

3) *Function spaces.* (These are perhaps the most important examples and the ones to which our ensuing theory has the most important and immediate applications - some of which will be taken up in later chapters.)

Recall that the set of all real valued functions with common domain $D \subseteq \mathbb{R}$, usually a closed interval $[a,b]$, forms a vector space with respect to the point-wise defined operations:

$$"f+g": D \rightarrow \mathbb{R} : x \mapsto f(x)+g(x)$$

$$"\lambda f": D \rightarrow \mathbb{R} : x \mapsto \lambda f(x) .$$

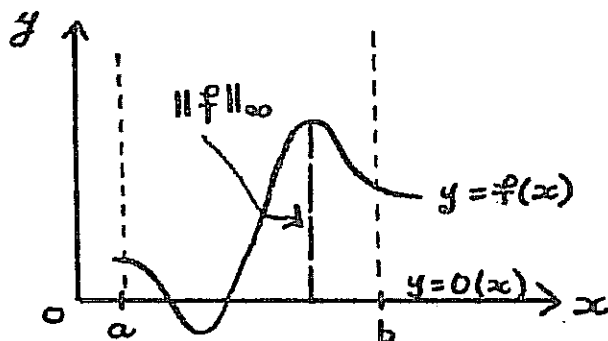
Let \mathcal{B} denote the vector subspace of all bounded functions on D ; that is, functions f for which there exists a constant M_f such that

$$|f(x)| \leq M_f, \text{ for all } x \in D.$$

Defining a norm on \mathcal{B} amounts to providing a measure of the proximity of a function f to the zero function $0[0(x) = 0, \text{ all } x \in D]$. Now at any $x \in D$ the function f differs in value from the zero function by $|f(x)|$. In many applications it seems reasonable to take the "largest" such difference in value to be the norm of f .

Accordingly we define the uniform (or Tchebyscheff) norm on \mathcal{B} by

$$\|f\|_{\infty} = \sup_{x \in [a,b]} |f(x)|$$



This is the analogue of the $\|\cdot\|_\infty$ in the last two examples. To see that it is indeed a norm function we proceed as follows.

Clearly: $\|f\|_\infty \geq 0$; $\|f\|_\infty = 0 \Leftrightarrow |f(x)| = 0$ all x

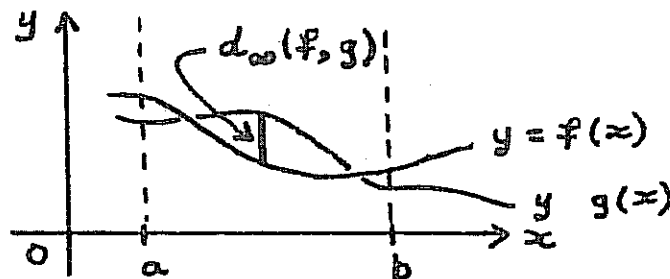
$$\Leftrightarrow f = 0;$$

$$\|\lambda f\|_\infty = \sup_{x \in D} |\lambda f(x)| = \sup_{x \in D} |\lambda| |f(x)| = |\lambda| \|f\|_\infty;$$

$$\begin{aligned} \|f+g\|_\infty &= \sup_{x \in D} |f(x)+g(x)| \\ &\leq \sup_{x \in D} (|f(x)| + |g(x)|) \\ &\leq \sup_{x \in D} |f(x)| + \sup_{x \in D} |g(x)| \\ &= \|f\|_\infty + \|g\|_\infty \end{aligned}$$

and so $\|\cdot\|_\infty$ is indeed a norm on B , inducing the *uniform metric*

$$d_\infty(f,g) = \|f-g\|_\infty = \sup_{x \in D} |f(x) - g(x)|.$$



We will later show that any continuous function defined on a closed interval $[a,b]$ is bounded. Thus an important subspace of $(B, \|\cdot\|_\infty)$ is $(C[a,b], \|\cdot\|_\infty)$ where $C[a,b]$ denotes the space of all continuous real valued functions with domain $[a,b]$. By replacing summation with integration it is possible to define other norms on $C[a,b]$ in analogy

with $\|\cdot\|_p$ ($p=1$, or 2) of the last two examples. Thus for $C[a,b]$ we have the uniform norm defined above,

$$\|f\|_\infty = \max_{x \in [a,b]} |f(x)|$$

and in addition:

$$\|f\|_1 = \int_a^b |f(x)| dx ;$$

$$\|f\|_2 = \sqrt{\int_a^b |f(x)|^2 dx}$$

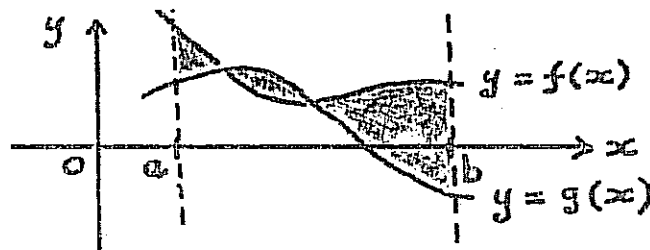
A proof that $\|\cdot\|_2$ is a norm will be given in the next section on inner-products. That $\|\cdot\|_1$ satisfies (N1) to (N4) is left as an easy EXERCISE.

NOTE: $\|\cdot\|_1$ takes the "absolute area" between f and the zero function 0 (the x -axis) as a measure of the distance between these two functions.

The metric induced by this norm,

$$d_1(f,g) = \int_a^b |f - g|$$

is represented by the area of the shaded region in the following sketch



EXAMPLE: For $f(x) = x^3 + x + 1$ and $g(x) = x^3 + x^2 + \frac{1}{2}x + 1$ in $C[0,1]$

we have

$$d_\infty(f,g) = \max_{0 \leq x \leq 1} \left| \frac{1}{2}x - x^2 \right|$$

$$= \frac{1}{4} \quad (\text{check this})$$

while,

$$\begin{aligned} d_1(f,g) &= \int_0^1 |\frac{1}{2}x - x^2| dx = \int_0^{\frac{1}{2}} (\frac{1}{2}x - x^2) dx + \int_{\frac{1}{2}}^1 (x^2 - \frac{1}{2}x) dx \\ &= \frac{1}{8}. \end{aligned}$$

EXERCISE: (a) In $C[0,1]$ determine the values of $d_\infty(f,g)$ and $d_1(f,g)$ when $f(x) = \sin x$ and $g(x) = x$, and also when $g(x) = x - x^3/6$.

(b) Using Taylor's Theorem with remainder, obtain estimates for $d_\infty(f, p_n)$ and $d_1(f, p_n)$ in $C[0,1]$

when

$$f = \exp \text{ and } p_n(x) = \sum_{m=0}^n x^m/m!$$

INNER PRODUCT SPACES.

Just as metrics are induced by the richer structure of a norm, a norm itself sometimes results because of other structure carried by the space. In particular this is so when the space has an inner-product defined on it.

DEFINITION. An inner-product for the vector space X is a mapping from ordered pairs of elements of X into the real field;

$X \times X \rightarrow \mathcal{R}: (x,y) \mapsto \langle x,y \rangle$, which satisfies:

- (IP1) $\langle x,x \rangle > 0$ for all $x \in X$ and $x \neq 0$. (positivity)
- (IP2) $\langle x,y \rangle = \langle y,x \rangle$ for all $x,y \in X$. (symmetry)
- (IP3) $\langle \lambda x,y \rangle = \lambda \langle x,y \rangle$ for all $x,y \in X$ and $\lambda \in \mathcal{R}$ (homogeneity)
- (IP4) $\langle x+y,z \rangle = \langle x,z \rangle + \langle y,z \rangle$ for all $x,y,z \in X$. (additivity)

EXAMPLES

(1) On $X = \mathcal{R}^n$ the usual "dot" or scalar product of two vectors

$$\langle \underline{x}, \underline{y} \rangle \text{ is an inner-product: } \langle \underline{x}, \underline{y} \rangle = x_1 y_1 + x_2 y_2 + \dots + x_n y_n$$

where $\underline{x} = (x_1, x_2, \dots, x_n)$ and $\underline{y} = (y_1, y_2, \dots, y_n)$

[You should verify this as an EXERCISE.]

This is not the only inner-product which can be defined on \mathbb{R}^n , indeed for any set of n strictly positive numbers w_1, w_2, \dots, w_n an inner-product is defined by

$$\langle \underline{x}, \underline{y} \rangle = \sum_{i=1}^n w_i x_i y_i .$$

Such "weighted" inner-products are of considerable importance in some areas of statistics.

(2) An inner-product on ℓ_2 the space of all square summable sequences (see Example 2 of normed linear spaces) may be defined by

$$\langle \underline{x}, \underline{y} \rangle = \sum_{n=1}^{\infty} x_n y_n$$

[That this expression is finite, for all sequences $\underline{x} = x_1, x_2, \dots, x_n, \dots$ and $\underline{y} = y_1, y_2, \dots, y_n, \dots$ for which $\sum_{n=1}^{\infty} x_n^2$ and $\sum_{n=1}^{\infty} y_n^2$ are finite, follows from the Cauchy - Schwarz - Bunyakowski inequality to be established below. That it satisfies the four axioms of an inner-product is readily checked and so is left as an (optional) EXERCISE.]

(3) For $X = C[a, b]$ we can define

$$\langle f, g \rangle = \int_a^b f(x)g(x)dx \text{ for all } f, g \in C[a, b].$$

With the exception of (IP1) the axioms of an inner-product are easily verified:

(IP1) If $f \in C[a, b]$ is not the zero function, then there exists some $x_0 \in [a, b]$ for which $f(x_0) \neq 0$. By the continuity of f there exists $\delta > 0$ such that for $x \in [a, b]$

$$|x - x_0| < \delta \Rightarrow |f(x) - f(x_0)| < |f(x_0)|/2.$$

Consequently, for $x_0 - \delta < x < x_0 + \delta$ we have

$$f(x)^2 > \frac{1}{4} |f(x_0)|^2 > 0 ,$$

and so

$$\begin{aligned}
 \langle f, f \rangle &= \int_a^b f(x)^2 dx \\
 &= \int_a^{x_0 - \delta} f(x)^2 dx + \int_{x_0 - \delta}^{x_0 + \delta} f(x)^2 dx + \int_{x_0 + \delta}^b f(x)^2 dx \\
 &\geq \int_{x_0 - \delta}^{x_0 + \delta} f(x)^2 dx && \text{(as } f(x)^2 \geq 0 \text{ for all } x) \\
 &> \frac{1}{4} |f(x_0)|^2 \times 2\delta \\
 &> 0 .
 \end{aligned}$$

(IP2) For $f, g \in C[a, b]$ we have

$$\begin{aligned}
 \langle g, f \rangle &= \int_a^b g(x) f(x) dx \\
 &= \int_a^b f(x) g(x) dx \\
 &= \langle f, g \rangle .
 \end{aligned}$$

(IP3) For $f, g \in C[a, b]$ and $\lambda \in \mathbb{R}$

$$\begin{aligned}
 \langle \lambda f, g \rangle &= \int_a^b \lambda f(x) g(x) dx \\
 &= \lambda \int_a^b f(x) g(x) dx \\
 &= \lambda \langle f, g \rangle
 \end{aligned}$$

(IP4) For $f, g, h \in C[a, b]$

$$\begin{aligned}
 \langle f+g, h \rangle &= \int_a^b [f+g](x) h(x) dx \\
 &= \int_a^b [f(x)+g(x)] h(x) dx \\
 &= \int_a^b f(x) h(x) dx + \int_a^b g(x) h(x) dx \\
 &= \langle f, h \rangle + \langle g, h \rangle .
 \end{aligned}$$

The following useful properties of an inner-product are immediate consequences of the above axioms, which you should prove for yourself

as EXERCISES.

$$(1) \quad \langle x, y+z \rangle = \langle x, y \rangle + \langle x, z \rangle$$

$$(2) \quad \langle x, \lambda y \rangle = \lambda \langle x, y \rangle$$

$$(3) \quad \langle x, x \rangle = 0 \quad \text{if and only if } x = 0$$

$$(4) \quad y = 0 \quad \text{if and only if } \langle x, y \rangle = 0 \quad \text{for all } x \in X.$$

A vector space X together with an inner-product $\langle \cdot, \cdot \rangle$ will be referred to as an inner-product space.

Inner-product spaces were implicitly studied by many mathematicians [For example; the two German mathematicians, David Hilbert (1862-1943) and Erhard Schmidt (1876-1959) and the Hungarian Friederich Riesz (1880-1956)] during the first three decades of the twentieth century, however, the axioms were not made explicit until 1929 when they were expounded by John von Neumann (1903-1957) as a basis for his axiomatic development of quantum mechanics.

The importance of an inner-product space for our purposes is that the formula

$$\|x\| = \sqrt{\langle x, x \rangle}$$

defines a norm on X . The axioms (N1), (N2) and (N3) are easily established:

$$(N1) \quad \langle x, x \rangle \text{ is greater than } 0 \text{ if } x \neq 0 \text{ and equals } 0 \text{ if } x = 0$$

consequently, for all x , $\langle x, x \rangle \geq 0$ and so $\|x\| \geq 0$.

$$(N2) \quad \|x\| = 0 \Leftrightarrow \langle x, x \rangle = 0 \Leftrightarrow x = 0$$

$$(N3) \quad \begin{aligned} \|\lambda x\| &= \sqrt{\langle \lambda x, \lambda x \rangle} = \sqrt{\lambda^2 \langle x, x \rangle} \\ &= \sqrt{\lambda^2} \sqrt{\langle x, x \rangle} \\ &= |\lambda| \|x\|. \end{aligned}$$

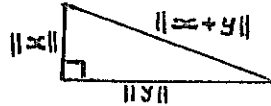
Thus, it only remains to establish (N4), the triangle inequality. This is done as part (ii) of the following theorem.

Theorem 1. *In any inner-product space the following are true:*

$$(i) \quad |\langle x, y \rangle| \leq \|x\| \|y\| \quad (\text{Cauchy-Schwarz-Bunyakovski inequality})$$

$$(ii) \quad \|x + y\| \leq \|x\| + \|y\| \quad (\text{triangle inequality})$$

$$(iii) \quad \text{If } \langle x, y \rangle = 0 \quad \|x + y\|^2 = \|x\|^2 + \|y\|^2 \quad (\text{Pythagorean identity})$$



$$(iv) \quad \|x + y\|^2 + \|x - y\|^2 = 2\|x\|^2 + 2\|y\|^2 \quad (\text{Parallelogram law})$$



Proof.

(i) If $y = 0$ both sides are zero and so the result is immediate.

If $y \neq 0$ we proceed as follows. For any scalar α we have

$$\begin{aligned} 0 \leq \|x + \alpha y\|^2 &= \langle x + \alpha y, x + \alpha y \rangle \\ &= \langle x, x \rangle + \langle \alpha y, x \rangle + \langle x, \alpha y \rangle + \langle \alpha y, \alpha y \rangle \\ &= \|x\|^2 + \alpha \langle x, y \rangle + \alpha [\langle x, y \rangle + \alpha \|y\|^2] \end{aligned}$$

So, choosing $\alpha = -\langle x, y \rangle / \|y\|^2$ (possible as $\|y\| \neq 0$) we see that the term in square brackets is zero and $0 \leq \|x\|^2 + \alpha \langle x, y \rangle = \|x\|^2 - \frac{\langle x, y \rangle^2}{\|y\|^2}$.

Rearranging we therefore have

$$\langle x, y \rangle^2 \leq \|x\|^2 \|y\|^2$$

Taking square roots we therefore have

$$|\langle x, y \rangle| \leq \|x\| \|y\| \quad \text{as required.}$$

$$\begin{aligned} (ii) \quad \|x + y\|^2 &= \langle x + y, x + y \rangle = \langle x, x \rangle + 2\langle x, y \rangle + \langle y, y \rangle \\ &= \|x\|^2 + 2\langle x, y \rangle + \|y\|^2 \\ &\leq \|x\|^2 + 2\|x\| \|y\| + \|y\|^2 \quad (\text{by i}) \\ &= (\|x\| + \|y\|)^2 \end{aligned}$$

so taking the square root of both sides we obtain the triangle inequality, which is otherwise known as Minkowski's inequality in this context.

(iii) Follows immediately from the first two lines of the proof in

(ii) with $\langle x, y \rangle = 0$.

(iv) From the first two lines of the proof in (ii)

$$\|x + y\|^2 = \|x\|^2 + 2\langle x, y \rangle + \|y\|^2 .$$

Similarly by expanding $\langle x-y, x-y \rangle$ we obtain

$$\|x - y\|^2 = \|x\|^2 - 2\langle x, y \rangle + \|y\|^2 .$$

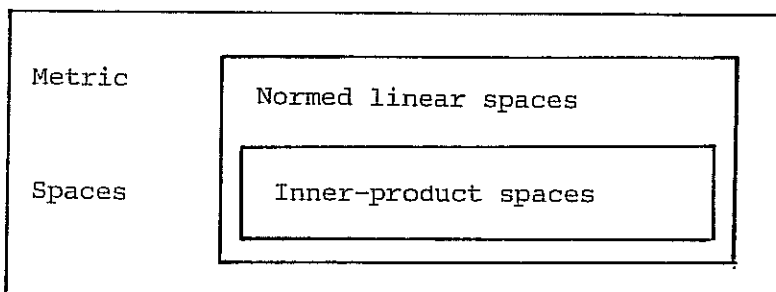
Adding these two identities yields the parallelogram law. ■

The Euclidean norm in \mathbb{R}^n and the norms denoted by $\|\cdot\|_2$ in ℓ_2 and in $C[a,b]$ arise from the above formula for the appropriately defined inner-products. Thus, we have finally arrived at a proof of the fact that these are indeed norm functions. All of these norms have special properties; for example, they satisfy the parallelogram law, which are not satisfied by all norm functions;

For example in ℓ_1^2 the norm $\|x\|_1 = |x_1| + |x_2|$ does not satisfy the parallelogram rule. To see this observe that for $x = (1,0)$ and $y = (0,1)$ we have

$$2(\|x\|^2 + \|y\|^2) = 4 \quad \text{while} \quad \|x + y\|^2 + \|x - y\|^2 = 8 .$$

This shows that $\|\cdot\|_1$ does not arise from any inner-product according to the formula $\sqrt{\langle x, x \rangle}$. Diagrammatically we have the following situation.



Indeed, any norm satisfying the parallelogram law may be shown to arise from a suitably defined inner-product. Thus the parallelogram law characterizes inner-product spaces (the Jordan-von Neumann characterization). Honours students may like to attempt proving this as an exercise. (See exercise 2 on the next page for a hint.)

EXERCISES

- Motivated by (iii) of the previous theorem and the ordinary idea of perpendicularity in \mathbb{R}^n , viz. $x \cdot y = 0$, we define two vectors x, y in the inner-product space X to be *orthogonal* if $\langle x, y \rangle = 0$.

- i) Show that x and y are orthogonal if [and only if - by iii) of the theorem] they satisfy the Pythagorean identity

$$\|x + y\|^2 = \|x\|^2 + \|y\|^2 .$$

- *ii) Show that x is orthogonal to y if and only if $\|x\| \leq \|x + \lambda y\|$ for all scalars $\lambda \in \mathbb{R}$. (Can you give a geometric interpretation to this result.)

The condition $\|x\| \leq \|x + \lambda y\|$ all λ is often used as a generalized definition of x being orthogonal to y in any normed linear space (R.C. James, 1947).

- iii) If $x \neq 0$ and y are elements of X verify that x and $y - \frac{\langle x, y \rangle}{\langle x, x \rangle} x$ are orthogonal.

- 2) For any inner-product space, verify the "polarization identity":

$$\langle x, y \rangle = \frac{1}{4} (\|x+y\|^2 - \|x-y\|^2)$$

- 3) In the space $C[0,1]$, show that

$$\begin{aligned} d_2(x, \sin x) &= \sqrt{\langle x - \sin x, x - \sin x \rangle} \\ &= \sqrt{\int_0^1 (x - \sin x)^2 dx} \doteq 0.061 \end{aligned}$$

while

$$d_2\left(x - \frac{x^3}{6}, \sin x\right) \doteq 0.002.$$

- 4) In terms of the Euclidean norm $\|\cdot\|_2$ on \mathbb{R}^2 note that the "post-office" metric is given by

$$d(\underline{x}, \underline{y}) = \begin{cases} 0 & \text{if } \underline{x} = \underline{y} \\ \|\underline{x}\|_2 + \|\underline{y}\|_2 & \text{otherwise.} \end{cases}$$

Use this to verify that d is indeed a metric.

[Hint. When proving (M4) consider the cases: $\underline{x} = \underline{z}$; $\underline{x} = \underline{y}$ or $\underline{y} = \underline{z}$; $\underline{x}, \underline{y}, \underline{z}$ are all distinct.]

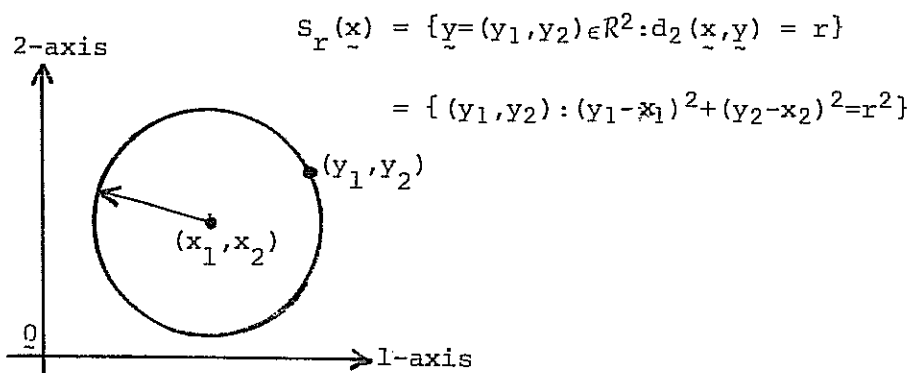
§1.2 Geometry in Metric spaces. Balls, convexity and boundedness.

In ordinary (Euclidean) geometry a circle (or sphere) is defined to be the set of points equidistant from a given point - the centre.

It is possible to generalise this notion into any metric space.

DEFINITION: Let (X,d) be a metric space. The sphere of radius $r > 0$ and centre x is the set $\{y \in X : d(x,y) = r\}$ which we denote by $S_r(x)$.

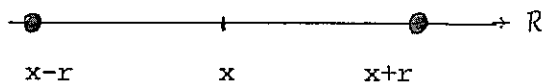
NOTES 1) Under this definition we continue to use the term sphere regardless of the dimension of the space. Thus for example, in ℓ_2^2 (\mathbb{R}^2 with the Euclidean metric) the "sphere" centre $\underline{x} = (x_1, x_2)$ and radius r is what would conventionally be referred to as the circle with centre x and radius r :



While, in \mathbb{R} with the usual metric, $d_1(x,y) = |x-y|$, we see that the "sphere" $S_r(x)$ consists of two points

$$S_r(x) = \{y \in \mathbb{R} : |x-y| = r\}$$

$$= \{x-r, x+r\}$$

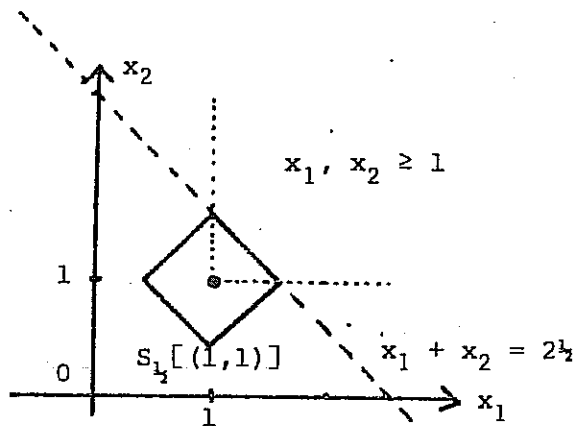


2) The "shape" of a sphere varies with the particular norm used. For example, in ℓ_1^2 the sphere (circle) with centre $(1,1)$ and radius $\frac{1}{2}$ is

$$S_{\frac{1}{2}}(1,1) = \{x = (x_1, x_2) \in \mathbb{R}^2 : d_1(x, (1,1)) = \|x - (1,1)\|_1 = \frac{1}{2}\}$$

$$= \{(x_1, x_2) : |x_1 - 1| + |x_2 - 1| = \frac{1}{2}\}$$

which is the diamond illustrated below



[To see this; consider the case $x_1, x_2 \geq 1$ then

$$|x_1 - 1| = x_1 - 1 \quad \text{and} \quad |x_2 - 1| = x_2 - 1 \quad \text{and so we need}$$

$x_1 - 1 + x_2 - 1 = \frac{1}{2}$ or $x_1 + x_2 = 2\frac{1}{2}$. Similarly consider the other three cases: $x_1 \geq 1, x_2 < 1$; $x_1 < 1, x_2 \geq 1$; $x_1, x_2 < 1$.]

EXERCISES: 1) For any set X with the discrete metric d , show that, for

any $x \in X$ we have

$$S_r(x) = \begin{cases} \emptyset \text{ (the empty set)} & \text{if } r \neq 1 \\ X \setminus \{x\} & \text{if } r = 1 \end{cases}$$

2) In ℓ_∞^2 sketch the sphere $S_{\frac{1}{2}}(1,1)$.

3) In \mathbb{R}^2 with the "post-office" metric sketch each of the following spheres:

$$S_1(0,0), S_3(2,0), S_{\frac{1}{2}}(1,0), S_1(1,0)$$

Of particular importance when $(X, \|\cdot\|)$ is a normed linear space is the *unit sphere of X* , $S_1(0)$ which we sometimes denote by $S(X)$. Thus

$$S(X) = \{x \in X : \|x\| = 1\}.$$

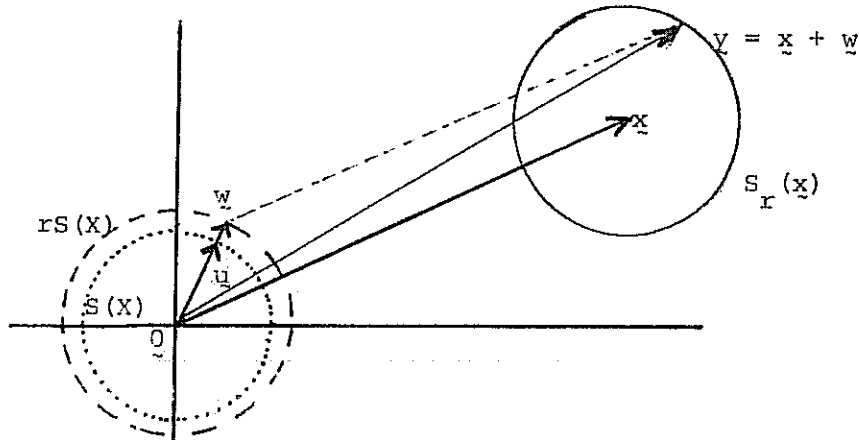
Once the unit sphere $S(X)$ is known all other spheres in the normed linear space $(X, \|\cdot\|)$ are essentially determined. Indeed all other spheres are translates of dilates of the unit sphere.

In $(X, \|\cdot\|)$

$$y \in S_r(x) \Leftrightarrow \|y - x\| = r$$

$$\Leftrightarrow y = x + w \text{ where } w (= y - x) \in S_r(0), \text{ i.e. } \|w\| = r$$

$$\Leftrightarrow y = x + ru \text{ where } u (= w/r) \text{ has } \|u\| = 1 \text{ i.e. } u \in S_1(x)$$



Thus if we write $rS(X)$ for the dilate $\{ru : u \in S(X)\}$ and $x + rS(X)$ for the translate $\{x + w : w \in rS(X)\}$ we have

$$S_r(x) = rS(X) + x$$

EXERCISE: Sketch the unit sphere in ℓ_1^2 . Hence deduce that the sketch of $S_{\frac{1}{2}}(1,1)$ obtained above is essentially correct.

More important than the concept of a sphere for the study of metric spaces is that of a "ball".

DEFINITION: Let (X, d) be a metric space. The open ball of radius r and centre x is

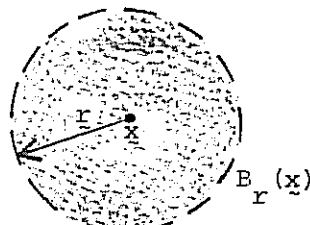
$$B_r(x) = \{y \in X : d(y, x) < r\}.$$

The closed ball of radius r and centre x is

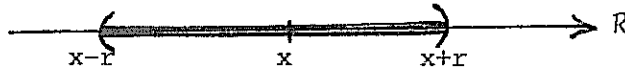
$$B_r[x] = \{y \in X : d(y, x) \leq r\}.$$

Thus in \mathbb{R}^2 with the Euclidean metric $B_r(x)$ is the "disk"

$$\{(y_1, y_2) : (y_1 - x_1)^2 + (y_2 - x_2)^2 < r^2\}$$



In \mathbb{R} with the usual metric $B_r(x)$ is the open interval $(x - r, x + r)$ - verify this.



[Note, $B_r[x] = B_r(x) \cup S_r(x)$ or equivalently

$$B_r(x) = B_r[x] \setminus S_r(x).]$$

The (closed) unit ball of a normed linear space $(X, \|\cdot\|)$ is

$$B[X] = \{x \in X : \|x\| \leq 1\} \quad (= B_1[0]).$$

As with spheres, in the case of a normed linear space we have

$$\begin{aligned} B_r[x] &= rB[X] + x \\ &= \{y \in X : y = ru + x \text{ for } u \in B[X]\}. \end{aligned}$$

Similarly, $B_r(x) = rB_1(0) + x$.

EXERCISES:

1) Verify the claim that in the normed linear space $(X, \|\cdot\|)$

$$B_r[x] = x + rB[X].$$

2) (a) Sketch the (closed) unit ball for ℓ_1^2 and ℓ_∞^2

(b) Sketch the (closed) unit ball in the normed linear space resulting from \mathbb{R}^2 equipped with the inner-product

$$\langle \underline{x}, \underline{y} \rangle = x_1y_1 + 2x_2y_2$$

where

$$\underline{x} = (x_1, x_2) \text{ and } \underline{y} = (y_1, y_2).$$

(c) Show that the (closed) unit ball in $(C[a, b], \|\cdot\|_\infty)$ consists of all continuous functions on $[a, b]$ whose graphs lie entirely between the two lines $y = 1$ and $y = -1$.

3) Sketch the ball $B_2(1, 1)$ in each of the following spaces.

$$\ell_1^2, \ell_2^2, \ell_\infty^2, \mathbb{R}^2 \text{ with the "post-office" metric.}$$

Because of the vector space structure present when $(X, \|\cdot\|)$ is a normed linear space, we can define the notion of a line as well as those of spheres and balls.