Final Exam

The exam is open book. You may complete this exam by any means you wish. However, **you and only you are to work on this exam**. Any help from the internet should be reported in the appendix. There are a total of 100 points.

Read each problem carefully. Show enough work to prove that you know what you are doing and to receive as much partial credit as possible. If computations are required, you must show supporting calculations or/and computer output. State all conclusions in the context of the problem.

Problem 1

The goal of this problem is to generate a sample from the following normal location mixture

$$p_1 N(7, 0.5^2) + (1 - p_1) N(10, 0.5^2).$$

a. Implement a Metropolis-Hastings algorithm to simulate the above mixture distribution with $p_1 = 0.7$, using $N(x^{(t)}, 0.01^2)$ as the proposal distribution. For each of the three starting values, $x^{(0)} = 0$, 7, and 15, run the chain for 10,000 iterations. Plot the sample path of the output from each chain.

b. For each of the simulations, create a histogram of the realizations with the true density superimposed on the histogram. Based on your output from all three chains, what can you say about the behavior of the chain?

c. Use the Gelman-Rubin method to investigate convergence of the chain, and run the chain until the chain has converged approximately to the target distribution according to $\bar{R} < 1.2$.

Problem 2

In 1986, the space shuttle Challenger exploded during takeoff, killing the seven astronauts aboard. The explosion was the result of an O-ring failure, a splitting of a ring of rubber that seals the parts of the ship together. The accident was believed to have been caused by the unusually cold weather (31° F or 0° C) at the time of launch, as there is reason to believe that the O-ring failure probabilities increase as temperature decreases. Data on previous space shuttle launches and O-ring failures is given in the dataset challenger provided with the "mcsm" package of R. The first column corresponds to the failure indicators y_i and the second column to the corresponding temperature x_i, $(1 \leq i \leq 24)$.

(a) The goal is to obtain MLEs for β_0 and β_1 in the following logistic regression model

$$log\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 x$$

where p is the probability that at least one O-ring is damaged and x is the temperature. Create computer programs using Newton-Raphson algorithm with the 'Iterative Reweighted Least Squares' algorithm to find MLEs $\hat{\beta}_0, \hat{\beta}_1$

(b) We are also interested in predicting O-ring failure. *Challenger* was launched at 31° F. What is the predicted probability of O-ring damage at 31° F? How many O-ring failures should be expected at 31° F? What can you conclude?

Problem 3

Part 1

The function g is given by

$$g(x, y) = 4xy + (x + y^2)^2.$$

The goal is to find the minimum of g.

a. Minimize g using Newton's method

b. Minimize g using the steepest descent method with golden section search method as line search.

Part 2

The function

$$g(x) = \frac{\log x}{1+x}$$

was discussed in class. The goal is to maximize g with respect to x.

a. Maximize g using Newton's method

b. Maximize g using the secant method

Problem 4

In this problem, we will use our improved and fixed support vectors/matrices machines codes from Exam 1 (after correcting for the value of "b"). Also, we use a different value for C instead of the default value of $C = \infty$ and different values for "gamma" rather than the default value of 1.5. We will use different types of kernels other than the well known Gaussian kernel. Our goal is to write R or Python codes of these kernels. Then apply these codes with your SVM/SMM codes to find the number of support vectors/matrices. In the support vectors case, we will use the data "pb2.txt" on the class webcourse as well as the classification of observation $\mathbf{z} = (18, 17, 33, 26)'$ from exam 1. For the SMM case, you can use the data from exam 1 or generate 13 matrices 4×5 from class $y = +1$ and 12 matrices 4×5 from class $y = -1$. Make sure to include your codes for generating these matrices. Also, use your SMM code to classify one new matrix of your choice.
Let $\mathbf{x}$ and $\mathbf{y}$ be $p-$ dimensional vectors

(a) The first kernel to be coded is the Marr wavelet kernel also known as the Mexican Hat wavelet kernel defined as:

$$K(\mathbf{x}, \mathbf{y}) = \prod_{i=1}^{p} \left(1 - \left(\frac{x_i - y_i}{\sigma}\right)^2\right) \exp\left(-\frac{1}{2}\left(\frac{x_i - y_i}{\sigma}\right)^2\right)$$

(b) The second kernel to be coded is the Morlet wavelet kernel defined as:

$$K(\mathbf{x}, \mathbf{y}) = \prod_{i=1}^{p} \left(\cos\left(1.75 \times \frac{x_i - y_i}{\sigma}\right)\right) \exp\left(-\frac{1}{2}\left(\frac{x_i - y_i}{\sigma}\right)^2\right)$$

(c) The third kernel to be coded is the Morlet-RBF wavelet kernel. Let K_1 be the Morlet kernel as defined in (b) and K_2 representing our traditional Gaussian kernel. The Morlet-RBF kernel is:

$$K(\mathbf{x}, \mathbf{y}) = \exp\left\{-\gamma\left[2 - 2K_1\left(\mathbf{x}, \mathbf{y}\right)\right]\right\},$$

where γ is the width parameter of the Gaussian kernel.

(d) This next kernel is to be adapted to matrix. Let $\mathbf{X}$ and $\mathbf{Y}$ be two $n \times p$ matrices, the Morlet wavelet kernel for SMM is defined as:

$$K(\mathbf{X}, \mathbf{Y}) = \prod_{i=1}^{n}\prod_{j=1}^{p}\left(\cos\left(1.75 \times \frac{x_{ij} - y_{ij}}{\sigma}\right)\right)\exp\left(-\frac{1}{2}\left(\frac{x_{ij} - y_{ij}}{\sigma}\right)^2\right)$$