Assignment 3

Yang Can, HKUST

2020/4/5

Problem

For a K-class classification problem, we can recode the class label c with a K-dimensional vector $\mathbf{y}$ with all entries equal to $-\frac{1}{K-1}$ except a 1 in position k if $c = k$, i.e.,

$$y_k = \begin{cases} 1, & \text{if } c = k, \\ -\frac{1}{K-1}, & \text{if } c \neq k. \end{cases}$$

Let $\mathbf{f} = (f_1(\mathbf{x}), \ldots, f_K(\mathbf{x}))^T$ with $\sum_{k=1}^K f_k(\mathbf{x}) = 0$, and define

$$L(\mathbf{y}, \mathbf{f}(\mathbf{x})) = \exp(-\frac{1}{K}\mathbf{y}^T\mathbf{f}(\mathbf{x})).$$

(a) Using Lagrange multipliers, derive the population minimizer $\mathbf{f}^*$ of $\mathbb{E}_{\mathbf{y}|\mathbf{x}}[L(\mathbf{y}, \mathbf{f}(\mathbf{x}))]$, subject to $\sum_{k=1}^K f_k(\mathbf{x}) = 0$, and relate these to the class probabilities.

(b) Derive a multiclass boosting algorithm using this loss function and verify that it covers the Adaboost algorithm as a special case ($K = 2$).

(c) Implement your derived algorithm, where you are allowed to call package of trees. Compare your implementation with the existing standard gradient boosting package on a multiclass classification problem. Make some discussion about what you observed.

Requirement

- You need to submit a report, in which you should clearly describe your method and explain your idea. The code should also be included.
- You can use R or Python for coding.
- Your report should be in the **pdf** or **html** format, which is automatically generated by either R markdown or Jupyter notebook.
- The report is due to April, 20, 23:59 pm, 2020 (HK time).