APMA E4990.002: FINAL QUIZ
DUE TUESDAY, DECEMBER 13 AT 11:59PM

Please justify your answers, proving the statements you make. You are allowed to refer to results shown in lecture (or that are in the textbook) as long as you state them precisely, meaning that you should say exactly which hypothesis are needed in the result you use.

This quiz is open book, but you are not allowed to share or discuss this quiz with any person. If you have any questions or find errors please email me at vak2116@columbia.edu.

If you need to impose extra conditions on a problem to make it easier (or consider specific cases of the question, like taking n to be 2, e.g.), state explicitly that you have done so. Solutions where extra conditions were assumed, or where only special cases where treated, will also be graded (probably scored as a partial answer).

Note that it may take some time to upload the quiz on Gradescope. Please plan accordingly. I suggest uploading your quiz at least 30 minutes before the deadline so you will have ample time to select the correct pages for your submission. Make sure your answers to each problem are clearly stated in the submitted PDF.

(1) This problem considers soft-margin SVM introduced in lecture.
  (a) Show that the following unconstrained version of the soft-margin SVM
$$\text{minimize}_{w,b}\|w\|_2^2 + \frac{C}{n}\sum_{i=1}^{n}(1 - y_i(w^T v_i - b))_+$$
  is equivalent to the constrained version derived on in the lecture.
  (b) Show that the unconstrained version of the soft-margin SVM with the kernel ϕ
$$J(w) = \frac{c}{n}\sum_{i=1}^{n}(1 - y_i(\phi(x_i)^T w))_+ + \|w\|_2^2$$
  is equivalent to the following kernelized version (up to a constant prefactor)
$$J(\alpha) = \frac{1}{n}\sum_{i=1}^{n}(1 - y_i(K\alpha)_i)_+ + \frac{\lambda}{2}\alpha^T K\alpha$$
  where $w = \sum_i \alpha_i\phi(x_i)$, and determine the subgradient of this objective.
  (c) Write down stochastic (sub)gradient descent updates for kernelized soft-margin SVM using the subgradient determined in the previous question.

(2) In this exercise you will use the code in *final_quiz.zip*. Complete the functions *RBF_kernel* and *sgd_for_soft_svm* in *final_quiz.ipynb* using your results from the previous problem. The specifications of those functions are described in the comments. Include the output plots from the program in your submission.

(3) Consider the two-armed bandit problem with 0/1 rewards where the safe arm is a Bernoulli distribution with mean $(1+\epsilon)/2$ and the risky arm is a Bernoulli distribution with mean $(1-\epsilon)/2$, both distributed independently of each other and the history.

Date: December 6, 2022.

(a) Show that the maximum likelihood estimator of the safe arm given the revealed rewards is determined by the sign of $2G_1 - 2G_2 + s_2 - s_1$ (positive sign corresponds to arm 1 and negative sign corresponds to arm 2) where G_1 and G_2 are the cumulative revealed rewards of arm 1 and arm 2 respectively, and s_1 and s_2 are the total number of times arm 1 and arm 2 respectively were previously chosen by the player.
[Hint: by the independence assumptions you can write down the probability p_i of observing G_1 ones and $s_1 - G_1$ zeros from arm 1 and G_2 one and $s_2 - G_2$ zeros from arm 2 assuming arm i is safe as the PDFs of a binomial random variable; then consider the ratio p_1/p_2.]
Note the policy that chooses the arm according to this MLE entails no exploration (greedy policy).

(b) Denote by a_1 and a_2 the Bernoulli random variables distributed according to the distributions of arm 1 and 2 respectively. Given that in this problem $a_1 = 1 - a_2$, can you suggest a simple procedure for converting a sample of reward from arm 1 into a sample of reward from arm 2? Can you (informally) argue that therefore no exploration is not needed in this simplified problem?