Homework 4
Stat 215A, Fall 2024

Due: this homework is not required and will not count towards your grade. But it is good practice for the midterm.

1 EM Algorithm

Suppose $X_1, \ldots, X_n$ are i.i.d. observations from a mixture of two Poisson distributions, Pois (μ_0) and Pois (μ_1), with mixing probabilities of π and $1-\pi$, respectively. That is, there is an initial probability π that an observation X_i is drawn from Pois (μ_0) and probability $1-\pi$ from Pois (μ_1). Recall that if $X \sim \text{Pois}(\mu)$, the probability density function is given by

$$p_\mu(x) = \frac{\mu^x e^{-\mu}}{x!} \tag{1.1}$$

- Define the observed data vector, the latent variable vector, and the distribution of the latent variable.
- Write down the E step for estimating μ_0, μ_1, π.
- Write down the M step for estimating μ_0, μ_1, π.
- Give an initial estimator to start the EM algorithm.
- Write R code to implement the E and M steps.
- Simulate data from a mixture of two Poisson distributions, where you know the true parameters, and run your EM algorithm on the simulated data. Show the accuracy of EM clustering as you vary the values of μ_0 and μ_1.
- Now simulate data from a mixture of two binomial distributions, Binom $(100, \pi_0)$ and Binom $(100, \pi_1)$. Show the accuracy of your Poisson EM clustering algorithm as you vary the values of π_0 and π_1 in this misspecified model.

2 Classification

Read section 7.2 (p. 121) on probit models in Freedman and complete questions 1, 2, and 3 in Exercise set B on page 124.

3 The Hat Matrix

The Sherman Morrison formula gives us an expression for the inverse of a rank-1 update to a matrix. if A is an invertible $n \times n$ matrix, and $u, v \in \mathbb{R}^n$, then:

$$(A + uv^\perp)^{-1} = A^{-1} - \frac{A^{-1}uv^\top A^{-1}}{1 + v^\top A^{-1}u} \tag{3.1}$$

Use this to prove equations (5.1) and (5.5) in Hoaglin, David C., and Roy E. Welsch. "The hat matrix in regression and ANOVA." Note that they consider $\mathbf{x}_i$ a column vector—you can use either convention. Also, equation (5.5) contains a typo, and should be:

$$\hat{\beta} - \hat{\beta}_{(i)} = \cdots$$

4 Miscellaneous

- In the two examples introduced by Freedman, the practitioners seek to perform inference for some parameters of interest. How do the scientific questions the practitioners wish to answer compare to the standard statistical formulation of significance tests?
- Summarize a critique Freedman has for the "subjective" camp of statisticians. What are counterpoints this camp would respond with to this critique?
- Summarize a critique Freedman has for the "objective" camp of statisticians. What does Freedman propose as a possible remedy for this critique?