EE 516: Mathematical Foundations of Machine Learning **Winter 2023**

Homework 2

Due: January 27, 2023, 11:59PM PT

Student Name: *Instructor Name: John Lipor*

Problem 1 (5 pts)

Describe how the eigenvalues and eigenvectors of $B = A - 10I$ are related to the eigenvalues and eigenvectors of A.

Problem 2 (6 pts)

Let $v_1, v_2, \ldots, v_n$ be orthonormal vectors in $\mathbb{R}^n$. Show that $Av_1, Av_2, \ldots, Av_n$ are also orthonormal if and only if $A \in \mathbb{R}^{n \times n}$ is an orthogonal matrix.

Note: Proving a statement of the form "A if and only if B" is equivalent to showing that B implies A (if) *and* A implies B (only if).

Problem 3 (5 pts)

Let $A \in \mathbb{R}^{n \times n} = U\Sigma V^T$ with rank $r = n$. Express A^{-1} in terms of the SVD of A.

Problem 4 (8 pts)

Let $B = T^{-1}AT$ for an invertible matrix T. Determine the relationship between the eigenvalues and eigenvectors of B and those of A. Explain. The matrices A and B, when thus related, are said to be "similar".

Note: Be sure to follow the convention that eigenvectors are normalized to have unit norm.

Problem 5 (5 pts)

Let $X \in \mathbb{R}^{m \times n}$. Using the SVD only, show that if $X^T X = 0$, then $X = 0$.

Problem 6 (8 pts)

The Frobenius norm of $A \in \mathbb{R}^{m \times n}$ is defined as $\|A\|_F = \sqrt{\sum_{i,j} |a_{ij}|^2}$. Express the Frobenius norm of A in terms of its singular values. Hint: How can $\|A\|_F$ be computed from the matrix $A^T A$?

Problem 7 (6 pts, 2 pts, 2 pts)

Let $A \in \mathbb{R}^{m\times n}$ with singular values $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_{\min(m,n)}$. Recall that the "operator norm" of a matrix $\|A\|_2 = \sigma_1$. There are many other interesting and useful matrix norms, and we will continue to see these throughout the course. Bounding these norms in terms of each other is an important problem.

(a) Show that
$$\sigma_1 \leq \|A\|_F \leq \sqrt{\min(m,n)}\sigma_1.$$

(b) If the Frobenius norm is small, what does this tell us about the operator norm?

(c) If the operator norm is small, what does this tell us about the Frobenius norm?

Problem 8 (6 pts, 2 pts, 2 pts)

This problem develops a tool that will be used in later homework assignments for an application called *photometric stereo*. To approximate the derivatives of a function $f(x)$ that is sampled on a grid $x_1, \ldots, x_n$ where $x_{i+1} = x_i + \delta$, a typical *finite difference* approach is:

$$\left.\frac{\partial f(x)}{\partial x}\right|_{x=x_i} \approx \frac{f(x_{i+1}) - f(x_i)}{\delta}.$$

When the sample spacing is $\delta = 1$, this approximation simplifies to

$$f'(x_i) := \left.\frac{\partial f(x)}{\partial x}\right|_{x=x_i} \approx f(x_{i+1}) - f(x_i).$$

We can express this relation for all x_i samples via the matrix-vector product

$$\begin{bmatrix} f'(x_1) \\ f'(x_2) \\ \vdots \\ f'(x_n) \end{bmatrix} \approx D_n \begin{bmatrix} f(x_1) \\ f(x_2) \\ \vdots \\ f(x_n) \end{bmatrix},$$

where D_n is the so-called *first-difference* matrix

$$D_n = \begin{bmatrix} -1 & 1 & & & \\ & -1 & 1 & & \\ & & \ddots & \ddots & \\ & & & -1 & 1 \\ 1 & & & & -1 \end{bmatrix}.$$

Here we choose to set $D_n(n, 1) = 1$, which corresponds to the (perhaps unexpected) approximation $f'(x_n) \approx f(x_1) - f(x_n)$. This choice is called a periodic boundary condition because essentially we are assuming that the domain wraps around. We make this assumption because the resulting D_n is a circulant matrix, so its eigenvectors can be computed in closed form!

The goal of this problem is for you to derive and implement the analog of D_n for two-dimensional differentiation. Let $f(x, y)$ be a function of two variables. We can approximate its partial derivatives using finite differences as follows:

$$\frac{\partial f(x,y)}{\partial x} \approx \frac{f(x+1,y) - f(x,y)}{(x+1) - x} = f(x+1,y) - f(x,y) \tag{1}$$

$$\frac{\partial f(x,y)}{\partial y} \approx \frac{f(x,y+1) - f(x,y)}{(y+1) - y} = f(x,y+1) - f(x,y) \tag{2}$$

To simplify notation, define the $m \times n$ matrices `FXY`, `DFDX`, and `DFDY` having elements as follows:

$$\texttt{FXY}(i,j) = f(i,j)$$

$$\texttt{DFDX}(i,j) = \frac{\partial f(i,j)}{\partial x}$$

$$\texttt{DFDY}(i,j) = \frac{\partial f(i,j)}{\partial y}.$$

The x coordinate is along the column of `FXY` and the y coordinate is along the row of `FXY`, so we can think in terms of $\texttt{FXY}(x, y)$. Define the corresponding $mn \times 1$ vectors `fxy`, `dfdx`, and `dfdy` to be the vectorized versions obtianed using the vec(·) operation (HW1, Problem 9). With this notation, we can succinctly express equations (1) and (2) as

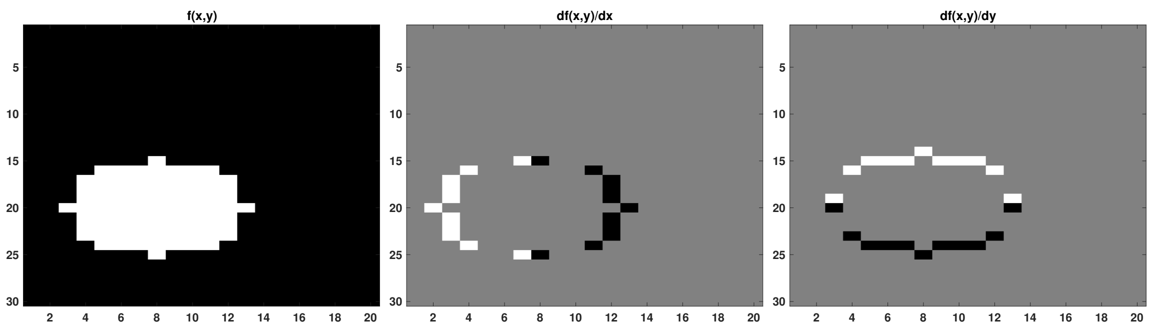
$$\begin{bmatrix} \texttt{dfdx} \\ \texttt{dfdy} \end{bmatrix} = A\ \texttt{fxy},$$

where A is a $2mn \times mn$ matrix.

(a) Find an expression for A in terms of the first difference matrices D_n and D_m, appropriately sized identity matrices, and appropriate Kronecker products of these matrices. Use periodic boundary conditions.

(b) Once you have determined A, write the function `first_diffs_2d_matrix` that takes as input the dimensions m and n of `FXY` and returns the appropriate A matrix, stored in **sparse** format, since the resulting matrix will be prohibively large to store as a full double-precision matrix. PYTHON users should make use of the `scipy.sparse` library.

(c) In this problem your function is designed to compute finite-difference approximations to derivatives along x and y. If you create an $m \times n$ array `X` that is a picture of a disk, then the finite derivatives will be mostly zero except near the edges of the disk. The code for generating such a disk is on the course website. Test your code using your function and turn in plots of the resulting images, which should match those below. Note that the PYTHON `reshape` requires setting `order='F'`.



Problem 9 Reflection (5 pts)

Read Chapter 1 of *Solving Mathematical Problems* by Terence Tao, available here, and answer the following questions.

- Among the problem solving strategies discussed, have you used any (whether implicitly or explicitly) in the course so far?
- Among the problem solving strategies discussed, which do you think would be helpful to use going forward?
- Comment on whether you think this article will be helpful to further problem solving in the course and beyond.