Subject

Date δ

Attendants

6(c) Describe difference between

MAP Estimation (maximum posterior prediction)

and Baysion Prediction

MAP Estimation can be describe as a special case

of Baysian Prediction

$$p(\theta|x) \approx q(\theta|\hat{\theta}) = \delta(\theta - \hat{\theta})$$

where $\delta(\theta) = \begin{cases} +\infty & \text{if } \theta = 0, \\ 0 & \text{otherwise} \end{cases}$

$$\int_{-\infty}^{\infty} \delta(\theta)\, d\theta = 1$$

$$\int_{-\infty}^{+\infty} \delta(\theta) f(\theta)\, d\theta = \langle f(\theta) \rangle_{\delta(\theta)} = f(0)$$

$$KL\left(q(\theta|\hat{\theta}) \,\|\, p(\theta|x) \right) = \mathbb{E}_{q(\theta|\hat{\theta})}\left[\log q(\theta|\hat{\theta}) - \log p(\theta|x) \right]$$

$$\int q \log \frac{q}{p}\, dx$$

Subject $KL(q \| p) = \mathbb{E}_q[\log q - \log p]$

Date 9

Attendants $= \mathbb{E}_q[\log q(\theta|\hat{\theta})] - \mathbb{E}_q[\log p(x|\theta)]$

$$KL(q \| p) = \log q(\hat{\theta}|\hat{\theta}) - \log p(x|\hat{\theta}) - \log p(\hat{\theta})$$

$$\underbrace{\mathbb{E}_{q(\theta|\hat{\theta})}}[\log q(\theta|\hat{\theta})] = \log q(\hat{\theta}|\hat{\theta})$$

$\theta = \hat{\theta}$のときのみ無限大に

尖った形状のような確率分布

$$\mathbb{E}_{q(\theta|\hat{\theta})}[\underbrace{\log p(x|\theta)}] = \log p(x|\hat{\theta}) + \log p(\hat{\theta})$$

$\theta = \hat{\theta}$のときしか

評価されない

$$\int_{-\infty}^{\infty} \underbrace{\log p(x|\theta)}_{f(\theta)} \delta(\theta - \hat{\theta})\, d\theta = f(0)$$

$$\downarrow$$

$$\log p(x|0)$$

Subject　　　　　　　　　　　　　　　　Date　10

Attendants

事後分布を近似する

$q(\theta) = \delta(\theta - \hat{\theta})$ と仮定

$$KL(q(\theta) \| p(\theta|x)) = KL(\delta(\theta-\hat{\theta}) \| p(\theta|x))$$

$$= \int \delta(\theta-\hat{\theta}) \log \frac{\delta(\theta-\hat{\theta})}{p(\theta|x)} d\theta$$

$$= \underbrace{\int \delta(\theta-\hat{\theta}) \log \delta(\theta-\hat{\theta}) d\theta} - \int \delta(\theta-\hat{\theta}) p(\theta|x) d\theta$$

クロネッカーのデルタ関数の性質を利用している

$$= \underbrace{\log \delta(\hat{\theta}-\hat{\theta})} - \log p(\hat{\theta}|x)$$

$$= \log \delta(\hat{\theta}-\hat{\theta}) - \log p(\hat{\theta}|x)$$

$$\int \delta(\theta-\hat{\theta}) f(\theta) d\theta = f(\hat{\theta})$$

$$= \log \delta(\hat{\theta}-\hat{\theta}) - \int \log \frac{p(x|\hat{\theta})p(\hat{\theta})}{p(x)}$$

$$= -\log p(x|\hat{\theta})p(\hat{\theta}) + \log \delta(\hat{\theta}-\hat{\theta}) + \log p(x)$$

Subject

Date 11/

Attendants

$KL(q(\theta) \| p(\theta|x))$

$= -\log p(x|\theta)p(\theta) + \underbrace{\log \delta(\hat{\theta}-\hat{\theta})}_{=\delta(0)=\infty} + \underbrace{\log p(x)}$

けど潜在変数

xにも

パラメータθにも

依存しないので

除外

θに

依存しない

KLを最小にする $\hat{\theta}$ は

$$\hat{\theta} = \underset{\hat{\theta}}{\arg\min} \{ -\log p(x|\theta)p(\theta) \}$$

$$= \underset{\hat{\theta}}{\arg\max} \{ \log p(x|\theta)p(\theta) \}$$

$$= \hat{\theta}_{MAP}$$

Subject

Date 12

Attendants

すなわち MAP推定は

ベイズ推定に対し

事後分布に $g(\theta|\hat{\theta}) = f(\theta - \hat{\theta})$

delta distribution function

where $f(\theta) = +\infty$ if $\theta = 0$

$f(\theta) = 0$ otherwise

仮定している