Math 351 Notes

Instructor Arkady Etkin

Metrics and Norms

METRIC SPACES

Definition:

A function d on $M \times M$ satisfying the following properties is called a metric on M:

i) $0 \leq d(x, y) < \infty$ for all pairs $x, y \in M$.

ii) $d(x, y) = 0$ iff $x = y$.

iii) $d(x, y) = d(y, x)$ for all pairs $x, y \in M$.

iv) $d(x, y) \leq d(x, z) + d(z, y)$ for all $x, y, z \in M$.

The couple (M, d), consisting of a set M together with a metric d defined on M, is called a metric space.

Example:

a) Every set M admits at least one metric. For example, check that the function defined by $d(x, y) = 1$ for any $x \neq y$ in M, and $d(x, x) = 0$ for all $x \in M$, is a metric. This mundane but always available metric is called the dicrete metric on M. A set supplied with its discrete metric is called a discrete space.

b) An important example is the real line $\mathbb{R}$ together with its usual metric $d(a, b) = |a - b|$. Any time we refer to $\mathbb{R}$ without explicitly naming a metric, the absolute value metric is always understood to be the one that we have in mind.

c) Any subset of a metric space is also a metric space in its own right. If d is a metric on M, and $A \subset M$, then $d(x, y)$ is defined for any pair $x, y \in A$. Moreover, the restriction of d to $A \times A$ obviously still satisfies properties (i)–(iv). That is, the metric defined on M automatically defines a metric on A by restriction. We will even use the same letter d and simply refer to the metric space (A, d). Of particular interest in this regard is that $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, and $\mathbb{R}\backslash\mathbb{Q}$ each come already supplied with a natural metric, namely the restriction of the usual absolute value metric on $\mathbb{R}$. In each case, we will refer to this restriction as the usual metric.

How can we enrich our arsenal of metric functions?

$\bullet\!\!\rightarrow$ Suppose that d is a metric on M and $f : M \longrightarrow M$ is a bijective function. Then, $\rho : M \times M \longrightarrow \mathbb{R}$ defined by $\rho(x, y) = d(f(x), f(y))$ is also a metric on M, as we can verify.

$\bullet\!\!\rightarrow$ Suppose that $g : M \longrightarrow \mathbb{R}$ is an injective function (not necessarily surjective). Then, $G : M \times M \longrightarrow \mathbb{R}$ defined by $G(x, y) = \left|g(x) - g(y)\right|$ is a metric on M.

To see this, observe that

i) $0 \leq G(x, y) < \infty$ for all pairs $x, y \in M$ ✓

ii) $G(x, y) = 0$ iff $\left|g(x) - g(y)\right| = 0$, which is true only iff $g(x) = g(y)$, which happens only when $x = y$, since we assumed that g is injective. ✓

iii) $G(x, y) = \left|g(x) - g(y)\right| = \left|g(y) - g(x)\right| = G(y, x)$ ✓

iv) $$G(x, y) = \left|g(x) - g(y)\right| = \left|g(x) - g(z) + g(z) - g(y)\right| \leq \left|g(x) - g(z)\right| + \left|g(z) - g(y)\right| = G(x, z) + G(z, y) \quad ✓$$

Example:

a) Define $d_1, d_2, d_3 : \mathbb{R}^2 \longrightarrow [0, \infty)$ by

$$d_1(x, y) = \left|\tan^{-1}(x) - \tan^{-1}(y)\right| \;, \quad d_2(x, y) = \left|x^3 - y^3\right|, \quad \text{and} \quad d_3(x, y) = \left|e^x - e^y\right|$$

Then, d_1, d_2, and d_3 are all metric functions on $\mathbb{R}$.

b) Let $M = (0, \infty)$. Then $d_1, d_2, d_3 : (0, \infty) \times (0, \infty) \longrightarrow [0, \infty)$ defined by

$$d_1(x, y) = \left|\sqrt{x} - \sqrt{y}\right| \;, \quad d_2(x, y) = \left|\ln(x) - \ln(y)\right| = \left|\ln\left(\frac{x}{y}\right)\right| \;, \quad d_3(x, y) = \left|\frac{1}{x} - \frac{1}{y}\right|$$

are all metric functions on M.

c) Note that a function can be a metric on one set and fail to be a metric on another. Take, for instance, the function $d(x, y) = \left|x^2 - y^2\right|$. Then d defines a metric on $[0, \infty)$, but fails to be a metric on $\mathbb{R}$. (we can easily see why, as it violates some properties of metric spaces).

Note: We can expand our collection of metrics even further. To do this, we first prove the following lemma.

• Lemma:

Let $f : [0, \infty) \longrightarrow [0, \infty)$ be any function with the following two properties:

a) $f(x) = 0$ iff $x = 0$. Otherwise $f(x) > 0$.

b) f' is decreasing. That is, if $x < y$, then $f'(x) > f'(y)$.
Then for any pair $x, y \in [0, \infty)$, $f(x+y) \leq f(x) + f(y)$.

Proof:
Let $g(x) = f(x+y)$ and $p(x) = f(x) + f(y)$, where we regard y as a fixed number. We wish to show that $g(x) \leq p(x)$ or, equivalently, that $0 \leq p(x) - g(x)$.
Notice that $\frac{d}{dx}\left(p(x) - g(x)\right) = p'(x) - g'(x) = f'(x) - f'(x+y) \geq 0$ by property b) of f.
Thus, by the first derivative test, $p(x) - g(x)$ is increasing for all $x \in [0, \infty)$, attaining its smallest value when $x = 0$. Now, $p(0) - g(0) = f(0) + f(y) - f(y) = f(y) - f(y) = 0$.
Thus, $p(x) - g(x) \geq 0$ for all x and the desired result follows. ■

• Theorem:
Let $d : M \times M \longrightarrow [0, \infty)$ be a metric function on M and suppose $f : [0, \infty) \longrightarrow [0, \infty)$ satisfies properties a) and b) of the above lemma.
If $f'(t) > 0 \;\; \forall \; t \in (0, \infty)$, then $\rho : M \times M \longrightarrow [0, \infty)$ given by $\rho(x, y) = f(d(x, y))$ defines another metric on M.

Proof:
We have to check if all four properties of metrics are satisfied:
i) Clearly $0 \leq f(d(x, y)) < \infty \;\; \forall \; x, y \in M$. ✓

ii) Suppose $\rho(x, y) = 0$, then $f(d(x, y)) = 0$. By property a) of f, this implies that $d(x, y) = 0$, or $x = y$. Obviously $\rho(x, x) = 0$. ✓

iii) Clearly $\rho(x, y) = \rho(y, x) \;\; \forall \; x, y \in M$.

iv) $$\begin{aligned}\rho(x, y) = f(d(x, y)) &\leq f(d(x, z) + d(z, y)) \\ &\leq f(d(x, z)) + f(d(z, y)) \\ &= \rho(x, z) + \rho(z, y) \quad \forall \; x, y, x \in M. \quad ✓\end{aligned}$$
Note that on property iv) the first inequality comes from the assumption that f is increasing and $d(x, y) \leq d(x, z) + d(z, y)$, and the second inequality is a consequence of the above lemma.
Hence, since ρ satisfies all the required properties, we conclude that ρ is a metric function. ■

Example:
a) We should verify that

$$\rho(a, b) = \sqrt{|a-b|}\,, \qquad \sigma(a, b) = \frac{|a-b|}{1+|a-b|}, \qquad z(a, b) = \ln(|a-b|+1)$$

each define metrics on $\mathbb{R}$.

b) If d is any metric on M, verify that

$$\rho(x, y) = \sqrt{d(x, y)}\,, \qquad \sigma(x, y) = \frac{d(x,y)}{1+d(x,y)}, \qquad z(x, y) = \ln(d(x, y)+1)$$

are also metrics on M.

Comprehension check: Is $\rho(x, y) = \sqrt{\ln(\left|x^3 - y^3\right| + 1)}$ a metric function on $\mathbb{R}$?

How about $\sigma(x, y) = \dfrac{\sqrt{\ln(\left|x^3-y^3\right|+1)}}{1+\sqrt{\ln(\left|x^3-y^3\right|+1)}}$?

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.

Definition: A norm on a vector space V is a function $\|\cdot\| : V \longrightarrow [0, \infty)$ satisfying
i) $0 \leq \|x\| < \infty \quad \forall\, x \in V$
ii) $\|x\| = 0$ iff $x = 0$
iii) $\|\alpha x\| = |\alpha|\, \|x\|$ for any scalar α and any $x \in V$
iv) $\|x + y\| \leq \|x\| + \|y\| \quad \forall\, x, y \in V$
A function $\|\cdot\| : V \longrightarrow [0, \infty)$ satisfying all of the above properties except ii) is called a pseudonorm. That is, a pseudonorm allows nonzero vectors to have zero 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 $(V, \|\cdot\|)$.

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

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

For $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$,

$\bullet\!\!\to \; \| x \|_1 = \sum_{i=1}^{n} |x_i|$

$\bullet\!\!\to \; \| x \|_2 = \left(\sum_{i=1}^{n} |x_i|^2 \right)^{1/2}$

...As it happens, for $1 \le p < \infty$, the expression $\| x \|_p = \left(\sum_{i=1}^{n} |x_i|^p \right)^{1/p}$ defines a norm on $\mathbb{R}^n$.

$\bullet\!\!\to \; \| x \|_\infty = \max_{1 \le i \le n} |x_i|$

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 in $\mathbb{R}^n$

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

$\bullet\!\!\to \; \| f \|_1 = \int_a^b | f(t)| \, dt$

$\bullet\!\!\to \; \| f \|_2 = \left(\int_a^b | f(t)|^2 \, dt \right)^{1/2}$

$\bullet\!\!\to \; \| f \|_\infty = \max_{a \le t \le b} | f(t)|$

Again, the second expression is the 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

$$\| x \|_p = \left(\sum |x_i|^p \right)^{1/p} \quad \text{if } 1 \leq p < \infty$$

or

$$\| x \|_\infty = \sup_{n \in \mathbb{N}} |x_n| \qquad \text{if } p = \infty$$

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):
$\sum_{i=1}^{n} |x_i\, y_i| \leq \| x \|_2 \; \| y \|_2$ for any $x,\ y \in \ell_2$.

Proof:
To simplify notation, we write $\langle x,\ y \rangle = \sum x_i\, y_i$.
We first consider the case where $x,\ y \in \mathbb{R}^n$ (that is, $x_i = 0 = y_i \quad \forall\, i > n$). 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

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

Since this (nontrivial) quadratic in t is always nonnegative, it must have a nonpositive discriminant. 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 $\sum_{i=1}^{n} |x_i\, y_i| \leq \| x \|_2 \; \| y \|_2$. ∎

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

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

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

$$\begin{aligned} \sum_{i=1}^{n} |x_i + y_i|^2 &= \sum_{i=1}^{n} |x_i|^2 + 2 \sum_{i=1}^{n} x_i\, y_i + \sum_{i=1}^{n} |y_i|^2 \\ &\leq \| x \|_2^2 + 2 \| x \|_2 \| y \|_2 + \| y \|_2^2 = (\| x \|_2 + \| y \|_2)^2 . \end{aligned}$$

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

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

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

Proof:
$(a + b)^p \leq (2 \max\{a, b\})^p = 2^p \max\{a^p, b^p\} \leq 2^p (a^p + b^p)$.
Thus, if $x, y \in \ell_p$, then

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

• 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 $ab \leq \frac{a^p}{p} + \frac{b^p}{q}$, with equality occurring iff $a^{p-1} = b$.

Proof:
Since the inequality trivially holds if either a or b is 0, we may suppose $a, b > 0$.
Since $\frac{1}{p} + \frac{1}{q} = 1$, we see that $p\left(\frac{1}{p} + \frac{1}{q}\right) = p \Longrightarrow 1 + \frac{p}{q} = p$. In particular, $\frac{p}{q} = q - 1$.

Similarly, notice that $q\left(\frac{1}{p}+\frac{1}{q}\right)=q$, implying that $\frac{q}{p}=q-1$.

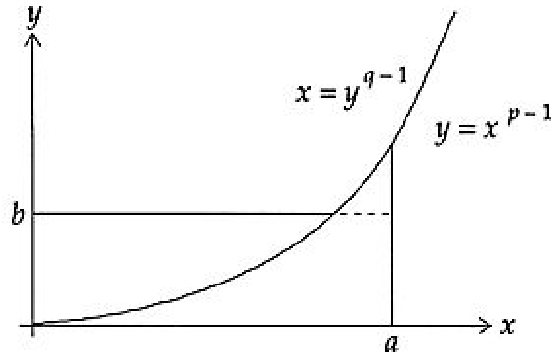
Thus,

$$\frac{1}{p-1}=\frac{1}{p/q}=\frac{q}{p}=q-1\,.$$

Also notice that $q=\frac{1}{p-1}+1$, implying that, just like p, q is also 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):



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 $a^{p-1}=b$. ■

Note: When $p=q=2$, Young's inequality reduces to the arithmetic–geometric mean inequality (although it is usually stated in the form $\sqrt{ab}\leq\frac{a+b}{2}$). Young's inequality will supply the extension of the Cauchy-Schwarz inequality that we need. So now we present a more generalized 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 $x\in\ell_p$ and $y\in\ell_q$, 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:

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

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

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

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

• Theorem (Minkowski's (General) Inequality):
Let $1 < p < \infty$. If $x, y \in \ell_p$, then $x + y \in \ell_p$. Moreover, $\|x + y\|_p \leq \|x\|_p + \|y\|_p$.

Proof:
We have already shown in a previous lemma that $x + y \in \ell_p$. 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:

$$\begin{aligned}
\sum_{i=1}^{\infty} |x_i + y_i|^p &= \sum_{i=1}^{\infty} |x_i + y_i| \cdot |x_i + y_i|^{p-1} \\
&\leq \sum_{i=1}^{\infty} |x_i| \cdot |x_i + y_i|^{p-1} + \sum_{i=1}^{\infty} |y_i| \cdot |x_i + y_i|^{p-1} \\
&\leq \|x\|_p \cdot \|(|x_n + y_n|^{p-1})\|_q + \|y\|_p \cdot \|(|x_n + y_n|^{p-1})\|_q \\
&= \|x + y\|_p^{p-1} \left(\|x\|_p + \|y\|_p\right).
\end{aligned}$$

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

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

Definition: 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 also need to refer to the metric d, then we write $B_r^d(x)$. We may occasionally refer to the set $C_r^d(x) = \{y \in M : d(x, y) \leq r\}$ as the closed ball about x of radius r.

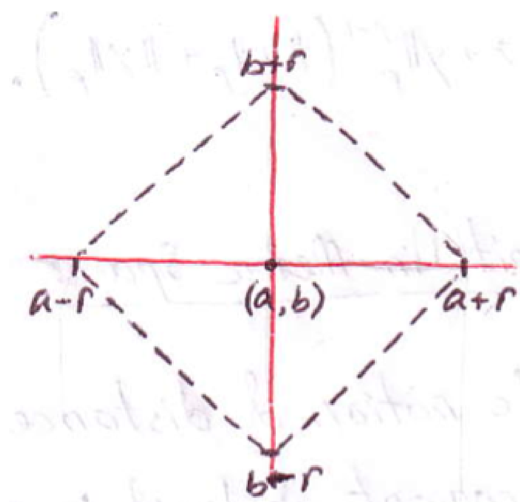
Example:

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.

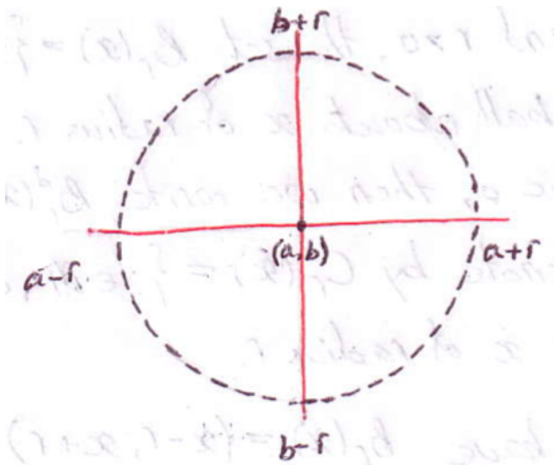
b) In $\mathbb{R}^2$ the set $B_r(x)$ is the open disk of radius r centered at x.

The appereance of $B_r(x)$ in fact depends on the metric at hand:

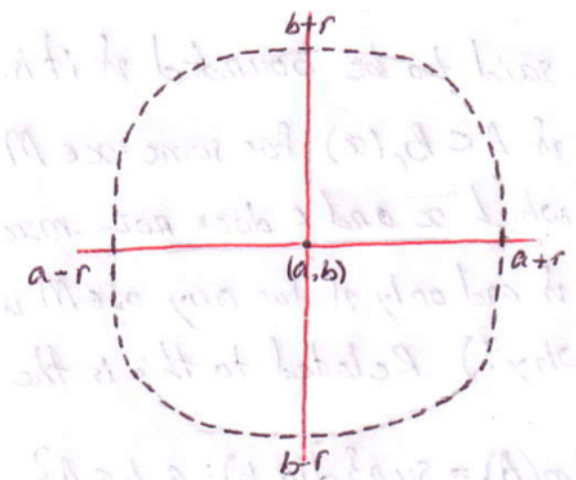
If d is generated by the norm $\|\cdot\|_1$, then $B_r(x)$ will look like a square of diameter $2r$ centered at $x = (a, b)$.



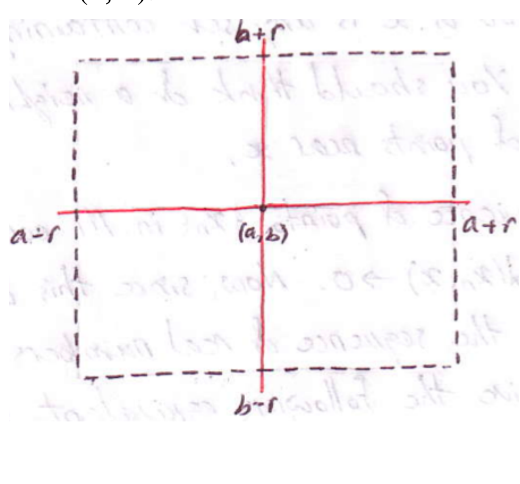
If d is generated by the norm $\|\cdot\|_2$, then $B_r(x)$ will look like a disk of radius r centered at $x = (a, b)$.



If d is generated by the norm $\|\cdot\|_p$, with $1 < p < \infty$, 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. In this setting $B_r(x) = x + B_r(0) = \{y \in V : y = x + z \wedge \|z\| < r\}$.

Note: 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 iff for any $x \in M$ we have $\sup_{a \in A}(d(x, a)) < \infty$. Related to this is the diameter of A, defined by $\operatorname{diam}(A) = \sup\{d(a, b) : a, b \in A\}$. The diameter of A is a convenient measure of size because it does not refer to points outside of A.

<u>Definition:</u> A neighborhood of x is any set containing an open ball about x. We 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 $d(x_n, x) \longrightarrow 0$.
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:

$$\left\{\begin{array}{l}(x_n) \text{ converges to } x \text{ if and only if, given any } \varepsilon > 0, \text{ there is} \\ \text{an integer } N \geq 1 \text{ such that } d(x_n, x) < \varepsilon \text{ whenever } n \geq N,\end{array}\right.$$

or

$$\left\{\begin{array}{l}(x_n) \text{ converges to } x \text{ if and only if, given any } \varepsilon > 0, \text{ there is} \\ \text{an integer } N \geq 1 \text{ such that } \{x_n : n \geq N\} \subset B_\varepsilon(x).\end{array}\right.$$

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

$$\begin{cases} (x_n) \text{ converges to } x \text{ if and only if, given any } \varepsilon > 0, \\ \text{the sequence } (x_n) \text{ is eventually in } B_\varepsilon(x) \end{cases}$$

or, in yet another incarnation,

$$\begin{cases} (x_n) \text{ converges to } x \text{ if and only if the sequence} \\ (x_n) \text{ is eventually in every neighborhood of } x. \end{cases}$$

This final version is blessed by a total lack of N ś and ε ś! In any event, just as with real sequences, we usually settle for the shorthand $x_n \longrightarrow x$ 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 $x_n \overset{d}{\longrightarrow} x$ or $x_n \longrightarrow x$ in (M, d).

Definition: A sequence $\{x_n\}$ is a Cauchy sequence in (M, d) if, given any $\varepsilon > 0$, there is an integer $N \geq 1$ such that $d(x_m, x_n) < \varepsilon$ whenever $m, n \geq N$.
We can reword this just a bit to read: $\{x_n\}$ is Cauchy iff, given $\varepsilon > 0$, there is an integer $N \geq 1$ such that $\text{diam}(\{x_n : n \geq N\}) < \varepsilon$.

Related to the concept of clustering (i.e. Cauchy) sequences, but in no way identical to it, is the concept of convergent sequences (Cauchy sequences need not necessarily converge). For convenience, the definition of convergence is stated below.

Definition: 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 $N > 0$ that satisfies $d(x_n, x) < \epsilon$ whenever $n \geq N$.

Example:
Let $M = [0, \infty)$ be endowed with its usual $|\cdot|$ metric and let $\left\{\frac{1}{n}\right\}_{n=1}^{\infty}$ be a sequence in M.
a) Is $\left\{\frac{1}{n}\right\}_{n=1}^{\infty}$ a convergent sequence?

Solution:
Recall that vaguely familiar expression from calculus: $\underset{n \longrightarrow \infty}{\text{limit}} \frac{1}{n} = 0$. We will show that $\left\{\frac{1}{n}\right\}_{n=1}^{\infty}$ converges to 0.

Since $\mathbb{R}$ is an ordered field, $\left|\frac{1}{n} - 0\right| = \frac{1}{n} < \epsilon$ iff $n > \frac{1}{\epsilon}$. It follows from the Archimidean property of $\mathbb{R}$ that $n > \frac{1}{\epsilon}$ can be achieved for some sufficiently large integer N. Thus, if $n \geq N$, $\frac{1}{n} < \epsilon$ as desired.

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

Solution:
The sequence is Cauchy. For $\epsilon > 0$, let N be a positive integer such that if $n \geq N$, $\left|\frac{1}{n} - 0\right| \leq \frac{\epsilon}{2}$. Then for $m, n \geq N$,

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

Example:
Consider the sequence $\left\{\frac{1}{n}\right\}_{n=1}^{\infty}$, but this time in $M = (0, \infty)$.
Then $\left\{\frac{1}{n}\right\}_{n=1}^{\infty}$ 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, $\left\{\frac{1}{n}\right\}_{n=1}^{\infty}$ does not converge in M.

Note: As we ll see now, whether a sequence converges or not depends on the metric function as well.

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

Note: As we ll see now, whether a given sequence is Cauchy depends on the metric function.

Example:
Let $M = (0, \infty)$ and d be defined by $d(x, y) = \left|\frac{1}{x} - \frac{1}{y}\right|$. Then the sequence $\left\{\frac{1}{n}\right\}_{n=1}^{\infty} \subset M$ is

not Cauchy: $d\left(\frac{1}{n}, \frac{1}{m}\right) = |n - m| \geq 1$ for $m \neq n$.

• Proposition:

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

Proof:
We will show that $d(x, y) = 0$ by proving that $d(x, y) < \epsilon$ for any $\epsilon > 0$.
Since $x_n \overset{d}{\longmapsto} y$, there is some $N > 0$ such that $d(x_n, y) < \frac{\epsilon}{2}$ whenever $n \geq N$. Similarly, since $x_n \overset{d}{\longmapsto} x$, 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 $n \geq k$. ■

The proposition above is reassuring. It tells us that when we go somewhere, we will arrive to one place and one place only. It would ve been rather confusing if we had arrived at different 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 $x_n \overset{d}{\longmapsto} x$. Then, for any $\epsilon > 0$, there is a positive integer N such that $d(x_n, x) < \frac{\epsilon}{2}$ 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.

Now 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 $d(y_n, y) \leq r \; \forall \; n \geq 1$.

Let $\epsilon > 0$. Then, for some $N > 0$, $d(y_n, y_m) < \epsilon$ whenever $n, m \geq N$.
Now we set

$$y = y_N \qquad \text{and} \qquad r = \sum_{i=1}^{N-1} d(y_i,\, y_N) + \epsilon\,.$$

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, $d(y_n,\, y_N) < r \;\; \forall\, n \in \mathbb{N}$. ■

Note: Although every Cauchy sequence is bounded, not every bounded sequence is, in turn, Cauchy. For an easy example, consider $\{n\}_{n=1}^{\infty} \subset \mathbb{R}$ under the discrete metric. This sequence is definitely bounded, 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 a finite range. However this is not a Cauchy sequence.