9.1: Vector Spaces, linear Mappings, and Convexity

The euclidean space $\mathbb{R}^n$ has already made an appearance in the metric space chapter. In this chapter, we will extend the differential calculus we created for one variable to several variables. The key idea in differential calculus is to approximate functions by lines and linear functions. In several variables we must introduce a little bit of linear algebra before we can move on. So let us start with vector spaces and linear functions on vector spaces. While it is common to use $\vec{x}$ or the bold $\mathbf{x}$ for elements of $\mathbb{R}^n$, especially in the applied sciences, we use just plain $x$, which is common in mathematics. That is $x \in \mathbb{R}^n$ is a vector, which means that $x = (x^1,x^2,\ldots,x^n)$ is an $(n)$-tuple of real numbers. We use upper indices for identifying components, leaving us the lower index for sequences of vectors. For example, we can have vectors $x_1$ and $x_2$ in $\mathbb{R}^n$ and then $x_1 = (x_1^1,x_1^2,\ldots,x_1^n)$ and $x_2 = (x_2^1,x_2^2,\ldots,x_2^n)$. It is common to write vectors as column vectors, that is, $(n \times 1)$ matrices:

$$x = (x^1,x^2,\ldots,x^n) = \begin{bmatrix} x^1 \\ x^2 \\ \vdots \\ x^n \end{bmatrix}.$$  

We will do so when convenient. We call real numbers scalars to distinguish them from vectors. Let $X$ be a set together with operations of addition, $+ : X \times X \to X$, and multiplication, $\cdot : \mathbb{R} \times X \to X$ (we write $ax$ instead of $a \cdot x$). $X$ is called a vector space (or a real vector space) if the following conditions are satisfied: (Addition is associative) If $u, v, w \in X$, then $(u+(v+w) = (u+v)+w)$. (Addition is commutative) If $u, v \in X$, then $u+v = v+u$. (Additive identity) There is a $0 \in X$ such that $v+0=v$ for all $v \in X$. (Additive inverse) For every $v \in X$, there is a $-v \in X$, such that $v+(-v)=0$. (Distributive law) If $(a, v \in X)$, then $a(u+v) = au+av$. (Distributive law) If $(a, b \in \mathbb{R}, v \in X)$, then $(a+b)v = av+bv$. (Multiplication is associative) If $(a, b \in \mathbb{R})$, then $(ab)v = a(bv)$. (Multiplicative identity) $(1v = v)$ for all $v \in X$. Elements of a vector space are usually called vectors, even if they are not elements of $\mathbb{R}^n$ (vectors in the “traditional” sense). An example vector space is $\mathbb{R}^n$, where addition and multiplication by a constant is done...
componentwise: if \(\alpha \in \mathbb{R}\) and \(x,y \in \mathbb{R}^n\), then \[\begin{aligned}
& x+y := (x^1,x^2,\ldots,x^n) + (y^1,y^2,\ldots,y^n) = (x^1+y^1,x^2+y^2,\ldots,x^n+y^n) ,
& \alpha x := \alpha (x^1,x^2,\ldots,x^n) = (\alpha x^1, \alpha x^2,\ldots, \alpha x^n) .
\end{aligned}\]
In this book we mostly deal with vector spaces that can be regarded as subsets of \(\mathbb{R}^n\), but there are other vector spaces that are useful in analysis. For example, the space \(C([0,1],\mathbb{R}^n)\) of continuous functions on the interval \([0,1]\) is a vector space. A trivial example of a vector space (the smallest one in fact) is just \(\mathbb{R}^0\). The operations are defined in the obvious way. You always need a zero vector to exist, so all vector spaces are nonempty sets. It is also possible to use other fields than \(\mathbb{R}\) in the definition (for example it is common to use the complex numbers \(\mathbb{C}\)), but let us stick with the real numbers. A function \(f : \mathcal{V} \to \mathcal{Y}\), when \(\mathcal{Y}\) is not \(\mathbb{R}\), is often called a mapping or a map rather than a function.

Linear combinations and dimension If we have vectors \(x_1, \ldots, x_k \in \mathbb{R}^n\) and scalars \(a^1, \ldots, a^k \in \mathbb{R}\), then \[a^1 x_1 + a^2 x_2 + \cdots + a^k x_k\] is called a linear combination of the vectors \(x_1, \ldots, x_k\). If \(\mathcal{Y} \subset \mathbb{R}^n\) is a set then the span of \(\mathcal{Y}\), or in notation \(\text{span}(\mathcal{Y})\), is the set of all linear combinations of some finite number of elements of \(\mathcal{Y}\). We also say \(\mathcal{Y}\) spans \(\text{span}(\mathcal{Y})\).

Let \(\mathcal{Y} := \{ (1,1) \} \subset \mathbb{R}^2\). Then \[\text{span}(\mathcal{Y}) = \{ (x,x) \in \mathbb{R}^2 : x \in \mathbb{R} \} .\]
That is, \(\text{span}(\mathcal{Y})\) is the line through the origin and the point \((1,1)\).

Let \(\mathcal{Y} := \{ (1,1), (0,1) \} \subset \mathbb{R}^2\). Then \[\text{span}(\mathcal{Y}) = \mathbb{R}^2 ,\]
as any point \((x,y) \in \mathbb{R}^2\) can be written as a linear combination \((x,y) = x (1,1) + (y-x) (0,1) .\] A sum of two linear combinations is again a linear combination, and a scalar multiple of a linear combination is a linear combination, which proves the following proposition. Let \(\mathcal{X}\) be a vector space. For any \(\mathcal{Y} \subset \mathcal{X}\), the set \(\text{span}(\mathcal{Y})\) is a vector space itself. If \(\mathcal{Y}\) is already a vector space then \(\text{span}(\mathcal{Y}) = \mathcal{Y}\). A set of vectors \(\{ x_1, x_2, \ldots, x_k \}\) is linearly independent, if the only solution to \[a^1 x_1 + a^2 x_2 + \cdots + a^k x_k = 0\] is the trivial solution \((a^1 = a^2 = \cdots = a^k = 0)\). A set that is not linearly independent, is linearly dependent. A linearly independent set \(B\) of vectors such that \(\text{span}(B) = \mathcal{X}\) is called a basis of \(\mathcal{X}\). For example the set \(\{ (1,1) \}\) of the two vectors in is a basis of \(\mathbb{R}^2\). If a vector space \(\mathcal{X}\) contains a linearly independent set of \(d\) vectors, but no linearly independent set of \(d+1\) vectors then we say the dimension of \(\mathcal{X}\) is \(d\). If for all \(d \in \mathbb{N}\) the vector space \(\mathcal{X}\) contains a set of \(d\) linearly independent vectors, we say \(\mathcal{X}\) is infinite dimensional and write \(\dim \, \mathcal{X} := \infty\). Clearly for the trivial vector space, \(\dim \, \{ 0 \} = 0\). We will see in a moment that any vector space that is a subset of \(\mathbb{R}^n\) has a finite dimension, and that dimension is less than or equal to \(n\). If a set is linearly dependent, then one of the vectors is a linear combination of the others. In other words, if \(\forall \ a^j \ \forall \not= 0\), then we can solve for \(x_j\)
\[
[x_j = \frac{a^1 x_1 + a^2 x_2 + \cdots + a^k x_k}{a^j} \}
\]
Clearly then the vector \( x_j \) has at least two different representations as linear combinations of \( \{ x_1, x_2, \ldots, x_k \} \). If \( B = \{ x_1, x_2, \ldots, x_k \} \) is a basis of a vector space \( X \), then every point \( y \in X \) has a unique representation of the form \( y = \sum_{j=1}^k \alpha^j \alpha^j x_j \) for some numbers \( \alpha^1, \alpha^2, \ldots, \alpha^k \). Every \( y \in X \) is a linear combination of elements of \( B \) since \( X \) is the span of \( B \). For uniqueness suppose \( y = \sum_{j=1}^k \alpha^j \alpha^j x_j \neq 0 \) then \( \sum_{j=1}^k \alpha^j \alpha^j x_j = 0 \). By linear independence of the basis \( \{ \alpha^1, \alpha^2, \ldots, \alpha^k \} \) for all \( j \). For \( \{ \text{mathbb\{R\}} \}^n \) we define \( e_1 := (1,0,0,\ldots,0) \), \( e_2 := (0,1,0,\ldots,0) \), \( \ldots \), \( e_n := (0,0,0,\ldots,1) \), and call this the standard basis of \( \{ \text{mathbb\{R\}} \}^n \). We use the same letters \( e_j \) for any \( \{ \text{mathbb\{R\}} \}^n \), and which space \( \{ \text{mathbb\{R\}} \}^n \) we are working in is understood from context.

A direct computation shows that \( \{ \{ e_1, e_2, \ldots, x_d \} \} \) is really a basis of \( \{ \text{mathbb\{R\}} \}^n \); it is easy to show that it spans \( \{ \text{mathbb\{R\}} \}^n \) and is linearly independent. In fact,

\[
x = (x^1, x^2, \ldots, x^n) = \sum_{j=1}^n x^j e_j
\]

Let \( (X) \) be a vector space. If \( (X) \) is spanned by \( \{ \text{dim} \} \) vectors, then \( (\{ \text{dim} \} \leq \text{dim} \} \) if and only if \( (X) \) has a basis of \( \{ \text{dim} \} \) vectors (and so every basis has \( \{ \text{dim} \} \) vectors). In particular, \( (\{ \text{dim} \} \leq \{ \text{mathbb\{R\}} \}^n) \) if \( (Y \subset X) \) is a vector space and \( (\{ \text{dim} \} \leq \text{dim} \} \). If \( (\{ \text{dim} \} \leq \text{dim} \} \) and \( \{ \text{dim} \} \) spans \( (X) \), then \( \{ \text{dim} \} \) is linearly independent. If \( \{ \text{dim} \} \leq \text{dim} \} \) and \( \{ \text{dim} \} \) is linearly independent, then there is a set \( \{ \text{dim} \} \) such that \( (T \cup \{ S \}) \) is a basis of \( (X) \). Let us start with (i). Suppose \( \{ S = \{ x_1, x_2, \ldots, x_d \} \} \) spans \( (X) \), and \( (T = \{ y_1, y_2, \ldots, y_m \} \) is a set of linearly independent vectors of \( (X) \). We wish to show that \( \{ \text{dim} \} \) spans \( (X) \).

Write \( y_1 = \sum_{k=1}^d \alpha^1 x_k \), which we can do as \( (S) \) spans \( (X) \). One of the \( \{ \alpha^1 x_k \} \) is nonzero (otherwise \( y_1 = 0 \) would be zero), so suppose without loss of generality that this is \( \{ \alpha^1 x_1 \} \). Then we can solve \( x_1 = \frac{1}{\alpha^1} y_1 - \sum_{k=2}^d \frac{\alpha^1 x_k}{\alpha^1} \). In particular \( \{ y_1, x_2, \ldots, x_d \} \) spans \( (X) \), since \( x_1 \) can be obtained from \( \{ y_1, x_2, \ldots, x_d \} \). Next, \( y_2 = \alpha^2 y_1 = \sum_{k=2}^d \alpha^2 x_k \). As \( (T) \) is linearly independent, we must have that one of the \( \{ \alpha^2 x_k \} \) for \( k \geq 2 \) must be nonzero. Without loss of generality, we can assume \( \alpha^2 x_2 \) to be the one. Proceed to solve for \( x_2 = \frac{1}{\alpha^2} y_2 - \sum_{k=3}^d \frac{\alpha^2 x_k}{\alpha^2} \). In particular \( \{ y_1, y_2, x_3, x_4, \ldots, x_d \} \) spans \( (X) \). The astute reader will think back to linear algebra and notice that we are row-reducing a matrix. We continue this procedure. If \( m < d \), then we are done. So suppose \( m \geq d \). After \( d \) steps we obtain that \( \{ y_1, y_2, \ldots, y_m \} \) spans \( (X) \). Any other vector \( \{ v \} \) in \( (X) \) is a linear combination of \( \{ y_1, y_2, \ldots, y_m \} \), and hence cannot be in \( (T) \) as \( (T) \) is linearly independent. So \( \{ y_m \} \). Let us look at (ii). First notice that if we have a set \( \{ T \} \) of \( k \) linearly independent vectors that do not span \( (X) \), then we can always choose a vector \( \{ v \} \in X \) so that \( \{ v \} \) is linearly independent (exercise). If \( \{ \text{dim} \} \leq \text{dim} \} \), then there must exist some linearly independent set of \( \{ d \} \) vectors \( \{ T \} \), and it must span \( (X) \), otherwise we could choose a larger set of linearly independent vectors. So we have a basis of \( \{ d \} \) vectors. On the other hand if we have a basis of \( \{ d \} \) vectors, it is linearly independent and spans \( (X) \). By (i) we know there is no set of \( \{ d+1 \} \) linearly independent vectors, so dimension must be \( \{ d \} \). For (iii) notice that \( \{ e_1, e_2, \ldots, e_n \} \) is a basis of \( \{ \text{mathbb\{R\}} \}^n \). To see (iv), suppose \( (Y) \) is a vector space and \( (Y \subset X) \), where \( \{ \text{dim} \} \leq \text{dim} \} \). As \( (X) \) cannot contain \( \{ d+1 \} \) linearly independent vectors, neither can \( (Y) \). For (v) suppose \( (T) \) is a set of \( \{ m \} \) vectors that is linearly dependent and spans \( (X) \). Then one of the vectors is a linear combination of the others. Therefore we remove it from \( (T) \) we obtain a set of \( \{ m-1 \} \) vectors that still span \( (X) \) and hence \( \{ \text{dim} \} \leq \text{dim} \} \). For (vi) suppose \( (T = \{ x_1, \ldots, x_m \} \) is a linearly independent set. We follow the procedure above in the proof of (ii) to keep adding vectors while keeping the set linearly independent. As the dimension is
we can add a vector exactly \(d-m\) times. Linear mappings \(A\) mapping \(\langle A \colon X \to Y \rangle\) of vector spaces \(\langle X \rangle\) and \(\langle Y \rangle\) is linear (or a linear transformation) if for every \((a \in \mathbb{R}, x, y \in X)\) we have \(\langle A(a x) = a A(x) \rangle\) \(\langle A(x+y) = A(x)+A(y) \rangle\). We usually write \(\langle A(x) \rangle\) instead of \(\langle A(x) \rangle\) if \(\langle A \rangle\) is linear. If \(\langle A \rangle\) is one-to-one an onto then we say \(\langle A \rangle\) is invertible and we denote the inverse by \(\langle A^{-1} \rangle\). If \(\langle A \colon X \to X \rangle\) is linear then we say \(\langle A \rangle\) is a linear operator on \(\langle X \rangle\). We write \(\langle L(X,Y) \rangle\) for the set of all linear transformations from \(\langle X \rangle\) to \(\langle Y \rangle\), and just \(\langle L(X) \rangle\) for the set of linear operators on \(\langle X \rangle\). If \(\langle (a,b) \in \mathbb{R} \times \mathbb{R} \rangle\) and \(\langle (A, B) \in L(X,Y) \rangle\), define the transformation \(\langle (aA+bB) \rangle\) \(\langle (aA+bB)(x) = aAx+bBx \rangle\). If \(\langle (A \in L(Y,Z)) \rangle\) and \(\langle (B \in L(X,Y)) \rangle\), define the transformation \(\langle (AB) \rangle\) as \(\langle ABx := A(Bx) \rangle\). Finally denote by \(\langle (I \in L(X)) \rangle\) the identity: the linear operator such that \(\langle I(x) = x \rangle\) for all \(\langle x \rangle\). It is not hard to see that \(\langle (aA+bB \in L(X,Y)) \rangle\), and that \(\langle (AB \in L(X,Z)) \rangle\). In particular, \(\langle (L(X,Y)) \rangle\) is a vector space. It is obvious that if \(\langle (A) \rangle\) is linear then \(\langle (A0 = 0) \rangle\). If \(\langle (A \colon X \to Y) \rangle\) is invertible, then \(\langle (A^{-1} \rangle\) is linear. Let \(\langle (a \in \mathbb{R}, y \in Y) \rangle\). As \(\langle (A) \rangle\) is onto, then there is an \(\langle x \rangle\) such that \(\langle y = Ax \rangle\), and further as it is also one-to-one \(\langle (A^{-1} \langle Az = z \rangle) \rangle\) for all \(\langle z \in X \rangle\). So

\[
\langle A^{-1}(ay) = A^{-1}(aAx) = A^{-1}
\bigl(A(ax)\bigr) = ax = aA^{-1}(y)\rangle.
\]

Similarly let \(\langle (y_1, y_2 \in Y) \rangle\), and \(\langle (x_1, x_2 \in X) \rangle\) such that \(\langle Ax_1 = y_1 \rangle\) and \(\langle Ax_2 = y_2 \rangle\), then \(\langle (A^{-1}(y_1+y_2) = A^{-1}(x_1+x_2) = A^{-1}\langle y_1 \rangle + A^{-1}\langle y_2 \rangle \rangle\). 

If \(\langle (A \col X \to Y) \rangle\) is linear then it is completely determined by its values on a basis of \(\langle X \rangle\). Furthermore, if \(\langle (B) \rangle\) is a basis, then any function \(\langle (\tilde{A} \colon \col B \to Y) \rangle\) extends to a linear function on \(\langle X \rangle\). For infinite dimensional spaces, the proof is essentially the same, but a little trickier to write, so let us stick with finitely many dimensions. We leave the infinite dimensional case to the reader. Let \(\langle \{ x_1, x_2, \ldots, x_n \} \rangle\) be a basis and suppose \(\langle (Ax_j = y_j) \rangle\). Then every \(\langle (x \in X) \rangle\) has a unique representation \(\langle x = \sum_{j=1}^{\langle n \rangle} b^j x_j \rangle\) for some \(\langle (b^1, b^2, \ldots, b^n) \rangle\). Then by linearity \(\langle Ax = A\sum_{j=1}^{\langle n \rangle} b^j x_j = \sum_{j=1}^{\langle n \rangle} b^j Ax_j = \sum_{j=1}^{\langle n \rangle} b^j y_j \rangle\). The “furthermore” follows by defining the extension \(\langle Ax = \sum_{j=1}^{\langle n \rangle} b^j y_j \rangle\), and noting that this is well defined by uniqueness of the representation of \(\langle x \rangle\). If \(\langle (X) \rangle\) is a finite dimensional vector space and \(\langle (A \colon \col X \to X) \rangle\) is linear, then \(\langle (A) \rangle\) is one-to-one if and only if it is onto. Let \(\langle \{ x_1, x_2, \ldots, x_n \} \rangle\) be a basis for \(\langle X \rangle\). Suppose \(\langle (A) \rangle\) is one-to-one. Now suppose

\[
\langle \sum_{j=1}^{\langle n \rangle} c^j Ax_j = A\sum_{j=1}^{\langle n \rangle} c^j x_j = 0 \rangle.
\]

As \(\langle (A) \rangle\) is one-to-one, the only vector that is taken to 0 is 0 itself. Hence, \(\langle 0 = \sum_{j=1}^{\langle n \rangle} c^j x_j \rangle\) and \(\langle (c^j = 0) \rangle\) for all \(\langle j \rangle\). Therefore, \(\langle \{ Ax_1, Ax_2, \ldots, Ax_n \} \rangle\) is linearly independent. By an above proposition and the fact that the dimension is \(\langle \langle n \rangle \rangle\), we have that \(\langle \{ Ax_1, Ax_2, \ldots, Ax_n \} \rangle\) span \(\langle X \rangle\). As any point \(\langle x \in X \rangle\) can be written as \(\langle x = \sum_{j=1}^{\langle n \rangle} a^j x_j \rangle\) \(\langle Ax_j = A\sum_{j=1}^{\langle n \rangle} a^j x_j \rangle\) so \(\langle (A) \rangle\) is onto. Now suppose \(\langle (A) \rangle\) is onto. As \(\langle (A) \rangle\) is determined by the action on the basis we see that every element of \(\langle X \rangle\) has to be in the span of \(\langle \{ Ax_1, Ax_2, \ldots, Ax_n \} \rangle\). Suppose \(\langle (Ax = Ay) \rangle\), then \(\langle (A(x-y) = 0) \rangle\) and so \(\langle (x=y) \rangle\). Convexity A subset \(\langle U \rangle\) of a vector space is convex if whenever \(\langle (x, y \in U) \rangle\), the line segment from \(\langle x \rangle\) to \(\langle y \rangle\) lies in \(\langle U \rangle\). That is, if the convex combination \(\langle (t=1tx+ty) \rangle\) is in \(\langle U \rangle\) for all \(\langle t \in [0,1] \rangle\). See . Note that in \(\langle (\mathbb{R}) \rangle\), every connected interval is convex. In \(\langle (\mathbb{R}^2) \rangle\) (or higher dimensions) there are lots of nonconvex connected sets. For example the set \(\langle (\mathbb{R}^2 \setminus \{0\}) \rangle\) is not convex but it is connected. To see this simply take any \(\langle (x \in \mathbb{R}^2 \setminus \{0\}) \rangle\) and let \(\langle (y=-x) \rangle\). Then \(\langle (\text{nicefrac}{1}{2})x + \langle (\text{nicefrac}{1}{2})y \rangle \rangle = \langle (0,0) \rangle\), which is not in the set. On the other hand, the ball \(\langle (B(x,r) \setminus \{0\}) \rangle\) (using the
standard metric on \( (\mathbb{R}^n) \) is always convex by the triangle inequality. Show that in \( (\mathbb{R}^n) \) any ball \( B(x,r) \) for \( x \in \mathbb{R}^n \) and \( r > 0 \) is convex. Any subspace \( V \) of a vector space \( X \) is convex. A somewhat more complicated example is given by the following. Let \( C([0,1], \{ \mathbb{R} \}) \) be the vector space of continuous real valued functions on \( (\mathbb{R}) \). Let \( X \subset C([0,1], \{ \mathbb{R} \}) \) be the set of those \( f \) such

\[
\int_0^1 f(x) \, dx \leq 1 \quad \text{and} \quad f(x) \geq 0 \quad \text{for all } x \in [0,1].
\]

Then \( X \) is convex. Take \( t \in [0,1] \) and note that if \( f,g \in X \) then \( t f(x) + (1-t) g(x) \geq 0 \) for all \( x \). Furthermore

\[
\int_0^1 (tf(x) + (1-t)g(x)) \, dx = t \int_0^1 f(x) - dx + (1-t) \int_0^1 g(x) \, dx \leq 1.
\]

Note that \( X \) is not a subspace of \( (C([0,1], \{ \mathbb{R} \})) \). The intersection two closed sets is convex. In fact, If \( \{ \lambda \}_{\lambda \in \mathbb{R}} \) is an arbitrary collection of convex sets, then

\[
\bigcap_{\lambda \in \mathbb{R}} C_{\lambda} = \{ x \mid \lambda \in \mathbb{R} \}
\]

is convex. The proof is easy. If \( x,y \in C_{\lambda} \), then \( x,y \in C_{\lambda} \) for all \( \lambda \in \mathbb{R} \), and hence if \( \lambda \in [0,1] \), then \( (t x + (1-t)y) \in C_{\lambda} \) for all \( \lambda \in \mathbb{R} \). Therefore \( (t x + (1-t)y) \in C \) and \( C \) is convex. Let \( T : V \to W \) be a linear mapping between two vector spaces and let \( C \subset V \) be a convex set. Then \( T(C) \) is convex. Take any two points \( (p,q) \in T(C) \). Then pick \( (x,y) \in C \) such that \( T(x) = p \) and \( T(y) = q \). As \( C \) is convex then for all \( t \in [0,1] \) we have \( tx + (1-t)y \in C \), so \( T(tx + (1-t)y) = T(x) + (1-t)T(y) = T(p + (1-t)q) \) is in \( T(C) \). For completeness, let us A very useful construction is the convex hull. Given any set \( S \subset V \) of a vector space, define the convex hull of

\[
\text{co}(S) := \bigcap \{ C \subset V : S \subset C, \text{ and $C$ is convex} \}
\]

is convex. The proof is easy. If \( x,y \in C \), then \( x,y \in C \) for all \( \lambda \in \mathbb{R} \), and hence if \( \lambda \in [0,1] \), then \( (t x + (1-t)y) \in C \) for all \( \lambda \in \mathbb{R} \). Therefore \( (t x + (1-t)y) \in C \) and \( C \) is convex. Let \( T : V \to W \) be a linear mapping between two vector spaces and let \( C \subset V \) be a convex set. Then \( T(C) \) is convex. Take any two points \( (p,q) \in T(C) \). Then pick \( (x,y) \in C \) such that \( T(x) = p \) and \( T(y) = q \). As \( C \) is convex then for all \( t \in [0,1] \) we have \( tx + (1-t)y \in C \), so \( T(tx + (1-t)y) = T(x) + (1-t)T(y) = T(p + (1-t)q) \) is in \( T(C) \). For completeness, let us A very useful construction is the convex hull. Given any set \( S \subset V \) of a vector space, define the convex hull of

\[
\text{co}(S) := \bigcap \{ C \subset V : S \subset C, \text{ and $C$ is convex} \}
\]

is convex. That is, the convex hull is the smallest convex set containing \( S \). Note that by a proposition above, the intersection of convex sets is convex and hence, the convex hull is convex. The convex hull of 0 and 1 in \( (\mathbb{R}^n) \) is \( (0,1) \). Proof: Any convex set containing 0 and 1 must contain \( (0,1) \). The set \( (0,1) \) is convex, therefore it must be the convex hull. Exercises Verify that \( (\mathbb{R}^n) \) is a vector space. Let \( X \) be a vector space. Prove that a finite set of vectors \( \{ x_1, \ldots, x_n \} \subset X \) is linearly independent if and only if for every \( j=1,2,\ldots,n \) \( \text{operatormname}[\text{span}](\{ x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n \}) \subset \text{operatormname}[\text{span}](\{ x_1, \ldots, x_n \}) \). That is, the span of the set with one vector removed is strictly smaller. Prove that \( C([0,1], \{ \mathbb{R} \}) \) is an infinite dimensional vector space where the operations are defined in the obvious way: \( \langle s=f,g \rangle \) and \( \langle m=fg \rangle \) are defined as \( \langle s(x) := f(x)+g(x) \rangle \) and \( \langle m(x) := f(x)g(x) \rangle \). Hint: for the dimension, think of functions that are only nonzero on the interval \( \langle \{ \text{nicefrac} \{ \{ \text{nicefrac} \{ \{ n \} \} \} \} \} \rangle \). Let \( k : [0,1] \to [0,1] \) be continuous. Show that \( L \subset C([0,1], \{ \mathbb{R} \}) \) defined by \( L(y) := \int_0^x k(x,y)f(x) \, dx \) is a linear operator. That is, show that \( L \) is well defined (that \( L(f) \) is continuous), and that \( L \) is linear.