Problem Set 6

Due Nov 28, before class

1. Likelihood for probit model

Write down the likelihood function for binary probit model, and derive the steps for Newton's method for computing the MLE.

2. Quadratic discriminant analysis

Assume that $y_i \sim \text{Bernoulli}(q)$ and

$$x_i \mid y_i = 1 \sim \text{N}(\mu_1, \Sigma_1), \quad x_i \mid y_i = 0 \sim \text{N}(\mu_0, \Sigma_0)$$

where $x_i \in \mathcal{R}^p$ does not contain 1. Prove that

$$\text{logit}\{\text{pr}(y_i = 1 \mid x_i)\} = \alpha + x_i^T \beta + x_i^T \Lambda x_i,$$

where

$$\begin{aligned}
\alpha &= \log \frac{q}{1-q} - \frac{p}{2} \log \frac{\det(\Sigma_1)}{\det(\Sigma_0)} - \frac{1}{2} \left(\mu_1^T \Sigma_1^{-1} \mu_1 - \mu_0^T \Sigma_0^{-1} \mu_0 \right), \\
\beta &= \Sigma_1^{-1} \mu_1 - \Sigma_0^{-1} \mu_0, \\
\Lambda &= -\frac{1}{2} (\Sigma_1^{-1} - \Sigma_0^{-1}).
\end{aligned}$$

Remark: This problem extends the linear discriminant model to the quadratic discriminant model by allowing for heteroskedasticity in the conditional Normality of x given y. It implies the logistic model with the linear, quadratic, and interaction terms of the basic covariates.

3. Inverse model for the multinomial logit model

Assume that $y_i \sim \text{Multinomial}(1; q_1, \dots, q_K)$ and $x_i \mid y_i = k \sim \text{N}(\mu_k, \Sigma)$, where x_i does not contain 1. Verify that $y_i \mid x_i$ follows a multinomial logit model:

$$\text{pr}(y_i = k \mid x_i) \;=\; \frac{e^{\alpha_k + x_i^T \beta_k}}{\sum_{l=1}^K e^{\alpha_l + x_i^T \beta_l}},$$

where

$$\alpha_k \;=\; \log q_k - \frac{1}{2}\mu_k^T \Sigma^{-1} \mu_k, \quad \beta_k \;=\; \Sigma^{-1}\mu_k.$$

4. Iteratively reweighted least squares algorithm for the multinomial logit model

Similar to the binary logistic model, Newton's method for computing the MLE for the multinomial logit model can be written as iteratively re-weighted least squares. Give the details.