1. Let $\boldsymbol{x}_1, \dots, \boldsymbol{x}_n \in \mathbb{R}^p$ be known vectors (all vectors are column vectors in this problem) with $p \leq n$, let $\boldsymbol{\beta} \in \mathbb{R}^p$ be an unknown parameter vector, and let $Y_1, \dots, Y_n$ be independent Bernoulli random variables such that

$$P_{\boldsymbol{\beta}}(Y_i = 1) = 1 - P_{\boldsymbol{\beta}}(Y_i = 0) = \Phi(\boldsymbol{x}_i'\boldsymbol{\beta}), \quad i = 1, \dots, n, \tag{1}$$

where Φ denotes the c.d.f. of the standard normal distribution and "prime" denotes transpose. This problem concerns using the EM algorithm to find the MLE of $\boldsymbol{\beta}$ from the data $\boldsymbol{Y} = (Y_1, Y_2, \dots, Y_n)'$.

(a) Show that if $\varepsilon_1, \dots, \varepsilon_n$ are i.i.d. standard normals and $W_i = \boldsymbol{x}_i'\boldsymbol{\beta} + \varepsilon_i$, $i = 1, \dots, n$, then

$$Y_i = \mathbf{1}\{W_i > 0\}, \quad i = 1, \dots, n \tag{2}$$

have the distribution (1). For the rest of this problem assume the Y_i are defined by (2) and observed but that the W_i and ε_i are unobserved.

(b) Show that the complete log-likelihood function $\ell_c(\boldsymbol{\beta})$ of $W_1, \dots, W_n$, if they were observed, is given by

$$\ell_c(\boldsymbol{\beta}) = -\frac{1}{2}\sum_{i=1}^{n}(W_i - \boldsymbol{x}_i'\boldsymbol{\beta})^2$$

up to additive constants.

(c) The E step of the EM algorithm involves computing

$$Q(\widetilde{\boldsymbol{\beta}}, \boldsymbol{\beta}) = E_{\boldsymbol{\beta}}\left[\ell_c(\widetilde{\boldsymbol{\beta}}) \middle| \boldsymbol{Y}\right].$$

Show that

$$E_{\boldsymbol{\beta}}\left[(W_i - \boldsymbol{x}_i'\widetilde{\boldsymbol{\beta}})^2 \middle| \boldsymbol{Y}\right] = E_{\boldsymbol{\beta}}\left[(W_i - \boldsymbol{x}_i'\widetilde{\boldsymbol{\beta}})^2 \middle| Y_i\right], \tag{3}$$

and, up to additive terms C, C' that don't depend on $\widetilde{\boldsymbol{\beta}}$, that

$$E_{\boldsymbol{\beta}}\left[(W_i - \boldsymbol{x}_i'\widetilde{\boldsymbol{\beta}})^2 \middle| Y_i = 1\right] = (\boldsymbol{x}_i'\widetilde{\boldsymbol{\beta}})^2 - 2\boldsymbol{x}_i'\widetilde{\boldsymbol{\beta}}\left(\boldsymbol{x}_i'\boldsymbol{\beta} + \frac{\phi(\boldsymbol{x}_i'\boldsymbol{\beta})}{\Phi(\boldsymbol{x}_i'\boldsymbol{\beta})}\right) + C, \quad \text{and} \tag{4}$$

$$E_{\boldsymbol{\beta}}\left[(W_i - \boldsymbol{x}_i'\widetilde{\boldsymbol{\beta}})^2 \middle| Y_i = 0\right] = (\boldsymbol{x}_i'\widetilde{\boldsymbol{\beta}})^2 - 2\boldsymbol{x}_i'\widetilde{\boldsymbol{\beta}}\left(\boldsymbol{x}_i'\boldsymbol{\beta} - \frac{\phi(\boldsymbol{x}_i'\boldsymbol{\beta})}{\Phi(-\boldsymbol{x}_i'\boldsymbol{\beta})}\right) + C', \tag{5}$$

where ϕ is the standard normal density function. You don't have to explicitly find C, C'.

(d) Let $\boldsymbol{X}$ be the $(n \times p)$ matrix with rows $\boldsymbol{x}_1', \dots, \boldsymbol{x}_n'$, assumed to be of full rank p, and let $\boldsymbol{v} = \boldsymbol{v}(\boldsymbol{\beta})$ be the n-long vector with entries

$$v_i = \begin{cases} \phi(\boldsymbol{x}_i'\boldsymbol{\beta})/\Phi(\boldsymbol{x}_i'\boldsymbol{\beta}), & \text{if } Y_i = 1 \\ -\phi(\boldsymbol{x}_i'\boldsymbol{\beta})/\Phi(-\boldsymbol{x}_i'\boldsymbol{\beta}), & \text{if } Y_i = 0. \end{cases}$$

Show that the recursion for the EM iterates $\{\boldsymbol{\beta}^{(k)}\}$ given by maximizing $Q(\widetilde{\boldsymbol{\beta}}, \boldsymbol{\beta})$ over $\widetilde{\boldsymbol{\beta}}$ is

$$\boldsymbol{\beta}^{(k+1)} = \boldsymbol{\beta}^{(k)} + (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{v}(\boldsymbol{\beta}^{(k)}). \tag{6}$$

Hint: Write $Q(\widetilde{\boldsymbol{\beta}}, \boldsymbol{\beta})$ in matrix notation using $\boldsymbol{X}$ and $\boldsymbol{v}$ and then maximize over $\widetilde{\boldsymbol{\beta}}$.

2. Let $\theta \sim N(0,1)$ and, conditional on θ, let $Y_i \sim N(\theta, 1), i = 1, 2, \cdots, n$ be i.i.d. Let $\overline{Y} = \sum_{i=1}^n Y_i/n$.

(a) Compute the density function of $\overline{Y}$.

(b) Compute the posterior distribution of θ given $\overline{Y}$.

(c) Compute the conditional expectation $E[\theta|\overline{Y}]$, and determine the behavior of the posterior distribution of θ as $n \to \infty$.