STAT 798c, HW Problem Set #4, due Monday 3/31/03

These problems involve putting together some ideas of numerical maximization of likelihoods *along with* verifications that the maximum achieved is the right one. For both problems, the ASCII datasets mentioned can be found on the public course-directory **/usr/local/StatData/SplusCrs** or at the Data link on the course web-page.

(A). Using dataset **Dset3A.asc** , which has 300 records each consisting of 3 covariates Z_{i1}, Z_{i2}, Z_{i3} and one binary response-variable Y_i :

find the MLE for the parameter-vector β in the generalized linear model

$$P(Y = 1 \,|\, Z_1,\, Z_2,\, Z_3) \;=\; F(\beta_0 \,+\, \sum_{j=1}^{3} \beta_j \, Z_j) \quad , \qquad F(x) \;=\; \Phi(x^{1/3})$$

Verify using any methods you choose that your reported MLE does indeed globally maximize the likelihood. (Note that the derivative(s) of the log-likelihood do not behave well near $\beta = \mathbf{0}$.) Give as convincing pictorial or numerical evidence as you can that you have attained the global maximum. Preferably, maximize using more than one method (nlmin, Gradient Ascent with or without optimized step-length, Newton-Raphson, etc.)

(B). Dataset **Dset3B.asc** contains 1000 values $T_j,\ j = 1, \ldots, 1000$ simulated from the mixture density

$$f(t) \;=\; 2p\, e^{-2t} \,+\, (1-p)\, \lambda\, e^{-\lambda\, t} \quad , \qquad 0 < \lambda < 2\ , \ \ 0 < p < 1$$

with $\vartheta = (\lambda, p)$ as unknown parameter. Find the Maximum Likelihood Estimator for ϑ by two distinct methods:

(i). Steepest-ascent or Quasi-Newton direct maximization of the likelihood.
(ii). The EM algorithm.

For each method display all iterations required, starting with the initial guess $\vartheta_1 = (1, 0.5)$, to converge so that at the last iteration the log-likelihood $\max(l(\vartheta_n) - l(\vartheta_{n-1}),\ \frac{1}{3}\,\|\vartheta_n - \vartheta_{n-1}\|_2) \leq\ 10^{-3}$.