# Applied Math HW 2

Colin Williams

September 21, 2021

Prove the following properties of matrix norms for $A$ an $m \times n$ matrix for properties 1-4 and an $n \times n$ matrix for properties 5-8.

## Property 1

$$||A||_1 := \max_{x \neq 0} \frac{||Ax||_1}{||x||_1} = \max_j \sum_{i=1}^m |a_{ij}| = ||A^T||_\infty$$

*Proof.*
Recall that in class, we have proven that

$$||A||_\infty = \max_i \sum_{j=1}^n |a_{ij}|$$

Thus, $||A||_\infty$ is the maximum sum over the absolute values of elements of each row of $A$. Therefore, since $A^T$ has rows equal to the columns of $A$, then $||A^T||_\infty$ should be equal to the maximum sum over the absolute values of elements of each column of $A$. In other words,

$$||A^T||_\infty = \max_j \sum_{i=1}^m |a_{ij}|$$

which proves the last equality above. Next, moving to the left side of the desired equality, we get

$$\begin{aligned} ||A||_1 = \max_{x \neq 0} \frac{||Ax||_1}{||x||_1} &= \max_{x \neq 0} \frac{\left|\left|\sum_{i=1}^n a_i x_i\right|\right|_1}{||x||_1} \\ &\leq \max_{x \neq 0} \frac{\sum_{i=1}^n ||a_i||_1 |x_i|}{||x||_1} && \text{by Triangle Inequality} \\ &\leq \max_{x \neq 0} \frac{\max_j ||a_j||_1 \sum_{i=1}^n |x_i|}{||x||_1} \\ &= \max_{x \neq 0} \frac{\max_j ||a_j||_1 ||x||_1}{||x||_1} \\ &= \max_j ||a_j||_1 \end{aligned}$$

where $(a_j)$ denote the $j$-th column of $A$. Thus, using the definition of the 1-norm, we get

$$||A||_1 \leq \max_j \sum_{i=1}^m |a_{ij}|$$

Next, let $k$ be the index in the above expression such that $\sum_{i=1}^m |a_{ij}|$ is maximized when $j = k$ and let $e_k$ denote the $k$-th standard unit vector. If that is the case, then

$$\begin{aligned}\max_j \sum_{i=1}^{m} |a_{ij}| &= \max_j ||a_j||_1 \\ &= ||Ae_k||_1 \\ &= \frac{||Ae_k||_1}{||e_k||_1} \\ &\leq \max_{x\neq 0} \frac{||Ax||_1}{||x||_1} \\ &= ||A||_1\end{aligned}$$

Thus, I have shown that both inequalities hold, so in fact

$$||A||_1 = \max_j \sum_{i=1}^{m} |a_{ij}|$$

□

# Property 2

$$||A||_2 := \max_{x\neq 0} \frac{||Ax||_2}{||x||_2} = \sqrt{\lambda_{max}(A^T A)} \qquad \text{where } \lambda_{max}(B) \text{ denotes the largest eigenvalue of } B$$

*Proof.*

□

# Property 3

$$||A||_2 = ||A^T||_2$$

*Proof.*
Taking Property 2 as given, we see that

$$||A||_2 = \sqrt{\lambda_{max}(A^T A)} \qquad \text{and} \qquad ||A^T||_2 = \sqrt{\lambda_{max}(AA^T)}$$

Furthermore, as we have shown in class, the matrix $A^T A$ has real and non-negative eigenvalues. Also, we have shown that $AB$ and $BA$ have the same non-zero eigenvalues for any matrices $A$ and $B$. Therefore, $A^T A$ and $AA^T$ have the same non-zero eigenvalues. Thus, if $\lambda_{max}(A^T A) > 0$, then we know that $\lambda_{max}(A^T A) = \lambda_{max}(AA^T)$. If $\lambda_{max}(A^T A) = 0$. Then we know that all eigenvalues of $A^T A$ are zero. This also means that $AA^T$ has all zero eigenvalues since otherwise the set of non-zero eigenvalues wouldn't be the same. Thus $\lambda_{max}(AA^T) = 0$. Therefore, in either case $\lambda_{max}(A^T A) = \lambda_{max}(AA^T)$. This proves that $||A||_2 = ||A^T||_2$. □

# Property 4

$$||QAZ|| = ||A||$$

where $Q \in \mathbb{R}^{m\times m}$ and $Z \in \mathbb{R}^{n\times n}$ are orthogonal matrices. The matrix norm here is either the Frobenius norm or the operator norm induced by $||\cdot||_2$.

*Proof.* Frobenius.
Let's recall the definition of the Frobenius Norm

$$||A||_F = \left(\sum_{j=1}^{n}\sum_{i=1}^{m} |a_{ij}|^2\right)^{1/2} = \left(\sum_{j=1}^{n} ||a_j||_2^2\right)^{1/2} = \left(\sum_{i=1}^{m} ||a_j||_2^2\right)^{1/2}$$

where $(a_j)$ is the $j$-th column of $A$. The second expression is what I will start using with the LHS of the desired equality:

$$||QAZ||_F^2 = \sum_{j=1}^{n} ||r_j||_2^2$$

where $(r_j)$ is the $j$-th column of $QAZ$. Let $(q_j) \subset \mathbb{R}^m$ be the $j$-th column of $Q$ and let $(y_{ij})$ be the $i,j$-th element of $AZ$, then we have that

$$r_j = \sum_{k=1}^{m} q_k y_{kj}$$

Thus, we get

$$\begin{aligned}
||QAZ||_F^2 &= \sum_{j=1}^{n} \left|\left| \sum_{k=1}^{m} q_k y_{kj} \right|\right|_2^2 \\
&= \sum_{j=1}^{n} \left\langle \sum_{k=1}^{m} q_k y_{kj}, \sum_{k=1}^{m} q_k y_{kj} \right\rangle \\
&= \sum_{j=1}^{n} \sum_{k=1}^{m} |y_{kj}|^2 ||q_k||_2^2 && \text{since the colums of } Q \text{ are orthogonal} \\
&= \sum_{j=1}^{n} \sum_{k=1}^{m} |y_{kj}|^2 && \text{since the columns of } Q \text{ are normal} \\
&= ||AZ||_F^2
\end{aligned}$$

Similarly, let $(a_{ij})$ be the elements of $A$ and let $(z_i)$ be the rows of $Z$. Then, I will use the third expression for the Frobenius norm as expressed as the beginning to say that

$$||AZ||_F^2 = \sum_{i=1}^{m} ||y_i||_2^2$$

where $(y_i)$ is the $i$-th row of $AZ$. Note we can express $y_i$ as

$$y_i = \sum_{k=1}^{n} z_k a_{ik}$$

Therefore, we get

$$\begin{aligned}
||AZ||_F^2 &= \sum_{i=1}^{m} \left|\left| \sum_{k=1}^{n} z_k a_{ik} \right|\right|_2^2 \\
&= \sum_{i=1}^{m} \left\langle \sum_{k=1}^{n} z_k a_{ik}, \sum_{k=1}^{n} z_k a_{ik} \right\rangle \\
&= \sum_{i=1}^{m} \sum_{k=1}^{n} |a_{ik}|^2 ||z_k||^k && \text{since the rows of } Z \text{ are orthogonal} \\
&= \sum_{i=1}^{m} \sum_{k=1}^{n} |a_{ik}|^2 && \text{since the rows of } Z \text{ are normal} \\
&= ||A||_F^2
\end{aligned}$$

□

*Proof.* Operator Norm.
Notice that

$$\begin{aligned}
||QAZ||_2 &= \max_{x \neq 0} \frac{||QAZx||_2}{||x||_2} \\
&= \max_{x \neq 0} \frac{||Q(AZx)||_2}{||x||_2} \\
&= \max_{x \neq 0} \frac{||AZx||_2}{||x||_2} && \text{since } Q \text{ is orthogonal, i.e. satisfies } ||Qv|| = ||v|| \\
&= ||AZ||_2
\end{aligned}$$

Next, use property 3 to conclude that $||AZ||_2 = ||(AZ)^T||_2 = ||Z^T A^T||_2$. Note that $Z^T$ must also be an orthogonal matrix, so we can use an equivalent proof as above to conclude that

$$||Z^T A^T||_2 = ||A^T||_2$$

Thus, once again using property 3, we get the sequence of equalities

$$||QAZ||_2 = ||AZ||_2 = ||Z^T A^T||_2 = ||A^T||_2 = ||A||_2$$

□

## Property 5

$$||A||_2 = \max_{||x||_2=1} |x^T Ax| \quad \text{if } A \text{ is symmetric}$$

*Proof.*
Let $x$ be a unit vector. Since $A$ is symmetric, it has an orthonormal eigenbasis $\{v_i\}_{i=1}^n$. Furthermore, each eigenvalue of $A$ must be non-negative. Thus, expressing $x$ in this eigenbasis gives:

$$\begin{aligned}
x &= \sum_{i=1}^n \alpha_i v_i \\
\implies x^T Ax &= \sum_{i=1}^n \alpha_i v_i^T \sum_{i=1}^n \alpha_i A v_i \\
&= \sum_{i=1}^n \alpha_i v_i^T \sum_{i=1}^n \alpha_i \lambda_i v_i \\
&= \sum_{i=1}^n \alpha_i^2 \lambda_i && \text{since the basis is orthonormal}
\end{aligned}$$

Notice that this sum here is always positive and is maximized when $\alpha_i = 1$ for $i$ the index of the maximal eigenvalue and $\alpha_i = 0$ for all other indices. Thus, we get that

$$\max_{||x||_2=1} |x^T Ax| = \lambda_{max}(A)$$

On the other hand, by property 2, we have

$$\begin{aligned}
||A||_2 &= \sqrt{\lambda_{max}(A^T A)} \\
&= \sqrt{\lambda_{max}(A^2)}
\end{aligned}$$

Furthermore, let $\lambda_0$ be the maximal eigenvalue of $A$ with eigenvector $v_0$. Then, we have that

$$A^2 v_0 = A(Av_0) = A(\lambda_0 v_0) = \lambda_0 A v_0 = \lambda_0^2 v_0$$

so that $\lambda_0^2$ is an eigenvalue of $A^2$ with the same eigenvector. Furthermore, $\lambda_0^2$ must be the maximal eigenvalue of $A^2$ since we can find that if $\lambda_i$ is an eigenvalue for $A$, then $\lambda_i^2$ is an eigenvalue for $A^2$ by the argument above. Thus, since each $\lambda_i$ is non-negative, the maximum of $\{\lambda_i^2\}$ corresponds to the square of the maximum of $\{\lambda_i\}$. Thus, we get

$$\sqrt{\lambda_{max}(A^2)} = \sqrt{\lambda_{max}(A)^2} = |\lambda_{max}(A)| = \lambda_{max}(A)$$

Thus, we get

$$||A||_2 = \lambda_{max}(A) = \max_{||x||_2=1} |x^T Ax|$$

which finishes the proof □

## Property 6

$$\frac{1}{\sqrt{n}}||A||_2 \leq ||A||_1 \leq \sqrt{n}||A||_2$$

*Proof.*
Recall the similar inequality for vector norms:

$$||x||_2 \leq ||x||_1 \leq \sqrt{n}||x||_2 \tag{0.1}$$

Thus, we have that

$$\begin{aligned}
||A||_1 = \max_{x \neq 0} \frac{||Ax||_1}{||x||_1} &\leq \max_{x \neq 0} \frac{\sqrt{n}||Ax||_2}{||x||_1} && \text{by (0.1)} \\
&\leq \max_{x \neq 0} \frac{\sqrt{n}||Ax||_2}{||x||_2} && \text{by (0.1)} \\
&= \sqrt{n}||A||_2 \\
||A||_1 = \max_{x \neq 0} \frac{||Ax||_1}{||x||_1} &\geq \max_{x \neq 0} \frac{||Ax||_1}{\sqrt{n}||x||_2} && \text{by (0.1)} \\
&\geq \max_{x \neq 0} \frac{||Ax||_2}{\sqrt{n}||x||_2} && \text{by (0.1)} \\
&= \frac{1}{\sqrt{n}}||A||_2
\end{aligned}$$

$\square$

## Property 7

$$\frac{1}{\sqrt{n}}||A||_2 \leq ||A||_\infty \leq \sqrt{n}||A||_2$$

*Proof.*
Recall the similar property we have for vector norms:

$$||x||_\infty \leq ||x||_2 \leq \sqrt{n}||x||_\infty \tag{0.2}$$

Thus, we have that

$$\begin{aligned}
||A||_\infty = \max_{x\neq 0}\frac{||Ax||_\infty}{||x||_\infty} &\leq \max_{x\neq 0}\frac{||Ax||_2}{||x||_\infty} && \text{by (0.2)}\\
&\leq \max_{x\neq 0}\frac{||Ax||_2}{(1/\sqrt{n})||x||_2} && \text{by (0.2)}\\
&= \sqrt{n}||A||_2\\
||A||_\infty = \max_{x\neq 0}\frac{||Ax||_\infty}{||x||_\infty} &\geq \max_{x\neq 0}\frac{(1/\sqrt{n})||Ax||_2}{||x||_\infty} && \text{by (0.2)}\\
&\geq \max_{x\neq 0}\frac{||Ax||_2}{\sqrt{n}||x||_2} && \text{by (0.2)}\\
&= \frac{1}{\sqrt{n}}||A||_2
\end{aligned}$$

$\square$