# Applied Math HW 1

Colin Williams

September 11, 2021

## Question 1

Show that any linear transformation $f : \mathbb{R}^n \to \mathbb{R}^m$ can be expressed as

$$f(x) = Ax$$

where $A$ is an $m \times n$ matrix.

*Proof.*
Let $x = (x_1, x_2, \ldots, x_n)^T$ be a column vector in $\mathbb{R}^n$. Then, recall that $x$ can be expressed as

$$x = x_1 e_1 + x_2 e_2 + \ldots + x_n e_n$$

where $(e_j)$ is the standard basis of $\mathbb{R}^n$. Thus, by the linearity of $f$, we can say

$$\begin{aligned} f(x) &= f(x_1 e_1 + x_2 e_2 + \ldots + x_n e_n) \\ &= f(x_1 e_1) + f(x_2 e_2) + \ldots + f(x_n e_n) \\ &= x_1 f(e_1) + x_2 f(e_2) + \ldots + x_n f(e_n) \\ &= \sum_{j=1}^{n} x_j f(e_j) \end{aligned}$$

Next, denote $f(e_j)$ as the following vector in $\mathbb{R}^m$:

$$f(e_j) = \begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix}$$

Then, substitute this expression into the previous summation to get:

$$\begin{aligned} f(x) &= \sum_{j=1}^{n} x_j f(e_j) \\ &= \sum_{j=1}^{n} x_j \begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix} \\ &= x_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} + \cdots + x_n \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix} \\ &= \begin{bmatrix} x_1 a_{11} + x_2 a_{12} + \ldots + x_n a_{1n} \\ x_1 a_{21} + x_2 a_{22} + \ldots + x_n a_{2n} \\ \vdots \\ x_1 a_{m1} + x_2 a_{m2} + \ldots + x_n a_{mn} \end{bmatrix} \\ &= \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & \ddots & \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \\ &= Ax \end{aligned}$$

which finishes the proof. $\square$

## Question 2

Show that any matrix $A \in \mathbb{R}^{m \times n}$ with $m < n$ has full rank if and only if the map $x \mapsto Ax$ is surjective.

*Proof.*
First, assume $A \in \mathbb{R}^{m \times n}$ is of full rank, i.e. $\text{Rank}(A) = m$. By definition of Rank, this means that $A$ has $m$ linearly independent columns. Let $S = \{a_1, a_2, \ldots, a_m\} \subset \mathbb{R}^m$ be these linearly independent columns (Note that since $m < n$, this is not all columns of $A$). Note that $S$ must be a basis for $\mathbb{R}^m$ since it is a set of $m$ linearly independent vectors of $\mathbb{R}^m$. Therefore, any $y \in \mathbb{R}^m$ can be expressed as a linear combination of vectors in $S$. Thus, since the matrix-vector multiplication $Ax$ can be thought of as a linear combination of the columns of $A$, we know that there exists some solution to the equation $Ax = y$ which means that the map $x \mapsto Ax$ is surjective.

Next, assume that the map $x \mapsto Ax$ is surjective. This means that for any $y \in \mathbb{R}^m$, we can find a solution to the equation $Ax = y$. Since this matrix-vector multiplication can be thought of as a linear combination of the columns of $A$, we can say that $y$ is an element of the column space of $A$. Since this can be done for any $y \in \mathbb{R}^m$, we can say that the column space of $A$ is equal to $\mathbb{R}^m$. In other words, we have that $\dim(\text{Column Space}) = \text{Rank}(A) = m$. Thus, $A$ is full rank. □

## Question 3

Let $A \in \mathbb{R}^{n \times n}$, with $\text{rank}(A) = 1$.

(a) Show that $A$ can be written as an outer product $uv^T$, with $u, v \in \mathbb{R}^n \backslash \{0\}$.

- *Proof.* Since $\text{Rank}(A) = 1$, we know that $A$ only has 1 linearly independent column. In other words, every column of $A$ is equal to a scalar multiple of the first column of $A$:

$$\begin{aligned} A &= \begin{bmatrix} | & | & & | \\ a_1 & a_2 & \ldots & a_n \\ | & | & & | \end{bmatrix} \\ &= \begin{bmatrix} | & | & & | \\ \beta_1 a_1 & \beta_2 a_1 & \ldots & \beta_n a_1 \\ | & | & & | \end{bmatrix} \end{aligned}$$

  Where $\beta_i \in \mathbb{R}$ for $i = 1, 2, \ldots, n$ (Clearly, $\beta_1 = 1$ but the presentation seemed nicer to leave it as is). Thus, if we define

$$u := a_1 = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} \qquad\qquad v := \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{bmatrix}$$

  Then we get that

$$\begin{aligned} uv^T &= \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} \cdot \begin{bmatrix} \beta_1 & \beta_2 & \cdots & \beta_n \end{bmatrix} \\ &= \begin{bmatrix} u_1\beta_1 & u_1\beta_2 & \cdots & u_1\beta_n \\ u_2\beta_1 & u_2\beta_2 & \cdots & u_2\beta_n \\ \vdots & \vdots & \ddots & \vdots \\ u_n\beta_1 & u_n\beta_2 & \cdots & u_n\beta_n \end{bmatrix} \\ &= \begin{bmatrix} | & | & & | \\ \beta_1 a_1 & \beta_2 a_1 & \ldots & \beta_n a_1 \\ | & | & & | \end{bmatrix} \\ &= A \end{aligned}$$

  Notice that $a_1$ is nonzero and that not every $\beta_i$ is zero or otherwise $A$ would be the zero matrix which has Rank of zero. Thus, $u, v \neq 0$ and we have shown the desired statement. □

(b) Give an interpretation of $\text{null}(A)$ and $\text{range}(A)$ in terms of the vectors $u$ and $v$ in part (a).

- Recall that $\text{null}(A)$ is the set of all vectors $x$ such that $Ax = 0$. However, in our case, this is equivalent to evaluating $uv^T x = 0$. Since Matrix multiplication is an associative operation, we can evaluate this as

$u(v^T x) = 0$. Note that $v^T x \in \mathbb{R}$ so $u(v^T x) = 0 \iff v^T x = 0$. Therefore, $x$ must be an orthogonal vector to $v$, meaning we can interpret null$(A)$ as

$$\text{null}(A) = \{x \in \mathbb{R}^n \mid x \text{ is orthogonal to } v\}$$

- Recall that range$(A)$ is the span of the columns of $A$. However in our case, every column is a multiple of $u$, so we can say that

$$\text{range}(A) = \{\beta u \mid \beta \in \mathbb{R}\}$$

# Question 4

For $A, B \in \mathbb{C}^{n \times n}$, show that $(AB)^* = B^* A^*$.

*Proof.*
Let us do this by direct computation. I will use the notation $[C]_{ij}$ to represent the $(i,j)$-th component of the matrix $C$.

$$\begin{aligned}
[AB]_{ij} &= \sum_{k=1}^{n} [A]_{ik}[B]_{kj} && \text{by definition of matrix product} \\
\implies [(AB)^*]_{ij} &= \overline{\sum_{k=1}^{n} [A]_{jk}[B]_{ki}} && \text{since adjoint swaps indices then conjugates} \\
&= \sum_{k=1}^{n} \overline{[A]_{jk}[B]_{ki}} \\
&= \sum_{k=1}^{n} \overline{[A]_{jk}}\ \overline{[B]_{ki}} \\
&= \sum_{k=1}^{n} \overline{[B]_{ki}}\ \overline{[A]_{jk}} \\
&= \sum_{k=1}^{n} [B^*]_{ik}[A^*]_{kj} && \text{since adjoint swaps indices then conjugates} \\
&= [B^* A^*]_{ij} && \text{by definition of matrix product}
\end{aligned}$$

Therefore, I have shown that the $(i,j)$-th component of $(AB)^*$ is equal to the $(i,j)$-th component of $B^* A^*$ which proves the two matrices are equal. □

# Question 5

(a) Show that

$$||x||_\infty \leq ||x||_1 \leq n||x||_\infty$$

.

*Proof.*

- First, note that

$$||x||_\infty = \max_{1 \leq j \leq n} |x_j| \qquad \text{and} \qquad ||x||_1 = \sum_{j=1}^{n} |x_j|$$

Thus, let $k$ be the index of the maximum coordinate of $x$ so that $||x||_\infty = |x_k|$. Then, we clearly have that

$$\begin{aligned}
||x||_\infty &= |x_k| \\
&\leq |x_1| + |x_2| + \ldots + |x_k| + \ldots + |x_n| \\
&= ||x||_1
\end{aligned}$$

which proves the first inequality.

- Next, notice that since $|x_k|$ is the maximum of all coordinates of $x$ that $|x_i| \leq |x_k|$ for all $i = 1, 2, \ldots, n$. Thus,

$$\begin{aligned}
||x||_1 &= |x_1| + |x_2| + \ldots + |x_n| \\
&\leq |x_k| + |x_k| + \ldots + |x_k| \\
&= n|x_k| \\
&= n||x||_\infty
\end{aligned}$$

Thus, I have shown both inequalities hold true.

□

(b) Show that

$$||A|| = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}^2}, \quad A = [a_{ij}] \in \mathbb{R}^{m\times n}$$

defines a matrix norm on $\mathbb{R}^{m\times n}$.

*Proof.*

- Notice that $a_{ij}^2$ is guaranteed to be non-negative since $a_{ij} \in \mathbb{R}$. Thus, $||A||$ is the square root of the sum of non-negative numbers, so $||A|| \geq 0$ which clearly has equality when $A$ is the zero matrix. On the other hand, if $||A|| = 0 \implies ||A||^2 = 0$, then we can conclude that each $a_{ij}^2 = 0$ since we're evaluating a sum and each term is non-negative. Thus, $a_{ij} = 0$ for all $i, j$.
- Next, notice the following:

$$\begin{aligned}
||A||^2 &= \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}^2 \\
&= \sum_{j=1}^{n}\sum_{i=1}^{m} a_{ij}^2 \\
&= \sum_{j=1}^{n} ||a_j||_2^2 && \text{by definition of vector 2-norm} \\
&= \left|\left| \begin{bmatrix} | \\ a_1 \\ | \\ \\ | \\ a_2 \\ | \\ \vdots \\ | \\ a_n \end{bmatrix} \right|\right|_2^2
\end{aligned}$$

  The final expression is simply the vector 2-norm of the vector composed of each column of $A$ stacked on top of each other. Thus, the matrix norm we are given can be thought of as the vector 2-norm for a vector $\widetilde{A} \in \mathbb{R}^{m\cdot n}$ where the tilde simply represents a re-indexing of the matrix $A$. Thus, we can use the fact that the vector 2-norm satisfies the Triangle Inequality to say:

$$\begin{aligned}
||A + B|| &= ||\widetilde{A + B}||_2 \\
&= ||\widetilde{A} + \widetilde{B}||_2 \\
&\leq ||\widetilde{A}||_2 + ||\widetilde{B}||_2 \\
&= ||A|| + ||B||
\end{aligned}$$

  Thus, the Triangle Inequality is satisfied for the Matrix Norm
- Lastly, let $\alpha \in \mathbb{R}$ be a scalar, then

$$\begin{aligned}
||\alpha A||^2 &= \sum_{i=1}^{m}\sum_{j=1}^{n} (\alpha a_{ij})^2 \\
&= \sum_{i=1}^{m}\sum_{j=1}^{n} \alpha^2 a_{ij}^2 \\
&= \alpha^2 \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}^2 \\
&= \alpha^2 ||A|| \\
\implies ||\alpha A|| &= |\alpha| \; ||A||
\end{aligned}$$

  Thus, this function satisfies all three required properties of a norm.

□