Likelihood Theory and Extensions (BIOS:7110)
Breheny

Assignment 3
Due: Monday, September 23

1. *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?

2. *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 $\mathbf{A}$.

3. *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 $\mathbf{P}$.

4. *Logistic regression.* 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. For (b)-(d), 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) Let ℓ_i denote the contribution to the log-likelihood from observation i. Find the partial derivative of ℓ_i with respect to η_i. Simplify your answer as much as possible.

   (b) Let $\ell : \mathbb{R}^n \to \mathbb{R}$ denote the log-likelihood as a function of the linear predictors $\boldsymbol{\eta}$. Find $\nabla_{\boldsymbol{\eta}}\ell$.

   (c) Find $\nabla_{\boldsymbol{\beta}}\boldsymbol{\eta}$.

   (d) Find $\nabla_{\boldsymbol{\beta}}\ell$.

5. *Exponential Taylor series.* Note: (b) and (c) are not trick questions; they take little effort to derive, but the results are useful to know.

(a) Show that for any x,

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}.$$

Note: one can partially prove this using the Poisson distribution, but this proof would only work for $x > 0$.

(b) Starting with the result in (a), derive the infinite series for e^{ax}.

(c) Starting with the result in (a), derive the infinite series for b^x, where $b > 0$.

(d) Let $f : \mathbb{R}^d \to \mathbb{R}$. What is the second-order Taylor series for $f(\mathbf{x}) = \exp(\mathbf{a}^\top \mathbf{x})$ about $\mathbf{x} = \mathbf{0}$? Give both the o-notation and Lagrange forms.

(e) Suppose $\mathbf{a} = [2 \quad -1]^\top$ and $\mathbf{x} = [1 \quad 1]^\top$. Find the point $\bar{\mathbf{x}}$ on the line segment connecting $\mathbf{x}$ and $\mathbf{0}$ that satisfies the Lagrange form of Taylor's theorem.