John Nachbar
Washington University in St. Louis
November 4, 2017

Vector Spaces and Norms[1]

1 The Vector Space $\mathbb{R}^N$.

1.1 $\mathbb{R}^N$ basics.

$\mathbb{R}^N$ is the N-fold Cartesian product of $\mathbb{R}$. Thus, for example, $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$. If $x = (x_1, \dots, x_n, \dots, x_N) \in \mathbb{R}^N$, then x_n is the value of *coordinate* n. I write 0 both for $0 \in \mathbb{R}$ and for $(0, \dots, 0) \in \mathbb{R}^N$.

An element x of $\mathbb{R}^N$ can be interpreted either as a *point* in $\mathbb{R}^N$ or as a *vector* (a directed line segment with base at the origin and head at x). I use these interpretations interchangeably.

I sometimes represent an elements of $\mathbb{R}^N$ as a matrix. If I do so then the matrix is a column:

$$x = (x_1, \dots, x_N) = \begin{bmatrix} x_1 \\ \vdots \\ x_N \end{bmatrix}.$$

Associated with $\mathbb{R}^N$ are two basic operations.

- *Vector addition.* If $x, y \in \mathbb{R}^N$ then

$$x + y = (x_1 + y_1, \dots, x_N + y_N) \in \mathbb{R}^N.$$

- *Scalar multiplication.* If $a \in \mathbb{R}$ and $x \in \mathbb{R}^N$ then

$$ax = (ax_1, \dots, ax_N) \in \mathbb{R}^N.$$

  The point $a \in \mathbb{R}$ is called a *scalar.*

$\mathbb{R}^N$, when combined with vector addition and scalar multiplication, is a vector space. I define general vector spaces in Section 2. $\mathbb{R}^N$ is the canonical example of a vector space.

$\mathbb{R}^N$ has the following standard partial order. For any $x, y \in \mathbb{R}^N$, write $x \geq y$ iff $x_n \geq y_n$ for all n, write $x > y$ iff $x \geq y$ and $x \neq y$, and write $x \gg y$ iff $x_n > y_n$ for all n. It bears emphasis that $\geq$ is not a complete order on $\mathbb{R}^N$; for example, $(2, 1) \not\geq (1, 2)$ and $(1, 2) \not\geq (2, 1)$. It also bears emphasis that $\gg$ does *not* mean "much greater than." We could have $x \gg y$ even if x were only a tiny bit greater than y in every coordinate (by whatever standard we are using for tiny).

The set $\mathbb{R}^N_+ = \{x \in \mathbb{R}^N : x \geq 0\}$ (i.e., the set of points in $\mathbb{R}^N$ where all coordinates are non-negative) is called the *non-negative orthant*. $\mathbb{R}^N_+$ is the standard mathematical setting for a number of economic models, including classical consumer theory with N commodities. The set $\mathbb{R}^N_{++} = \{x \in \mathbb{R}^N : x \gg 0\}$ is called the *strictly positive orthant*.

Two points $x, y \in \mathbb{R}^N$, $x, y \neq 0$, are *collinear* (lie on the same line through the origin) iff there is an $a \in \mathbb{R}$ such that $x = ay$.

1.2 Inner product.

Definition 1. *If $x, y \in \mathbb{R}^N$,*

$$x \cdot y = \sum_{n=1}^{N} x_n y_n.$$

$x \cdot y$ is the inner product *of x and y.*

Recall that a point $x \in \mathbb{R}^N$ can be represented as a (column) matrix. Recall also that if x is an $N \times 1$ column matrix then its transpose x' is a $1 \times N$ row matrix:

$$x' = \begin{bmatrix} x_1 & \cdots x_N \end{bmatrix}.$$

It then follows, by the standard rules of matrix multiplication, that for any $x, y \in \mathbb{R}^N$,

$$x \cdot y = x'y = y'x.$$

Remark 1. There is also something called the *outer product*: the outer product of x and y is the $N \times N$ matrix xy'. I will not be using outer products. □

As discussed in Remark 3 in Section 1.3, the inner product $x \cdot y$ is closely related to the angle between x and y (when x and y are viewed as vectors).

1.3 The Euclidean Norm.

Definition 2. *The* Euclidean norm *of $x \in \mathbb{R}^N$, written $\|x\|$, is*

$$\|x\| = \sqrt{x \cdot x}.$$

Alternatively,

$$\|x\| = \left(\sum_{n=1}^{N} x_n^2 \right)^{1/2}.$$

If $N = 1$, then the Euclidean norm is simply absolute value: $\|x\| = |x|$.

The standard interpretation of the Euclidean norm is that it is a measure of distance to the origin. In particular, consider $(x_1, x_2) \in \mathbb{R}^2$ and the right triangle formed by the vertices $(0,0)$, $(x_1,0)$, and (x_1,x_2). The side with endpoints $(0,0)$ and $(x_1,0)$, call it A, has length x_1. The side with endpoints $(x_1,0)$ and (x_1,x_2), call it B, has length x_2. By the Pythagorean Theorem, the hypotenuse, with endpoints $(0,0)$ and (x_1,x_2), has length equal to the square root of the sum of the squared lengths of A and B, namely $\sqrt{x_1^2+x_2^2} = \|x\|$. A similar argument applies in higher dimensions.

The Euclidean norm is named for the Greek mathematician Euclid, who studied "flat" geometries in which the Pythagorean theorem holds. (The Pythagorean Theorem does not hold, for example, for triangles inscribed on a sphere.) The vector space $\mathbb{R}^N$ with the Euclidean norm is called *Euclidean space*.

The Euclidean norm in $\mathbb{R}^N$ has the following properties.

Theorem 1. *For any $x, y \in \mathbb{R}^N$ and any $a \in \mathbb{R}$ the following hold.*

1. $\|x\| \geq 0$. $\|x\| = 0$ *iff* $x = 0$.
2. $\|ax\| = |a|\|x\|$.
3. $\|x+y\| \leq \|x\| + \|y\|$.

The third property is called the triangle inequality.

The motivation for the name "triangle inequality" is the following. Consider vectors x and y, $x, y \neq 0$, and suppose that x and y are *not* collinear. Then the points 0, x, and $x+y$ form a triangle with sides given by the line segment from 0 to x, the line segment from x to $x+y$, and the line segment from 0 to $x+y$. The triangle inequality says that the sum of the lengths of two sides of this triangle is greater than the length of the third side.

The difficult step in proving Theorem 1 is proving the triangle inequality. As an intermediate step, I first prove Theorem 2, called the Cauchy-Schwartz Inequality, which is of independent interest.

Theorem 2 (Cauchy-Schwartz Inequality)**.** *If $x, y \in \mathbb{R}^N$ then*

$$|x \cdot y| \leq \|x\| \; \|y\|.$$

Moreover, for $x, y \neq 0$, this weak inequality holds with equality iff x and y are collinear.

Proof. If $x \cdot y = 0$ then the inequality holds, since $\|x\|, \|y\| \geq 0$. Therefore, assume $x \cdot y \neq 0$. Note that this implies $x \neq 0$.

For any $\lambda \in \mathbb{R}$, since a sum of squares is always non-negative,

$$\begin{aligned} 0 &\leq (x - \lambda y) \cdot (x - \lambda y) \\ &= \|x\|^2 - 2\lambda(x \cdot y) + \lambda^2 \|y\|^2. \end{aligned}$$

Set

$$\lambda = \frac{\|x\|^2}{x \cdot y},$$

which is well defined since $x \cdot y \neq 0$. Therefore,

$$0 \leq \|x\|^2 - 2\frac{\|x\|^2}{(x \cdot y)}(x \cdot y) + \frac{\|x\|^4}{(x \cdot y)^2}\|y\|^2.$$

Collecting terms and rearranging yields,

$$\|x\|^2(x \cdot y)^2 \leq \|x\|^4\|y\|^2.$$

Since $\|x\| > 0$ ($\|x\| = 0$ iff $x = 0$), one can divide both sides by $\|x\|^2$ to get

$$(x \cdot y)^2 \leq \|x\|^2\|y\|^2.$$

Taking the square root of both sides yields the desired weak inequality.

Finally, assume $x, y \neq 0$. If x and y are collinear then there is a $\lambda \in \mathbb{R}$ such that $x = \lambda y$. It is easy to verify that the inequality holds with equality. Conversely, if x and y are not collinear then for any λ, in particular for $\lambda = \|x\|^2/(x \cdot y)$, $x - \lambda y \neq 0$, hence $0 < (x - \lambda y) \cdot (x - \lambda y)$, and the above argument then implies $|x \cdot y| < \|x\| \; \|y\|$. ■

Proof of Theorem 1. The first two properties are almost immediate. As for the third, the triangle inequality, the argument is as follows. By the Cauchy-Schwartz inequality, $x \cdot y \leq |x \cdot y| \leq \|x\| \; \|y\|$. Therefore,

$$\begin{aligned}
\|x + y\|^2 &= (x + y) \cdot (x + y) \\
&= x \cdot x + 2x \cdot y + y \cdot y \\
&\leq \|x\|^2 + 2\|x\| \; \|y\| + \|y\|^2 \\
&= (\|x\| + \|y\|)^2.
\end{aligned}$$

Taking the square root of both sides yields the desired inequality. ■

Remark 2. A stronger version of the triangle inequality holds for the Euclidean norm. Assume throughout that $x, y \neq 0$. Say that x, y are *positively collinear* iff there is an $\alpha \in \mathbb{R}$, $\alpha > 0$, such that $x = \alpha y$. Viewed as directed line segments, positively collinear vectors point in the same direction. I claim that

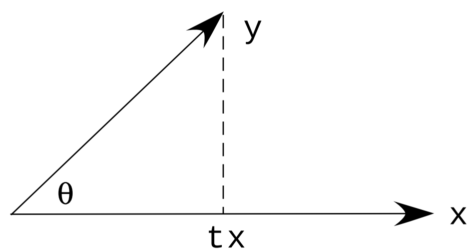
$$\|x + y\| = \|x\| + \|y\|$$

iff x and y are positively collinear.

Suppose $x, y \neq 0$. If x, y are positively collinear, with $x = \alpha y$, then $x \cdot y > 0$, since $x \cdot y = \alpha\|y\|^2 > 0$. Hence $x \cdot y = |x \cdot y| = \|x\|\|y\|$, with the last equality coming from the Cauchy-Schwartz Inequality. It then follows from the proof of Theorem

1 that the desired weak inequality holds with equality. If x, y are collinear but not positively collinear (i.e., $\alpha < 0$), then $x \cdot y < 0$, hence $x \cdot y < |x \cdot y|$ and the desired weak inequality holds strictly. Finally, by the Cauchy-Schwartz inequality, the desired weak inequality holds strictly if x and y are not collinear. □

Remark 3. Suppose $x, y \neq 0$ and label x and y so that the angle, call it θ, between x and y is non-negative. Assume further that θ lies strictly between 0 and 90 degrees. Let t be the scalar such that the line segment from 0 to tx is at right angles to the line segment from tx to y.



Then the points 0, tx, and y form a right triangle, with hypotenuse given by the line segment from 0 to y. By the Pythagorean Theorem, $\|y\|^2 = \|tx\|^2 + \|y - tx\|^2$. Hence $y \cdot y = t^2(x \cdot x) + (y - tx) \cdot (y - tx) = t^2(x \cdot x) + y \cdot y - 2t(x \cdot y) + t^2(x \cdot x)$. Cancelling the $y \cdot y$ and collecting terms yields

$$t = \frac{x \cdot y}{x \cdot x}.$$

By the assumption that θ lies strictly between 0 and 90 degrees, t is positive.

By the definition of cosine (length of the side adjacent to the angle divided by the length of the hypotenuse):

$$\cos(\theta) = \frac{\|tx\|}{\|y\|} = \frac{x \cdot y}{x \cdot x} \frac{(x \cdot x)^{\frac{1}{2}}}{\|y\|} = \frac{x \cdot y}{\|x\|\|y\|}.$$

Rewriting,

$$x \cdot y = \|x\|\|y\| \cos(\theta).$$

In particular, if $\|x\| = \|y\| = 1$ (so that both x and y lie on the unit circle centered on the origin), then $x \cdot y = \cos(\theta)$.

This implies the Cauchy-Schwartz inequality. Specifically, for θ strictly between 0 and 90 degrees, $0 < \cos(\theta) < 1$, hence $0 < x \cdot y < \|x\|\|y\|$.

One can extend this argument to θ in other ranges. In particular, $x \cdot y = 0$ iff x and y are *orthogonal* (are 90 degrees apart). □

2 General Vector Spaces.

A general vector space consists of a non-empty set X together with the following.

- There is a binary operation on X, *vector addition*, and an element $0 \in X$ such that for any $x, \hat{x} \in X$, $x + \hat{x} \in X$ (closure), $x + \hat{x} = \hat{x} + x$ (commutativity), and $x + 0 = x$ (0 is the additive identity). In addition, for any $x \in X$, one requires that there exists an element $-x$ such that $x + (-x) = 0$; $-x$ is the additive inverse of x.

- There is a field $\mathcal{F}$ (recall that $\mathbb{Q}$ and $\mathbb{R}$ are fields) and a binary operation, *scalar multiplication*, mapping from $\mathcal{F} \times X$ to X such that for any $a, b \in \mathcal{F}$ and any $x, \hat{x} \in X$

  1. $a(x + \hat{x}) = ax + a\hat{x}$
  2. $(a + b)x = ax + bx$
  3. $a(bx) = (ab)x$
  4. $1x = x$, where $1 \in \mathcal{F}$ is the multiplicative identity for the field $\mathcal{F}$.

  Note that for any $x \in X$, if 0 is the zero element (additive identify) of $\mathcal{F}$, then $0x = (1 - 1)x = x + (-x) = 0$. Once again, I am using 0 in two different senses, both as the zero element of $\mathcal{F}$ and as the zero element of X.

For nearly all applications in economics, the field $\mathcal{F}$ is taken to be $\mathbb{R}$, in which case the vector space is said to be over the reals. In what follows, I assume in every case that the vector field is over the reals and I omit writing "over the reals".

If X is a vector space and $A \subseteq X$ is a vector space with the same definition of vector multiplication and scalar multiplication as for X, then A is a *vector subspace* of X.

Example 1. For any $w \in \mathbb{R}^N$, the line through w and the origin is a vector subspace of $\mathbb{R}^N$. In this case $X = \{x \in \mathbb{R}^N : \exists a \in \mathbb{R} \text{ s.t. } x = aw\}$. In particular, the x_n axis is a vector subspace of $\mathbb{R}^N$ for any n. □

Example 2. $\mathbb{R}^N_+$ is not a vector subspace of $\mathbb{R}^N$. In particular, no point in $\mathbb{R}^N_+$ other than the origin has an additive inverse. □

3 General Norms.

Definition 3. *A function $f : X \to \mathbb{R}$ is a* norm *on X iff for any $x, y \in X$ and any $a \in \mathbb{R}$,*

1. *$f(x) \geq 0$ and $f(x) = 0$ iff $x = 0$,*
2. *$f(ax) = |a|f(x)$,*
3. *$f(x + y) \leq f(x) + f(y)$. This property is called the* triangle inequality.

For $\mathbb{R}^N$, I have already discussed one norm, the Euclidean norm. There are other possible norms for $\mathbb{R}^N$, infinitely many in fact, but the Euclidean norm is the default: unless specified explicitly otherwise, the norm in $\mathbb{R}^N$ is understood to be the Euclidean norm. In other vector spaces, there is often no default norm, and one has to be explicit as to which norm one is using.

Remark 4. If X is a vector space, say that a function $f : X \to \mathbb{R}$ is *convex* iff for any $x, y \in X$, and any $a \in [0, 1]$, $f(ax + (1-a)y) \leq af(x) + (1-a)f(y)$. A simple example of a convex function is the standard norm on $\mathbb{R}$, namely absolute value: $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = |x|$.

More generally, any norm is convex: for any $x, y \in X$, and any $a \in [0, 1]$, if f is a norm then $f(ax + (1-a)y) \leq f(ax) + f((1-a)y) = af(x) + (1-a)f(y)$, where the inequality is from the third property of norms (the triangle inequality) and the equality is from the second property of norms. □

If X is a normed vector space and A is a vector subspace of X, then, unless I state explicitly otherwise, I assume that A inherits the norm of X, with domain restricted to A. A is then a *normed vector subspace* of X.

4 Examples.

4.1 The max norm on $\mathbb{R}^N$.

For any $x \in \mathbb{R}^N$, define

$$\|x\|_{\max} = \max_n |x_n|.$$

Thus, for example, if $x = (2, -7)$ then $\|x\|_{\max} = 7$.

Theorem 3. *$\|\cdot\|_{\max}$ is a norm on $\mathbb{R}^N$.*

Proof. Of the three norm properties, only the triangle inequality is non-trivial. Consider any $x, y \in \mathbb{R}^N$. For each n, $|x_n + y_n| \leq |x_n| + |y_n|$ (this is just the Euclidean triangle inequality for $N = 1$). And, for each n, $|x_n| \leq \max_m |x_m|$ (where I've changed the index on the right from n to m to avoid confusion) and $|y_n| \leq \max_m |y_m|$. Thus, for all n, $|x_n + y_n| \leq \max_m |x_m| + \max_m |y_m|$. Hence $\max_n |x_n + y_n| \leq \max_m |x_m| + \max_m |y_m|$, from which the result follows. ■

Given any norm f, one can form a new norm g by multiplying by a constant $c > 0$: $g(x) = cf(x)$. This amounts to a change in units (e.g., yards versus meters). There are an infinity of norms that differ in this trivial sense.

The Euclidean norm and the max norm are different in a more substantive sense. Consider the points $x = (100, 0)$ and $y = (90, 90)$. Then $\|x\|_{\max} = 100 > \|y\|_{\max} = 90$. But under the Euclidean norm, $\|x\| = 100 < \|y\| = 90\sqrt{2} \approx 127$. Thus, the Euclidean norm and the max norm give different answers to the question, "which point is closer to the origin, x or y?" The Euclidean and max norms are, however, nevertheless similar in another sense, discussed below in the section on norm equivalence.

Remark 5. The min function $f : \mathbb{R}^N \to \mathbb{R}$, $f(x) = \min_n |x_n|$, is not a norm. For example if $x = (1, 0)$ and $y = (0, 1)$ then $f(x) = f(y) = 0$ but $f(x + y) = 1$. □

4.2 The sup norm on ℓ^∞.

Recall that $\mathbb{R}^\omega$ is the set of elements of the form $(x_1, x_2, \dots)$, with $x_n \in \mathbb{R}$ for all n.[2] As discussed in the Set Theory notes, I can construct $\mathbb{R}^\omega$ formally by identifying it with the set of functions from $\{1, 2, \dots\}$ to $\mathbb{R}$: if g is such a function, then $x_1 = g(1)$, $x_2 = g(2)$ and so forth. Make $\mathbb{R}^\omega$ a vector space by defining vector addition and scalar multiplication for $\mathbb{R}^\omega$ as for $\mathbb{R}^N$.

For $\mathbb{R}^\omega$, the natural analog of the max norm is the sup norm. A problem, however, is that sup is not defined (or is defined to be infinite) for points such as $x = (1, 2, 3, \dots)$. For this reason, restrict attention to the set of x in $\mathbb{R}^\omega$ that are bounded in the sense that all coordinates of x have absolute values less than some real number M, where M is allowed to vary with x. This subset of $\mathbb{R}^\omega$ is called ℓ^∞. Formally,

$$\ell^\infty = \{x \in \mathbb{R}^\omega : \exists M \in \mathbb{R} \text{ s.t. } \forall n, |x_n| < M\}.$$

Thus $(x_1, 0, \dots) \in \ell^\infty$ no matter how large x_1 is (just take $M > x_1$) but $(1, 2, 3, \dots) \notin \ell^\infty$.

ℓ^∞ is a vector subspace of $\mathbb{R}^\omega$. Although this may be obvious, in infinite dimensional settings such as this it can be important to check vector space properties explicitly. In particular, then, note that ℓ_∞ is closed under vector addition (i.e., if $x, y \in \ell^\infty$ then $x + y \in \ell^\infty$) since if M_x is a bound for x and M_y is a bound for y then for any n, $|x_n + y_n| \leq |x_n| + |y_n| < M_x + M_y$, and hence $M_x + M_y$ is a bound for $x + y$. Similarly, ℓ^∞ is closed under scalar multiplication.

By the LUB property, for any $x \in \ell^\infty$,

$$\|x\|_{\text{sup}} = \sup_n |x_n|$$

is well defined. For example, if $x = (3, 3.1, 3.14, \dots)$ then $\|x\|_{\text{sup}} = \pi$. $\|\cdot\|_{\text{sup}}$ is called the sup norm.

Theorem 4. *$\|\cdot\|_{\text{sup}}$ is a norm on ℓ^∞.*

Proof. Of the three norm properties, only the triangle inequality is non-trivial. The proof is almost the same as the proof for $\|\cdot\|_{\max}$. Consider any $x, y \in \ell^\infty$. For each n, $|x_n + y_n| \leq |x_n| + |y_n|$ (this is just the Euclidean triangle inequality for $N = 1$). And, for each n, $|x_n| \leq \sup_m |x_m|$ (where I've changed the index on

[2] $\mathbb{R}^\infty$ might seem to be more natural notation, but $\mathbb{R}^\infty$ is standard notation for a different vector space, namely the vector subspace of $\mathbb{R}^\omega$ consisting of points that are zero in all but a finite number of coordinates. $\mathbb{R}^\infty$ can be thought of as the union of all possible $\mathbb{R}^N$.

the right from n to m to avoid confusion) and $|y_n| \leq \sup_m |y_m|$. Thus, for all n, $|x_n + y_n| \leq \sup_m |x_m| + \sup_m |y_m|$. Hence $\sup_n |x_n + y_n| \leq \sup_m |x_m| + \sup_m |y_m|$, from which the result follows. ■

4.3 Other examples.

There are many other important examples of vector spaces and norms. I discuss two briefly, to give a flavor of what is possible.

1. Let ℓ^2 be the subset of $\mathbb{R}^\omega$ that is square summable: if $x \in \mathbb{R}^\omega$ then $x \in \ell^2$ iff the infinite sum $\sum_{n=1}^{\infty} x_n^2$ is finite.[3] For example, $(1, 1/2, 1/3, \dots) \in \ell^2$ but $(1, 1, 1, \dots) \notin \ell^2$. It should be evident, in particular, that ℓ^2 is a proper subset of ℓ^∞.

   One can show that ℓ^2 is a vector space. One can also show that for any $x, y \in \ell^2$ there is a well-defined inner product given by

   $$x \cdot y = \sum_{n=1}^{\infty} x_n y_n.$$

   Finally, one can then define the ℓ^2 norm by

   $$\|x\|_2 = \sqrt{x \cdot x}.$$

   The proof that the ℓ^2 norm is indeed a norm is identical to the proof that the Euclidean norm is a norm.

   The vector space ℓ^2 with the ℓ^2 norm is an infinite dimensional analog of $\mathbb{R}^N$ with the Euclidean norm. Note that the ℓ^2 norm is not well defined for ℓ^∞. For example, the ℓ^2 norm of the point $(1, 1, 1, \dots)$ is undefined (or infinite, if you prefer). There is no natural analog of the Euclidean norm on ℓ^∞.

2. Let $C[0, 1]$ be the set of continuous functions from $[0, 1]$ to $\mathbb{R}$. I define continuity formally later in the course; here I assume that you have an informal understanding of what continuity means. For $f, g \in C[0, 1]$, define $h = f + g$ to be $h(x) = f(x) + g(x)$ for all $x \in [0, 1]$. For $f \in C[0, 1]$ and $a \in \mathbb{R}$, define $h = af$ to be $h(x) = af(x)$ for all $x \in [0, 1]$. Then one can show that $C[0, 1]$ is a vector space (for example, one can prove that the sum of continuous functions is continuous).

[3] I have not defined what it means for an infinite sum to be finite. For present purposes, the following suffices. Since $x_n^2 \geq 0$, the partial sum $S_N = \sum_{n=1}^{N} x_n^2$ is weakly increasing in N. If there is an M such that $S_N < M$ for all N, then by the LUB property, $\sup_N \{S_N\}$ is well defined. In this case, $\sum_{n=1}^{\infty} x_n^2$ is said to be finite and equals $\sup_N \{S_N\}$. Otherwise the infinite sum is either undefined or said to be infinite.

One obvious norm for $C[0,1]$ is the sup norm: $\|f\|_{\sup} = \sup_x |f(x)|$. This (candidate) norm is well defined thanks to a basic fact, proved later in the course, that any continuous function defined on $[0,1]$ is bounded. The proof that $\|\cdot\|_{\sup}$ is indeed a norm is then almost the same as for ℓ^∞.

5 Finite-Dimensional Vector Spaces.

Let X be a vector space. Given a *finite* set $Z \subseteq X$, the *span* of Z is the set A such that for every $x \in A$ there are numbers $\lambda_z \in \mathbb{R}$, one for each $z \in Z$, such that $x = \sum_{z \in Z} \lambda_z z$: x is a weighted sum of elements of Z. Z is said to span A. It is easy to verify that if $Z \subseteq X$ spans A then A is a vector subspace of X.

$A \subseteq X$ is *finite dimensional* iff it can be spanned by a finite subset of X. In particular, X is finite dimensional iff it can be spanned by a finite set of its vectors. If X is not finite dimensional then it is *infinite dimensional.* ℓ^∞ is an example of an infinite dimensional vector space.

Suppose that a vector space X is spanned by the finite set Z. Z is a *basis* for X iff no set with lower cardinality spans X. If $Z = \{z^1, \dots, z^N\}$ is a basis for X then the dimension of X is said to be N and, for any $x \in X$, the representation

$$x = \sum_{n=1}^{N} \lambda_n z^n$$

is unique. To see the latter, I argue by contraposition. Suppose that for some $x \in X$,

$$x = \sum_{n=1}^{N} \lambda_n z^n = \sum_{n=1}^{N} \hat{\lambda}_n z^n.$$

with $\lambda = (\lambda_1, \dots, \lambda_N) \in \mathbb{R}^N$ not equal to $\hat{\lambda} = (\hat{\lambda}_1, \dots, \hat{\lambda}_N) \in \mathbb{R}^N$. Then,

$$0 = \sum_{n=1}^{N} (\lambda_n - \hat{\lambda}_n) z^n.$$

If $\lambda \neq \hat{\lambda}$ then $\lambda_n - \hat{\lambda}_n \neq 0$ for at least one n; for notational convenience (alternatively, by reindexing the z^n), suppose this occurs for $n = 1$. Then,

$$z^1 = -\sum_{n=2}^{N} \frac{\lambda_n - \hat{\lambda}_n}{\lambda_1 - \hat{\lambda}_1} z^n.$$

This implies $Z \setminus \{z^1\}$ spans X, so that Z was not a basis.

A finite-dimensional vector space with dimension greater than 1 will have an infinity of bases (for one thing, if $Z = \{z^1, \dots, z^N\}$ is a basis, then $\{2z^1, \dots, 2z^N\}$ is also a basis) and there is, in general, no default or standard basis. But $\mathbb{R}^N$ does have a standard basis, with typical element $z^n = (0, \dots, 0, 1, 0, \dots, 0)$, where the 1 appears in coordinate n.

For a general finite-dimensional vector space X with basis Z, on can define the max norm as before, with

$$\|x\|_{\max} = \max_n |\lambda_n|$$

where $x = \sum_n \lambda_n z^n$. The proof that the max norm is a norm is the same as in the earlier proof that the max norm on $\mathbb{R}^N$ is a norm (Theorem 3), so I won't provide it as a separate theorem.

The following is a simple example of a finite-dimensional vector space that is not $\mathbb{R}^N$ or some vector subspace of $\mathbb{R}^N$.

Example 3. Let X be the set of polynomials on $\mathbb{R}$ with degree at most $N-1$: an element $P \in X$ is a function $P : \mathbb{R} \to \mathbb{R}$ defined by,

$$P(b) = \lambda_1 + \lambda_2 b + \cdots + \lambda_N b^{N-1},$$

where $\lambda_n \in \mathbb{R}$ for each n. It is easy to verify that X is a vector space and that a basis for X is the set of the first N monomials $z^n : \mathbb{R} \to \mathbb{R}$ given by, for $b \in \mathbb{R}$, $z^n(b) = b^n$; I include here the trivial monomial z^0 given by $z^0(b) = 1$. With this basis, any $P \in X$ is characterized by the associated vector $(\lambda_1, \dots, \lambda_N) \in \mathbb{R}^N$. □

As noted in this last example, if X is the vector space of polynomials on $\mathbb{R}$ with degree at most $N-1$, then, using the first N monomials as a basis, any element of X is uniquely identified with an element of $\mathbb{R}^N$, namely the element of $\mathbb{R}^N$ giving the polynomial weights, and vice versa. Moreover, if we endow both X and $\mathbb{R}^N$ with the max norm, then norm is preserved: for any $P \in X$, if $\lambda \in \mathbb{R}^N$ is the associated vector of polynomial weights, then $\|P\|_{\max} = \|\lambda\|_{\max}$, and vice versa for any $\lambda \in \mathbb{R}^N$. We say that X and $\mathbb{R}^N$, both with the max norm, are *isometrically isomorphic*. At an abstract level, there isn't really any difference between X and $\mathbb{R}^N$: they are just two different interpretations of the same underlying mathematical object. This can easily be shown to be general, although I won't do so explicitly; in particular, any finite-dimensional vector space of dimension N with the max norm is isometrically isomorphic to $\mathbb{R}^N$.

6 Equivalent Norms.

Definition 4. *Let X be a vector space. Two norms, f and g, are* equivalent *iff there are $0 < a \leq b$ such that, for any $x \in X$,*

$$ag(x) \leq f(x) \leq bg(x).$$

For any norm f and any $r \in \mathbb{R}$, $r > 0$, consider the set of points with norm less than or equal to r. In $\mathbb{R}^3$ with the Euclidean norm, this set is a solid ball, centered at the origin, of radius r. Call this set a *closed f-norm ball of radius r*.

If f and g are equivalent then the closed f- and g-norm balls are nested, like matryoshka dolls. In particular, for any $r > 0$, a closed g-norm ball of radius r/b is contained inside a closed f-norm ball of radius r, which is contained inside a closed g-norm ball of radius r/a.

Theorem 5. *In $\mathbb{R}^N$, the max norm and the Euclidean norm are equivalent.*

Proof. For any $x \in X$, $\|x\|^2 = \sum_n x_n^2 \geq \max_n x_n^2$, hence $\|x\| \geq \max_n |x_n| = \|x\|_{\max}$.

On the other hand, for any $x \in X$,

$$\begin{aligned}\|x\| &= \sqrt{\sum_n x_n^2} \\ &\leq \sqrt{N \max_n x_n^2} \\ &= \sqrt{N} \max_n |x_n| \\ &= \sqrt{N}\|x\|_{\max}.\end{aligned}$$

Combining, for any $x \in X$,

$$\|x\|_{\max} \leq \|x\| \leq \sqrt{N}\|x\|_{\max},$$

which shows that the max norm and the Euclidean norm are equivalent. (In the definition of equivalence, let f be the Euclidean norm, g be the max norm, and let $a = 1$ and $b = \sqrt{N}$.) ■

Theorem 5 is a special case of the fact, proved in the notes on Existence of Optima, that in any finite-dimensional vector space, all norms are equivalent.

In infinite-dimensional vectors spaces, on the other hand, norms need not be equivalent.

Example 4. For ℓ^∞, consider the norm pw* defined by

$$\|x\|_{\text{pw}^*} = \sup_n \frac{|x_n|}{n}.$$

(This norm has no standard name. I am calling it pw* because it is a variation on another bit of mathematical machinery that I will be introducing later and that I will call the pw metric.)

It is not hard to verify that the pw* norm is defined for all $x \in \ell^\infty$ and satisfies the norm properties.

The pw* and sup norms are not equivalent. In particular, consider points of the form $(1, 0, 0, \dots)$, $(0, 2, 0, 0, \dots)$, $(0, 0, 3, 0, 0, \dots)$, and so on. These points are all contained in the pw*-norm ball of radius 2. But no sup-norm ball of any radius, no matter how large, contains all of these points. □