The title of my master's thesis

Christian Peters

April 8, 2021

Contents

Situation: We have n data points $(x_i, y_i),\ i = 1, ..., n$ with $x_i \in \mathbb{R}^d$ and $y \in \{-1, 1\}$.

Probit Model: y_i is a realization of the random variable Y_i. $Y_1, ..., Y_n$ are independent. The distribution of Y_i is as follows:

$$P(Y_i = 1 | x_i; \beta) = \Phi(x_i^T \beta)$$
$$P(Y_i = -1 | x_i; \beta) = 1 - \Phi(x_i^T \beta)$$

where $\beta \in \mathbb{R}^d$. It follows that

$$P(Y_i = y_i | x_i; \beta) = \Phi(y_i x_i^T \beta)$$

Likelihood: The likelihood of a parameter vector β is given as follows:

$$L(\beta) = \prod_{i=1}^{n} P(Y_i = y_i | x_i; \beta) = \prod_{i=1}^{n} \Phi(y_i x_i^T \beta)$$

The negative log-likelihood that we wish to minimize is:

$$\mathcal{L}(\beta) = -\sum_{i=1}^{n} \log \Phi(y_i x_i^T \beta)$$

The weighted case: We introduce sample weights $w_i \in \mathbb{R}_{>0}$ comprising a weight vector $w \in \mathbb{R}^n_{>0}$. Further, let $g(z) = -\log \Phi(-z)$. The objective function now becomes:

$$f_w(\beta) = \sum_{i=1}^{n} w_i g(-y_i x_i^T \beta)$$

Bounds on $g(z)$: The following bounds on $g(z)$ will be useful.

Theorem 1. *Let $g(z) = -\log \Phi(-z)$. Then it holds for all $z \geq 0$ that:*

$$\frac{1}{2} z^2 \leq g(z) \leq 2z^2$$

Proof. TODO. □

References

List of Figures