Supplement 1

Metric Space and Topological Properties of \( \mathbb{R}^n \)

In the next few lectures we will discuss \( n \)-dimensional Euclidean space. We do so however by alluding to a more general concept, that of a metric space. This will further emphasize some abstraction of basic properties that have wider application, and hints at the synthesis of basic ideas and their development in different contexts, the core of mathematics. The material for the first few lectures will come from Chapter One and parts of Chapter Two of NOTES and \$13/\$22 of Ross. The notations used differ slightly, as does the terminology. We shall try to indicate where statements appear in each reference (e.g. 13.1R, would mean statement 13.1 in Ross, and p27N, will indicate page 27 of Notes).

**Definitions:** \( n \)-dimensional Euclidean Space, denoted \( \mathbb{R}^n \), is the set of all \( n \)-tuples, of real numbers:

\[
\mathbb{R}^n := \{(x_1, \ldots, x_n) : x_i \in \mathbb{R}, i = 1, \ldots, n\}
\]

where \( x = (x_1, \ldots, x_n) \) is a point or a vector or an element of \( \mathbb{R}^n \). The entries \( x_i, i = 1, \ldots, n \), are the components of the vector \( x \). For arbitrary points, \( x = (x_1, \ldots, x_n) \), \( y = (y_1, \ldots, y_n) \), in \( \mathbb{R}^n \), and real number \( \lambda \), the operations

\[
\begin{align*}
x + y &:= (x_1 + y_1, \ldots, x_n + y_n) \quad \text{addition, and} \\
\lambda x &:= (\lambda x_1, \ldots, \lambda x_n) \quad \text{scalar multiplication}
\end{align*}
\]

satisfy the properties to make \( \mathbb{R}^n \) into a vector space. That is

(i) \( x + y = y + x \)
(ii) \( (x + y) + p = x + (y + p) \)
(iii) For \( \mathbb{O} := (0, \ldots, 0) \), \( \mathbb{O} + x = x + \mathbb{O} = x \).
(iv) for \( x \) there is an element \( -x \) so that \( x + (-x) = \mathbb{O} \)
(v) \( -1x = -x \), \( 1x = x \) and \( 0x = \mathbb{O} \)
(vi) \( \lambda(\mu x) = (\lambda \mu)x \)
(vii) \( \lambda(x + y) = \lambda x + \lambda y \), \( (\lambda + \mu)x = \lambda x + \mu x \)

There is an inner product defined on the vector space \( \mathbb{R}^n \) by

\[
\langle x, y \rangle := \sum_{i=1}^{n} x_i y_i = x_1 y_1 + \ldots + x_n y_n.
\]

The following properties of the inner product follow fairly easily from the definition (and you should have seen them before in linear algebra):

(i) \( \langle x, x \rangle \geq 0 \), with equality if and only if \( x = \mathbb{O} \)
(ii) \( \langle x, y \rangle = \langle y, x \rangle \)
(iii) \( \langle x, y + p \rangle = \langle x, y \rangle + \langle x, p \rangle \)
(iv) \( \langle \lambda x, y \rangle = \langle x, \lambda y \rangle = \lambda \langle x, y \rangle \)
The first of these properties can be used to define the **absolute value** or **modulus** of a point $x \in \mathbb{R}^n$, or a **norm** on $\mathbb{R}^n$:

$$|x| := \sqrt{(x,x)} = \left(\sum_{i=1}^{n} x_i^2\right)^{1/2} = \left(x_1^2 + \ldots + x_n^2\right)^{1/2}.$$  

**Theorem S.1.1.** (Cauchy-Schwartz Inequality, p7N)

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

with equality $\iff x = \lambda y$, $\lambda > 0$, or one of $x$ or $y$ is $\emptyset$.

**Proof.** Look at the point $p = \lambda x - \mu y$ for $\lambda, \mu \in \mathbb{R}$, and note that

$$0 \leq \langle p, p \rangle = \langle \lambda x - \mu y, \lambda x - \mu y \rangle$$

$$= \lambda^2 \langle x, x \rangle - \lambda \mu \langle x, y \rangle - \mu \lambda \langle y, x \rangle + \mu^2 \langle y, y \rangle$$

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

If we choose $\lambda = |y|$ and $\mu = |x|$, the last inequality becomes

$$0 \leq |y|^2 |x|^2 - 2|x||y|\langle x, y \rangle + |x|^2 |y|^2 = 2|y||x||y||x| - \langle x, y \rangle.$$  

Hence, $\langle x, y \rangle \leq |y||x|$. Moreover, equality can hold if and only if

$$\langle \lambda x - \mu y, \lambda x - \mu y \rangle = 0 \iff \lambda x - \mu y = \emptyset \quad \text{when } \lambda = |y| \text{ and } \mu = |x|.$$  

That is, equality holds if and only if $|y|x = |x|y = \emptyset$, and the result follows. \qed

**Corollary S.1.2.** (Triangle Inequality p8N) For any $x$ and $y$ in $\mathbb{R}^n$, we have

$$||x| - |y|| \leq |x \pm y| \leq |x| + |y|.$$  

**Proof.** As in the last proof with $\mu = \lambda = 1$

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

$$\begin{cases} \leq |x|^2 + 2|x||y| + |y|^2 = (|x| + |y|)^2, & \text{if } \pm = +; \\ \geq |x|^2 - 2|x||y| + |y|^2 = (|x| - |y|)^2, & \text{if } \pm = -, \end{cases}$$

where the last two inequalities came from the Cauchy-Schwartz inequality. Therefore,

$$||x| - |y|| \leq |x \pm y| \leq |x| + |y|. \quad \Box$$

With the triangle inequality, the norm (absolute value, modulus) satisfies the requirements for a **distance function** or **metric** on sets in $\mathbb{R}^n$.

**Definition:** (13.1R) Let $S$ be a set, and a metric on $S$ is any real-valued function $d$ defined on $S \times S := \{(x, y) : x, y \in S\}$, which satisfies the three properties:

D1. $d(x, x) = 0$ for all $x \in S$ and $d(x, y) > 0$ if $x \neq y$. 

2
D2. \( d(x, y) = d(y, x) \) for all \( x, y \in S \).

D3. \( d(x, y) \leq d(x, z) + d(z, x) \) for all \( x, y, z \in S \) (triangle inequality).

If the set has a metric \( d \), the pair \((S, d)\) is said to be a **metric space**. Note: \( S \) is a set, it need not be a vector space (in spite of the use of the word “space” in metric space). We have seen that \( d(x, y) = |x - y| : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R} \) satisfies the requirements for a metric. Hence, \((\mathbb{R}^n, | |)\) is a metric space, but so is \((S, | |)\) for **any** subset \( S \) of \( \mathbb{R}^n \).

In the case of the familiar \( \mathbb{R}^2 \) and \( \mathbb{R}^3 \), the set

\[
B_S(x, \rho) := \{ y \in S : |x \cdot y| < \rho \}
\]  

(1.1)

consists of all points inside the circle (respectively, sphere or ball) of radius \( \rho \). The terminology is carried over to general metric spaces, and (1.1) is called the **ball of radius** \( \rho \) **about the point** \( x \) in \( S \).

There are other “distance” functions on \( \mathbb{R}^n \). Two of the more important ones are given by the \( \ell_1 \)-norm, \(| \cdot |_1\), and the \( \ell_\infty \)-norm, \(| \cdot |_\infty\):

\[
|x|_1 := \sum_{k=1}^{n} |x_k|, \quad \text{and} \quad |x|_\infty := \max_{1 \leq k \leq n} |x_k|.
\]

The Euclidean norm is sometimes referred to as the \( \ell_2 \)-norm.

The “ball” of radius \( \rho = 1 \) about the origin in \( \mathbb{R}^2 \) are the diamond and square respectively.

**Theorem S.1.3.** Let \( d_1(x, y) := |x - y|_1 \) and \( d_\infty(x, y) := |x - y|_\infty \). Then the distances are related by

\[
\begin{align*}
(i) \quad & d_\infty(x, y) \leq d(x, y) \leq d_1(x, y) \\
(ii) \quad & d_1(x, y) \leq n d_\infty(x, y) \\
(iii) \quad & d_1(x, y) \leq \sqrt{n} d(x, y) \\
(iv) \quad & d(x, y) \leq \sqrt{n} d_\infty(x, y)
\end{align*}
\]  

(1.2)
Inequalities (ii) and (iv) follow from replacing $|x_j - y_j|$ by the max, namely $d_\infty(x, y)$ in each term of the sums defining $d_1(x, y)$ and $d(x, y)$ respectively. The first inequality of (i) follows because the max of $|x_j - y_j|$ is achieved for some $j_0$, and hence appears at least once in the sum for $d(x, y)$, and second inequality (i) follows because when the square of a sum of positive terms is expanded, we also pick up cross terms. Finally, (iii) follow by the Cauchy-Schwarz inequality after noting that $d_1(x, y) = \langle x - y, (\text{sgn}(x_1 - y_1), \ldots, \text{sgn}(x_n - y_n)) \rangle$, where

$$\text{sgn}(x_i - y_i) = \begin{cases} 1, & \text{if } x_i - y_i > 0; \\ 0, & \text{if } x_i - y_i = 0; \\ -1, & \text{if } x_i - y_i < 0; \end{cases}$$

and $|x_i - y_i| = (x_i - y_i) \text{sgn}(x_i - y_i)$.

Once you have a distance, then you can naturally define convergence for sequences. Recall, that a sequence is a mapping of the natural numbers to points in $S$, a “labelling” of the points.

**Definition:** (13.2R) A sequence $(s_k)$ converges to an element $s$ in the metric space $(S, d)$ if

$$\lim_{k \to \infty} d(s_k, s) = 0, \quad \text{(as a sequence of real numbers).}$$

Thus, $s_k \to s$ in $(S, d)$ if and only if, for each given $\epsilon > 0$, there is a $N = N_\epsilon$ so that

$$k > N \implies d(s_k, s) < \epsilon.$$

Geometrically, the sequence $(s_k)$ converges to $s$, if given any ball of radius $\epsilon$ centered on $s$, then there is a $N$ so that all the points $s_k$ with $k > N$ belong to that ball.
A sequence in the metric space \((S,d)\) is a **Cauchy sequence** if for each given \(\epsilon > 0\), there is a \(N = N_\epsilon\) so that

\[ k, m > N \implies d(s_k, s_m) < \epsilon. \]

Now for the real numbers, the Completeness Axiom allowed us to say that every Cauchy sequence of real numbers converged to some real number. In a general metric space, a Cauchy sequence may not have a limit in the set \(S\). For example, we may take \(S\) to be the open interval \((0,1)\), and \(d(x,y) = |x - y|\). Then \(((0,1), |·|)\) is a metric space, the sequence \((1/k)\) is a Cauchy sequence in \(S\), but its limit is not in \(S\). A metric space for which every Cauchy sequence from \(S\) converges to a limit in \(S\) is called a **complete metric space**.

We will consider sequences in \(\mathbb{R}^n\). We denote a sequence of points in \(\mathbb{R}^n\) by \(x^{(k)} = (x_1^{(k)}, \ldots, x_n^{(k)})\), \(k = 1, 2, \ldots\). Thus from the definition, we see that every sequence in \(\mathbb{R}^n\) corresponds to a collection of \(n\)-sequences \((x_j^{(k)})\), \(j = 1, \ldots, n\), namely the sequences of its components. Not surprisingly, the convergence of \((x^{(k)})\) is determined by the convergence of the sequences of its components.

**Lemma S.1.4.** (13.3R) The sequence \((x^{(k)})\) is a Cauchy sequence in \(\mathbb{R}^n\) if and only if each of the component sequences \((x_j^{(k)})\) is a Cauchy sequence in \(\mathbb{R}\), \(j = 1, \ldots, n\).

**Proof.** This will follow from the inequalities in (1.2). Indeed, by (i) of (1.2)

\[ d(x^{(k)}, x^{(m)}) < \epsilon \implies d_\infty(x^{(k)}, x^{(m)}) < \epsilon \implies |x_j^{(k)} - x_j^{(m)}| < \epsilon, \]

for each \(j = 1, \ldots, n\). Hence, the same \(N\) that corresponds to \(\epsilon\) for the sequence \((x^{(k)})\) will give the Cauchy criterion for each of the sequences \((x_j^{(k)})\), \(j = 1, \ldots, n\).

On the other hand, if each of the sequences \((x_j^{(k)})\), \(j = 1, \ldots, n\), is a Cauchy sequence, then given \(\epsilon\), we let \(N_j\) be such that

\[ k, m > N_j \implies |x_j^{(k)} - x_j^{(m)}| < \epsilon / \sqrt{n}. \]

Let \(N = \max\{N_1, N_2, \ldots, N_n\}\). By (iv) of (1.2) and the choice of \(N_j\),

\[ k, m > N_j \implies \max_j |x_j^{(k)} - x_j^{(m)}| < \epsilon \implies d(x^{(k)}, x^{(m)}) < \epsilon. \]

Hence, \((x^{(k)})\) satisfies the Cauchy criterion. \(\square\)

**Corollary S.1.5.** The metric spaces \((\mathbb{R}^n, d)\), \((\mathbb{R}^n, d_1)\), and \((\mathbb{R}^n, d_\infty)\) are complete metric spaces.

**Proof.** If \((x^{(k)})\) is a Cauchy sequence in \(\mathbb{R}^n\), then by the lemma, each \((x_j^{(k)})\) is a Cauchy sequence in \(\mathbb{R}\), \(j = 1, \ldots, n\). Since Cauchy sequences in \(\mathbb{R}\) converge to some real number, there exists numbers \(y_j\) so that

\[ \lim_{k \to \infty} |x_j^{(k)} - y_j| = 0, \quad j = 1, \ldots, n. \]
Then for $\mathbf{y} = (y_1, \ldots, y_n)$, we must have
\[
d(\mathbf{x}^{(k)}, \mathbf{y}) \leq \sqrt{n} d_\infty(\mathbf{x}^{(k)}, \mathbf{y}) = \sqrt{n} \max_j |x_j^{(k)} - y_j| \to 0, \quad \text{as } k \to \infty. \quad \Box
\]

Finally, we have the analog of the Bolzano-Weierstrass Theorem:

**Theorem S.1.6.** *(Bolzano-Weierstrass 13.5R, p13N)* Every bounded sequence in $\mathbb{R}^n$ has a convergent subsequence.

**Proof.** A subset $E$ of a metric space is said to be bounded if it is contained in some ball of finite radius about some point. For $(\mathbb{R}^n, d)$, this is equivalent to, there exists an $M$, $0 < M < \infty$, so that
\[
E \subseteq B(\mathbf{0}, M).
\]
Let $\mathbf{x}^{(k)}$ be a bounded sequence in $\mathbb{R}^n$, with say $\mathbf{x}^{(k)} \subseteq B(\mathbf{0}, M)$. By definition of the ball and (1.2), we have
\[
\max_j |x_j^{(k)}| \leq d(\mathbf{x}^{(k)}, \mathbf{0}) \leq M.
\]
Hence, each sequence $(x_j^{(k)})$ is bounded in $\mathbb{R}$. Beginning with $(x_1^{(k)})$, we extract a convergent subsequence $(x_1^{(k_m)})$ from $(x_1^{(k)})$ (possible by the Bolzano-Weierstrass Theorem for $\mathbb{R}$). Then, $(x_2^{(k_m)})$ will be a bounded subsequence of $\mathbb{R}$ and we may extract a convergent subsequence from it. Repeating this argument, we get a common subsequence for each of the $(x_j^{(k)})$ which converges. Then the subsequence of $(\mathbf{x}^{(k)})$ corresponding to that common subsequence must converge in $\mathbb{R}^n$. \quad \Box