Prob. ML — Spring 2020 **Midterm Exam** Due Feb 21, 2020

Honor code:

As a student and citizen of the Duke University Community:
I will not lie, cheat, or steal in my academic endeavors.
I will forthrightly oppose each and every instance of academic dishonesty.
I will communicate directly with any person or persons I believe to have been dishonest. Such communication may be oral or written. Written communication may be signed or anonymous.
I will give prompt written notification to the appropriate faculty member and to the Dean of Trinity College or the Dean of the School of Engineering when I observe academic dishonesty in any course.
I will let my conscience guide my decision about whether my written report will name the person or persons I believe to have committed a violation of this code.

Problem 1

Linear regression.

1. (a) Write down the minimization problem for multivariate logistic regression and state the Newton-Raphson solution for the coefficients.

   (b) Write down the standard likelihood for multivariate logistic regression. Under what assumptions is the MLE stable ?

2. (a) Write down the minimization problem for multivariate logistic regression with a penalized loss, specifically the standard ℓ_2 shrinkage.

   (b) Write down the posterior distribution for the standard logistic regression problem with a standard shrinkage prior, you can leave the denominator as an integral.

Problem 2

Nonlinear regression.

1. Derive support vector machines from a geometric perspective and what is a support vector ?

2. Write out the posterior predictive distribution for a vector of unseen responses $\mathbf{y}_*$ in a Gaussian process.

3. Write out the posterior predictive distribution for a vector of unseen responses $\mathbf{y}_*$ in a Gaussian process with a linear kernel.

Problem 3

Sparse regression.

1. Write out the penalized loss function for LASSO and the ℓ_0 ideal penalized loss.
2. Write out pseudo code for LARS.
3. Provide an explanation as to why LASSO solutions can be sparse.
4. Derive the double exponential distribution as a scale mixture of normals.

Problem 4

Data Analysis.
This question is based the hw2 geyser.csv data available on the course website

1. Use cross-validation to set the tuning parameter σ for a Gaussian process regression model using the following kernel

$$k(u, v) = \exp(-\sigma \|u - v\|^2).$$

2. Compute the cross-validation error for LASSO, Elastic net, and ridge regression. How did you set the regularization parameters.
3. Do nonlinear or linear models predict more accurately on this data set.