21.10.16 бммо   VIII

MCMC

$$\int_{\mathcal{D}} f(x)\,dx \approx \frac{1}{N}\sum_n f(x_n), \quad x_n \sim U(\mathcal{D})$$

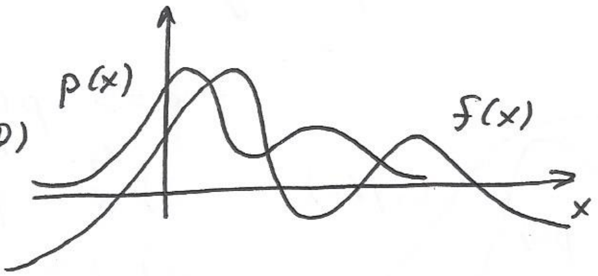
1) $\int p(T|X,w)\,p(w)\,dw$ | $\int_{\mathcal{D}} f(x)\,dx = \mathbb{E} f(x) \approx \frac{1}{N}\sum_n f(x_n)$

2) $p(w|X,T)$, $\int p(t|x,w)\,p(w|X,T)\,dw$

$\mathbb{E}_p f = \int p(x) f(x)\,dx \approx$

$\approx \frac{1}{N}\sum_n p(x_n) f(x_n), \; x_n \sim U(\mathcal{D})$

$\approx \underbrace{\frac{1}{N}\sum_n f(x_n)}_{= J}, \; x_n \sim p(x)$



$\mathbb{E}_p J = \mathbb{E}_p \frac{1}{N}\sum_n f(x_n) = \frac{1}{N}\sum_n \mathbb{E}_p f = \mathbb{E}_p f$

несмещённая оценка на интеграл

$\mathbb{D}_p J = \mathbb{D}_p \frac{1}{N}\sum_n f(x_n) = \frac{1}{N^2}\sum_n \mathbb{D}_p f = \frac{1}{N}\mathbb{D}_p f$

сходимость $O(\frac{1}{N})$, $\mathbb{D}_p f = \int p(x)\left(f(x) - \mathbb{E}_p f(x)\right)^2 dx$

генерирование случайных величин

$\xi \sim U[0,1]$, $\eta = \sum_{i=1}^{12} \xi_i - 6 \approx N(\eta|0,1)$

$F_X(x) = P(X < x)$

$P(F_X(x) < x) = P(X < F_X^{-1}(x)) = F(F^{-1}(x)) = x, \; x \in [0,1]$

$X = F^{-1}(\xi), \; \xi \sim U[0,1]$ | $\xi = F(x)$ $\forall$ ф.р. F $\xi \sim U[0,1]$

1) Показательное распределение, $p(x) = \lambda e^{-\lambda x}, \; x \geq 0, \; \lambda > 0$

~~2)~~ $F(x) = 1 - e^{-\lambda x}$, $1 - e^{-\lambda x} = \xi$, $e^{-\lambda x} = 1 - \xi$, $-\lambda x = \ln(1-\xi)$

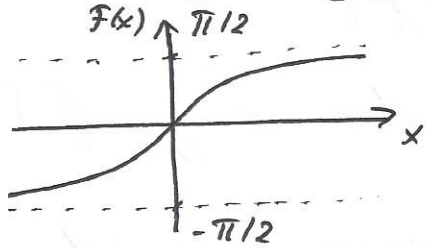
$x = -\frac{1}{\lambda}\ln(1-\xi)$

$x_i = -\frac{1}{\lambda}\ln \xi_i$, $\xi_i \sim U[0,1]$

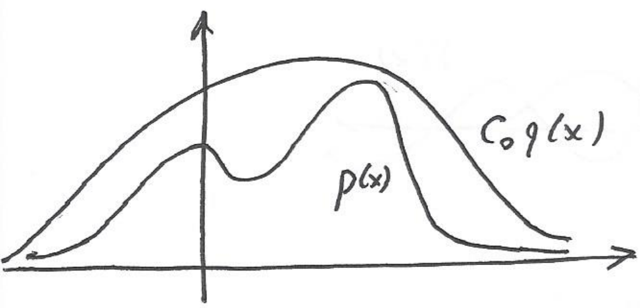
$x_i \sim Exp[\lambda]$

2) Распределение Коши

$$p(x) = \left(\frac{1}{1+x^2}\right)\frac{1}{\pi}, \quad F(x) = \frac{1}{\pi}\left(\operatorname{arctg}(x) + \frac{\pi}{2}\right)$$



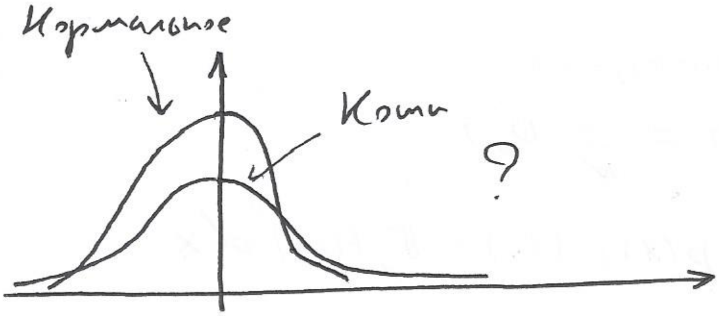
① Сэмплирование с отклонением
Rejection Sampling

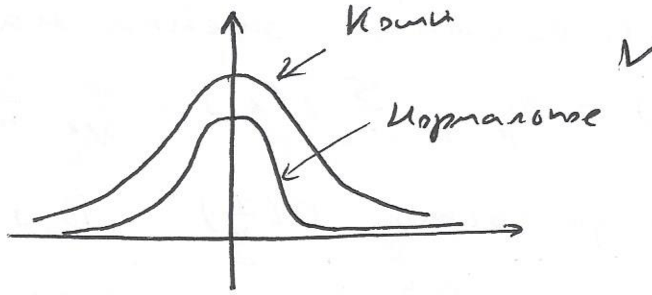


$\forall x \Rightarrow p(x) \leq C_0 q(x)$

$x_1 \dots x_n \sim q(x)$

1) $x' \sim q(x)$
2) $\xi \sim U(0, C_0 q(x'))$
3) Если $\xi < p(x')$, то $x_{n+1} = x'$
   иначе переход к 1)

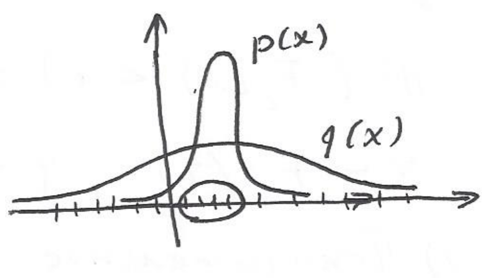




② Importance Sampling. Сэмплирование по важности

$$\mathbb{E}_p f(x) = \int p(x) f(x)\,dx = \int q(x) \frac{p(x)}{q(x)} f(x)\,dx \approx$$

$$\approx \frac{1}{N} \sum_n \underbrace{\frac{p(x_n)}{q(x_n)}}_{w_n} f(x_n), \quad x_n \sim q(x)$$



$x_1 \dots x_N \sim p(x)$, $\quad x_1 \dots x_N \sim q(x)$

$w_1 \dots w_N$

$q(x)$ – proposal distributional

предложенное распределение

$$\mathbb{E}_p f(x) \approx \frac{1}{N} \sum_n f(x_n), \quad x_n \sim p(x)$$

$$p(x) = \frac{1}{Z_p}\hat{p}(x)$$

$$\mathbb{E}_p f(x) = \int p(x) f(x)dx = \int \frac{1}{Z_p}\hat{p}(x) f(x)dx =$$

$$= \frac{1}{Z_p}\int \hat{p}(x) f(x) dx = \frac{\int \hat{p}(x) f(x)dx}{\int \hat{p}(x)dx} = \frac{\int q(x)\frac{\hat{p}(x)}{q(x)} f(x)dx}{\int q(x)\frac{\hat{p}(x)}{q(x)}dx} \approx$$

$$\approx \frac{\sum_n \cancel{q(x_n)} \frac{\hat{p}(x_n)}{q(x_n)} f(x_n) \cancel{dx}}{\sum_n \frac{\hat{p}(x_n)}{q(x_n)}} = \sum_n w_n f(x_n)$$

$$x_n \sim q(x)$$

Марковская цепь $x_1 \dots x_n \dots$

$x_1 \sim p_1(x_1)$, ~~$x_n \sim p_n$~~

$x_n \sim \tau_n(x_n | x_{n-1})$ 1-го порядка

$\tau_n(x_n | x_{n-1}) = \tau(x_n | x_{n-1})$ однородная цепь

$$p(x_1 \dots x_n) = p_1(x_1)\,\tau_2(x_2|x_1)\,\tau_3(x_3|x_2)\dots \tau_n(x_n|x_{n-1})$$

$$p_2(x_2) - ? = \int p(x_1, x_2)dx_1 = \int p_1(x_1)\,\tau_2(x_2|x_1)dx_2$$

действие линейного оператора

$\exists ? \lim\limits_{n\to\infty} p_n(x_n)$ 1) гомогенная цепь
2) эргодичность

$\;\;\;\; \parallel$

$\;\;\; \pi(x)$ $\qquad \int \pi(x_n)\,\tau(x_{n+1}|x_n)dx_n = \pi(x_{n+1})$

предельное стационарное состояние

$\tau(x'|x'') > 0 \;\; \forall x', x'' \in \mathcal{D} \Rightarrow$ эргодичность достаточное условие

1) и 2) $\Rightarrow \forall p_1(x) \; \exists!\, \pi(x)$

Уравнение детального баланса. Достаточное условие.

Гомогенная, эргодичная м.ц. $\tau(x|y)$

Если выполнено у.д.б.: $p(y)\tau(x|y) = p(x)\tau(y|x)$, то $p(x)$ – стационарное распределение марковской цепи.

$$q(x_{n+1}) = \int p(x_n)\,\tau(x_{n+1}|x_n)dx_n = \int p(x_{n+1})\,\tau(x_n|x_{n+1})dx_n =$$

$$= p(x_{n+1})\int \tau(x_n|x_{n+1})dx_n = p(x_{n+1})$$

Схема получения $\pi(x) = \hat{p}(x)$

1) Метрополис - Гастингс

$q(x'|x)$ - proposal distribution > 0

$N(x'|x, \sigma^2 I)$

1) $x' \sim q(x'|x_n)$

2) Принимаем её с вер-ю $A = \min\left(1, \dfrac{\hat{p}(x')\, q(x_n|x')}{\hat{p}(x_n)\, q(x'|x_n)}\right)$

у.д.б. $p(x) r(y|x) = p(y) r(x|y)$ выполнено

$r(y|x) = q(y|x) \cdot A$

$$p(x) \cdot r(y|x) = p(x)\, q(y|x) \cdot \min\left(1, \frac{p(y)\, q(x|y)}{p(x)\, q(y|x)}\right) =$$

$$= \min\left(p(x)\, q(y|x),\ p(y)\, q(x|y)\right) =$$

$$= p(y)\, q(x|y) \cdot \min\left(\frac{p(x)\, q(y|x)}{p(y)\, q(x|y)}, 1\right) = p(y) \cdot r(x|y)$$

Значит, $p(x)$ - стационарное распределение.

2) Гиббс

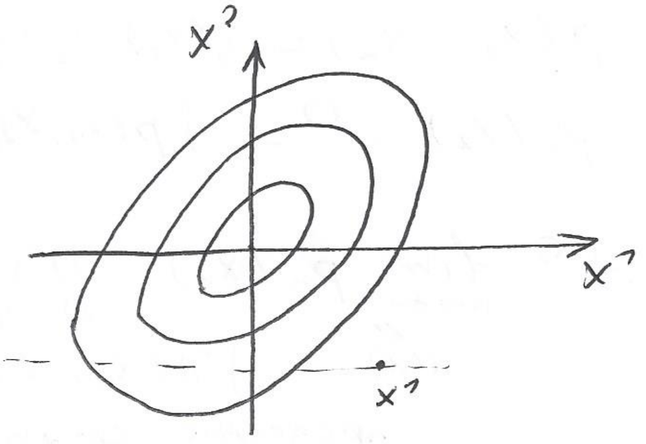
$x_n \in \mathbb{R}^d \qquad x_1 \dots x_N$

$x^1_{n+1} \sim \hat{p}(x^1 | x_n^2 \dots x_n^d)$

$x^2_{n+1} \sim \hat{p}(x^2 | x^1_{n+1}\, x^3_n \dots x^d_n)$

$\dots$

$x^d_{n+1} \sim \hat{p}(x^d | x^1_{n+1} \dots x^{d-1}_{n+1})$



$x_2^1 \sim p(x^1 | x_1^2)$

x

21. 10. 16 бимо сем

$$\tilde{x} \sim q(\tilde{x} \mid x), \quad A = \min\left(1, \frac{p(\tilde{x})\, q(x \mid \tilde{x})}{p(x)\, q(\tilde{x} \mid x)}\right)$$

$$x_{n+1} = \begin{cases} \tilde{x}, & A \\ x_n, & 1-A \end{cases}$$ Метрополис - Гастингс

$$x^m = (x_1^m \dots x_n^m)$$

Гиббс

$$p(x_i^{m+1} \mid x_1^{m+1} \dots x_{i-1}^{m+1}, x_{i+1}^{m+1} \dots x_n^{m+1})$$

$$A \begin{pmatrix} \hat{x} \sim p(\hat{x} \mid y, z) \\ \hat{y} \sim p(\hat{y} \mid \hat{x}, z) \\ \hat{z} \sim p(\hat{z} \mid \hat{x}, \hat{y}) \end{pmatrix} : (x, y, z) \longrightarrow (\hat{x}, \hat{y}, \hat{z})$$

$$p = Ap, \quad p(\hat{x}, \hat{y}, \hat{z}) = \int q(\hat{x}, \hat{y}, \hat{z} \mid x, y, z)\, p(x, y, z)\, dx\, dy\, dz \ominus$$

$p(x, y, z)$ – предельное стационарное распределение.

$$q(\hat{x}, \hat{y}, \hat{z} \mid x, y, z) = p(\hat{x} \mid y, z)\, p(\hat{y} \mid \hat{x}, z) \cdot p(\hat{z} \mid \hat{x}, \hat{y})$$

$$\ominus \int p(\hat{x} \mid y, z) \cdot p(\hat{y} \mid \hat{x}, z) \cdot p(\hat{z} \mid \hat{x}, \hat{y}) \cdot p(x, y, z)\, dx\, dy\, dz =$$

$$= \int \underline{p(\hat{x} \mid y, z)}\, p(\hat{y} \mid \hat{x}, z)\, p(\hat{z} \mid \hat{x}, \hat{y})\, \underbrace{\overset{p(x \mid y, z)\, p(y, z)}{\underline{p(y, z)}}}_{\cancel{p(y \mid z) \cdot p(z)}}\, dy\, dz =$$

$$= \int p(\hat{x}, y, z)\, p(\hat{y} \mid \hat{x}, z)\, p(\hat{z} \mid \hat{x}, \hat{y})\, dy\, dz =$$

$$= \int p(\hat{x}, z)\, p(\hat{y} \mid z, \hat{x})\, p(\hat{z} \mid \hat{x}, \hat{y})\, \cancel{dy}\, dz =$$

$$= \int p(\hat{x}, z, \hat{y})\, p(\hat{z} \mid \hat{x}, \hat{y})\, dz = p(\hat{x}, \hat{y}, \hat{z}) \quad \blacksquare$$

1) $\hat{x} \sim p(\hat{x} \mid y)$, 2) $\hat{y} \sim p(\hat{y} \mid \hat{x})$

$$(x, y) \to (\tilde{x}, y) \to (\tilde{x}, \tilde{y})$$

$$q_1(\tilde{x}, \tilde{y} \mid x, y) = p(\tilde{x} \mid \tilde{y})\, I[y = \tilde{y}]$$

$$q_2(\tilde{x}, \tilde{y} \mid x, y) = p(\tilde{y} \mid \tilde{x})\, I[x = \tilde{x}]$$

\# $$p(X, T, w \mid \theta, \alpha_0) = \prod_{n,k} [p(x_n \mid \theta_k)^{t_{nk}} w_k]^{t_{nk}}\, p(w \mid \alpha_0)$$

$$p(w \mid \alpha_0) = \frac{\Gamma(\alpha_0 K)}{\Gamma^K(\alpha_0)} \prod_k w_k^{\alpha_0 - 1}$$

$p(T, w \mid \Theta, X, \alpha_0)$ $\qquad P(x_i^{m+1} \mid x_1^{m+1} \dots x_{i-1}^{m+1}, x_{i+1}^{m} \dots x_n^{m})$

$\cancel{X = (X_1, X_1 \dots X_n)}$ $\qquad X = (x_1 \dots x_n)$

$p(T, w \mid \Theta, Y, \alpha_0)$, $X = (\underset{\{w''\}}{X_0}\ \underset{\{t_1''\}}{X_1}\ X_2 \dots \underset{\{t_n''\}}{X_n})$

$p(w \mid t_1 \dots t_n, \Theta, Y, \alpha_0)$ ✓

$p(t_i \mid w, t_1, \dots, t_{i-1}, t_{i+1}, \dots, t_n, \Theta, Y, \alpha_0)$ ✓

$p(w \mid T, \Theta, Y, \alpha_0) \propto p(w, T, Y \mid \Theta, \alpha_0) =$

$$= \prod_{n,k} [p(y_n \mid \Theta_k) w_k]^{t_{nk}} \frac{\Gamma(\alpha_0 K)}{\Gamma^K(\alpha_0)} \prod_k w_k^{\alpha_0 - 1} =$$

$$= \left(\prod_k w_k^{\sum_n t_{nk} + \alpha_0 - 1}\right) \frac{1}{Z_1}$$

$$p(t_i \mid w, t_1 \dots t_{i-1}, t_{i+1} \dots t_n, \Theta, Y, \alpha_0) = \prod_k [p(y_i \mid \Theta_k) w_k]^{t_i} \frac{1}{Z_2}$$

$$p(t_{ik} = 1 \mid w, t_1 \dots t_{i-1}, t_{i+1} \dots t_n, \Theta, Y, \alpha_0) = \frac{p(y_i \mid \Theta_k) w_k}{\sum_m p(y_i \mid \Theta_m) w_m}$$

$p(X \mid \Theta) \to p(X, T \mid \Theta)$

$E: p(T \mid X, \Theta_{old})$, $M: \mathbb{E}_{p(T \mid X, \Theta_{old})} p(X, T \mid \Theta) \to \max$

$T_1 \dots T_M$ $\qquad \approx \frac{1}{M} \sum_m p(X, T_m \mid \Theta)$

$\qquad T_m \sim p(T \mid X, \Theta_{old})$

$\ln p(X \mid \Theta) = \int \ln p(X \mid \Theta)\ \cancel{\ln}\ q(T, w)\, dw\, dT$