Příklady ke cvičení z MCMC – markovské řetězce a MCMC algoritmy

Příklad 1: Nechť P_i jsou markovská jádra a π je invariantní míra pro každé i. Ukažte, že následující operace zachovávají invariantní míru:
a) skládání jader: $P_1P_2(x,A) = \int P_2(y,A)P_1(x,dy)$,
b) míchání jader: $P = \sum_i p_i P_i$, kde $\sum_i p_i = 1, p_i > 0$.

Příklad 2: Nerozložitelnost se nemusí zachovávat při skládání. Uvažujme stavový prostor $S = \{1,2,3\}$ a $P_1(1,2) = P_1(2,3) = P_2(1,3) = P_2(2,1) = 1$, $P_i(3,i) = P_i(3,3) = \frac{1}{2}$ pro $i = 1,2$. Markovské řetězce s přechodovými jádry P_1 i P_2 jsou neperiodické a nerozložitelné se stacionárním rozdělením $\pi(1) = \pi(2) = \frac{1}{4}$, $\pi(3) = \frac{1}{2}$, ale $P_1P_2(1,1) = 1$, takže P_1P_2 negeneruje nerozložitelný řetězec.

Naproti tomu pro míchání jader platí, že směs $P = \sum_i p_i P_i$, $\sum_i p_i = 1, p_i > 0$ φ-nerozložitelných jader je zase φ-nerozložitelná. Dokažte.

A jak je tomu s periodicitou, resp. aperiodicitou? Zachovává se při skládání či míchání jader?

Příklad 3: Uvažujme MH algoritmus se symetrickou náhodnou procházkou na $\mathcal{X} = \mathbb{R}^d$ s přírůstkovou hustotou q_0, která je kladná na nějakém okolí počátku. Budiž $\mathcal{X}^+$ otevřená souvislá podmnožina $\mathbb{R}^d$. MH algoritmem generovaný řetězec je potom π-nerozložitelný a aperiodický. Dokažte.

Příklad 4: Uvažujme pravděpodobnostní rozdělení π na prostoru $\mathcal{X} = \{0,1\}^2$.

$$\pi(\{i\} \times \{j\}) = \pi_{ij}, \quad i,j, \in \{0,1\}.$$

Definujte Gibbsův výběrový plán pro generování z π a spočítejte pravděpodobnosti přechodu vzniklého markovského řetězce. Určete stacionární rozdělení tohoto markovského řetězce. Za jakých podmínek bude stacionární rozdělení i rozdělením limitním?

Příklad 5: Nechť $X = (X_1, X_2)$ má rozdělení z předchozí úlohy s $\pi_{00} = \pi_{11} = \frac{p}{2}$ a $\pi_{10} = \pi_{01} = \frac{1-p}{2}$. Určete marginální rozdělení X_1 a X_2. Spočtěte korelační koeficient ρ složek X_1 a X_2.

Uvažujte markovský řetězec $\{X^{(j)}\}$ vytvořený pomocí Gibbsova výběrového plánu. Ověřte, že platí: $P(X_1^{(j)} = 1 | X_1^{(j-1)} = 1) = P(X_1^{(j)} = 0 | X_1^{(j-1)} = 0) = p^2 + (1-p)^2$. Označme $p_j = P(X_1^{(j)} = 1)$ a ukažte, že $p_j = \rho^2 p_{j-1} + b$, kde $b = 2p(1-p)$, a tedy $p_j = \rho^{2j} p_0 + b(1-\rho^{2j})/(1-\rho^2)$.

Ukažte, že matice pravděpodobností přechodu odpovídající první složce $\{X_1^{(j)}\}$ má vlastní čísla 1 a ρ^2. Takže řád konvergence řetězce je ρ^2 a ergodicita řetězce je zajištěna pro $p > 0$. Ale pokud je p blízko 0 nebo 1, bude i řád konvergence blízko 1 a tedy bude konvergence k limitnímu rozdělení velmi pomalá.

Ukažte, že v limitě $p_j \to \frac{b}{1-\rho^2} = \frac{1}{2}$, ale je-li např. $p = 0.999$ je po sto iteracích $p_{100} = 0.67 p_0 + 0.164$, což může být (v závislosti na volbě p_0) dost daleko od očekávané limity.

Příklad 6: Nechť π je pravděpodobnostní rozdělení na $R = \{1,\dots,n\}$. Metropolisův-Hastingsův algoritmus s návrhovou maticí $Q = (q_{ij})$ má matici přechodu $P = (p_{ij})$. Označme $\alpha_{ij} = \min\left\{\frac{\pi_j q_{ji}}{\pi_i q_{ij}}, 1\right\}$, potom $p_{ij} = \alpha_{ij} q_{ij}$ pro $i \neq j$ a $p_{ii} = 1 - \sum_{j \neq i} \alpha_{ij} q_{ij}$. Pro $q_{ij} = q_j^*$ (nezávislý výběr) označme $w_i = \frac{\pi_i}{q_i^*}$ a předpokládejme, že stavy jsou uspořádány tak, že $w_1 \geq w_2 \geq \cdots \geq w_n$. Výpočtem ověřte, že vlastní čísla matice P pravděpodobností přechodu jsou $\lambda_1 = 1$ a $\lambda_{k+1} = \sum_{j=k+1}^n \pi_j \left(\frac{1}{w_j} - \frac{1}{w_k}\right)$ pro $k \geq 1$. Příslušné pravé vlastní vektory jsou $v_1 = (1,\dots,1)^T$ a $v_{k+1} = (\underbrace{0,\dots,0}_{k-1}, \sum_{j=k+1}^n \pi_j, -\pi_k, \dots - \pi_k)^T$ pro $k \geq 1$. Jaký je geometrický řád konvergence algoritmu? Které volby návrhového rozdělení vedou k rychlé konvergenci?

Příklad 7: Pro cílovou hustotu $f(x) = e^{-x}$ uvažujme Metropolisův algoritmus se symetrickou náhodnou procházkou a návrhovou hustotou $q(x, y) = \frac{1}{2}$ pro $|x - y| < 1$. Ukažte, že algoritmus je geometricky ergodický (můžete použít driftovou funkci $V(x) = e^{sx}$ pro vhodné s). Ukažte, že algoritmus není stejnoměrně ergodický.

Příklad 8: Uvažujme jako cílové rozdělení centrované dvourozměrné normální rozdělení s jednotkovými rozptyly a korelací ρ ($\rho \neq 0, |\rho| < 1$). Označme $X^{(t)} = (X_1^{(t)}, X_2^{(t)})$ markovský řetězec generovaný odpovídajícím Gibbsovým výběrovým plánem. Potom platí $X_1^{(t)}|X_1^{(t-1)} \sim N(\rho^2 X^{(t-1)}, \sigma^2)$ pro nějakou konstantu σ^2. Určete tuto konstantu a ukažte, že markovský řetězec $\{X_1^{(t)}\}$ tvořený jen první souřadnicí generovaných vektorů není stejnoměrně ergodický, ale je geometricky ergodický (můžete použít driftovou funkci $V(x) = x^2$).

Příklad 9: Uvažujme ergodická jádra P_1 a P_2 se stacionárním rozdělením π a budiž P_1 stejnoměrně ergodické. Ukažte, že pak směs jader $\alpha P_1 + (1 - \alpha)P_2$ pro $\alpha \in (0, 1)$ je stejnoměrně ergodická. Navíc, pokud P_1 splňuje minorizační podmínku $M(1, \beta, \mathcal{X}, \nu)$, tak $P_1 P_2$ i $P_2 P_1$ jsou také stejnoměrně ergodická.

Příklad 10:
a) Buď π rozdělení na S^k, kde S je spočetná nebo konečná. Ukažte, že Gibbsův výběrový plán má přechodové jádro odpovídající složení k Metropolis-Hastingsových jader s pravděpodobnostmi přijetí rovnými 1.
b) Pro π rozdělení s hustotou $f(x)$ vzhledem k součinové míře $\mu = \mu_1 \times \cdots \times \mu_d$ na součinovém prostoru $\mathcal{X} = \mathcal{X}_1 \times \cdots \times \mathcal{X}_d$, kde μ_i jsou Lebesgueovy míry na $\mathbb{R}$, neplatí tvrzení z bodu a) pro GVP – ověřte. Nicméně i v tomto případě lze dokázat, že π je stacionární rozdělení pro přechodové jádro odpovídající aktualizaci i-té souřadnice

$$Q_i(x, A) = Q_i((x_1, \ldots, x_d), A) = \int_{\{y_i \in \mathcal{X}_i : (x_1, \ldots, x_{i-1}, y_i, x_{i+1}, \ldots x_d) \in A\}} f(y_i|x_{-i}) \mu_i(\mathrm{d}y_i), \quad A \in \mathfrak{X}.$$

Dokažte.

Příklad 11: (variable-at-a-time MH algoritmus) Chceme generovat z pravděpodobnostního rozdělení s hustotou f na součinovém prostoru $\mathcal{X} = \mathcal{X}_1 \times \cdots \times \mathcal{X}_d$ vzhledem k součinové míře $\mu = \mu_1 \times \cdots \times \mu_d$. Postupně aktualizujeme jednotlivé souřadnice v pevném pořadí následujícím způsobem: generujeme $y \in X_i$ z návrhového rozdělení s hustotou $q_i(x, \cdot)$ vzhledem k μ_i, návrh y přijmeme s pravděpodobností

$$\alpha(x, y) = \min \left\{ \frac{f(x_y) q_i(x_y, x_i)}{f(x) q_i(x, y)}, 1 \right\},$$

kde $x = (x_1, \ldots, x_d)$ a $x_y = (x_1, \ldots, x_{i-1}, y, x_{i+1}, \ldots, x_d)$.
Ukažte, že jednotlivé kroky algoritmu jsou reverzibilní vzhledem k úplným podmíněným rozdělením $f(x_i|x_{-i})$. Stačí toto zjištění k tomu, aby variable-at-a-time MH algoritmus měl stacionární rozdělení s hustotou f?