Assignment 8

Due Monday 11/25/19

Reading:

- Required: Course Notes 4.5-4.6,6.4-6.6,8.1,9.2-9.3
- Recommended: Web link on Four Fundamental Subspaces
- Required: LADR p. 198-200

Problems:

1. (EF: 3.7.3) (5 pts) Find the orthogonal complement (using the standard inner-product) of the range (i.e., column space) of
$$A = \begin{bmatrix} 1 & 2 \\ 3 & -1 \\ 0 & 1 \\ 5 & 7 \end{bmatrix}.$$

2. (EF: 5.3.3) (5 pts) Let $V = \mathbb{F}^n$ be the standard inner-product space and T be the linear operator on V defined by $T\underline{v} = A\underline{v}$, where A is an n by n matrix. Show that the space $\mathcal{R}(A^H)$ is the orthogonal complement of the null space $\mathcal{N}(A)$.

3. (MMA: 4.9.38) (10 pts) Explain why $AA^\dagger$ and $A^\dagger A$ are projection operators. What fundamental subspaces do they project onto? Hint: Given the compact SVD $A = U\Sigma V^H$, where Σ is an invertible $r \times r$ diagonal matrix, the pseudoinverse of A is defined by $A^\dagger \triangleq V\Sigma^{-1}U^H$.

4. (EF: 8.1.1) (5 pts each) In this problem, we use the singular value decomposition $A = U\Sigma V^T$ to compute the pseudoinverse and projection matrices for
$$A = \begin{bmatrix} 1 & 1 & 3 \\ 2 & 3 & 8 \\ 3 & 5 & 13 \\ 1 & -2 & -3 \end{bmatrix}.$$
   (a) Use a computer to compute the SVD of A and give the numerical result. Determine the numerical rank of A.
   (b) Give formulas for the pseudoinverse $A^\dagger$ and projection matrix P_A in terms of U_1,Σ_1, and V_1. Compute these matrices on a computer and give the numerical results.
   (c) Compute the projection of $[1\ 2\ 3\ 4]^T$ onto $\mathcal{R}(A)$ and $\mathcal{R}(A)^\perp$.

5. (EF: 4.6.1) (2.5 pts each) Let V be a finite-dimensional Hilbert space and $C \subseteq V$ be a convex subset. The projection of $\underline{v} \in V$ onto C is defined to be
$$P_C(\underline{v}) \triangleq \arg\min_{\underline{u}\in C} \|\underline{u} - \underline{v}\|.$$

(a) For some $\underline{w} \in V$, let $A = C + \underline{w} = \{\underline{u} + \underline{w} \,|\, \underline{u} \in C\}$. Show that the projections P_A and P_C are related by the translation formula $P_A(\underline{v}) = P_C(\underline{v} - \underline{w}) + \underline{w}$. Hint: Draw a picture.

(b) Let $V = \mathbb{R}^2$ and $C = \{s(1, 2) \,|\, s \in \mathbb{R}\}$ be a subspace of V. Argue that C is convex and give a formula for $P_C(\underline{v})$.

(c) Using the setup of part (b), let $A = C + (1, 0)$ be an affine subspace of V. Show that A is convex and use part (a) to give a formula for $P_A(\underline{v})$. Draw a picture illustrating the projection of $\underline{v} = (0, 0)$ onto A.

(d) Let $C = \{(x, y) \in \mathbb{R}^2 \,|\, x^2 + y^2 \leq 1\}$. Compute $P_C\big((3, 4)\big)$ and draw a picture illustrating this projection.

Practice Problems (do not hand in):

1. (MMA: 4.2.18) Let P be a projection operator on a finite-dimensional normed space. Show that $\|P\| = 1$ for the induced operator norm.

2. (MMA: 3.8.10) Define an inner product between matrices $X, Y \in \mathbb{C}^{l \times n}$ as

$$\langle X, Y \rangle = \operatorname{tr}\left(XY^H\right),$$

where $\operatorname{tr}(\cdot)$ is the sum of the diagonal elements (see section C.3). We want to approximate the matrix Y by the scalar linear combination of matrices $X_1, X_2, \ldots, X_m$, as

$$Y = c_1 X_1 + c_2 X_2 + \cdots + c_m X_m + E.$$

Using the orthogonality principle, determine a set of normal equations that can be used to find $c_1, c_2, \ldots, c_m$ that minimize the induced norm of E.

3. (MMA: 7.3.5) Consider a matrix A with SVD $A = U \Sigma V^H$, where $U = [U_1 \; U_2]$, $V = [V_1 \; V_2]$, and Σ is block diagonal with blocks Σ_1 and Σ_2. Show that the squared error of the least-squares solution (i.e., the minimum of $\|A\mathbf{x} - \mathbf{b}\|_2^2$) is

$$E_{\min}^2 = \|U_2^H \mathbf{b}\|_2^2.$$

Interpret this result in light of the four fundamental subspaces.

4. (MMA: 2.15.80) (Modified Gram-Schmidt) Let $A \in \mathbb{C}^{m \times n}$ be any complex matrix and $Q = [\mathbf{q}_1, \mathbf{q}_2, \cdots, \mathbf{q}_n]$ be the matrix whose columns are the normalized output of the Gram-Schmidt algorithm. Then, $R = Q^{-1}A = Q^H A$ is the upper triangular matrix that completes the QR decomposition of A. The standard Gram-Schmidt process produces a column of Q and a column of R during each step, but is not well-conditioned numerically. The modified Gram-Schmidt (MGS) process improves stability by reordering the computation to produce a column of Q and a row of R during each step.

One can think of the MGS process as a sequence of elementary column operations that transform A into Q. The kth step consists of normalizing the kth column and then subtracting it from columns $k+1, k+2, \ldots, n$ to make them orthogonal to the kth column.

Let the kth column of Q be denoted as $\mathbf{q}_k$, and let the kth row of R be denoted as $\mathbf{r}_k^T$.

(a) Show that for an $m \times n$ matrix A,

$$A - \sum_{i=1}^{k-1} \mathbf{q}_i \mathbf{r}_i^T = \sum_{i=k}^{n} \mathbf{q}_i \mathbf{r}_i^T = \left[\begin{array}{cc} \mathbf{0} & A^{[k]} \end{array} \right],$$

where $A^{[k]}$ is $m \times (n-k+1)$.

(b) Let $A^{[k]} = [\ \mathbf{z}_k \quad B^{[k]}\ ]$, where $B^{[k]}$ is $m \times n - k$, and explain why the kth column of Q and the kth row of R are given by

$$r_{kk} = \|\mathbf{z}_k\| \quad \mathbf{q}_k = \mathbf{z}_k / r_{kk} \quad (r_{k,k+1}, \ldots, r_{k,n}) = \mathbf{q}_k^H B^{[k]}.$$

(c) Then, show that the next step starts by computing

$$A - \sum_{i=1}^{k} \mathbf{q}_i \mathbf{r}_i^T = \left[\begin{array}{cc} \mathbf{0} & A^{[k+1]} \end{array} \right],$$

where $A^{[k+1]} = B - \mathbf{q}_k (r_{k,k+1}, \ldots, r_{k,n})$.

(d) Code both Gram-Schmidt algorithms in MATLAB.

(e) Use both GS algorithms and the MATLAB's built-in `qr(A,0)` function to compute the QR decomposition of the 15 by 10 "Hilbert matrix" A defined by $[A]_{ij} = \frac{1}{i+j-1}$. Evaluate the performance of each algorithm by computing the distances $d(A, QR)$ and $d(Q^H Q, I)$ when $d(A, B) \triangleq \max_{i,j} |A_{ij} - B_{ij}|$.