Coursework **Fall 2019**
Machine Learning (B555) **Roni Khardon**
Department of Computer Science, IU Bloomington

Written Assignment 3

This assignment is due by 10/23 11:59pm. Please submit solutions as a single PDF file through Canvas. Handwritten or typed solutions are fine, so long as the solution is legible and is submitted as a PDF. You might find LaTeX to be a useful tool for typesetting text and math expressions.

1. Bayesian Inference for Univariate Normal: Consider an IID sample for a Normal variable

$$p(X|\mu, \lambda) = \prod_{i=1}^{N} (\frac{\lambda}{2\pi})^{1/2} e^{-\frac{\lambda}{2}(x_i-\mu)^2}$$

with mean μ and precision λ, and consider a Normal-Gamma prior given by

$$p(\mu, \lambda) = \mathcal{N}(\mu|\mu_0, \frac{1}{\beta_0 \lambda}) \text{ Gamma}(\lambda|a_0, b_0).$$

Calculate the posterior for μ and λ and show that it is also decomposable as a Normal-Gamma distribution, with new parameters a_N, b_N, β_N, μ_N. In particular explicitly calculate and show the parameters of the posterior when given in this form.

Notes: (1) The text on page 101 of the textbook is relevant for this question but the notation there can be confusing. It is therefore recommended to use the notation given here. (2) This is a special case of the Bayesian linear regression result for (w, β) derived in class but in this assignment you are asked to develop the result from first principles via completing the square.

2. Calculate one iteration of the Newton-Raphson method for minimizing the function $f(x) = x_1^3 + 5x_1x_2^2 - 7x_1^2x_2$ where $x = (x_1, x_2)^T$. Use $x = (1, 1)^T$ as the initial value.

3. Solve problem 4.18 (page 223) in the textbook.

4. In this question we develop the properties for the scaled-variants of exponential family distributions that were used in the derivation of GLM.
(i) Show that if $p_1(x)$ is a normalized probability density function then $p_2(x) = \frac{1}{s}p_1(\frac{x}{s})$ is also a normalized probability density function.
(ii) Now consider the scaled variant of the one dimensional exponential family distribution $p(x|s, \eta) = \frac{1}{s}h(\frac{x}{s})g(\eta)e^{\eta x/s}$. Show that the expectation parameter $\theta = E[x|\eta, s]$ satisfies $\theta = E[x|\eta, s] = -s\frac{d}{d\eta}\ln g(\eta)$.
(iii) Show that the variance satisfies $\mathbb{V}\text{ar}(x|\eta) = -s^2\frac{d^2}{d\eta^2}\ln g(\eta)$.

5. Consider a Poisson likelihood function, $\text{Poisson}(x|\lambda) = \frac{1}{x!}e^{-\lambda}\lambda^x$, with prior, $\text{Gamma}(\lambda|a, b)$, on λ. Develop a Gaussian approximation to $p(\lambda|x, a, b)$ using the Laplace approximation. First develop the formula in general and then apply it to the case $a = 1$, $b = 1$, $x = 3$. For this case, plot the approximation and the true function for $0 \leq \lambda \leq 10$.