Economics 204 Summer/Fall 2010
Lecture 3–Wednesday July 28, 2010

Section 2.1. Metric Spaces and Normed Spaces

Here we seek to generalize notions of distance and length in $\mathbf{R}^n$ to abstract settings.

Definition 1 A *metric space* is a pair (X, d), where X is a set and $d : X \times X \to \mathbf{R}_+$ a function satisfying

1. $d(x, y) \geq 0,\ d(x, y) = 0 \Leftrightarrow x = y\ \forall x, y \in X$
2. $d(x, y) = d(y, x)\ \forall x, y \in X$
3. *triangle inequality*:
$$d(x, z) \leq d(x, y) + d(y, z)\ \ \forall x, y, z \in X$$
$$\begin{array}{ccccc} & & y & & \\ & \nearrow & & \searrow & \\ x & & \rightarrow & & z \end{array}$$

A function $d : X \times X \to \mathbf{R}_+$ satisfying 1-3 is called a *metric* on X.

A metric gives a notion of distance between elements of X.

Definition 2 Let V be a vector space over $\mathbf{R}$. A *norm* on V is a function $\| \cdot \| : V \to \mathbf{R}_+$ satisfying

1. $\|x\| \geq 0\ \forall x \in V$
2. $\|x\| = 0 \Leftrightarrow x = 0\ \forall x \in V$
3. *triangle inequality*:
$$\|x + y\| \leq \|x\| + \|y\|\ \forall x, y \in V$$
$$\begin{array}{ccc} & x & \\ x \nearrow & & \searrow y \\ 0 & \rightarrow & x + y \\ y \searrow & & \nearrow x \\ & y & \end{array}$$
4. $\|\alpha x\| = |\alpha| \|x\|\ \forall \alpha \in \mathbf{R}, x \in V$

A *normed vector space* is a vector space over $\mathbf{R}$ equipped with a norm.

A norm gives a notion of length of a vector in V.

Example: In $\mathbf{R}^n$, the standard notion of distance between two vectors x and y measures the length of the difference $x-y$, i.e., $d(x,y)=\|x-y\|=\sqrt{\sum_{i=1}^n (x_i-y_i)^2}$.

In an abstract normed vector space, the norm can be used analogously to define a notion of distance.

Theorem 3 *Let $(V,\|\cdot\|)$ be a normed vector space. Let $d: V\times V \Rightarrow \mathbf{R}_+$ be defined by*

$$d(v,w)=\|v-w\|$$

Then (V,d) is a metric space.

Proof: We must verify that d satisfies all the properties of a metric.

1. Let $v,w\in V$. Then by definition, $d(v,w)=\|v-w\|\geq 0$ (why?), and

$$\begin{aligned} d(v,w)=0 &\Leftrightarrow \|v-w\|=0 \\ &\Leftrightarrow v-w=0 \\ &\Leftrightarrow (v+(-w))+w=w \\ &\Leftrightarrow v+((-w)+w)=w \\ &\Leftrightarrow v+0=w \\ &\Leftrightarrow v=w \end{aligned}$$

2. First, note that for any $x\in V$, $0\cdot x=(0+0)\cdot x=0\cdot x+0\cdot x$, so $0\cdot x=0$. Then $0=0\cdot x=(1-1)\cdot x=1\cdot x+(-1)\cdot x=x+(-1)\cdot x$, so we have $(-1)\cdot x=(-x)$. Then let $v,w\in V$.

$$\begin{aligned} d(v,w) &= \|v-w\| \\ &= |-1|\|v-w\| \\ &= \|(-1)(v+(-w))\| \\ &= \|(-1)v+(-1)(-w)\| \\ &= \|-v+w\| \\ &= \|w+(-v)\| \\ &= \|w-v\| \\ &= d(w,v) \end{aligned}$$

3. Let $u,w,v\in V$.

$$\begin{aligned} d(u,w) &= \|u-w\| \\ &= \|u+(-v+v)-w\| \\ &= \|u-v+v-w\| \\ &\leq \|u-v\|+\|v-w\| \\ &= d(u,v)+d(v,w) \end{aligned}$$

Thus d is a metric on V. ∎

Examples of Normed Vector Spaces

- $\mathbf{E}^n$: n-dimensional Euclidean space.

$$V = \mathbf{R}^n, \ \|x\|_2 = |x| = \sqrt{\sum_{i=1}^{n} (x_i)^2}$$

- $V = \mathbf{R}^n, \ \|x\|_1 = \sum_{i=1}^{n} |x_i|$ (the "taxi cab" norm or L^1 norm)
- $V = \mathbf{R}^n, \ \|x\|_\infty = \max\{|x_1|, \ldots, |x_n|\}$ (the maximum norm, or sup norm, or L^∞ norm)
- $C([0,1]), \ \|f\|_\infty = \sup\{|f(t)| : t \in [0,1]\}$
- $C([0,1]), \ \|f\|_2 = \sqrt{\int_0^1 (f(t))^2 \, dt}$
- $C([0,1]), \ \|f\|_1 = \int_0^1 |f(t)| \, dt$

Theorem 4 (Cauchy-Schwarz Inequality)
If $v, w \in \mathbf{R}^n$, then

$$\begin{aligned} \left(\sum_{i=1}^{n} v_i w_i \right)^2 &\leq \left(\sum_{i=1}^{n} v_i^2 \right) \left(\sum_{i=1}^{n} w_i^2 \right) \\ |v \cdot w|^2 &\leq |v|^2 |w|^2 \\ |v \cdot w| &\leq |v||w| \end{aligned}$$

Proof: Read the proof in de La Fuente. ∎

The Cauchy-Schwarz Inequality is essential in proving the triangle inequality in $\mathbf{E}^n$. Deriving the triangle inequality in $\mathbf{E}^n$ from the Cauchy-Schwarz inequality is a good exercise. The Cauchy-Schwarz inequality can also be viewed as a consequence of geometry in $\mathbf{R}^2$, in particular the law of cosines. Note that for $v, w \in \mathbf{R}^2$, $v \cdot w = |v||w| \cos\theta$ where θ is the angle between v and w; see Figure 1.[1]

Notice that a given vector space may have many different norms. As a trivial example, if $\|\cdot\|$ is a norm on a vector space V, so are $2\|\cdot\|$ and $3\|\cdot\|$ and $k\|\cdot\|$ for any $k > 0$. Less trivially, $\mathbf{R}^n$ supports many different norms as in the examples above. Different norms on a given vector space yield different geometric properties; for example, see Figure 2 for different norms on $\mathbf{R}^2$.

[1] From the law of cosines, $(v-w)\cdot(v-w) = v\cdot v + w\cdot w - 2|v||w|\cos\theta$. On the other hand, $(v-w)\cdot(v-w) = v\cdot v - 2v\cdot w + w\cdot w$, so $v \cdot w = |v||w|\cos\theta$.

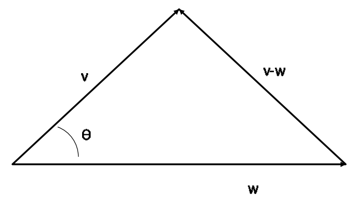

Figure 1: θ is the angle between v and w.

Definition 5 Two norms $\|\cdot\|$ and $\|\cdot\|^*$ on the same vector space V are said to be *Lipschitz-equivalent* (or *equivalent*) if $\exists m, M > 0$ s.t. $\forall x \in V$,

$$m\|x\| \leq \|x\|^* \leq M\|x\|$$

Equivalently, $\exists m, M > 0$ s.t. $\forall x \in V, x \neq 0$,

$$m \leq \frac{\|x\|^*}{\|x\|} \leq M$$

If two norms are equivalent, then they define the same notions of convergence and continuity. For topological purposes, equivalent norms are indistinguishable. For example, suppose two norms $\|\cdot\|$ and $\|\cdot\|^*$ on the vector space V are equivalent, and fix $x \in V$. Let $B_\varepsilon(x, \|\cdot\|)$ denote the $\|\cdot\|$-ball of radius ε about x; similarly, let $B_\varepsilon(x, \|\cdot\|^*)$ denote the $\|\cdot\|^*$-ball of radius ε about x. That is,

$$\begin{aligned} B_\varepsilon(x, \|\cdot\|) &= \{y \in V : \|x - y\| < \varepsilon\} \\ B_\varepsilon(x, \|\cdot\|^*) &= \{y \in V : \|x - y\|^* < \varepsilon\} \end{aligned}$$

Then for any $\varepsilon > 0$,

$$B_{\frac{\varepsilon}{M}}(x, \|\cdot\|) \subseteq B_\varepsilon(x, \|\cdot\|^*) \subseteq B_{\frac{\varepsilon}{m}}(x, \|\cdot\|)$$

See Figure 3.

In $\mathbf{R}^n$ (or any finite-dimensional normed vector space), all norms are equivalent. This says roughly that, up to a difference in scaling, for topological purposes there is a unique norm in $\mathbf{R}^n$.

Theorem 6 *All norms on $\mathbf{R}^n$ are equivalent.*[2]

[2] The statement of the theorem in de la Fuente (Theorem 10.8, p. 107) is correct, but the proof has a problem.

However, infinite-dimensional spaces support norms that are not equivalent. For example, on $C([0,1])$, let f_n be the function

$$f_n(t) = \begin{cases} 1 - nt & \text{if } t \in \left[0, \frac{1}{n}\right] \\ 0 & \text{if } t \in \left(\frac{1}{n}, 1\right] \end{cases}$$

Then

$$\frac{\|f_n\|_1}{\|f_n\|_\infty} = \frac{\frac{1}{2n}}{1} = \frac{1}{2n} \to 0$$

Definition 7 In a metric space (X,d), a subset $S \subseteq X$ is *bounded* if $\exists x \in X, \beta \in \mathbf{R}$ such that $\forall s \in S$, $d(s,x) \leq \beta$.

In a metric space (X,d), define

$$\begin{aligned} B_\varepsilon(x) &= \{y \in X : d(y,x) < \varepsilon\} \\ &= \textit{open ball with center } x \textit{ and radius } \varepsilon \\ B_\varepsilon[x] &= \{y \in X : d(y,x) \leq \varepsilon\} \\ &= \textit{closed ball with center } x \textit{ and radius } \varepsilon \end{aligned}$$

We can use the metric d to define a generalization of "radius". In a metric space (X,d), define the *diameter* of a subset $S \subseteq X$ by

$$\text{diam}\,(S) = \sup\{d(s,s') : s, s' \in S\}$$

Similarly, we can define the distance from a point to a set, and distance between sets, as follows:

$$\begin{aligned} d(A,x) &= \inf_{a \in A} d(a,x) \\ d(A,B) &= \inf_{a \in A} d(B,a) \\ &= \inf\{d(a,b) : a \in A, b \in B\} \end{aligned}$$

Note that $d(A,x)$ cannot be a metric (since a metric is a function on $X \times X$, the first and second arguments must be objects of the same type); in addition, $d(A,B)$ does not define a metric on the space of subsets of X (why?).[3]

Section 2.2. Convergence of Sequences in Metric Spaces

Definition 8 Let (X,d) be a metric space. A sequence $\{x_n\}$ *converges* to x (written $x_n \to x$ or $\lim_{n\to\infty} x_n = x$) if

$$\forall \varepsilon > 0 \ \exists N(\varepsilon) \in \mathbf{N} \text{ s.t. } n > N(\varepsilon) \Rightarrow d(x_n, x) < \varepsilon$$

[3]Another, more useful notion of the distance between sets is the Hausdorff distance, given by $d(A,B) = \max\{\sup_{a\in A} \inf_{b\in B} d(a,b), \sup_{b\in B} \inf_{a\in A} d(a,b)\}$.

Notice that this is exactly the same as the definition of convergence of a sequence of real numbers, except we replace the standard measure of distance $|\cdot|$ in $\mathbf{R}$ by the general metric d.

Theorem 9 (Uniqueness of Limits) *In a metric space (X, d), if $x_n \to x$ and $x_n \to x'$, then $x = x'$.*

$$\begin{array}{ccccccccc} & & & & & & \cdot x & & \\ & & & & & \cdot & \downarrow & & \varepsilon \\ & & x_n & & \cdot & & \downarrow & & \\ \cdot & \cdot & \cdot & \cdot & & & \asymp & \varepsilon = \frac{d(x,x')}{2} & \\ & & & & \cdot & & \uparrow & & \\ & & & & & \cdot & \uparrow & & \varepsilon \\ & & & & & & \cdot x' & & \end{array}$$

Proof: Suppose $\{x_n\}$ is a sequence in X, $x_n \to x$, $x_n \to x'$, $x \neq x'$. Since $x \neq x'$, $d(x, x') > 0$. Let

$$\varepsilon = \frac{d(x, x')}{2}$$

Then there exist $N(\varepsilon)$ and $N'(\varepsilon)$ such that

$$\begin{array}{rcl} n > N(\varepsilon) & \Rightarrow & d(x_n, x) < \varepsilon \\ n > N'(\varepsilon) & \Rightarrow & d(x_n, x') < \varepsilon \end{array}$$

Choose

$$n > \max\{N(\varepsilon), N'(\varepsilon)\}$$

Then

$$\begin{array}{rcl} d(x, x') & \leq & d(x, x_n) + d(x_n, x') \\ & < & \varepsilon + \varepsilon \\ & = & 2\varepsilon \\ & = & d(x, x') \\ d(x, x') & < & d(x, x') \end{array}$$

a contradiction.■

Definition 10 An element c is a *cluster point* of a sequence $\{x_n\}$ in a metric space (X, d) if $\forall \varepsilon > 0$, $\{n : x_n \in B_\varepsilon(c)\}$ is an infinite set. Equivalently,

$$\forall \varepsilon > 0, N \in \mathbf{N}\ \exists n > N \text{ s.t. } x_n \in B_\varepsilon(c)$$

Example:

$$x_n = \begin{cases} 1 - \frac{1}{n} & \text{if } n \text{ even} \\ \frac{1}{n} & \text{if } n \text{ odd} \end{cases}$$

For n large and odd, x_n is close to zero; for n large and even, x_n is close to one. The sequence does not converge; the set of cluster points is $\{0, 1\}$.

If $\{x_n\}$ is a sequence and $n_1 < n_2 < n_3 < \cdots$ then $\{x_{n_k}\}$ is called a *subsequence.*

Note that a subsequence is formed by taking some of the elements of the parent sequence, *in the same order.*

Example: $x_n = \frac{1}{n}$, so $\{x_n\} = \left(1, \frac{1}{2}, \frac{1}{3}, \ldots\right)$. If $n_k = 2k$, then $\{x_{n_k}\} = \left(\frac{1}{2}, \frac{1}{4}, \frac{1}{6}, \ldots\right)$.

Theorem 11 (2.4 in De La Fuente, plus ...) *Let (X, d) be a metric space, $c \in X$, and $\{x_n\}$ a sequence in X. Then c is a cluster point of $\{x_n\}$ if and only if there is a subsequence $\{x_{n_k}\}$ such that $\lim_{k\to\infty} x_{n_k} = c$.*

Proof: Suppose c is a cluster point of $\{x_n\}$. We inductively construct a subsequence that converges to c. For $k = 1$, $\{n : x_n \in B_1(c)\}$ is infinite, so nonempty; let

$$n_1 = \min\{n : x_n \in B_1(c)\}$$

Now, suppose we have chosen $n_1 < n_2 < \cdots < n_k$ such that

$$x_{n_j} \in B_{\frac{1}{j}}(c) \text{ for } j = 1, \ldots, k$$

$\{n : x_n \in B_{\frac{1}{k+1}}(c)\}$ is infinite, so it contains at least one element bigger than n_k, so let

$$n_{k+1} = \min\left\{n : n > n_k,\ x_n \in B_{\frac{1}{k+1}}(c)\right\}$$

Thus, we have chosen $n_1 < n_2 < \cdots < n_k < n_{k+1}$ such that

$$x_{n_j} \in B_{\frac{1}{j}}(c) \text{ for } j = 1, \ldots, k, k+1$$

Thus, by induction, we obtain a subsequence $\{x_{n_k}\}$ such that

$$x_{n_k} \in B_{\frac{1}{k}}(c)$$

Given any $\varepsilon > 0$, by the Archimedean property, there exists $N(\varepsilon) > 1/\varepsilon$.

$$\begin{aligned} k > N(\varepsilon) \quad &\Rightarrow \quad x_{n_k} \in B_{\frac{1}{k}}(c) \\ &\Rightarrow \quad x_{n_k} \in B_\varepsilon(c) \end{aligned}$$

so

$$x_{n_k} \to c \text{ as } k \to \infty$$

Conversely, suppose that there is a subsequence $\{x_{n_k}\}$ converging to c. Given any $\varepsilon > 0$, there exists $K \in \mathbf{N}$ such that

$$k > K \Rightarrow d(x_{n_k}, c) < \varepsilon \Rightarrow x_{n_k} \in B_\varepsilon(c)$$

Therefore,

$$\{n : x_n \in B_\varepsilon(c)\} \supseteq \{n_{K+1}, n_{K+2}, n_{K+3}, \ldots\}$$

Since $n_{K+1} < n_{K+2} < n_{K+3} < \cdots$, this set is infinite, so c is a cluster point of $\{x_n\}$. ■

Section 2.3. Sequences in R and R^m

Definition 12 A sequence of real number $\{x_n\}$ is *increasing* (*decreasing*) if $x_{n+1} \geq x_n$ ($x_{n+1} \leq x_n$) for all n.

Definition 13 If $\{x_n\}$ is a sequence of real numbers, $\{x_n\}$ *tends to infinity* (written $x_n \to \infty$ or $\lim x_n = \infty$) if

$$\forall K \in \mathbf{R}\ \exists N(K) \text{ s.t. } n > N(K) \Rightarrow x_n > K$$

Similarly define $x_n \to -\infty$ or $\lim x_n = -\infty$.

Notice we don't say the sequence *converges* to infinity; the term "converge" is limited to the case of finite limits.

Theorem 14 (Theorem 3.1') *Let $\{x_n\}$ be an increasing (decreasing) sequence of real numbers. Then $\lim_{n\to\infty} x_n = \sup\{x_n : n \in \mathbf{N}\}$ ($\lim_{n\to\infty} x_n = \inf\{x_n : n \in \mathbf{N}\}$). In particular, the limit exists.*

Proof: Read the proof in the book, and figure out how to handle the unbounded case. ■

Lim Sups and Lim Infs:[4]

Consider a sequence $\{x_n\}$ of real numbers. Let

$$\begin{aligned} \alpha_n &= \sup\{x_k : k \geq n\} \\ &= \sup\{x_n, x_{n+1}, x_{n+2}, \ldots\} \\ \beta_n &= \inf\{x_k : k \geq n\} \end{aligned}$$

Either $\alpha_n = +\infty$ for all n, or $\alpha_n \in \mathbf{R}$ and $\alpha_1 \geq \alpha_2 \geq \alpha_3 \geq \cdots$. Either $\beta_n = -\infty$ for all n, or $\beta_n \in \mathbf{R}$ and $\beta_1 \leq \beta_2 \leq \beta_3 \leq \cdots$.

[4] See the handout for this material.

Definition 15

$$\begin{aligned}
\limsup_{n\to\infty} x_n &= \begin{cases} +\infty & \text{if } \alpha_n = +\infty \text{ for all } n \\ \lim \alpha_n & \text{otherwise.} \end{cases} \\
\liminf_{n\to\infty} x_n &= \begin{cases} -\infty & \text{if } \beta_n = -\infty \text{ for all } n \\ \lim \beta_n & \text{otherwise.} \end{cases}
\end{aligned}$$

Theorem 16 *Let $\{x_n\}$ be a sequence of real numbers. Then*

$$\begin{gathered} \lim_{n\to\infty} x_n = \gamma \in \mathbf{R} \cup \{-\infty, \infty\} \\ \Leftrightarrow \limsup_{n\to\infty} x_n = \liminf_{n\to\infty} x_n = \gamma \end{gathered}$$

Theorem 17 (Theorem 3.2, Rising Sun Lemma) *Every sequence of real numbers contains an increasing subsequence or a decreasing subsequence or both.*

$$\begin{array}{ccccccccccccccccc}
 & & & \circ & \leftarrow & \leftarrow & \leftarrow & \leftarrow & \leftarrow & \leftarrow & \leftarrow & \leftarrow & \leftarrow & \leftarrow & \leftarrow & \leftarrow & \mathrm{S} \\
\bullet & & \bullet & & \bullet & & \bullet & \circ & \leftarrow & \leftarrow & \leftarrow & \leftarrow & \leftarrow & \leftarrow & \leftarrow & \leftarrow & \mathrm{U} \\
 & \bullet & & & & \bullet & & & \bullet & & \bullet & \bullet & \circ & \leftarrow & \leftarrow & \leftarrow & \mathrm{N} \\
 & & & & & & & & & \bullet & & & & \bullet & & \bullet & \\
 & & & & & & & & & & & & & & \bullet & &
\end{array}$$

Proof: Let

$$S = \{s \in \mathbf{N} : x_s > x_n \quad \forall n > s\}$$

Either S is infinite, or S is finite.

If S is infinite, let

$$\begin{aligned}
n_1 &= \min S \\
n_2 &= \min\left(S \setminus \{n_1\}\right) \\
n_3 &= \min\left(S \setminus \{n_1, n_2\}\right) \\
&\vdots \\
n_{k+1} &= \min\left(S \setminus \{n_1, n_2, \ldots, n_k\}\right)
\end{aligned}$$

Then $n_1 < n_2 < n_3 < \cdots$.

$$\begin{array}{rl}
x_{n_1} > x_{n_2} & \text{since } n_1 \in S \text{ and } n_2 > n_1 \\
x_{n_2} > x_{n_3} & \text{since } n_2 \in S \text{ and } n_3 > n_2 \\
 & \vdots \\
x_{n_k} > x_{n_{k+1}} & \text{since } n_k \in S \text{ and } n_{k+1} > n_k \\
 & \vdots
\end{array}$$

so $\{x_{n_k}\}$ is a strictly decreasing subsequence of $\{x_n\}$.

If S is finite and nonempty, let $n_1 = (\max S) + 1$; if $S = \emptyset$, let $n_1 = 1$. Then

$$\begin{array}{lll} n_1 \notin S & \text{so} & \exists n_2 > n_1 \text{ s.t. } x_{n_2} \geq x_{n_1} \\ n_2 \notin S & \text{so} & \exists n_3 > n_2 \text{ s.t. } x_{n_3} \geq x_{n_2} \\ & \vdots & \\ n_k \notin S & \text{so} & \exists n_{k+1} > n_k \text{ s.t. } x_{n_{k+1}} \geq x_{n_k} \\ & \vdots & \end{array}$$

so $\{x_{n_k}\}$ is a (weakly) increasing subsequence of $\{x_n\}$. ■

Theorem 18 (Thm. 3.3, Bolzano-Weierstrass) *Every bounded sequence of real numbers contains a convergent subsequence.*

Proof: Let $\{x_n\}$ be a bounded sequence of real numbers. By the Rising Sun Lemma, find an increasing or decreasing subsequence $\{x_{n_k}\}$. If $\{x_{n_k}\}$ is increasing, then by Theorem 3.1', $\lim x_{n_k} = \sup\{x_{n_k} : k \in \mathbf{N}\} \leq \sup\{x_n : n \in \mathbf{N}\} < \infty$, since the sequence is bounded; since the limit is finite, the subsequence converges. Similarly, if the subsequence is decreasing, it converges. ■

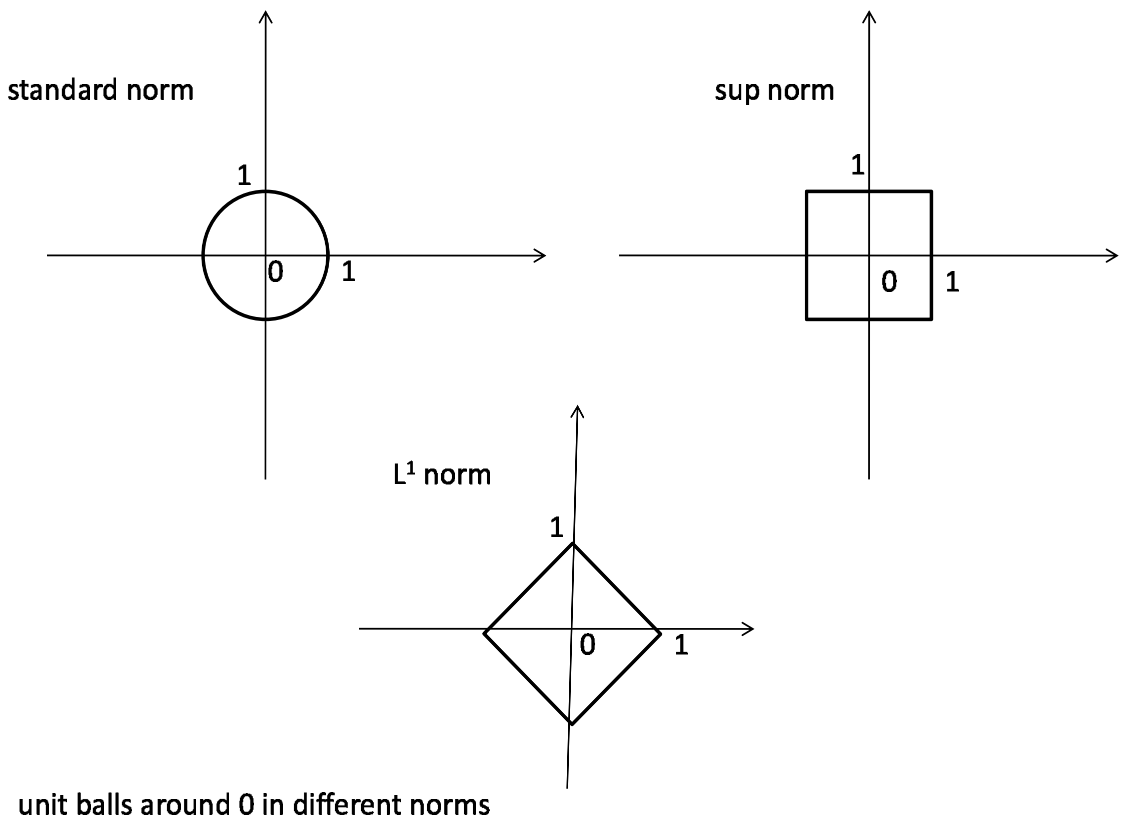

Figure 2: The unit ball around 0 in different norms on $\mathbf{R}^2$: standard $\|\cdot\|_2$, $\|\cdot\|_1$ (L^1 or taxi cab norm) and $\|\cdot\|_\infty$ (sup norm or L^∞ norm).

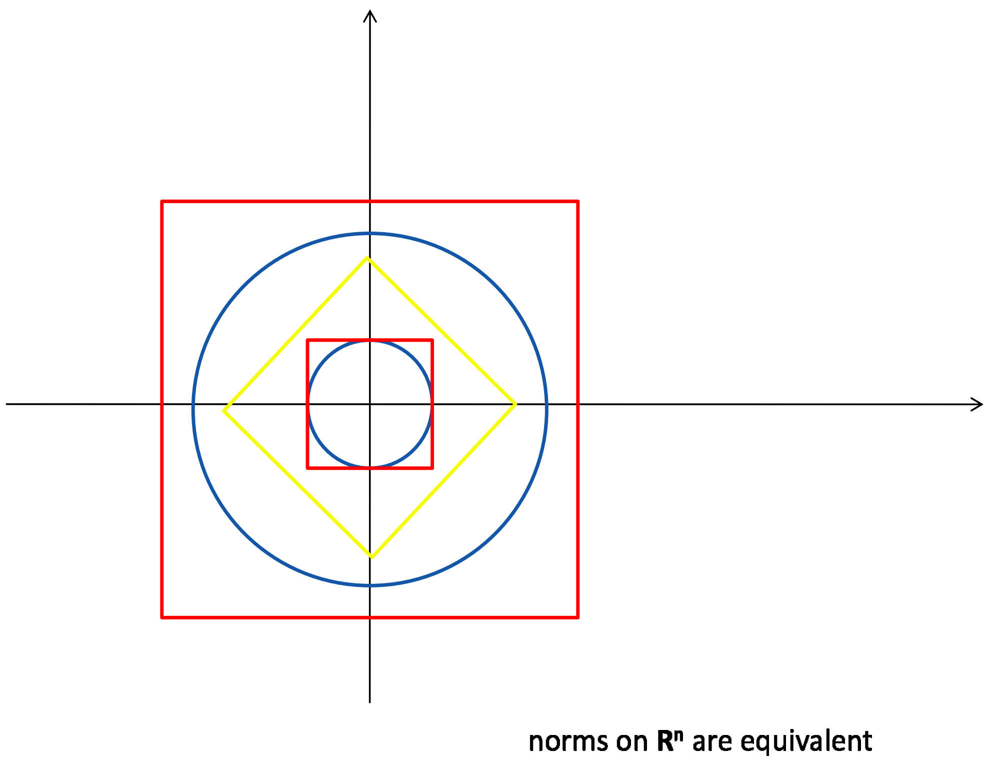

Figure 3: All norms on $\mathbf{R}^n$ are equivalent.