This homework mostly serves as a linear algebra refresher. We will recall some definitions. The null space or kernel of a matrix $A \in \mathbb{R}^{m \times n}$ is the set
$$\ker(A) = \{ x \in \mathbb{R}^n : Ax = 0 \}$$
while the range space or image is the set
$$\text{im}(A) = \{ y \in \mathbb{R}^m : y = Ax \text{ for some } x \in \mathbb{R}^n \}.$$  

The rank and nullity of $A$ are defined as the dimensions of these spaces,
$$\text{rank}(A) = \text{dim } \text{im}(A) \quad \text{and} \quad \text{nullity}(A) = \text{dim } \ker(A).$$

By convention we write all vectors in $\mathbb{R}^n$ as column vectors.

1. (a) For $A \in \mathbb{R}^{m \times n}$ and $B \in \mathbb{R}^{n \times p}$, show that
$$\text{im}(AB) \subseteq \text{im}(A) \quad \text{and} \quad \ker(AB) \supseteq \ker(B).$$

When does equality occur in each of these inclusions?

(b) For $A, B \in \mathbb{R}^{n \times n}$, show that
$$\text{rank}(AB) \leq \min\{ \text{rank}(A), \text{rank}(B) \},$$
$$\text{nullity}(AB) \leq \text{nullity}(A) + \text{nullity}(B),$$
$$\text{rank}(A + B) \leq \text{rank}(A) + \text{rank}(B).$$

(c) For $A, B \in \mathbb{R}^{n \times n}$, show that if $AB = 0$, then
$$\text{rank}(A) + \text{rank}(B) \leq n.$$

2. (a) Let $A \in \mathbb{R}^{m \times n}$ and $B \in \mathbb{R}^{p \times q}$. Show that
$$\text{rank} \left( \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} \right) = \text{rank}(A) + \text{rank}(B).$$

We have used the block matrix notation here. For example if $A = \begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix} \in \mathbb{R}^{2 \times 3}$ and $B = \begin{bmatrix} \beta \end{bmatrix} \in \mathbb{R}^{2 \times 1}$, then
$$\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} = \begin{bmatrix} a & b & c & 0 \\ d & e & f & 0 \\ 0 & 0 & 0 & \alpha \\ 0 & 0 & 0 & \beta \end{bmatrix} \in \mathbb{R}^{4 \times 4}.$$

This is sometimes also denoted as $A \oplus B$. It is a direct sum of operators induced by a direct sum of vector spaces.

(b) For $x = [x_1, \ldots, x_m]^\top \in \mathbb{R}^m$ and $y = [y_1, \ldots, y_n]^\top \in \mathbb{R}^n$, observe that $xy^\top \in \mathbb{R}^{m \times n}$. Let $A \in \mathbb{R}^{m \times n}$. Show that $\text{rank}(A) = 1$ iff $A = xy^\top$ for some nonzero $x \in \mathbb{R}^m$ and $y \in \mathbb{R}^n$. 

---

*Date: September 30, 2020 (Version 1.0); due: October 11, 2020.*
3. Let \( A \in \mathbb{R}^{m \times n} \).
   (a) Show that \( \ker(A^TA) = \ker(A) \) and \( \text{im}(A^TA) = \text{im}(A^T) \).
   Give an example to show this is not true over a finite field (e.g. a field of two elements \( \mathbb{F}_2 = \{0, 1\} \) with binary arithmetic).
   (b) Show that \( A^TAx = A^Tb \) always has a solution (even if \( Ax = b \) has no solution). Give an example to show that this is not true over a finite field.

4. Let \( v_1, \ldots, v_n \in \mathbb{R}^n \). Let \( G_r = [g_{ij}] \in \mathbb{R}^{r \times r} \) be the matrix with \( g_{ij} = v_i^Tv_j \) for \( i, j = 1, \ldots, r \). This is called a Gram matrix.
   (a) Show that \( v_1, \ldots, v_r \) are linearly independent iff \( \text{nullity}(G_r) = 0 \).
   (b) Show that \( G_r = I_r \) iff \( v_1, \ldots, v_r \) are pairwise orthogonal unit vectors, i.e., \( \|v_i\|_2 = 1 \) for all \( i = 1, \ldots, r \), and \( v_i^Tv_j = 0 \) for all \( i \neq j \). If this holds, show that
   \[
   \sum_{i=1}^{r}(v^Tv_i)^2 \leq \|v\|_2^2
   \]
   (4.1)
   for all \( v \in \mathbb{R}^n \). What can you say about \( v_1, \ldots, v_r \) if equality always holds in (4.1) for all \( v \in \mathbb{R}^n \)?

5. Let \( A \in \mathbb{C}^{n \times n} \). Recall that \( A \) is diagonalizable iff there exists an invertible \( X \in \mathbb{C}^{n \times n} \) such that \( X^{-1}AX = \Lambda \), a diagonal matrix.
   (a) Show that \( A \) is diagonalizable if and only if its minimal polynomial is of the form \( m_A(x) = (x - \lambda_1) \cdots (x - \lambda_d) \)
   where \( \lambda_1, \ldots, \lambda_d \in \mathbb{C} \) are all distinct. Hence deduce for a diagonalizable matrix, the degree of its minimal polynomial equals the number of distinct eigenvalues.
   (b) Let \( A \) be diagonalizable. Let \( x_1, \ldots, x_n \in \mathbb{C}^n \) be \( n \) linearly independent right eigenvectors, i.e., \( Ax_i = \lambda_i x_i \); and \( y_1, \ldots, y_n \in \mathbb{C}^n \) be \( n \) linearly independent left eigenvectors, i.e., \( y_i^TA = \lambda_i y_i^T \). Show that there is a choice of left and right eigenvectors of \( A \) such that any vector \( v \in \mathbb{C}^n \) can be expressed as
   \[
   v = \sum_{i=1}^{n}(y_i^Tv)x_i.
   \]
   If we write \( X = [x_1, \ldots, x_n] \in \mathbb{C}^{n \times n} \) and \( Y = [y_1, \ldots, y_n] \in \mathbb{C}^{n \times n} \). What is the relation between \( X \) and \( Y \)?