Likelihood Theory and Extensions (BIOS:7110)
Breheny

Assignment 3

Due: Tuesday, September 15

1. *The Riemann-Stieltjes Integral.* Suppose μ is a bounded, nondecreasing function on $[a, b]$ and that μ is continuous at a point $x_0 \in [a, b]$. Further suppose that $g(x_0) = 1$ and $g(x) = 0$ if $x \neq x_0$. Prove that $\int g d\mu = 0$. Hint: In order to show that $\inf_P U(P, g, \mu) = 0$, you need to show that (i) for all P, $0 \leq U(P, g, \mu)$ and that (ii) for any $\epsilon > 0$, we can find a partition P such that $U(P, g, \mu) < \epsilon$.

2. *Matrix square root.* Let $\mathbf{A}$ be a symmetric, positive definite matrix with eigendecomposition $\mathbf{A} = \mathbf{Q}\boldsymbol{\Lambda}\mathbf{Q}^\top$.

   (a) Find $\mathbf{A}^{1/2}$ and show that it is a matrix square root.
   - If $\mathbf{A}$ was not symmetric, would your derivation still work?
   - If $\mathbf{A}$ was positive semidefinite, would your derivation still work?

   (b) Find $\mathbf{A}^{-1/2}$ and show that it is both the square root of $\mathbf{A}^{-1}$ and the inverse of $\mathbf{A}^{1/2}$.
   - If $\mathbf{A}$ was not symmetric, would your derivation still work?
   - If $\mathbf{A}$ was positive semidefinite, would your derivation still work?

   (c) From (b), it follows that $\mathbf{A}^{-1/2}\mathbf{A}\mathbf{A}^{-1/2} = \mathbf{I}$. If $\mathbf{A}$ is not full rank, but we take generalized inverses where needed, what does $\mathbf{A}^{-1/2}\mathbf{A}\mathbf{A}^{-1/2}$ equal?

3. *Trace and eigenvalues.* Let $\mathbf{A}$ be a $d \times d$ symmetric matrix. Prove that $\text{tr}(\mathbf{A}) = \sum_i \lambda_i$, where $\{\lambda_i\}_{i=1}^d$ are the eigenvalues of $\mathbf{A}$.

4. *Projection matrices and rank.* A matrix $\mathbf{P}$ satisfying $\mathbf{P} = \mathbf{P}\mathbf{P}$ is known as an *idempotent matrix*, or *projection matrix*. Below, suppose that $\mathbf{P}$ is a symmetric projection matrix.

   (a) Show that every eigenvalue of $\mathbf{P}$ must be either 1 or 0.

   (b) Show that the rank of $\mathbf{P}$ equals the trace of $\mathbf{P}$.

5. *Logistic regression derivatives.* The logistic regression model states that Y_i is equal to 1 with probability π_i and 0 otherwise, with π_i related to a set of linear predictors $\{\eta_i\}$ by the following model:

$$\log \frac{\pi_i}{1 - \pi_i} = \eta_i \qquad \text{for } i = 1, 2, \ldots, n$$
$$\boldsymbol{\eta} = \mathbf{X}\boldsymbol{\beta}$$

   where $\boldsymbol{\eta} \in \mathbb{R}^n$, $\boldsymbol{\beta} \in \mathbb{R}^d$, and $\mathbf{X}$ is an $n \times d$ matrix. Let $\ell : \mathbb{R}^n \to \mathbb{R}$ denote the log-likelihood, $\ell = \sum_{i=1}^n \ell_i$, where ℓ_i is the contribution to the log-likelihood from observation i. For (c), (e), and (f), express your answer in vector/matrix notation, not as a collection of scalar terms (i.e., something like $\mathbf{a} + \mathbf{b}$, not $z_1 = 1, z_2 = 3, \ldots$).

   (a) Find $\partial \ell_i / \partial \eta_i$. Simplify your answer as much as possible.

(b) Find $\partial^2 \ell_i / \partial \eta_i^2$. Simplify your answer as much as possible.

(c) Find $\partial \ell / \partial \boldsymbol{\eta}$.

(d) Find $\partial^2 \ell / \partial \boldsymbol{\eta}^2$.

(e) Find $\partial \boldsymbol{\eta} / \partial \boldsymbol{\beta}$.

(f) Find $\partial \ell / \partial \boldsymbol{\beta}$.