Chapter 2

Normed Linear Spaces

2.1 The Norm Topology

In order to do analysis on vector spaces, we need to endow these spaces with a topological structure which is compatible with the linear structure. This is made precise in the following definition.

Definition 2.1.1 *A* **topological vector space** *is a vector space V that is endowed with a* Hausdorff *topology such that the maps*

$$(x, y) \in V \times V \mapsto x + y \in V \quad \textit{and} \quad (\alpha, x) \in \mathbb{F} \times V \mapsto \alpha x \in V$$

are continuous, each product space being endowed with the appropriate product topology using the given topology of V and the usual topology on the scalar field $\mathbb{F}$ $(= \mathbb{R}$ or $\mathbb{C})$. ■

We will restrict our attention to a class of topological vector spaces called normed linear spaces, which we now proceed to define.

Definition 2.1.2 *A* **norm** *on a vector space V is a function $\|.\| : V \to [0, \infty)$ such that*
(i) $\|x\| = 0$ *if, and only if,* $x = \mathbf{0}$;
(ii) $\|\alpha x\| = |\alpha|\|x\|$ *for every* $\alpha \in \mathbb{F}$ *and every* $x \in V$;
(iii) *(Triangle Inequality) for every x and $y \in V$, we have*

$$\|x + y\| \;\leq\; \|x\| + \|y\|. \; ■ \tag{2.1.1}$$

Associated to a norm on a vector space V, we have a metric defined by

$$d(x, y) \;=\; \|x - y\|.$$

It is immediate to verify that this defines a metric. The triangle inequality (2.1.1) yields the inequality (1.2.1) (which is also called by the same name) via the relation

$$x - z \;=\; (x-y)+(y-z).$$

(Thus, the norm of a vector is its distance from the origin and is a generalization of the notion of the length of a vector, as we know it, in Euclidean space.)

Thus, V is endowed with a metric topology. In this topology, a sequence $\{x_n\}$ converges to x in V if, and only if

$$\|x_n - x\| \;\to\; 0.$$

Now if $\{x_n\}$ and $\{y_n\}$ are sequences in V and $\{\alpha_n\}$ a sequence in $\mathbb{F}$, we have, for $x, y \in V$ and $\alpha \in \mathbb{F}$,

$$\begin{aligned} \|(x_n+y_n)-(x+y)\| &\leq \|x_n - x\| + \|y_n - y\|, \\ \|\alpha_n x_n - \alpha x\| &\leq |\alpha_n|\|x_n - x\| + |\alpha_n - \alpha|\|x\|. \end{aligned}$$

Thus, if $x_n \to x$, $y_n \to y$ in V and if $\alpha_n \to \alpha$ in $\mathbb{F}$, it immediately follows that $x_n + y_n \to x + y$ and that $\alpha_n x_n \to \alpha x$ in V. Thus addition and scalar multiplication are continuous and so V becomes a topological vector space with this metric topology.

Definition 2.1.3 *A* **normed linear space** *is a vector space V endowed with a norm. The metric topology induced by the norm is called its* **norm topology**. ■

The norm itself is a continuous function with respect to this topology. Indeed, if x and $y \in V$, then, since $x = (x-y)+y$, the triangle inequality yields

$$\|x\| \;\leq\; \|x-y\| + \|y\|$$

which we rewrite as

$$\|x\| - \|y\| \;\leq\; \|x-y\|.$$

Interchanging the roles of x and y we finally obtain

$$|\;\|x\| - \|y\|\;| \;\leq\; \|x-y\|$$

from which the continuity of the function $x \in V \mapsto \|x\| \in \mathbb{R}$ follows.

Definition 2.1.4 *A normed linear space is said to be a* **Banach space** *if it is complete under the norm topology.* ■

2.2 Examples

We will now look at several examples of normed linear spaces. Essentially, they can be classified into three groups - finite dimensional spaces, sequence spaces and function spaces. In the examples that follow, we will set $\mathbb{F} = \mathbb{R}$. The reader can easily make the necessary changes to cover the case when $\mathbb{R}$ is replaced by $\mathbb{C}$.

Example 2.2.1 We can consider $\mathbb{R}$ as a vector space over itself. The map $x \in \mathbb{R} \mapsto |x|$ is easily seen to define a norm which generates the usual topology on $\mathbb{R}$. Since $\mathbb{R}$ is complete, it thus becomes a Banach space. ■

Example 2.2.2 Let $1 \leq p < \infty$. For $x = (x_1, x_2, ..., x_N) \in \mathbb{R}^N$, define

$$\|x\|_p \ = \ \left(\sum_{i=1}^{N} |x_i|^p\right)^{\frac{1}{p}}.$$

It is easy to see that conditions (i)-(ii) of Definition 2.1.2 are verified. We will presently prove the triangle inequality and thus $\mathbb{R}^N$ with the norm $\|.\|_p$ will become a normed linear space. It is, again, immediate to see that a sequence $\{x^{(n)}\}$ in $\mathbb{R}^N$ converges in this norm to $x \in \mathbb{R}^N$ if, and only if, for every $1 \leq i \leq N$, we have $x_i^{(n)} \to x_i$. Similarly, $\{x^{(n)}\}$ is Cauchy in this norm if, and only if, for every $1 \leq i \leq N$, the sequences $\{x_i^{(n)}\}$ are Cauchy in $\mathbb{R}$. Since $\mathbb{R}$ is complete, it now follows that $\mathbb{R}^N$ is also complete with respect to each of the norms $\|.\|_p$ defined above. Thus for each of these norms, $\mathbb{R}^N$ is a Banach space. ■

We now proceed to prove the triangle inequality for each of the norms $\|.\|_p$ for $1 \leq p < \infty$.

Definition 2.2.1 *Let $1 \leq p \leq \infty$. If $p = 1$, set $p^* = \infty$ and vice-versa. Otherwise, let $1 < p^* < \infty$ be such that*

$$\frac{1}{p} + \frac{1}{p^*} \ = \ 1 \tag{2.2.1}$$

The number p^* defined thus is called the **conjugate exponent** *of p.* ■

Lemma 2.2.1 *Let $1 < p < \infty$. Let p^* be its conjugate exponent. Then, if a and b are non-negative real numbers, we have*

$$a^{1/p} b^{1/p^*} \ \leq \ \frac{a}{p} + \frac{b}{p^*}. \tag{2.2.2}$$

Proof: Let $t \geq 1$ and consider the function

$$f(t) = k(t-1) - t^k + 1$$

for some $k \in (0,1)$. Then $f'(t) = k(1 - t^{k-1}) \geq 0$ since $k < 1$. Thus, f is an increasing function on $[1, \infty)$ and, since $f(1) = 0$, we immediately deduce that

$$t^k \leq k(t-1) + 1 \tag{2.2.3}$$

for $t \geq 1$ and $0 < k < 1$.

Now, if a or b is zero, then (2.2.2) is obviously true. So let us assume that $a \geq b > 0$.

The inequality (2.2.2) now follows by setting $t = a/b$ and $k = 1/p$ in (2.2.3) and using the definition of p^*. ■

Lemma 2.2.2 (**Hölder's inequality**) *Let $1 < p < \infty$. Let p^* be its conjugate exponent. Then, for $x, y \in \mathbb{R}^N$,*

$$\sum_{i=1}^{N} |x_i y_i| \leq \|x\|_p \|y\|_{p^*}. \tag{2.2.4}$$

Proof: Since the result is trivially true for $x = \mathbf{0}$ or $y = \mathbf{0}$, we can assume, without loss of generality, that both x and y are non-zero vectors. Then, set

$$a = \frac{|x_i|^p}{\|x\|_p^p} \text{ and } b = \frac{|y_i|^{p^*}}{\|y\|_{p^*}^{p^*}}$$

for a fixed $1 \leq i \leq N$. Then (2.2.2) yields

$$\frac{|x_i y_i|}{\|x\|_p \|y\|_{p^*}} \leq \frac{1}{p} \frac{|x_i|^p}{\|x\|_p^p} + \frac{1}{p^*} \frac{|y_i|^{p^*}}{\|y\|_{p^*}^{p^*}}.$$

Summing over the range of the index i, we get

$$\frac{\sum_{i=1}^N |x_i y_i|}{\|x\|_p \|y\|_{p^*}} \leq \frac{1}{p} + \frac{1}{p^*} = 1$$

which proves (2.2.4). ■

Lemma 2.2.3 (**Minkowski's Inequality**) *Let $1 \leq p < \infty$. Let $x, y \in \mathbb{R}^N$. Then*

$$\|x + y\|_p \leq \|x\|_p + \|y\|_p. \tag{2.2.5}$$

Proof: The proof is obvious if $p = 1$. Let us, therefore, assume that $1 < p < \infty$. Let p^* be the conjugate exponent. Then,

$$\begin{aligned}\sum_{i=1}^N |x_i + y_i|^p &\leq \sum_{i=1}^N |x_i + y_i|^{p-1}|x_i| + \sum_{i=1}^N |x_i + y_i|^{p-1}|y_i| \\ &\leq (\|x\|_p + \|y\|_p)\left(\sum_{i=1}^N |x_i + y_i|^{(p-1)p^*}\right)^{\frac{1}{p^*}}\end{aligned}$$

by a simple application of Hölder's inequality (2.2.4). But $(p-1)p^* = p$ by definition and so,

$$\|x + y\|_p^p \leq (\|x\|_p + \|y\|_p)\|x + y\|_p^{p/p^*}.$$

Since the result is obviously true when $x + y = \mathbf{0}$, we can assume, without loss of generality, that $x + y \neq \mathbf{0}$ and so, dividing both sides of the above inequality by $\|x + y\|^{p/p^*}$ and using, once again, the definition of p^*, we get (2.2.5). ■

Since Minkowski's inequality is exactly the triangle inequality for the norm $\|.\|_p$, our proof that $\mathbb{R}^N$ is a Banach space for each of these norms is complete.

Remark 2.2.1 The inequalities of Hölder and Minkowski are clearly true when x and $y \in \mathbb{C}^N$ and so $\mathbb{C}^N$ is also a Banach space for each of the norms $\|.\|_p$, $1 \leq p < \infty$. ■

Remark 2.2.2 When $p = 2$, we have $p^* = 2$ as well. In this case Hölder's inequality is known as the **Cauchy-Schwarz inequality**. The inequality (2.2.2), in this case, turns out to be the familiar inequality relating the arithmetic and geometric means of two positive real numbers. The norm $\|.\|_2$ is also called the *Euclidean norm* since it corrsponds to the usual Euclidean distance in $\mathbb{R}^N$. ■

Example 2.2.3 For $x = (x_1, ..., x_N) \in \mathbb{R}^N$, define

$$\|x\|_\infty = \max_{1 \leq i \leq N} |x_i|.$$

It is easy to verify that this also defines a norm on $\mathbb{R}^N$. Again convergence and the Cauchy criterion hold if and only if they hold componentwise and so $\mathbb{R}^N$ is a Banach space for this norm as well. Again all these assertions hold for $\mathbb{C}^N$ as well. It is immediate to see that Hölder's inequality is true when $p = 1$ as well. ■

Remark 2.2.3 The spaces $\mathbb{R}^N$ (or $\mathbb{C}^N$ when the base field is $\mathbb{C}$) with the norm $\|.\|_p$, where $1 \leq p \leq \infty$ are usually denoted by ℓ_p^N in the literature. ■

Remark 2.2.4 The notation $\|.\|_\infty$ for the norm defined in Example 2.2.3 can be 'justified' as follows. Let $x \in \mathbb{R}^N$. Assume that the maximum for $|x_i|$ is attained for a *single* index, say, i_0. Then,

$$\|x\|_p = |x_{i_0}| \left[1 + \sum_{i \neq i_0} \left(\frac{|x_i|}{|x_{i_0}|}\right)^p\right]^{\frac{1}{p}}.$$

Thus, since $|x_i|/|x_{i_0}| < 1$ for $i \neq i_0$, we get that

$$\|x\|_p \to \|x\|_\infty$$

when $p \to \infty$. ■

Example 2.2.4 We now consider sets of real (or complex) sequences

$$x = (x_1, x_2, ..., x_i, ...).$$

Let $1 \leq p < \infty$. We define the space

$$\ell_p = \left\{x \mid \sum_{i=1}^{\infty} |x_i|^p < \infty\right\}.$$

We define vector addition and scalar multiplication (over the corresponding field) componentwise, *i.e.* if $x = (x_i)$ and $y = (y_i)$ are sequences in ℓ_p and if α is a scalar, we set

$$x + y = (x_i + y_i) \text{ and } \alpha x = (\alpha x_i).$$

We also define

$$\|x\|_p = \left(\sum_{i=1}^{\infty} |x_i|^p\right)^{\frac{1}{p}}.$$

We will prove the triangle inequality for $\|.\|_p$ (which will also simultaneously show that ℓ_p is closed under vector addition and hence that it is a vector space). Since properties (i) and (ii) of Definition 2.1.2 are obvious, it will follow that $\|.\|_p$ defines a norm on ℓ_p.

Let x and $y \in \ell_p$. Then, for any positive integer N, we have

$$\begin{aligned}\textstyle\sum_{i=1}^{N} |x_i + y_i|^p &\leq \left[\left(\textstyle\sum_{i=1}^{N} |x_i|^p\right)^{\frac{1}{p}} + \left(\textstyle\sum_{i=1}^{N} |y_i|^p\right)^{\frac{1}{p}}\right]^p \\ &\leq [\|x\|_p + \|y\|_p]^p\end{aligned}$$

using Minkowski's inequality (2.2.5) for the integer N. Thus, since N was arbitrary, we deduce that

$$\sum_{i=1}^{\infty} |x_i + y_i|^p \leq [\|x\|_p + \|y\|_p]^p < \infty$$

which shows that $x + y \in \ell_p$ and also proves the triangle inequality for $\|.\|_p$. ■

Remark 2.2.5 The triangle inequality

$$\|x + y\|_p \leq \|x\|_p + \|y\|_p$$

for x and $y \in \ell_p$ is again referred to as Minkowski's inequality. We can also prove Hölder's inequality: if $x \in \ell_p$ and $y \in \ell_{p^*}$ where p^* is the conjugate exponent, then

$$\sum_{i=1}^{\infty} |x_i y_i| \leq \|x\|_p \|y\|_{p^*}. \ ■$$

Proposition 2.2.1 *Let $1 \leq p < \infty$. Then ℓ_p is a Banach space.*

Proof: We just need to prove the completeness of the space. Let $\{x^{(n)}\}$ be a Cauchy sequence in ℓ_p, *i.e.* given $\varepsilon > 0$, there exists N such that

$$\sum_{i=1}^{\infty} |x_i^{(m)} - x_i^{(l)}|^p < \varepsilon \tag{2.2.6}$$

for all $m \geq N$, $l \geq N$. Thus, it is clear that for each fixed subscript i, the sequence $\{x_i^{(n)}\}$ is Cauchy in $\mathbb{R}$ (or $\mathbb{C}$, as the case may be). Thus, there exists x_i such that $x_i^{(n)} \to x_i$ for each i. Set

$$x = (x_1, x_2, ..., x_i, ...).$$

We will first show that $x \in \ell_p$. Since $\{x^{(n)}\}$ is a Cauchy sequence, it is bounded. Thus, there exists a $C > 0$ such that

$$\|x^{(n)}\|_p^p \leq C, \quad \text{for all } n.$$

Let k be any fixed positive integer. Then,

$$\sum_{i=1}^{k} |x_i^{(n)}|^p \leq C$$

which implies that

$$\sum_{i=1}^{k} |x_i|^p \leq C.$$

Since k is arbitrary, this shows that

$$\sum_{i=1}^{\infty} |x_i|^p \leq C < \infty.$$

This shows that $x \in \ell_p$.

Now, for any positive integer k and all $m, l \geq N$, it follows from (2.2.6) that

$$\sum_{i=1}^{k} |x_i^{(m)} - x_i^{(l)}|^p < \varepsilon.$$

Passing to the limit as $l \to \infty$, we get that for any $m \geq N$ and for any k,

$$\sum_{i=1}^{k} |x_i^{(m)} - x_i|^p \leq \varepsilon.$$

Since k is arbitrary, we deduce that for $m \geq N$,

$$\|x^{(m)} - x\|_p^p \leq \varepsilon$$

i.e. $x^{(n)} \to x$ in ℓ_p. This completes the proof. ■

Example 2.2.5 Set

$$\ell_\infty = \left\{ x = (x_i) \mid \sup_{1 \leq i < \infty} |x_i| < +\infty \right\}$$

i.e. the space of all bounded real (or complex) sequences. This is clearly a vector space under componentwise addition and scalar multiplication. Define

$$\|x\|_\infty = \sup_{1 \leq i < \infty} |x_i|.$$

This makes ℓ_∞ a Banach space (check!). ■

Remark 2.2.6 Once again, Hölder's inequality holds for $p = 1$ as well. ■

Remark 2.2.7 For those readers who are acquainted with measure theory, the spaces ℓ_p^N and ℓ_p, for $1 \le p \le \infty$ are particular cases of the Lebesgue spaces $L^p(\mu)$ where $\{X, \mathcal{S}, \mu\}$ is a measure space. In the case of ℓ_p^N, we have $X = \{1, 2, ..., N\}$ and in the case of ℓ_p we have $X = \mathbb{N}$, the set of all natural numbers. In either case, the σ-algebra is the collection of all subsets of X and the measure μ is the counting measure. We will study L^p spaces in detail in Chapter 6. ■

Our final example is that of a function space.

Example 2.2.6 Let $C[0,1]$ denote the set of all continuous real valued functions on the closed interval $[0,1]$. This becomes a vector space under the operations of addition and scalar multiplications defined pointwise, *i.e.*

$$(f+g)(x) = f(x) + g(x) \text{ and } (\alpha f)(x) = \alpha f(x)$$

for f and $g \in C[0,1], \alpha \in \mathbb{R}$ and for $x \in [0,1]$. Define

$$\|f\| = \sup_{x\in[0,1]} |f(x)| \left(= \max_{x\in[0,1]} |f(x)| \right).$$

This is well defined since $[0,1]$ is compact and so every continuous function is bounded and attains its maximum. The verification that this defines a norm on $C[0,1]$ is routine and is left to the reader.

Let $\{f_n\}$ be a Cauchy sequence in $C[0,1]$. This implies that for every $\varepsilon > 0$, there exists a positive integer N such that, for all $x \in [0,1]$ and for all $n \ge N$ and $m \ge N$, we have

$$|f_n(x) - f_m(x)| < \varepsilon. \tag{2.2.7}$$

Thus the pointwise sequences $\{f_n(x)\}$ are all Cauchy and hence convergent. Define

$$f(x) = \lim_{n\to\infty} f_n(x).$$

We will show that the function f thus defined is in $C[0,1]$ and that $\|f_n - f\| \to 0$. This will show that $C[0,1]$, with the given norm, is a Banach space.

Let ε and N be as above. Then, keeping $n \geq N$ fixed and passing to the limit as $m \to \infty$ in (2.2.7), we get

$$|f_n(x) - f(x)| \ \leq \ \varepsilon \tag{2.2.8}$$

for all $n \geq N$ and for all $x \in [0,1]$. Fix a point $x_0 \in [0,1]$. Since f_N is continuous, there exists $\delta > 0$ such that for all $|x_0 - y| < \delta$, we have

$$|f_N(x_0) - f_N(y)| \ < \ \varepsilon.$$

Thus, if $|x_0 - y| < \delta$, we get, using the above inequality and also the inequality (2.2.8),

$$|f(x_0) - f(y)| \ \leq \ |f(x_0) - f_N(x_0)| + |f_N(x_0) - f_N(y)| + |f_N(y) - f(y)| \ \leq 3\varepsilon.$$

This proves that f is continuous and from (2.2.8), we see that $\|f_n - f\| \to 0$. ■

Remark 2.2.8 The convergence described in the preceding example is what is known in the literature as *uniform convergence*. The norm is often referred to as the 'sup-norm'. ■

We conclude this section by showing a standard method of producing new normed linear spaces from existing ones.Let V be a normed linear space and let W be a closed subspace of V, *i.e.* W is a linear subspace of V and is closed under the norm topology. We define an equivalence relation on V by

$$x \sim y \ \Leftrightarrow x - y \in W.$$

The equivalence class containing a vector $x \in V$ is called a coset and is denoted as $x + W$. It consists of all elements of the form $x + w$ where $w \in W$. The set of all cosets is called the **quotient space** and is denoted V/W. Addition and scalar multiplication on V/W are defined by

$$(x+W) + (y+W) \ = \ (x+y) + W \ \text{ and } \ \alpha(x+W) \ = \ \alpha x + W.$$

If $x \sim x'$ and $y \sim y'$, then, clearly, $x + y \sim x' + y'$ and $\alpha x \sim \alpha x'$, since W is a linear subspace of V. Thus, addition and scalar multiplication are well defined. Thus the quotient space becomes a vector space. On this, we define

$$\|x + W\|_{V/W} \ = \ \inf_{w \in W} \|x + w\|.$$

In other words, the 'norm' defined above is the infimum of the norms of all the elements in the coset and so, clearly, it is well defined.

Proposition 2.2.2 *Let V be a normed linear space and let W be a closed subspace. Then, $\|.\|_{V/W}$ defined above is a norm on the quotient space V/W. Further, if V is a Banach space, so is V/W.*

Proof: Clearly $\|x+W\|_{V/W} \geq 0$ for all $x \in V$. If $x+W = \mathbf{0}+W$ in V/W, we have $x \in W$; then $-x \in W$ and so $0 \leq \|x+W\|_{V/W} \leq \|x+(-x)\| = 0$ and so $\|x + W\|_{V/W} = 0$. Conversely, if $\|x + W\|_{V/W} = 0$, then, by definition, there exists a sequence $\{w_n\}$ in W such that $\|x + w_n\| \to 0$. This means that $w_n \to -x$ in V and, since W is closed, it follows that $-x \in W$ and so $x \in W$ as well. This means that $x \sim \mathbf{0}$, *i.e.* $x + W$ is the zero element of V/W.

If $\alpha \neq 0$, then $\alpha x + w = \alpha(x + w')$ where $w' = \alpha^{-1}w \in W$. From this it is easy to see that $\|\alpha x + W\|_{V/W} = |\alpha|\|x + W\|_{V/W}$. The case $\alpha = 0$ is obvious.

Finally, we prove the triangle inequality.

$$\begin{aligned}
\|x + y + W\|_{V/W} &= \inf\{\|x + y + w\| \mid w \in W\} \\
&= \inf\{\|x + y + w + w'\| \mid w, w' \in W\} \\
&\leq \inf\{\|x + w\| + \|y + w'\| \mid w, w' \in W\} \\
&= \inf\{\|x + w\| \mid w \in W\} + \inf\{\|y + w'\| \mid w' \in W\} \\
&= \|x + W\|_{V/W} + \|y + W\|_{V/W}.
\end{aligned}$$

Thus, V/W is a normed linear space. Now assume that V is complete. Let $\{x_n + W\}$ be a Cauchy sequence in V/W. Then, we can find a subsequence such that

$$\|(x_{n_k} + W) - (x_{n_{k+1}} + W)\|_{V/W} \ < \ \frac{1}{2^k}. \text{ (why?)}$$

Now choose $y_k \in x_{n_k} + W$ such that $\|y_k - y_{k+1}\| < 1/2^k$. Then the sequence $\{y_k\}$ is Cauchy (why?) and so, since V is complete, $y_k \to y$ in V. Thus

$$\|(x_{n_k} + W) - (y + W)\|_{V/W} \ \leq \ \|y_k - y\| \ \to \ 0.$$

Thus, the Cauchy sequence $\{x_n + W\}$ has a convergent subsequence $\{x_{n_k} + W\}$ and so the Cauchy sequence itself must be convergent and converge to the same limit (why?). Hence V/W is complete. ■

2.3 Continuous Linear Transformations

An important aspect of functional analysis is to study mappings between normed linear spaces which 'respect' the linear and topological structures. We make this notion precise in the following definition.

Definition 2.3.1 *Let V and W be normed linear spaces. A linear transformation $T : V \to W$ is said to be a* **continuous linear transformation** *or, a* **continuous linear operator**, *if it is continuous as a map between the topological spaces V and W (endowed with their norm topologies). If W is the base field, then a continuous linear transformation is called a* **continuous linear functional.** ■

Definition 2.3.2 *A subset of a normed linear space is* **bounded** *if it can be contained in a ball.* ■

The following proposition gives an important characterization of continuous linear transformations.

Proposition 2.3.1 *Let V and W be normed linear spaces and let $T : V \to W$ be a linear transformation. The following are equivalent:*
(i) *T is continuous.*
(ii) *T is continuous at $\mathbf{0}$.*
(iii) *There exists a constant $K > 0$ such that, for all $x \in V$,*

$$\|T(x)\|_W \;\leq\; K\|x\|_V \tag{2.3.1}$$

where $\|.\|_V$ and $\|.\|_W$ denote the respective norms in the spaces V and W.
(iv) *If $B = \{x \in V \mid \|x\|_V \leq 1\}$ is the (closed) unit ball in V, then $T(B)$ is a bounded set in W.*

Proof: (i) $\Leftrightarrow$ (ii) If T is continuous, then, clearly, it is continuous at $\mathbf{0} \in V$. Conversely, let T be continuous at $\mathbf{0} \in V$. Let $x \in V$ be arbitrary and let $x_n \to x$ in V. Then $x_n - x \to \mathbf{0}$ in V and so, $T(x_n - x) \to \mathbf{0}$ in W, *i.e.* $T(x_n) \to T(x)$ in W. Thus, T is continuous.

(ii) $\Leftrightarrow$ (iii) If T is continuous at $\mathbf{0} \in V$, there exists a $\delta > 0$ such that $\|x\|_V < \delta$ implies that $\|T(x)\|_W < 1$. For any $x \in X$, set $y = \frac{\delta}{2\|x\|_V}x$ so that $\|y\|_V = \delta/2 < \delta$ and so $\|T(y)\|_W < 1$. By linearity, it follows that

$$\|T(x)\|_V \;\leq\; \frac{2}{\delta}\|x\|_V$$

which proves (2.3.1) with $K = 2/\delta$. Conversely, if (2.3.1) is true, then whenever $x_n \to \mathbf{0}$ in V, it follows that $T(x_n) \to \mathbf{0}$ in W, *i.e.* T is continuous at $\mathbf{0}$.

(iii) $\Leftrightarrow$ (iv) By virtue of (2.3.1), it follows that $\|T(x)\|_W \leq K$ for all $x \in B$. Thus, $T(B)$ is bounded in W. Conversely, if $T(B)$ is bounded in W, there exists a $K > 0$ such that $\|T(x)\|_W \leq K$ for all $x \in B$. Now, if $\mathbf{0} \neq x \in V$ is arbitrary, set $y = x/\|x\|$. Then $\|T(y)\|_W \leq K$ from which (2.3.1) follows, by linearity. ■

Remark 2.3.1 Continuous linear transformations are also known as **bounded linear transformations** since they map bounded sets into bounded sets. ■

The above proposition inspires the following definition.

Definition 2.3.3 *Let V and W be normed linear spaces and let $T : V \to W$ be a continuous linear transformation. Let B be the closed unit ball in V. The* **norm** *of T, denoted $\|T\|$, is given by*

$$\|T\| \;=\; \sup_{x \in B} \|T(x)\|_W. \blacksquare \tag{2.3.2}$$

The following proposition gives alternative characterizations of the norm of a continuous linear transformation.

Proposition 2.3.2 *Let V and W be normed linear spaces and let $T : V \to W$ be a continuous linear transformation. Then*

$$\begin{aligned}
\|T\| &= \sup\{\|T(x)\|_W \mid \|x\|_V = 1\} \\
&= \sup\{\|T(x)\|_W/\|x\|_V \mid \mathbf{0} \neq x \in V\} \\
&= \inf\{K > 0 \mid \|T(x)\|_W \;\leq\; K\|x\|_V \;\; \textit{for all} \;\; x \in V\}.
\end{aligned}$$

Proof: Let us set

$$\begin{aligned}
\alpha &= \sup\{\|T(x)\|_W \mid \|x\|_V = 1\}, \\
\beta &= \sup\{\|T(x)\|_W/\|x\|_V \mid \mathbf{0} \neq x \in V\} \text{ and} \\
\gamma &= \inf\{K > 0 \mid \|T(x)\|_W \;\leq\; K\|x\|_V \;\; \text{for all} \;\; x \in V\}.
\end{aligned}$$

Clearly, $\alpha \leq \beta$. If x is a non-zero vector in V, then $x/\|x\|_V$ has unit norm and $\|T(x)\|_W/\|x\|_V = \|T(x/\|x\|_V)\|_W$. This shows that we also have $\beta \leq \alpha$. If $K > 0$ is any number in the set defining γ, then it follows immediately that $\beta \leq K$ and so, *a fortiori*, we have $\beta \leq \gamma$. Now, we also have that $\|T(x)\|_W \leq \beta\|x\|_V$ for all $x \in V$ and so, by definition, $\gamma \leq \beta$. Thus, we have

$$\alpha \;=\; \beta \;=\; \gamma.$$

Clearly $\|T\| \geq \alpha$ by definition. If K is in the set defining γ, then, for all $x \in V$ such that $\|x\|_V \leq 1$, we have $\|T(x)\|_W \leq K$ and so $\|T\| \leq K$. Thus, we get that $\|T\| \leq \gamma = \alpha$. Thus we get that

$$\|T\| \;=\; \alpha \;=\; \beta \;=\; \gamma. \blacksquare$$

Corollary 2.3.1 *If V and W are normed linear spaces and if $T : V \to W$ is a continuous linear transformation, then*

$$\|T(x)\|_W \;\leq\; \|T\|\|x\|_V \tag{2.3.3}$$

for all $x \in V$. $\blacksquare$

Let V and W be normed linear spaces. Let us denote by $\mathcal{L}(V, W)$, the set of all continuous linear maps from V into W. If T_1 and T_2 are such maps, let us define $T_1 + T_2$ by

$$(T_1 + T_2)(x) \;=\; T_1(x) + T_2(x)$$

for all $x \in V$. Clearly, $T_1 + T_2$ is also a linear transformation. Now,

$$\|(T_1 + T_2)(x)\|_W \;\leq\; \|T_1(x)\|_W + \|T_2(x)\|_W \;\leq\; (\|T_1\| + \|T_2\|)\|x\|_V$$

by virtue of the triangle inequality and the above corollary. Thus, it follows that $T_1 + T_2$ is also a continuous linear transformation and that

$$\|T_1 + T_2\| \;\leq\; \|T_1\| + \|T_2\|.$$

Similarly, if T is a continuous linear transformation and if α is a scalar, we define

$$(\alpha T)(x) \;=\; \alpha T(x)$$

for all $x \in V$. It is then easy to see that αT is also continuous and that

$$\|\alpha T\| \;=\; |\alpha|\|T\|.$$

The zero element of $\mathcal{L}(V, W)$ is the trivial map which maps every element of V into the null vector of W. The element $-T$ is defined by $(-T)(x) = -T(x)$. Thus, $\mathcal{L}(V, W)$ is a vector space; in fact, it is a normed linear space for the norm of a continuous linear transformation defined above.

Proposition 2.3.3 *Let V and W be normed linear spaces. If W is complete, then $\mathcal{L}(V, W)$ is also complete.*

Proof: Let $\{T_n\}$ be a Cauchy sequence in $\mathcal{L}(V, W)$. Then, given $\varepsilon > 0$, we can find a positive integer N such that, for all m and $n \geq N$, we have

$$\|T_n - T_m\| \ < \ \varepsilon.$$

Let $x \in V$. Then,

$$\|T_n(x) - T_m(x)\|_W \ \leq \ \|T_n - T_m\|\|x\|_V$$

and so it follows that the sequence $\{T_n(x)\}$ is Cauchy in W. Since W is complete, this sequence is convergent. Let us define

$$T(x) \ = \ \lim_{n\to\infty} T_n(x).$$

Clearly, the map $x \mapsto T(x)$ is linear. We will show that it is continuous and that $\|T_n - T\| \to 0$. This will complete the proof.

Since the sequence $\{T_n\}$ is Cauchy, it is bounded, *i.e.* there exists $M > 0$ such that, for all positive integers n, we have $\|T_n\| \leq M$. Now, since for any $x \in V$, we have $\|T_n(x)\|_W \leq \|T_n\|\|x\|_V \leq M\|x\|_V$, it follows, on passing to the limit as $n \to \infty$, that, for all $x \in V$,

$$\|T(x)\|_W \ \leq M\|x\|_V.$$

Thus, T is continuous and so $T \in \mathcal{L}(V, W)$.

Let $\varepsilon > 0$ and let N be as defined earlier by the Cauchy property of the given sequence. Let B be the closed unit ball in V. For all $x \in B$, we have

$$\|T_n(x) - T_m(x)\|_W \ \leq \ \|T_n - T_m\| \ < \ \varepsilon.$$

Keeping n fixed and letting m tend to infinity, we get

$$\|T_n(x) - T(x)\|_W \ \leq \ \varepsilon$$

for all $x \in B$. This shows that, for $n \geq N$, we have $\|T_n - T\| \leq \varepsilon$, *i.e.* $T_n \to T$ in $\mathcal{L}(V, W)$. This completes the proof. ■

In particular, since the scalar field is a Banach space over itself (cf. Example 2.2.1), the set of all continuous linear functionals $\mathcal{L}(V, \mathbb{R})$ (or, $\mathcal{L}(V, \mathbb{C})$, as the case may be) is always a Banach space.

Definition 2.3.4 *Let V be a normed linear space. The space of all continuous linear functionals on V is a Banach space and is called the* **dual space** *of V. It is denoted by V^*.* ■

Another particular case is when $W = V$. In this case we write $\mathcal{L}(V)$ for the space of all continuous linear operators instead of $\mathcal{L}(V, V)$. This space is Banach if V is Banach. On this space, we have a third operation (after addition and scalar multiplication) namely, composition of operators: if T_1 and T_2 are continuous linear operators, we define T_1T_2 by

$$(T_1T_2)(x) \;=\; T_1(T_2(x)).$$

Now, for any $x \in V$, we have

$$\|(T_1T_2)(x)\|_V \;\leq\; \|T_1\|\|T_2(x)\|_V \;\leq\; \|T_1\|\|T_2\|\|x\|_V.$$

Thus, T_1T_2 is also a continuous linear operator and, further,

$$\|T_1T_2\| \;\leq\; \|T_1\|\|T_2\|. \tag{2.3.4}$$

Further, multiplication is a continuous operation. Indeed, if $T_n \to T$ and $T'_n \to T'$ in $\mathcal{L}(V)$, we have

$$\|T_nT'_n - TT'\| \;\leq\; \|T_n\|\|T'_n - T'\| + \|T'\|\|T_n - T\|.$$

Since $\|T_n\|$ is bounded independent of n, it follows that $T_nT'_n \to TT'$. Finally, if I is the identity mapping *i.e.* $I(x) = x$ for all $x \in V$, we have

$$\|I\| = 1.$$

Definition 2.3.5 *A Banach space V on which we have a multiplication operation $(x, y) \in V \times V \mapsto xy \in V$ such that addition and multiplication make it a ring and such that*

$$\|xy\| \;\leq\; \|x\|\|y\| \quad \textit{and} \quad \|\mathbf{1}\| \;=\; 1$$

where $\mathbf{1}$ is the multiplicative identity in V, is called a **Banach algebra**. ■

Thus, $\mathcal{L}(V)$, where V is a Banach space, is a Banach algebra.

Let us now study various examples of continuous linear transformations.

Example 2.3.1 Any linear transformation $T : \mathbb{R}^N \to \mathbb{R}^M$ is given by an $M \times N$ matrix. Assume that $\mathbb{R}^N$ is the space ℓ_1^N. Then, for any norm on $\mathbb{R}^M$, any linear transformation is continuous. Let $\{\mathbf{e}_1, ..., \mathbf{e}_N\}$ be the standard basis of $\mathbb{R}^N$ (cf. Example 1.1.2). If $x = (x_1, ..., x_N) \in \mathbb{R}^N$, then $x = \sum_{i=1}^N x_i \mathbf{e}_i$. Then $T(x) = \sum_{i=1}^N x_i T(\mathbf{e}_i)$. Thus,

$$\|T(x)\|_{\mathbb{R}^M} \;\le\; K \sum_{i=1}^{N} |x_i| \;=\; K\|x\|_1$$

where

$$K \;=\; \max_{1 \le i \le N} \|T(\mathbf{e}_i)\|_{\mathbb{R}^M}. \;\blacksquare$$

Example 2.3.2 Let $a_1, ..., a_N$ be scalars. For $x = (x_1, ..., x_N) \in \mathbb{R}^N$, define

$$f(x) \;=\; \sum_{i=1}^{N} a_i x_i.$$

Then f is a linear functional on $\mathbb{R}^N$ which is continuous if $\mathbb{R}^N = \ell_1^N$. $\blacksquare$

We will see later that these transformations and functionals are continuous for *any* norm defined on $\mathbb{R}^N$.

Example 2.3.3 Let $x = (x_i) \in \ell_2$. Define

$$T(x) \;=\; \left(\frac{x_1}{1}, \frac{x_2}{2}, ..., \frac{x_i}{i}, ...\right).$$

Then, since

$$\sum_{i=1}^{\infty} \left|\frac{x_i}{i}\right|^2 \;\le\; \sum_{i=1}^{\infty} |x_i|^2 \;<\; \infty,$$

we have that T is a continuous linear operator on ℓ_2 and that $\|T\| \le 1$. The map T is not onto. In fact, the range of T consists of all square summable sequences (y_i) such that

$$\sum_{i=1}^{\infty} i^2 |y_i|^2 \;<\; \infty. \;\blacksquare$$

Example 2.3.4 Let $1 \le p \le \infty$ and let p^* be its conjugate exponent. For $x \in \ell_p$ and $y \in \ell_{p^*}$, define

$$f_y(x) \;=\; \sum_{i=1}^{\infty} x_i y_i.$$

Then, by Hölder's inequality, we have

$$|f_y(x)| \;\le\; \sum_{i=1}^{\infty} |x_i y_i| \;\le\; \|x\|_p \|y\|_{p^*}.$$

Thus f_y defines a continuous linear functional on ℓ_p and

$$\|f_y\| \;\le\; \|y\|_{p^*}.$$

We will see later that, when $1 \le p < \infty$, all continuous linear functionals on ℓ_p occur only in this way and that the last inequality is, in fact, an equality. ■

Example 2.3.5 Consider an infinite matrix $(a_{ij})_{i,j=1}^{\infty}$. This can be used to define a linear mapping on ℓ_p as follows. Let $x = (x_i) \in \ell_p$. Define a sequence $A(x)$ by

$$A(x)_i \;=\; \sum_{j=1}^{\infty} a_{ij} x_j.$$

Showing that A defines a continuous linear map of ℓ_p into itself is usually a non-trivial problem and some examples are given in the exercises at the end of this chapter. We now give an example (due to Schur).

Assume that $a_{ij} \ge 0$ for all i and j. Assume further there exists a sequence $\{p_i\}$ of positive real numbers and $\beta > 0$ and $\gamma > 0$ such that

$$\sum_{i=1}^{\infty} a_{ij} p_i \;\le\; \beta p_j$$

for all $j \in \mathbb{N}$ and also such that

$$\sum_{j=1}^{\infty} a_{ij} p_j \;\le\; \gamma p_i$$

for all $i \in \mathbb{N}$. Then $A \in \mathcal{L}(\ell_2)$ and $\|A\|^2 \le \beta\gamma$.

To see this, let $x = (x_i) \in \ell_2$. We write

$$\sum_{j=1}^{\infty} a_{ij} x_j \;=\; \sum_{j=1}^{\infty} \sqrt{a_{ij}}\sqrt{p_j}\,\frac{\sqrt{a_{ij}}\,x_j}{\sqrt{p_j}}.$$

Applying the Cauchy-Schwarz inequality, this yields

$$|A(x)_i|^2 \;\le\; \left(\sum_{j=1}^{\infty} a_{ij} p_j\right)\left(\sum_{j=1}^{\infty} \frac{a_{ij}|x_j|^2}{p_j}\right).$$

It follows from the hypotheses that

$$\begin{aligned}\sum_{i=1}^{\infty}|A(x)_i|^2 &\leq \sum_{i=1}^{\infty}\gamma p_i\sum_{j=1}^{\infty}\frac{a_{ij}|x_j|^2}{p_j}\\ &= \gamma\sum_{j=1}^{\infty}\frac{|x_j|^2}{p_j}\sum_{i=1}^{\infty}a_{ij}p_i\\ &\leq \gamma\sum_{j=1}^{\infty}\frac{|x_j|^2}{p_j}\beta p_j\\ &= \beta\gamma\|x\|_2^2\end{aligned}$$

which establishes the claim.

An interesting particular case is that of the *Hilbert matrix.* Set

$$a_{ij} = \frac{1}{i+j+1}$$

for $0 \leq i, j \leq \infty$. Set $p_i = 1/\sqrt{i+\frac{1}{2}}$. Since the matrix is symmetric, it suffices to check one of the two conditions. Now,

$$\begin{aligned}\sum_{i=0}^{\infty}a_{ij}p_i &= \sum_{i=0}^{\infty}\frac{1}{(i+\frac{1}{2}+j+\frac{1}{2})\sqrt{i+\frac{1}{2}}}\\ &< \int_0^{\infty}\frac{dx}{(x+j+\frac{1}{2})\sqrt{x}}\\ &= 2\int_0^{\infty}\frac{dt}{t^2+j+\frac{1}{2}}\\ &= \frac{\pi}{\sqrt{j+\frac{1}{2}}}.\end{aligned}$$

Thus, by Schur's test, the matrix defines a continuous linear operator A on ℓ_2 whose norm is less than, or equal to π. (In fact, it has been shown by Hardy, Littlewood and Polya that the norm is exactly π.) ■

Example 2.3.6 (Cesàro Operator) Let $x = (x_i) \in \ell_p$ where $1 < p < \infty$. Define

$$(T(x))_n = \frac{x_1 + \cdots + x_n}{n}.$$

We show that $T \in \mathcal{L}(\ell_p)$ and that

$$\|T\| \leq \frac{p}{p-1}.$$

Indeed,

$$|T(x)_n| \leq \frac{|x_1| + \cdots + |x_n|}{n}.$$

Set $A_n = |x_1| + \cdots + |x_n|$ and $\alpha_n = A_n/n$, for $n \geq 1$ and set $x_0 = A_0 = 0$. Then

$$\begin{aligned}
\alpha_n^p - \tfrac{p}{p-1}\alpha_n^{p-1}|x_n| &= \alpha_n^p - \tfrac{p}{p-1}\alpha_n^{p-1}(n\alpha_n - (n-1)\alpha_{n-1}) \\
&= \left(1 - \tfrac{np}{p-1}\right)\alpha_n^p + \tfrac{(n-1)p}{p-1}\alpha_n^{p-1}\alpha_{n-1} \\
&= \left(1 - \tfrac{np}{p-1}\right)\alpha_n^p + \tfrac{(n-1)p}{p-1}(\alpha_n^p)^{\frac{p-1}{p}}(\alpha_{n-1}^p)^{\frac{1}{p}}.
\end{aligned}$$

Recall that $p/(p-1) = p^*$, the conjugate exponent of p. Thus by Lemma 2.2.1, we get

$$\begin{aligned}
\alpha_n^p - \tfrac{p}{p-1}\alpha_n^{p-1}|x_n| &\leq \left(1 - \tfrac{np}{p-1}\right)\alpha_n^p + \tfrac{(n-1)p}{p-1}\left(\tfrac{p-1}{p}\alpha_n^p + \tfrac{1}{p}\alpha_{n-1}^p\right) \\
&= \tfrac{1}{p-1}[(n-1)\alpha_{n-1}^p - n\alpha_n^p].
\end{aligned}$$

Fix a positive integer N. Then summing both sides over n running between 1 and N, and noticing that the right-hand side is a telescoping sum, we get

$$\sum_{n=1}^{N}\alpha_n^p - \frac{p}{p-1}\sum_{n=1}^{N}\alpha_n^{p-1}|x_n| \;\leq\; -\frac{N}{p-1}\alpha_N^p \;\leq\; 0.$$

By an application of Hölder's inequality, we now deduce that

$$\sum_{n=1}^{N}\alpha_n^p \;\leq\; \frac{p}{p-1}\sum_{n=1}^{N}\alpha_n^{p-1}|x_n| \;\leq\; \frac{p}{p-1}\left(\sum_{n=1}^{N}\alpha_n^p\right)^{\frac{p-1}{p}}\left(\sum_{n=1}^{N}|x_n|^p\right)^{\frac{1}{p}}.$$

Dividing both sides by $(\sum_{n=1}^{N}\alpha_n^p)^{((p-1)/p)}$, which.is strictly positive for non-zero x, we get

$$\left(\sum_{n=1}^{N}\alpha_n^p\right)^{\frac{1}{p}} \;\leq\; \frac{p}{p-1}\left(\sum_{n=1}^{N}|x_n|^p\right)^{\frac{1}{p}}.$$

Since N was arbitrarily chosen, we deduce that

$$\|T(x)\|_p \;\leq\; \frac{p}{p-1}\|x\|_p$$

which establishes our claim. (In fact, Hardy, Littlewood and Polya also show that $\|T\| = p/(p-1)$.) ■

Example 2.3.7 (Volterra integral operator) Let $K : [0,1] \times [0,1] \to \mathbb{R}$ be a continuous function. For $f \in \mathcal{C}[0,1]$ and $s \in [0,1]$, define

$$T(f)(s) \;=\; \int_0^s K(s,t)f(t)dt.$$

Since K is continuous on the compact set $[0,1] \times [0,1]$, it is bounded and uniformly continuous. Assume that, for all s and t in $[0,1]$, we have

$$|K(s,t)| \;\leq\; \kappa.$$

Further, given $\varepsilon > 0$, there exists $\delta > 0$ such that, whenever $|s_1 - s_2| < \delta$, we have

$$|K(s_1,t) - K(s_2,t)| \;<\; \varepsilon$$

for all $t \in [0,1]$, by virtue of the uniform continuity. Without loss of generality, we can assume that $\delta \leq \varepsilon$. Thus,

$$T(f)(s_1) - T(f)(s_2) \;=\; \int_0^{s_1} (K(s_1,t) - K(s_2,t))f(t)dt + \int_{s_2}^{s_1} K(s_2,t)f(t)dt.$$

If $\|f\|$ denotes the norm of $f \in \mathcal{C}[0,1]$ as defined in Example 2.2.6, we have

$$|T(f)(s_1) - T(f)(s_2)| \;\leq\; \varepsilon\|f\|s_1 + \delta\kappa\|f\| \;<\; (1+\kappa)\|f\|\varepsilon$$

whenever $|s_1 - s_2| < \delta$. This shows that $T(f)$ is a continuous function. The mapping T being clearly linear, it thus defines a linear operator on $\mathcal{C}[0,1]$. Further,

$$|T(f)(s)| \;\leq\; \kappa\|f\|s \;\leq\; \kappa\|f\|.$$

Thus T is a continuous linear operator on $\mathcal{C}[0,1]$ and $\|T\| \leq \kappa$. ■

So far, we have been seeing examples of continuous linear transformations. We now give an example of a linear transformation which is *not* continuous.

Example 2.3.8 Consider the space $\mathcal{C}^1[0,1]$ of continuous functions on $[0,1]$ which are continuously differentiable on $(0,1)$ and whose derivatives can be extended continuously to $[0,1]$. This is a subspace of $\mathcal{C}[0,1]$. Let both these spaces be endowed with the 'sup-norm' (cf. Example 2.2.6). Then, the map $T : \mathcal{C}^1[0,1] \to \mathcal{C}[0,1]$ defined by $T(f) = f'$, where f' denotes the derivative of f, is not continuous. To see this, consider the sequence of functions $\{f_n\}$ defined by $f_n(t) = t^n$ for $n \geq 1$. Then, it is easy to see that $\|f'_n\| = n$ while $\|f_n\| = 1$. Hence there can be no constant $C > 0$ such that $\|T(f)\| \leq C\|f\|$ for all $f \in \mathcal{C}^1[0,1]$. Thus, T is not continuous. ■

Definition 2.3.6 *Let V be a normed linear space and let $T \in \mathcal{L}(V)$ be a bijection. If T^{-1} is also continuous, then T is said to* **invertible** *or an* **isomorphism**. ■

Note: When dealing with normed linear spaces, the word isomorphism is understood in the topological sense: not only is it an isomorphism in the usual algberaic sense *i.e.* it is linear and is a bijection, but it also implies that both the mapping and its inverse are continuous.

Definition 2.3.7 *Two norms defined on the same vector space are said to be* **equivalent** *if the topologies induced by these two norms coincide.* ■

Proposition 2.3.4 *Let V be a vector space and let $\|.\|_{(1)}$ and $\|.\|_{(2)}$ be two norms defined on it. The two norms are equivalent if, and only if, there exist two constants $C_1 > 0$ and $C_2 > 0$ such that, for all $x \in V$, we have*

$$C_1\|x\|_{(1)} \ \leq\ \|x\|_{(2)} \ \leq\ C_2\|x\|_{(1)}.$$

Proof: The topologies induced by the two norms coincide if, and only if, the identity mapping

$$I : \{V, \|\cdot\|_{(1)}\} \to \{V, \|\cdot\|_{(2)}\}$$

is an isomorphism. This is equivalent to saying that there exist two constants $K_1 > 0$ and $K_2 > 0$ such that

$$\|x\|_{(2)} \ \leq\ K_2\|x\|_{(1)} \quad \text{and} \quad \|x\|_{(1)} \ \leq\ K_1\|x\|_{(2)}$$

for all $x \in V$. This proves the proposition on setting $C_1 = K_1^{-1}$ and $C_2 = K_2$. ■

Example 2.3.9 Let $x = (x_1, ..., x_N) \in \mathbb{R}^N$ (or $\mathbb{C}^N$). Then, clearly,

$$\|x\|_\infty \ \leq\ \|x\|_1 \ \leq\ N\|x\|_\infty.$$

Thus these two norms are equivalent and the topologies induced on $\mathbb{R}^N$ (respectively, $\mathbb{C}^N$) by the norms $\|.\|_1$ and $\|.\|_\infty$ coincide. It is a simple matter to check that the topology induced by $\|.\|_\infty$ is none other than the product topology on $\mathbb{R}^N$ (respectively $\mathbb{C}^N$) when $\mathbb{R}$ (respectively $\mathbb{C}$) is given its usual topology for every component. ■

We will, in fact, now prove a much stronger result.

Proposition 2.3.5 *Any two norms on a finite dimensional vector space are equivalent.*

Proof: Let V be a finite dimensional normed linear space with dimension N. We will show that V is isomorphic to the space ℓ_1^N. Thus, given two norms on V, it will be isomorphic to ℓ_1^N for each of those norms and from this we will deduce the equivalence of the norms.

Step 1. Let $\{\mathbf{e}_1, ..., \mathbf{e}_N\}$ be the standard basis of ℓ_1^N. Fix a basis $\{v_1, ..., v_N\}$ for V. Define $T : \ell_1^N \to V$ by setting $T(\mathbf{e}_i) = v_i$ for all $1 \le i \le N$ and then extending T linearly to all of ℓ_1^N. Clearly T is a bijection and an identical argument as in Example 2.3.1 shows that T is continuous.

Step 2. Assume, if possible, that T^{-1} is not continuous. Then the continuity must fail at $\mathbf{0}$ and so we can find a sequence $\{y_n\}$ in V and a real number $\varepsilon > 0$ such that $\|T^{-1}(y_n)\|_1 \ge \varepsilon > 0$ while $y_n \to \mathbf{0}$. Set $z_n = y_n/\|T^{-1}(y_n)\|_1$. Then, $z_n \to \mathbf{0}$ and $\|T^{-1}(z_n)\|_1 = 1$. Now, the set

$$B = \{x \in \ell_1^N \mid \|x\|_1 \le 1\}$$

is compact. To see this, observe that B is a closed set and is contained in the set $\Pi_{i=1}^N[-1,1]$, which, being the product of compact sets, is compact in the topology induced by the norm $\|.\|_\infty$, and hence, in the topology induced by the norm $\|.\|_1$ as well (cf. Example 2.3.9). Thus, B is also compact. Consequently, it is also sequentially compact and so, there exists a subsequence $\{z_{n_k}\}$ such that $\{T^{-1}(z_{n_k})\}$ is convergent. Let $T^{-1}(z_{n_k}) \to x$, where $\|x\|_1 = 1$. Since T is continuous, we then deduce that $z_{n_k} \to T(x)$ which then implies that $T(x) = \mathbf{0}$. But T is a one-one map and $\|x\|_1 = 1$ implies that $x \neq \mathbf{0}$ which shows that $T(x) \neq \mathbf{0}$ as well. This gives us a contradiction. Hence T^{-1} must also be continuous.

Step 3. Thus, whatever be the norm on V, the same map T is always an isomorphism between ℓ_1^N and V which implies in turn that the identity map on V, considered as a map of normed linear spaces when V is provided with two different norms, must be an isomorphism as well. Hence any two norms on V are equivalent. ■

Remark 2.3.2 We mentioned in Proposition 1.2.5 that a set in $\mathbb{R}^N$ (or $\mathbb{C}^N$) is compact if, and only if, it is closed and bounded. However, the topology given there is the 'usual' topolgy, *viz.* that of ℓ_2^N. It is for this reason that we needed to give a separate argument to show that S (which is also closed and bounded) is compact in ℓ_1^N. However, thanks to the preceding proposition, we now know that the topology is the same for all the spaces ℓ_p^N or, for any other norm on $\mathbb{R}^N$ (respectively, $\mathbb{C}^N$) and so a subset thereof will be compact if, and only if, it is bounded and closed. ■

Corollary 2.3.2 *Any finite dimensional normed linear space is complete. In particular, any finite dimensional subspace of a normed linear space is closed.* ■

Example 2.3.10 Let $f \in \mathcal{C}[0,1]$. Define

$$\|f\|_1 \;=\; \int_0^1 |f(t)|\; dt.$$

It is simple to check that this defines a norm on $\mathcal{C}[0,1]$. Consider the sequence $\{f_n\}$ defined by

$$f_n(x) = \begin{cases} 1 - nx, & \text{for } 0 \le x \le \frac{1}{n} \\ 0, & \text{for } \frac{1}{n} \le x \le 1. \end{cases}$$

Clearly, $\|f_n\| = 1$ for all n (where $\|.\|$ denotes the usual 'sup' norm) while $\|f_n\|_1 = \int_0^1 f_n(t)\; dt = 1/2n$ which tends to zero as n tends to infinity. Thus it is clear that these two norms cannot be equivalent. Thus in infinite dimensional spaces, two norms are not, in general, equivalent. ■

Since all norms on $\mathbb{R}^N$ (respectively, $\mathbb{C}^N$) generate the same topology, it is now clear that any matrix generates a continuous linear transformation, whatever the norm on that space may be. Thus, if $\mathbf{T}$ is an $N \times N$ matrix and if $\|.\|$ is a norm on $\mathbb{R}^N$ (respectively, $\mathbb{C}^N$), we can define

$$\|T\| \;=\; \sup_{\|x\| \le 1,\; x \ne \mathbf{0}} \frac{\|Tx\|}{\|x\|} \tag{2.3.5}$$

or, via any of the other equivalent formulations as in Proposition 2.3.2. Since the unit ball (and hence the unit sphere) is compact, the 'sup' above is, in fact, a 'max'. If $\mathbf{T}$ and $\mathbf{S}$ are matrices of order N, then $\mathbf{TS}$ represents the composition of the corresponding linear transformations and so we also have

$$\|\mathbf{TS}\| \leq \|\mathbf{T}\|.\|\mathbf{S}\|. \tag{2.3.6}$$

Let $\mathcal{M}_N$ denote the set of all matrices of order N with entries from the corresponding field. This itself (under the operations of matrix addition and scalar multiplication) is a vector space (of dimension N^2). Any norm on this space which satisfies (2.3.6) is called **a matrix norm**. If such a norm were induced by a vector norm on $\mathbb{R}^N$ (respectively, $\mathbb{C}^N$) via (2.3.5), then we always have

$$\|\mathbf{I}\| = 1$$

for the identity matrix $\mathbf{I}$.

In particular, for the vector norms $\|.\|_p$ defining the spaces ℓ_p^N for $1 \leq p \leq \infty$, we denote the induced matrix norms by $\|.\|_{p,N}$.

Example 2.3.11 Since $\mathcal{M}_N$ is a vector space of dimension N^2 over the corresponding field, we can string out its rows to form a vector of that dimension and define the usual Euclidean norm. Thus, if $\mathbf{T} = (t_{ij})$, then define

$$\|\mathbf{T}\|_E = \left(\sum_{i,j=1}^{N} |t_{ij}|^2\right)^{\frac{1}{2}} = \sqrt{\mathrm{tr}(\mathbf{T}^*\mathbf{T})}.$$

This obviously defines a norm on $\mathcal{M}_N$. It is also a matrix norm. For this, we only need to check the validity of (2.3.6). It $\mathbf{T} = (t_{ij})$ and $\mathbf{S} = (s_{ij})$, then

$$\begin{aligned} \|\mathbf{TS}\|_E^2 &= \textstyle\sum_{i,j=1}^N \left|\sum_{k=1}^N t_{ik}s_{kj}\right|^2 \\ &\leq \textstyle\sum_{i,j=1}^N \left(\sum_{k=1}^N |t_{ik}|^2\right)\left(\sum_{k=1}^N |s_{kj}|^2\right) \end{aligned}$$

by the Cauchy-Schwarz inequality. Thus,we get

$$\|\mathbf{TS}\|_E^2 \leq \sum_{i,k=1}^{N} |t_{ik}|^2 \sum_{k,j=1}^{N} |s_{kj}|^2 = \|\mathbf{T}\|_E^2\|\mathbf{S}\|_E^2$$

which shows that $\|.\|_E$ is indeed a matrix norm. This is also known as the *Hilbert-Schmidt norm*.

However, notice that this norm is not induced by any vector norm when $N \geq 2$. Indeed to see this, observe that $\|\mathbf{I}\|_E = \sqrt{N} \neq 1$. ■

The proof of the Proposition 2.3.5 depended crucially on the fact that the unit ball in a finite dimensional space is compact. In fact this property characterizes finite dimensional spaces, which we now proceed to show. We begin with a very useful technical result.

Let V be a normed linear space. If $E \subset V$, we define the distance of a vector $x \in V$ from E as

$$d(x, E) = \inf_{y \in E} \|x - y\|.$$

This is the same notion of the distance of a point from a subset in a metric space if we look at V as a metric space for the metric $d(x, y) = \|x - y\|$.

Lemma 2.3.1 (**Riesz' Lemma**) *Let V be a normed linear space and let $W \subset V$ be a closed and proper subspace. Then, for every $\varepsilon > 0$, we can find a vector $u \in V$ (depending on ε) such that*

$$\|u\| = 1 \ \textit{ and } \ d(u, W) \geq 1 - \varepsilon.$$

Proof: Since W is a proper subspace, there exists $v \in V \backslash W$, so that $\delta = d(v, W) > 0$. Now, choose $w \in W$ such that

$$\delta \leq \|v - w\| \leq \frac{\delta}{1 - \varepsilon}.$$

Set $u = (v - w)/\|v - w\|$ so that $\|u\| = 1$. Let $z \in W$ be an arbitrary element. Then

$$\|u - z\| = \|v - (w + \|v - w\|z)\|/\|v - w\| \geq \delta/(\delta/(1 - \varepsilon)) = 1 - \varepsilon$$

by the definition of δ (since $w + \|v - w\|z \in W$) and the choice of w. This completes the proof. ■

Proposition 2.3.6 *A normed linear space V is finite dimensional if, and only if, the closed unit ball in V, i.e. the set*

$$B = \{x \in V \mid \|x\| \leq 1\},$$

is compact.

Proof: Assume that V is finite dimensional, with dimension N. Let $T : \ell_1^N \to V$ be the canonical mapping as defined in Proposition 2.3.5. We have seen that T is an isomorphism. It then follows that $T^{-1}(B)$ is bounded and closed in ℓ_1^N and so it is compact (cf. Remark 2.3.2). Consequently $B = T(T^{-1}(B))$ is compact as well.

Conversely, let us suppose that B is compact. Then, there exists a positive integer n and points $x_i \in B,\ 1 \le i \le n$, such that

$$B \ \subset \ \cup_{i=1}^{n} B(x_i, 1/2), \tag{2.3.7}$$

where $B(x_i, 1/2) = \{x \in V \mid \|x - x_i\| < 1/2\}$ is the open ball centered at x_i and of radius $1/2$. Set $W = \operatorname{span}\{x_1, ..., x_n\}$. We claim that $W = V$ and this will prove that the dimension of V is less than, or equal to, n and so V has to be finite dimensional. Assume the contrary. Then W will be a proper and closed (since it is finite dimensional) subspace of V. Now, by the preceding lemma of Riesz, we have the existence of $u \in V$ such that $\|u\| = 1$ (and so $u \in B$) and such that $d(u, W) \ge 2/3$. In particular, it follows that $u \in B$ is such that $\|u - x_i\| \ge 2/3$ for all $1 \le i \le n$ which contradicts (2.3.7). This completes the proof. ■

Example 2.3.12 Consider the space ℓ_2 of all square summable sequences. Consider the sequence $\mathsf{e}_i \in \ell_2$ which has its i-th component equal to unity and all other components equal to zero. Then, since $\|\mathsf{e}_i\|_2 = 1$, it belongs to the closed unit ball in that space. Now, if $i \neq j$, we have

$$\|\mathsf{e}_i - \mathsf{e}_j\|_2 \ = \ \sqrt{2}$$

and so the the sequence $\{\mathsf{e}_i\}$ can never have a convergent subsequence. Thus, in the infinite dimensional space ℓ_2, we directly see that the unit ball is not sequentially compact and hence it is not compact. ■

2.4 Applications to Differential Equations

One of the famous results in analysis is Banach's contraction mapping theorem (also known as Banach's fixed point theorem), which is stated as follows.

Theorem 2.4.1 *(Contraction Mapping Theorem) Let (X, d) be a complete metric space and let $F : X \to X$ be a* contraction, *i.e. there exists a constant $0 < c < 1$ such that, for all x and $y \in X$, we have*

$$d(F(x), F(y)) \ \le \ c\ d(x, y).$$

Then F has a unique fixed point, i.e. there exists a unique point $x^* \in X$ such that

$$F(x^*) = x^*.$$

Further, given any $x_0 \in X$, the sequence $\{x_n\}$ defined by

$$x_{n+1} = F(x_n), \ n \geq 0,$$

converges to x^*. ■

Proof: Obviously, F is continuous. Let $x_0 \in X$ and let the x_n be as defined in the statement of the theorem. Then, by hypothesis,

$$\begin{aligned} d(x_{n+1}, x_n) &= d(F(x_n), F(x_{n-1})) \\ &\leq cd(x_n, x_{n-1}) \end{aligned}$$

and, proceeding recursively, we deduce that

$$d(x_{n+1}, x_n) \leq c^n d(x_1, x_0).$$

Thus, if $n < m$, we have

$$\begin{aligned} d(x_n, x_m) &\leq d(x_n, x_{n+1}) + d(x_{n+1}, x_{n+2}) + \cdots + d(x_{m-1}, x_m) \\ &\leq (c^n + c^{n+1} + \cdots + c^{m-1}) d(x_1, x_0) \end{aligned}$$

which can be made arbitrarily small for large n and m since the geometric series $\sum_{k=1}^{\infty} c^k$ is convergent for $0 < c < 1$. Thus $\{x_n\}$ is a Cauchy sequence and, since X is complete, it converges to some $x^* \in X$. The continuity of F and the definition of the x_n now imply that $x^* = F(x^*)$.

If there were two distinct fixed points of F, say, x and y, then, we have that

$$0 < d(x, y) = d(F(x), F(y)) \leq cd(x, y)$$

which is a contradiction, since $0 < c < 1$. This completes the proof. ■

We now give a well known application of this result.

Theorem 2.4.2 *(***Picard's Theorem***) Let R be a closed rectangle in the plane $\mathbb{R}^2$ whose sides are parallel to the coordinate axes. Let $f : R \to \mathbb{R}$ be a function which is continuous and which is such that $\frac{\partial f}{\partial y}$ exists and* *is continuous on R. Let (x_0, y_0) be a point in the interior of R. Then, there exists $h > 0$ such that the initial value problem*

$$\begin{aligned} \frac{dy}{dx} &= f(x,y) \\ y(x_0) &= y_0 \end{aligned}$$

has a unique solution in the interval $(x_0 - h, x_0 + h)$.

Proof: Since R is compact and f and $\frac{\partial f}{\partial y}$ are continuous on R, there exist $K > 0$ and $M > 0$ such that

$$|f(x,y)| \leq K \text{ and } \left|\frac{\partial f}{\partial y}(x,y)\right| \leq M$$

for all $(x,y) \in R$. Then, by the mean value theorem, it follows that, for all (x, y_1) and (x, y_2) in R, we have

$$|f(x,y_1) - f(x,y_2)| \leq M|y_1 - y_2|.$$

It is easy to see that a function $y = y(x)$ is a solution of the initial value problem above if, and only if, it satisfies the following:

$$y(x) = y_0 + \int_{x_0}^{x} f(t, y(t))\, dt$$

for all x.

Now, choose $h > 0$ such that $Mh < 1$. Consider the rectangle R' defined as follows:

$$R' = \{(x,y) \mid |x - x_0| \leq h,\ |y - y_0| \leq Kh\}.$$

By choosing h small enough, we can ensure that $R' \subset R$. Now consider

$$X = \{g \in C[x_0 - h, x_0 + h] \mid |g(x) - y_0| \leq Kh \text{ for all } x\}.$$

Then, we see that X is a closed subspace of $C[x_0 - h, x_0 + h]$ and so is a complete metric space (with the distance induced by the 'sup-norm'). Let $g \in X$. Define

$$F(g)(x) = y_0 + \int_{x_0}^{x} f(t, g(t))\, dt$$

for all $x \in [x_0 - h, x_0 + h]$. Then, clearly, $F(g)$ is continuous on that interval and

$$|F(g)(x) - y_0| \leq K|x - x_0| \leq Kh.$$

Further, if g_1 and g_2 are in X, we have

$$\begin{aligned}|F(g_1)(x) - F(g_2)(x)| &\le \int_{x_0}^{x} |f(t, g_1(t)) - f(t, g_2(t))|\, dt \\ &\le M\|g_1 - g_2\||x - x_0| \\ &\le Mh\|g_1 - g_2\|\end{aligned}$$

from which it follows that $\|F(g_1) - F(g_2)\| \le Mh\|g_1 - g_2\|$ which shows that F maps the complete metric space X into itself and is a contraction. Thus F has a unique fixed point $y \in X$ which solves the initial value problem. This completes the proof. ■

We now prove a corollary of the contraction mapping theorem which will also be useful in proving the existence of solutions to higher order initial value problems.

Corollary 2.4.1 *Let (X, d) be a complete metric space and let $F : X \to X$ be a mapping such that, for some positive integer n, the map $F^n : X \to X$ is a contraction. Then F has a unique fixed point.*

Proof: Since $F^n : X \to X$ is a contraction, this mapping has a unique fixed point x^* by the preceding theorem. Now,

$$F(x^*) = F(F^n(x^*)) = F^{n+1}(x^*) = F^n(F(x^*))$$

and thus, $F(x^*)$ is also a fixed point for F^n. By the uniqueness of the fixed point, it follows that

$$F(x^*) = x^*$$

and so F has a fixed point, *viz.* x^*.

On the other hand, any fixed point y of F is also a fixed point of F^n since

$$F^n(y) = F^{n-1}(F(y)) = F^{n-1}(y) = \cdots = F(y) = y.$$

Thus F also must have a unique fixed point. ■

Example 2.4.1 Consider the Volterra integral operator $T : \mathcal{C}[0,1] \to \mathcal{C}[0,1]$ defined in Example 2.3.7, *viz.*

$$T(f)(s) = \int_0^s K(s,t) f(t)\, dt$$

where $K : [0,1] \times [0,1] \to \mathbb{R}$ is a continuous function. Consider the following problem: find $u \in \mathcal{C}[0,1]$ such that

$$u(s) \;=\; w(s) + \lambda \int_0^s K(s,t)u(t)\; dt \tag{2.4.1}$$

where $w \in \mathcal{C}[0,1]$ is a given function and $\lambda \in \mathbb{R}$. The equation (2.4.1) is called a *Volterra integral equation.* It is clear that a solution u of (2.4.1) is a fixed point of the (affine linear) mapping $F : \mathcal{C}[0,1] \to \mathcal{C}[0,1]$ defined by

$$F(u) \;=\; w + \lambda T(u).$$

Now, for any $s \in [0,1]$ and for any u_1 and $u_2 \in \mathcal{C}[0,1]$,

$$\begin{array}{rcl} |F(u_1)(s) - F(u_2)(s)| & = & |\lambda \int_0^s K(s,t)((u_1(t) - u_2(t))\; dt| \\ & \leq & |\lambda|\kappa\|u_1 - u_2\|s \end{array}$$

where

$$\kappa \;=\; \max_{[0,1]\times[0,1]} |K(s,t)|.$$

Hence, we get

$$\|F(u_1) - F(u_2)\| \;\leq\; |\lambda|\kappa\|u_1 - u_2\|.$$

Again,

$$\begin{array}{rcl} |F^2(u_1)(s) - F^2(u_2)(s)| & = & |\lambda \int_0^s K(s,t)(F(u_1)(t) - F(u_2)(t))\; dt| \\ & \leq & |\lambda|\kappa \int_0^s |F(u_1)(t) - F(u_2)(t)|\; dt \\ & \leq & |\lambda|^2\kappa^2\|u_1 - u_2\| \int_0^s t\; dt \\ & = & |\lambda|^2\kappa^2\|u_1 - u_2\|\frac{t^2}{2} \end{array}$$

whence we deduce that

$$\|F^2(u_1) - F^2(u_2)\| \;\leq\; \frac{|\lambda|^2\kappa^2}{2}\|u_1 - u_2\|.$$

Proceeding in this way, we get, for any positive integer n,

$$\|F^n(u_1) - F^n(u_2)\| \;\leq\; \frac{|\lambda|^n\kappa^n}{n!}\|u_1 - u_2\|.$$

Since $|\lambda|^n\kappa^n/n!$ is the general term of the convergent exponential series $\exp(|\lambda|\kappa)$, it tends to zero as n tends to infinity and so, for sufficiently large n, we have

$$\frac{|\lambda|^n\kappa^n}{n!} \;<1$$

and hence F^n is a contraction. Thus, by the preceding corollary, F has a unique fixed point. In other words, the Volterra integral equation (2.4.1) has a unique solution. ■

We now show that the study of certain differential equations can be reduced to the study of the Volterra integral equation.

Consider the equation

$$x''(s) + p(s)x'(s) + q(s)x(s) = f(s) \tag{2.4.2}$$

for $s \in (0,1)$, where p, q and f are given continuous functions. Consider the initial conditions

$$x(0) = \alpha \text{ and } x'(0) = \beta \tag{2.4.3}$$

where α and β are two given real numbers. Set $u(s) = x''(s)$. Then,

$$x'(s) = \beta + \int_0^s u(t)\, dt.$$

Again,

$$\begin{aligned} x(s) &= \alpha + \textstyle\int_0^s x'(t)\, dt \\ &= \alpha + \beta s + \textstyle\int_0^s \int_0^t u(\tau)\, d\tau\, dt \\ &= \alpha + \beta s + \textstyle\int_0^s \int_\tau^s u(\tau)\, dt\, d\tau \\ &= \alpha + \beta s + \textstyle\int_0^s u(\tau)(s-\tau)\, d\tau. \end{aligned}$$

Thus, (2.4.2) can be written in the form (2.4.1) with

$$w(s) = f(s) - [\beta p(s) + \alpha q(s) + \beta s q(s)],$$

$\lambda = 1$ and

$$K(s,t) = -[p(s) + (s-t)q(s)].$$

Hence, the initial value problem (2.4.2)-(2.4.3) has a unique solution.

2.5 Exercises

Note: In all the function spaces which occur below, it is assumed that vector addition and scalar multiplication are defined pointwise.

2.1 Let $f \in C[0,1]$. Define

$$\|f\|_p = \left(\int_0^1 |f(t)|^p\, dt\right)^{\frac{1}{p}}$$

where $1 \leq p < \infty$. Show that this defines a norm on $\mathcal{C}[0,1]$.

2.2 Show that the space $\mathcal{C}[0,1]$ with the norm $\|.\|_1$ defined in the previous exercise is not complete by producing a Cauchy sequence which is not convergent.

2.3 Let $f \in \mathcal{C}(\mathbb{R})$. The *support* of f is the closure of the set of points where f does not vanish. Let $\mathcal{C}_c(\mathbb{R})$ denote the space of all continuous real valued functions on $\mathbb{R}$ whose support is a compact subset of $\mathbb{R}$. Show that it is a normed linear space with the 'sup-norm' and that it is not complete.

2.4 Let $\mathcal{C}_0(\mathbb{R})$ denote the space of all continuous real valued functions on $\mathbb{R}$ which *vanish at infinity, i.e* if $f \in \mathcal{C}_0(\mathbb{R})$, then, given any $\varepsilon > 0$, there exists a compact subset $K \subset \mathbb{R}$ such that

$$|f(x)| \ < \ \varepsilon$$

for all $x \in \mathbb{R}\backslash K$. Show that $\mathcal{C}_0(\mathbb{R})$ is a Banach space with the 'sup-norm'. Show also that the space $\mathcal{C}_c(\mathbb{R})$ defined in the previous exercise is dense in $\mathcal{C}_0(\mathbb{R})$.

2.5 Let $\mathcal{C}^1[0,1]$ denote the space of all continuous real valued functions on $[0,1]$ which are continuously differentiable on $(0,1)$ and whose derivatives can be continuously extended to $[0,1]$. For $f \in \mathcal{C}^1[0,1]$,define

$$\|f\| \ = \ \max_{t\in[0,1]}\{|f(t)|, |f'(t)|\}$$

where f' denotes the derivative of f. Show that $\mathcal{C}^1[0,1]$ is a Banach space for this norm. State and prove an analogous result for $\mathcal{C}^k[0,1]$, the space of all continuous real valued functions on $[0,1]$ which are k times continuously differentiable on $(0,1)$ and all those derivatives possessing continuous extensions to $[0,1]$.

2.6 Let $f \in \mathcal{C}^1[0,1]$ and let f' denote its derivative. Define

$$\|f\|_1 \ = \ \left(\int_0^1 (|f(t)|^2 + |f'(t)|^2)dt\right)^{\frac{1}{2}}.$$

Show that $\|.\|_1$ defines a norm on $\mathcal{C}^1[0,1]$. If we define

$$|f|_1 \ = \ \left(\int_0^1 |f'(t)|^2 dt\right)^{\frac{1}{2}},$$

does $|.|_1$ define a norm on $\mathcal{C}^1[0,1]$?

2.7 Let

$$V = \{f \in \mathcal{C}^1[0,1] \mid f(0) = 0\}.$$

Show that $|.|_1$ defines a norm on V.

2.8 Let V be a Banach space with norm $\|.\|_V$. Set

$$X = \mathcal{C}([0,1]; V)$$

to be the space of all continuous functions from $[0,1]$ into the space V. Define, for $f \in X$,

$$\|f\|_X = \sup_{t\in[0,1]} \|f(t)\|_V = \left(\max_{t\in[0,1]} \|f(t)\|_V\right)$$

Show that $\|.\|_X$ is well defined and that it defines a norm on X. Show also that, under this norm, X is a Banach space.

2.9 Let V and W be normed linear spaces and let $T : V \to W$ be a linear transformation. Show that T is continuous if, and only if, T maps Cauchy sequences in V into Cauchy sequences in W.

2.10 Let $\mathcal{C}^1[0,1]$ be endowed with the norm as in Exercise 2.5 above. Let $\mathcal{C}[0,1]$ be endowed with the usual 'sup-norm'. Show that $T : \mathcal{C}^1[0,1] \to \mathcal{C}[0,1]$ defined by $T(f) = f'$ is a continuous linear transformation and that $\|T\| = 1$.

2.11 Let $\mathcal{C}[0,1]$ be endowed with its usual norm. For $f \in \mathcal{C}[0,1]$, define

$$T(f)(t) = \int_0^t f(s)\, ds, \ t \in [0,1].$$

For every positive integer n, show that

$$\|T^n\| = 1/n!$$

2.12 Consider the space $\mathcal{C}_c(\mathbb{R})$ defined in Exercise 2.3 above. For $f \in \mathcal{C}_c(\mathbb{R})$, define

$$\varphi(f) = \int_{-\infty}^{\infty} f(t)\, dt.$$

Show that φ is well defined and that it is a linear functional on this space. Is it continuous?

2.13 Let $\{t_i\}_{i=1}^n$ be given points in the closed interval $[0,1]$. Let $\{\omega_i\}_{i=1}^n$ be given real numbers. Let $f \in \mathcal{C}[0,1]$. Define

$$\varphi(f) = \sum_{i=1}^{n} \omega_i f(t_i).$$

Show that φ defines a continuous linear functional on $\mathcal{C}[0,1]$ and that

$$\|\varphi\| = \sum_{i=1}^{n} |\omega_i|.$$

2.14 Let $\mathcal{M}_n$ denote the set of all $n \times n$ matrices with complex entries. Let $\|.\|_{p,n}$ denote the matrix norm induced by the vector norm $\|.\|_p$ on $\mathbb{C}^n$, for $1 \leq p \leq \infty$. If $\mathbf{A} = (a_{ij}) \in \mathcal{M}_n$, show that

$$\|\mathbf{A}\|_{1,n} = \max_{1\leq j\leq n} \left\{ \sum_{i=1}^{n} |a_{ij}| \right\}.$$

State and prove an analogous result for $\|\mathbf{A}\|_{\infty,n}$.

2.15 With the notations introduced in the preceding exercise, show that

$$\|\mathbf{A}\|_{2,n} = \sqrt{\rho(\mathbf{A}^*\mathbf{A})}$$

where $\rho(\mathbf{T})$ denotes the spectral radius of a matrix $\mathbf{T}$. (Hint: Use Proposition 1.1.8). If $\mathbf{A}$ is a normal matrix, show that $\|\mathbf{A}\|_{2,n} = \rho(\mathbf{A})$.

2.16 With the notations introduced above, show that, for any matrix $\mathbf{A} \in \mathcal{M}_n$, we have

$$\|\mathbf{A}\|_{2,n} \leq \|\mathbf{A}\|_E \leq \sqrt{n}\|\mathbf{A}\|_{2,n}$$

where $\|.\|_E$ is the norm introduced in Example 2.3.11.

2.17 If $\|.\|$ defines a matrix norm on $\mathcal{M}_n$, show that $\rho(\mathbf{A}) \leq \|\mathbf{A}\|$ for all $\mathbf{A} \in \mathcal{M}_n$.

2.18 Let $\mathbf{A} \in \mathcal{M}_n$ be invertible and let $\|.\|$ be a matrix norm. The *condition number* of $\mathbf{A}$ is defined as

$$\text{cond}(\mathbf{A}) = \|\mathbf{A}\|.\|\mathbf{A}^{-1}\|.$$

Show that
(a) $\text{cond}(\mathbf{A}) \geq 1$ for any invertible matrix $\mathbf{A} \in \mathcal{M}_n$;
(b) $\text{cond}(\alpha\mathbf{A}) = \text{cond}(\mathbf{A})$ for any invertible matrix $\mathbf{A}$ and for any scalar $\alpha \neq 0$;
(c) for any invertible and normal matrix $\mathbf{A}$,

$$\text{cond}_{2,n}(\mathbf{A}) = \frac{\max_{1\leq i\leq n} |\lambda_i(\mathbf{A})|}{\min_{1\leq i\leq n} |\lambda_i(\mathbf{A})|}$$

where $\{\lambda_i(\mathbf{A})\}_{i=1}^{n}$ are the eigenvalues of $\mathbf{A}$ and $\text{cond}_{2,n}(\mathbf{A})$ denotes the condition number of $\mathbf{A}$ with respect to the norm $\|.\|_{2,n}$.

2.19 For what class of matrices does $\text{cond}_{2,n}$ attain its minimum value?

2.20 Let $\mathbf{A} = (a_{ij})$ be a 2×2 matrix which is invertible. Show that

$$\text{cond}_{2,2}(\mathbf{A}) = \sigma + (\sigma^2 - 1)^{\frac{1}{2}}$$

where

$$\sigma = \frac{\sum_{i,j=1}^{2} |a_{ij}|^2}{2|\det(\mathbf{A})|}.$$

2.21 Let $\mathcal{M}_n$ be endowed with the topology generated by any matrix norm. Let $GL_n(\mathbb{C})$ denote the set of all invertible matrices in $\mathcal{M}_n$. Show that $GL_n(\mathbb{C})$ is an open and dense set. Is it connected?

2.22 (a) Let $\mathcal{D}_n$ be the subset of all $n \times n$ matrices with distinct eigenvalues. Show that $\mathcal{D}_n$ is dense in $\mathcal{M}_n$ (endowed with any matrix norm).
(b) Prove the Cayley-Hamilton theorem for any diagonalizable matrix:'Every $n \times n$ matrix satisfies its characteristic equation'.
(c) Deduce the Cayley-Hamilton theorem for all $n \times n$ matrices.

2.23 Let $\mathbf{A} \in \mathcal{M}_n$ be an invertible matrix. Show that

$$\inf_{\mathbf{B} \text{ is singular}} \|\mathbf{A} - \mathbf{B}\|_{2,n} = \frac{1}{\|\mathbf{A}^{-1}\|_{2,n}}.$$

2.24 Show that the set of all orthogonal matrices in the space of all $n \times n$ real matrices (endowed with any norm topology) is compact.

2.25 Let $1 \le p < q \le \infty$. Show that $\ell_p \subset \ell_q$ and that, for all $x \in \ell_p$,

$$\|x\|_q \;\le\; \|x\|_p.$$

2.26 Consider an infinite matrix (a_{ij}), $i, j \in \mathbb{N}$ of scalars. Let $x = (x_i) \in \ell_p, 1 \le p \le \infty$. Define a sequence $A(x)$ whose i-th component is given by

$$\sum_{j=1}^{\infty} a_{ij} x_j.$$

(a) Assume that

$$\alpha \;=\; \sup_i \sum_{j=1}^{\infty} |a_{ij}| \;<\; \infty.$$

Show that $A \in \mathcal{L}(\ell_1)$ and that $\|A\| = \alpha$.
(b) Assume that

$$\alpha \;=\; \sup_j \sum_{i=1}^{\infty} |a_{ij}| \;<\; \infty.$$

Show that $A \in \mathcal{L}(\ell_\infty)$ and that $\|A\| = \alpha$.
(c) Assume that

$$\alpha \;=\; \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} |a_{ij}|^2 \;<\; \infty.$$

Show that $A \in \mathcal{L}(\ell_2)$ and that $\|A\|^2 \le \alpha$.
(d) Assume that $A \in \mathcal{L}(\ell_p)$ and that $A \in \mathcal{L}(\ell_q)$ where $1 \le p < q < \infty$. Let $\theta \in (0, 1)$. Set $r = \theta p + (1 - \theta) q$. Show that $A \in \mathcal{L}(\ell_r)$ and that

$$\|A\|_{\mathcal{L}(\ell_r)} \;\le\; \|A\|^{\theta}_{\mathcal{L}(\ell_p)} \|A\|^{1-\theta}_{\mathcal{L}(\ell_q)}.$$

2.27 Let $\{a_k\}_{k=0}^{\infty}$ be a sequence of real numbers such that $\sum_{k=0}^{\infty} |a_k| < \infty$. Consider the infinite lower triangular matrix

$$\begin{bmatrix} a_0 & 0 & 0 & \cdots & \cdots \\ a_1 & a_0 & 0 & \cdots & \cdots \\ a_2 & a_1 & a_0 & \cdots & \cdots \\ a_3 & a_2 & a_1 & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \end{bmatrix}.$$

Let A be the linear map defined on ℓ_2 by this matrix (as in the preceding exercise). Show that $A \in \mathcal{L}(\ell_2)$ and that

$$\|A\| \;\leq\; \sum_{k=0}^{\infty} |a_k|.$$

2.28 Let V be a Banach space. Let $\{A_n\}$ be a sequence of continuous linear operators on V. Let

$$S_n \;=\; \sum_{k=1}^{n} A_k.$$

If $\{S_n\}$ is a convergent sequence in $\mathcal{L}(V)$, we say that the *series*

$$\sum_{k=1}^{\infty} A_k$$

is *convergent* and the limit of the sequence $\{S_n\}$ is called the sum of the series. If $\sum_{k=1}^{\infty} \|A_k\| < \infty$, we say that the series $\sum_{k=1}^{\infty} A_k$ is *absolutely convergent.* Show that any absolutely convergent series in convergent.

2.29 Let V be a Banach space. If $A \in \mathcal{L}(V)$ is such that $\|A\| < 1$, show that the series

$$I + \sum_{k=1}^{\infty} A^k$$

is convergent and that its sum is $(I - A)^{-1}$.

2.30 (a) Let V be a Banach space and let $A \in \mathcal{L}(V)$. Show that the series

$$I + \sum_{k=1}^{\infty} \frac{A^k}{k!}$$

is convergent. The sum is denoted $\exp(A)$.
(b) If A and $B \in \mathcal{L}(V)$ are such that $AB = BA$, show that

$$\exp(A + B) \;=\; \exp(A)\exp(B).$$

(c) Deduce that $\exp(A)$ is invertible for any $A \in \mathcal{L}(V)$.
(d) Let

$$A \;=\; \begin{bmatrix} \alpha & -\omega \\ \omega & \alpha \end{bmatrix},$$

where α and ω are real numbers. Show that, for any $t \in \mathbb{R}$,

$$\exp(tA) \;=\; e^{\alpha t}\begin{bmatrix} \cos\omega t & -\sin\omega t \\ \sin\omega t & \cos\omega t \end{bmatrix}.$$

2.31 Let V be a Banach space. Show that $\mathcal{G}$, the set of invertible linear operators in $\mathcal{L}(V)$ is an open subset of $\mathcal{L}(V)$ (endowed with its usual norm topology).

2.32 (a) Define $T:\ell_2 \to \ell_2$ and $S:\ell_2 \to \ell_2$ by

$$\begin{aligned} T(x) &= (0, x_1, x_2, ...) \\ S(x) &= (x_2, x_3, ...) \end{aligned}$$

where $x = (x_1, x_2, ...) \in \ell_2$. Show that T and S define continuous linear operators on ℓ_2 and that $ST = I$ while $TS \neq I$. (Thus, T and S, which are called the *right* and *left shift operators* respectively, are not invertible).
(b) If A is a continuous linear operator on ℓ_2 such that $\|A - T\| < 1$, show that A is also not invertible. Deduce that, in general, $\mathcal{G}$, defined in Exercise 2.31 above, is not dense in $\mathcal{L}(V)$, if V is infinite dimensional. (Compare this with the finite dimensional case, cf. Exercise 2.21).

2.33 Let $\mathcal{P}$ denote the space of all polynomials in one variable with real coefficients. Let $\mathsf{p} \in \mathcal{P}$ and let $\mathsf{p} = \sum_{i=1}^{n} a_i x^i$, where $a_i \in \mathbb{R}$ for $1 \leq i \leq n$. Define

$$\|\mathsf{p}\|_1 \;=\; \sum_{i=1}^{n} |a_i| \quad \text{and} \quad \|\mathsf{p}\|_\infty \;=\; \max_{1\leq i\leq n} |a_i|.$$

Show that $\|.\|_1$ and $\|.\|_\infty$ define norms on $\mathcal{P}$ and that they are not equivalent.

2.34 Let V be a normed linear space and let W be a finite dimensional (and hence, closed) subspace of V. Let $x \in V$. Show that there exists $w \in W$ such that

$$\|x + W\| \;=\; \|x + w\|.$$

2.35 Let E_i be Banach spaces for $1 \leq i \leq 3$. Let $A \in \mathcal{L}(E_1, E_2)$ and $B \in \mathcal{L}(E_1, E_3)$. Assume, further, if K is any bounded set in E_1, then $\overline{B(K)}$ is compact in E_3. Finally, assume that $x \in E_1 \mapsto \|A(x)\|_{E_2} + \|B(x)\|_{E_3}$ defines a norm on E_1 which is equivalent to the norm $\|.\|_{E_1}$.

(a) Show that Ker(A), the kernel of A, is finite dimensional.
(b) Let $\mathcal{R}(A)$ denote the range of A. Show that the canonical mapping $\overline{A} : E_1/(\mathrm{Ker}(A)) \to \mathcal{R}(A)$ defined by $\overline{A}(x + \mathrm{Ker}(A)) = A(x)$ for $x \in E_1$, is an isomorphism.
(c) Deduce that $\mathcal{R}(A)$ is a closed subspace of E_3.

2.36 Let V and Y be normed linear spaces and let W be a dense subspace of V. Let $A \in \mathcal{L}(W, Y)$. Show that there exists a unique continuous extension $\widetilde{A} \in \mathcal{L}(V, Y)$ (*i.e.* $\widetilde{A}|_W = A$) and that $\|\widetilde{A}\|_{\mathcal{L}(V,Y)} = \|A\|_{\mathcal{L}(W,Y)}$.

2.37 (Completion of a normed linear space) Let V be a normed linear space. We say that two Cauchy sequences $\{x_n\}$ and $\{y_n\}$ in V are equivalent if $\|x_n - y_n\| \to 0$ as $n \to \infty$.
(a) Show that this defines an equvalence relation. Let the set of all equivalence classes be denoted by $\overline{V}$.
(b) Let $\overline{x}$ and $\overline{y}$ denote the equivalence classes of the Cauchy sequences $\{x_n\}$ and $\{y_n\}$ respectively. Let $\mathbf{0}$ denote the equivalence class of the sequence all of whose terms are zero. Let α be a scalar. Define $\overline{x} + \overline{y}$ to be the equivalence class of the sequence $\{x_n + y_n\}$ and $\alpha\overline{x}$ to be that of the sequence $\{\alpha x_n\}$. Show that these operations are well defined and make $\overline{V}$ a vector space.
(c) With the above notations, define $\|\overline{x}\|_{\overline{V}} = \lim_{n\to\infty} \|x_n\|_V$. Show that this is well defined and that it defines a norm on $\overline{V}$.
(d) Define $\mathsf{i} : V \to \overline{V}$ by setting $\mathsf{i}(x)$ to be the equivalence class of the sequence all of whose terms are equal to x, for any $x \in V$. Show that $\mathsf{i} \in \mathcal{L}(V, \overline{V})$ and that it is an injection. Show also that the image $\mathsf{i}(V)$ is dense in $\overline{V}$.
(e) Show that $\overline{V}$ is complete (and this space is called the *completion* of V). (Hint: Given a Cauchy sequence $\{\overline{x}^{(n)}\}$ in $\overline{V}$, choose $x_n \in V$ such that $\|\overline{x}^{(n)} - \mathsf{i}(x_n)\|_{\overline{V}} < 1/n$. Show that $\{x_n\}$ is a Cauchy sequence in V and if $\overline{x}$ denotes its equivalence class, show that $\overline{x}^{(n)} \to \overline{x}$ in $\overline{V}$.)

2.38 Let V and W be normed linear spaces and let $U \subset V$ be an open subset. Let $J : U \to W$ be a mapping. We say that J is *(Fréchet) differentiable* at $u \in U$ if there exists $T \in \mathcal{L}(V, W)$ such that

$$\lim_{h\to\mathbf{0}} \frac{\|J(u+h) - J(u) - T(h)\|}{\|h\|} \;=\; 0.$$

(Equivalently,

$$J(u+h) - J(u) - T(h) \;=\; \varepsilon(h), \;\lim_{h\to 0} \frac{\|\varepsilon(h)\|}{\|h\|} \;=\; 0.)$$

(a) If such a T exists, show that it is unique. (We say that T is the *(Fréchet) derivative* of J at $u \in U$ and write $T = J'(u)$.)
(b) If J is differentiable at $u \in U$, show that J is continuous at $u \in U$.

2.39 Let V and W be normed linear spaces and let $U \subset V$ be an open subset. Let $J : U \to W$ be a mapping. We say that J is *Gâteau differentiable* at $u \in U$ along a vector $w \in V$ if

$$\lim_{t\to 0} \frac{1}{t}(J(u+tv) - J(u))$$

exists. (We call the limit the *Gâteau derivative* of J at u along w.)

If J is Fréchet differentiable at $u \in U$, show that it is Gâteau differentiable at u along any vector $w \in V$ and that the corresponding Gâteau derivative is given by $J'(u)w$.

2.40 Let V and W be normed linear spaces and let $A \in \mathcal{L}(V, W)$. Let $w_0 \in W$ be given. define $J : V \to W$ by $J(u) = A(u) + w_0$. Show that J is differentiable at every $u \in V$ and that $J'(u) = A$.

2.41 Let $U = GL_n(\mathbb{C}) \subset \mathcal{M}_n$ (cf. Exercise 2.21). Define $J(\mathbf{A}) = \mathbf{A}^{-1}$ for $\mathbf{A} \in U$. Show that J is differentiable at every $\mathbf{A} \in U$ and that, if $\mathbf{H} \in \mathcal{M}_n$, we have

$$J'(\mathbf{A})(\mathbf{H}) \;=\; -\mathbf{A}^{-1}\mathbf{H}\mathbf{A}^{-1}.$$

2.42 (a) Let $\mathbf{A} \in GL_n(\mathbb{C})$. Show that

$$\det(\mathbf{I}+\mathbf{A}) \;=\; 1 + \mathrm{tr}(\mathbf{A}) + \varepsilon(A)$$

where

$$\lim_{\mathbf{A}\to 0} \frac{|\varepsilon(\mathbf{A})|}{\|\mathbf{A}\|} = 0$$

(for any matrix norm).
(b) Deduce that if we define $J(\mathbf{A}) = \det(\mathbf{A})$ for $\mathbf{A} \in GL_n(\mathbb{C})$, then J is differentiable at all such $\mathbf{A}$ and that, if $\mathbf{H} \in \mathcal{M}_n$, then

$$J'(\mathbf{A})(\mathbf{H}) \;=\; \det(\mathbf{A})\mathrm{tr}(\mathbf{A}^{-1}\mathbf{H}).$$

2.43 (Chain Rule) Let V, W and Z be normed linear spaces and let $f : V \to W$ and $g : W \to Z$ be mappings such that f is differentiable at a point $v \in V$ and g is differentiable at $f(v) = w \in W$. Show that the map $g \circ f : V \to Z$ is differentiable at $v \in V$ and that

$$(g \circ f)'(v) \;=\; g'(f(v)) \circ f'(v).$$

2.44 (a) Let V be a real normed linear space and let $J : V \to \mathbb{R}$ be a given mapping. A subset K of V is said to be *convex* if, for every u and $v \in K$ and for every $t \in [0, 1]$, we have that

$$tu + (1 - t)v \in K.$$

Let $K \subset V$ be a closed convex set. Assume that J attains its minimum over K at $u \in K$. If J is differentiable at u, then show that

$$J'(u)(v - u) \;\geq\; 0$$

for every $v \in K$.
(b) Let $K = V$. If J attains its minimum at $u \in V$ and if J is differentiable at u, show that $J'(u) = \mathbf{0}$.

2.45 Let V be a real normed linear space. A mapping $J : V \to \mathbb{R}$ is said to be *convex* if, for every u and $v \in V$ and for every $t \in [0, 1]$, we have

$$J(tu + (1 - t)v) \;\leq\; tJ(u) + (1 - t)J(v).$$

(a) If $J : V \to \mathbb{R}$ is convex and differentiable at every point, show that

$$J(v) - J(u) \;\geq\; J'(u)(v - u)$$

for every u and $v \in V$.
(b) Let $J : V \to \mathbb{R}$ be convex and differentiable at every point of V. Let $K \subset V$ be a closed convex set. Let $u \in K$ be such that

$$J'(u)(v - u) \;\geq\; 0$$

for every $v \in K$. Show that

$$J(u) \;=\; \min_{v \in K} J(v).$$

(c) If $J : V \to \mathbb{R}$ is convex and differentiable at every point of V, and if $u \in V$ is such that $J'(u) = \mathbf{0}$, show that J attains its minimum (over all of V) at u.

Remark 2.5.1 Exercise 2.44 gave necessary conditions for a differentiable function J to attain a minimum at a point u. The preceding exercise shows that these conditions are also sufficient in the case of convex functions. ■

2.46 Let $m > n$. Let $\mathbf{A}$ be an $m \times n$ matrix and let $\mathbf{b} \in \mathbb{R}^m$. Consider the linear system of equations

$$\mathbf{Ax} = \mathbf{b}.$$

This system may not have a solution since the number of equations exceeds the number of unknowns. A *least squares approximate* solution is a vector $\mathbf{x}_0 \in \mathbb{R}^n$ such that

$$\|\mathbf{Ax}_0 - \mathbf{b}\|_2 = \min_{\mathbf{x}\in\mathbb{R}^n} \|\mathbf{Ax} - \mathbf{b}\|_2.$$

Show that such a solution must satisfy the linear system

$$\mathbf{A}^*\mathbf{Ax}_0 = \mathbf{A}^*\mathbf{b}$$

and that this system has a unique solution if the rank of $\mathbf{A}$ is n.