BIOS 560R Advanced Statistical Computing
Fall 2012
Homework 2

Due 10/09/2012 at 4pm before the class

Problem 1: In the logistic regression model of Example 5 (lecture "MM Algorithm"), it is possible to separate parameters and avoid matrix inversion altogether.

1. In constructing a minorizing function, first prove the inequality

$$-\log\left\{1+\exp(X_i^{\mathrm{T}}\theta)\right\} \geq -\log\left\{1+\exp(X_i^{\mathrm{T}}\theta^{(k)})\right\} - \frac{\exp(X_i^{\mathrm{T}}\theta)-\exp(X_i^{\mathrm{T}}\theta^{(k)})}{1+\exp(X_i^{\mathrm{T}}\theta^{(k)})},$$

   with equality when $\theta = \theta^{(k)}$. This eliminates the log terms.

2. Now apply the arithmetic-geometric mean inequality to the exponential function $\exp(X_i^{\mathrm{T}}\theta)$ to separate parameters. Assuming that θ has p components and that there are n observations, show that these maneuvers lead to the minorizing function

$$g(\theta|\theta^{(k)}) = -\frac{1}{p}\sum_{i=1}^{n}\frac{\exp(X_i^{\mathrm{T}}\theta^{(k)})}{1+\exp(X_i^{\mathrm{T}}\theta^{(k)})}\sum_{j=1}^{p}\exp\{pX_{ij}(\theta_j-\theta_j^{(k)})\} + \sum_{i=1}^{n}Y_iX_i^{\mathrm{T}}\theta$$

   up to a constant that does not depend on θ.

3. Finally, prove that maximizing $g(\theta|\theta^{(k)})$ consists in solving the equation

$$-\sum_{i=1}^{n}\frac{\exp(X_i^{\mathrm{T}}\theta^{(k)})X_{ij}\exp(-pX_{ij}\theta_j^{(k)})}{1+\exp(X_i^{\mathrm{T}}\theta^{(k)})}\exp(pX_{ij}\theta_j) + \sum_{i=1}^{n}Y_iX_{ij} = 0$$

   for each j. This can be accomplished numerically and you do not need to show that.

Problem 2: Compare different algorithms for the logistic regression model of Example 5 (lecture "MM Algorithm").

1. Outline the (standard) Newton-Raphson algorithm and the Fisher Scoring algorithm. Show the relation between Newton-Raphson and Fisher Scoring.

2. Implement the Newton-Raphson algorithm in R.

3. Implement the MM algorithm of Example 5 in R.

4. Conduct a simulation study. For 1,000 individuals, generate the binary response from

$$\Pr(Y_i = 1) = \frac{\exp(X_i\theta)}{1 + \exp(X_i\theta)},$$

   where $\theta = 0.3$ and $X_i \sim N(0, 1)$. Apply your NR and MM algorithms to this data set; select your own starting value and stopping criterion, but make sure they are the same for the two algorithms. For each algorithm,

   (a) report MLE $\widehat{\theta}$ and number of iteration,

   (b) make a table similar to that on Page 13 of the lecture notes "EM Algorithm". If the table is too long, you can just show the first and last few rows.

   (c) compare the two algorithms in terms of convergence rate.

Problem 3: The standard linear regression model can be written as

$$Y_i = X_i^{\mathrm{T}}\beta + \epsilon_i,$$

where $\epsilon_i \sim N(0, \sigma^2)$. Suppose that X_i is observed, but that rather than observing Y_i, we observe $Y_i^* = \min\{Y_i, c\}$, where the censoring value c is known and is constant for all i. The EM algorithm offers a vehicle for estimating the parameters (θ, σ^2) in the presence of censoring. Suppose we reorder the data so that the first m subjects have observed Y_i ($Y_i^* = Y_i$) and the rest $n - m$ subjects have censored Y_i ($Y_i^* = c$).

1. Write out the complete-data log likelihood. Identify the conditional expectations to be evaluated in the E step.

2. Show that

$$\begin{aligned}
\mathrm{E}(Y_i | Y_i^* = c, \beta^{(k)}, \sigma^{(k)}) =& X_i^{\mathrm{T}}\beta^{(k)} + \sigma^{(k)}\psi\left(\frac{c - X_i^{\mathrm{T}}\beta^{(k)}}{\sigma^{(k)}}\right) \\
\mathrm{E}[(Y_i - X_i^{\mathrm{T}}\beta^{(k)})^2 | Y_i^* = c, \beta^{(k)}, \sigma^{(k)}] =& \sigma^{2(k)}\left\{1 + \frac{c - X_i^{\mathrm{T}}\beta^{(k)}}{\sigma^{(k)}}\psi\left(\frac{c - X_i^{\mathrm{T}}\beta^{(k)}}{\sigma^{(k)}}\right)\right\},
\end{aligned}$$

   where $\psi(z) = \phi(z)/(1 - \Phi(z))$, and $\phi(z)$ and $\Phi(z)$ are corresponding pdf and cdf of a standard normal random variable.

   Hint:

$$\begin{aligned}
&\int_u^\infty z\phi(z)dz = (2\pi)^{-1/2}\int_u^\infty \exp\left(-\frac{z^2}{2}\right)d\frac{z^2}{2} = (2\pi)^{-1/2}\exp(-\frac{u^2}{2}) = \phi(u); \\
&\int_u^\infty z^2\phi(z)dz = -(2\pi)^{-1/2}\int_u^\infty z d\exp\left(-\frac{z^2}{2}\right) = (2\pi)^{-1/2}u\exp\left(-\frac{u^2}{2}\right) + (2\pi)^{-1/2}\int_u^\infty \exp\left(-\frac{z^2}{2}\right)dz \\
&\qquad = u\phi(u) + 1 - \Phi(u)
\end{aligned}$$

3. Write out the formulas for updating parameters in the M step.