Math 281C Qualifying Exam – Spring 2023

Let $(Y_1, X_1), \ldots, (Y_n, X_n)$ be i.i.d. copies from the random vector $(Y, X) \in \mathbb{R}^p \times \{-1, 1\}$. Assume $X \sim P_0$ for some probability measure P_0 on $\mathbb{R}^p$, and the conditional distribution of Y given X is

$$P(Y = 1|X = x) = \frac{1}{1 + e^{-x^T\theta_0}}, \quad \text{for every } x \in \mathbb{R}^p, \tag{1}$$

where $\theta_0 \in \mathbb{R}^p$ is the unkown parameter of interest. Moreover, assume that $\Sigma = \mathbb{E}(XX^T)$ is positive definite and $\mathbb{E}\|X\|_2^4 \leq C$ for some $C > 0$. Define the empirical loss/risk function

$$L_n(\theta) = \frac{1}{n}\sum_{i=1}^{n} \log(1 + e^{-Y_i \cdot X_i^T\theta}),$$

and its population counterpart $L(\theta) = \mathbb{E}L_n(\theta)$, where the expectation is over the underlying distribution of (Y, X). Consider the M-estimator (or maximum likelihood estimator)

$$\hat{\theta}_n \in \arg\min_{\theta\in\mathbb{R}^p} L_n(\theta).$$

(a) [5 pts] Show that θ_0 in model (1) is the unique minimizer of $\theta \mapsto L(\theta)$, that is, $\theta_0 = \arg\min_{\theta\in\mathbb{R}^p} L(\theta)$.

(b) [5 pts] Show that for any convex function $h : \mathbb{R}^p \to \mathbb{R}$, if there is some $r > 0$ and a point θ_0 such that $h(\theta) > h(\theta_0)$ for all θ satisfying $\|\theta - \theta_0\|_2 = r$. Then $h(\theta') > h(\theta_0)$ for all θ' with $\|\theta' - \theta_0\|_2 > r$.

(c) [5 pts] Show that $\theta \mapsto L(\theta)$ and $\theta \mapsto L_n(\theta)$ are convex.

(d) [10 pts] Show that $\hat{\theta}_n$ is consistent, that is, $\hat{\theta}_n \to \theta_0$ in probability as $n \to \infty$.

(e) [10 pts] Derive the convergence rate of $\hat{\theta}_n$, that is, find a positive sequence δ_n satisfying $\|\hat{\theta}_n - \theta_0\|_2 = O_P(\delta_n)$.

(f) [5 pts] Find the asymptotic distribution of $\sqrt{n}(\hat{\theta}_n - \theta_0)$, and describe its asymptotic covariance matrix. For this sub-question you only need to provide a heuristic or informal argument.