Normed Vector Spaces

A large and important class of metric spaces are also vector spaces over $\mathbb{R}$ or $\mathbb{C}$. Notice, for example, that $C[0,1]$ is a vector space. An easy way to build a metric on a vector space is by way of a length function or norm.

Def: A norm on a vector space V is a function $\|\cdot\|: V \to [0,\infty)$ satisfying

(i) $0 \le \|x\| < \infty$ for all $x \in V$

(ii) $\|x\| = 0$ if and only if $x = 0$

(iii) $\|\alpha x\| = |\alpha| \|x\|$ for any scalar α and any $x \in V$

(iv) the triangle inequality: $\|x+y\| \le \|x\| + \|y\|$ for all $x, y \in V$

A function $\|\cdot\|: V \to [0,\infty)$ satisfying all of the above properties except (ii) is called a pseudonorm; that is, a pseudonorm permits nonzero vectors to have 0 length.

The pair $(V, \|\cdot\|)$, consisting of a vector space V together with a norm on V, is called a normed vector space.

It is easy to see that any norm induces a metric on V by setting $d(x,y) = \|x-y\|$. We will refer to this particular metric as the usual metric on $(V, \|\cdot\|)$.

Ex.

(a) The absolute value function $|\cdot|$ clearly defines a norm on $\mathbb{R}$

(b) Each of the following defines a norm on $\mathbb{R}^n$:

$$\|x\|_1 = \sum_{i=1}^{n} |x_i| , \quad \|x\|_2 = \left(\sum_{i=1}^{n} |x_i|^2 \right)^{1/2}$$

and $\|x\|_\infty = \max_{1 \le i \le n} |x_i|$, where $x = (x_1, \dots, x_n) \in \mathbb{R}^n$.

The first and last expressions are very easy to check while the second takes a bit more work.

The function $\|\cdot\|_2$ is often called the Euclidean norm and is generally accepted as the norm of choice on $\mathbb{R}^n$. As it happens, for $1 \le p < \infty$, the expression $\|x\|_p = \left(\sum |x_i|^p \right)^{1/p}$ defines a norm on $\mathbb{R}^n$.

(c) Each of the following defines a norm on $C[a,b]$:

$$\|f\|_1 = \int_a^b |f(t)|\,dt , \quad \|f\|_2 = \left(\int_a^b |f(t)|^2 dt \right)^{1/2},$$

and $\|f\|_\infty = \max_{a \le t \le b} |f(t)|$

Again, the second expression is hardest to check. The last expression is generally taken as "the" norm on $C[a,b]$

(d) If $(V, \|\cdot\|)$ is a normed vector space, and if W is a linear subspace of V, then W is also normed by $\|\cdot\|$. That is, the restriction of $\|\cdot\|$ to W defines a norm on W.

(e) We might also consider the sequence space analogues of the "scale" of norms on $\mathbb{R}^n$ given in (b). For $1 \le p < \infty$, we define l_p to be the collection of all real sequences $x = \{x_n\}$ for which $\sum_{n=1}^{\infty} |x_n|^p < \infty$, and we define l_∞ to be the collection of all bounded real sequences. Each l_p is a vector space under "coordinatewise" addition and scalar multiplication. Moreover, the expression $\|x\|_p = \left(\sum |x_n|^p\right)^{1/p}$ if $1 \le p < \infty$ or $\|x\|_\infty = \sup_{n \in \mathbb{N}} |x_n|$ if $p = \infty$ defines a norm on l_p. The cases $p = 1$ and $p = \infty$ are easy to check. We will verify the other results shortly.

Lemma (The Cauchy-Schwarz Inequality):

$$\sum_{i=1}^{\infty} |x_i y_i| \le \|x\|_2 \, \|y\|_2 \quad \text{for any } x, y \in l_2$$

Proof: To simplify notation, write $\langle x, y \rangle = \sum x_i y_i$. We first consider the case where $x, y \in \mathbb{R}^n$ (that is, $x_i = 0 = y_i$ for all $i > n$). In this case, $\langle x, y \rangle$ is the usual "dot product" in $\mathbb{R}^n$. Also notice that we may suppose that $x, y \neq 0$. (There is nothing to show if either is 0). Now let $t \in \mathbb{R}$ and consider

$$0 \leq \|x+ty\|_2^2 = \langle x+ty, x+ty\rangle = \|x\|_2^2 + 2t\langle x,y\rangle + t^2\|y\|_2^2$$

Since this (nontrivial) quadratic in t is always nonnegative, it must have a nonpositive discriminant. (Why?) Thus, $(2\langle x,y\rangle)^2 - 4\|x\|_2^2\|y\|_2^2 \leq 0$ or, after simplifying, $|\langle x,y\rangle| \leq \|x\|_2\|y\|_2$. That is, $\left|\sum_{i=1}^n x_i y_i\right| \leq \|x\|_2\|y\|_2$.

Now this isn't quite what we wanted, but it actually implies the stronger inequality in the statement of the lemma. Why? Because the inequality that we have shown must also hold for the vectors $(|x_1|, |x_2|, \ldots, |x_n|)$ and $(|y_1|, |y_2|, \ldots, |y_n|)$. That is,

$$\sum_{i=1}^n |x_i||y_i| \leq \|(|x_1|, \ldots, |x_n|)\|\,\|(|y_1|, \ldots, |y_n|)\| = \|x\|_2\|y\|_2$$

Finally, let $x, y \in \ell_2$. Then for each n we have

$$\sum_{i=1}^n |x_i y_i| \leq \left(\sum_{i=1}^n |x_i|^2\right)^{1/2}\left(\sum_{i=1}^n |y_i|^2\right)^{1/2} \leq \|x\|_2\|y\|_2$$

Thus, $\sum_{i=1}^{\infty} x_i y_i$ must be absolutely convergent and satisfy $\sum_{i=1}^{\infty} |x_i y_i| \le \|x\|_2 \|y\|_2$.

Now we are ready to prove the triangle inequality for the ℓ_2-norm.

Thm (Minkowski's Inequality): If $x, y \in \ell_2$, then $x+y \in \ell_2$. Moreover, $\|x+y\|_2 \le \|x\|_2 + \|y\|_2$

Proof: It follows from the Cauchy-Schwarz inequality that for each n we have

$$\sum_{i=1}^{n} |x_i + y_i|^2 = \sum_{i=1}^{n} |x_i|^2 + 2 \sum_{i=1}^{n} x_i y_i + \sum_{i=1}^{n} |y_i|^2$$

$$\le \|x\|_2^2 + 2\|x\|_2 \|y\|_2 + \|y\|_2^2 = \left(\|x\|_2 + \|y\|_2\right)^2.$$

Thus, since n is arbitrary, we have $x+y \in \ell_2$ and $\|x+y\|_2 \le \|x\|_2 + \|y\|_2$.

We now proceed to show that $\|\cdot\|_p$ is a norm on ℓ_p.

Lemma: Let $1 < p < \infty$ and let $a, b \ge 0$. Then, $(a+b)^p \le 2^p (a^p + b^p)$. Consequently, $x+y \in \ell_p$ whenever $x, y \in \ell_p$.

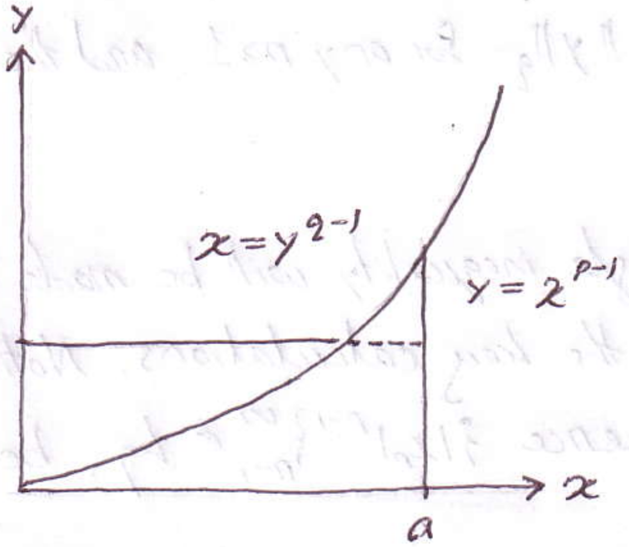
Proof: $(a+b)^p \le (2 \max\{a, b\})^p = 2^p \max\{a^p, b^p\} \le 2^p (a^p + b^p)$. Thus, if $x, y \in \ell_p$, then $\sum_{n=1}^{\infty} |x_n + y_n|^p \le$

$$2^p \sum_{n=1}^{\infty} |x_n|^p + 2^p \sum_{n=1}^{\infty} |y_n|^p < \infty.$$

Lemma (Young's Inequality): Let $1 < p < \infty$ and let q be defined by $\frac{1}{p} + \frac{1}{q} = 1$. Then, for any $a, b \geq 0$, we have $ab \leq a^p/p + b^q/q$, with equality occurring if and only if $a^{p-1} = b$.

Proof: Since the inequality trivially holds if either a or b is 0, we may suppose that $a, b > 0$. Since $\frac{1}{p} + \frac{1}{q} = 1$, we see that $p\left(\frac{1}{p} + \frac{1}{q}\right) = p$ or $1 + \frac{p}{q} = p$. In particular, $\frac{p}{q} = p - 1$. Notice that $q\left(\frac{1}{p} + \frac{1}{q}\right) = q$, implying that $\frac{q}{p} = q - 1$. Thus $\frac{1}{p-1} = \frac{1}{\left(\frac{p}{q}\right)} = \frac{q}{p} = q - 1$. Also, notice that $q = \frac{1}{p-1} + 1$, implying that, like p, $q \in (1, \infty)$. Thus, the functions $f(x) = x^{p-1}$ and $g(x) = x^{q-1}$ are inverses for $x \geq 0$.

The proof of the inequality follows from a comparison of areas (see figure below). The area of the rectangle with sides of lengths a and b is at most the sum of the areas under the graphs of the functions $y = x^{p-1}$ for $0 \leq x \leq a$ and $x = y^{q-1}$ for $0 \leq y \leq b$.



That is,

$$ab \leq \int_0^a x^{p-1}dx + \int_0^b y^{q-1}dy = \frac{a^p}{p} + \frac{b^q}{q}.$$

Clearly, equality can occur only if $a^{p-1} = b$.

We can now generalize the Cauchy-Schwarz Inequality.

Lemma: (Hölder's Inequality): Let $1 < p < \infty$ and let q be defined by $\frac{1}{p} + \frac{1}{q} = 1$. Given $x \in L_p$ and $y \in L_q$, we have $\sum_{i=1}^{\infty} |x_i y_i| \leq \|x\|_p \|y\|_q$.

Proof: We may suppose that $\|x\|_p > 0$ and $\|y\|_q > 0$ (since, otherwise, there is nothing to show). Now, for $n \geq 1$ we use Young's inequality to estimate:

$$\sum_{i=1}^{n} \left|\frac{x_i y_i}{\|x\|_p \|y\|_q}\right| \leq \frac{1}{p}\sum_{i=1}^{n}\left|\left(\frac{x_i}{\|x\|_p}\right)\right|^p + \frac{1}{q}\sum_{i=1}^{n}\left|\left(\frac{y_i}{\|y\|_q}\right)\right|^q \leq \frac{1}{p}+\frac{1}{q}=1$$

Thus, $\sum_{i=1}^{n} |x_i y_i| \leq \|x\|_p \|y\|_q$ for any $n \geq 1$, and the result follows.

Our proof of the triangle inequality will be made easier if we first isolate one of the key calculations. Notice that if $x \in l_p$, then the sequence $\{|x_n|^{p-1}\}_{n=1}^{\infty} \in l_q$, because $(p-1)q = p$. Moreover,

$$\|\{|x_i|^{p-1}\}\|_q = \left(\sum_{i=1}^{\infty} |x_i|^p\right)^{1/q} = \|x\|_p^{p-1}$$

Thm (Minkowski's Inequality): Let $1<p<\infty$. If $x, y \in l_p$, then $x+y \in l_p$ and $\|x+y\|_p \leq \|x\|_p + \|y\|_p$

Proof: We have already shown that $x+y \in l_p$. To prove the triangle inequality, we once again let q be defined by $\frac{1}{p}+\frac{1}{q}=1$, and we now use Hölder's inequality to estimate:

$$\sum_{i=1}^{\infty} |x_i+y_i|^p = \sum_{i=1}^{\infty} |x_i+y_i||x_i+y_i|^{p-1} \leq$$

$$\sum_{i=1}^{\infty} |x_i||x_i+y_i|^{p-1} + \sum_{i=1}^{\infty} |y_i||x_i+y_i|^{p-1} \leq$$

$$\|x\|_p \, \|\{|x_n+y_n|^{p-1}\}\|_q + \|y\|_p \, \|\{|x_n+y_n|^{p-1}\}\|_q =$$
$$= \|x+y\|_p^{p-1}\left(\|x\|_p + \|y\|_p\right)$$

That is, $\|x+y\|_p^p \leq \|x+y\|_p^{p-1}\left(\|x\|_p+\|y\|_p\right)$, and the triangle inequality follows.

Limits in Metric Spaces

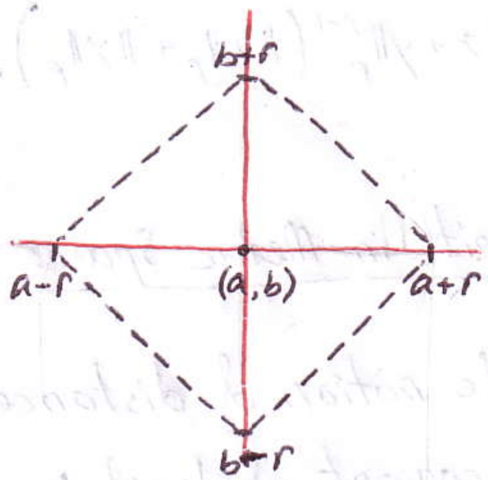
Having generalized the notion of distance, we are now ready to define the concept of limit in an abstract metric space. Throughout this section, unless otherwise specified, we will assume that we are always dealing with a generic metric space (M,d).

Def. Given $x \in M$ and $r > 0$, the set $B_r(x) = \{y \in M : d(x,y) < r\}$ is called the open ball about x of radius r. If we need to refer to the metric d, then we write $B_r^d(x)$.
We will sometimes denote by $C_r^d(x) = \{y \in M : d(x,y) \leq r\}$ the closed ball about x of radius r.

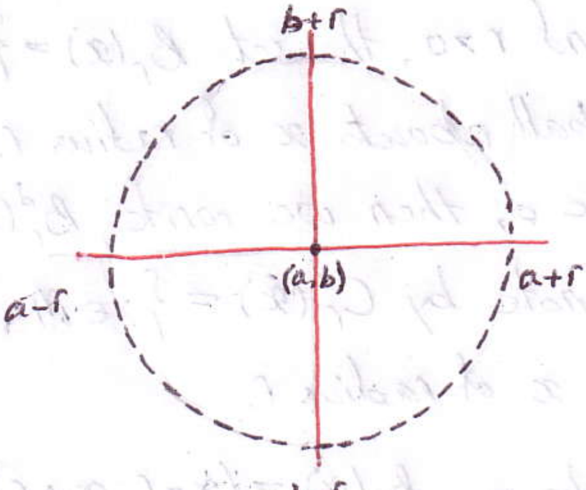
Ex. (a) In $\mathbb{R}$ we have $B_r(x) = (x-r, x+r)$, the open interval of radius r about x and $C_r(x) = [x-r, x+r]$, the closed interval of radius r about x.

(b) In $\mathbb{R}^2$, the set $B_r(x)$ is the open disk of radius r centered at x. The appearance of $B_r(x)$ depends on the metric at hand.

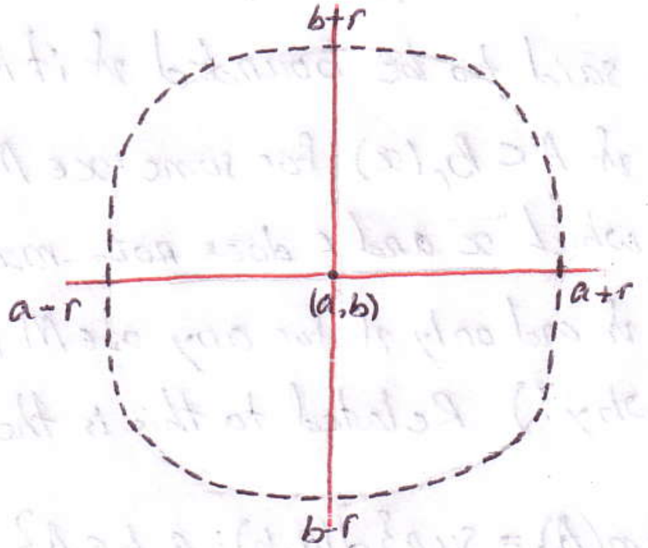
If d is generated by the norm $\| \|_1$, then $B_r(x)$ will look like a square of diameter $2r$ centered at $x=(a,b)$.



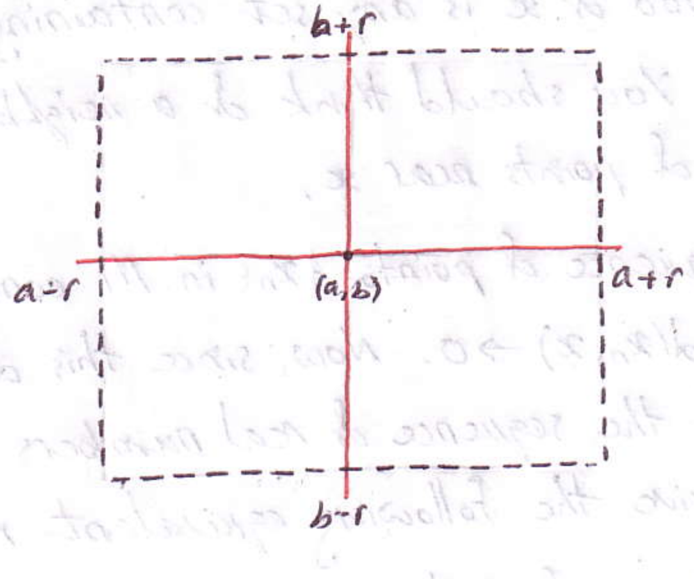
If d is generated by the norm $\|\cdot\|_2$, then $B_r(x)$ will look like a disc of radius r centered at $x=(a,b)$



If d is generated by the norm $\|\cdot\|_p$, $1<p<\infty$, then $B_r(x)$ will look like a brick with rounded corners. As p gets larger, the brick will assume the appearance of a regular square.



Finally, if d is generated by the norm $\|\cdot\|_\infty$, then $B_r(x)$ will look like a square with diameter $2\sqrt{2}\,r$, centered at $x=(a,b)$.



(c) In a discrete space $B_1(x) = \{x\}$ and $B_2(x) = M$.

(d) In a normed vector space $(V, \|\cdot\|)$ the balls centered at 0 play a special role: $B_r(x) = x + B_r(0) = \{y \in V;\ y = x + z,\ \|z\| < r\}$.

A subset A of M is said to be bounded if it is contained in some ball, that is, if $A \subset B_r(x)$ for some $x \in M$ and some $r > 0$. But exactly which x and r does not much matter. In fact, A is bounded if and only if for any $x \in M$ we have $\sup_{a \in A} d(x, a) < \infty$ (Why?) Related to this is the diameter of A, defined by $\text{diam}(A) = \sup\{d(a,b) : a, b \in A\}$. The diameter of A is a convenient measure of size because it does not refer to points outside of A.

Def: A neighborhood of x is any set containing an open ball about x. You should think of a neighborhood of x as a "thick" set of points near x.

We say that a sequence of points $\{x_n\}$ in M converges to a point $x \in M$ if $d(x_n, x) \to 0$. Now, since this definition is stated in terms of the sequence of real numbers $\{d(x_n, x)\}_{n=1}^{\infty}$, we can easily derive the following equivalent reformulations; $\{x_n\}$ converges to x if and only if, given any $\epsilon > 0$, there is an integer $N \geq 1$ such that $d(x_n, x) < \epsilon$ whenever $n \geq N$.

or $\{x_n\}$ converges to x if and only if, given any $\epsilon>0$, there is an integer $N\geq 1$ such that $\{x_n : n\geq N\}\subset B_\epsilon(x)$

If it should happen that $\{x_n : n\geq N\}\subset A$ for some N, we say that the sequence $\{x_n\}$ is eventually in A. Thus, our last formulation can be written

$\{x_n\}$ converges to x if and only if the sequence $\{x_n\}$ is eventually in every neighborhood of x.

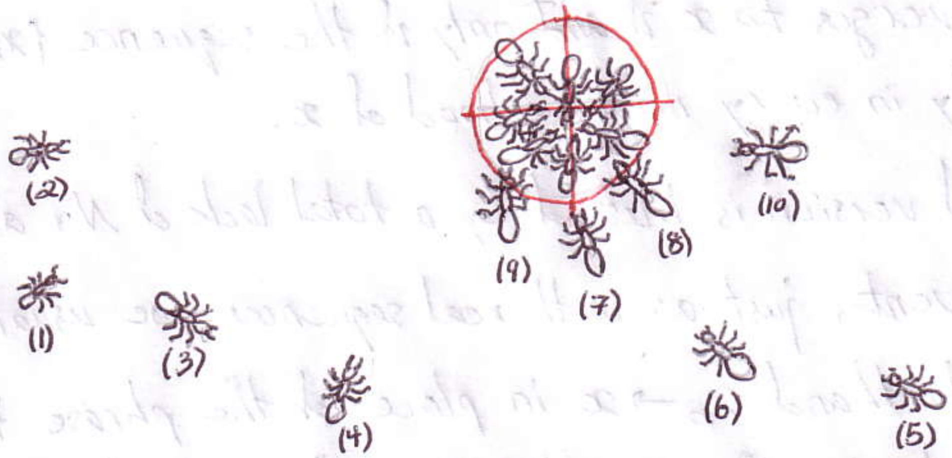
This final version is blessed by a total lack of N's and ϵ's!

In any event, just as with real sequences, we usually settle for the shorthand $x_n\to x$ in place of the phrase $\{x_n\}$ converges to x. On occasion we will want to display the set M, or d, or both, and so we may also write $x_n \xrightarrow{d} x$ or $x_n\to x$ in (M,d).

Def: A sequence $\{x_n\}$ is called Cauchy in (M,d) if, given any $\epsilon>0$, there is an integer $N\geq 1$ such that $d(x_m,x_n)<\epsilon$ whenever $m,n\geq N$.

We can reword this just a bit to read: $\{x_n\}$ is Cauchy if and only if, given $\epsilon>0$, there is an integer $N\geq 1$ such that $\operatorname{diam}\{x_n : n\geq N\}\leq\epsilon$ (How?)

The last formulation characterizes Cauchy sequences as sequences, whose elements tend to cluster together. To visualize how a Cauchy sequence might look like in nature, simply take an ant nest and enumerate every ant that lives in it. Now place a sugar cube in the vicinity of the nest. You will see the following pattern:



Suppose the sugar cube is in the center of the red target scope of radius $\epsilon = 1$ (Open ball of radius $\epsilon = 1$). Then all the ants with the number level $n \geq 11$ are apparently within the red scope. In particular, if one ant is labled n and another is labled m, where $m, n \geq 11$, the distance between them is at most 2. As the red circle is shrunk about the sugar cube, we will still find that all but finitely many insects are inside the red circle.

Related to the concept of clustering (i.e. Cauchy) sequences, but in no way identical to it is the concept of convergent sequences. For convenience, the definition of convergence is stated below.

Def: Let (M,d) be a metric space. A sequence $\{x_n\}_{n=1}^{\infty} \subset M$ is said to converge in M if there is some $x \in M$ such that for every $\epsilon > 0$, there is an integer $N > 0$ such that $d(x_n, x) < \epsilon$ whenever $n \geq N$.

Ex. Let $M = [0, \infty)$ be equipped with its usual $|\cdot|$ metric and let $\{\frac{1}{n}\}_{n=1}^{\infty}$ be a sequence in M.

(a) Is $\{\frac{1}{n}\}_{n=1}^{\infty}$ a convergent sequence?

(b) Is $\{\frac{1}{n}\}_{n=1}^{\infty}$ a Cauchy sequence?

Solution:

(a) Recall that vaguely familiar expression from Calculus: $\lim_{n \to \infty} \frac{1}{n} = 0$. We will show that $\{\frac{1}{n}\}_{n=1}^{\infty}$ converges to 0.

Since $\mathbb{R}$ is an ordered field, $|\frac{1}{n} - 0| = \frac{1}{n} < \epsilon$ if and only if $n > \frac{1}{\epsilon}$ (Why?). It follows from the archimedean property of $\mathbb{R}$ that $n > \frac{1}{\epsilon}$ is achieved for some sufficiently large integer N. Thus, if $n \geq N$, $\frac{1}{n} < \epsilon$ as desired.

(b) The sequence is Cauchy. For $\epsilon > 0$, let N be a positive integer such that if $n \geq N$, $|\frac{1}{n} - 0| < \epsilon/2$. Then for $m, n \geq N$

$$\left|\frac{1}{m} - \frac{1}{n}\right| \leq \left|\frac{1}{n} - 0\right| + \left|\frac{1}{m} - 0\right| < \epsilon/2 + \epsilon/2 = \epsilon.$$

Ex. Let $M \subset \mathbb{R}^2$ be the set of position coordinates of ants with the usual, $\|\cdot\|_2$ metric. Is the sequence of ants a Cauchy sequence? Is it convergent?

Solution:

Since the ants eventually cluster together (i.e. ants designated by large positive integers are close to one another), the sequence of their position coordinates form a Cauchy sequence. The ants cluster, however, around a sugar cube, which is evidently not an ant. Hence the sequence does not converge (in M).

Ex. Consider the sequence $\{\frac{1}{n}\}_{n=1}^{\infty}$, but this time in $M = (0,\infty)$. Then $\{\frac{1}{n}\}_{n=1}^{\infty}$ is still a Cauchy sequence under the usual metric inherited by M from $\mathbb{R}$. Notice, however, that $\lim_{n\to\infty} \frac{1}{n} = 0 \notin (0,\infty)$. Thus $\{\frac{1}{n}\}_{n=1}^{\infty}$ does not converge (in M)

Ex. Whether a sequence converges or not depends on the metric function as well. Let (M,d) be the metric space $[0,\infty)$ under the discrete metric $d(x,y) = \begin{cases} 1 & \text{if } x \neq y \\ 0 & \text{if } x = y. \end{cases}$

Then the sequence $\{\frac{1}{n}\}_{n=1}^{\infty}$ does not converge; for any $x \in M$, $B_1^d(x) = \{y \in M : d(x,y) < 1\} = \{x\}$. In other words, $\{\frac{1}{n}\}_{n=1}^{\infty}$ fails to cluster around x.

Ex. Whether or not a given sequence is Cauchy depends on the metric function. Let $M=(0,\infty)$ and d be defined by $d(x,y) = \left|\frac{1}{x}-\frac{1}{y}\right|$. Then the sequence $\{\frac{1}{n}\}_{n=1}^{\infty} \subset M$ is not Cauchy: $d(\frac{1}{n},\frac{1}{m}) = |n-m| \geq 1$ for $m \neq n$.

Lets try to clarify the relationship between Cauchy and convergent sequences.

Proposition: Limits are unique. That is, if $x_n \overset{d}{\mapsto} x$ and $x_n \overset{d}{\mapsto} y$ then $x=y$

Proof: We will show that $d(x,y)=0$ by proving that $d(x,y)<\epsilon$ for any $\epsilon>0$. Since $x_n \overset{d}{\mapsto} y$, there is some $N>0$ such that $d(x_n,y)<\epsilon/2$ whenever $n \geq N$. Similarly, since $x_n \overset{d}{\mapsto} x$, there is some $M>0$ such that $d(x_n,x)<\frac{\epsilon}{2}$ whenever $n\geq M$. Letting $K=\max\{M,N\}$ we see that $d(x,y) \leq d(x,x_n)+d(x_n,y) < \frac{\epsilon}{2}+\frac{\epsilon}{2} = \epsilon$ whenever $n \geq K$.

This proposition is reassuring. It tells us that when we go somewhere, we will arrive in only one place. It would have been rather confusing if we arrived in several places at once.

Proposition: Every convergent sequence is Cauchy, and a Cauchy sequence is bounded. That is, the set $\{x_n : n \geq 1\}$ is bounded.

Proof: Suppose $x_n \xrightarrow{d} x$. Then for any $\epsilon > 0$, there is a positive integer N such that $d(x_n, x) < \epsilon/2$ whenever $n \geq N$.
Now $d(x_n, x_m) \leq d(x_n, x) + d(x, x_m) < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$ whenever $n, m \geq N$. Thus $\{x_n\}_{n=1}^{\infty}$ is Cauchy.
Suppose $\{y_n\}_{n=1}^{\infty}$ is Cauchy in (M, d). We would like to show that $\{y_n : n \geq 1\}$ is bounded. That is, we need to find $y \in M$ and $r \in \mathbb{R}$ such that $d(y_n, y) \leq r$ for all $n \geq 1$. Let $\epsilon > 0$. Then for some $N > 0$, $d(y_n, y_m) < \epsilon$ whenever $n, m \geq N$. Set $y = y_N$ and $r = \sum_{i=1}^{N-1} d(y_i, y_N) + \epsilon$. Observe that $d(y_n, y_N) < \epsilon$ whenever $n \geq N$ and $d(y_n, y_N) \leq \sum_{i=1}^{N-1} d(y_i, y_N)$ whenever $n \leq N-1$. Thus $d(y_n, y_N) < r$ for all $n \in \mathbb{N}$.

Although every Cauchy sequence is bounded, not every bounded sequence is, in turn, Cauchy. For easy examples, consider $\{n\}_{n=1}^{\infty} \subset \mathbb{R}$ under the discrete metric. This sequence is bounded (Why?), but it is not Cauchy. As another example, notice that the sequence $\{(-1)^n\}_{n=1}^{\infty} \subset \mathbb{R}$ is bounded in any metric, as it has finite range.