Exercise E (Separability in logistic regression)

The log-likelihood of a logistic regression model is

$$\ell(\beta; y) = \sum_{i=1}^{m} y_i (x_i^T \beta) - \sum_{i=1}^{m} \log(1 + \exp(x_i^T \beta)) = \log \prod_{i=1}^{m} p(y_i; x_i^T \beta).$$

Suppose there exist a vector β such that

$$x_i^T \beta > 0 \quad \text{if} \quad y_i = 1.$$

$$x_i^T \beta < 0 \quad \text{if} \quad y_i = 0.$$

for $i = 1, \dots, m$. Let A the set of value of β satisfying the above relationship. Note that any $\beta(\kappa) = \kappa\beta \in A$, for any $\kappa > 0$. Then, the log-likelihood in κ is

$$\ell(\beta; y) = \sum_{i: y_i = 1} x_i^T \beta - \sum_{i: y_i = 1} \log(1 + \exp(x_i^T \beta)) - \sum_{i: y_i = 0} \log(1 + \exp(x_i^T \beta))$$

for any $\beta \in B$. Hence, let us evaluate the above in $\beta(\kappa) = \kappa\beta$.

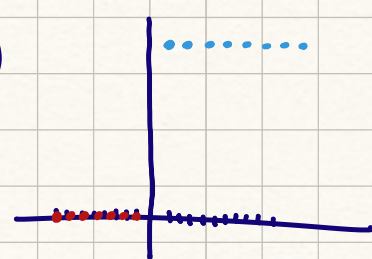
$$\ell(\beta(\kappa); y) = \sum_{i: y_i = 1} \kappa \underbrace{x_i^T \beta}_{\text{Positive}} - \sum_{i: y_i = 1} \log(1 + \exp(\kappa \cdot \underbrace{x_i^T \beta}_{\text{Positive}})) - \sum_{i: y_i = 0} \log(1 + \exp(\kappa \cdot \underbrace{x_i^T \beta}_{\text{Negative}}))$$

Now, let us consider $\kappa \to \infty$. We get that the last term vanishes, hence:

$$\lim_{\kappa \to \infty} \ell(\beta(\kappa); y) = \lim_{\kappa \to \infty} \sum_{i: y_i = 1} \kappa(x_i^T \beta) - \log(1 + \exp(\kappa x_i^T \beta))$$

$$= \lim_{\kappa \to \infty} \sum_{i: y_i = 1} \log\left(\underbrace{\frac{\exp(\kappa x_i^T \beta)}{1 + \exp(\kappa x_i^T \beta)}}_{\text{probabilities}} \right)$$

$$= \sum_{i: y_i = 1} \log(1) = 0.$$



Note that as $\kappa \to \infty$, the function is also increasing, therefore

$$\sup_{\beta} \ell(\beta; y) = 0.$$

(Note that $\mathcal{L}(\beta; y) = \prod_{i=1}^{m} p(y_i; x_i^T \beta)$ and the maximum is 1.)