EndTerm Test

Monday, April 20 | 80 minutes | 100 points

NAME: FINAL SCORE:

The endterm test is open notes/book. But do not use the full internet. These notes are restricted to those directly linked off of the course webpage; especially the M4D book.

Show your work and you will be eligible for partial credit on wrong answers (or just get them all right).

Take Home Final Rules

Any violation of these rules will result in an Academic Misconduct filing.

1. Spend at most 80 minutes on the test. You should attempt to make these consecutive minutes; but if life happens and you get interrupted (e.g., a child wakes up and needs your attention), that is ok to "pause the clock."
2. Do not communicate with anyone else (especially other students in the course) about the test. The only exception are emails to profs/TAs about how to interpret questions. You may "pause the clock" while waiting for a reply; all general clarifications will be posted in a canvas thread.
3. Do not use resources beyond those directly linked off of the course website. Nothing should be required beyond the M4D book.
4. If you are aware of any efforts to violate these rules (including others communicating with you), it is your duty to report those violations to the professor. A failure to communicate to the professor a violation of these rules is itself a first class Academic Misconduct.

By signing below,

I pledge that I did not violate any of the above take home final rules.

sign here

1 Regression (30 points)

We studied different formulations of regression, and they each have different advantages. Consider these traits:

(a) : provides unbiased estimate

(b) : simple closed-form optimal solution

(c) : if parameter chosen correctly, can generalize better than least mean squared solution

(d) : provides sparse solution (many non-zero coefficients)

(e) : can use explanatory variables with different units

Assign each trait to one of the following models (some traits can fit more than one model). Each model must be assigned at least one trait.

Ordinary Least Squares $\alpha^* = \arg\min_\alpha \|X\alpha - y\|$

Ridge Regression $\alpha^\circ = \arg\min_\alpha \|X\alpha - y\|^2 + \gamma\|\alpha\|_2^2$

Lasso $\alpha^\diamond = \arg\min_\alpha \|X\alpha - y\| + \gamma\|\alpha\|_1.$

2 Dimensionality Reduction (40 points)

Consider a single data point $a = (10, 0, -2, 8, -4, 0)$ in 6 dimensions.

Let $u_1 = (\frac{1}{2}, 0, \frac{1}{2}, -\frac{1}{2}, \frac{1}{2}, 0)$ and $u_2 = (-\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, 0, 0, \frac{1}{2})$ be basis vectors.

A: (10 points) Which of a, u_1 and u_2 are unit vectors? Explain why.

B: (5 points) Are u_1 and u_2 orthogonal? Explain why.

C: (15 points) Show how to project a to a two-dimensional space spanned by u_1 and u_2. That is, write the two coordinates of a in this new two-dimensional space (show your work).

D: (5 points) Calculate the L_2 norm of the new 2-dimensional vector (show your work).

E: (5 points) Can the L_2 norm of vector ever be larger after a projection using two orthogonal unit vectors? If so, provide an example.

3 Graphs (30 points)

Consider the following matrix P (I have labeled columns/rows with nodes A, B, C, D, E)

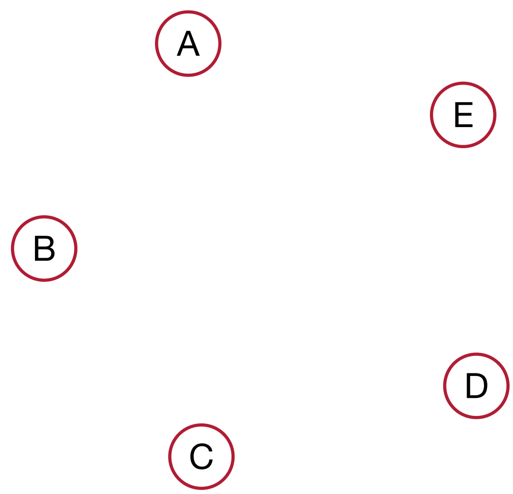
$$P = \begin{pmatrix} 0 & 0 & 0 & 0 & 1/2 \\ 1/3 & 0 & 1/2 & 0 & 0 \\ 1/3 & 1 & 0 & 1/2 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 1/3 & 0 & 1/2 & 1/2 & 1/2 \end{pmatrix}$$

and with labeled rows/columns

	A	B	C	D	E
A	0	0	0	0	1/2
B	1/3	0	1/2	0	0
C	1/3	1	0	1/2	0
D	0	0	0	0	0
E	1/3	0	1/2	1/2	1/2

A: (5 points) Is P a valid *probability transition matrix*? (If not, why not.)

B: (15 points) Draw the associated directed graph using the labeled nodes below; in particular (1) draw edges, and (2) label them with transition probabilities.



C: (5 points) Is the associated Markov Chain ergodic? And if not, identify a property it violates.

D: (5 points) If I want to calculate the pageRank vector $r \in \mathbb{R}^5$ associated with the above graph, I can (step 1) take the top eigenvector of a matrix R, and then (step 2) L_1 normalize it. How should I derive matrix R from P? (And why not just take the top eigenvector of P?)