CSC411 Assignment 2

Yue Guo

November 7, 2017

1 Gaussian

1.1 P(y)

$P(\boldsymbol{x}|\ mu, \sigma) = \sum_{k=1}^{K} P(\boldsymbol{x}|y=k,\ mu,\sigma)P(y=k|\mu,\sigma)$
$= \sum_{k=1}^{K} \alpha_k P(\boldsymbol{x}|y=k,\ mu,\sigma)$
$= \sum_{k=1}^{K} \alpha_k (\prod_{i=1}^{D} 2\pi\sigma_i)^{\frac{1}{2}}$

1.2 visualization

$P(y|\boldsymbol{x},\ mu,\sigma) = \frac{P(\boldsymbol{x},y=k|\mu,\sigma)}{P(\boldsymbol{x}|\mu,\sigma)}$
$= \frac{P(\boldsymbol{x}|y=k,\ mu,\sigma)P(y=k|\mu,\sigma)}{P(\boldsymbol{x}|\mu,\sigma)}$
$= \frac{\alpha_k (\prod_{i=1}^{D} 2\pi\sigma_i^2)^{-\frac{1}{2}} exp\{-\sum_{i=1}^{D}\frac{1}{2\sigma_i^2}(x_i-\mu_{ki})^2\}}{\sum_{k=1}^{K}\alpha_k (\prod_{i=1}^{D} 2\pi\sigma_i^2)^{-\frac{1}{2}} exp\{-\sum_{i=1}^{D}\frac{1}{2\sigma_i^2}(x_i-\mu_{ki})^2\}}$

1.3 visualization

$-logP(y^{(1)}, x^{(1)}, ..., y^{(N)}, x^{(N)}|\theta)$
$= -logP(y^{(1)}, x^{(1)}|\theta) - logP(y^{(2)}, x^{(2)}|\theta) ... - log(y^{(N)}, x^{(N)}|\theta)$
$= \sum_{k=1}^{K} -log[\alpha_k(\prod_{i=1} D2\pi\sigma_i^2)^{-\frac{1}{2}} exp\{-\sum_{i=1}^{D}\frac{1}{2\sigma_i^2}(x_i-\mu_{ki})^2\}]$
$= \sum_{k=1}^{K}\sum_{i=1}^{D}\frac{1}{2\sigma_i^2}(x_i-\mu_{ki})^2 + \sum_{k=1}^{K} -log[\alpha_k(\prod_{i=1}^{D} 2\pi\sigma_i^2)^{-\frac{1}{2}}$
$= \sum_{k=1}^{K}\sum_{i=1}^{D}\frac{1}{2\sigma_i^2}(x_i-\mu_{ki})^2 + \sum_{k=1}^{K} \frac{1}{2}log[\alpha_k(\prod_{i=1}^{D} 2\pi\sigma_i^2)$
$= \sum_{k=1}^{K}\sum_{i=1}^{D}\frac{1}{2\sigma_i^2}(x_i-\mu_{ki})^2 + \sum_{k=1}^{K}\sum_{i=1}^{D}[\frac{1}{2}\alpha_k \log(2\pi\sigma_i^2)]$

1.4 visualization

$\frac{\partial}{\partial\mu_{ki}} = \sum_{i=1}^{D}\frac{1}{2\sigma_i^2}2(-x_i+\mu_{ki})\mathbb{1}(i=k)$
$= \sum_{i=1}^{D}\frac{1}{\sigma_i^2}(-x_i+\mu_{ki})\mathbb{1}(i=k)$
Setting $\frac{\partial}{\partial\mu_{ki}} = 0$, we have

$$\sum_{i=1}^{D}\frac{1}{\sigma_i^2}(-x_i+\mu_{ki})\mathbb{1}(i=k) = 0$$

$x_i = \frac{\sum_{i=1}^{D}\mu_{ki}\mathbb{1}(i=k)}{\sum_{i=1}^{D}\mathbb{1}(i=k)}$

1.5 visualization

$\frac{\partial}{\partial\sigma_i^2} = \sum_{k=1}^{K}\sum_{i=1}^{D} -\frac{1}{2}(x_i-\mu_{ki})^2\frac{1}{\sigma_i^4} + \frac{1}{2}\sum_{k=1}^{K}\sum_{i=1}^{D}\frac{2\pi}{2\pi\sigma_i^2}$
$= \sum_{k=1}^{K}\sum_{i=1}^{D} -\frac{1}{2}(x_i-\mu_{ki})^2\frac{1}{\sigma_i^4} + \frac{1}{2}\sum_{k=1}^{K}\sum_{i=1}^{D}\frac{1}{\sigma_i^2}$

$= \sum_{i=1}^{D} \frac{1}{2}[-(x_i - \mu_{ki})^2 \frac{1}{\sigma_i^4} + \frac{1}{\sigma_i^2}]\mathbb{1}(i = k)$
Setting $\frac{\partial}{\partial \sigma_i^2}$ to zero, we have
$\sum_{i=1}^{D} \mathbb{1}(i = k)(x_i - \mu_{ki})^2 \frac{1}{\sigma_i^4} = \sum_{i=1}^{D} \mathbb{1}(i = k) \frac{1}{\sigma_i^2}$
$\sum_{i=1}^{D} \mathbb{1}(i = k)(x_i - \mu_{ki})^2 = \sum_{i=1}^{D} \mathbb{1}(i = k)\sigma_i^2$
$\sigma_i^2 = \frac{\sum_{i=1}^{D} \mathbb{1}(i=k)(x_i - \mu_{ki})^2}{\sum_{i=1}^{D} \mathbb{1}(i=k)}$

1.5.1 Feature weights