1

演習課題 8

- 対数尤度関数 $\ln L(\boldsymbol{\theta})$ をバイアス b_i で偏微分したときの $\frac{\partial \ln L(\boldsymbol{\theta})}{\partial b_i}$ を求めよ.

10/19/2023

2

演習課題 8 解答

- $\frac{\partial \ln L(\boldsymbol{\theta})}{\partial b_i} = \frac{\partial \sum_{n=1}^{N}(-\Phi(\boldsymbol{x}_n,\boldsymbol{\theta}) - \ln z(\boldsymbol{\theta}))}{\partial b_i}$
$= \sum_{n=1}^{N}\left(\boxed{-\frac{\partial \Phi(\boldsymbol{x}_n,\boldsymbol{\theta})}{\partial b_i}}\boxed{-\frac{\partial \ln z(\boldsymbol{\theta})}{\partial b_i}}\right)$

- $\boxed{-\frac{\partial \Phi(\boldsymbol{x}_n,\boldsymbol{\theta})}{\partial b_i}} = \frac{\partial\left(\sum_{(i,j)} w_{j,i} x_i^n x_j^n + \sum_i b_i x_i^n\right)}{\partial b_i} = x_i^n$

10/19/2023

演習課題８解答

- $\boxed{-\frac{\partial \Phi(x_n,\boldsymbol{\theta})}{\partial b_i}} = \frac{\partial\left(\sum_{(i,j)} w_{j,i} x_i^n x_j^n + \sum_i b_i x_i^n\right)}{\partial b_i} = x_i^n$

- $\boxed{-\frac{\partial \ln z(\boldsymbol{\theta})}{\partial b_i}} = -\frac{1}{z(\boldsymbol{\theta})}\frac{\partial z(\boldsymbol{\theta})}{\partial b_i} = -\frac{1}{z(\boldsymbol{\theta})}\frac{\partial \sum_x e^{-\Phi(x,\boldsymbol{\theta})}}{\partial b_i}$

$$= -\frac{1}{z(\boldsymbol{\theta})}\sum_x \frac{\partial e^{-\Phi(x,\boldsymbol{\theta})}}{\partial b_i} = -\frac{1}{z(\boldsymbol{\theta})}\sum_x e^{-\Phi(x,\boldsymbol{\theta})}\frac{\partial(-\Phi(\boldsymbol{x},\boldsymbol{\theta}))}{\partial b_i}$$

$$= -\sum_x \frac{1}{z(\boldsymbol{\theta})} e^{-\Phi(\boldsymbol{x},\boldsymbol{\theta})} x_i = -\boxed{\sum_x P(\boldsymbol{x}|\boldsymbol{\theta})x_i}$$

全ての $\boldsymbol{x}$ を考えたときの x_i の期待値
（$E_{\boldsymbol{\theta}}[x_i]$と表す）

演習課題８解答

- $\frac{\partial \ln L(\boldsymbol{\theta})}{\partial b_i} = \frac{\partial \sum_{n=1}^{N}(-\Phi(x_n,\boldsymbol{\theta}) - \ln z(\boldsymbol{\theta}))}{\partial b_i}$

$$= \sum_{n=1}^{N}\left(\boxed{-\frac{\partial \Phi(x_n,\boldsymbol{\theta})}{\partial b_i}}\boxed{-\frac{\partial \ln z(\boldsymbol{\theta})}{\partial b_i}}\right)$$

$$= \sum_{n=1}^{N}(x_i^n - E_{\boldsymbol{\theta}}[x_i])$$

$$= \sum_{n=1}^{N}(x_i^n) - NE_{\boldsymbol{\theta}}[x_i]$$