CS 3220: Prelim 1

Instructor: Anil Damle

Name and netID:

Policies and details

This document constitutes the first preliminary exam for 3220. It contains 3 questions and you have one hour and thirty minutes to complete the exam. Please neatly write your solutions directly onto the exam booklet (scratch paper is available if necessary). All of the work on this exam must be your own and this exam is governed by the Cornell Academic Integrity Code. It is a violation of the Academic Integrity Code to look at any exam other than your own, utilize any references, or otherwise give or receive any unauthorized assistance. Please avoid discussing this exam with any students scheduled to take the exam at an alternative time.

This exam is constructed to assess your grasp of the material in the class over several levels of difficulty—some of the questions may be difficult. Please do your best to answer all of the questions and provide partial solutions if you have them, partial credit will be awarded as appropriate. If you have any questions during the exam please ask, I will not provide hints but if something is unclear I am happy to clarify.

Question 1

For the entirety of this problem you should assume all matrices involved are symmetric. In class we talked about the power method, an algorithm (that under mild assumptions) allowed us to compute the largest magnitude eigenvalue of A. For a matrix A, starting guess for the desired eigenvector $v^{(0)}$, desired accuracy $\epsilon > 0$, and maximal number of iterations K the power method is given in Algorithm 1

Algorithm 1 The power method

input: A, $v^{(0)}$, and ϵ
output: Eigenvector estimate v and eigenvalue estimate λ

$\lambda^{(0)} = (v^{(0)})^T A v^{(0)}$
for $k = 1, 2, \ldots, K$ **do**
  $v^{(k)} = Av^{(k-1)}$
  $v^{(k)} = v^{(k)}/\|v^{(k)}\|_2$
  $\lambda^{(k)} = (v^{(k)})^T A v^{(k)}$
  if $\|Av^{(k)} - \lambda^{(k)} v^{(k)}\|_2 \leq \epsilon$ **then**
    return: $v \equiv v^{(k)}$ and $\lambda \equiv \lambda^{(k)}$
  end if
end for
return: $v \equiv v^{(k)}$ and $\lambda \equiv \lambda^{(k)}$

Now, lets say that we have a symmetric non-singular matrix A and we would like to compute the eigenvalue smallest in magnitude. Denote the eigenvalue of A as $\lambda_1, \ldots, \lambda_n$ with associated eigenvectors $v_1, \ldots, v_n$ and assume they satisfy $|\lambda_n| < |\lambda_{n-1}| \leq \cdots \leq |\lambda_1|$. Address the following questions:

(a) In terms of the eigenvalues of A, what is the largest magnitude eigenvalue of A^{-1}? (Recall that we have assumed A is non-singular.)

(b) In light of the previous question, can you use the power method (perhaps applied to a matrix related to A rather than A itself) to find v_n and λ_n? Rewrite Algorithm 1 to accomplish this goal.

(c) We can break down the computational cost of the power method into two parts—fixed costs and any additional cost per iteration. For example, in Algorithm 1 there are no fixed costs and the cost per iteration is $\mathcal{O}(n^2)$ (as a consequence of the matrix vector multiplication). What is the cost (broken down in this manner) for computing λ_n and v_n? Clearly explain your answer. If your algorithm as written costs more than $\mathcal{O}(n^2)$ per iteration can you rewrite the algorithm so the per iteration cost is $\mathcal{O}(n^2)$? If so, do this and explain what the fixed cost is to make this happen.

(d) For Algorithm 1 we argued that the convergence rate of $v^{(k)}$ to v_1 is $\mathcal{O}\left(|\lambda_2/\lambda_1|^k\right)$ and the convergence rate of $\lambda^{(k)}$ to λ_1 is $\mathcal{O}\left(|\lambda_2/\lambda_1|^{2k}\right)$. For your method in the prior part, what are the convergence rates of $v^{(k)}$ to v_n and $\lambda^{(k)}$ to λ_n? Clearly justify your answers.

QUESTION 2

While discussed the definition of an induced matrix norm in general, we only really talked about it in detail when the vector norm we used was $\|\cdot\|_2$. Recall the definition of an induced matrix norm given any vector norm $\|\cdot\|$ is

$$\|A\| = \max_{\|x\|=1} \|Ax\|.$$

We will now consider $\|A\|_\infty$ and see how it relates to the entries of A. (Turns out that this is not a hard matrix norm to compute given the entries of A.)

(a) Prove that for any x with $\|x\|_\infty \leq 1$

$$\|Ax\|_\infty \leq \max_{j=1,\ldots,n} \sum_{i=1}^{n} |A(j,i)|.$$

(b) Prove that for any $j \in \{1, \ldots, n\}$ there exists a vector x with $\|x\|_\infty = 1$ such that

$$(Ax)_j = \sum_{i=1}^{n} |A(j,i)|$$

where $(Ax)_j$ is entry j in the vector Ax.

(c) Using the following two results prove that

$$\|Ax\|_\infty = \max_{j=1,\ldots,n} \sum_{i=1}^{n} |A(j,i)|.$$

Question 3

We have spent a fair bit of time discussing the SVD, nevertheless there remain additional interpretations to explore. As we discussed in class a square matrix $A \in \mathbb{R}^{n \times n}$ is singular if at least one of its singular values is zero. It turns out that the smallest singular value of a matrix A actually tells us how close we are to a singular matrix, in other words what the smallest perturbation to A is that can make it singular—this seems useful to know.

(a) Given two vectors x and y with $\|x\|_2 = 1$ and $\|y\|_2 = 1$ prove that

$$\|xy^T\|_2 = 1.$$

(b) Given a non-singular matrix A with SVD $A = U\Sigma V^T$ show that there is a rank-one perturbation E such that $A + E$ is singular and $\|E\|_2 = \sigma_n$.

(c) Prove that given any matrix E with $\|E\|_2 < \sigma_n$ $A + E$ is non-singular. Here, σ_n is the smallest singular value of A. (Hint: we talked about how the largest singular value of a matrix bounds how big A can make a vector with unit norm; can you cook up a lower bound on how small A can make a vector with unit norm in terms of singular values?)