Homework 3

Aukkawut Ammartayakun
CS 539 Machine Learning

Spring 2023

Problem 1

Consider a hidden Markov model in which the emission densities are represented by a parametric model $p(x|z, w)$, such as a linear regression model or a neural network, in which w is a vector of adaptive parameters. Describe how the parameters w can be learned from data using maximum likelihood.

Solution. Given that w is the parameters of the parametric model, one can define the likelihood using the structure of HMM (from emission density, which we assume to be a parametric model) as product of the emission density and the transition density. Now, we can just use EM algorithm to find the maximum likelihood of w.

Problem 2

Show that the finite sample estimator f defined by (11.2)

$$\hat{f} = \frac{1}{L}\sum_{l=1}^{L} f(\mathbf{z}^{(l)})$$

has mean equal to $\mathbb{E}[f]$ and variance given by (11.3)

$$\mathrm{Var}[\hat{f}] = \frac{1}{L}\mathbb{E}[(f - \mathbb{E}[f])^2]$$

Solution. It is clear that,

$$\begin{aligned}
\mathbb{E}[\hat{f}] &= \mathbb{E}\left[\frac{1}{L}\sum_{l=1}^{L} f(\mathbf{z}^{(l)})\right] \\
&= \frac{1}{L}\sum_{l=1}^{L} \mathbb{E}[f(\mathbf{z}^{(l)})] \\
&= \frac{1}{L}\sum_{l=1}^{L} \mathbb{E}[f] \\
&= \mathbb{E}[f]
\end{aligned}$$

We can also show that

$$\begin{aligned}
\mathrm{Var}[\hat{f}] &= \mathbb{E}[(\hat{f} - \mathbb{E}[\hat{f}])^2] \\
&= \mathbb{E}\left[\left(\frac{1}{L}\sum_{l=1}^{L} f(\mathbf{z}^{(l)}) - \mathbb{E}[f]\right)^2\right]
\end{aligned}$$

Since $\hat{f}$ is an unbiased estimator of f, the summation terms after distributing the quadratic form, will be $\frac{1}{L}\mathbb{E}[f^2]$ which then leads to the desired result.

Problem 3

Suppose that z has a uniform distribution over the interval $[0, 1]$. Show that the variable $y = b \tan z + c$ has a Cauchy distribution given by (11.16).

$$q(z) = \frac{k}{1 + (z - c)^2/b^2}$$

Solution. We will show that y has a Cauchy distribution. We know that the transformation of $u(z)$ to $q(y)$ can be done with the Jacobian determinant. We know that $y = b \tan z + c$ or $z = \arctan\left(\frac{y-c}{b}\right)$ and $z \sim U(0, 1)$. We can calculate the Jacobian determinant and transform that as

$$\begin{aligned} q(y) &= u(z)\left|\frac{dz}{dy}\right| \\ &= \left|\frac{dz}{dy}\right| \\ &= \frac{1}{b}\frac{1}{1 + \left(\frac{y-c}{b}\right)^2} \end{aligned}$$

Without loss of generality, we can assume that $\frac{1}{b} = k$ and get what we desired.