Homework 3. Due Wednesday, Sept. 18.

1. (**8 pts**) *Additional reading for this problem: J. Demmel, Applied Numerical Linear Algebra, available online via the UMCP library.*

   The goal of this exercise is to understand how one can compute a QR decomposition using *Householder reflections.*

   (a) Let u be a unit vector in $\mathbb{R}^n$, i.e., $\|u\|_2 = 1$. Let $P = I - 2uu^\top$. This matrix performs reflection with respect to the hyperplane orthogonal to the vector u. Show that $P = P^\top$ and $P^2 = I$.

   (b) Let $x \in \mathbb{R}^n$ be any vector, $x = [x_1, \ldots, x_n]^\top$. Let u be defined as follows:

$$\tilde{u} := \begin{bmatrix} x_1 + \mathsf{sign}(x_1)\|x\|_2 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \equiv x + \mathsf{sign}(x_1)\|x\|_2 e_1, \quad u = \frac{\tilde{u}}{\|\tilde{u}\|_2}, \tag{1}$$

   where $e_1 = [1, 0, \ldots, 0]^\top$. The matrix with the vector u constructed according to (1) will be denoted $\mathsf{House}(x)$:

$$P = I - 2uu^\top \equiv I - 2\frac{\tilde{u}\tilde{u}^\top}{\tilde{u}^\top\tilde{u}} \equiv \mathsf{House}(x).$$

   Calculate Px.

   (c) Let A be an $m \times n$ matrix, $m \geq n$, with columns a_j, $j = 1 \ldots, n$. Let $A_0 = A$. Let $P_1 = \mathsf{House}(a_1)$. Then $A_1 := P_1 A_0$ has the first column with the first entry nonzero and the other entries being zero. Next, we define P_2 as

$$P_2 = \begin{bmatrix} 1 & 0 \\ 0 & \tilde{P}_2 \end{bmatrix}.$$

   where the matrix $\tilde{P}_2 = \mathsf{House}\left(A_1(2:m, 2)\right)$. The notation $A_1(2 : m, 2)$ is Matlab's syntax indicating this is the vector formed by entries 2 through m of the 2nd column on A_1. Then we set $A_2 = P_2 A_1$. And so on.
   This algorithm can be described as follows. Let $A_0 = A$. Then for $j = 1, 2, \ldots, n$ we set

$$P_j = \begin{bmatrix} I_{(j-1)\times(j-1)} & 0 \\ 0 & \tilde{P}_j \end{bmatrix}; \quad \tilde{P}_j = \mathsf{House}\left(A_{j-1}(j : m, j)\right), \quad A_j = P_j A_{j-1}.$$

   Check that the resulting matrix A_n is upper triangular, its entries $(A_n)_{ij}$ are all zeros for $i > j$. Propose an `if`-statement in this algorithm that will guarantee that A_n has positive entries $(A_n)_{jj}$, $1 \leq j \leq n$.

(d) Extract the QR decomposition of A given the matrices P_j, $1 \le j \le n$, and A_n.

2. **(6 pts)** Prove items (1)–(6) of Theorem 3 on page 14 of `LinearAlgebra.pdf`.

3. **(4 pts)** Let A be an $m \times n$ matrix where $m < n$ and rows of A are linearly independent. Then the system of linear equations $Ax = b$ is underdetermined, i.e., infinitely many solutions. Among them, we want to find the one that has the minimum 2-norm. Check that the minimum 2-norm solution is given by

$$x^* = A^\top (AA^\top)^{-1} b.$$

Hint. One way to solve this problem is the following. Check that x^* is a solution to $Ax = b$. Show that is $x^* + y$ is also a solution of $Ax = b$ then $Ay = 0$. Then check that the 2-norm of $x^* + y$ is minimal if $y = 0$.

4. **(3 pts)** Let A be a 3×3 matrix, and let T be its Schur form, i.e., there is a unitary matrix Q (i.e., $Q^*Q = QQ^* = I$ where Q^* denotes the transpose and complex conjugate of Q) such that

$$A = QTQ^*, \quad \text{where} \quad T = \begin{bmatrix} \lambda_1 & t_{12} & t_{13} \\ 0 & \lambda_2 & t_{23} \\ 0 & 0 & \lambda_3 \end{bmatrix}.$$

Assume that λ_j, $j = 1, 2, 3$ are all distinct.

(a) Show that if v is an eigenvector of T then Qv is the eigenvector of A corresponding to the same eigenvalue.

(b) Find eigenvectors of T. *Hint: Check that $v_1 = [1, 0, 0]^\top$. Look for v_2 of the form $v_2 = [a, 1, 0]^\top$, and then for v_3 of the form $v_3 = [b, c, 1]^\top$, where a, b, c are to be expressed via the entries of the matrix T.*

(c) Write out eigenvectors of A in terms of the found eigenvectors of T and the columns of Q: $Q = [q_1, q_2, q_3]$.

5. **(4 pts)** Download the MNIST dataset. For your convenience, I prepared it as the `mnist.mat` file. This file contains 60,000 training images and 10,000 test images of handwritten digits from 0 to 9, and labels for the training and test images. Each image is 20-by-20 pixels because I stripped off paddings with zeros. You can use Matlab or Python.

(a) Convert the set of test images into a matrix A of size $10^4 \times 400$. Compute an SVD of this matrix. Project the data onto the space spanned by the first two right singular vectors v_1 and v_2. Display only points corresponding to digits 0 and 1 in this 2D space and color the points corresponding to 1 and 0 in different colors. Check if 0s and 1s cluster in this 2D space.

(b) Do this task for k = 10, 20, and 50. Compute $A_k = U_k \Sigma_k V_k^\top$, where U_k (V_k) is comprised of the first k left (right) singular vectors, and $\Sigma_k = \mathsf{diag}\{\sigma_1, \dots, \sigma_k\}$. Then take the first four rows of A and A_k, reshape each of these rows back to 20×20 images, and display them. The result for $k = 3$ is shown in Fig. 1.

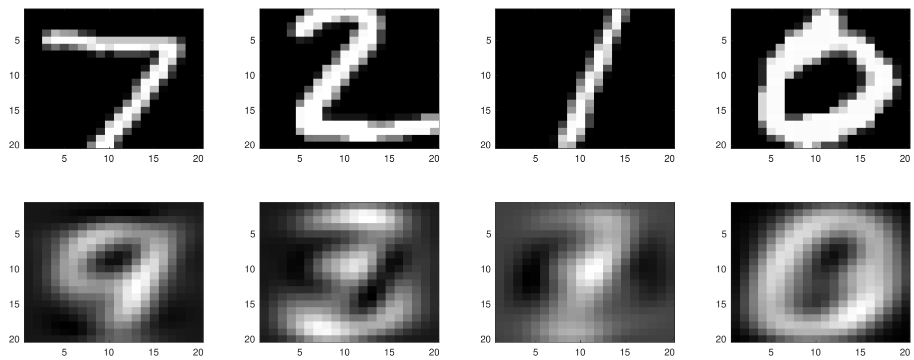

Figure 1: The first four MNIST digits from the test set (the top row) and their best 3D representation (the bottom row).