MATH 36 FINAL EXAM
DUE 11/21 AT 11AM

Instructions: This is a take home, open notes, open book exam. It is due by 11am on Tuesday, 11/21, and may be submitted electronically or in person. The exam is separated into four sections, each of which requires you to perform a mathematical analysis and construct an argument proposing and defending your methods, particularly with respect to the data provided. Your writing should be focused on making a coherent and well–supported argument. Exposition will be graded on clarity, conciseness, and relevance, as well as the correctness of your mathematical content.

The .zip folder on the assignments page of the website contains a copy of these directions as well as supporting papers and data for the questions. Unless specifically directed in the problem statement below, You need not read any of the particular papers. However, depending on your choice of analytical techniques, they may provide useful, additional context. If and when you do decide to rely on these (or other) materials in your analyses, please follow standard citation conventions.

The Honor Principle requires that you neither give nor receive aid on this exam. You may, as always, ask me clarifying questions.

1. Age Class Systems

Summary: Write a 1 page technical report proposing a method for modeling the distribution of ages in the US over time and analyzing the consequences of the long term behavior of your model.

Details: In class, we discussed examples of rigid East African age–class systems as well as demographic models of populations. Your goal is to formulate a model for the evolution and stability of age distributions in the US population over time. As an example, a common variant of several of the techniques that we discussed in class is called the Leslie matrix, which is described in detail in the attached papers. This method shares properties and structure with both the difference equation and Markov models that we discussed.

The links in the attached folder are to the US census data, with age distributions broken down by year from 1900 to 2000 and you should use this data to help construct your model parameters. You may use any type of model and you may choose to aggregate the data into convenient time periods, e.g., 5 or 10 year windows.

Your essay should describe your model and justify the decisions you made in the modeling process. Additionally, you should discuss the limiting behavior of your model, describing your prediction for the long term distribution of ages in the US population. You should consider the issue of stability of age distributions in a population and discuss how you might define a measure of stability. How do your model's conclusions align with this concept? Are your conclusions reasonable in light of current demographic trends? The final .pdf in the folder provides some context for models of the aging population.

Date: November 13, 2017.

2. Voting Methods

Summary: Write a 1 page technical report proposing a method for analyzing incomplete voting data and apply your method to a specific dataset.

Details: As we have seen, the data that we are presented with is not always ideal for the methods we have. Thus, constructing appropriate modifications of standard methods is a necessary task. Imagine that a library is holding a pumpkin carving contest and has purchased prizes to award to the first, second, and third place finishers. Ten pumpkins, labeled A, B, C, D, E, F, G, H, I, and J, are entered into the contest and each of 200 library patrons submitted ranked votes for their three favorite pumpkins. The methods that we studied in class assumed that we had access to a complete list of rankings for each voter, so they can't be directly applied to this data.

Your task is to propose a method for determining the top three pumpkins (i.e. which pumpkins should win the prizes) based on the attached data. Your discussion should describe the useful features of your method and why the results that it will return are meaningful. After selecting a method, apply your approach to the voting data in the attached folder to determine the winners. Following this analysis, do you still believe that your method is appropriate for this type of data? If you were a participant in the voting would you feel the method was fair? Is it possible to vote strategically, by disguising your true preferences to achieve a particular desired outcome?

3. Information Diffusion

Summary: Write a 1 page technical report proposing a method for selecting important nodes in a network in order to maximize information diffusion.

Details: A common application of network models is identifying important nodes in communities, in order to understand the spread of information. This is particularly important for public health interventions in developing countries, where given resources may not be sufficient for contacting all of the inhabitants directly. Recent examples of this type of analysis are presented in this folder.

Your task is to propose and defend a method for identifying which subset of 5% of the nodes in a given social network should be given a piece of useful information in order to have the information spread throughout the network (i.e. reach every node) as quickly as possible. You should describe your method and explain its desirable features and properties. Below are some specific questions to help guide your analysis. You need not address all of these questions in your write–up, but they may provide helpful considerations.

a) What model of information flow seems most reasonable to you? Does the type of information matter for selecting an appropriate method? How can you measure this?
b) If the network has a well–defined community structure, does your method guarantee that at least one person in each community is selected? Does this matter?
c) Can you identify places in the network that it is difficult for information to travel to/from? Does your method take these areas in to account?
d) How does your method perform on the specific networks we looked at in class? Do the results seem reasonable to you?
e) What results would your method return on the various null models we considered? How would changing the parameters affect these results?

4. Opinion Dynamics

Summary: Write a 1–2 page technical report discussing opinion propagation, social influence, and fission in the setting of social networks.

Details: Read the paper "Social Positions in Influence Networks," focusing on the discussions in Section 2 and 4 (don't stress over the mathematical details in Section 3). Begin by describing the authors' model of opinion formation and comment on how the structure of the network plays a role in this process. Next, describe the mathematical formulation presented in Section 2, specifically, the social process encoded in equation (2). Explain how the Markov chain matrix, W, affects the long term behavior of the process and interpret it in terms of the social model.

Next, use this model to explore opinion propagation in karate club network. One of Zachary's hypotheses in the original paper was that the fission that occurred within this group was aided by unequal information flow. Use the attached MATLAB function opinionprop.m to simulate opinion propagation in a network. The syntax for this function is:

>> `output=opinionprop(y1,W,`α`,staticinds,n);`

where y1 is a column vector of initial opinions, W is the matrix in equation (2), α is the constant in equation (2), `staticinds` is a list of indices associated with nodes whose opinions do not change (input [] here if there are no such nodes), and n is the number of iterations you wish to run. The variable output is an $k \times n$ matrix (where k is the number of nodes in the network) where column i gives the group's opinion at time i. It may be helpful to visualize these matrices using `imagesc(output)`.

Using the karate club network as a basis, form the associated matrix W. Also form an initial opinion vector that reflects the opinion dynamics before the fission of the group. Justify your choices and explain their expected consequences. You may wish to experiment with the dolphin network as well, in order to help develop your intuition.

Experiment with different values of the parameter α and initial vectors y_i in order to to explore the opinion dynamics and interpret your results in terms of the model from Zachary's paper. When $\alpha = 1$, who are the last people to change their minds? What is the interpretation of this behavior? Based on this result, is it likely that one subgroup of the network would form a coherent network on its own? How does this affect the fission process?

Finally, assume that the two leaders, nodes 1 and 34, will not change their opinions. Using statinds = [1; 34], rerun the opinion dynamics process. What are the results and how do you interpret them in terms of the questions concerning the fission of the group? How reasonable is this approach for modeling interactions between people with diverse political opinions?

Department of Mathematics, Dartmouth College
E-mail address: ddeford@math.dartmouth.edu