CS/ECE/ME 532
Homework 1: Vectors and Matrices

1. **Matrix multiplication.** The local factory makes widgets and gizmos. Making one widget requires 3 lbs of materials, 4 parts, and 1 hour of labor. Making one gizmo requires 2 lbs of materials, 3 parts, and 2 hours of labor.

   a) Write the information above in a matrix. What do the rows represent? What do the columns represent?

   SOLUTION: One possible representation is:

   $$A = \begin{bmatrix} 3 & 4 & 1 \\ 2 & 3 & 2 \end{bmatrix}$$

   The rows represent the requirements for widgets (first row) and gizmos (second row). Each is a row vector that contains the lbs of materials, number of parts, and hours of labor required. The columns represent the lbs of materials (first column), the number of parts (second column) and the hours of labor (third column). Each column is a vector where the first component corresponds to widgets and the second component corresponds to gizmos.

   b) Suppose materials cost \$1/lb, parts cost \$10 each, and labor costs \$100/hr. Write this information in a vector. Write out a matrix-vector multiplication that calculates the total cost of making widgets and gizmos.

   SOLUTION: Write the material costs in a column vector: $c = \begin{bmatrix} 1 & 10 & 100 \end{bmatrix}^T$. The matrix multiplication that calculates total cost is:

   $$t = Ac = \begin{bmatrix} 3 & 4 & 1 \\ 2 & 3 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 10 \\ 100 \end{bmatrix} = \begin{bmatrix} 143 \\ 232 \end{bmatrix}$$

   So the total cost of making widgets and gizmos is \$143 and \$232 respectively.

   c) Suppose the factory receives an order for 3 widgets and 4 gizmos. Again using matrix multiplication, find the total material, parts, and labor required to fill the order.

   SOLUTION: Write the order in a column vector: $g = \begin{bmatrix} 3 & 4 \end{bmatrix}^T$. The matrix multiplication that calculates the requirements is:

   $$r = g^T A = \begin{bmatrix} 3 & 4 \end{bmatrix} \begin{bmatrix} 3 & 4 & 1 \\ 2 & 3 & 2 \end{bmatrix} = \begin{bmatrix} 17 & 24 & 11 \end{bmatrix}$$

   So to make 3 widgets and 4 gizmos, we would require 17 lbs of materials, 24 parts, and 11 hours of labor.

d) Calculate the total cost for the order (using, you guessed it, matrix multiplication)

SOLUTION: The total cost is $g^T Ac = 1357$

e) Get up and running with either Matlab or Python. In your language of choice, write a script that computes the matrix multiplications in the previous parts of this problem.

SOLUTION: We will show the code for part c); the other are similar.

```
% Matlab
g = [3;4];
A = [3 4 1; 2 3 2];
r = g'*A
```

```
# Python 3.4
import numpy as np
g = np.matrix([[3],[4]])
A = np.matrix([[3,4,1],[2,3,2]])
r = g.T * A
print(r)
```

2. Let $\boldsymbol{X} = [\boldsymbol{x}_1\ \boldsymbol{x}_2\ \cdots\ \boldsymbol{x}_n] \in \mathbb{R}^{p\times n}$, where $\boldsymbol{x}_i \in \mathbb{R}^p$ is the ith column of $\boldsymbol{X}$. Consider the matrix

$$\boldsymbol{C} = \frac{\boldsymbol{X}\boldsymbol{X}^T}{n}.$$

a) Express $\boldsymbol{C}$ as a sum of rank-1 matrices (i.e., columns of $\boldsymbol{X}$ times rows of $\boldsymbol{X}^T$).

SOLUTION: Split X into its columns and carry out the multiplication:

$$\tfrac{1}{n}XX^T = \tfrac{1}{n}\begin{bmatrix} x_1 & \dots & x_n \end{bmatrix}\begin{bmatrix} x_1^T \\ \vdots \\ x_n^T \end{bmatrix} = \tfrac{1}{n}\sum_{k=1}^{n} x_k x_k^T$$

each $x_k x_k^T$ is an $p \times p$ rank-1 matrix.

b) Assuming $\boldsymbol{x}_1, \boldsymbol{x}_2, \dots, \boldsymbol{x}_n$ are linearly independent, what is the rank of $\boldsymbol{C}$?

SOLUTION: We will prove something more general, that

$$\text{rank}(X) = \text{rank}(X^T X) = \text{rank}(XX^T) = n$$

We made a change to this problem to only ask for $\text{rank}(X^T X)$, but we have included the full proof anyway. We will prove the result in steps.

(i) $\text{rank}(X) = n$. This is true because the columns of X are independent.

(ii) $\text{rank}(X^T X) = n$. Since $X^T X \in \mathbb{R}^{n\times n}$, We can prove the result by showing that $X^T X$ has independent columns. If a linear combination of the columns is zero, then $X^T X w = 0$ for some vector w. This implies that $0 = w^T X^T X w = \|Xw\|^2$. If a norm is zero, its argument must be zero. So $Xw = 0$. Using the fact that X has independent columns, we conclude that $w = 0$. Therefore, $X^T X$ has independent columns, which means $\text{rank}(X^T X) = n$.

(iii) $\text{rank}(XA) \le \text{rank}(X)$ for any matrix A. This is because each column of XA is a linear combination of columns of X. So the span of the columns of XA is a subset (possibly equal) of the span of the columns of X.

(iv) $\text{rank}(AX) \le \text{rank}(X)$. Same argument as in (ii) applied to rows instead of columns.

(v) $\text{rank}(X^T X X^T X) = n$. To see why this is true, notice that from (ii), $X^T X$ is full-rank and therefore invertible. Since $X^T X X^T X (X^T X)^{-1}(X^T X)^{-1} = I$, we conclude that $X^T X X^T X$ is invertible as well, and therefore has rank n.

(vi) $\text{rank}(XX^T) = n$. Using (iii)–(iv), we have:

$$n = \text{rank}(X^T X X^T X) \le \text{rank}(XX^T) \le \text{rank}(X) = n$$

Therefore $\text{rank}(XX^T) = n$ and we are done.

Note: there is a much easier way to solve this problem using the SVD—we will learn it later on in the course!

3. Define the mapping $\Phi : \mathbb{R}^n \to \mathbb{R}$ as follows. For every $\boldsymbol{x} \in \mathbb{R}^n$ let

$$\Phi(\boldsymbol{x}) \;=\; \sum_{i=1}^{n} \sum_{j=i+1}^{n} \max(|x_i|, |x_j|) \;,$$

where x_i and x_j are the ith and jth entries in $\boldsymbol{x}$. Is $\Phi(\boldsymbol{x})$ a norm?

SOLUTION: $\Phi(x)$ is a norm whenever $n \ge 2$. To prove this, we must verify the three axioms.

(i) $\Phi(x) = 0$ if and only if $x = 0$.
Suppose $\Phi(x) = 0$. Let k correspond to the index of the largest entry of x in absolute value. In other words, $|x_k| \ge |x_m|$ for all $m = 1, \dots, n$. Therefore, we also have $|x_k| = \max(|x_k|, |x_i|)$ for all $i = 1, \dots, n$. We can lower-bound $\Phi(x)$ by only considering the parts of the sum where $i = k$ or $j = k$.

$$\Phi(x) = \sum_{i=1}^{n} \sum_{j=i+1}^{n} \max(|x_i|, |x_j|) \ge \sum_{i\neq k} \max(|x_k|, |x_i|) = (n-1)|x_k| \ge 0$$

Since $\Phi(x) = 0$, it must be that $|x_k| = 0$. If the largest absolute entry in x is zero, all entries in x are zero. So $x = 0$. In the case where $n = 1$, $\Phi(x) = 0$ for all x and this argument is invalid. So Φ is not a norm when $n = 1$.

(ii) $\Phi(\alpha x) = |\alpha|\Phi(x)$.
Note that for any x_i, we have $|\alpha x_i| = |\alpha||x_i|$. Therefore:

$$\begin{aligned}
\Phi(\alpha x) &= \sum_{i=1}^{n} \sum_{j=i+1}^{n} \max(|\alpha x_i|, |\alpha x_j|) \\
&= \sum_{i=1}^{n} \sum_{j=i+1}^{n} \max(|\alpha||x_i|, |\alpha||x_j|) \\
&= \sum_{i=1}^{n} \sum_{j=i+1}^{n} |\alpha| \max(|x_i|, |x_j|) \\
&= |\alpha| \sum_{i=1}^{n} \sum_{j=i+1}^{n} \max(|x_i|, |x_j|) \\
&= |\alpha|\Phi(x)
\end{aligned}$$

(iii) $\Phi(x + y) \le \Phi(x) + \Phi(y)$, the triangle inequality. One way to do this is to notice that the maximum is an infinity-norm in disguise, and then to apply the triangle inequality to each term of the sum.

$$\begin{aligned}
\Phi(x + y) &= \sum_{i=1}^{n} \sum_{j=i+1}^{n} \max(|x_i + y_i|, |x_j + y_j|) \\
&= \sum_{i=1}^{n} \sum_{j=i+1}^{n} \left\| \begin{bmatrix} x_i \\ x_j \end{bmatrix} + \begin{bmatrix} y_i \\ y_j \end{bmatrix} \right\|_\infty \\
&\le \sum_{i=1}^{n} \sum_{j=i+1}^{n} \left\| \begin{bmatrix} x_i \\ x_j \end{bmatrix} \right\|_\infty + \left\| \begin{bmatrix} y_i \\ y_j \end{bmatrix} \right\|_\infty \\
&= \sum_{i=1}^{n} \sum_{j=i+1}^{n} \max(|x_i|, |x_j|) + \max(|y_i|, |y_j|) \\
&= \Phi(x) + \Phi(y)
\end{aligned}$$

4. Equivalence of norms. For each case below find positive constants a and b (possibly different in each case) so that for every $\boldsymbol{x} \in \mathbb{R}^n$

(i) $a\|\boldsymbol{x}\|_1 \;\le\; \|\boldsymbol{x}\|_2 \;\le\; b\|\boldsymbol{x}\|_1$

SOLUTION: The lower bound follows from the Cauchy-Schwarz inequality, which states that for any vectors u, v, we have: $u^T v \le \|u\|_2\|v\|_2$. Apply it here by letting

$$u = \begin{bmatrix} 1 & 1 & \dots & 1 \end{bmatrix}^T \qquad \text{and} \qquad v = \begin{bmatrix} |x_1| & |x_2| & \dots & |x_n| \end{bmatrix}^T$$

Then we have: $\|x\|_1 = u^T v$, and $\|u\|_2 = \sqrt{n}$, and $\|v\|_2 = \|x\|_2$. By Cauchy-Schwarz, we then have $\|x\|_1 \le \sqrt{n}\|x\|_2$. Equality occurs when all components of x are equal. The upper bound follows from the triangle inequality. Write:

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} x_1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ x_2 \\ \vdots \\ 0 \end{bmatrix} + \cdots + \begin{bmatrix} 0 \\ 0 \\ \vdots \\ x_n \end{bmatrix}$$

Then the triangle inequality gives: $\|x\|_2 \le |x_1| + \cdots + |x_n| = \|x\|_1$. Equality occurs when x has a single nonzero component. Assembling the upper and lower bounds, we obtain:

$$\frac{1}{\sqrt{n}}\|x\|_1 \le \|x\|_2 \le \|x\|_1$$

(ii) $a\|\boldsymbol{x}\|_1 \ \le \ \|\boldsymbol{x}\|_\infty \ \le \ b\|\boldsymbol{x}\|_1$

SOLUTION: Recall that $\|x\|_\infty = \max_{1\le i\le n} |x_i|$. It is the largest component of x in absolute value. The upper bound is simply $\|x\|_\infty \le \|x\|_1$. Equality occurs when x has a single nonzero component. The lower bound is found by reasoning that given a list of nonnegative numbers, the largest number is always greater than (or equal to) the average of the numbers. Equality occurs when all components are equal. Assembling both bounds, we obtain:

$$\frac{1}{n}\|x\|_1 \le \|x\|_\infty \le \|x\|_1$$

(iii) $a\|\boldsymbol{x}\|_1 \ \le \ \Phi(\boldsymbol{x}) \ \le \ b\|\boldsymbol{x}\|_1$ where $\Phi(\boldsymbol{x})$ is defined in problem 3 above.

SOLUTION: The norm $\Phi(x)$ is a sum of maxima taken over all pairs of components of x. It's easy to convince yourself that $\Phi(\cdot)$ as well as $\|\cdot\|_1$ are invariant under permutation of components. So without loss of generality, we may assume that $x_1 \ge x_2 \ge \cdots \ge x_n$. In this scenario,

$$\begin{aligned} \|x\|_1 &= |x_1| + |x_2| + \cdots + |x_{n-1}| + |x_n| \\ \Phi(x) &= (n-1)|x_1| + (n-2)|x_2| + \cdots + |x_{n-1}| \end{aligned}$$

The upper bound is $\Phi(x) \le (n-1)\|x\|_1$, which follows by substituting in the expansions of the two norms. Equality occurs in this case when x_1 is nonzero and all other components are zero. The lower bound is $\Phi(x) \ge \frac{1}{2}(n-1)\|x\|_1$. Again, it can be verified by substitution. Equality occurs in this case when all components of x are equal. Assembling both bounds, we obtain:

$$\frac{1}{2}(n-1)\|x\|_1 \le \Phi(x) \le (n-1)\|x\|_1$$

5. Let $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ be a non-singular matrix. Suppose that $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}$.

a) Given $\boldsymbol{A}$ and $\boldsymbol{y}$, write an expression for $\boldsymbol{x}$.

SOLUTION: $x = A^{-1}y$

b) Bound the 2-norm of $\boldsymbol{x}$ in terms of $\|\boldsymbol{y}\|_2$ and a function of the matrix $\boldsymbol{A}$.

SOLUTION: By the definition of the induced 2-norm (spectral norm), we have:

$$\begin{aligned} y &= Ax &&\implies && \|y\|_2 \le \|A\|\|x\|_2 \\ x &= A^{-1}y &&\implies && \|x\|_2 \le \|A^{-1}\|\|y\|_2 \end{aligned}$$

Rearranging these inequalities, we obtain both lower and upper bounds:

$$\|A\|^{-1}\|y\|_2 \;\le\; \|x\|_2 \;\le\; \|A^{-1}\|\|y\|_2$$

Note that $\|A\|^{-1} \ne \|A^{-1}\|$ in general.

c) Suppose that instead we are given $\boldsymbol{A}$ and $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{e}$ where both $\boldsymbol{x}$ and $\boldsymbol{e}$ are unknown. Bound the error between $\boldsymbol{x}$ and $\boldsymbol{A}^{-1}\boldsymbol{y}$ in terms of the norm of $\boldsymbol{e}$.

SOLUTION: if $y = Ax + e$, then rearranging, we obtain:

$$A(A^{-1}y - x) = e$$

This is the same as part (b) but with different x and y vectors. The bound is:

$$\|A\|^{-1}\|e\|_2 \;\le\; \|A^{-1}y - x\|_2 \;\le\; \|A^{-1}\|\|e\|_2$$

6. Let

$$\boldsymbol{A} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}.$$

a) What is the rank of $\boldsymbol{A}$?

SOLUTION: A has rank 3 (full rank) because of the triangular structure.

b) Suppose that $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}$. Derive an explicit formula for $\boldsymbol{x}$ in terms of $\boldsymbol{y}$.

SOLUTION: Writing out the equation $y = Ax$, we have:

$$y = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} x \quad\iff\quad \begin{aligned} y_1 &= x_1 + x_2 + x_3 \\ y_2 &= x_1 + x_2 \\ y_3 &= x_1 \end{aligned}$$

From the last equation, $x_1 = y_3$. Substituting this into the second equation, $x_2 = y_2 - y_3$. Substituting all of this into the first equation, $x_3 = y_1 - y_2$. So we conclude that:

$$\begin{aligned} x_1 &= y_3 \\ x_2 &= y_2 - y_3 \\ x_3 &= y_1 - y_2 \end{aligned} \quad\iff\quad x = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & -1 \\ 1 & -1 & 0 \end{bmatrix} y$$

7. Let

$$\boldsymbol{X} = \begin{bmatrix} 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 \end{bmatrix}.$$

a) What is the rank of $\boldsymbol{X}$?

b) What is the rank of $\boldsymbol{X}\boldsymbol{X}^T$?

c) Find a set of linearly independent columns in $\boldsymbol{X}$.

SOLUTION: The rank of X is 3. The first three columns are clearly linearly independent, and the last two columns are linear combinations of the first three. Namely:

$$x_4 = x_3 + x_2 - x_1 \quad \text{and} \quad x_5 = x_3 + x_2$$

Note that we can deduce a rank bound simply by examining the rows. Since there are only three nonzero rows, the rank can't be any larger than 3. As for the matrix XX^T, it is equal to:

$$XX^T = \begin{bmatrix} 3 & 1 & 0 & 2 \\ 1 & 3 & 0 & 2 \\ 0 & 0 & 0 & 0 \\ 2 & 2 & 0 & 3 \end{bmatrix}$$

Again, this matrix has rank at most 3 because of the zero column. The rank of XX^T is actually exactly 3, because the remaining three columns are linearly independent.