Chapter 2

Vector spaces

In this chapter we introduce the notion of a vector space which is fundamental for the approximation methods that we will later develop, in particular through the orthogonal projection onto a subspace representing the best possible approximation in that subspace. We use the Euclidian space as an illustrative example but the concept of a vector space is much more general than that, forming the basis for the theory of function approximation and partial differential equations.

2.1 Vector spaces

Vector space

We denote the elements of $\mathbb{R}$, the real numbers, as *scalars*, and a *vector space*, or *linear space*, is then defined by a set V together with two basic operations on V: *vector addition* and *scalar multiplication*,

(i) $x, y \in V \Rightarrow x + y \in V$,

(ii) $x \in V, \alpha \in \mathbb{R} \Rightarrow \alpha x \in V$.

A vector space defined over $\mathbb{R}$ is a *real vector space*. More generally we may define vector spaces over the complex numbers $\mathbb{C}$, or any *algebraic field* $\mathbb{F}$.

The Euclidian space $\mathbb{R}^n$

The Euclidian space $\mathbb{R}^n$ is a vector space consisting of the set of column vectors $x = (x_1, ..., x_n)^T$, where $(x_1, ..., x_n)$ is a row vector with $x_j \in \mathbb{R}$, and where v^T denotes the transpose of the vector v. In $\mathbb{R}^n$ the basic operations are defined by component-wise addition and multiplication, such that,

(i) $x + y = (x_1 + y_1, ..., x_n + y_n)^T$,

(ii) $\alpha x = (\alpha x_1, ..., \alpha x_n)^T$.

A geometrical interpretation of a vector space will prove to be useful. For example, the vector space $\mathbb{R}^2$ can be interpreted as the vector arrows in the Euclidian plane, defined by: (i) a direction with respect to a fixed point (origo), and (ii) a magnitude (the Euclidian length).

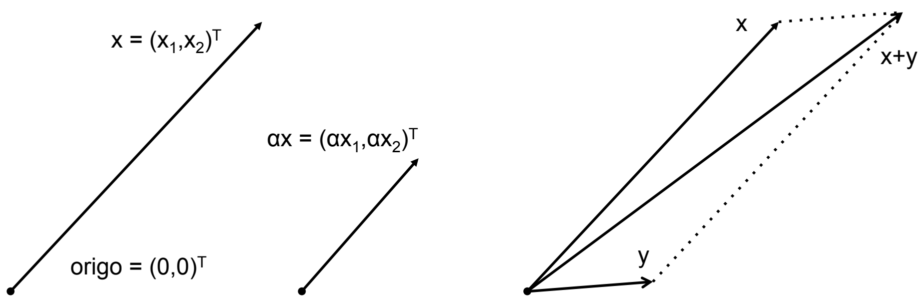

Figure 2.1: Geometric interpretation of a vector $x = (x_1, x_2)^T$ in the Euclidian plane $\mathbb{R}^2$ (left), scalar multiplication αx with $\alpha = 0.5$ (center), and vector addition $x + y$ (right).

Vector subspace

A *subspace* of a vector space V is a subset $S \subset V$, such that S is a also vector space. For example, the planes $S_1 = \{x \in \mathbb{R}^3 : x_3 = 0\}$ and $S_2 = \{x \in \mathbb{R}^3 : ax_1 + bx_2 + cx_3 + d = 0 : a, b, c, d \in \mathbb{R}\}$ are subspaces of $\mathbb{R}^3$, see Figure 2.2.

Basis

The sum $\sum_{i=1}^n \alpha_i v_i$ is referred to as a *linear combination* of the set of vectors $\{v_i\}_{i=1}^n$ in V. All possible linear combinations defines a subspace $S = \{v \in V : v = \sum_{i=1}^n \alpha_i v_i, \alpha_i \in \mathbb{R}\}$, and we say that the vector space S is *spanned* by the set of vectors $\{v_i\}_{i=1}^n$, denoted by $S = \text{span}\{v_i\}_{i=1}^n = \langle v_1, ..., v_n \rangle$.

The set $\{v_i\}_{i=1}^n$ is *linearly independent*, if

$$\sum_{i=1}^n \alpha_i v_i = 0 \quad \Rightarrow \quad \alpha_i = 0, \quad i = 1, ..., n. \tag{2.1}$$

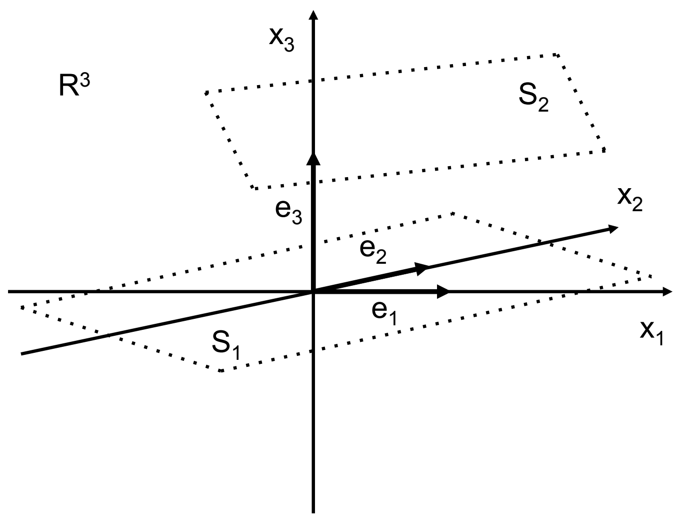

Figure 2.2: Illustration of the Euclidian space $\mathbb{R}^3$ with the three coordinate axes in the direction of the standard basis vectors e_1, e_2, e_3, and two subspaces S_1 and S_2, where S_1 is the x_1x_2-plane and S_2 a generic plane in $\mathbb{R}^3$, with the indicated planes extending to infinity.

A linearly independent set $\{v_i\}_{i=1}^n$ is a *basis* for the vector space V, if all $v \in V$ can be expressed as a linear combination of the vectors in the basis,

$$v = \sum_{i=1}^{n} \alpha_i v_i, \tag{2.2}$$

where α_i are the *coordinates* of v with respect to the basis $\{v_i\}_{i=1}^n$. The *dimension* of V, $\dim(V)$, is the number of vectors in any basis for V.

The *standard basis* $\{e_1, ..., e_n\} = \{(1, 0, ..., 0)^T, ..., (0, ..., 0, 1)^T\}$ spans $\mathbb{R}^n$, such that all $x \in \mathbb{R}^n$ can be expressed as $x = \sum_{i=1}^n x_i e_i$. We refer to the coordinates $x_i \in \mathbb{R}$ in the standard basis as *Cartesian coordinates*, and $\dim \mathbb{R}^n = n$

Norm

To measure the size of vectors we introduce the *norm* $\|\cdot\|$ of a vector in the vector space V. A norm must satisfy the following conditions:

(i) $\|x\| \geq 0$, $\forall x \in V$, and $\|x\| = 0 \Leftrightarrow x = 0$,

(ii) $\|\alpha x\| = |\alpha|\|x\|$, $\forall x \in V, \alpha \in \mathbb{R}$,

(iii) $\|x + y\| \leq \|x\| + \|y\|$, $\forall x, y \in V$,

where (iii) is the *triangle inequality*.

A *normed vector space* is a vector space on which a norm is defined. For example, $\mathbb{R}^n$ is a normed vector space on which the l_2*-norm* is defined,

$$\|x\|_2 = (\sum_{i=1}^{n} x_i^2)^{1/2} = (x_1^2 + ... + x_n^2)^{1/2}, \tag{2.3}$$

which corresponds to the *Euclidian length* of the vector x.

Inner product

A function $(\cdot, \cdot) : V \times V \to \mathbb{R}$ on the vector space V is an *inner product* if

(i) $(\alpha x + \beta y, z) = \alpha(x, z) + \beta(y, z)$,

(ii) $(x, \alpha y + \beta z) = \alpha(x, y) + \beta(x, z)$,

(iii) $(x, y) = (y, x)$,

(iv) $(x, x) \geq 0$, $\forall x \in V$, and $(x, x) = 0 \Leftrightarrow x = 0$,

for all $x, y, z \in V$ and $\alpha, \beta \in \mathbb{R}$.

An *inner product space* is a vector space on which an inner product is defined. An inner product induces an associated norm by $\|x\| = (x, x)^{1/2}$, and thus an inner product space is also a normed space. An inner product and its associated norm satisfies the *Cauchy-Schwarz inequality*.

Theorem 1 (Cauchy-Schwarz inequality)**.**

$$|(x, y)| \leq \|x\|\|y\|, \quad \forall x, y \in V \tag{2.4}$$

Proof. Let $s \in \mathbb{R}$ so that

$$0 \leq \|x + sy\|^2 = (x + sy, x + sy) = \|x\|^2 + 2s(x, y) + s^2\|y\|^2,$$

and then choose s as the minimizer of the right hand side of the inequality, that is, $s = -(x, y)/\|y\|^2$, which proves the theorem. $\square$

The Euclidian space $\mathbb{R}^n$ is an inner product space with the *Euclidian inner product*, also referred to as scalar product or dot product, defined by

$$(x, y)_2 = x \cdot y = (x_1 y_1 + ... + x_n y_n), \tag{2.5}$$

which induces the l_2-norm $\|x\|_2 = (x, x)_2^{1/2}$. In $\mathbb{R}^n$ we often drop the subscript for the Euclidian inner product and norm, with the understanding that $(x, y) = (x, y)_2$ and $\|x\| = \|x\|_2$.

We can also define general l_p-norms as

$$\|x\|_p = \left(\sum_{i=1}^{n} |x_i|^p \right)^{1/2}, \tag{2.6}$$

for $1 \leq p < \infty$. For example, the l_1-norm is defined as $\|x\|_1 = |x_1| + ... + |x_n|$. For $p = \infty$, we define the l_∞-norm as

$$\|x\|_\infty = \max_{1 \leq p \leq n} |x_i|. \tag{2.7}$$

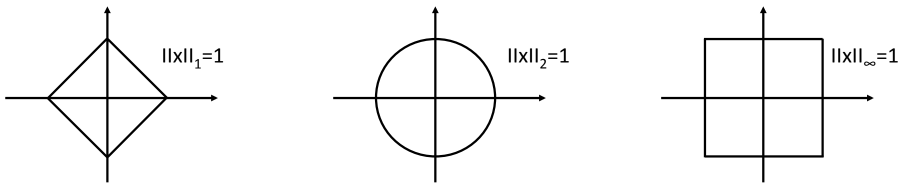

Figure 2.3: Illustration of l_p-norms in $\mathbb{R}^n$ through the unit circles $\|x\|_p = 1$, for $p = 1, 2, \infty$ (from left to right).

In fact, the Cauchy-Schwarz inequality is a special case of the *Hölder inequality* for general l_p-norms in $\mathbb{R}^n$.

Theorem 2 (Hölder inequality)**.** *For $1/p + 1/q = 1$, we have that*

$$|(x, y)| \leq \|x\|_p \|y\|_q, \quad \forall x, y \in \mathbb{R}^n \tag{2.8}$$

In particular, we have that $|(x, y)| \leq \|x\|_1 \|y\|_\infty$, $\forall x, y \in \mathbb{R}^n$.

2.2 Orthogonal projections

Orthogonality

An inner product space V provides a means to generalize the concept of measuring angles between vectors, where in particular two vectors $x, y \in V$ are *orthogonal* if $(x, y) = 0$.

If a vector $x \in V$ is orthogonal to all vectors s in a subspace $S \subset V$, so that

$$(x, s) = 0, \quad \forall s \in S,$$

then x is said to be orthogonal to S. For example, the vector $(0, 0, 1)^T \in \mathbb{R}^3$ is orthogonal to the subspace spanned in $\mathbb{R}^3$ by the vectors $(1, 0, 0)^T$ and $(0, 1, 0)^T$.

We denote by $S^\perp$ the *orthogonal complement* of S in V, that is $S^\perp = \{v \in V : (v, s) = 0, \forall s \in S\}$. The only vector in V that is an element of both S and $S^\perp$ is the zero vector, and any vector $v \in V$ can be decomposed into two orthogonal components as $v = s_1 + s_2$, with $s_1 \in S$ and $s_2 \in S^\perp$.

Orthogonal projection

The *orthogonal projection* of a vector $x \in V$ in the direction of another vector $y \in V$, is the vector βy with $\beta = (x, y)/\|y\|^2 \in \mathbb{R}$, such that the difference between the two vectors is orthogonal to y, that is $(x - \beta y, y) = 0$.

The orthogonal projection of a vector $v \in V$ onto the subspace $S \subset V$ is a vector $v_s \in S$ such that

$$(v - v_s, s) = 0, \quad \forall s \in S. \tag{2.9}$$

The orthogonal projection is the best approximation in the subspace $S \subset V$, with respect to the norm induced by the inner product of V.

Theorem 3 (Best approximation property)**.**

$$\|v - v_s\| \leq \|v - s\|, \quad \forall s \in S \tag{2.10}$$

Proof. For any vector $s \in S$ we have that

$$\|v - v_s\|^2 = (v - v_s, v - v_s) = (v - v_s, v - s) + (v - v_s, s - v_s) = (v - v_s, v - s),$$

since $(v - v_s, s - v_s) = 0$, by (2.9) and the fact that $s - v_s \in S$. The result then follows from Cauchy-Schwarz inequality and division of both sides by the factor $\|v - v_s\|$,

$$(v - v_s, v - s) \leq \|v - v_s\|\|v - s\| \Rightarrow \|v - v_s\| \leq \|v - s\|.$$

□

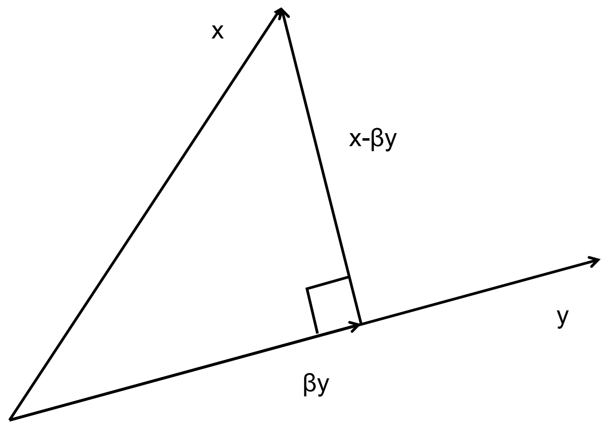

Figure 2.4: Illustration of βy, the projection of the x in the direction of y.

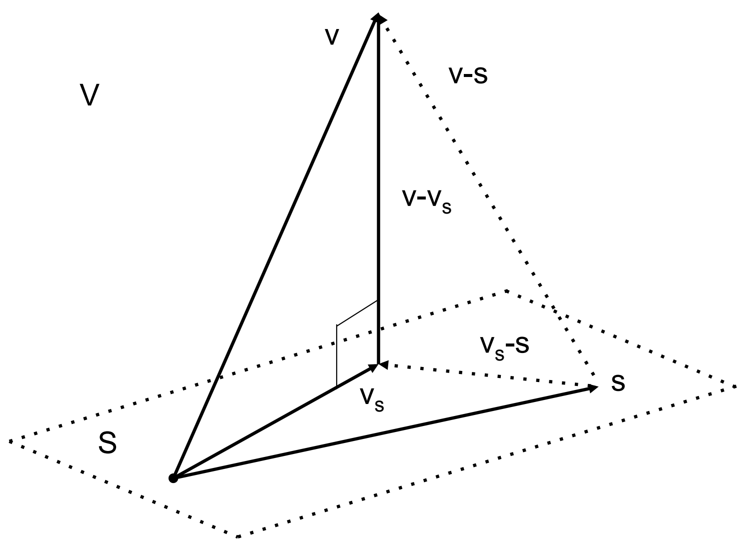

Figure 2.5: The orthogonal projection v_s is the best approximation in the subspace $S \subset V$.

Orthonormal basis

We refer to a set of non-zero vectors $\{v_i\}_{i=1}^n$ in the inner product space V as an *orthogonal set*, if all vectors v_i are pairwise orthogonal, that is if $(v_i, v_j) = 0$ for all $i \neq j$. If $\{v_i\}_{i=1}^n$ is an orthogonal set in the subspace $S \subset V$ and $\dim(S) = n$, then $\{v_i\}_{i=1}^n$ is a basis for S, that is all $v_s \in S$ can be expressed as

$$v_s = \alpha_1 v_1 + ... + \alpha_n v_n = \sum_{i=1}^n \alpha_i v_i, \tag{2.11}$$

with the coordinate $\alpha_i = (v_s, v_i)/\|v_i\|^2$ being the projection of v_s in the direction of the basis vector v_i.

If $Q = \{q_i\}_{i=1}^n$ is an orthogonal set, and $\|q_i\| = 1$ for all i, we say that Q is an *orthonormal set*. Let Q be an orthonormal basis for S, then

$$v_s = (v_s, q_1)q_1 + ... + (v_s, q_n)q_n = \sum_{i=1}^n (v_s, q_i)q_i, \quad \forall v_s \in S, \tag{2.12}$$

where the coordinate (v_s, q_i) is the projection of the vector v_s onto the basis vector q_i. An arbitrary vector $v \in V$ can be written

$$v = r + \sum_{i=1}^n (v, q_i)q_i, \tag{2.13}$$

where $r = v - \sum_{i=1}^n (v, q_i)q_i$. With $v_s = \sum_{i=1}^n (v, q_i)q_i$, the vector $r = v - v_s$ is orthogonal to Q, and thus orthogonal to S. By (2.9), the vector $r \in V$ satisfies the orthogonality condition

$$(r, s) = 0, \quad \forall s \in S, \tag{2.14}$$

and from (2.10) we know that r is the vector that corresponds to the minimal projection error of the vector v onto S.

2.3 Excercises

Problem 1. *Prove that the planes S_1 and S_2 are subspaces of $\mathbb{R}^3$, where $S_1 = \{x \in \mathbb{R}^3 : x_3 = 0\}$ and $S_2 = \{x \in \mathbb{R}^3 : ax_1 + bx_2 + cx_3 + d = 0 : a, b, c, d \in \mathbb{R}\}$.*

Problem 2. *Prove that the standard basis in $\mathbb{R}^n$ is linearly independent.*

Problem 3. *Prove that the Euclidian l_2-norm $\|\cdot\|_2$ is a norm.*

Problem 4. *Prove that the Euclidian scalar product* $(\cdot,\cdot)_2$ *is an inner product.*

Problem 5. *Prove that* $|(x,y)| \leq \|x\|_1\|y\|_\infty$, $\forall x, y \in \mathbb{R}^n$.

Problem 6. *Prove that the vector* $(0,0,1)^T \in \mathbb{R}^3$ *is orthogonal to the subspace spanned in* $\mathbb{R}^3$ *by the vectors* $(1,0,0)^T$ *and* $(0,1,0)^T$.

Problem 7. *Let* $\{w_i\}_{i=1}^n$ *be a basis for the subspace* $S \subset V$*, so that all* $s \in S$ *can be expressed as* $s = \sum_{i=1}^n \alpha_i w_i$.

(a) Prove that (2.9) is equivalent to finding the vector $v_s \in S$ *that satisfies the* n *equations of the form*

$$(v - v_s, w_i) = 0, \quad i = 1, ..., n.$$

(b) Since $v_s \in S$*, we have that* $v_s = \sum_{j=1}^n \beta_j w_j$*. Prove that (2.9) is equivalent to finding the set of coordinates* β_i *that satisfies*

$$\sum_{j=1}^n \beta_j(w_j, w_i) = (v, w_i), \quad i = 1, ..., n.$$

(c) Let $\{q_i\}_{i=1}^n$ *be an orthonormal basis for the subspace* $S \subset V$*, so that we can express* $v_s = \sum_{j=1}^n \beta_j q_j$*. Use the result in (b) to prove that (2.9) is equivalent to the condition that the coordinates are given as* $\beta_j = (v, q_j)$.