Homework set 2 — MATH 393C — Spring 2019

Due on Thursday March 7. Please hand in solutions to two problems of your choice in the set $\{3, 4, 5, 6\}$, as well as to problems 7 and 8.

Problem 1: Suppose that $\mathbf{A}$ is a real symmetric $n \times n$ matrix. Let $\{\mathbf{v}_j\}_{j=1}^n$ denote an orthonormal set of eigenvectors so that $\mathbf{A}\mathbf{v}_j = \lambda_j \mathbf{v}_j$ for some numbers λ_j. Let us use an ordering where $|\lambda_1| \geq |\lambda_2| \geq \dots \geq |\lambda_n|$. Define a sequence of vectors $\mathbf{x}_p = \mathbf{A}^p\mathbf{g}$, where $\mathbf{g}$ is an $n \times 1$ random vector whose entries are drawn independently from a standard Gaussian distribution.

(a) Set $\beta = |\lambda_2|/|\lambda_1|$ and $\mathbf{y}_p = (1/\|\mathbf{x}_p\|)\mathbf{x}_p$. Assume $\lambda_1 = 1$ and $\beta < 1$. Prove that as $p \to \infty$, the vectors $\{\mathbf{y}_p\}$ converge either to $\mathbf{v}_1$ or $-\mathbf{v}_1$.
(b) What is the speed of convergence of $\{\mathbf{y}_p\}$?
(c) Assume again that $\beta < 1$, but now drop the assumption that $\lambda_1 = 1$. Prove that your answers in (a) and (b) are still correct, with the exception that if $\lambda_1 < 0$, then it is the vector $(-1)^p\mathbf{y}_p$ that converges instead.

Problem 2: Let $\mathbf{A} \in \mathbb{R}^{m\times n}$, J be an $n \times 1$ permutation vector, and $J_s = J(1 : k)$ for some $k \in \mathbb{N}$. Now suppose that

$$\underset{m\times n}{\mathbf{A}} = \underset{m\times k}{\mathbf{E}}\ \underset{k\times n}{\mathbf{F}}, \tag{1}$$

and suppose that for some matrix $\mathbf{T}$ of size $k \times (n - k)$, it holds that

$$\mathbf{F}(:, J) = \mathbf{F}(:, J_s)\begin{bmatrix}\mathbf{I}_k & \mathbf{T}\end{bmatrix},$$

where $\mathbf{I}_k$ is the $k \times k$ identity matrix. Prove that

$$\mathbf{A}(:, J) = \mathbf{A}(:, J_s)\begin{bmatrix}\mathbf{I}_k & \mathbf{T}\end{bmatrix}.$$

Problem 3: Suppose that $\mathbf{A}$ is an $m \times n$ matrix of approximate rank k, and that we have identified two index sets I_s and J_s such that the matrices

$$\mathbf{C} = \mathbf{A}(:, J_s), \quad \mathbf{R} = \mathbf{A}(I_s, :)$$

hold k columns/rows that span the column/row space of $\mathbf{A}$. Then

$$\mathbf{A} \approx \mathbf{C}\mathbf{C}^\dagger\mathbf{A}\mathbf{R}^\dagger\mathbf{R},$$

and the optimal choice for the "U" factor in the CUR decomposition is

$$\mathbf{U} = \mathbf{C}^\dagger\mathbf{A}\mathbf{R}^\dagger.$$

Set $\mathbf{X} = \mathbf{C}\mathbf{C}^\dagger$.

(a) Suppose that $\mathbf{C}$ has the SVD $\mathbf{C} = \mathbf{W}\mathbf{D}\mathbf{V}^*$. Prove that $\mathbf{X} = \mathbf{W}\mathbf{W}^*$.
(b) Suppose that $\mathbf{C}$ has the QR factorization $\mathbf{C}\mathbf{P} = \mathbf{Q}\mathbf{S}$. Prove that $\mathbf{X} = \mathbf{Q}\mathbf{Q}^*$.
(c) Prove that $\mathbf{X}$ is the orthogonal projection onto Col($\mathbf{C}$).
(d) Suppose that $\mathbf{A}$ has precisely rank k and that $\mathbf{C}$ and $\mathbf{R}$ are both of rank k. Prove that then $\mathbf{C}^\dagger\mathbf{A}\mathbf{R}^\dagger = (\mathbf{A}(I_s, J_s))^{-1}$.

Problem 4: Let $\mathbf{A} \in \mathbb{R}^{m\times n}$ have rank exactly k. In this problem, we will prove that $\mathbf{A}$ admits a factorization $\mathbf{A} = \mathbf{A}(:, J_s)\mathbf{Z}$, where $\mathbf{A}(:, J_s) \in \mathbb{R}^{m\times k}$ and $\mathbf{Z} \in \mathbb{R}^{k\times n}$ such that $\mathbf{Z}(:, J_s) = \mathbf{I}_k$ and $\max_{i,j} |\mathbf{Z}(i, j)| \leq 1$.

(a) **case 1:** $m = k$.

  (a) Pick a permutation vector J_s such that $|\det(\mathbf{A}(:, J_s))|$ is maximized, and let J_r denote the remaining indices so that $[J_s \quad J_r]$ is some permutation of the vector $[1\ 2 \quad \cdots \quad n]$. Then we have that

$$\mathbf{A}(:, [J_s \quad J_r]) = [\mathbf{A}(:, J_s) \quad \mathbf{A}(:, J_r)]$$

  can be written as $\mathbf{AP}$ for some permutation matrix $\mathbf{P}$. Find an interpolative decomposition $\mathbf{A} = \mathbf{CZ}$ of $\mathbf{A}$, where the columns of $\mathbf{C}$ are some of the columns of $\mathbf{A}$. $\mathbf{C}$ and $\mathbf{Z}$ should be in terms of $\mathbf{A}(:, J_s), \mathbf{A}(:, J_r), \mathbf{P}$, and the identity matrix $\mathbf{I}$.

  (b) Consider the matrix $\mathbf{T} = \mathbf{A}(:, J_s)^{-1}\mathbf{A}(:, J_r)$. If we can show that

(2) $$\max_{i,j} |\mathbf{T}(i, j)| \leq 1,$$

  then we will be done with the case $m = k$ (why?). Find a way to show (2) by applying Cramer's Rule to our definition of $\mathbf{T}$.

  Cramer's Rule: Consider the linear system $\mathbf{Ax} = \mathbf{b}$. The i-th entry of the solution $\mathbf{x}$ is given by

$$x_i = \frac{\det(\mathbf{A}_i)}{\det(\mathbf{A})},$$

  where $\mathbf{A}_i$ is matrix formed by replacing the i-th column of $\mathbf{A}$ with $\mathbf{b}$.

(b) **case 2:** $m \geq k$.

Then $\mathbf{A}$ admits a factorization $\mathbf{A} = \mathbf{EF}$, where $\mathbf{E}$ is $m \times k$ and $\mathbf{F}$ is $k \times n$. Apply case 1 to $\mathbf{F}$ to show the result for this case (something we proved in a previous problem may help, too...).

Problem 5: The purpose of this exercise is to prove the equivalence of subspace iteration and the "power" version of the RSVD. Suppose you are given an $m \times n$ matrix $\mathbf{A}$ of rank at least k, and an $n \times k$ matrix $\mathbf{G}$ for which $\mathbf{AG}$ is of full rank. Then set

$$\mathbf{Y} = (\mathbf{AA}^*)^q\mathbf{AG}$$

for some positive integer q. Also define $\mathbf{Z}$ as the output of the iteration

$\mathbf{Z} \leftarrow \text{orth}(\mathbf{AG})$
for $i = 1 : q$
  $\mathbf{Z} \leftarrow \text{orth}(\mathbf{A}^*\mathbf{Z})$
  $\mathbf{Z} \leftarrow \text{orth}(\mathbf{AZ})$
end

The output of orth($\mathbf{A}$) for a matrix $\mathbf{A}$ is a matrix $\mathbf{Z}$ with orthonormal columns such that ran($\mathbf{Z}$) = ran($\mathbf{A}$). Show that ran($\mathbf{Y}$) = ran($\mathbf{Z}$).

Problem 6: Let $\mathbf{R}$ be an $m \times n$ random matrix. Assume the entries of $\mathbf{R}$ are independent, and $\mathbb{E}[\mathbf{R}_{ij}] = 0$ and $\text{Var}(\mathbf{R}_{ij}) = 1\ \forall i, j$. Let $\mathbf{x} \in \mathbb{R}^n$. Show that $\mathbb{E}[\|\mathbf{Rx}\|^2] = m\|\mathbf{x}\|^2$.

Problem 7:

(a) Let $\mathbf{B} \in \mathbb{R}^{n\times n}$ be tridiagonal, and let $\mathbf{x} \in \mathbb{R}^n$ be a given vector. Write a function `y = solve_tridiag(B,x)` that computes the solution to $\mathbf{B}\mathbf{y} = \mathbf{x}$.

(b) Let $\mathbf{B} \in \mathbb{R}^{n\times n}$ be tridiagonal. Write a function `[L,U] = LU_tridiag(B)` that computes the LU factorization of $\mathbf{B}$.

(c) Let $\mathbf{A} \in \mathbb{R}^{n\times n}$ be semi-separable, and let $\mathbf{y} \in \mathbb{R}^n$ be a given vector. Write a function `x = solve_SS(A,y)` that computes the solution to $\mathbf{A}\mathbf{x} = \mathbf{y}$.

(d) Let $\mathbf{A} \in \mathbb{R}^{n\times n}$ be semi-separable. Write a function `B = inv_SS(A)` that computes the inverse of $\mathbf{A}$.

(e) Verify that all your functions have linear complexity through numerical experiments. In other words, measure the execution time t_n required to solve the problem, and plot t_n against n. (Or, even better, plot t_n/n against n.)

You need to use data-efficient formats to represent the matrices. For instance, you could represent a semiseparable matrix $\mathbf{A}$ using four vectors a, b, c, d such that

$$\mathbf{A}(i,j) = \begin{cases} a(i)b(j), & \text{for } i \le j, \\ c(i)d(j), & \text{for } i \ge j. \end{cases}$$

Ensure that you pick vectors that satisfy $a(i)b(i) = c(i)d(i)$. Or, even better, specify just three vectors and compute the dependent variable on the fly. A tridiagonal matrix, can be specified by giving the three vectors that hold the diagonal and the offdiagonal entries. For instance, the matrix $\mathbf{A} = \begin{bmatrix} f_1 & g_1 & 0 & 0 & \cdots \\ e_1 & f_2 & g_2 & 0 & \cdots \\ 0 & e_2 & f_3 & g_3 & \cdots \\ 0 & 0 & e_3 & f_4 \cdots & \\ \vdots & \vdots & \vdots & \vdots & \end{bmatrix}$ is specified by giving just the vectors $\mathbf{e}$, $\mathbf{f}$, $\mathbf{g}$.

Write the functions as simple `for` loops in your first implementation. Then see if you can accelerate them using more efficient programming techniques. (Using, for instance, built-in routines for sparse operations.)

Various pathological cases may arise that involve division by zero. You are welcome to disregard this, and just assume that no divisions by zero happen.

Problem 8: In this problem, n and k are positive integers such that $k < n$, $\mathbf{A}$ is an $N \times N$ invertible matrix, and $\mathbf{B} = \mathbf{A}^{-1}$. Let us further assume that every diagonal block of $\mathbf{A}$ is invertible.

(a) Suppose that $N = 2n$, and that we can write $\mathbf{A}$ and $\mathbf{B}$ as

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix}, \qquad \text{and} \qquad \mathbf{B} = \begin{bmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} \\ \mathbf{B}_{21} & \mathbf{B}_{22} \end{bmatrix},$$

where each block is of size $n \times n$. Suppose further that $\mathbf{A}_{12}$ and $\mathbf{A}_{21}$ have rank k. What is the highest possible value for the rank of $\mathbf{B}_{12}$?

(b) Suppose that $N = 4n$, and that we can write $\mathbf{A}$ and $\mathbf{B}$ as

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} & \mathbf{A}_{13} & \mathbf{A}_{14} \\ \mathbf{A}_{21} & \mathbf{A}_{22} & \mathbf{A}_{23} & \mathbf{A}_{24} \\ \mathbf{A}_{31} & \mathbf{A}_{32} & \mathbf{A}_{33} & \mathbf{A}_{34} \\ \mathbf{A}_{41} & \mathbf{A}_{42} & \mathbf{A}_{43} & \mathbf{A}_{44} \end{bmatrix}, \qquad \text{and} \qquad \mathbf{B} = \begin{bmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} & \mathbf{B}_{13} & \mathbf{B}_{14} \\ \mathbf{B}_{21} & \mathbf{B}_{22} & \mathbf{B}_{23} & \mathbf{B}_{24} \\ \mathbf{B}_{31} & \mathbf{B}_{32} & \mathbf{B}_{33} & \mathbf{B}_{34} \\ \mathbf{B}_{41} & \mathbf{B}_{42} & \mathbf{B}_{43} & \mathbf{B}_{44} \end{bmatrix},$$

where each block is of size $n \times n$. Suppose further that $\mathbf{A}_{12}$, $\mathbf{A}_{21}$, $\mathbf{A}_{34}$, $\mathbf{A}_{43}$, $\begin{bmatrix} \mathbf{A}_{13} & \mathbf{A}_{14} \\ \mathbf{A}_{23} & \mathbf{A}_{24} \end{bmatrix}$, and $\begin{bmatrix} \mathbf{A}_{31} & \mathbf{A}_{32} \\ \mathbf{A}_{41} & \mathbf{A}_{42} \end{bmatrix}$ all have rank k. What is the highest possible value for the rank of $\mathbf{B}_{12}$?

(c) *[Optional:]* Consider the natural generalization to a matrix consisting of 8×8 blocks. What is the maximal rank of $\mathbf{B}_{12}$? What about a matrix with $2^p \times 2^p$ blocks?

Please motivate your answers rigorously if you can.