11

Basis and Dimension

In chapter 10, the notions of a linearly independent set of vectors in a vector space V, and of a set of vectors that span V were established; any set of vectors that span V can be reduced to some minimal collection of linearly independent vectors; such a minimal set is called a *basis* of the subspace V.

Definition Let V be a vector space. Then a set S is a **basis** for V if S is linearly independent and $V = \operatorname{span} S$.

If S is a basis of V and S has only finitely many elements, then we say that V is **finite-dimensional**. The number of vectors in S is the **dimension** of V.

Suppose V is a *finite-dimensional* vector space, and S and T are two different bases for V. One might worry that S and T have a different number of vectors; then we would have to talk about the dimension of V in terms of the basis S or in terms of the basis T. Luckily this isn't what happens. Later in this chapter, we will show that S and T must have the same number of vectors. This means that the dimension of a vector space is basis-independent. In fact, dimension is a very important characteristic of a vector space.

Example 121 $P_n(t)$ (polynomials in t of degree n or less) has a basis $\{1, t, \ldots, t^n\}$, since every vector in this space is a sum

$$a^0\,1 + a^1\,t + \cdots + a^n\,t^n, \qquad a^i \in \mathbb{R}\,,$$

so $P_n(t) = \operatorname{span}\{1, t, \ldots, t^n\}$. This set of vectors is linearly independent; If the polynomial $p(t) = c^0 1 + c^1 t + \cdots + c^n t^n = 0$, then $c^0 = c^1 = \cdots = c^n = 0$, so $p(t)$ is the zero polynomial. Thus $P_n(t)$ is finite dimensional, and $\dim P_n(t) = n + 1$.

Theorem 11.0.1. *Let $S = \{v_1, \ldots, v_n\}$ be a basis for a vector space V. Then every vector $w \in V$ can be written* uniquely *as a linear combination of vectors in the basis S:*

$$w = c^1 v_1 + \cdots + c^n v_n.$$

Proof. Since S is a basis for V, then $\operatorname{span} S = V$, and so there exist constants c^i such that $w = c^1 v_1 + \cdots + c^n v_n$.

Suppose there exists a second set of constants d^i such that

$$w = d^1 v_1 + \cdots + d^n v_n\,.$$

Then

$$\begin{aligned} 0_V &= w - w \\ &= c^1 v_1 + \cdots + c^n v_n - d^1 v_1 - \cdots - d^n v_n \\ &= (c^1 - d^1) v_1 + \cdots + (c^n - d^n) v_n. \end{aligned}$$

If it occurs exactly once that $c^i \neq d^i$, then the equation reduces to $0 = (c^i - d^i) v_i$, which is a contradiction since the vectors v_i are assumed to be non-zero.

If we have more than one i for which $c^i \neq d^i$, we can use this last equation to write one of the vectors in S as a linear combination of other vectors in S, which contradicts the assumption that S is linearly independent. Then for every i, $c^i = d^i$. □

Proof Explanation

Remark This theorem is the one that makes bases so useful–they allow us to convert abstract vectors into column vectors. By ordering the set S we obtain $B = (v_1, \ldots, v_n)$ and can write

$$w = (v_1, \ldots, v_n) \begin{pmatrix} c^1 \\ \vdots \\ c^n \end{pmatrix} = \begin{pmatrix} c^1 \\ \vdots \\ c^n \end{pmatrix}_B .$$

Remember that in general it makes no sense to drop the subscript B on the column vector on the right–most vector spaces are not made from columns of numbers!

Worked Example

Next, we would like to establish a method for determining whether a collection of vectors forms a basis for $\mathbb{R}^n$. But first, we need to show that any two bases for a finite-dimensional vector space has the same number of vectors.

Lemma 11.0.2. *If $S = \{v_1, \ldots, v_n\}$ is a basis for a vector space V and $T = \{w_1, \ldots, w_m\}$ is a linearly independent set of vectors in V, then $m \leq n$.*

The idea of the proof is to start with the set S and replace vectors in S one at a time with vectors from T, such that after each replacement we still have a basis for V.

Reading homework: problem 1

Proof. Since S spans V, then the set $\{w_1, v_1, \ldots, v_n\}$ is linearly dependent. Then we can write w_1 as a linear combination of the v_i; using that equation, we can express one of the v_i in terms of w_1 and the remaining v_j with $j \neq i$. Then we can discard one of the v_i from this set to obtain a linearly independent set that still spans V. Now we need to prove that S_1 is a basis; we must show that S_1 is linearly independent and that S_1 spans V.

The set $S_1 = \{w_1, v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n\}$ is linearly independent: By the previous theorem, there was a unique way to express w_1 in terms of the set S. Now, to obtain a contradiction, suppose there is some k and constants c^i such that

$$v_k = c^0 w_1 + c^1 v_1 + \cdots + c^{i-1} v_{i-1} + c^{i+1} v_{i+1} + \cdots + c^n v_n.$$

Then replacing w_1 with its expression in terms of the collection S gives a way to express the vector v_k as a linear combination of the vectors in S, which contradicts the linear independence of S. On the other hand, we cannot express w_1 as a linear combination of the vectors in $\{v_j | j \neq i\}$, since the expression of w_1 in terms of S was unique, and had a non-zero coefficient for the vector v_i. Then no vector in S_1 can be expressed as a combination of other vectors in S_1, which demonstrates that S_1 is linearly independent.

The set S_1 spans V: For any $u \in V$, we can express u as a linear combination of vectors in S. But we can express v_i as a linear combination of vectors in the collection S_1; rewriting v_i as such allows us to express u as a linear combination of the vectors in S_1. Thus S_1 is a basis of V with n vectors.

We can now iterate this process, replacing one of the v_i in S_1 with w_2, and so on. If $m \leq n$, this process ends with the set $S_m = \{w_1, \ldots, w_m, v_{i_1}, \ldots, v_{i_{n-m}}\}$, which is fine.

Otherwise, we have $m > n$, and the set $S_n = \{w_1, \ldots, w_n\}$ is a basis for V. But we still have some vector w_{n+1} in T that is not in S_n. Since S_n is a basis, we can write w_{n+1} as a combination of the vectors in S_n, which contradicts the linear independence of the set T. Then it must be the case that $m \leq n$, as desired. □

Worked Example

Corollary 11.0.3. *For a finite-dimensional vector space V, any two bases for V have the same number of vectors.*

Proof. Let S and T be two bases for V. Then both are linearly independent sets that span V. Suppose S has n vectors and T has m vectors. Then by the previous lemma, we have that $m \leq n$. But (exchanging the roles of S and T in application of the lemma) we also see that $n \leq m$. Then $m = n$, as desired. □

Reading homework: problem 2

11.1 Bases in $\mathbb{R}^n$.

In review question 2, chapter 10 you checked that

$$\mathbb{R}^n = \operatorname{span}\left\{ \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \ldots, \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} \right\},$$

and that this set of vectors is linearly independent. (If you didn't do that problem, check this before reading any further!) So this set of vectors is a basis for $\mathbb{R}^n$, and $\dim \mathbb{R}^n = n$. This basis is often called the *standard* or *canonical basis* for $\mathbb{R}^n$. The vector with a one in the ith position and zeros everywhere else is written e_i. (You could also view it as the function $\{1, 2, \ldots, n\} \to \mathbb{R}$ where $e_i(j) = 1$ if $i = j$ and 0 if $i \neq j$.) It points in the direction of the ith coordinate axis, and has unit length. In multivariable calculus classes, this basis is often written $\{\hat{\imath}, \hat{\jmath}, \hat{k}\}$ for $\mathbb{R}^3$.

Note that it is often convenient to order basis elements, so rather than writing a set of vectors, we would write a list. This is called an ordered basis. For example, the canonical ordered basis for $\mathbb{R}^n$ is $(e_1, e_2, \ldots, e_n)$. The possibility to reorder basis vectors is not the only way in which bases are non-unique.

Bases are not unique. While there exists a unique way to express a vector in terms of any particular basis, bases themselves are far from unique. For example, both of the sets

$$\left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\} \text{ and } \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\}$$

are bases for $\mathbb{R}^2$. Rescaling any vector in one of these sets is already enough to show that $\mathbb{R}^2$ has infinitely many bases. But even if we require that all of the basis vectors have unit length, it turns out that there are still infinitely many bases for $\mathbb{R}^2$ (see review question 3).

To see whether a set of vectors $S = \{v_1, \ldots, v_m\}$ is a basis for $\mathbb{R}^n$, we have to check that the elements are linearly independent and that they span $\mathbb{R}^n$. From the previous discussion, we also know that m must equal n, so lets assume S has n vectors. If S is linearly independent, then there is no non-trivial solution of the equation

$$0 = x^1 v_1 + \cdots + x^n v_n .$$

Let M be a matrix whose columns are the vectors v_i and X the column vector with entries x^i. Then the above equation is equivalent to requiring that there is a unique solution to

$$MX = 0 .$$

To see if S spans $\mathbb{R}^n$, we take an arbitrary vector w and solve the linear system

$$w = x^1 v_1 + \cdots + x^n v_n$$

in the unknowns x^i. For this, we need to find a unique solution for the linear system $MX = w$.

Thus, we need to show that M^{-1} exists, so that

$$X = M^{-1}w$$

is the unique solution we desire. Then we see that S is a basis for $\mathbb{R}^n$ if and only if $\det M \neq 0$.

Theorem 11.1.1. *Let $S = \{v_1, \ldots, v_m\}$ be a collection of vectors in $\mathbb{R}^n$. Let M be the matrix whose columns are the vectors in S. Then S is a basis for V if and only if m is the dimension of V and*

$$\det M \neq 0.$$

Remark Also observe that S is a basis if and only if $\mathrm{RREF}(M) = I$.

Example 122 Let

$$S = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\} \text{ and } T = \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\}.$$

Then set $M_S = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. Since $\det M_S = 1 \neq 0$, then S is a basis for $\mathbb{R}^2$.

Likewise, set $M_T = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$. Since $\det M_T = -2 \neq 0$, then T is a basis for $\mathbb{R}^2$.

11.2 Matrix of a Linear Transformation (Redux)

Not only do bases allow us to describe arbitrary vectors as column vectors, they also permit linear transformations to be expressed as matrices. This is a very powerful tool for computations, which is covered in chapter 7 and reviewed again here.

Suppose we have a linear transformation $L \colon V \to W$ and ordered input and output bases $E = (e_1, \ldots, e_n)$ and $F = (f_1, \ldots, f_m)$ for V and W respectively (of course, these need not be the standard basis–in all likelihood V is *not* $\mathbb{R}^n$). Since for each e_j, $L(e_j)$ is a vector in W, there exist unique numbers m^i_j such that

$$L(e_j) = f_1 m^1_j + \cdots + f_m m^m_j = (f_1, \ldots, f_m) \begin{pmatrix} m^1_j \\ \vdots \\ m^m_j \end{pmatrix}.$$

The number m^i_j is the ith component of $L(e_j)$ in the basis F, while the f_i are vectors (note that if α is a scalar, and v a vector, $\alpha v = v\alpha$, we have used the latter—rather uncommon—notation in the above formula). The numbers m^i_j naturally form a matrix whose jth column is the column vector displayed above. Indeed, if

$$v = e_1 v^1 + \cdots + e_n v^n \,,$$

Then

$$\begin{aligned} L(v) &= L(v^1 e_1 + v^2 e_2 + \cdots + v^n e_n) \\ &= v^1 L(e_1) + v^2 L(e_2) + \cdots + v^n L(e_n) \;=\; \sum_{j=1}^{m} L(e_j) v^j \\ &= \sum_{j=1}^{m} (f_1 m^1_j + \cdots + f_m m^m_j) v^j \;=\; \sum_{i=1}^{n} f_i \left[\sum_{j=1}^{m} M^i_j v^j \right] \\ &= \begin{pmatrix} f_1 & f_2 & \cdots & f_m \end{pmatrix} \begin{pmatrix} m^1_1 & m^1_2 & \cdots & m^1_n \\ m^2_1 & m^2_2 & & \\ \vdots & & \ddots & \vdots \\ m^m_1 & & \cdots & m^m_n \end{pmatrix} \begin{pmatrix} v^1 \\ v^2 \\ \vdots \\ v^n \end{pmatrix} \end{aligned}$$

In the column vector-basis notation this equality looks familiar:

$$L \begin{pmatrix} v^1 \\ \vdots \\ v^n \end{pmatrix}_E = \left(\begin{pmatrix} m^1_1 & \ldots & m^1_n \\ \vdots & & \vdots \\ m^m_1 & \ldots & m^m_n \end{pmatrix} \begin{pmatrix} v^1 \\ \vdots \\ v^n \end{pmatrix} \right)_F .$$

The array of numbers $M = (m^i_j)$ is called the matrix of L in the input and output bases E and F for V and W, respectively. This matrix will change if we change either of the bases. Also observe that the columns of M are computed by examining L acting on each basis vector in V expanded in the basis vectors of W.

Example 123 Let $L\colon P_1(t) \mapsto P_1(t)$, such that $L(a + bt) = (a + b)t$. Since $V =$

$P_1(t) = W$, let's choose the same ordered basis $B = (1-t, 1+t)$ for V and W.

$$\begin{aligned} L(1-t) &= (1-1)t = \ 0 \ = (1-t)\cdot 0 + (1+t)\cdot 0 = \big(1-t, 1+t\big)\begin{pmatrix}0\\0\end{pmatrix} \\ L(1+t) &= (1+1)t = 2t \ = (1-t)\cdot -1 + (1+t)\cdot 1 = \big(1-t, 1+t\big)\begin{pmatrix}-1\\1\end{pmatrix} \\ &\Rightarrow \ L\begin{pmatrix}a\\b\end{pmatrix}_B = \left(\begin{pmatrix}0 & -1\\0 & 1\end{pmatrix}\begin{pmatrix}a\\b\end{pmatrix}\right)_B . \end{aligned}$$

When the vector space is $\mathbb{R}^n$ and the standard basis is used, the problem of finding the matrix of a linear transformation will seem almost trivial. It is worthwhile working through it once in the above language though.

Example 124 Any vector in $\mathbb{R}^n$ can be written as a linear combination of the *standard (ordered) basis* $(e_1, \ldots e_n)$. The vector e_i has a one in the ith position, and zeros everywhere else. $I.e.$

$$e_1 = \begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix}, \quad e_2 = \begin{pmatrix}0\\1\\\vdots\\0\end{pmatrix}, \ldots, \quad e_n = \begin{pmatrix}0\\0\\\vdots\\1\end{pmatrix}.$$

Then to find the matrix of any linear transformation $L\colon \mathbb{R}^n \to \mathbb{R}^n$, it suffices to know what $L(e_i)$ is for every i.

For any matrix M, observe that Me_i is equal to the ith column of M. Then if the ith column of M equals $L(e_i)$ for every i, then $Mv = L(v)$ for every $v \in \mathbb{R}^n$. Then the matrix representing L in the standard basis is just the matrix whose ith column is $L(e_i)$.

For example, if

$$L\begin{pmatrix}1\\0\\0\end{pmatrix} = \begin{pmatrix}1\\4\\7\end{pmatrix}, \quad L\begin{pmatrix}0\\1\\0\end{pmatrix} = \begin{pmatrix}2\\5\\8\end{pmatrix}, \quad L\begin{pmatrix}0\\0\\1\end{pmatrix} = \begin{pmatrix}3\\6\\9\end{pmatrix},$$

then the matrix of L in the standard basis is simply

$$\begin{pmatrix}1 & 2 & 3\\4 & 5 & 6\\7 & 8 & 9\end{pmatrix}.$$

Alternatively, this information would often be presented as

$$L\begin{pmatrix}x\\y\\z\end{pmatrix}=\begin{pmatrix}x+2y+3z\\4x+5y+6z\\7x+8y+9z\end{pmatrix}.$$

You could either rewrite this as

$$L\begin{pmatrix}x\\y\\z\end{pmatrix}=\begin{pmatrix}1&2&3\\4&5&6\\7&8&9\end{pmatrix}\begin{pmatrix}x\\y\\z\end{pmatrix},$$

to immediately learn the matrix of L, or taking a more circuitous route:

$$\begin{aligned}L\begin{pmatrix}x\\y\\z\end{pmatrix}&=L\left[x\begin{pmatrix}1\\0\\0\end{pmatrix}+y\begin{pmatrix}0\\0\\1\end{pmatrix}+z\begin{pmatrix}0\\0\\1\end{pmatrix}\right]\\&=x\begin{pmatrix}1\\4\\7\end{pmatrix}+y\begin{pmatrix}2\\5\\8\end{pmatrix}+z\begin{pmatrix}3\\6\\9\end{pmatrix}=\begin{pmatrix}1&2&3\\4&5&6\\7&8&9\end{pmatrix}\begin{pmatrix}x\\y\\z\end{pmatrix}.\end{aligned}$$

11.3 Review Problems

Webwork:

Reading Problems	1,2
Basis checks	3,4
Computing column vectors	5,6

1. (a) Draw the collection of all unit vectors in $\mathbb{R}^2$.

   (b) Let $S_x=\left\{\begin{pmatrix}1\\0\end{pmatrix},x\right\}$, where x is a unit vector in $\mathbb{R}^2$. For which x is S_x a basis of $\mathbb{R}^2$?

   (c) Sketch all unit vectors in $\mathbb{R}^3$.

   (d) For which $x\in\mathbb{R}^3$ is $S_x=\left\{\begin{pmatrix}1\\0\\0\end{pmatrix},\begin{pmatrix}0\\1\\0\end{pmatrix},x\right\}$ a basis for $\mathbb{R}^3$.

   (e) Discuss the generalization of the above to $\mathbb{R}^n$.

2. Let B^n be the vector space of column vectors with bit entries $0,1$. Write down every basis for B^1 and B^2. How many bases are there for B^3? B^4? Can you make a conjecture for the number of bases for B^n?

(Hint: You can build up a basis for B^n by choosing one vector at a time, such that the vector you choose is not in the span of the previous vectors you've chosen. How many vectors are in the span of any one vector? Any two vectors? How many vectors are in the span of any k vectors, for $k \leq n$?)

Hint

3. Suppose that V is an n-dimensional vector space.

   (a) Show that any n linearly independent vectors in V form a basis.

   (Hint: Let $\{w_1, \ldots, w_m\}$ be a collection of n linearly independent vectors in V, and let $\{v_1, \ldots, v_n\}$ be a basis for V. Apply the method of Lemma 11.0.2 to these two sets of vectors.)

   (b) Show that any set of n vectors in V which span V forms a basis for V.

   (Hint: Suppose that you have a set of n vectors which span V but do not form a basis. What must be true about them? How could you get a basis from this set? Use Corollary 11.0.3 to derive a contradiction.)

4. Let $S = \{v_1, \ldots, v_n\}$ be a subset of a vector space V. Show that if every vector w in V can be expressed uniquely as a linear combination of vectors in S, then S is a basis of V. In other words: suppose that for every vector w in V, there is exactly one set of constants $c^1, \ldots, c^n$ so that $c^1 v_1 + \cdots + c^n v_n = w$. Show that this means that the set S is linearly independent and spans V. (This is the converse to theorem 11.0.1.)

5. Vectors are objects that you can add together; show that the set of all linear transformations mapping $\mathbb{R}^3 \to \mathbb{R}$ is itself a vector space. Find a basis for this vector space. Do you think your proof could be modified to work for linear transformations $\mathbb{R}^n \to \mathbb{R}$? For $\mathbb{R}^{\mathbb{N}} \to \mathbb{R}^m$? For $\mathbb{R}^{\mathbb{R}}$?

   Hint: Represent $\mathbb{R}^3$ as column vectors, and argue that a linear transformation $T \colon \mathbb{R}^3 \to \mathbb{R}$ is just a row vector.

6. Let S_n denote the vector space of all $n \times n$ symmetric matrices;

$$S_n := \{M : \mathbb{R}^n \to \mathbb{R}^n \mid M = M^T\}.$$

Let A_n denote the vector space of all $n \times n$ anti-symmetric matrices;

$$A_n = \{M : \mathbb{R}^n \to \mathbb{R}^n \mid M = -M^T\}.$$

(a) Find a basis for S_3.

(b) Find a basis for A_3.

(c) Can you find a basis for S_n? For A_n?

Hint: Describe it in terms of combinations of the matrices F^i_j which have a 1 in the i-th row and the j-th column and 0 everywhere else. Note that $\{F^i_j \mid 1 \leq i \leq r, 1 \leq j \leq k\}$ is a basis for M^r_k.

7. Give the matrix of the linear transformation L with respect to the input and output bases B and B' listed below:

(a) $L : V \to W$ where $B = (v_1, \ldots, v_n)$ is a basis for V and $B' = (L(v_1), \ldots, L(v_n))$ is a basis for W.

(b) $L : V \to V$ where $B = B' = (v_1, \ldots, v_n)$ and $L(v_i) = \lambda_i v_i$ for all $1 \leq i \leq n$.