Pattern Classification Homework 5 Solutions

1. The log-likelihood function is

$$\mathrm{l}(\theta) = \sum_{k=1}^{n} \ln p(x_k|\theta) = \sum_{k=1}^{n} [\ln\theta - \theta x_k] = n \ln\theta \ - \theta \sum_{k=1}^{n} x_k$$

We solve $\nabla_\theta l(\theta) = 0$ to find $\hat{\theta}$

$$\nabla_\theta l(\theta) = \frac{d}{d\theta}[n \ln\theta \ - \theta \textstyle\sum_{k=1}^{n} x_k = \frac{n}{\theta} - \sum_{k=1}^{n} x_k = 0$$

Thus the maximum-likelihood solution is

$$\hat{\theta} = \frac{1}{\frac{1}{n}\sum_{k=1}^{n} x_k}$$

2. Consider MAP estimates, that is ones that maximize $\mathrm{l}(\theta)\mathrm{p}(\theta)$

In this problem, the parameter needed to be estimated is μ. Given the training data we have

$$\mathrm{l}(\mu)\mathrm{p}(\mu) = \ln[\mathrm{p}(\mathrm{D}|\mu)\mathrm{p}(\mu)]$$

where for the Gaussian

$$\ln[\mathrm{p}(\mathrm{D}|\mu)] = \ln(\prod_{k=1}^{n} p(x_k|\mu) = \sum_{k=1}^{n} \ln[p(x_k|\mu)] = -\frac{n}{2}\ln[(2\pi)^d|\Sigma|] - \sum_{k-1}^{n} \frac{1}{2}(x_k - \mu)\Sigma^{-1}(x_k - \mu)$$

and

$$\mathrm{p}(\mu) = \frac{1}{(2\pi)^{d/2}\,|\Sigma_0|^{1/2}} \exp[-\frac{1}{2}(\mu - m_0)^t \Sigma_0^{-1}(\mu - m_0)]$$

The Map estimator for the mean is then

$$\hat{\mu} = \underset{\mu}{\operatorname{argmax}}\{\ [-\frac{n}{2}\ln[(2\pi)^d|\Sigma|] - \sum_{k=1}^{n} \frac{1}{2}(x_k - \mu)^t \Sigma^{-1}(x_k - \mu)]$$

$$+\ln[\tfrac{1}{(2\pi)^{d/2}\,|\Sigma_0|^{1/2}} \exp[-\tfrac{1}{2}(\mu - m_0)^t \Sigma_0^{-1}(\mu - m_0)]]\ \}$$