report

April 2, 2020

0.1 Question 1

1.

$$P(y=k \mid x, \sigma, \mu) = \frac{P(x \mid y=k, \sigma, \mu)\, P(y=k)}{P(x \mid \sigma, \mu)}$$

= by law of total probability

$$p(x \mid \sigma, \mu) = \sum_{k'} p(x \mid y=k', \sigma, \mu) \cdot p(y=k')$$

$$= \frac{\exp(\ln[p(x \mid y=k, \sigma, \mu)]^{①} + \ln[p(y=k)])}{\exp(\ln[p(x \mid y=k', \sigma, \mu)] + \ln[p(y=k')])}$$

① $\ln[p(x \mid y=k, \sigma, \mu)]$

$$= -\frac{1}{2}\left(\sum_{i=1}^{D} \ln(2\pi\sigma_i^2)\right) - \sum_{i=1}^{D} \frac{1}{2\sigma_i^2}(x_i - \mu_{ki})^2$$

$$= \frac{\exp\left(-\frac{1}{2}\left[\sum_{i=1}^{D} \ln(2\pi\sigma_i^2)\right] - \sum_{i=1}^{D} \frac{1}{2\sigma_i^2}(x_i - \mu_{ki})^2 + a_k\right)}{\sum_{k'} \exp\left(-\frac{1}{2}\left[\sum_{i=1}^{D} \ln(2\pi\sigma_i^2)\right] - \sum_{i=1}^{D} \frac{1}{2\sigma_i^2}(x_i - \mu_{k'i})^2 + a_{k'}\right)}$$

$$\propto \exp\left(-\frac{1}{2}\left[\sum_{i=1}^{D} \ln(2\pi\sigma_i^2)\right] - \sum_{i=1}^{D} \frac{1}{2\sigma_i^2}(x_i - \mu_{ki})^2 + a_k\right)$$

2.

$$l(\theta; D) = -\log\left[p(y^{(1)}|\theta) \cdot p(x^{(1)}|\theta) \cdot \ldots \cdot p(y^{(N)}|\theta) \cdot p(x^{(N)}|\theta)\right]$$

"Assume data are iid"

$(y^{(1)}, x^{(1)}) \perp (y^{(2)}, x^{(2)}) \perp \ldots \perp (y^{(N)}, x^{(N)})$

$$= -\log \left(p(y^{(1)}, x^{(1)}|\theta) \cdot p(y^{(2)}, x^{(2)}|\theta) \cdot \ldots \cdot p(y^{(N)}, x^{(N)}|\theta)\right)$$

$$= -\sum_{n=1}^{N} \left[\log(p(y^{(n)}|x^{(n)}, \theta)) + \log(p(x^{(n)}|\theta)) \right]$$

$$= -\sum^{N} \log(\underset{①}{p(x|y, \theta)} \cdot \underset{②}{p(y|\theta)})$$

② $\log p(y|\theta) = \log \sum_{k=1}^{d} \alpha_k \cdot \mathbb{I}[y^{(i)} = k]$

① $p(x|y, \theta) = \log \left(\prod_{i=1}^{D} 2\pi \sigma_i^2 \right)^{-1/2} \exp\left\{ -\sum_{i=1}^{D} \frac{1}{2\sigma_i} (x_i - \mu_{ki})^2 \right\}$

$$= -1/2 \left(\sum_{i=1}^{D} \log(2\pi \sigma^2) \right) + \left\{ -\sum_{i=1}^{D} \frac{1}{2\sigma_i} (x_i - \mu_{ki})^2 \right\}$$

$$l(D; \theta) = -\sum_{n=1}^{N} \left[\log\left(\sum_{k=1}^{d} \alpha_k \cdot \mathbb{I}[y^{(n)} = k]\right) - 1/2 \left(\sum_{i=1}^{D} \log(2\pi \sigma_i^2) \right) + \left(-\sum_{i=1}^{D} \frac{(x_i - \mu_{ki})^2}{2\sigma_i^2} \right) \right]$$

3.

$$\frac{\partial \mathcal{L}}{\partial \mu_{ki}} = -\sum_{n=1}^{N}\left[-\sum_{i=1}^{D}\frac{(x_i-\mu_{ki})}{\sigma_i^2}\right]$$

$$\frac{\partial \mathcal{L}}{\partial \sigma_i^2} = -\sum_{n=1}^{N}\left[-\tfrac{1}{2}\left(\sum_{i=1}^{D}\frac{1}{2\pi\sigma_i^2}\cdot 2\sigma_i\right) + \left(-\sum_{i=1}^{D}\frac{(x_i-\mu_{ki})^2}{2\sigma_i^3}\cdot -2\right)\right]$$

$$= -\sum_{n=1}^{N}\left[-\tfrac{1}{2}\left(\sum_{i=1}^{D}\frac{1}{2\pi\sigma_i}\right) + \sum_{i=1}^{D}\frac{(x_i-\mu_{ki})^2}{\sigma_i^3}\right]$$

4.

$$\mathcal{L}(x;\mu,\sigma) = \sum_{i=2}^{N} -\frac{1}{2}\log 2\pi - \log\sigma - \frac{1}{2\sigma^2}(x^{(i)}-\mu)^2$$

$$\frac{\partial \mathcal{L}}{\partial \mu} = -\frac{1}{\sigma^2}\sum_{i=2}^{N}(x^{(i)}-\mu) = 0$$

$$= \sum_{i=1}^{N}x^{(i)} - n\mu = 0$$

$$\mu = \frac{\sum_{i=1}^{N}x^{(i)}}{n}$$

$$\frac{\partial L}{\partial \sigma} = \frac{\partial}{\partial \sigma}\left[\sum_{i=1}^{N} -\frac{1}{2}\log 2\pi - \log\sigma - \frac{1}{2\sigma^2}(x^{(i)}-\mu)^2\right] = 0$$

$$= \sum_{i=1}^{N} \frac{1}{\sigma} - \frac{1}{\sigma^3}(x^{(i)}-\mu)^2 = 0$$

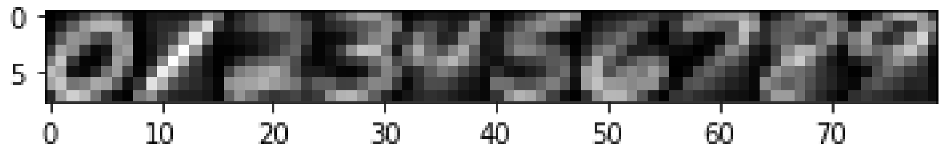
$$= \frac{N}{\sigma} - \frac{1}{\sigma^3}\sum_{i=1}^{N}(x^{(i)}-\mu)^2 = 0$$

$$= \sigma^2 N = \sum_{i=1}^{N}(x^{(i)}-\mu)^2$$

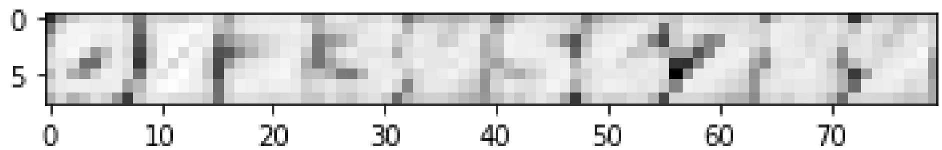
$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x^{(i)}-\mu)^2}$$

0.2 Question 2

0.2.1 2.0



0.2.2 2.1.1



0.2.3 2.1.2

Average Conditional Log-Likelihood

Training Set: -0.45150334150843136

Test Set -1.574084784152816

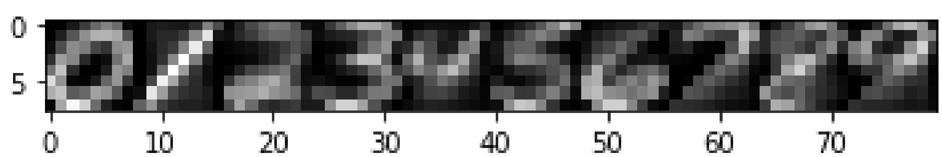
0.2.4 2.1.3

Accuracy of Model:

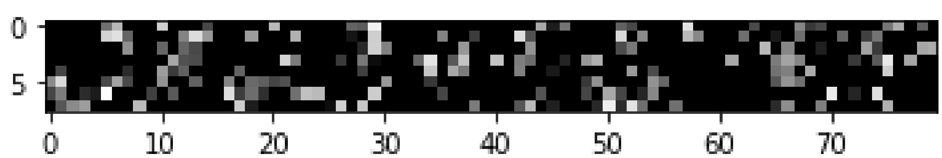
Training Set: 0.9812857142857143

Test Set 0.95925

0.2.5 2.2.3



0.2.6 2.2.4



0.2.7 2.2.5

Average Conditional Log-Likelihood

Training Set: -30.78815907276509

Test Set -30.73534308421527

0.2.8 2.2.6

Accuracy of Model:

Training Set: 0.775

Test Set 0.76525

[ ]:
```

```