Linear combinations of linearly independent vectors. Suppose a vector x is a linear combination of $a_1, \ldots, a_k$,

$$x = \beta_1 a_1 + \cdots + \beta_k a_k.$$

When the vectors $a_1, \ldots, a_k$ are linearly independent, the coefficients that form x are *unique*: If we also have

$$x = \gamma_1 a_1 + \cdots + \gamma_k a_k,$$

then $\beta_i = \gamma_i$ for $i = 1, \ldots, k$. This tells us that, in principle at least, we can find the coefficients that form a vector x as a linear combination of linearly independent vectors.

To see this, we subtract the two equations above to get

$$0 = (\beta_1 - \gamma_1) a_1 + \cdots + (\beta_k - \gamma_k) a_k.$$

Since $a_1, \ldots, a_k$ are linearly independent, we conclude that $\beta_i - \gamma_i$ are all zero.

The converse is also true: If each linear combination of a list of vectors can only be expressed as a linear combination with one set of coefficients, then the list of vectors is linearly independent. This gives a nice interpretation of linear independence: A list of vectors is linearly independent if and only if for any linear combination of them, we can infer or deduce the associated coefficients. (We will see later how to do this.)

Supersets and subsets. If a collection of vectors is linearly dependent, then any superset of it is linearly dependent. In other words: If we add vectors to a linearly dependent collection of vectors, the new collection is also linearly dependent. Any nonempty subset of a linearly independent collection of vectors is linearly independent. In other words: Removing vectors from a collection of vectors preserves linear independence.

5.2 Basis

Independence-dimension inequality. If the n-vectors $a_1, \ldots, a_k$ are linearly independent, then $k \leq n$. In words:

A linearly independent collection of n-vectors can have at most n elements.

Put another way:

Any collection of $n+1$ or more n-vectors is linearly dependent.

As a very simple example, we can conclude that any three 2-vectors must be linearly dependent. This is illustrated in figure 5.1.

We will prove this fundamental fact below; but first, we describe the concept of basis, which relies on the independence-dimension inequality.

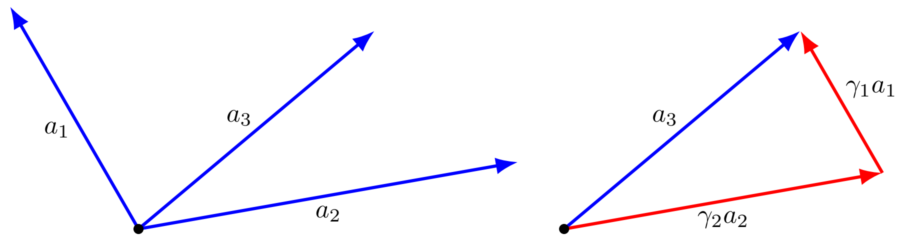

Figure 5.1 *Left.* Three 2-vectors. *Right.* The vector a_3 is a linear combination of a_1 and a_2, which shows that the vectors are linearly dependent.

Basis. A collection of n linearly independent n-vectors (*i.e.*, a collection of linearly independent vectors of the maximum possible size) is called a *basis*. If the n-vectors $a_1, \ldots, a_n$ are a basis, then any n-vector b can be written as a linear combination of them. To see this, consider the collection of $n+1$ n-vectors $a_1, \ldots, a_n, b$. By the independence-dimension inequality, these vectors are linearly dependent, so there are $\beta_1, \ldots, \beta_{n+1}$, not all zero, that satisfy

$$\beta_1 a_1 + \cdots + \beta_n a_n + \beta_{n+1} b = 0.$$

If $\beta_{n+1} = 0$, then we have

$$\beta_1 a_1 + \cdots + \beta_n a_n = 0,$$

which, since $a_1, \ldots, a_n$ are linearly independent, implies that $\beta_1 = \cdots = \beta_n = 0$. But then all the β_i are zero, a contradiction. So we conclude that $\beta_{n+1} \neq 0$. It follows that

$$b = (-\beta_1/\beta_{n+1}) a_1 + \cdots + (-\beta_n/\beta_{n+1}) a_n,$$

i.e., b is a linear combination of $a_1, \ldots, a_n$.

Combining this result with the observation above that any linear combination of linearly independent vectors can be expressed in only one way, we conclude:

> *Any n-vector b can be written in a unique way as a linear combination of a basis $a_1, \ldots, a_n$.*

Expansion in a basis. When we express an n-vector b as a linear combination of a basis $a_1, \ldots, a_n$, we refer to

$$b = \alpha_1 a_1 + \cdots + \alpha_n a_n,$$

as the *expansion of b in the $a_1, \ldots, a_n$ basis*. The numbers $\alpha_1, \ldots, \alpha_n$ are called the *coefficients* of the expansion of b in the basis $a_1, \ldots, a_n$. (We will see later how to find the coefficients in the expansion of a vector in a basis.)

Examples.

- The n standard unit n vectors $e_1, \ldots, e_n$ are a basis. Any n-vector b can be written as the linear combination

$$b = b_1 e_1 + \cdots + b_n e_n.$$

(This was already observed on page 17.) This expansion is unique, which means that there is no other linear combination of $e_1, \ldots, e_n$ that equals b.

- The vectors

$$a_1 = \begin{bmatrix} 1.2 \\ -2.6 \end{bmatrix}, \qquad a_2 = \begin{bmatrix} -0.3 \\ -3.7 \end{bmatrix}$$

are a basis. The vector $b = (1, 1)$ can be expressed in only one way as a linear combination of them:

$$b = 0.6513\, a_1 - 0.7280\, a_2.$$

(The coefficients are given here to 4 significant digits. We will see later how these coefficients can be computed.)

Cash flows and single period loans. As a practical example, we consider cash flows over n periods, with positive entries meaning income or cash in and negative entries meaning payments or cash out. We define the single-period loan cash flow vectors as

$$l_i = \begin{bmatrix} 0_{i-1} \\ 1 \\ -(1+r) \\ 0_{n-i-1} \end{bmatrix}, \quad i = 1, \ldots, n-1,$$

where $r \geq 0$ is the per-period interest rate. The cash flow l_i represents a loan of \$1 in period i, which is paid back in period $i+1$ with interest r. (The subscripts on the zero vectors above give their dimensions.) Scaling l_i changes the loan amount; scaling l_i by a negative coefficient converts it into a loan *to* another entity (which is paid back in period $i+1$ with interest).

The vectors $e_1, l_1, \ldots, l_{n-1}$ are a basis. (The first vector e_1 represents income of \$1 in period 1.) To see this, we show that they are linearly independent. Suppose that

$$\beta_1 e_1 + \beta_2 l_1 + \cdots + \beta_n l_{n-1} = 0.$$

We can express this as

$$\begin{bmatrix} \beta_1 + \beta_2 \\ \beta_3 - (1+r)\beta_2 \\ \vdots \\ \beta_n - (1+r)\beta_{n-1} \\ -(1+r)\beta_n \end{bmatrix} = 0.$$

The last entry is $-(1+r)\beta_n = 0$, which implies that $\beta_n = 0$ (since $1 + r > 0$). Using $\beta_n = 0$, the second to last entry becomes $-(1+r)\beta_{n-1} = 0$, so we conclude that $\beta_{n-1} = 0$. Continuing this way we find that $\beta_{n-2}, \ldots, \beta_2$ are all zero. The first entry of the equation above, $\beta_1 + \beta_2 = 0$, then implies $\beta_1 = 0$. We conclude that the vectors $e_1, l_1, \ldots, l_{n-1}$ are linearly independent, and therefore a basis.

This means that any cash flow n-vector c can be expressed as a linear combination of (*i.e.*, replicated by) an initial payment and one period loans:

$$c = \alpha_1 e_1 + \alpha_2 l_1 + \cdots + \alpha_n l_{n-1}.$$

It is possible to work out what the coefficients are (see exercise 5.3). The most interesting one is the first coefficient

$$\alpha_1 = c_1 + \frac{c_2}{1+r} + \cdots + \frac{c_n}{(1+r)^{n-1}},$$

which is exactly the net present value (NPV) of the cash flow, with interest rate r. Thus we see that *any* cash flow can be replicated as an income in period 1 equal to its net present value, plus a linear combination of one-period loans at interest rate r.

Proof of independence-dimension inequality. The proof is by induction on the dimension n. First consider a linearly independent collection $a_1, \ldots, a_k$ of 1-vectors. We must have $a_1 \neq 0$. This means that every element a_i of the collection can be expressed as a multiple $a_i = (a_i/a_1)a_1$ of the first element a_1. This contradicts linear independence unless $k = 1$.

Next suppose $n \geq 2$ and the independence-dimension inequality holds for dimension $n-1$. Let $a_1, \ldots, a_k$ be a linearly independent list of n-vectors. We need to show that $k \leq n$. We partition the vectors as

$$a_i = \left[\begin{array}{c} b_i \\ \alpha_i \end{array}\right], \quad i = 1, \ldots, k,$$

where b_i is an $(n-1)$-vector and α_i is a scalar.

First suppose that $\alpha_1 = \cdots = \alpha_k = 0$. Then the vectors $b_1, \ldots, b_k$ are linearly independent: $\sum_{i=1}^k \beta_i b_i = 0$ holds if and only if $\sum_{i=1}^k \beta_i a_i = 0$, which is only possible for $\beta_1 = \cdots = \beta_k = 0$ because the vectors a_i are linearly independent. The vectors $b_1, \ldots, b_k$ therefore form a linearly independent collection of $(n-1)$-vectors. By the induction hypothesis we have $k \leq n-1$, so certainly $k \leq n$.

Next suppose that the scalars α_i are not all zero. Assume $\alpha_j \neq 0$. We define a collection of $k-1$ vectors c_i of length $n-1$ as follows:

$$c_i = b_i - \frac{\alpha_i}{\alpha_j} b_j, \quad i = 1, \ldots, j-1, \qquad c_i = b_{i+1} - \frac{\alpha_{i+1}}{\alpha_j} b_j, \quad i = j, \ldots, k-1.$$

These $k-1$ vectors are linearly independent: If $\sum_{i=1}^{k-1} \beta_i c_i = 0$ then

$$\sum_{i=1}^{j-1} \beta_i \left[\begin{array}{c} b_i \\ \alpha_i \end{array}\right] + \gamma \left[\begin{array}{c} b_j \\ \alpha_j \end{array}\right] + \sum_{i=j+1}^{k} \beta_{i-1} \left[\begin{array}{c} b_i \\ \alpha_i \end{array}\right] = 0 \tag{5.2}$$

with

$$\gamma = -\frac{1}{\alpha_j}\left(\sum_{i=1}^{j-1} \beta_i \alpha_i + \sum_{i=j+1}^{k} \beta_{i-1} \alpha_i\right).$$

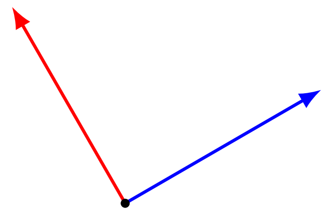

Figure 5.2 Orthonormal vectors in a plane.

Since the vectors $a_i = (b_i, \alpha_i)$ are linearly independent, the equality (5.2) only holds when all the coefficients β_i and γ are all zero. This in turns implies that the vectors $c_1, \ldots, c_{k-1}$ are linearly independent. By the induction hypothesis $k-1 \leq n-1$, so we have established that $k \leq n$.

5.3 Orthonormal vectors

A collection of vectors $a_1, \ldots, a_k$ is *orthogonal* or *mutually orthogonal* if $a_i \perp a_j$ for any i, j with $i \neq j$, $i, j = 1, \ldots, k$. A collection of vectors $a_1, \ldots, a_k$ is *orthonormal* if it is orthogonal and $\|a_i\| = 1$ for $i = 1, \ldots, k$. (A vector of norm one is called *normalized*; dividing a vector by its norm is called *normalizing* it.) Thus, each vector in an orthonormal collection of vectors is normalized, and two different vectors from the collection are orthogonal. These two conditions can be combined into one statement about the inner products of pairs of vectors in the collection: $a_1, \ldots, a_k$ is orthonormal means that

$$a_i^T a_j = \begin{cases} 1 & i = j \\ 0 & i \neq j. \end{cases}$$

Orthonormality, like linear dependence and independence, is an attribute of a collection of vectors, and not an attribute of vectors individually. By convention, though, we say "The vectors $a_1, \ldots, a_k$ are orthonormal" to mean "The collection of vectors $a_1, \ldots, a_k$ is orthonormal".

Examples. The standard unit n-vectors $e_1, \ldots, e_n$ are orthonormal. As another example, the 3-vectors

$$\begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix}, \qquad \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \qquad \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \tag{5.3}$$

are orthonormal. Figure 5.2 shows a set of two orthonormal 2-vectors.

Linear independence of orthonormal vectors. Orthonormal vectors are linearly independent. To see this, suppose $a_1, \ldots, a_k$ are orthonormal, and

$$\beta_1 a_1 + \cdots + \beta_k a_k = 0.$$