Equivalence of Norms on $\mathbb{R}^n$

Next, we discuss a result for l^∞, which then can be easily extended to other norms on $\mathbb{R}^n$ via another theorem (the equivalence theorem).

Theorem 1. *The sequence of vectors $\left\{\boldsymbol{x}^{(k)}\right\}$ converges to $\boldsymbol{x}$ in $\mathbb{R}^n$ with respect to the l^∞-norm if and only if $\lim_{k\to\infty} x_i^{(k)} = x_i$, for each $i = 1, 2, \ldots, n$.*

Proof. When proving statements involving if and only if such as $A \iff B$, we need to prove both directions of the statement, that is, $A \implies B$ and $B \implies A$.

(1) $\implies$: assume that $\boldsymbol{x}^{(k)} \to \boldsymbol{x}$ in l^∞, prove that $\lim_{k\to\infty} x_i^{(k)} = x_i$, for each $i = 1, 2, \ldots, n$.

Proof. Given $\epsilon > 0$, there exists an integer $N(\epsilon)$ such that for all $k \geq N(\epsilon)$,

$$\left\|\boldsymbol{x}^{(k)} - \boldsymbol{x}\right\|_\infty = \max_{i=1,2,\ldots,n} \left|x_i^{(k)} - x_i\right| < \epsilon.$$

This result implies that $\left|x_i^{(k)} - x_i\right| < \epsilon$, for each $i = 1, 2, \ldots, n$, so $x_i^{(k)} \to x_i$ for each i. □

(2) $\impliedby$: assume that $\lim_{k\to\infty} x_i^{(k)} = x_i$, for each $i = 1, 2, \ldots, n$, prove that $\boldsymbol{x}^{(k)} \to \boldsymbol{x}$ in l^∞.

Proof. Given $\epsilon > 0$, there exists an integer $N_i(\epsilon)$ such that for all $k \geq N_i(\epsilon)$,

$$\left|x_i^{(k)} - x_i\right| < \epsilon,$$

for each $i = 1, 2, \ldots, n$.

Define $N(\epsilon) = \max_{i=1,2,\ldots,n} N_i(\epsilon)$. If $k \geq N(\epsilon)$, then

$$\left\|\boldsymbol{x}^{(k)} - \boldsymbol{x}\right\|_\infty = \max_{i=1,2,\ldots,n} \left|x_i^{(k)} - x_i\right| < \epsilon$$

which implies $\left\{\boldsymbol{x}^{(k)}\right\}$ converges to $\boldsymbol{x}$ with respect to the l^∞-norm. □

□

Remark. You may wonder if this theorem only applies to the l^∞-norm. The next theorem will tell us the answer.

Theorem 2. *For $\boldsymbol{x} \in \mathbb{R}^n$,*

$$\|\boldsymbol{x}\|_\infty \leq \|\boldsymbol{x}\|_2 \leq \sqrt{n}\,\|\boldsymbol{x}\|_\infty .$$

Proof. See proof in your textbook. Theorem 7.7. □

Remark. This theorem also tells us that all norms on $\mathbb{R}^n$ are equivalent with respect to convergence. That is, if $\|\cdot\|$ and $\|\cdot\|'$ are any two norms on $\mathbb{R}^n$, and $\left\{\boldsymbol{x}^{(k)}\right\}$ has the limit $\boldsymbol{x}$ with respect to $\|\cdot\|$, then $\left\{\boldsymbol{x}^{(k)}\right\}$ also has the **same** limit $\boldsymbol{x}$ with respect to $\|\cdot\|'$.

Remark. To connect with the previous theorem that component-wise convergence and convergence are always true together, here we simply see that Theorem 1 is true for all l^p-norms in $\mathbb{R}^n$ (the converse of Theorem 2 becomes untrue when the vectors live in infinite-dimensional space).

Example. Consider the sequence

$$\boldsymbol{x}^{(k)} = \left(1, 2 + \frac{1}{k}, \frac{3}{k^3}, e^{-k} \sin k\right)^{\mathrm{T}} \in \mathbb{R}^4.$$

Find the limit of $\left\{\boldsymbol{x}^{(k)}\right\}$ in l^∞-norm and l^2-norm respectively.

Solution. l^∞-norm is straightforward. Theorem 1 tells us that the l^∞ limit of a sequence, if exists, is the limit of individual components. Thus,

$$\begin{aligned} \lim_{k\to\infty} 1 &= 1; \\ \lim_{k\to\infty} 2 + \frac{1}{k} &= 2; \\ \lim_{k\to\infty} \frac{3}{k^2} &= 0; \\ \lim_{k\to\infty} e^{-k} \sin(k) &= 0, \end{aligned}$$

and therefore, $\boldsymbol{x}^{(k)}$ converges to $(1, 2, 0, 0)^{\mathrm{T}}$ with respect to the l^∞-norm.

To show that $\boldsymbol{x}^{(k)}$ converges to the same limit in l^2, we use Theorem 2 and the definition of convergence.

Since $\boldsymbol{x}^{(k)} \to \boldsymbol{x} = (1, 2, 0, 0)^{\mathrm{T}}$ in l^∞, for any $\epsilon > 0$, we can furnish an integer $N(\epsilon)$ such that

$$\left\|\boldsymbol{x}^{(k)} - \boldsymbol{x}\right\|_\infty < \frac{\epsilon}{2}$$

whenever $k \geq N(\epsilon)$. By the equivalent theorem (Theorem 2), for exactly the same $N(\epsilon)$, we have

$$\left\|\boldsymbol{x}^{(k)} - \boldsymbol{x}\right\|_2 \overset{\text{Theorem 2 for } n=4}{\leq} \sqrt{4} \left\|\boldsymbol{x}^{(k)} - \boldsymbol{x}\right\|_\infty \leq 2\left(\frac{\epsilon}{2}\right) = \epsilon$$

for $k \geq N(\epsilon)$. Thus $\boldsymbol{x}^{(k)}$ also converges to $\boldsymbol{x}$ in l^2-norm.

Example. Sublevel sets of $\|\boldsymbol{x}\|_\infty$ and $\|\boldsymbol{x}\|_2$ for $\boldsymbol{x} \in \mathbb{R}^2$.

Consider $f(\boldsymbol{x}) = \|\boldsymbol{x}\|_\infty \leq 1$. This function outputs a square with sidelength 2, centered at $(0, 0)$.

$g(\boldsymbol{x}) = \|\boldsymbol{x}\|_2 \leq 1$ covers the circle of radius 1.

Matrix Norm

We have now seen a measure of distance between vectors in $\mathbb{R}^n$. Is there such a thing for matrices, namely, arrays that lives in $\mathbb{R}^{n\times n}$ (square matrices)? And even if there is, what geometric meaning does it have? We can surely visualize l^2 or l^∞ norms of vectors, but what about the "norm" of a matrix?

Definition. A **matrix norm** on the set of all $n \times n$ matrices is a real-valued function $\|\cdot\|$, defined on this set, satisfying for all $n \times n$ matrices A and B and all real numbers α:

(1) $\|A\| \geq 0$;
(2) $\|A\| = 0$ if and only if A is the zero matrix;
(3) $\|\alpha A\| = |\alpha| \|A\|$;
(4) $\|A + B\| \leq \|A\| + \|B\|$;
(5) $\|AB\| \leq \|A\| \|B\|$.

Then, the distance between $n \times n$ matrices A and B with respect to this matrix norm is $\|A - B\|$.

Still, all of these criteria are very abstract. Can we utilize the notion of a vector norm to induce a norm on a matrix?

Theorem. *If $\|\cdot\|$ is a vector norm on $\mathbb{R}^n$, then*

$$\|A\| = \max_{\|\boldsymbol{x}\|=1} \|A\boldsymbol{x}\| = \max_{\boldsymbol{z}\neq\boldsymbol{0}} \frac{\|A\boldsymbol{z}\|}{\|\boldsymbol{z}\|}$$

is a matrix norm.

Proof. We have nothing to prove if A is the zero matrix. The first three criteria are simple. We work with the first definition.

$$\begin{aligned}
\|A+B\| &= \max_{\|\boldsymbol{x}\|=1} \|(A+B)\,\boldsymbol{x}\| \\
&= \max_{\|\boldsymbol{x}\|=1} \|A\boldsymbol{x} + B\boldsymbol{x}\| \\
&\overset{\text{triangle ineq.}}{\leq} \max_{\|\boldsymbol{x}\|=1} \left(\|A\boldsymbol{x}\| + \|B\boldsymbol{x}\|\right) \\
&\leq \max_{\|\boldsymbol{x}\|=1} \|A\boldsymbol{x}\| + \max_{\|\boldsymbol{x}\|=1} \|B\boldsymbol{x}\| \\
&= \|A\| + \|B\|\,.
\end{aligned}$$

The last one requires some more work.

$$\|AB\| = \max_{\|\boldsymbol{x}\|=1} \|AB\boldsymbol{x}\|$$

and now we are stuck because we wish we can bound the $\|AB\boldsymbol{x}\|$ by something nice.

Lemma. $\|A\boldsymbol{x}\| \leq \|A\|\,\|\boldsymbol{x}\|$ *for* $\boldsymbol{x} \neq \boldsymbol{0}$.

Proof. Define $\boldsymbol{y} = \frac{\boldsymbol{x}}{\|\boldsymbol{x}\|}$. Then,

$$\|A\boldsymbol{x}\| \overset{\text{absolute homogeneity}}{=} \left\|A\frac{\boldsymbol{x}}{\|\boldsymbol{x}\|}\right\| \|\boldsymbol{x}\| \leq \max_{\|\boldsymbol{y}\|=1} \|A\boldsymbol{y}\|\,\|\boldsymbol{x}\| = \|A\|\,\|\boldsymbol{x}\|\,.$$

□

Now, we go on with the proof for the (submultiplicative) property:

$$\begin{aligned}
\|AB\| &= \max_{\|\boldsymbol{x}\|=1} \|AB\boldsymbol{x}\| \\
&\leq \max_{\|\boldsymbol{x}\|=1} \|A\|\,\|B\boldsymbol{x}\| \\
&= \|A\| \max_{\|\boldsymbol{x}\|=1} \|B\boldsymbol{x}\| \\
&= \|A\|\,\|B\|\,.
\end{aligned}$$

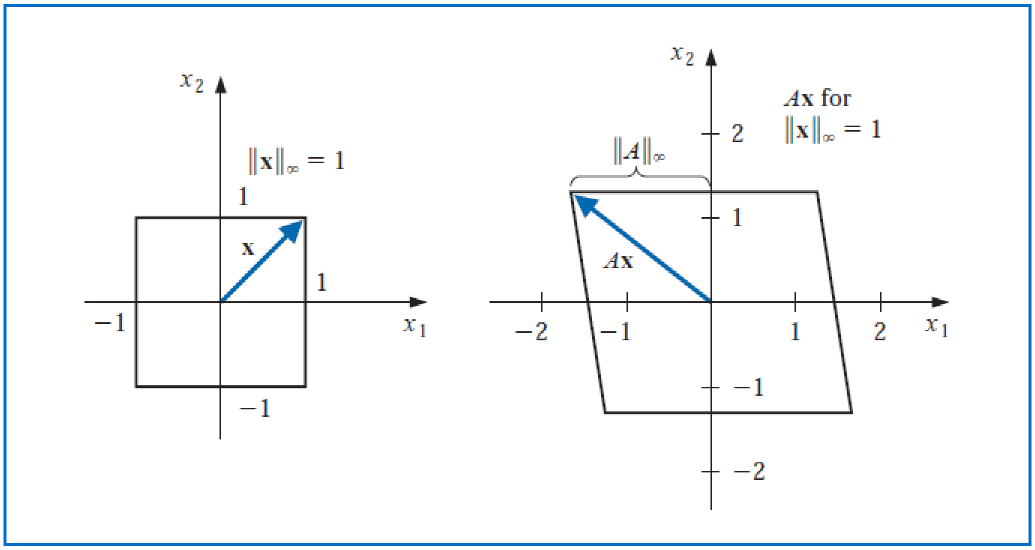
□

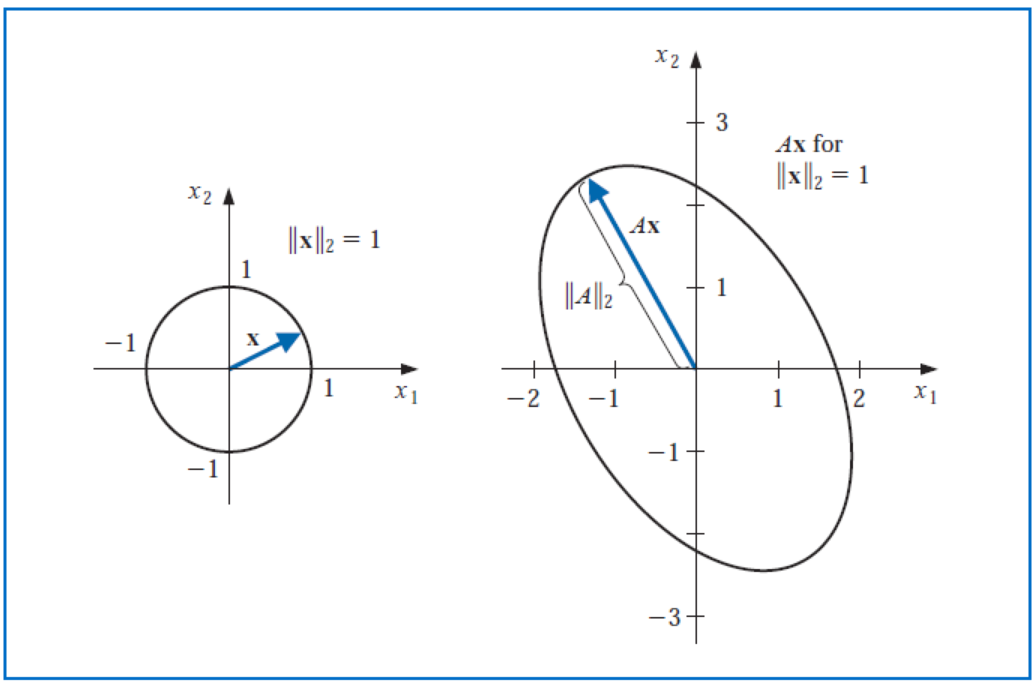
Any vector norm on $\mathbb{R}^n$ **induces** matrix norm. So, we can consider

$$\|A\|_\infty = \max_{\|\boldsymbol{x}\|_\infty=1} \|A\boldsymbol{x}\|_\infty\,, \quad l^\infty\text{-norm;}$$

$$\|A\|_2 = \max_{\|\boldsymbol{x}\|_2=1} \|A\boldsymbol{x}\|_2\,, \quad l^2\text{-norm.}$$

But again, what is geometrically meaningful of these matrix norms?





We learn a great deal from the action of a matrix on a vector. In principle, $A\boldsymbol{x}$ simply changes the direction and magnitude of $\boldsymbol{x}$. Therefore, if we consider the vectors that satisfy $\|\boldsymbol{x}\|_2 = 1$, namely, the vectors that live on the unit circle, a matrix A will transform this circle to an ellipse, but stretching (or compressing) the length, while rotating the principle axes of the circle. The figure above showcases the matrix norm where

$$A = \begin{bmatrix} 0 & -2 \\ 2 & 0 \end{bmatrix}.$$

In fact, this matrix rotates any initial vector $\boldsymbol{x}$ by 90 degrees counterclockwise, and then stretches twice as long.

Eigenvalues and Eigenvectors

Matrix operations on vectors have deep geometric meanings. For each matrix, there may be certain vectors that don't change direction at all, but only gets stretched or compressed. In other words,

$$A\boldsymbol{x} = \lambda \boldsymbol{x}$$

where $\lambda \in \mathbb{R}$ is a scaling factor that describes the extent of stretching or compression. We call λ the eigenvalue of A corresponding to the eigenvector $\boldsymbol{x}$.

Rearranging the equation, we have

$$A\boldsymbol{x} - \lambda \boldsymbol{x} = 0 \implies A\boldsymbol{x} - \lambda I \boldsymbol{x} = 0$$

where I is the identity matrix in $\mathbb{R}^n$. Then, we can factor out $\boldsymbol{x}$ such that

$$(A - \lambda I)\,\boldsymbol{x} = 0,$$

which is just another system of equations $\tilde{A}\boldsymbol{x} = 0$. This equation has a unique solution if and only if $\det \tilde{A} = \det (A - \lambda I) = 0$, i.e., the determinant of $(A - \lambda I)$ is zero. Note that $A - \lambda I$ only differs from A on the diagonal entries (since the identity only affects the diagonal entries).

Example. (Section 7.2 Example 2).