}^{(\alpha)}$  je však v obou případech její splnění zřejmé.

### 4. Příklad

Množinu  $\mathcal{X}$  tvoří všechny  $a$ -členné posloupnosti  $x = (x^{(1)}, \dots, x^{(a)})$  čísel 1 a  $-1$ . Rozdělení pravděpodobností  $\pi$  je definováno vzorcem

$$\pi_x = c \exp \left\{ \beta \sum_{\alpha=0}^a x^{(\alpha)} x^{(\alpha+1)} \right\},$$

kde  $c$  je normující konstanta,  $\beta > 0$  je parametr rozdělení a  $x^{(0)} = x^{(\alpha+1)} = 1$ . Parametr  $\beta$  je jakousi mírou závislosti (setrvačnosti) mezi sousedními souřadnicemi; snadno nahlédneme, že pro  $\beta \rightarrow 0$  se rozdělení  $\pi$  blíží rovnoměrnému rozdělení na  $\mathcal{X}$ , kdežto pro  $\beta \rightarrow \infty$  se  $\pi$  blíží k rozdělení soustředěnému v bodě  $(1, 1, \dots, 1)$ .

Dosadíme-li do vzorce pro Gibbsův výběr a označíme-li

$$E = \exp\{-2\beta(x^{(\alpha-1)} + x^{(\alpha+1)})\},$$

dostaneme

$$p_{xy}^{(\alpha)} = \begin{cases} (1 + E)^{-1} & \text{pro } y^{(\alpha)} = 1 \\ E(1 + E)^{-1} & \text{pro } y^{(\alpha)} = -1 \end{cases}$$

a ostatní souřadnice  $y$  shodné s  $x$ .

Pro Metropolisův algoritmus dostaneme (s tímž významem  $E$ )

$$p_{xy}^{(\alpha)} = \min(E^{x^{(\alpha)}}, 1) \quad \text{pro } y^{(\alpha)} = -x^{(\alpha)}$$

a ostatní souřadnice  $y$  shodné s  $x$ ;

$$p_{xx}^{(\alpha)} = 1 - p_{xy}^{(\alpha)} \quad \text{s hořejším } y.$$

Popsaný model má fyzikální význam hlavně ve své dvoj- a trojrozměrné podobě (ferromagnetismus, rozpoznávání obrazů).

## 5. Splynutí trajektorií v konečném čase

Jak jsme již vyložili, metoda MCMC dává hodnoty  $\Phi_{t_0}^T(x)$ , které mají rozdělení  $\pi$  jen asymptoticky, pro  $T \rightarrow \infty$ . V konečném čase  $T$  je stále přítomen vliv počátečního stavu  $x$ . Označme  $\Phi_{t_0}^t = (\Phi_{t_0}^t(1), \dots, \Phi_{t_0}^t(N))$ , kde jsme opět očíslovali stavy  $x \in \mathcal{X}$  čísly  $1, 2, \dots, N$ . Realizovat posloupnost  $(\Phi_{t_0}^t, t \geq t_0)$  pak znamená simulovat současně  $N$  trajektorií řetězce, začínajících v čase  $t_0$  ve stavech  $x = 1, 2, \dots, N$ , a to pomocí téže posloupnosti náhodných čísel  $(U_t, t \geq t_0)$ .

Všechny tyto trajektorie s pravděpodobností 1 v konečném čase splynou a nadále se pak pohybují jako jediná trajektorie. Přesvědčme se o tom; pro jednoduchost nechť  $t_0 = 0$ . Podle vlastnosti (iv) existuje  $L \geq 1$  tak, že matice  $\mathbf{P}^L$  má všechny prvky kladné. Je tedy kladná pravděpodobnost, řekněme  $\varepsilon$ , že bude

$$\Phi_0^L(1) = \Phi_0^L(2) = \dots = \Phi_0^L(N),$$

tj. že zobrazení  $\Phi_0^L$  bude konstantní. Avšak  $\Phi_0^L$  má totéž rozdělení jako  $\Phi_{2L}^{2L}, \Phi_{2L}^{3L}$  atd., přičemž tato náhodná zobrazení jsou navzájem nezávislá. Je tedy s pravděpodobností 1 aspoň jedno z nich, řekněme  $\Phi_{(K-1)L}^{KL}$ , konstantní, a tedy i

$$\Phi_0^{KL} = \Phi_{(K-1)L}^{KL} \circ \dots \circ \Phi_L^{2L} \circ \Phi_0^L$$

je konstantní, neboli od náhodného okamžiku  $KL$ , ale možná už dříve, je zobrazení  $\Phi_0^t$  konstantní.

Je-li zobrazení  $\Phi$  (s libovolnými indexy) konstantní, pak tímž symbolem  $\Phi$  budeme značit i tuto konstantu. Nyní vyslovme domněnku:

*Od prvního okamžiku  $T$  splynutí všech trajektorií má již náhodná veličina  $Z = \Phi_0^T$  (stejně jako  $Z' = \Phi_0^{T+1}$ ,  $Z'' = \Phi_0^{T+2}$  atd.) rozdělení  $\pi$ .*

Tato domněnka je však *mylná*, jak ukazuje následující protipříklad:

Nechť

$$\mathcal{X} = \{0, 1\}, \quad \mathbf{P} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ 1 & 0 \end{pmatrix};$$

potom  $\pi = (\frac{2}{3}, \frac{1}{3})$ , avšak rozdělení  $Z$  je  $\mathcal{L}(Z) = (1, 0)$ . Je totiž  $T = 1$  a  $Z = 0$ , začínají-li obě trajektorie (vycházející z 0 a z 1) takto:

$$\begin{matrix} 0, & 0 \\ 1, & 0 \end{matrix},$$

to má pravděpodobnost  $\frac{1}{2}$ ; nebo je  $T = 2$  a  $Z = 0$ , pokud začínají

$$\begin{matrix} 0, & 1, & 0 \\ 1, & 0, & 0 \end{matrix},$$

to má pravděpodobnost  $\frac{1}{4}$ ; nebo je  $T = 3$  a  $Z = 0$ , začínají-li

$$\begin{matrix} 0, & 1, & 0, & 0 \\ 1, & 0, & 1, & 0 \end{matrix},$$

to má pravděpodobnost  $\frac{1}{8}$ , atd. Je tedy

$$P(Z = 0) = \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \dots = 1.$$

## 6. Simulace z minula

Přesto lze v konečném čase simulovat náhodný výběr z rozdělení  $\pi$  *přesně*. Umožňuje to elegantní obrat, jehož autory jsou Propp a Wilson [3].

Posloupnost zobrazení  $(\Phi_0^t, t \geq t_0)$  jsme dosud simulovali tak, že jsme všech  $N$  trajektorií (vždy na základě dalšího náhodného čísla) prodloužili od času  $t$  do  $t + 1$ . Formálně zapsáno, postupovali jsme podle předpisu

$$\Phi_0^{t+1} = \varphi_t \circ \Phi_0^t, \quad t \geq 0.$$

Na rozdíl od reálného času může však čas simulací plynout i opačně, od  $t + 1$  do  $t$ , a to až do libovolně velkých záporných hodnot. Vytvářejme tedy náhodná čísla postupně pro  $t = -1, -2, \dots$  a posloupnost zobrazení  $(\Phi_t^0, t < 0)$  podle předpisu

$$\Phi_t^0 = \Phi_{t+1}^0 \circ \varphi_t, \quad t < 0. \quad (*)$$

Opět tak simulujeme  $N$  trajektorií, tentokrát však v každém okamžiku  $t < 0$  začínají všechny ve stavech  $1, 2, \dots, N$  a teprve pak každá z nich naváže na některou z trajektorií začínajících v čase  $t + 1$ . Nyní platí

**Věta.** *S pravděpodobností 1 existuje  $T < \infty$  tak, že  $\Phi_{-T}^0$  je konstantní, přičemž  $T$  je nejmenší takové číslo. Náhodná veličina  $Z = \Phi_{-T}^0 = \Phi_{-T-1}^0 = \Phi_{-T-2}^0 = \dots$  má již rozdělení  $\pi$ .*

*Důkaz.* Rozdělení  $\Phi_{-M}^0$ ,  $M \in \mathbb{N}$ , jsou ovšem stejná jako rozdělení  $\Phi_0^M$ ,  $M \in \mathbb{N}$ , pro která jsme již tvrzení o splynutí trajektorií dokázali; odtud plyne existence  $T$ . Rovnost  $\Phi_{-T}^0 = \Phi_{-T-1}^0 = \dots$  je zřejmá z (\*), kde za  $t$  klademe postupně  $-T-1$ ,  $-T-2$ ,  $\dots$

Zbývá najít rozdělení  $Z$ . Pro  $M \in \mathbb{N}$  definujme  $Z_M = Z \cdot 1_{\{T \leq M\}}$ , kde  $1_A$  značí indikátor jevu  $A$ . Je  $Z_M = \Phi_{-T}^0 \cdot 1_{\{T \leq M\}} = \Phi_{-M}^0 \cdot 1_{\{T \leq M\}}$ . Pro  $M \rightarrow \infty$  je  $\mathcal{L}(\Phi_{-M}^0) \rightarrow \pi$  a  $1_{\{T \leq M\}} \rightarrow 1$  s pravděpodobností 1, tedy  $\mathcal{L}(Z_M) \rightarrow \pi$ . Avšak  $Z_M \rightarrow Z$  s pravděpodobností 1, tedy také  $\mathcal{L}(Z_M) \rightarrow \mathcal{L}(Z)$ . Odtud  $\mathcal{L}(Z) = \pi$ .  $\square$

## 7. Příklad

Nechť je opět

$$\mathcal{X} = \{0, 1\}, \quad \mathbf{P} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ 1 & 0 \end{pmatrix};$$

dále nechť  $\mathcal{U} = \{0, 1\}$ ,  $P(U = 0) = P(U = 1) = \frac{1}{2}$ ,

$$\varphi(x, u) = \begin{cases} 0 & \text{pro } (x, u) = (0, 0), (1, 0), (1, 1), \\ 1 & \text{pro } (x, u) = (0, 1). \end{cases}$$

V následující tabulce, v poli nadepsaném  $U$ , představuje každý řádek všechny posloupnosti náhodných čísel  $(U_t, t < 0)$  začínající tak, jak je vyznačeno;  $T$  a  $Z$  mají význam jako ve větě,  $P$  jsou odpovídající pravděpodobnosti.

TABULKA

$t$	-1	-2	-3	-4	-5	...	$T$	$Z$	$P$
$U$	0						1	0	$\frac{1}{2}$
	1	0					2	1	$\frac{1}{4}$
	1	1	0				3	0	$\frac{1}{8}$
	1	1	1	0			4	1	$\frac{1}{16}$
	1	1	1	1	0		5	0	$\frac{1}{32}$
	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$

Výpočet tabulky:

$$\begin{aligned} U_{-1} = 0 &\implies \Phi_{-1}^0 = \varphi_{-1} = (0, 0) \\ U_{-1} = 1 &\implies \Phi_{-1}^0 = \varphi_{-1} = (1, 0) \\ (U_{-1}, U_{-2}) = (1, 0) &\implies \Phi_{-2}^0 = \Phi_{-1}^0 \circ \varphi_{-2} = (1, 1) \\ (U_{-1}, U_{-2}) = (1, 1) &\implies \Phi_{-2}^0 = \Phi_{-1}^0 \circ \varphi_{-2} = (0, 1) \\ (U_{-1}, U_{-2}, U_{-3}) = (1, 1, 0) &\implies \Phi_{-3}^0 = \Phi_{-2}^0 \circ \varphi_{-3} = (0, 0) \quad \text{atd.} \end{aligned}$$

Odtud  $P(Z = 0) = \frac{1}{2} + \frac{1}{8} + \frac{1}{32} + \dots = \frac{2}{3}$ ,  $P(Z = 1) = \frac{1}{4} + \frac{1}{16} + \dots = \frac{1}{3}$ , tj.  $\mathcal{L}(Z) = \pi$ .



## 8. Monotónní zobrazení

Zdánlivě je popsána metoda v praxi jen stěží použitelná, protože generovat  $N$  trajektorií až do jejich splnutí je pro velká  $N$  časově velmi náročné. V mnoha případech však stačí místo  $N$  trajektorií generovat jen *dvě*. Je to tehdy, platí-li současně:

- (i) V množině  $\mathcal{X}$  lze zavést částečné uspořádání „ $\leq$ “ a existují prvky  $\hat{0}$  a  $\hat{1}$  takové, že  $\hat{0} \leq x \leq \hat{1}$  pro všechna  $x \in \mathcal{X}$ .
- (ii) Zobrazení  $\varphi$  je monotónní, tj. pro  $x \leq y$  je

$$\varphi(x) \leq \varphi(y) \quad \text{s pravděpodobností 1,}$$

neboli  $\varphi(x, u) \leq \varphi(y, u)$  až na  $u$  z množiny  $\mathcal{U}_0$  takové, že  $P(U \in \mathcal{U}_0) = 0$ .

Z definice složených zobrazení  $\Phi_t^0$  vyplývá, že i ona jsou monotónní, a tedy nejmenší  $T$ , pro které je  $\Phi_{-T}^0(\hat{0}) = \Phi_{-T}^0(\hat{1})$ , je rovno nejmenšímu  $T$ , pro které je  $\Phi_{-T}^0$  konstantní. Stačí tedy generovat vždy jen dvě trajektorie, a to ty, které začínají ve stavech  $\hat{0}$  a  $\hat{1}$ .

Také je výhodné neprodłużovat simulace o jednotku času, tedy od  $t$  do  $t - 1$ , nýbrž od okamžiku  $-2^\tau$  rovnou do okamžiku  $-2^{\tau+1}$ ,  $\tau = 0, 1, 2, \dots$ ; přitom ovšem simulace začínající v čase  $-2^{\tau+1}$  musí v časovém intervalu  $[-2^\tau, 0]$  využít ta náhodná čísla, která již byla pro tento časový úsek v předchozím kroku vytvořena.

Byla však již také navržena modifikovaná simulace z minula, která vyžaduje jen jediné čtení posloupnosti náhodných čísel [4].

## 9. Příklad

$\mathcal{X}$  je množina všech pořadí  $a$  prvků, pro jednoduchost všech pořadí čísel  $1, 2, \dots, a$ . Máme generovat náhodný výběr z následujícího rozdělení (má význam v neparametrické statistice):

$$\pi_x \propto q^{\text{inv}_x}, \quad x \in \mathcal{X},$$

kde  $0 < q < \infty$ , symbol  $\propto$  značí úměrnost a  $\text{inv}_x$  značí počet inverzí v pořadí  $x = (x^{(1)}, x^{(2)}, \dots, x^{(a)})$ . K tomu cíli konstruujeme Markovův řetězec na množině stavů  $\mathcal{X}$  takto:

Je-li systém ve stavu  $x$ , pak zvolíme náhodně jeho dvě *sousední* souřadnice  $x^{(i)}, x^{(i+1)}$ ,  $i = 1, \dots, a - 1$ , a s pravděpodobností  $1/(q + 1)$  je uspořádáme podle velikosti vzestupně a s pravděpodobností  $q/(q + 1)$  sestupně; systém tak přejde do stavu  $\varphi(x)$ . Zobrazení  $\varphi$  tedy závisí na „dvojměrném“ náhodném čísle: první určí dvojici  $(i, i + 1)$  a druhé rozhodne, zda se  $x^{(i)}, x^{(i+1)}$  uspořádají vzestupně nebo sestupně. Řetězec je zřejmě nerozložitelný neperiodický. Ověříme, že  $\pi$  splňuje podmínku vyváženosti

$$\pi_x p_{xy} = \pi_y p_{yx} \quad \text{pro všechna } x, y \in \mathcal{X}, x \neq y,$$

kteřá je postačující pro to, aby  $\pi$  bylo stacionárním rozdělením tohoto řetězce.

Zřejmě stačí podmínku ověřit pro pevně zvolenou dvojici  $(i, i + 1)$ ; nechť nejprve  $x^{(i)} < x^{(i+1)}$ . Potom jediné  $y$ , pro které je třeba podmínku ověřit, je to, které vznikne z  $x$  vzájemnou záměnou  $x^{(i)}$  a  $x^{(i+1)}$ . Touto záměnou se však počet inverzí o jednu zvýší, takže  $\pi_y = q\pi_x$ , a po dosazení za  $p_{xy}$  a  $p_{yx}$  je podmínka splněna. Pro  $x^{(i)} > x^{(i+1)}$  je ověření obdobné.

Následující částečné uspořádání v  $\mathcal{X}$  se zdá nejpřirozenější:

$x \leq y$ , právě když pořadí  $x$  vznikne z pořadí  $y$  jedním nebo několikerým vzestupným uspořádáním jeho sousedních prvků.

Při tomto uspořádání však není zobrazení  $\varphi$  monotónní. Přesvědčme se o tom na případě, kdy  $\mathcal{X}$  je množina všech pořadí prvků 1, 2, 3. Nechť  $x = (2, 3, 1)$ ,  $y = (3, 2, 1)$ ; zřejmě je  $x \leq y$ . Uspořádáme-li poslední dva prvky v  $x$  i v  $y$  vzestupně, dostaneme  $(2, 1, 3) \not\leq (3, 1, 2)$ , takže je kladná pravděpodobnost  $1/(q + 1)$ , že nerovnost  $\varphi(x) \leq \varphi(y)$  neplatí.

Zvolíme proto v  $\mathcal{X}$  jiné uspořádání. Nejprve zapíšeme každé pořadí  $x$  pomocí matice  $\mathbf{X}$  rozměru  $a \times a$ , kde v  $i$ -tém sloupci,  $i = 1, \dots, a$ , stojí na prvních  $x^{(i)}$  místech jedničky a na zbývajících nuly. Např.  $(3, 1, 2)$  zapíšeme jako

$$\begin{pmatrix} 1, & 1, & 1 \\ 1, & 0, & 1 \\ 1, & 0, & 0 \end{pmatrix}.$$

Nyní definujeme:

$x \leq y$ , právě když částečné součty (jedniček) v každém řádku matice  $\mathbf{X}$  jsou menší nebo rovny částečným součtům v témže řádku matice  $\mathbf{Y}$ .

Je zřejmé, že jde o částečné uspořádání; roli  $\widehat{0}$  a  $\widehat{1}$  hrají, stejně jako v předchozím případě, pořadí  $(1, 2, \dots, a)$  a  $(a, a - 1, \dots, 1)$ . Zobrazení  $\varphi$  je tentokrát již monotónní:

V matici  $\mathbf{X}$  zvolme sloupce  $i$ ,  $i + 1$  a řádek  $j$ , v něm označme  $s$ ,  $s + \xi_0$ ,  $s + \xi_0 + \xi_1$  částečné součty prvních  $i - 1$ ,  $i$  a  $i + 1$  prvků; je-li  $i = 1$ , položme  $s = 0$ . Stejný význam nechť mají  $t$ ,  $t + \eta_0$ ,  $t + \eta_0 + \eta_1$  v matici  $\mathbf{Y}$ . Zobrazení  $\varphi$  znamená buď vzestupné uspořádání obou dvojic  $(x^{(i)}, x^{(i+1)})$  a  $(y^{(i)}, y^{(i+1)})$ , nebo sestupné; tedy v našem řádku vzestupné či sestupné uspořádání dvojic  $(\xi_0, \xi_1)$  a  $(\eta_0, \eta_1)$ . Operací  $\varphi$  tak přejdou částečné součty  $s + \xi_0$ ,  $t + \eta_0$  na  $s + \min(\xi_0, \xi_1)$ ,  $t + \min(\eta_0, \eta_1)$  při vzestupném uspořádání a na  $s + \max(\xi_0, \xi_1)$ ,  $t + \max(\eta_0, \eta_1)$  při sestupném. Ostatní součty zůstávají beze změny.

Nyní se již pro všechny dvojice  $(\xi_0, \xi_1)$  a  $(\eta_0, \eta_1)$  nul a jedniček snadno ověří, že z  $x \leq y$  plyne  $\varphi(x) \leq \varphi(y)$ . Např. pro  $(\xi_0, \xi_1) = (1, 1)$  a  $(\eta_0, \eta_1) = (1, 0)$  z  $x \leq y$  plyne

$$s \leq t, \quad s + 1 \leq t + 1, \quad s + 2 \leq t + 1,$$

tzn.  $s + 1 \leq t$  a odtud již  $\varphi(x) \leq \varphi(y)$ .

Také v příkladě z odst. 4 stačí simulovat vždy jen dvě trajektorie. Ponecháváme čtenáři jako cvičení ověřit, že Gibbsův výběr i Metropolisův algoritmus definují monotónní zobrazení, je-li v  $\mathcal{X}$  zavedeno částečné uspořádání jako v  $\mathbb{R}^a$ .

## 10. Simulace omezené délky

Generujeme-li skutečně tak daleko z minula, až všechny trajektorie v čase 0 splynou, pak  $Z$  má přesně rozdělení  $\pi$ . Prakticky však musíme délku simulace omezit, řekněme číslem  $M$ , a nenastane-li do té doby splynutí, pak neukončenou simulaci vyřadíme. Realizujeme tak náhodný výběr z rozdělení  $Z$ , podmíněného jevem  $\{T \leq M\}$ . Veličiny  $T$  a  $Z$  však nejsou obecně nezávislé a podmíněné rozdělení  $Z$  se tedy může od nepodmíněného, tj. od rozdělení  $\pi$ , lišit. Přesnost metody pak závisí na tom, jak malá je pravděpodobnost nedokončení simulace, tedy  $P(T > M)$ . Víme, že pro  $M \rightarrow \infty$  jde tato pravděpodobnost k nule; pro praktickou aplikaci metody však potřebujeme její numerický odhad. Ten lze získat experimentálně:

Připomeňme nejprve, že doba do splynutí  $T$  má stejné rozdělení při simulaci do budoucna,  $(\Phi_t^t, t \geq 0)$ , i při simulaci z minula,  $(\Phi_t^0, t < 0)$ , kdežto rozdělení  $\pi$  má jen náhodná veličina  $Z = \Phi_{-T}^0$ , obecně nikoli náhodná veličina  $\Phi_0^T$ .

Nechť nyní  $T_1, \dots, T_m, T$  jsou nezávislé náhodné veličiny, všechny stejně rozdělené jako doba do splynutí. Představujeme si, že  $T_1, \dots, T_m$  byly zjištěny při  $m$  zkušebních simulacích do budoucna a že  $T$  je zamýšlená simulace z minula, pro kterou chceme na základě  $T_1, \dots, T_m$  stanovit  $M$ . Pomůckou k tomu je následující

**Věta.**  $P(T > T_1 + \dots + T_m) \leq 2^{-m}$ .

*Důkaz.* Nechť  $K_1, K_2 \in \mathbb{N}$ . Jevy  $\{\Phi_0^{K_1} \text{ je konstantní}\}$  a  $\{\Phi_{K_1}^{K_1+K_2} \text{ je konstantní}\}$  jsou nezávislé a každý z nich implikuje jev  $\{\Phi_0^{K_1+K_2} \text{ je konstantní}\}$ . Je tedy

$$P(\Phi_0^{K_1} \text{ ani } \Phi_{K_1}^{K_1+K_2} \text{ nejsou konstantní}) \geq P(\Phi_0^{K_1+K_2} \text{ není konstantní}),$$

odtud

$$P(T > K_1) P(T > K_2) \geq P(T > K_1 + K_2);$$

to platí i tehdy, jsou-li  $K_1, K_2$  celočíselné náhodné veličiny, nezávislé na  $T$ . Je tedy

$$P(T > T_1 + \dots + T_m) \leq P(T > T_1)^m,$$

přičemž  $P(T > T_1) \leq \frac{1}{2}$ .  $\square$

Předpokládejme tedy, že délka simulací bude omezena číslem  $\widehat{M}$ , které se určí na základě  $m$  zkušebních běhů jako  $\widehat{M} = T_1 + \dots + T_m$ . Oč se liší rozdělení  $Z$  při takto omezené délce simulací od rozdělení  $\pi$ , udává nerovnost

$$\|\mathcal{L}(Z | T \leq \widehat{M}) - \pi\| \leq \frac{1}{2^m - 1},$$

kde vzdáleností dvou pravděpodobnostních rozdělení  $\varrho$  a  $\sigma$  se rozumí číslo  $\|\varrho - \sigma\| = \sup\{|\varrho(A) - \sigma(A)| : \text{všechny náhodné jevy } A\}$ . Nerovnost plyne z jiné, v počtu pravděpodobnosti známé nerovnosti

$$\|\mathcal{L}(Y | A) - \mathcal{L}(Y)\| \leq \frac{1 - P(A)}{P(A)},$$

kteřá platí pro kařždou náhodnou veličinu  $Y$  a kařždý jev  $A$  s  $P(A) > 0$ . Dosadíme-li zde  $Z$  za  $Y$  a  $\{T \leq \widehat{M}\}$  za  $A$ , máme

$$\|\mathcal{L}(Z | T \leq \widehat{M}) - \pi\| \leq \frac{P(T > \widehat{M})}{1 - P(T > \widehat{M})},$$

cořž je nanejvš rovno  $2^{-m}/(1 - 2^{-m})$  podle předchozí věty.

Místo nerovnosti z věty lze také použít nerovnost

$$P(T > j \max(T_1, \dots, T_m)) \leq \frac{1}{\binom{m+j}{j}}, \quad j \in \mathbb{N},$$

kteřá se dokáže kombinatorickou úvahou. Omezíme-li nyní délku simulací číslem  $\widetilde{M} = j \max(T_1, \dots, T_m)$ , dostáváme

$$\|\mathcal{L}(Z | T \leq \widetilde{M}) - \pi\| \leq \frac{1}{\binom{m+j}{j} - 1}.$$

Např. při volbě  $\widehat{M}$  s  $m = 10$  je odchylka od rozdělení  $\pi$  nejvš  $\frac{1}{1023}$ , při volbě  $\widetilde{M}$  s  $j = 4$ ,  $m = 10$  je odchylka nejvšše  $\frac{1}{1000}$ .

#### L i t e r a t u r a

- [1] GILKS, W. R., RICHARDSON, S., SPIEGELHALTER, D. J. (eds.): *Markov Chain Monte Carlo in Practice*. Chapman & Hall, London 1996.
- [2] METROPOLIS, N., ROSENBLUTH, A. W., ROSENBLUTH, M. N., TELLER, A. H., TELLER, E.: *Equation of state calculations by fast computing machines*. J. Chem. Phys. 21 (1953), 1087–1092.
- [3] PROPP, J., WILSON, D.: *Exact sampling with coupled Markov chains and applications to statistical mechanics*. Random Structures Alg. 9 (1996), 223–252.
- [4] WILSON, D.: *How to couple from the past using a read-once source of randomness*. Random Structures Alg. 16 (2000), 85–113.