Approximating distribution of input conditional on label using IFT of approximate characteristic function of its low dimensional embedding obtained using orthonormal projection of input

Dhruv Kohli [*1] **Praanshu Goyal** [*2]

Abstract

1. Proposed Method

Given dataset $\mathbf{D} = (\mathbf{x_i}, y_i)_{i=1}^{N}$ where $\mathbf{x_i} \in \mathbb{R}^m$ and $y_i \in \{1, 2, \ldots, C\}$ is the class to which $\mathbf{x_i}$ belongs. N is the number of samples in $\mathbf{D}$, m is the dimension of the input and C is the number of possible classes. In the subsequent detail, we show how to compute $P(X = \mathbf{x_i}|y = c)$ and thereafter compute $P(y = c|X = \mathbf{x_i})$ using Bayes therorem. First, we compute an approxiation of the characteristic function $\phi_{X|y=c}(\mathbf{t})$ as in (1). And using the inverse transform of the approximate characteristic function, we obtain an approximation of $P(X = \mathbf{x}|y = c)$ and therefore of $P(y = c|X = \mathbf{x})$.

$$\begin{aligned}
\phi_{X|y=c}(\mathbf{t}) &= \mathbb{E}(e^{i\mathbf{t}^T X}|y = c) \\
&= \sum_{k=0}^{\infty} \frac{i^k}{k!} \mathbb{E}((\mathbf{t}^T X)^k | y = c) \\
&= \sum_{k=0}^{\infty} \frac{i^k}{k!} \frac{\mathbb{E}((\mathbf{t}^T X)^k \delta(y - c))}{P(y = c)} \\
&\approx \sum_{k=0}^{K} \frac{i^k}{k!} \frac{\sum_{j=0}^{N} (\mathbf{t}^T \mathbf{x_j})^k \delta(y_j - c)}{NP(y = c)} \\
&= \frac{1}{NP(y = c)} \sum_{k=0}^{K} \frac{i^k}{k!} \sum_{j=0}^{N} (\mathbf{t}^T \mathbf{x_j})^k \delta(y_j - c)
\end{aligned} \tag{1}$$

$$\begin{aligned}
&2\pi NP(y = c)P(X = \mathbf{x}|y = c) = \\
&= 2\pi NP(y = c) \int_{t_1} \ldots \int_{t_m} \phi_{X|y=c}(\mathbf{t}) e^{-\mathbf{t}^T \mathbf{x}} \partial t_1 \ldots \partial t_m \\
&\approx \sum_{k=0}^{K} \frac{i^k}{k!} \sum_{j=0}^{N} \delta(y_j - c) \int_{t_1} \ldots \int_{t_m} (\mathbf{t}^T \mathbf{x_j})^k e^{-\mathbf{t}^T \mathbf{x}} \partial t_1 \ldots \partial t_m
\end{aligned} \tag{2}$$

[*]Equal contribution [1]Indian Instiute of Technology, Guwahati, India [2]Indian Instiute of Technology, Guwahati, India. Correspondence to: Dhruv Kohli <dhruvkohli99@gmail.com>, Praanshu Goyal <praanshugoyal13@gmail.com>.