

Chapter 20. Vector Spaces, Bases and Dimension

In this course, we have proceeded step-by-step through low-dimensional Linear Algebra. We have looked at lines, planes, hyperplanes, and have seen that there is no limit to this hierarchy of structures. We have indexed these objects by their “dimension” in an *ad hoc* way. It is time to unify these structures and ideas; this chapter gives a brief introduction to this more abstract viewpoint. Paradoxically, this abstraction also makes Linear Algebra more applicable to other areas of Mathematics and Science.

In this higher viewpoint, Linear Algebra is the study of Vector Spaces and Linear Mappings between vector spaces. A **vector space** is a set \mathbf{V} of objects \mathbf{v} called “vectors” which can be added and multiplied by scalars $t \in \mathbb{R}$, subject to the rules:

$$\begin{aligned}\mathbf{u} + \mathbf{v} &= \mathbf{v} + \mathbf{u}, & (\mathbf{u} + \mathbf{v}) + \mathbf{w} &= \mathbf{u} + (\mathbf{v} + \mathbf{w}), \\ t(\mathbf{u} + \mathbf{v}) &= t\mathbf{u} + t\mathbf{v}, & (t + s)\mathbf{v} &= t\mathbf{v} + s\mathbf{v}, & t(s\mathbf{v}) &= (ts)\mathbf{v},\end{aligned}$$

these holding for all $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbf{V}$ and $s, t \in \mathbb{R}$. There must also be a “zero vector” $\mathbf{0} \in \mathbf{V}$ with the properties that

$$\mathbf{0} + \mathbf{v} = \mathbf{v}, \quad 0\mathbf{v} = \mathbf{0}$$

for all $\mathbf{v} \in \mathbf{V}$.

Examples:

- \mathbb{R}^n is a vector space. Any line, plane, hyperplane,... in \mathbb{R}^n is a vector space. Even the origin $\mathbf{0}$ is a vector space (consisting of the single vector $\mathbf{0}$).
- The set \mathbf{P} of all polynomials $a_0 + a_1x + \cdots + a_nx^n$ with real coefficients is a vector space. Here the “vectors” are the polynomials. The infinitely many powers $1, x, x^2, \dots$ play the role of the standard basis $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ in \mathbb{R}^n .
- The set \mathbf{A} of all power series $a_0 + a_1x + \cdots$ with positive radius of convergence. The letter “ \mathbf{A} ” stands for “analytic functions”. The vector space \mathbf{A} is much larger than \mathbf{P} ; for example the analytic function

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots,$$

is a vector in the vector space \mathbf{A} .

Any nonzero vector space \mathbf{V} contains infinitely many vectors. To write them all down, we want to have something like the standard basis $\mathbf{e}_1, \dots, \mathbf{e}_n$ of \mathbb{R}^n . Here, we have

$$(x_1, x_2, \dots, x_n) = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \dots + x_n\mathbf{e}_n. \quad (1)$$

Thus, any vector $\mathbf{x} = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$ is a linear combination of the vectors $\mathbf{e}_1, \dots, \mathbf{e}_n$. Moreover, this linear combination is unique: Equation (1) is the only way to write \mathbf{x} as a combination of the \mathbf{e}_i . This is what we want to have in any vector space.

Definition of a Basis: A **basis** of a vector space \mathbf{V} is a set \mathbf{B} of vectors in \mathbf{V} with the property that every vector in \mathbf{V} can be uniquely expressed as a linear combination of vectors in \mathbf{B} .

Examples of Bases:

- The standard basis $\mathbf{B} = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ is a basis of \mathbb{R}^n .
- A basis of a line L is a nonzero vector in L .
- A basis of a plane P is a pair of vectors in P not lying on the same line.
- A basis of \mathbb{R}^3 is any set of three vectors not contained in a plane.
- A basis of a hyperplane $H \subset \mathbb{R}^4$ is a set of three vectors in H not contained in a plane.
- A basis of \mathbb{R}^4 is a set of four vectors not contained in a hyperplane.
- The infinite set $\mathbf{B} = \{1, x, x^2, x^3, \dots\}$ is a basis of \mathbf{P} .

The definition of Basis is simple to state, but the idea of a Basis takes some time to fully digest, since the phrase “can be uniquely” is actually two conditions in compressed form. We now uncompress them, starting with “can be”.

Definition of Span : *The span of a collection of vectors $\mathbf{B} \subset \mathbf{V}$ is the set*

$$\mathbb{R}\mathbf{B} = \{c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v} : c_1, c_2, \dots, c_n \in \mathbb{R}\}$$

of all linear combinations of vectors in \mathbf{B} . We say that \mathbf{B} spans \mathbf{V} if $\mathbb{R}\mathbf{B} = \mathbf{V}$.

In other words, \mathbf{B} spans \mathbf{V} if every vector in V can be written as a linear combination of vectors in \mathbf{B} . This is consistent with our previous use of “span” for lines and planes, as in “line $\mathbb{R}\mathbf{u}$ through \mathbf{u} ” or “plane spanned by \mathbf{u} and \mathbf{v} ”.

Now for the second part of the definition of Basis: the uniqueness.

Definition of Linear Independence: *A collection of vectors \mathbf{B} in a vector space \mathbf{V} is Linearly Independent if no vector in \mathbf{B} is a linear combination of the other vectors in \mathbf{B} .*

If some vector in \mathbf{B} is a linear combination of the others, we say that \mathbf{B} is Linearly Dependent. If we have vectors $\mathbf{v}_1, \dots, \mathbf{v}_n, \mathbf{u}_1, \dots, \mathbf{u}_m$ in \mathbf{B} and nonzero scalars $c_1, \dots, c_n, b_1, \dots, b_m$ such that

$$c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n = b_1\mathbf{u}_1 + \cdots + b_m\mathbf{u}_m,$$

Then \mathbf{v}_1 is a linear combination of $\mathbf{v}_2, \dots, \mathbf{v}_n, \mathbf{u}_1, \dots, \mathbf{u}_m$, so \mathbf{B} is linearly dependent. Conversely, if \mathbf{B} is linearly dependent, then we can write some \mathbf{v} in \mathbf{B} as a linear combination of the other vectors in \mathbf{B} :

$$\mathbf{v} = b_1\mathbf{u}_1 + \cdots + b_m\mathbf{u}_m,$$

which is two linear combinations giving the same vector. So Linear Independence is equivalent to Uniqueness of Linear Combinations.

As a practical test, \mathbf{B} is linearly independent exactly when the only way to write $\mathbf{0}$ as a linear combination

$$c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n = \mathbf{0}$$

of vectors $\mathbf{v}_i \in \mathbf{B}$ is if all the coefficients c_i are zero.

Definition of a Basis, rephrased: A Basis of a vector space V is a linearly independent subset $B \subset V$ which spans V .

To check that a given subset $B \subset V$ is a basis, there are two steps:

(i) *Show that B spans V :* Take an arbitrary vector $\mathbf{v} \in V$, and show that there are vectors $\mathbf{v}_1, \dots, \mathbf{v}_n$ in B and scalars c_1, \dots, c_n such that

$$\mathbf{v} = c_1\mathbf{v}_1 + \dots + c_n\mathbf{v}_n.$$

(ii) *Show that B is linearly independent:* Prove that an equation of the form

$$c_1\mathbf{v}_1 + \dots + c_n\mathbf{v}_n = \mathbf{0}$$

with all $\mathbf{v}_i \in B$ can only hold if all $c_1 = c_2 = \dots = c_n = 0$.

Example 1: Suppose we have a set B of n vectors $\mathbf{v}_1, \dots, \mathbf{v}_n$ in \mathbb{R}^n . Let B be the $n \times n$ matrix sending \mathbf{e}_i to \mathbf{v}_i . For any scalars c_1, \dots, c_n we have

$$B(c_1\mathbf{e}_1 + \dots + c_n\mathbf{e}_n) = c_1\mathbf{v}_1 + \dots + c_n\mathbf{v}_n.$$

Hence B is linearly independent exactly when $\ker B = \mathbf{0}$ and B spans \mathbb{R}^n exactly when $\text{im } B = \mathbb{R}^n$. In fact, the following conditions are equivalent:

1. $\det(B) \neq 0$;
2. B is invertible;
3. $\ker B = \mathbf{0}$;
4. $\text{im } B = \mathbb{R}^n$;
5. The columns of B are linearly independent
6. The columns of B span \mathbb{R}^n .

Indeed, the inverse formula that we have seen for $n \leq 4$ holds in general, so the first two are equivalent. It is easy to see that B invertible implies 3 and 4. Conversely 3 and 4 imply that B is one-to-one and onto, so is invertible. The Kernel-Image Theorem (see below) implies that 3 and 4 are equivalent to each other. Finally we have seen that 5 and 6 are just restatements of 3 and 4.

This result is not so much about \mathbb{R}^n as it is about changing bases. In the above discussion, replace \mathbb{R}^n by any vector space V and suppose $\mathbf{e}_1, \dots, \mathbf{e}_n$ is any basis of V . Then the above result holds without change.

Example 2: Let \mathbf{V} be the hyperplane in \mathbb{R}^4 with equation

$$x + y + z + w = 0.$$

The vectors

$$\mathbf{v}_1 = \mathbf{e}_1 - \mathbf{e}_2, \quad \mathbf{v}_2 = \mathbf{e}_2 - \mathbf{e}_3, \quad \mathbf{v}_3 = \mathbf{e}_3 - \mathbf{e}_4$$

are in \mathbf{V} . Let's show that the set $\mathbf{B} = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is a basis of \mathbf{V} . To check spanning, let $\mathbf{v} = (x, y, z, w)$ be an arbitrary vector in \mathbf{V} . We want to find scalars c_1, c_2, c_3 such that

$$\mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3.$$

Since $w = -x - y - z$, we have

$$\mathbf{v} = x\mathbf{v}_1 + (x + y)\mathbf{v}_2 + (x + y + z)\mathbf{v}_3.$$

This shows that \mathbf{B} spans \mathbf{V} . To check linear independence, suppose we have scalars c_1, c_2, c_3 such that

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3 = \mathbf{0}.$$

Writing both sides out in terms of components, we have

$$(c_1, c_2 - c_1, c_3 - c_2, -c_3) = (0, 0, 0, 0),$$

which amounts to the equations

$$\begin{aligned} c_1 &= 0 \\ c_2 - c_1 &= 0 \\ c_3 - c_2 &= 0 \\ -c_3 &= 0. \end{aligned}$$

The only solution to this equation is $c_1 = c_2 = c_3 = 0$. This shows that \mathbf{B} is linearly independent. Hence \mathbf{B} is a basis of \mathbf{V} .

Let us now take the vector space \mathbf{V}' of vectors (x, y, z, w, t) in \mathbb{R}^5 satisfying the same equation $x + y + z + w = 0$. Then the set \mathbf{B} above is no longer a basis of \mathbf{V}' : the vector \mathbf{e}_5 is in \mathbf{V}' since the equation does not care about the

fifth coordinate, but \mathbf{e}_5 is not in $\mathbb{R}\mathbf{B}$. It is easy to check that the enlarged set $\mathbf{B}' = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{e}_5\}$ is a basis of \mathbf{V}' .

Example 3: Let's show that the vector space \mathbf{P} of all polynomials has the basis $\mathbf{B} = \{1, x, x^2, x^3, \dots\}$. First, by definition, every vector in \mathbf{P} is a polynomial, which is of the form

$$a_0 + a_1x + \dots + a_nx^n$$

for some n . This shows that \mathbf{B} spans \mathbf{P} . A polynomial is the zero polynomial exactly when its coefficients a_i are all zero. This shows that \mathbf{B} is linearly independent and is therefore a basis of \mathbf{V} .

The set \mathbf{B} is the most obvious basis of \mathbf{P} , but for many purposes it is not the best one. For numerical integration, one prefers to use instead the basis

$$\mathbf{B}_L = \{P_0, P_1, P_2, \dots\}$$

where the P_n are the **Legendre Polynomials**. These are the unique polynomials satisfying the conditions

$$P_0 = 1, \quad \deg P_n = n, \quad P_n(1) = 1, \quad \int_{-1}^1 P_n P_m = 0 \text{ if } n \neq m.$$

The first few of these polynomials are

$$P_1 = x, \quad P_2 = \frac{1}{2}(3x^2 - 1), \quad P_3 = \frac{1}{2}(5x^3 - 3x). \quad (2)$$

Google can tell you more about the Legendre Polynomials.

Example 4: The set $\mathbf{B} = \{1, x, x^2, x^3, \dots\}$ is not a basis of the vector space \mathbf{A} of analytic functions because there are vectors in \mathbf{A} , such as e^x , that are not finite linear combinations of elements of \mathbf{B} . In fact, I don't know of an explicit basis for \mathbf{A} . However, let us consider the vector space

$$\mathbf{V} = \{f \in \mathbf{A} : f'' = f\}.$$

Any function $f \in \mathbf{A}$ is a power series:

$$f(x) = a_0 + a_1x + a_2x^2 + a_3x^3 + \dots$$

Its second derivative is

$$f''(x) = 2a_2 + 6a_3x + \dots$$

To get $f'' = f$, you compare coefficients and find that

$$f(x) = a_0e^x + a_1e^{-x}.$$

This shows that the set $\mathbf{B} = \{e^x, e^{-x}\}$ spans \mathbf{V} . To see that \mathbf{B} is linearly independent, suppose that

$$c_1e^x + c_2e^{-x} = 0,$$

for some numbers c_1, c_2 . Here 0 is the zero function, since that is the zero element of our vector space \mathbf{A} . Differentiating, we get

$$c_1e^x - c_2e^{-x} = 0.$$

Adding these two equations gives

$$2c_1e^x = 0,$$

hence $c_1 = 0$, so $c_2e^{-x} = 0$, so $c_2 = 0$. Therefore \mathbf{B} is linearly independent. This means that every solution of the Differential Equation $f'' = f$ can be written uniquely as a linear combination of the two functions e^x and e^{-x} . This is a typical example of how vector spaces and bases appear in other areas of Mathematics and Science.

Dimension

A vector space V is **finite dimensional** if it has a finite basis. So \mathbb{R}^n is finite dimensional, while \mathbf{P} and \mathbf{A} are infinite dimensional. The fundamental fact here is the following

Independence of Basis Theorem: *In a finite dimensional vector space, any two bases have the same number of elements.*

We have defined “finite dimensional” without yet defining dimension. However, the Basis Theorem allows us to define dimension precisely:

The **dimension** of a finite dimensional vector space V , denoted $\dim V$, is the number of vectors in any basis of V .

Clearly \mathbb{R}^n has dimension n . Returning to Example 2, the hyperplane \mathbf{V} defined by $x + y + z + w = 0$ in \mathbb{R}^4 has $\dim V = 3$, and the vector space \mathbf{V}' defined by $x + y + z + w = 0$ in \mathbb{R}^5 has $\dim \mathbf{V}' = 4$.

In Example 3, the vector space of all polynomials \mathbf{P} is infinite dimensional. The vector space \mathbf{P}_n of polynomials of degree $\leq n$ has (among many others) the bases

$$\mathbf{B} = \{1, x, \dots, x^n\}, \quad \mathbf{B}' = \{P_0, P_1, \dots, P_n\}.$$

All bases of \mathbf{P}_n have $n + 1$ elements, and we have $\dim \mathbf{P}_n = n + 1$.

In Example 4, the vector space \mathbf{V} of solutions to $f'' = f$ has (among many others) the bases

$$\mathbf{B} = \{e^x, e^{-x}\}, \quad \mathbf{B}' = \{\cosh x, \sinh x\}.$$

All bases of \mathbf{V} have two elements, and we have $\dim \mathbf{V} = 2$.

Subspaces

This is the generalization of a line, plane, hyperplane, etc. A **Subspace** of a vector space \mathbf{V} is a subset $\mathbf{U} \subset \mathbf{V}$ which is closed under addition and scalar multiplication. That is, if \mathbf{u}, \mathbf{v} are any two vectors in \mathbf{U} then $\mathbf{u} + \mathbf{v}$ is again in \mathbf{U} and if $t \in \mathbb{R}$, we have $t\mathbf{u} \in \mathbf{U}$. A subspace $\mathbf{U} \subset \mathbf{V}$ is itself a vector space, under the same operations as \mathbf{V} . Every vector space \mathbf{V} has at least two subspaces, namely $\mathbf{0}$ and \mathbf{V} itself.

An infinite dimensional vector space can have finite dimensional subspaces. In Example 2, the subspace \mathbf{P}_n of polynomials of degree $\leq n$ is a finite dimensional subspace of \mathbf{P} . In Example 3, the solutions of $f'' = f$ form a two-dimensional subspace of the infinite dimensional vector space \mathbf{A} .

For finite dimensional vector spaces we have the

Subspace Dimension Theorem *If \mathbf{U} is a subspace of a finite dimensional vector space \mathbf{V} , then $\dim \mathbf{U} \leq \dim \mathbf{V}$. If $\dim \mathbf{U} = \dim \mathbf{V}$ then $\mathbf{U} = \mathbf{V}$.*

Much of our previous work, for example computing kernels of matrices, is about interesections of subspaces. If \mathbf{U} and \mathbf{U}' are two subspaces of a vector space \mathbf{V} then the intersection $\mathbf{U} \cap \mathbf{U}'$ is also a subspace of \mathbf{V} . Assume that \mathbf{V} is finite dimensional, say $\dim \mathbf{V} = n$. Then we have the inequality

$$\dim(\mathbf{U} \cap \mathbf{U}') \geq \dim \mathbf{U} + \dim \mathbf{U}' - n, \quad (3)$$

but the exact dimension of $\mathbf{U} \cap \mathbf{U}'$ depends on the geometric relationship of \mathbf{U}

and U' . Let's suppose $U' = H$ is a hyperplane. Then we have

$$\dim(U \cap H) = \begin{cases} \dim U - 1 & \text{if } U \text{ is not contained in } H, \\ \dim U & \text{if } U \text{ is contained in } H. \end{cases} \quad (4)$$

The intersection of k hyperplanes

$$H_1 \cap \cdots \cap H_k \quad (5)$$

is obtained from a sequence of intersections

$$H_1 \cap H_2, \quad (H_1 \cap H_2) \cap H_3, \quad \dots,$$

where each time the dimension drops by 0 or 1, according to (4). So we have

$$\dim(H_1 \cap \cdots \cap H_k) \geq n - k,$$

with equality exactly when the dimension drops by 1 at each intersection, meaning that $H_1 \cap \cdots \cap H_i$ is not contained in H_{i+1} for $1 \leq i < k$.

If $V = \mathbb{R}^n$, where we have the dot product $\langle \cdot, \cdot \rangle$, each hyperplane H has a normal vector \mathbf{n} , such that

$$H = \{\mathbf{v} \in \mathbb{R}^n : \langle \mathbf{v}, \mathbf{n} \rangle = 0\}.$$

Take k hyperplanes H_1, \dots, H_k in \mathbb{R}^n and let \mathbf{n}_i be a normal vector for H_i . Then we have

$$\dim H_1 \cap \cdots \cap H_k \geq n - k,$$

with equality exactly when the normal vectors $\mathbf{n}_1, \dots, \mathbf{n}_k$ are linearly independent. (See Exercise 20.7 below.)

If our hyperplanes come from the rows of a $k \times n$ matrix, then the intersection of hyperplanes is exactly the kernel of A . Hence we have

$$\dim \ker A \geq n - k$$

with equality exactly when the rows of A are linearly independent.

Exercise 20.1 Find a basis of the hyperplane in \mathbb{R}^4 with equation

$$x + 2y + 3z + 4w = 0.$$

Exercise 20.2 Let \mathbf{P}_n be the vector space of polynomials of degree at most n . Let p_0, p_1, \dots, p_n be polynomials with $\deg(p_i) = i$. Show that $\{p_0, p_1, \dots, p_n\}$ is a basis of \mathbf{P}_n .

Exercise 20.3 On \mathbf{P}_n , we have the an analogue of the dot product, given by

$$\langle p, q \rangle = \int_{-1}^1 p(x)q(x) dx.$$

A basis $\{p_0, p_1, \dots, p_n\}$ of \mathbf{P}_n is orthogonal if $\langle p_i, p_j \rangle = 0$ for $i \neq j$. Show that the Legendre polynomials $\{P_0, P_1, P_2\}$ (see (2) above) are an orthogonal basis of \mathbf{P}_2 .

Exercise 20.4 Let \mathbf{U} be the subspace of analytic functions given by

$$\mathbf{U} = \{f \in \mathbf{A} : f'' - 2f' + f = 0\}.$$

(a) Show that the functions e^x and xe^x belong to \mathbf{U} .

(b) Show that the functions e^x and xe^x are linearly independent.

In fact $\{e^x, xe^x\}$ is a basis of \mathbf{U} , but that is a bit harder to prove.

Exercise 20.5 What are the possible dimensions of the intersection $\mathbf{H}_1 \cap \mathbf{H}_2 \cap \mathbf{H}_3$ of three hyperplanes in \mathbb{R}^6 ? Give an example of each case.

Exercise 20.6 Let \mathbf{U} be a subspace of \mathbb{R}^n , and let $\mathbf{u}_1, \dots, \mathbf{u}_k$ be an orthonormal basis of \mathbf{U} (i.e., $\langle \mathbf{u}_i, \mathbf{u}_j \rangle = 1$ if $i = j$ zero otherwise). The orthogonal space of \mathbf{U} is the subspace

$$\mathbf{U}^\perp = \{\mathbf{v} \in \mathbb{R}^n : \langle \mathbf{v}, \mathbf{u} \rangle = 0 \text{ for all } \mathbf{u} \in \mathbf{U}\}.$$

(a) Show that $\mathbf{U} \cap \mathbf{U}^\perp = \mathbf{0}$. Hint: If $\mathbf{u} \in \mathbf{U} \cap \mathbf{U}^\perp$, what is $\langle \mathbf{u}, \mathbf{u} \rangle$?

(b) Show that for any $\mathbf{v} \in \mathbb{R}^n$, the vector

$$\mathbf{w} = \mathbf{v} - \sum_{i=1}^k \langle \mathbf{v}, \mathbf{u}_i \rangle \mathbf{u}_i$$

belongs to the orthogonal space \mathbf{U}^\perp . Hint: show that $\langle \mathbf{w}, \mathbf{u}_i \rangle = 0$ for all i .

(c) Use parts (a) and (b) to show that any vector $\mathbf{v} \in \mathbb{R}^n$ can be written uniquely as $\mathbf{v} = \mathbf{u} + \mathbf{u}'$, with $\mathbf{u} \in \mathbf{U}$ and $\mathbf{u}' \in \mathbf{U}^\perp$.

Exercise 20.7 Take hyperplanes $\mathbf{H}_1, \dots, \mathbf{H}_k$ in \mathbb{R}^n , and let \mathbf{n}_i be a normal vector of \mathbf{H}_i . Let $\mathbf{U} = \mathbb{R}\{\mathbf{n}_i : i < k\}$ be the span of all these normal vectors except \mathbf{n}_k .

(a) Suppose that \mathbf{n}_k belongs to \mathbf{U} . Show that

$$\mathbf{H}_1 \cap \dots \cap \mathbf{H}_{k-1} \text{ is contained in } \mathbf{H}_k. \quad (6)$$

(b) Suppose that (6) holds. Show that \mathbf{n}_k belongs to \mathbf{U} .

Hint: The intersection in (6) is \mathbf{U}^\perp . Write $\mathbf{n}_k = \mathbf{u} + \mathbf{u}'$, with $\mathbf{u} \in \mathbf{U}$ and $\mathbf{u}' \in \mathbf{U}^\perp$, and show that $\langle \mathbf{u}', \mathbf{u}' \rangle = 0$.