베이즈통계 및 실습 HW3 2020-15709 나승한

1.9.12.

(a) 제안분포: 독립 메트로폴리스-헤이스팅스 알고리즘이므로 전이핵 $q(\theta^{(t-1)},\theta')=q(\theta')$이다. 즉 $q(\theta)=q(\theta_1,\theta_2)=Ca(0,1)\times Ca(0,1)$
$= \dfrac{I\{\theta_1\in\mathbb{R}\}I\{\theta_2\in\mathbb{R}\}}{\pi^2(1+\theta_1^2)(1+\theta_2^2)}$로 잡는다.

목표분포: $\pi = N_2\left(\begin{pmatrix}0\\0\end{pmatrix}, \begin{pmatrix}1&\rho\\ \rho&1\end{pmatrix}\right)$의 pdf는 $\pi\begin{pmatrix}\theta_1\\\theta_2\end{pmatrix} = \dfrac{\exp\left\{-\frac{\theta_1^2+\theta_2^2-2\rho\theta_1\theta_2}{2(1-\rho^2)}\right\}}{2\pi\sqrt{1-\rho^2}}$

이때 $\dfrac{\pi(\theta')q(\theta',\theta^{(t-1)})}{\pi(\theta^{(t-1)})q(\theta^{(t-1)},\theta')} = \dfrac{\pi(\theta')q(\theta^{(t-1)})}{\pi(\theta^{(t-1)})q(\theta')} = \alpha_0(\theta^{(t-1)},\theta',\rho)$

$$= \frac{(\theta_1'^2+1)(\theta_2'^2+1)}{(\theta_{1,t-1}^2+1)(\theta_{2,t-1}^2+1)}\times\exp\left\{\frac{\theta_{1,t-1}^2+\theta_{2,t-1}^2-\theta_1'^2-\theta_2'^2-2\rho(\theta_{1,t-1}\theta_{2,t-1}-\theta_1'\theta_2')}{2(1-\rho^2)}\right\}$$

즉 합격 확률 $\alpha(\theta^{(t-1)},\theta')=\min(1,\alpha_0(\theta^{(t-1)},\theta',\rho))$이다.
이제 알고리즘을 서술하면 다음과 같다.

(1) initialize $\theta^{(0)}=(\theta_1^{(0)},\theta_2^{(0)})^\dagger$

(2) 메트로폴리스-헤이스팅스 반복: for $i=1,2,\cdots,m^\dagger$

① $\theta'\sim q(\theta')=\dfrac{1}{\pi^2(1+\theta_1'^2)(1+\theta_2'^2)}$

$u\sim Unif(0,1)$, $u\perp\theta'$인 u,θ'을 추출한다.

② 합격 확률 계산: ①에서 추출된 θ'와 직전 단계에서 추출된 $\theta^{(t-1)}$을 통해 합격 확률 $\alpha(\theta^{(t-1)},\theta')=\min\{1,\alpha_0(\theta^{(t-1)},\theta',\rho^\dagger)\}$를 계산한다.

③ 합격-불합격 단계

$$\theta^{(t)}=\begin{cases}\theta' & \text{if } u\le\alpha(\theta^{(t-1)},\theta')\\ \theta^{(t-1)} & \text{if } u>\alpha(\theta^{(t-1)},\theta')\end{cases}$$

(3) MC $\{\theta^{(t)}: t=1,2,\cdots,m\}$으로 $\pi(\theta)$를 근사한다.

알고리즘에서 † 표시된 m, $\theta^{(0)}$, ρ는 적당한 숫자를 대입해 실행한다. 우선 $m=5000$을 수행하고, $\theta^{(0)}=\begin{pmatrix}0\\0\end{pmatrix}$으로 초기화하여 (C)를 풀이할 것이다. ρ는 0.3, 0.99 각각 풀이한다.

(b) ~ (d): R markdown 참고.

1.9.13.

목표분포 $\pi\begin{pmatrix}\theta_1\\\theta_2\end{pmatrix}=\dfrac{1}{2\pi\sqrt{1-\rho^2}}\exp\left\{-\dfrac{\theta_1^2+\theta_2^2-2\rho\theta_1\theta_2}{2(1-\rho^2)}\right\}$

(a) 제안분포: $q(\theta^{(t-1)},\theta')=q\left(\begin{pmatrix}\theta_{1,t-1}\\\theta_{2,t-1}\end{pmatrix},\begin{pmatrix}\theta_1'\\\theta_2'\end{pmatrix}\right)=q(\theta_{1,t-1},\theta_1')q(\theta_{2,t-1},\theta_2')$
$=g(\theta_1'-\theta_{1,t-1})h(\theta_2')$의 꼴로 나타나는 적절한 제안분포를 선택하면 된다. $g(\theta_1'-\theta_{1,t-1})=N(\theta_1'|\theta_{1,t-1},d^2)$, $h(\theta_2')=Cauchy(\theta_2'|0,1)$을 선택해 보자. d의 값은 시뮬레이션으로 결정한다.

합격 확률 중 q와 π의 몫으로 나타나는 부분의 값을 정리하면

$$\alpha_0(\theta^{(t-1)},\theta')=\frac{\pi(\theta')q(\theta',\theta^{(t-1)})}{\pi(\theta^{(t-1)})q(\theta^{(t-1)},\theta')}=\frac{\pi(\theta')h(\theta_{2,t-1})N(\theta_{1,t-1}|\theta_1',d^2)}{\pi(\theta^{(t-1)})h(\theta_2')N(\theta_1'|\theta_{1,t-1},d^2)}$$

(대칭)

$$=\frac{\cancel{dnorm(\theta_{1,t-1},\theta_1',d)}\times dcauchy(\theta_{2,t-1},0,1)\times dmvnorm\left(\theta',\begin{pmatrix}0\\0\end{pmatrix},\begin{pmatrix}1&\rho\\\rho&1\end{pmatrix}\right)}{\cancel{dnorm(\theta_1',\theta_{1,t-1},d)}\times dcauchy(\theta_2',0,1)\times dmvnorm\left(\theta^{(t-1)},\begin{pmatrix}0\\0\end{pmatrix},\begin{pmatrix}1&\rho\\\rho&1\end{pmatrix}\right)}$$

이제 알고리즘을 서술하면 다음과 같다.

(1) initialize $\theta^{(0)}=(\theta_1^{(0)},\theta_2^{(0)})=(0,0)$

(2) 메트로폴리스-헤이스팅스 반복: for $i=1,2,\cdots,m=5000$,

① $\theta'\sim q(\theta^{(t-1)},\theta')$, $u\sim Unif(0,1)$, $\theta'\perp u$ 추출.

② 합격 확률 계산: ①에서의 θ'와 직전 단계의 $\theta^{(t-1)}$로부터 합격 확률을 다음과 같이 계산

$\alpha(\theta^{(t-1)},\theta')=\min\{1,\alpha_0(\theta^{(t-1)},\theta')\}$

③ 합격-불합격 결정

$$\theta^{(t)}=\begin{cases}\theta' & \text{if } u\le\alpha(\theta^{(t-1)},\theta')\\ \theta^{(t-1)} & \text{if } u>\alpha(\theta^{(t-1)},\theta')\end{cases}$$

(3) MC $\{\theta^{(t)}: t=1,2,\cdots,m\}$으로 $\pi(\theta)$를 근사.

(b) ~ (d): R markdown 참고.

3

(a). 우선 편의를 위해 $\sigma^2 = \Delta$로 두고 풀이한다. 각 사전분포와 가능도의 밀도함수를 확인하라.

$\Delta \sim \text{Inv-Gamma}\left(\frac{\nu_0}{2}, \frac{\nu_0 s_0^2}{2}\right) \Rightarrow \pi(\Delta) \propto \Delta^{-\frac{\nu_0}{2}-1} e^{-\frac{\nu_0 s_0^2}{2\Delta}} \quad (\Delta > 0)$

$\theta | \Delta \sim t_\nu(\theta_0, \sqrt{\Delta}) \Rightarrow \pi(\theta|\Delta) \propto \Delta^{-\frac{1}{2}} \left\{1 + \frac{(\theta-\theta_0)^2}{\nu\Delta}\right\}^{-\frac{\nu+1}{2}} \quad (\theta \in \mathbb{R})$

$f(X_n|\theta, \Delta) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\Delta}} e^{-\frac{(x_i-\theta)^2}{2\Delta}} \propto \Delta^{-\frac{n}{2}} e^{-\frac{n}{2\Delta}\{ns_n^2 + (\bar{x}-\theta)^2\}}$

이때 감마 적분을 이용하면 $\text{Gamma}\left(\frac{\nu+1}{2}, 1 + \frac{(\theta-\theta_0)^2}{\nu\Delta}\right)$ 으로부터 $\left(1 + \frac{(\theta-\theta_0)^2}{\nu\Delta}\right)^{-\frac{\nu+1}{2}} = \frac{1}{\Gamma\left(\frac{\nu+1}{2}\right)} \int_0^\infty \xi^{\frac{\nu+1}{2}-1} e^{-\left(1+\frac{(\theta-\theta_0)^2}{\nu\Delta}\right)\xi} d\xi$

이다. 이를 통해 사후분포를 정리하고 ~~[illegible]~~ ξ을 추가하라.

$\pi(\theta, \Delta | X_{(n)}) \propto \Delta^{-\frac{\nu_0+n+1}{2}-1} e^{-\frac{1}{2\Delta}\{\nu_0 s_0^2 + ns_n^2 + (\bar{x}-\theta)^2\}} \left\{1 + \frac{(\theta-\theta_0)^2}{\nu\Delta}\right\}^{-\frac{\nu+1}{2}}$

$\Rightarrow \pi(\theta, \Delta, \xi | X_{(n)}) \propto \Delta^{-\frac{\nu_0+n+1}{2}-1} \xi^{\frac{\nu+1}{2}-1} e^{-\xi} e^{-\frac{1}{2\Delta}\{\nu_0 s_0^2 + ns_n^2 + (\bar{x}-\theta)^2 + \frac{2\xi}{\nu}(\theta-\theta_0)^2\}}$

이때 $(\theta-\bar{x})^2 + \frac{2\xi}{\nu}(\theta-\theta_0)^2 = \left(1+\frac{2\xi}{\nu}\right)\left(\theta - \frac{\bar{x}+\frac{2\xi\theta_0}{\nu}}{1+\frac{2\xi}{\nu}}\right)^2 + \frac{\frac{2\xi}{\nu}}{1+\frac{2\xi}{\nu}}(\theta_0-\bar{x})^2$

이고, $\theta_n := \frac{\nu\bar{x}+2\xi\theta_0}{\nu+2\xi}$, $\xi_n := \frac{\nu+2\xi}{\nu}$, $\nu_n := \nu_0 + n$,

$\nu_n \sigma_n^2 := \nu_0 s_0^2 + ns_n^2 + (\bar{x}-\theta_0)^2 \times \frac{2\xi}{\nu \cdot \xi_n}$ 으로 정의하면

$\pi(\theta, \Delta, \xi | X_{(n)}) \propto \Delta^{-\frac{\nu_n+1}{2}-1} \xi^{\frac{\nu+1}{2}-1} e^{-\xi} e^{-\frac{1}{2\Delta}\{\xi_n(\theta-\theta_n)^2 + \nu_n\sigma_n^2\}}$ 이다.

이제 θ, Δ, ξ 각각의 사후조건부분포를 구하면

$\pi(\theta | \Delta, \xi, X_{(n)}) \propto e^{-\frac{\xi_n}{2\Delta}(\theta-\theta_n)^2} \propto N\left(\theta_n, \frac{\Delta}{\xi_n}\right)$

$\pi(\Delta | \theta, \xi, X_{(n)}) \propto \Delta^{-\frac{\nu_n+1}{2}} e^{-\frac{1}{2\Delta}\{\xi_n(\theta-\theta_n)^2 + \nu_n\sigma_n^2\}}$

$\propto \text{Inv-Gamma}\left(\frac{\nu_n+1}{2}, \frac{\xi_n(\theta-\theta_n)^2 + \nu_n\sigma_n^2}{2}\right)$

$\pi(\xi | \Delta, \theta, X_{(n)}) \propto \xi^{\frac{\nu+1}{2}-1} e^{-\left(1+\frac{(\theta-\theta_0)^2}{\nu\Delta}\right)\xi}$

$\propto \text{Gamma}\left(\frac{\nu+1}{2}, 1 + \frac{(\theta-\theta_0)^2}{\nu\Delta}\right)$

3

(b) $\pi(\theta, \Delta, \xi | X_{(n)})$의 pdf와 θ, Δ, ξ 각각의 완전 조건부분포를 알고 있으므로 이로부터 깁스 샘플링 적용.

(1) 깁스 샘플링할 표본의 수 $m = 5000$과 초기값 $(\theta^{(0)}, \Delta^{(0)}, \xi^{(0)}) = (0, 1, 0)$ 결정

(2) $t = 1, 2, \cdots, m$에서 조건부 추출.

$\xi^{(t)} \sim \text{Gamma}\left(\frac{\nu+1}{2}, 1 + \frac{(\theta^{(t-1)}-\theta_0)^2}{\nu\Delta^{(t-1)}}\right)$

$\theta^{(t)} \sim N\left(\frac{\nu\bar{x}+2\theta_0\xi^{(t)}}{\nu+2\xi^{(t)}}, \frac{\nu\Delta^{(t-1)}}{\nu+2\xi^{(t)}}\right)$

$\Delta^{(t)} \sim \text{Inv-Gamma}\left(\frac{\nu_n+1}{2}, \frac{\frac{\nu+2\xi^{(t)}}{\nu}(\theta^{(t)}-\theta_n)^2 + \nu_0 s_0^2 + ns_n^2 + \frac{2\xi^{(t)}(\bar{x}-\theta_0)^2}{\nu+2\xi^{(t)}}}{2}\right)$

(3) $\{(\theta^{(t)}, \Delta^{(t)}), t = 1, 2, \cdots, m\}$으로 $\pi(\theta, \Delta | X_n)$을 근사한다. 이때 조건부분포를 계산하기 위한 ξ는 버려도 된다.