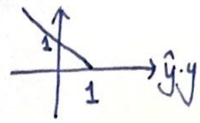
Zhe HUANG

22302183

Exercise: calibration of scoring losses

- Let $\ell^{hinge}(\hat{y}, y) = \max(0, 1-\hat{y}y)$



- We will consider 3 cases in this exercise: case 1: $\eta < \frac{1}{2}$

  case 2: $\eta = \frac{1}{2}$

  case 3: $\eta > \frac{1}{2}$

- Draw $C(\hat{y}, y)$ as a function of $\hat{y}$ in each of these three cases.
- In each case, show which value of $\hat{y}$ minimizes $C(\hat{y}, y)$
- Also, show which predicted class correspond to these $\hat{y}$
- If, instead of the hinge loss, we used the 0/1 loss, which class would be the optimal one in these 3 cases?
- Using the definition of calibration, show the hinge loss is calibrated.

Answer

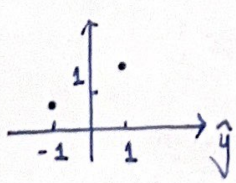
— $C(\hat{y}, y) = \eta\,\ell(\hat{y}, 1) + (1-\eta)\,\ell(\hat{y}, -1)$

$$= \eta\cdot\max(0, 1-\hat{y}) + (1-\eta)\cdot\max(0, 1+\hat{y})$$

$$= \max(0, \eta(1-\hat{y})) + \max(0, (1-\eta)(1+\hat{y}))$$

$$= \begin{cases} 2(1-\eta), & \hat{y}=1 \\ 2\eta, & \hat{y}=-1 \end{cases}$$

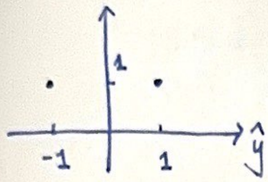
case 1° $\eta < \frac{1}{2}$



$\hat{y} = -1$ minimizes $C(\hat{y}, y)$

predicted class $\{-1\}$ correspond to the $\hat{y}$

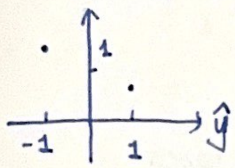
case 2° $\eta = \frac{1}{2}$



$\hat{y} = -1$ or 1 minimize $C(\hat{y}, y)$

predicted class $\{-1, 1\}$ corresponds to the $\hat{y}$

case 3° $\eta > \frac{1}{2}$



$\hat{y} = 1$ minimizes $C(\hat{y}, y)$

predicted class $\{1\}$ corresponds to the $\hat{y}$

— If we use the 0/1 loss instead

Then $C(\hat{y}, y) = \eta\, \ell(\hat{y}, 1) + (1-\eta)\, \ell(\hat{y}, -1)$ where $\ell(\hat{y}, y) = 1_{[y\hat{y} < 0]}$

$$= \eta\, 1_{[\hat{y} < 0]} + (1-\eta)\, 1_{[\hat{y} > 0]}$$

$$= \begin{cases} 1-\eta, & \hat{y} = 1 \\ \eta, & \hat{y} = -1 \end{cases}$$

case 1° $\eta < \frac{1}{2}$: class $\{-1\}$ would be the optimal one

case 2° $\eta = \frac{1}{2}$: class $\{-1, 1\}$ would be the optimal one

case 3° $\eta > \frac{1}{2}$: class $\{1\}$ would be the optimal one

— Proof of "hinge loss is calibrated"

- If $\eta \in [0, \frac{1}{2}[$, then $\inf_{\hat{y} \leq 0} C(\hat{y}, y) = 2\eta < 1 < 2(1-\eta) = \inf_{\hat{y} \geq 0} C(\hat{y}, y)$
- If $\eta \in]\frac{1}{2}, 1]$, then $\inf_{\hat{y} > 0} C(\hat{y}, y) = 2(1-\eta) < 1 < 2\eta = \inf_{\hat{y} \leq 0} C(\hat{y}, y)$

Thus hinge loss is calibrated.