✍1 Unitary matrices preserve the 2-norm.

Solution: The proof takes only one line if we use the result $(Ax, y) = (x, A^H y)$:

$$\|Qx\|_2^2 = (Qx, Qx) = (x, Q^H Qx) = (x, x) = \|x\|_2^2. \qquad \square$$

✍3 When do we have equality in Cauchy-Schwarz?

Solution: From the proof of Cauchy-Schwarz it can be seen that we have equality when $x = \lambda y$, i.e., when they are colinear.$\square$

✍4 Expand $(x + y, x + y)$ – What does Cauchy-Schwarz imply?

Solution: You will see that you can derive the triangle inequality from this expansion and the Cauchy-Schwarz inequality. $\square$.

- Proof of the Hölder inequality.

$$|(x, y)| \le \|x\|_p \|y\|_q \text{ , with } \frac{1}{p} + \frac{1}{q} = 1$$

Proof: For any z_i, v_i all nonnegative we have, setting $\zeta = \sum z_i$,

$$\begin{aligned}
\left(\sum (z_i/\zeta) v_i\right)^p &\le \sum (z_i/\zeta) v_i^p \text{ (convexity)} \rightarrow \\
\left(\sum z_i v_i\right)^p &\le \left[\sum (z_i/\zeta) v_i^p\right] \zeta^p = \left[\sum z_i v_i^p\right] \zeta^{p-1} \rightarrow \\
\sum z_i v_i &\le \left[\sum z_i v_i^p\right]^{1/p} \zeta^{(p-1)/p} \\
\sum z_i v_i &\le \left[\sum z_i v_i^p\right]^{1/p} \left[\sum z_i\right]^{1/q}
\end{aligned}$$

Now take $z_i = x_i^q$, and $v_i = y_i * x_i^{1-q}$. Then $z_i v_i = x_i y_i$ and:

$$z_i v_i^p = x_i^q * (y_i * x_i^{1-q})^p = y_i^p * x_i^{q+p-pq} = y_i^p * x_i^0 == y_i^p \quad \square$$

✍5 Second triangle inequality.

Solution: Start by invoking the triangle inequality to write:

$$\|x\| = \|(x-y)+y\| \le \|x-y\|+\|y\| \rightarrow \|x\|-\|y\| \le \|x-y\|$$

Next exchange the roles of x and y:

$$\|y\| - \|x\| \le \|y - x\| = \|x - y\|$$

The two inequalities $\|x\| - \|y\| \le \|x - y\|$ and $\|y\| - \|x\| \le \|x - y\|$ yield the result since they imply that

$$-\|x - y\| \le \|x\| - \|y\| \le \|x - y\|$$

$\square$

✍6 Consider the metric $d(x,y) = max_i|x_i - y_i|$. Show that any norm in $\mathbb{R}^n$ is a continuous function with respect to this metric.

Solution: We need to show that we can make $\|y\|$ arbitrarily close to $\|x\|$ by making y 'close' enough to x, where 'close' is measured in terms of the infinity norm distance $d(x,y) = \|x - y\|_\infty$. Define $u = x - y$ and write u in the canonical basis as $u = \sum_{i=1}^n \delta_i e_i$. Then:

$$\|u\| = \|\sum_{i=1}^n \delta_i e_i\| \le \sum_{i=1}^n |\delta_i| \, \|e_i\| \le \max |\delta_i| \sum_{i=1}^n \|e_i\|$$

Setting $M = \sum_{i=1}^n \|e_i\|$ we get

$$\|u\| \le M \max |\delta_i| = M\|x - y\|_\infty$$

Let ϵ be given and take x, y such that $\|x - y\|_\infty \le \frac{\epsilon}{M}$. Then, by using the second triangle inequality we obtain:

$$| \, \|x\| - \|y\| \, | \le \|x - y\| \le M \max \delta_i \le M \frac{\epsilon}{M} = \epsilon.$$

This means that we can make $\|y\|$ arbitrarily close to $\|x\|$ by making y close enough to x in the sense of the defined metric. Therefore $\|.\|$ is continuous. $\square$

✍7 In $\mathbb{R}^n$ (or $\mathbb{C}^n$) all norms are equivalent.

Solution: We will do it for $\phi_1 = \|.\|$ some norm, and $\phi_2 = \|.\|_\infty$ [and one can see that all other cases will follow from this one].

1. Need to show that for some α we have $\|x\| \le \alpha\|x\|_\infty$. Express x in the canonical basis of $\mathbb{R}^n$ as $x = \sum x_i e_i$ [look up canonical basis e_i from your csci2033 class.] Then

$$\|x\| = \|\sum x_i e_i\| \le \sum |x_i|\|e_i\| \le \max |x_i| \sum \|e_i\| = \|x\|_\infty \alpha$$

where $\alpha = \sum \|e_i\|$.

2. We need to show that there is a β such that $\|x\| \ge \beta\|x\|_\infty$. Assume $x \ne 0$ and consider $u = x/\|x\|_\infty$. Note that u has infinity norm equal to one. Therefore it belongs to the closed and bounded set $S_\infty = \{v \| \|v\|_\infty = 1\}$. Since norms are continuous (seen earlier), the minimum of the norm $\|u\|$ for all $u's$ in S_∞ is *reached*, i.e., there is a $u_0 \in S_\infty$ such that

$$\min_{u \in S_\infty} \|u\| = \|u_0\|.$$

Let us call β this minimum value, i.e., $\|u_0\| = \beta$. Note in passing that β cannot be equal to zero otherwise $u_0 = 0$ which would contradict the fact that u_0 belongs to S_∞ [all vectors in S_∞ have infinity norm equal to one.] The result follows because $u = x/\|x\|_\infty$, and so, remembering that $u = x/\|x\|_\infty$, we obtain

$$\left\| \frac{x}{\|x\|_\infty} \right\| \geq \beta \rightarrow \|x\| \geq \beta \|x\|_\infty$$

This completes the proof □

✍8 Show that for any x: $\frac{1}{\sqrt{n}}\|x\|_1 \leq \|x\|_2 \leq \|x\|_1$

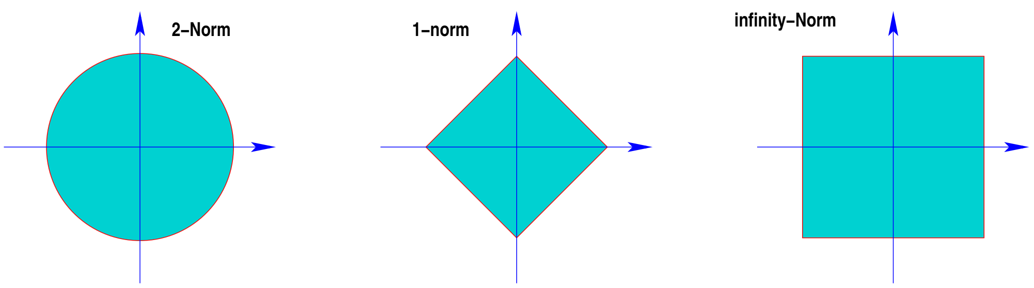
Solution: For the right inequality, it is easy to see that $\|x\|_2 \leq \|x\|_1$ because $\sum_i x_i^2 \leq [\sum_i |x_i|]^2$

For the left inequality, we rely on Cauchy-Schwarz. If we call $\mathbf{1}$ the vector of all ones, then:

$$\|x\|_1 = \sum_i |x_i|.1 \leq \|x\|_2 \|One\|_2 = \sqrt{n}\|x\|_2\|$$

□

✍9 Unit balls in $\mathbb{R}^2$.



✍14 Show that $\rho(A) \leq \|A\|$ for any matrix norm.

Solution: Let λ be the largest (in modulus) eigenvalue of A with associated eigenvector u. Then

$$Au = \lambda u \rightarrow \frac{\|Au\|}{\|u\|} = |\lambda| = \rho(A)$$

This implies that

$$\rho(A) \leq \max_{x \neq 0} \frac{\|Ax\|}{\|x\|} = \|A\|$$

□

✍15 Is $\rho(A)$ a norm?

Solution: This was answered in the notes.

✍16 Given a function $f(t)$ (e.g., e^t) how would you define $f(A)$? [You may limit yourself to the case when A is diagonalizable]

Solution: The easiest way would be through the Taylor series expansion..

$$f(A) = f(0)I + \frac{f'(0)}{1!}A + \frac{f''(0)}{2!}A^2 \cdots \frac{f^{(k)}(0)}{k!}A^k + \cdots$$

However, this will require a justification: Will this expression 'converge' as the number of terms goes to infinity? This is where norms are useful.

In the simplest case where A is diagonalizable you can write $A = XDX^{-1}$ and then consider the k-term part of the Taylor series expression above:

$$\begin{aligned} F_k &= f(0)I + \frac{f'(0)}{1!}A + \frac{f''(0)}{2!}A^2 + \cdots + \frac{f^{(k)}(0)}{k!}A^k \\ &= X\left[f(0)I + \frac{f'(0)}{1!}D + \frac{f''(0)}{2!}D^2 + \cdots + \frac{f^{(k)}(0)}{k!}D^k\right]X^{-1} \\ &\equiv XD_kX^{-1} \end{aligned}$$

where D_k is the matrix inside the brackets in line 2 of above equations. The $i-th$ diagonal entry of D_k is of the form

$$f_k(\lambda_i) = f(0) + \frac{f'(0)}{1!}\lambda_i + \frac{f''(0)}{2!}\lambda_i^2 + \cdots + \frac{f^{(k)}(0)}{k!}\lambda_i^k,$$

which is just the k-term part of the Taylor series expansion of $f(\lambda_i)$. Each of these will converge to $f(\lambda_i)$. Now it is easy to complete the argument. If we call D_f the diagonal matrix whose ith diagonal entry is $f(\lambda_i)$ and F_A the matrix defined by

$$F_A = XD_fX^{-1},$$

then clearly

$$\|F_k - F_A\|_2 = \|X(D_k - D_f)X^{-1}\|_2 \le \|X\|_2\|X^{-1}\|_2\|D_k - D_f\|_2$$
$$\le \|X\|_2\|X^{-1}\|_2 \max_i |f_k(\lambda_i) - f(\lambda_i)|$$

which converges to zero as k goes to infinity. ☐

✍17 The eigenvalues of A^HA and AA^H are real nonnegative.

Solution: Let us show it for A^HA [the other case is similar] If λ, u is an eigenpair of A^HA then $(A^HA)u = \lambda u$. Take inner products with u on both sides. Then:

$$\lambda(u, u) = ((A^HA)u, u) = (Au, Au) = \|Au\|^2$$

Therefore, $\lambda = \|Au\|^2/\|u\|^2$ which is a real nonnegative number. ☐

[Note: 1) Observe how simple the proof is for such an important fact. It is based on the result $(Ax, y) = (x, A^Hy)$. 2) The singular values of A are the square roots of the eigenvalues of A^HA if $m \ge n$ or those of the eigenvalues of AA^H if $m < n$. So there are always $\min(m, n)$ singular values. This is really just a preliminary definition as we need to refer to singular values often – but we will see singular values and the singular value decomposition in great detail later.]

✍18 Prove that when $A = uv^T$ then $\|A\|_2 = \|u\|_2 \|v\|_2$.

Solution: We start by dealing with the eigenvalues of an arbitrary matrix of the form $A = uv^T$ where both u and v are in $\mathbb{R}^n$. From $Ax = \lambda x$ we get:

$$uv^T x = \lambda x \rightarrow (v^T x) u = \lambda x$$

Notice that we did this because $v^T x$ is a scalar. We have 2 cases.

Case 1: $v^T x = 0$. In this case it is clear that the equation $Ax = \lambda x$ is satisfied with $\lambda = 0$. So any vector that is orthogonal to v is an eigenvector of A associated with the eigenvalue $\lambda = 0$. (It can be shown that the eigenvalue 0 is of multiplicity $n - 1$).

Case 2: $v^T x \neq 0$. In this case it is clear that the equation $Ax = \lambda x$ is satisfied with $\lambda = v^T u$ and $x = u$. So u is an eigenvector of A associated with the eigenvalue $v^T x$.

In summary the matrix uv^T has only two eigenvalues: 0, and $v^T u$.

Going back to the original question, we consider now $A = uv^T$ and we are interested in the 2-norm of A. We have

$$\|A\|_2^2 = \rho(A^T A) = \rho(vu^T uv^T) = \|u\|_2^2 \rho(vv^T) = \|u\|_2^2 \|v\|_2^2.$$

The last relation comes from what was done above to determine the eigenvalues of vv^T. So in the end, $\|A\|_2 = \|u\|_2 \|v\|_2$. □

✍19 In this case what is $\|A\|_F$?

Solution: Only the last part of the above answer changes (ρ is replaced by Tr) and you will find that actually the Frobenius norm of uv^T is again equal to $\|u\|_2 \|v\|_2$. □

Proof of Cauchy-Schwarz inequality: $|(x, y)|^2 \le (x, x)\ (y, y).$

Proof: We begin by expanding $(x - \lambda y, x - \lambda y)$ using properties of inner products:

$$(x - \lambda y, x - \lambda y) = (x, x) - \bar{\lambda}(x, y) - \lambda(y, x) + |\lambda|^2 (y, y).$$

If $y = 0$ then the inequality is trivially satisfied. Assume that $y \neq 0$ and take $\lambda = (x, y)/(y, y)$. Then, from the above equality, $(x - \lambda y, x - \lambda y) \ge 0$ shows that

$$\begin{aligned} 0 \le (x - \lambda y, x - \lambda y) &= (x, x) - 2\frac{|(x, y)|^2}{(y, y)} + \frac{|(x, y)|^2}{(y, y)} \\ &= (x, x) - \frac{|(x, y)|^2}{(y, y)}, \end{aligned}$$

which yields the result. □

PROOF OF EXPRESSION of $\| A \|_1 = \max_j \| A(:,j) \|_1$
$= \max$ of 1-norms of the columns of A

===

Definition:

$\| A \|_1 = \max_{\{\|x\|_1 = 1\}} \| Ax \|_1$

Let $\eta = \max_j \| A(:,j)\|_1 = \|A(:,j0)\|_1$ where j_0 = index of col that reaches the max

Part 1) $\|A\|_1 \le \eta$

let x any vector s.t. $\|x\|_1 = 1$

$\| Ax \|_1 = \| \sum x_j A(:,j) \|_1 \le \sum_j |x_j| \| A(:,j)\|_1$

$\le \sum_j |x_j| \max_j \| A(:,j)\|_1 = \sum |x_j| \eta$

$\| Ax\|_1 \le \eta \sum_j |x_j| = \eta$

==> $\|A\|_1 \le \eta$ □

Part 2) $\|A\|_1 \ge \eta$

Assume max reached for j0
let $x = e_{j0} = [\ 0;\ 0;\ ...,\ 0;\ 1;\ 0\ ....]$
^- position j0

$A e_j$ = jth column of A – note that $A e_{j0} = A(:,j_0) = j_0$-th column of A

for $x = e_{j0} \rightarrow$

Then $\| A x \|_1 = \| A(:,j0) \|_1 = \eta$

For one particular x $\| A x\|_1 = \eta \rightarrow$
so max $\| Ax \|_1$ over all x's with $\| x \|_1 = 1$ is $\ge \eta$

$\max_{\{x \mid \|x\|_1 = 1\}} \| A x\|_1 \ge \eta$

==> $\|A\|_1 \ge \eta$ □

EXPRESSION FOR INFINITY NORM OF A = max_i $\| A(i,:) \|_1$

===

NOTE: In the following ‖ . ‖_ is the infinity norm

To show ‖ A ‖_ = max of 1-norms of the *rows* of A

Definition: ‖ A ‖_ = max_{‖x‖_ = 1} ‖ Ax ‖_

Let $\eta = \max_i \| A(i,:)\|_1 = \|A(i0,:)\|_1$ where i0 = index of row that reaches the max

Part 1) ‖A‖_ $\le \eta$

Let x any vector s.t. ‖x‖_ = 1

$(Ax)_i = \sum_j A(i,j)\, x_j \rightarrow$

$|(Ax)_i| \le \sum_j |A(i,j)|\, |x_j| \rightarrow$

$\le \sum_j |A(i,j)| \max_j |x_j|$

$\le \sum_j |A(i,j)| = \| A(i,:) \|_1$

take max over i:

‖ Ax ‖_ $\le \eta$ for all x with ‖ x ‖_ = 1

===> ‖ A ‖_ $\le \eta$ □

Part 2) ‖A‖_ $\ge \eta$

Assume max reached for i0
let x_0 = vector of signs of the row A(i0,:)
= [±1; ±1,....,±1]

clearly ‖ x_0 ‖_ = 1

Then can show that

‖ A x_0 ‖_ = η

== similar argument as for $\| \|_1$ to complete proof

===