깁스샘플러

정의

- 몬테카를로 기법

1. Y ~ $f_y(y)$ 이고

   X ~ $f_{x|y}(x|y)$ 라고 하자.

2. T를 알고리즘에 의해 생성된 확률변수라고 가정하면

1) $P[T \le t]$ = $\mathrm{E}_Y[f_{x|y}(t)]$를 정의하면

(1) $\int_{-\infty}^{\infty}\left[\int_{-\infty}^{t} f_{x|y}(x|y)\,dx\right]\mathrm{F}_Y(y)\,dy$

$= \int_{-\infty}^{\infty}\left[\int_{-\infty}^{t} f_{x|y}(x|y)\,\mathrm{F}_Y(y)\,dy\right]dx$

$= \int_{-\infty}^{\infty}\left[\int_{-\infty}^{t} f_{x,y}(x,y)\,dy\right]dx$

$= \int_{-\infty}^{t} f_x(x)\,dx$

2) 즉, X ~ $f_{x|y}(x|y)$와 Y ~ $f_y(y)$ 에서 생성된 확률변수 T는 X의 분포를 갖는다.

(1) 이는, $f_x(x)$의 직접 생성은 여러 이유로 어렵지만

(2) Y ~ $f_y(y)$ 에 의존하는 X ~ $f_{x|y}(x|y)$의 생성은 쉽다면

(3) 이는 $f_x(x)$에 의해 생성하는 것과 동등함을 보여준다.

정의

- 몬테카를로 기법

3. 이 때, X에 대한 어떤 함수 $w(X)$의 기댓값 $\mathrm{E}[w(X)]$를 결정짓길 원한다고 하자.

1) $(X_1, Y_1) \dots, (X_n Y_n)$을 Y -> X의 순서로 재귀적으로 생성시킨다.

2) 이 때,

(1) $\overline{w} = \frac{1}{m}\sum_{i=1}^{m} w(X_i) \xrightarrow{p} \int_{-\infty}^{\infty} w(X) f(x)\, dx = \mathrm{E}[w(X)]$ 가 된다.

3) 한편, 중심극한정리에 의해

1) $\sqrt{m}(\overline{w} - \mathrm{E}[w(X)]) \xrightarrow{D} N(0, \sigma^2)$ 으로 분포 수렴하며

2) 이를 이용하여 근사 신뢰구간을 다음과 같이 정의할 수 있다.

(1) $\mathrm{P}\left[\overline{w} - \mathrm{z}\frac{\mathrm{S}}{\sqrt{m}} \le \mathrm{E}[w(X)] \le \overline{w} + \mathrm{z}\frac{\mathrm{S}}{\sqrt{m}}\right] = 1 - a$

정의

- 깁스샘플러

1. m을 양의 정수라 하고, X_0를 주어진 최초의 값이라 하자.
2. 이 때

1) $Y_i|X_{i-1} \sim f(y|x_{i-1})$

2) $X_i|Y_i \sim f(x_i|y_i)$

3) 재귀적으로 각각의 관측값을 생성한다.

3. 이 때, 통계량 w에 대하여

1) $\overline{w} = \frac{1}{m}\sum_{i=1}^{m} w(X_i) \xrightarrow{p} \int_{-\infty}^{\infty} w(X_i) f(x)\, dx = \mathrm{E}[w(X_i)]$ 이고

2) $\mathrm{P}\left[\overline{w} - \mathrm{z}\frac{\mathrm{S}}{\sqrt{m}} \leq \mathrm{E}[w(X_i)] \leq \overline{w} + \mathrm{z}\frac{\mathrm{S}}{\sqrt{m}}\right] = 1 - a$ 로 신뢰구간을 정의할 수 있다.

정의

- 계층적 베이지안

1. 하이퍼파라미터를 고려한 모수의 사전분포를 이용하는 방법
2. 임의의 초모수분포 → 모수사전분포 → 사후분포 순으로 계층적으로 영향을 미친다.
3. 즉, 다음과 같은 확률변수가 정의될 때

1) $X|\theta \sim f(\mathrm{x}|\theta)$ 표본모형

2) $\theta|\gamma \sim g(\theta|\gamma)$ 모수사전분포

3) $\gamma \sim \mathrm{h}(\gamma)$ 초모수분포

4. 이 때 사전 결합 pdf는

1) $\mathrm{r}(\theta,\gamma|X) = \frac{\mathrm{r}(X,\theta,\gamma)}{r(X)} = \frac{\mathrm{r}(X|\theta,\gamma)\mathrm{r}(\theta,\gamma)}{r(X)} = \frac{f(X|\theta)g(\theta|\gamma)\mathrm{h}(\gamma)}{r(X)}$

2) 따라서 사후 pdf $\mathrm{k}(X|\theta) = \frac{\int_{-\infty}^{\infty} f(X|\theta)g(\theta|\gamma)\mathrm{h}(\gamma)\, d\gamma}{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(X|\theta)g(\theta|\gamma)\mathrm{h}(\gamma)\, d\gamma\, d\theta}$

정의

- 계층적 베이지안

4. 이 때 사전 결합 pdf는

2) 따라서 사후 pdf $\mathrm{k}(X|\theta) = \frac{\int_{-\infty}^{\infty} f(X|\theta)g(\theta|\gamma)\mathrm{h}(\gamma)\,d\gamma}{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(X|\theta)g(\theta|\gamma)\mathrm{h}(\gamma)\,d\gamma\,d\theta}$

(1) 즉, 이는 기존의 $\mathrm{L}(\theta|X)\mathrm{g}(\theta)$ 를 $\int_{-\infty}^{\infty} f(X|\theta)g(\theta|\gamma)\mathrm{h}(\gamma)\,d\gamma$로 확장하고

(2) 따라서 $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathrm{L}(\theta|X)\mathrm{g}(\theta)\,d\theta$ 를 $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(X|\theta)g(\theta|\gamma)\mathrm{h}(\gamma)\,d\gamma\,d\theta$로 확장한 것이다.

5. 이 때, 이 사후분포의 베이지안 추정해는

1)$\overline{w} = \frac{1}{m}\sum_{i=1}^{m} w(\theta_i) \xrightarrow{p} \frac{\int_{-\infty}^{\infty} w(\theta)\left[\int_{-\infty}^{\infty} f(X|\theta)g(\theta|\gamma)\mathrm{h}(\gamma)\,d\gamma\right]d\theta}{\int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty} f(X|\theta)g(\theta|\gamma)\mathrm{h}(\gamma)\,d\gamma\right]d\theta} = \mathrm{E}_{사후}[w(\Theta)]$

정의

- 경험적 베이지안

1. 계층적 베이지안에서 초모수 분포를 직접 데이터에서 경험적으로 추정하는 것
2. 베이지안 모형이 다음과 같이 정의된다.

1) $X|\theta \sim f(\mathrm{x}|\theta)$ 표본모형

2) $\theta|\gamma \sim g(\theta|\gamma)$ 모수사전분포

3. $\mathrm{r}(X,\theta|\gamma) = \frac{\mathrm{r}(X,\theta,\gamma)}{\mathrm{h}(\gamma)} = \frac{\mathrm{r}(X|\theta,\gamma)\mathrm{r}(\theta,\gamma)}{\mathrm{h}(\gamma)} = \frac{f(X|\theta)g(\theta|\gamma)\mathrm{h}(\gamma)}{\mathrm{h}(\gamma)} = f(X|\theta)g(\theta|\gamma)$ 이고

1) 이 때, $\int_{-\infty}^{\infty} f(X|\theta)g(\theta|\gamma)\,d\theta = \mathrm{m}(\mathrm{x}|\gamma)$ 는 θ와 무관한 pdf이므로

2) $\frac{\partial \log[\mathrm{m}(\mathrm{X}|\gamma)]}{\partial\gamma} = 0$을 풀면, γ에 대한 MLE 추정량을 구할 수 있다.

4. $\mathrm{k}(X,\theta|\hat{\gamma}) = f(X|\theta)g(\theta|\hat{\gamma})$ 경험적 베이지안 모델의 사후 PDF가 되고

1\) $\overline{w} = \frac{1}{m}\sum_{i=1}^{m} w(\theta_i) \xrightarrow{p} \frac{\int_{-\infty}^{\infty} w(\theta)\left[\int_{-\infty}^{\infty} f(X|\theta)g(\theta|\hat{\gamma})\,d\hat{\gamma}\right]d\theta}{\int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty} f(X|\theta)g(\theta|\hat{\gamma})\,d\hat{\gamma}\right]d\theta} = \mathrm{E}_{사후}[w(\Theta)]$

예제

- 몬테카를로 기법

1. $\mathrm{Y} = \bar{x}$가 충분통계량일 때

1) $\mathrm{Y}|\theta \sim N(\theta, \frac{\sigma^2}{n})$과

2) $\Theta \sim h(\theta) \propto \frac{\mathrm{b}^{-1}\ \mathrm{exp}\left\{-\frac{\theta - a}{b}\right\}}{\left[1 + \mathrm{exp}\left\{-\frac{\theta - a}{b}\right\}\right]^2}$ (로지스틱 분포) 인 베이지안 모델을 정의하자.

2. 이 때, 사후 pdf는 $k(\theta|y) = \frac{\frac{1}{\sqrt{2\pi}\sigma/\sqrt{n}}\mathrm{exp}\left(-\frac{(x_i - \theta)^2}{2\sigma^2/\sqrt{n}}\right)\frac{\mathrm{b}^{-1}\ \mathrm{exp}\left\{-\frac{\theta - a}{b}\right\}}{\left[1 + \mathrm{exp}\left\{-\frac{\theta - a}{b}\right\}\right]^2}}{\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}\sigma/\sqrt{n}}\mathrm{exp}\left(-\frac{(x_i - \theta)^2}{2\sigma^2/\sqrt{n}}\right)\frac{\mathrm{b}^{-1}\ \mathrm{exp}\left\{-\frac{\theta - a}{b}\right\}}{\left[1 + \mathrm{exp}\left\{-\frac{\theta - a}{b}\right\}\right]^2}\, d\theta}$ 이고

1) 이 사후 pdf의 베이지안해는 손실함수가 MSE로 주어질 때, 사후 분포의 평균이다.

예제

- 몬테카를로 기법

2. 이 때, 사후 pdf는 $k(\theta|y) = \dfrac{\frac{1}{\sqrt{2\pi}\sigma/\sqrt{n}}\exp\left(-\frac{(x_i-\theta)^2}{2\sigma^2/\sqrt{n}}\right)\frac{\mathrm{b}^{-1}\exp\left\{-\frac{\theta-a}{b}\right\}}{\left[1+\exp\left\{-\frac{\theta-a}{b}\right\}\right]^2}}{\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}\sigma/\sqrt{n}}\exp\left(-\frac{(x_i-\theta)^2}{2\sigma^2/\sqrt{n}}\right)\frac{\mathrm{b}^{-1}\exp\left\{-\frac{\theta-a}{b}\right\}}{\left[1+\exp\left\{-\frac{\theta-a}{b}\right\}\right]^2}d\theta}$ 이고

1) 이 사후 pdf의 베이지안해는 손실함수가 MSE로 주어질 때, 사후 분포의 평균이다.

2) $w(\theta)$ = f(y|θ)= $\frac{1}{\sqrt{2\pi}\sigma/\sqrt{n}}\exp\left(-\frac{(x_i-\theta)^2}{2\sigma^2/\sqrt{n}}\right)$의 통계량으로 다시 정의할 때

(1) $\delta(y)$ = $\mathrm{E}_{사후}$ $[\Theta]$ = $\dfrac{\int_{-\infty}^{\infty}\theta w(\theta)\frac{\mathrm{b}^{-1}\exp\left\{-\frac{\theta-a}{b}\right\}}{\left[1+\exp\left\{-\frac{\theta-a}{b}\right\}\right]^2}d\theta}{\int_{-\infty}^{\infty}w(\theta)\frac{\mathrm{b}^{-1}\exp\left\{-\frac{\theta-a}{b}\right\}}{\left[1+\exp\left\{-\frac{\theta-a}{b}\right\}\right]^2}d\theta} = \dfrac{\mathrm{E}_{사전}[\Theta w(\Theta)]}{\mathrm{E}_{사전}[w(\Theta)]}$ 이다.

(2) 이 때, 무작위로 선출한 확률표본의 실현값을 $\Theta_1 \cdots \Theta_n$이라 하고, 이 때

- $\dfrac{m^{-1}\sum[\Theta_i w(\Theta_i)]}{m^{-1}\sum[w(\Theta_i)]} \xrightarrow{p} \dfrac{\mathrm{E}_{사전}[\Theta w(\Theta)]}{\mathrm{E}_{사전}[w(\Theta)]} = \delta(y)$는 대수의 약법칙에 따라 보장된다.

예제

- 계층적 베이지안

1. 다음과 같은 계층적 베이지안 모델이 주어졌다.

1) $X|\Theta \sim N(\theta, \frac{\sigma^2}{n})$

2) $\Theta|\Gamma^2 \sim N(0, \frac{\sigma^2}{n})$

3) $\frac{1}{\Gamma^2} \sim \Gamma(\mathrm{a}, \mathrm{b})$

2. 이 때, 이 모형의 베이지안 추정해를 구하라

1) 우선, θ 의 사후 pdf를 정의하면

(1) $\mathrm{k}(\theta|\mathrm{X}, \Gamma^2) \propto \mathrm{f}(\mathrm{X}|\theta)\mathrm{g}(\theta|\gamma^2)\mathrm{h}(\gamma^{-2})$

2) γ^2 를 일단 무시하면 $f(X|\theta)g(\theta|\gamma^2)$는 $\mathrm{N}\left[\frac{\gamma^2}{\gamma^2+{\sigma^2}/{\mathrm{n}}}, \frac{\sigma^2\gamma^2}{\sigma^2+n\gamma^2}\right]$ 을 따른다.

예제

- 계층적 베이지안

2. 이 때, 이 모형의 베이지안 추정해를 구하라.

3) Γ^2의 사후 pdf를 정의하면

(1) $\mathrm{k}(\frac{1}{\Gamma^2}|X,\theta) \propto f(X|\theta)g(\theta|\gamma^2)\mathrm{h}(\gamma^{-2})$

(2) 이 때 $f(X|\theta)g(\theta|\gamma^2)\mathrm{h}(\gamma^{-2}) = \frac{1}{\gamma}\exp\left(-\frac{\theta^2}{2\gamma^2}\right)\left(\frac{1}{\gamma^2}\right)^{(a-1)}\exp\left(-\frac{1}{\gamma^2}\cdot\frac{1}{b}\right)$

$= \left(\frac{1}{\gamma^2}\right)^{(a+(1/2)-1)}\exp\left(-\frac{1}{\gamma^2}\left[\frac{\theta^2}{2}+\frac{1}{b}\right]\right)$

(3) 이는 $\Gamma[a+\left(\frac{1}{2}\right),\left[\frac{\theta^2}{2}+\frac{1}{b}\right]^{-1}]$ 을 따른다.

3. 이를 이용해 베이지안 추정해를 구하면

1) $\hat{\theta} = \frac{1}{n-m}\sum_{i=1}^{m} w(\theta_i)$ 이다.

예제

- 경험적 베이지안

1. 다음과 같은 베이지안 모형이 주어졌다.

1) $X_i|\Theta \sim poisson(\theta)$

2) $\Theta|b \sim \Gamma(1, \mathrm{b})$

3) $\mathrm{X} = [X_1, \dots, X_n]$

2. 이 때

1) $\mathrm{f(X|\theta)} = \prod_{i=1}^{n} \frac{\theta^x e^{-\theta}}{x_i!} = \frac{\theta^{\sum x_i} e^{-n\theta}}{\prod x_i!}$

2) $g(\theta|\mathrm{b}) = \frac{1}{b}\mathrm{exp}(-\frac{\theta}{b})$

3) b 조건부 사후분포의 pdf를 구하면

(1) $\int_{-\infty}^{\infty} f(X|\theta) g(\theta|b)\, d\theta = \int_{-\infty}^{\infty} \frac{\theta^{\sum x_i} e^{-n\theta}}{\prod x_i!} \frac{1}{b}\mathrm{exp}(-\frac{\theta}{b})\ \ d\theta$

$= \frac{1}{b\prod x_i!} \int_{-\infty}^{\infty} \theta^{n\bar{x}} \mathrm{exp}[-\theta\left(1 - \frac{1}{b}\right)]\ \ d\theta\ = \frac{\Gamma(n\bar{x}+1)[\frac{\mathrm{b}}{nb+1}]^{(n\bar{x}+1)}}{b\prod x_i!} = \mathrm{m(x|}\gamma)$

예제

- 경험적 베이지안

3. b 조건부 사후분포를 b에 대해 편미분하면

1) $\frac{\partial \log[\mathrm{m}(\mathrm{X}|\gamma)]}{\partial b} = -\frac{1}{b} + \frac{(n\overline{x}+1)}{b(bn+1)} = 0$

2) 즉 $\frac{(n\overline{x}+1)}{b(bn+1)} = \frac{1}{b}$ 에서 b = $\overline{x}$

4. $\mathrm{k}(\mathrm{X}|\theta, \mathrm{b} = \overline{x}) \propto f(X|\theta)g(\theta|b) = \frac{\theta^{\sum x_i} e^{-n\theta}}{\prod x_i!}\frac{1}{\overline{x}}\exp\left(-\frac{\theta}{\overline{x}}\right)$

1) $\mathrm{k}(\mathrm{X}|\theta, \mathrm{b} = \overline{x}) \propto \theta^{n\overline{x}+1-1} e^{-\theta(n+\frac{1}{\overline{x}})}$

2) 이는 $\Gamma(n\overline{x} + 1, \frac{\overline{x}}{1+n\overline{x}})$ 의 분포를 암시한다.

4. 이 때, 베이지안 추정해는 MSE에서 사후분포의 평균이므로

1) 감마 분포의 평균은 $\alpha\beta = (n\overline{x} + 1)(\frac{\overline{x}}{1+n\overline{x}}) = \overline{x}$ 이다.