Homework 2: problems about exponential families

Rajdeep Brahma

Due: February 9th at 11:59 PM

This homework set will cover problems concerning exponential family theory. All derivations must be typed. Screenshots of work done with pen and paper will not be accepted.

Problem 1: Verify that displayed equation 7 in the exponential family notes holds for the binomial distribution, the Poisson distribution, and the normal distribution with both μ and σ^2 unknown.

Solution: First, we consider the binomial distribution. We have $c(\theta) = \mathrm{n}\log(1+e^{\theta})$ where $\theta = \log(\frac{p}{1-p})$. Now $\nabla c(\theta) = n\frac{e^{\theta}}{1+e^{\theta}} = np$ and $\nabla^2 c(\theta) = n\frac{e^{\theta}}{(1+e^{\theta})^2} = np(1-p)$. So, this matches the expectation and variance of the binomial distribution. Next, we consider the poisson distribution. We have $c(\theta) = e^{\theta}$ where $\theta = \log(\lambda)$. Now $\nabla c(\theta) = e^{\theta} = \lambda$ and $\nabla^2 c(\theta) = e^{\theta} = \lambda$. So, this matches the expectation and variance of the poisson distribution. Finally, we consider the normal distribution. We have $c(\theta) = -\frac{\theta_1^2}{4\theta_2} + \frac{1}{2}\log(-\frac{1}{2\theta_2})$ where $\theta = (\theta_1, \theta_2) = (\frac{\mu}{\sigma^2}, -\frac{1}{2\sigma^2})$. Now $\nabla c(\theta) = (-\frac{\theta_1}{2\theta_2}, \frac{\theta_1^2}{4\theta_2^2} - \frac{1}{2\theta_2}) = (\mu, \mu^2+\sigma^2)$ and $\nabla^2 c(\theta) = \begin{pmatrix} -\frac{1}{2\theta_2} & \frac{\theta_1}{2\theta_2^2} \\ \frac{\theta_1}{2\theta_2^2} & \frac{1}{2\theta_2^2} - \frac{\theta_1^2}{2\theta_2^3} \end{pmatrix} = \begin{pmatrix} \sigma^2 & 2\mu\sigma^2 \\ 2\mu\sigma^2 & 2\sigma^4+4\mu^2\sigma^2 \end{pmatrix}$. So, this matches the expectation and variance of (X, X^2) where $X \sim N(\mu, \sigma^2)$.

Problem 2: This problem concerns the proof of Theorem 3 in the exponential family notes. Do the following:

- **part a**: Show that the second derivative of the map h is equal to $-\nabla^2 c(\theta)$ and justify that this matrix is negative definite when the exponential family model is identifiable.
- **part a**: Finish the proof of Theorem 3.

Note that part a will be referenced later in this course. Hence, it is treated as its own sub-problem.

Solution:

$$h(\theta) = \langle \mu, \theta \rangle - c(\theta).$$

.Hence $\nabla h(\theta) = \mu - \nabla c(\theta)$ and $\nabla^2 h(\theta) = -\nabla^2 c(\theta) = -\mathrm{Var}_\theta(Y)$. This is a negative definite matrix because the exponential family model is identifiable implying that Y is not concentrated on a hyperplane.
To finish the proof of Theorem 3 we need to prove 2 more results. First one of them is "cumulant functions are infinitely differentiable and are therefore continuously differentiable". This is because $c(\theta) = \log\left(\int \exp\left(\langle Y(w), \theta\rangle\right)\lambda(dw)\right)$. Since log() and exp() functions are infinitely differentiable, $c(\theta)$ is too. The second result that needs to be proved is "$g^{-1}(\theta)$ is infinitely differentiable". This is because $\frac{\partial A^{-1}}{\partial x} = -A^{-1}\frac{\partial A}{\partial x}A^{-1}$. Hence, matrix inversion is infinitely differentiable and so is $g^{-1}(\theta)$ since $g(\theta) = \nabla c(\theta)$ is

infinitely differentiable.

Problem 3: Let y be a regular full exponential family with canonical parameter θ. Verify that y is sub-exponential.

Problem 4: In the notes it was claimed that the scalar products of $\sum_{i=1}^n \{y_i - \nabla c(\theta)\}$ are also sub-exponential (see the "Finite sample concentration of MLE" section in the exponential family notes). Show that this is in fact true when the observations y_i are iid realizations from a regular full exponential family.

Problem 5: Derive the MLEs of the canonical parameters of the binomial distribution and the normal distribution with both μ and σ^2 unknown.

Problem 6: Derive the asymptotic distribution for $\hat{\tau}$, the MLE of the submodel mean value parameter vector. Hint: use the Delta method.

Problem 7: Prove Lemma 1 in the exponential family notes.