

## LINEAR COMBINATIONS

### 1. Finite-dimensional vector spaces

#### 1.1. Overview

The definition from Table 1 of a vector space reflects everyday experience with vectors in Euclidean geometry, and it is common to refer to such vectors by descriptions in a Cartesian coordinate system. For example, a position vector  $\mathbf{r}$  within the plane can be referred through the pair of coordinates  $(x, y)$ . This intuitive understanding can be made precise through the definition of a vector space  $\mathcal{R}_2 = (\mathbb{R}^2, \mathbb{R}, +, \cdot)$ , called the real 2-space. Vectors within  $\mathcal{R}_2$  are elements of  $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R} = \{(x, y) \mid x, y \in \mathbb{R}\}$ , meaning that a vector is specified through two real numbers,  $\mathbf{r} \leftrightarrow (x, y)$ . Addition of two vectors,  $\mathbf{q} \leftrightarrow (s, t)$ ,  $\mathbf{r} \leftrightarrow (x, y)$  is defined by addition of coordinates  $\mathbf{q} + \mathbf{r} = (s + x, t + y)$ . Scaling  $\mathbf{r} \leftrightarrow (x, y)$  by scalar  $a$  is defined by  $a\mathbf{r} \leftrightarrow (ax, ay)$ . Similarly, consideration of position vectors in three-dimensional space leads to the defini-

tion of the  $\mathcal{R}_3 = (\mathbb{R}^3, \mathbb{R}, +, \cdot)$ , or more generally a real  $m$ -space  $\mathcal{R}_m = (\mathbb{R}^m, \mathbb{R}, +, \cdot)$ ,  $m \in \mathbb{N}$ ,  $m > 0$ .

Addition rules for	$\forall \mathbf{a}, \mathbf{b}, \mathbf{c} \in V$
$\mathbf{a} + \mathbf{b} \in V$	Closure
$\mathbf{a} + (\mathbf{b} + \mathbf{c}) = (\mathbf{a} + \mathbf{b}) + \mathbf{c}$	Associativity
$\mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a}$	Commutativity
$\mathbf{0} + \mathbf{a} = \mathbf{a}$	Zero vector
$\mathbf{a} + (-\mathbf{a}) = \mathbf{0}$	Additive inverse
Scaling rules for	$\forall \mathbf{a}, \mathbf{b} \in V, \forall x, y \in S$
$x\mathbf{a} \in V$	Closure
$x(\mathbf{a} + \mathbf{b}) = x\mathbf{a} + x\mathbf{b}$	Distributivity
$(x + y)\mathbf{a} = x\mathbf{a} + y\mathbf{a}$	Distributivity
$x(y\mathbf{a}) = (xy)\mathbf{a}$	Composition
$1\mathbf{a} = \mathbf{a}$	Scalar identity

**Table 1.** Vector space  $\mathcal{V} = (V, S, +, \cdot)$  properties for arbitrary  $\mathbf{a}, \mathbf{b}, \mathbf{c} \in V$

Note however that there is no mention of coordinates in the definition of a vector space as can be seen from the list of properties in Table 1. The intent of such a definition is to highlight that besides position vectors, many other mathematical objects follow the same rules. As an example, consider the set of all continuous functions  $C(\mathbb{R}) = \{f \mid f: \mathbb{R} \rightarrow \mathbb{R}\}$ , with function addition defined by the sum at each argument  $t$ ,  $(f + g)(t) = f(t) + g(t)$ , and scaling by  $a \in \mathbb{R}$  defined as  $(af)(t) = af(t)$ . Read this as: “given two continuous functions  $f$  and  $g$ , the function  $f + g$  is defined by stating that its value for argument  $x$  is the sum of the two real numbers  $f(t)$  and  $g(t)$ ”. Similarly: “given a continuous function  $f$ , the function  $af$  is defined by stating that its value for argument  $t$  is the product of the real numbers  $a$  and  $f(t)$ ”. Under such definitions  $\mathcal{C}^0 = (C(\mathbb{R}), \mathbb{R}, +, \cdot)$  is a vector space, but quite different from  $\mathcal{R}_m$ . Nonetheless, the fact that both  $\mathcal{C}^0$  and  $\mathcal{R}_m$  are vector spaces can be used to obtain insight into the behavior of continuous functions from Euclidean vectors, and vice versa. This correspondence principle between discrete and continuous formulations is a recurring theme in scientific computation.

#### 1.2. Real vector space $\mathcal{R}_m$

**Column vectors.** Since the real spaces  $\mathcal{R}_m = (\mathbb{R}^m, \mathbb{R}, +, \cdot)$  play such an important role in themselves and as a guide to other vector spaces, familiarity with vector operations in  $\mathcal{R}_m$  is necessary to fully appreciate the utility of linear algebra to a wide range of applications. Following the usage in geometry and physics, the  $m$  real numbers that specify a vector  $\mathbf{u} \in \mathbb{R}^m$  are called the *components* of  $\mathbf{u}$ . The one-to-one correspondence between a vector and its components  $\mathbf{u} \leftrightarrow (u_1, \dots, u_m)$ , is by convention taken to define an equality relationship,

$$\mathbf{u} = \begin{bmatrix} u_1 \\ \vdots \\ u_m \end{bmatrix}, \tag{1}$$

with the components arranged vertically and enclosed in square brackets. Given two vectors  $\mathbf{u}, \mathbf{v} \in \mathbb{R}^m$ , and a scalar  $a \in \mathbb{R}$ , vector addition and scaling are defined in  $\mathcal{R}_m$  by real number addition and multiplication of components

$$\mathbf{u} + \mathbf{v} = \begin{bmatrix} u_1 \\ \vdots \\ u_m \end{bmatrix} + \begin{bmatrix} v_1 \\ \vdots \\ v_m \end{bmatrix} = \begin{bmatrix} u_1 + v_1 \\ \vdots \\ u_m + v_m \end{bmatrix}, a\mathbf{u} = a \begin{bmatrix} u_1 \\ \vdots \\ u_m \end{bmatrix} = \begin{bmatrix} au_1 \\ \vdots \\ au_m \end{bmatrix}. \tag{2}$$

The vector space  $\mathcal{R}_m$  is defined using the real numbers as the set of scalars, and constructing vectors by grouping together  $m$  scalars, but this approach can be extended to any set of scalars  $S$ , leading to the definition of the vector spaces  $\mathcal{S}_n = (S^n, S, +, \cdot)$ . These will often be referred to as  $n$ -vector space of scalars, signifying that the set of vectors is  $V = S^n$ .

To aid in visual recognition of vectors, the following notation conventions are introduced:

- vectors are denoted by lower-case bold Latin letters:  $\mathbf{u}, \mathbf{v}$ ;
- scalars are denoted by normal face Latin or Greek letters:  $a, b, \alpha, \beta$ ;
- the components of a vector are denoted by the corresponding normal face with subscripts as in equation (1);
- related sets of vectors are denoted by indexed bold Latin letters:  $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n$ .

**Row vectors.** Instead of the vertical placement or components into one *column*, the components of could have been placed horizontally in one *row*  $[u_1 \dots u_m]$ , that contains the same data, differently organized. By convention vertical placement of vector components is the preferred organization, and  $\mathbf{u}$  shall denote a *column vector* henceforth. A transpose operation denoted by a  $T$  superscript is introduced to relate the two representations

$$\mathbf{u}^T = [u_1 \dots u_m],$$

and  $\mathbf{u}^T$  is the notation used to denote a *row vector*.

- In Julia, horizontal placement of successive components in a row is denoted by a space.

**Compatible vectors.** Addition of real vectors  $\mathbf{u}, \mathbf{v} \in \mathbb{R}^m$  defines another vector  $\mathbf{w} = \mathbf{u} + \mathbf{v} \in \mathbb{R}^m$ . The components of  $\mathbf{w}$  are the sums of the corresponding components of  $\mathbf{u}$  and  $\mathbf{v}$ ,  $w_i = u_i + v_i$ , for  $i = 1, 2, \dots, m$ . Addition of vectors with different number of components is not defined, and attempting to add such vectors produces an error. Such vectors with different number of components are called *incompatible*, while vectors with the same number of components are said to be *compatible*. Scaling of  $\mathbf{u}$  by  $a$  defines a vector  $\mathbf{z} = a\mathbf{u}$ , whose components are  $z_i = au_i$ , for  $i = 1, 2, \dots, m$ .

### 1.3. Working with vectors

**Ranges.** The vectors used in applications usually have a large number of components,  $m \gg 1$ , and it is important to become proficient in their manipulation. Previous examples defined vectors by explicit listing of their  $m$  components. This is impractical for large  $m$ , and support is provided for automated generation for often-encountered situations. First, observe that Table 1 mentions one distinguished vector, the zero element that is a member of any vector space  $\mathbf{0} \in V$ . The zero vector of a real vector space  $\mathcal{R}_m$  is a column vector with  $m$  components, all of which are zero, and a mathematical convention for specifying this vector is  $\mathbf{0}^T = [0 \ 0 \ \dots \ 0] \in \mathbb{R}^m$ . This notation specifies that transpose of the zero vector is the row vector with  $m$  zero components, also written through explicit indexing of each component as  $\mathbf{0}_i = 0$ , for  $i = 1, \dots, m$ . Keep in mind that the zero vector  $\mathbf{0}$  and the zero scalar  $0$  are different mathematical objects. The ellipsis symbol in the mathematical notation is transcribed in Julia by the notion of a range, with  $1 : m$  denoting all the integers starting from 1 to  $m$ , organized as a row vector. The notation is extended to allow for strides different from one, and the mathematical ellipsis  $i = m, m-1, \dots, 1$  is denoted as  $m : -1 : 1$ . In general  $r : s : t$  denotes the set of numbers  $\{r, r+s, \dots, r+ns\}$  with  $r+ns \leq t$ , and  $r, s, t$  real numbers and  $n$  a natural number,  $r, s, t \in \mathbb{R}$ ,  $n \in \mathbb{N}$ . If there is no natural number  $n$  such that  $r+ns \leq t$ , an empty vector with no components is returned.

## 2. Linear combinations

### 2.1. Linear combination as a matrix-vector product

The expression  $\mathbf{x} = x_1 \mathbf{e}_1 + x_2 \mathbf{e}_2 + \dots + x_m \mathbf{e}_m$  expresses the idea of scaling vectors within a set and subsequent addition to form a new vector  $\mathbf{x}$ . The matrix  $\mathbf{I} = [\mathbf{e}_1 \ \mathbf{e}_2 \ \dots \ \mathbf{e}_m]$  groups these vectors together in a single entity, and the scaling factors are the components of the vector  $\mathbf{x}$ . To bring all these concepts together it is natural to consider the notation

$$\mathbf{x} = \mathbf{I}\mathbf{x},$$

as a generalization of the scalar expression  $x = 1 \cdot x$ . It is clear what the operation  $\mathbf{I}\mathbf{x}$  should signify: it should capture the vector scaling and subsequent vector addition  $x_1 \mathbf{e}_1 + x_2 \mathbf{e}_2 + \dots + x_m \mathbf{e}_m$ . A specific meaning is now ascribed to  $\mathbf{I}\mathbf{x}$  by identifying two definitions to one another.

**Linear combination.** Repeatedly stating “vector scaling and subsequent vector addition” is unwieldy, so a special term is introduced for some given set of vectors  $\{\mathbf{a}_1, \dots, \mathbf{a}_n\}$ .

DEFINITION. (LINEAR COMBINATION). The *linear combination* of vectors  $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n \in V$  with scalars  $x_1, x_2, \dots, x_n \in S$  in vector space  $(V, S, +, \cdot)$  is the vector  $\mathbf{b} = x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \dots + x_n\mathbf{a}_n$ .

**Matrix-vector product.** Similar to the grouping of unit vectors  $\mathbf{e}_1, \dots, \mathbf{e}_m$  into the identity matrix  $\mathbf{I}$ , a more concise way of referring to arbitrary vectors  $\mathbf{a}_1, \dots, \mathbf{a}_n$  from the same vector space is the matrix  $\mathbf{A} = [\mathbf{a}_1 \ \mathbf{a}_2 \ \dots \ \mathbf{a}_n]$ . Combining these observations leads to the definition of a matrix-vector product.

DEFINITION. (MATRIX-VECTOR PRODUCT). In the vector space  $(V, S, +, \cdot)$ , the product of matrix  $\mathbf{A} = [\mathbf{a}_1 \ \mathbf{a}_2 \ \dots \ \mathbf{a}_n]$  composed of columns  $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n \in V$  with the vector  $\mathbf{x} \in S_n$  whose components are scalars  $x_1, x_2, \dots, x_n \in S$  is the linear combination  $\mathbf{b} = x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \dots + x_n\mathbf{a}_n = \mathbf{A}\mathbf{x} \in V$ .

**2.2. Linear algebra problem examples**

**Linear combinations in  $E_2$ .** Consider a simple example that leads to a common linear algebra problem: decomposition of forces in the plane along two directions. Suppose a force is given in terms of components along the Cartesian  $x, y$ -axes,  $\mathbf{b} = b_x\mathbf{e}_x + b_y\mathbf{e}_y$ , as expressed by the matrix-vector multiplication  $\mathbf{b} = \mathbf{I}\mathbf{b}$ . Note that the same force could be obtained by linear combination of other vectors, for instance the normal and tangential components of the force applied on an inclined plane with angle  $\theta$ ,  $\mathbf{b} = x_t\mathbf{e}_t + x_n\mathbf{e}_n$ , as in Figure 1. This defines an alternate reference system for the problem. The unit vectors along these directions are

- $$\mathbf{t} = \begin{bmatrix} \cos \theta \\ \sin \theta \end{bmatrix}, \mathbf{n} = \begin{bmatrix} -\sin \theta \\ \cos \theta \end{bmatrix},$$

```

∴ θ=π/6.; c=cos(θ); s=sin(θ); t=[c; s]; n=[-s; c];
∴
```

and can be combined into a matrix  $\mathbf{A} = [\mathbf{t} \ \mathbf{n}]$ . The value of the components  $(x_t, x_n)$  are the scaling factors and can be combined into a vector  $\mathbf{x} = [x_t \ x_n]^T$ . The same force must result irrespective of whether its components are given along the Cartesian axes or the inclined plane directions leading to the equality

- $$\mathbf{I}\mathbf{b} = \mathbf{b} = \mathbf{A}\mathbf{x}. \tag{9}$$

```

∴ b=[0.2; 0.4]; I*b
```

$$\begin{bmatrix} 0.2 \\ 0.4 \end{bmatrix} \tag{10}$$

```

∴
```

Interpret equation (9) to state that the vector  $\mathbf{b}$  could be obtained either as a linear combination of  $\mathbf{I}$ ,  $\mathbf{b} = \mathbf{I}\mathbf{b}$ , or as a linear combination of the columns of  $\mathbf{A}$ ,  $\mathbf{b} = \mathbf{A}\mathbf{x}$ . Of course the simpler description seems to be  $\mathbf{I}\mathbf{b}$  for which the components are already known. But this is only due to an arbitrary choice made by a human observer to define the force in terms of horizontal and vertical components. The problem itself suggests that the tangential and normal components are more relevant; for instance a friction force would be evaluated as a scaling of the normal force.

- The components of  $\mathbf{b}$  in this more natural reference system are not known, but can be determined by solving the vector equality  $\mathbf{A}\mathbf{x} = \mathbf{I}\mathbf{b} = \mathbf{b}$ , known as a *linear system of equations*, implemented in many programming environments (Julia, Matlab, Octave) through the backslash operator  $\mathbf{x} = \mathbf{A} \setminus \mathbf{b}$ .

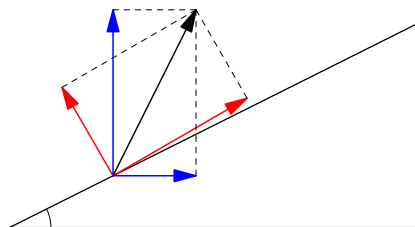


Figure 1. Alternative decompositions of force on inclined plane.

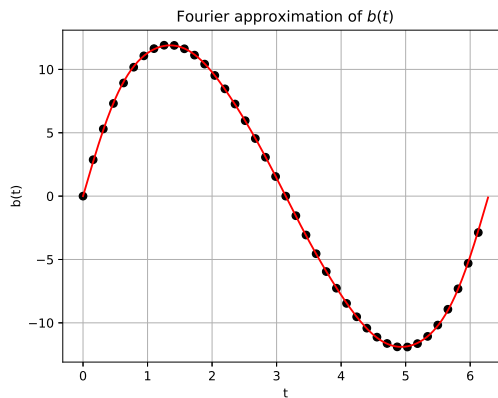
**Linear combinations in  $\mathcal{R}_m$  and  $\mathcal{C}^0[0, 2\pi]$ .** Linear combinations in a real space can suggest properties or approximations of more complex objects such as continuous functions. Let  $\mathcal{C}^0[0, 2\pi] = (C[0, 2\pi], \mathbb{R}, +, \cdot)$  denote the vector space of continuous functions that are periodic on the interval  $[0, 2\pi)$ ,  $C[0, \pi] = \{f \mid f: \mathbb{R} \rightarrow \mathbb{R}, f(t) = f(t + 2\pi)\}$ . Recall that vector addition is defined by  $(f + g)(t) = f(t) + g(t)$ , and scaling by  $(af)(t) = af(t)$ , for  $f, g \in C[0, 2\pi)$ ,  $a \in \mathbb{R}$ . Familiar functions within this vector space are  $\sin(kt)$ ,  $\cos(kt)$  with  $k \in \mathbb{N}$ , and these can be recognized to intrinsically represent periodicity on  $[0, 2\pi)$ , a role analogous to the normal and tangential directions in the inclined plane example. Define now another periodic function  $b(t + 2\pi) = b(t)$  by repeating the values  $b(t) = t(\pi - t)(2\pi - t)$  from the interval  $[0, 2\pi)$  on all intervals  $[2p\pi, 2(p + 1)\pi]$ , for  $p \in \mathbb{Z}$ . The function  $b$  is not given in terms of the “naturally” periodic functions  $\sin(kt)$ ,  $\cos(kt)$ , but could it thus be expressed? This can be stated as seeking a linear combination  $b(t) = \sum_{k=1}^{\infty} x_k \sin(kt)$ , as studied in Fourier analysis. The coefficients  $x_k$  could be determined from an analytical formula involving calculus operations  $x_k = \frac{1}{\pi} \int_0^{2\pi} b(t) \sin(kt) dt$ , but we'll seek an approximation using a linear combination of  $n$  terms

$$b(t) \cong \sum_{k=1}^n x_k \sin(kt), A(t) = [\sin(t) \sin(2t) \dots \sin(nt)], A: \mathbb{R} \rightarrow \mathbb{R}^n.$$

Organize this as a matrix vector product  $b(\mathbf{t}) \cong A(\mathbf{t}) \mathbf{x}$ , with

$$A(\mathbf{t}) = [\sin(t) \sin(2t) \dots \sin(nt)], \mathbf{x} = [x_1 \ x_2 \ \dots \ x_n]^T \in \mathbb{R}^n.$$

The idea is to sample the column vectors of  $A(\mathbf{t})$  at the components of the vector  $\mathbf{t} = [t_1 \ t_2 \ \dots \ t_m]^T \in \mathbb{R}^m$ ,  $t_j = (j - 1)h$ ,  $j = 1, 2, \dots, m$ ,  $h = \pi / m$ . Let  $\mathbf{b} = b(\mathbf{t})$ , and  $\mathbf{A} = A(\mathbf{t})$ , denote the so-sampled  $b, A$  functions leading to the definition of a vector  $\mathbf{b} \in \mathbb{R}^m$  and a matrix  $\mathbf{A} \in \mathbb{R}^{m \times n}$ . There are  $n$  coefficients available to scale the column vectors of  $\mathbf{A}$ , and  $\mathbf{b}$  has  $m$  components. For  $m > n$  it is generally not possible to find  $\mathbf{x}$  such that  $\mathbf{A}\mathbf{x}$  would exactly equal  $\mathbf{b}$ , but as seen later the condition to be as close as possible to  $\mathbf{b}$  leads to a well defined solution procedure. This is known as a least squares problem and is automatically applied in the `x=A\b` instruction when the matrix  $A$  is not square. As seen in the following numerical experiment and Figure 2, the approximation is excellent and the information conveyed by  $m = 1000$  samples of  $b(t)$  is now much more efficiently stored in the form chosen for the columns of  $\mathbf{A}$  and the  $n = 5$  scaling coefficients that are the components of  $\mathbf{x}$ .



**Figure 2.** Comparison of least squares approximation (red line) with samples (black dots) of exact function  $b(t) = t(\pi - t)(2\pi - t)$

### Summary.

- A widely used framework for constructing additive approximations is the vector space algebraic space structure in which scaling and addition operations are defined
- In a vector space linear combinations are used to construct more complicated objects from simpler ones

$$\mathbf{b} = \mathbf{A}\mathbf{x} = x_1 \mathbf{a}_1 + \dots + x_n \mathbf{a}_n$$