Chapter 2

Linear Algebra

Linear algebra is about linear systems of equations and their solutions. As a simple example, consider the following system of two linear equations with two unknowns:

$$\begin{aligned} 2x_1 - x_2 &= 1 \\ x_1 + x_2 &= 5. \end{aligned}$$

There is a unique solution, given by $x_1 = 2$ and $x_2 = 3$. One way to see this is by drawing a graph, as shown in Figure 2.1. One line in the figure represents the set of pairs (x_1, x_2) for which $2x_1 - x_2 = 1$, while the other represents the set of pairs for which $x_1 + x_2 = 5$. The two lines intersect only at $(2, 3)$, so this is the only point that satisfies both equations.

The above example illustrates one possible outcome with two equations and two unknowns. Two other possibilities are to have no solutions or an infinite number of solutions, as illustrated in Figure 2.2. The lines in Figure 2.2(a) are given by the equations

$$\begin{aligned} x_1 + x_2 &= 2 \\ x_1 + x_2 &= 5. \end{aligned}$$

Since they are parallel, they never intersect. The system of equations therefore does not have a solution. In Figure 2.2(b), a single line is given by both of the following equations

$$\begin{aligned} 2x_1 - x_2 &= 1 \\ -4x_1 + 2x_2 &= -2. \end{aligned}$$

Every point on this line solves the system of equations. Hence, there are an infinite number of solutions.

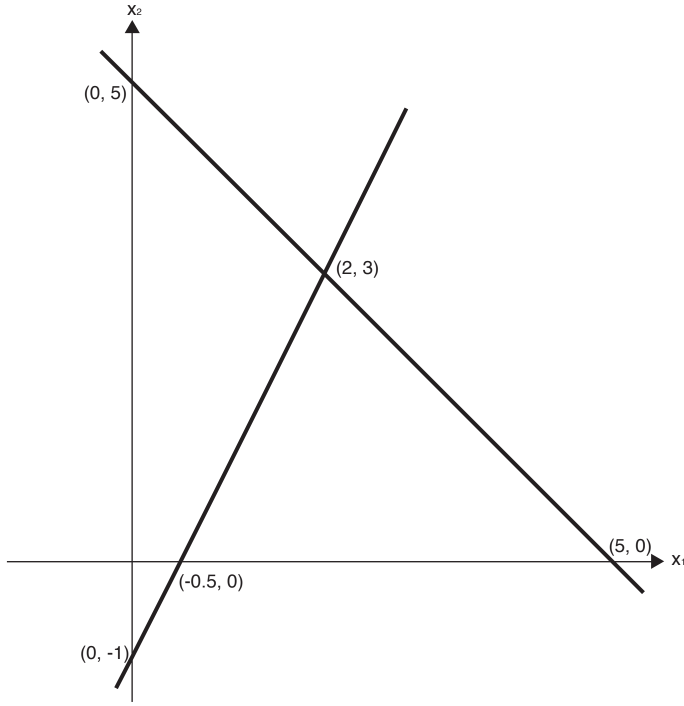


Figure 2.1: A system of two linear equations with two unknowns, with a unique solution.

There is no way to draw two straight lines on a plane so that they only intersect twice. In fact, there are only three possibilities for two lines on a plane: they never intersect, intersect once, or intersect an infinite number of times. Hence, a system of two equations with two unknowns can have either no solution, a unique solution, or an infinite number of solutions.

Linear algebra is about such systems of linear equations, but with many more equations and unknowns. Many practical problems benefit from insights offered by linear algebra. Later in this chapter, as an example of how linear algebra arises in real problems, we explore the analysis of contingent claims. Our primary motivation for studying linear algebra, however, is to develop a foundation for linear programming, which is the main topic of this book. Our coverage of linear algebra in this chapter is neither self-contained nor comprehensive. A couple of results are stated without any form of proof. The concepts presented are chosen based on their relevance to the rest of the book.

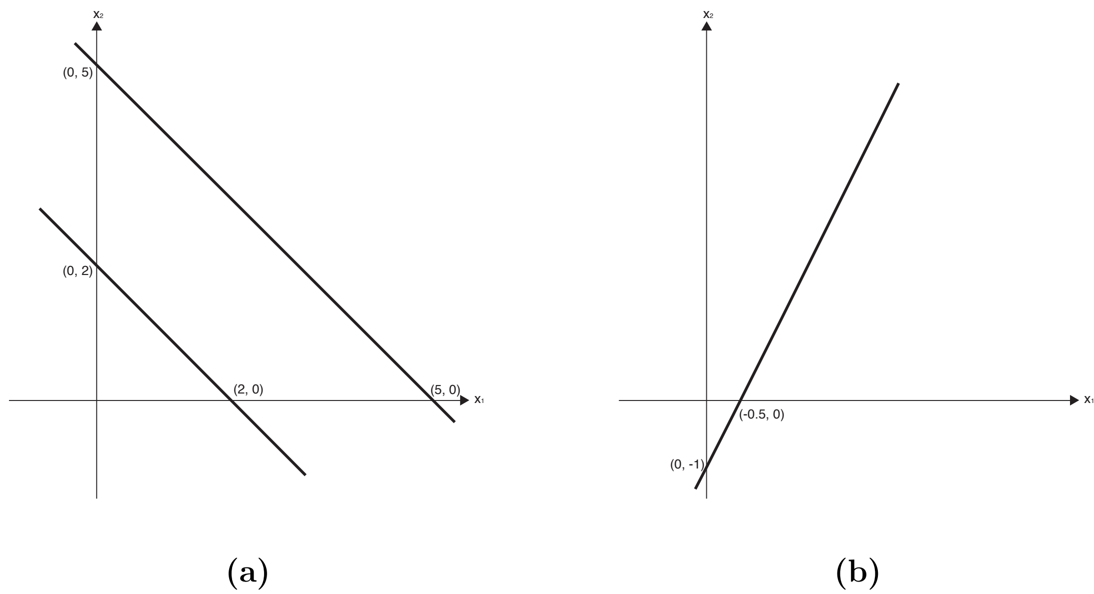


Figure 2.2: (a) A case with no solution. (b) A case with an infinite number of solutions.

2.1 Matrices

We assume that the reader is familiar with matrices and their arithmetic operations. This section provides a brief review and also serves to define notation that we will be using in the rest of the book. We will denote the set of matrices of real numbers with M rows and N columns by $\Re^{M\times N}$. Given a matrix $A \in \Re^{M\times N}$, we will use the notation A_{ij} to refer to the entry in the ith row and jth column, so that

$$A = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1N} \\ A_{21} & A_{22} & \cdots & A_{2N} \\ \vdots & \vdots & \cdots & \vdots \\ A_{M1} & A_{M2} & \cdots & A_{MN} \end{bmatrix}.$$

2.1.1 Vectors

A column vector is a matrix that has only one column, and a row vector is a matrix that only has one row. A matrix with one row and one column is a number. For shorthand, the set of column vectors with M components will be written as $\Re^M$. We will primarily employ column vectors, and as such, when we refer to a vector without specifying that it is a row vector, it should be taken to be a column vector. Given a vector $x \in \Re^M$, we denote the ith component by x_i. We will use the notation A_{i*} for the column vector consisting of entries of the ith row. Similarly, A_{*j} will be the column vector made up of the jth column of the matrix.

2.1.2 Addition

Two matrices can be added if they have the same number of rows and the same number of columns. Given $A, B \in \Re^{M\times N}$, entries of the matrix sum $A + B$ are given by the sums of entries $(A + B)_{ij} = A_{ij} + B_{ij}$. For example,

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} + \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix} = \begin{bmatrix} 6 & 8 \\ 10 & 12 \end{bmatrix}.$$

Multiplying a matrix by a scalar $\alpha \in \Re$ involves scaling each entry by α. In particular, given a scalar $\alpha \in \Re$ and a matrix $A \in \Re^{M\times N}$, we have $(\alpha A)_{ij} = (A\alpha)_{ij} = \alpha A_{ij}$. As an example,

$$2\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 2 & 4 \\ 6 & 8 \end{bmatrix}.$$

Note that matrix addition is commutative (so, $A+B = B+A$) and associative ($(A+B)+C = A+(B+C) =: A+B+C$) just like normal addition. This means we can manipulate addition-only equations with matrices in the same ways we can manipulate normal addition-only equations. We denote by 0 any matrix for which all of the elements are 0.

2.1.3 Transposition

We denote the *transpose* of a matrix A by A^T. If A is an element of $\Re^{M\times N}$, A^T is in $\Re^{N\times M}$, and each entry is given by $(A^T)_{ij} = A_{ji}$. For example,

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}^T = \begin{bmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{bmatrix}.$$

Clearly, transposition of a transpose gives the original matrix: $((A^T)^T)_{ij} = (A^T)_{ji} = A_{ij}$. Note that the transpose of a column vector is a row vector and vice-versa. More generally, the columns of a matrix become rows of its transpose and its rows become columns of the transpose. In other words, $(A^T)_{i*} = A_{*i}$ and $(A^T)_{*j} = A_{j*}$.

2.1.4 Multiplication

A row vector and a column vector can be multiplied if each has the same number of components. If $x, y \in \Re^N$, then the product $x^T y$ of the row vector x^T and column vector y is a scalar, given by $x^T y = \sum_{i=1}^N x_i y_i$. For example,

$$\begin{bmatrix} 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} = 1 \times 4 + 2 \times 5 + 3 \times 6 = 32.$$

Note that $x^T x = \sum_{i=1}^N x_i^2$ is positive unless $x = 0$ in which case it is 0. Its square root $(x^T x)^{1/2}$ is called the *norm* of x. In three-dimensional space, the norm of a vector is just its Euclidean length. For example, the norm of $[4\ 5\ 6]^T$ is $\sqrt{77}$.

More generally, two matrices A and B can be multiplied if B has exactly as many rows as A has columns. The product is a matrix $C = AB$ that has as many rows as A and columns as B. If $A \in \Re^{M \times N}$ and $B \in \Re^{N \times K}$ then the product is a matrix $C \in \Re^{M \times K}$, whose (i, j)th entry is given by the product of A_{i*}, the ith row of A and B_{*j}, the jth column of B; in other words,

$$C_{ij} = (A_{i*})^T B_{*j} = \sum_{k=1}^N A_{ik} B_{kj}.$$

For example,

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} = \begin{bmatrix} 32 \\ 77 \end{bmatrix}$$

Like scalar multiplication, matrix multiplication is associative (i.e., $(AB)C = A(BC)$) and distributive (i.e., $A(B + C) = AB + AC$ and $(A + B)C = AC + BC$), but unlike scalar multiplication, it is not commutative (i.e., AB is generally not equal to BA). The first two facts mean that much of our intuition about the multiplication of numbers is still valid when applied to matrices. The last fact means that not all of it is – the order of multiplication is important. For example,

$$\begin{bmatrix} 1 & 2 \end{bmatrix} \begin{bmatrix} 3 \\ 4 \end{bmatrix} = 11, \quad \text{but} \quad \begin{bmatrix} 3 \\ 4 \end{bmatrix} \begin{bmatrix} 1 & 2 \end{bmatrix} = \begin{bmatrix} 3 & 6 \\ 4 & 8 \end{bmatrix}.$$

A matrix is *square* if the number of rows equals the number of columns. The *identity* matrix is a square matrix with diagonal entries equal to 1 and all other entries equal to 0. We denote this matrix by I (for any number of rows/columns). Note that for any matrix A, $IA = A$ and $AI = A$, given identity matrices of appropriate dimension in each case. The identity matrix I plays a role in matrix multiplication analogous to that played by 1 in number multiplication.

A special case of matrix multiplication that arises frequently involves the second matrix being a column vector. Given a matrix $A \in \Re^{M\times N}$ and a vector $x \in \Re^N$, the product Ax is a vector in $\Re^M$. Each ith element of the product is the product of the vectors $(A_{i*})^T$ and x. Another useful way to view the product y is as a sum of columns of A, each multiplied by a component of x: $y = \sum_{j=1}^N x_j A_{*j}$.

2.1.5 Linear Systems of Equations

We will use matrix notation regularly to represent systems of linear equations. In particular, consider a system of M linear equations with N unknowns $x_1, \ldots, x_N$:

$$\begin{aligned} A_{11}x_1 + A_{12}x_2 + \cdots + A_{1N}x_N &= b_1 \\ A_{21}x_1 + A_{22}x_2 + \cdots + A_{2N}x_N &= b_2 \\ &\vdots \\ A_{M1}x_1 + A_{M2}x_2 + \cdots + A_{MN}x_N &= b_M. \end{aligned}$$

The ith equation can be rewritten as $(A_{i*})^T x = b_i$. Furthermore, the entire system can be written in a very compact form: $Ax = b$.

2.1.6 Partitioning of Matrices

It is sometimes convenient to view a matrix as being made up of smaller matrices. For example, suppose we have two matrix equations: $A^1 x = b^1$ and $A^2 x = b^2$, with $A^1 \in \Re^{M\times N}$, $A^2 \in \Re^{K\times N}$, $b^1 \in \Re^M$, and $b^2 \in \Re^K$. We can represent them as a single matrix equation if we define

$$A = \begin{bmatrix} A^1 \\ A^2 \end{bmatrix} \quad \text{and} \quad b = \begin{bmatrix} b^1 \\ b^2 \end{bmatrix}.$$

The two equations become one:

$$Ax = \begin{bmatrix} A^1 \\ A^2 \end{bmatrix} x = \begin{bmatrix} A^1 x \\ A^2 x \end{bmatrix} = \begin{bmatrix} b^1 \\ b^2 \end{bmatrix} = b. \tag{2.1}$$

Here, the first M rows of A come from A^1 and the last K rows come from A^2, so $A \in \Re^{(M+K)\times N}$. Similarly, the first M components of b come from b^1 and the last K components come from b^2, and $b \in \Re^{M+K}$. The representation of a matrix in terms of smaller matrices is called a *partition*.

We can also partition a matrix by concatenating two matrices with the same number of rows. Suppose we had two linear equations: $Ax^1 = b^1$ and $Ax^2 = b^2$. Note that the vectors x^1 and x^2 can take on different values, while the two equations share a common matrix A. We can concatenate x^1 and x^2 because they have the same number of components (they must to be multiplied by A). Similarly, we can concatenate b^1 and b^2. If we write $X = [x^1\ x^2]$ and $B = [b^1\ b^2]$ then we again have $AX = B$. Note that now X and B are not vectors; they each have 2 columns. Writing out the matrices in terms of partitions, we have

$$A \begin{bmatrix} x^1 & x^2 \end{bmatrix} = \begin{bmatrix} b^1 & b^2 \end{bmatrix}. \tag{2.2}$$

Given the same equations, we could alternatively combine them to form a single equation

$$\begin{bmatrix} A & 0 \\ 0 & A \end{bmatrix} \begin{bmatrix} x^1 \\ x^2 \end{bmatrix} = \begin{bmatrix} Ax^1 + 0x^2 \\ 0x^1 + Ax^2 \end{bmatrix} = \begin{bmatrix} Ax^1 \\ Ax^2 \end{bmatrix} = \begin{bmatrix} b^1 \\ b^2 \end{bmatrix},$$

where the 0 matrix has as many rows and columns as A.

Note that, for any partition, the numbers of rows and columns of component matrices must be compatible. For example, in Equation (2.1), the matrices A^1 and A^2 had to have the same number of columns in order for the definition of A to make sense. Similarly, in Equation (2.2), x^1 and x^2 must have the same number of components, as do b^1 and b^2. More generally, we can partition a matrix by separating rows, columns, or both. All that is required is that each component matrix has the same number of columns as any other component matrix it adjoins vertically, and the same number of rows as any other component matrix it adjoins horizontally.

2.2 Vector Spaces

The study of vector spaces comprises a wealth of ideas. There are a number of complexities that arise when one deals with *infinite-dimensional* vector spaces, and we will not address them. Any vector space we refer to in this book is implicitly assumed to be *finite-dimensional*.

Given a collection of vectors $a^1, \ldots, a^N \in \Re^M$, the term *linear combination* refers to a sum of the form

$$\sum_{j=1}^{N} x_j a^j,$$

for some real-valued coefficients $x_1, \ldots, x_N$. A finite-dimensional vector space is the set S of linear combinations of a prescribed collection of vectors. In particular, a collection of vectors $a^1, \ldots, a^N \in \Re^M$ generates a vector space

$$S = \left\{ \sum_{j=1}^{N} x_j a^j \;\middle|\; x \in \Re^M \right\}.$$

This vector space is referred to as the *span* of the vectors $a^1, \ldots, a^N$ used to generate it. As a convention, we will take the span of an empty set of vectors to be the 0-vector; i.e., the vector with all components equal to 0.

If a vector x is in the span of $a^1, \ldots, a^N$, we say x is *linearly dependent* on $a^1, \ldots, a^N$. On the other hand, if x is not in the span of $a^1, \ldots, a^N$, we say x is linearly independent of $a^1, \ldots, a^N$. A collection of vectors, $a^1, \ldots, a^N$ is called *linearly independent* if none of the vectors in the collection is linearly dependent on the others. Note that this excludes the 0-vector from any set of linearly independent vectors. The following lemma provides an equivalent definition of linear independence.

Lemma 2.2.1. *A collection of vectors $a^1, \ldots, a^N$ is linearly independent if $\sum_{j=1}^N \alpha_j a^j = 0$ implies that $\alpha_1 = \alpha_2 = \cdots = \alpha_N = 0$.*

One example of a vector space is the space $\Re^M$ itself. This vector space is generated by a collection of M vectors $e^1, \ldots, e^M \in \Re^M$. Each e^i denotes the M-dimensional vector with all components equal to 0, except for the ith component, which is equal to 1. Clearly, any element of $\Re^M$ is in the span of $e^1, \ldots, e^M$. In particular, for any $y \in \Re^M$,

$$y = \sum_{i=1}^{M} y_i e^i,$$

so it is a linear combination of $e^1, \ldots, e^M$.

More interesting examples of vector spaces involve nontrivial subsets of $\Re^M$. Consider the case of $M = 2$. Figure 2.3 illustrates the vector space generated by the vector $a^1 = [1\ 2]^T$. Suppose we are given another vector a^2. If a^2 is a multiple of a^1 (i.e., $a^2 = \alpha a^1$ for number α), the vector space spanned by the two vectors is no different from that spanned by a^1 alone. If a^2 is not a multiple of a^1 then any vector in $y \in \Re^2$ can be written as a linear combination

$$y = x_1 a^1 + x_2 a^2,$$

and therefore, the two vectors span $\Re^2$. In this case, incorporating a third vector a^3 cannot make any difference to the span, since the first two vectors already span the entire space.

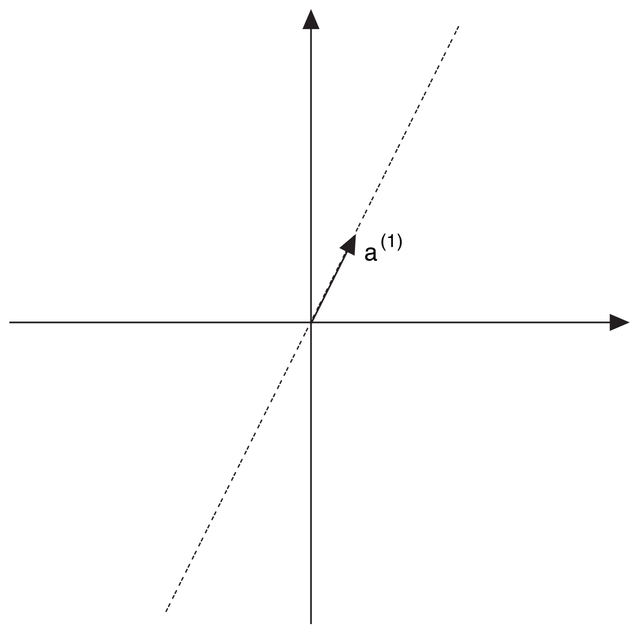


Figure 2.3: The vector space spanned by $a^1 = [1\ 2]^T$.

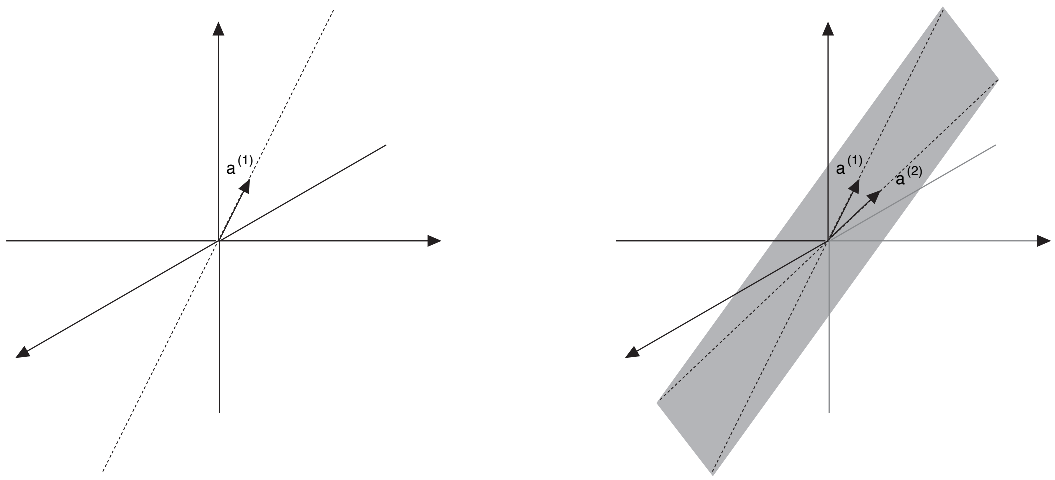
The range of possibilities increases when $M = 3$. As illustrated in Figure 2.4(a), a single vector spans a line. If a second vector is not a multiple of the first, the two span a plane, as shown in 2.4(b). If a third vector is a linear combination of the first two, it does not increase the span. Otherwise, the three vectors together span all of $\Re^3$.

Sometimes one vector space is a subset of another. When this is the case, the former is said to be a *subspace* of the latter. For example, the vector space of Figure 2.4(a) is a subspace of the one in Figure 2.4(b). Both of these vector spaces are subspaces of the vector space $\Re^3$. Any vector space is a subset and therefore a subspace of itself.

2.2.1 Bases and Dimension

For any given vector space S, there are many collections of vectors that span S. Of particular importance are those with the special property that they are also linearly independent collections. Such a collection is called a *basis*.

To find a basis for a space S, one can take any spanning set, and repeatedly remove vectors that are linear combinations of the remaining vectors.



(a) **(b)**

Figure 2.4: (a) The vector space spanned by a single vector $a^1 \in \Re^3$. (b) The vector space spanned by a^1 and $a^2 \in \Re^3$ where a^2 is not a multiple of a^1.

At each stage the set of vectors remaining will span S, and the process will conclude when the set of vectors remaining is linearly independent. That is, with a basis.

Starting with different spanning sets will result in different bases. They all however, have the following property, which is important enough to state as a theorem.

Theorem 2.2.1. *Any two bases of a vector space S have the same number of vectors.*

We prove this theorem after we establish the following helpful lemma:

Lemma 2.2.2. *If $A = \{a^1, a^2, \ldots, a^N\}$ is a basis for S, and if $b = \alpha_1 a^1 + \alpha_2 a^2 + \ldots + \alpha_N a^N$ with $\alpha_1 \neq 0$, then $\{b, a^2, \ldots, a^N\}$ is a basis for S.*

Proof of lemma: We need to show that $\{b, a^2, \ldots, a^N\}$ is linearly independent and spans S. Note that $a^1 = \frac{1}{\alpha_1} b - \frac{\alpha_2}{\alpha_1} a^2 - \ldots + \frac{\alpha_N a^N}{\alpha_1}$.

To show linear independence, note that the a^i's are linearly independent, so the only possibility for linear dependence is if b is a linear combination of $a^2, \ldots, a^N$. But a_1 is a linear combination of $b, a^2, \ldots, a^N$, so this would make a^1 a linear combination of $a^2, \ldots, a^N$, which contradicts the fact that

$a^1, \ldots, a^N$ are linearly independent. Thus b must be linearly independent of $a^2, \ldots, a^N$, and so the set is linearly independent.

To show that the set spans S, we just need to show that it spans a basis of S. It obviously spans $a^2, \ldots, a^n$, and because a^1 is a linear combination of $b, a^2, \ldots, a^n$ we know the set spans S. $\square$

Proof of theorem: Suppose $A = \{a^1, a^2, \ldots, a^N\}$ and $B = \{b^1, b^2, \ldots, b^M\}$ are two bases for S with $N > M$. Because A is a basis, we know that $b^1 = \alpha_1 a^1 + \alpha_2 a^2 + \ldots + \alpha_N a^N$ for some $\alpha_1, \alpha_2, \ldots, \alpha_N$ where not all the α_i's are equal to 0. Assume without loss of generality that $\alpha_1 \neq 0$. Then, by Lemma 2.2.2, $A^1 = \{b^1, a^2, \ldots, a^N\}$ is a basis for S.

Now, because A^1 is a basis for S, we know that $b^2 = \beta_1 b^1 + \alpha_2 a^2 + \ldots + \alpha_N a^N$ for some $\beta_1, \alpha_2, \ldots, \alpha_N$ where we know that not all of $\alpha_2, \ldots, \alpha_N$ are equal to 0 (otherwise b^2 would be linearly dependent on b^1, and so B would not be a basis). Note the α_i's here are not the same as those for b^1 in terms of $a^1, a^2, \ldots, a^N$. Assume that $\alpha_2 \neq 0$. The lemma now says that $A^2 = \{b^1, b^2, a^3 \ldots, a^N\}$ is a basis for S.

We continue in this manner, substituting b^i into A^{i-1} and using the lemma to show that A^i is a basis for S, until we arrive at $A^M = \{b^1, \ldots, b^M, a^{M+1}, \ldots, a^N\}$ this cannot be a basis for S because a^{M+1} is a linear combination of $\{b^1, \ldots, b^M\}$. This contradiction means that N cannot be greater than M. Thus any two bases must have the same number of vectors. $\square$

Let us motivate Theorem 2.2.1 with an example. Consider a plane in three dimensions that cuts through the origin. It is easy to see that such a plane is spanned by any two vectors that it contains, so long as they are linearly independent. This means that any additional vectors cannot be linearly independent. Since the first two were arbitrary linearly independent vectors, this indicates that any basis of the plane has two vectors.

Because the size of a basis depends only on S and not on the choice of basis, it is a fundamental property of S. It is called the dimension of S. This definition corresponds perfectly with the standard use of the word dimension. For example, $\dim(\Re^M) = M$ because $e^1, \ldots, e^M$ is a basis for $\Re^M$.

Theorem 2.2.1 states that any basis of a vector space S has $\dim(S)$ vectors. However, one might ask whether any collection of $\dim(S)$ linearly independent vectors in S spans the vector space. The answer is provided by the following theorem:

Theorem 2.2.2. *For any vector space $S \subseteq \Re^M$, each linearly independent collection of vectors $a^1, \ldots, a^{\dim(S)} \in S$ is a basis of S.*

Proof: Let $b^1, \ldots, b^K$ be a collection of vectors that span S. Certainly the set

$\{a^1, \ldots, a^{\dim(S)}, b^1, \ldots, b^K\}$ spans S. Consider repeatedly removing vectors b^i that are linearly dependent on remaining vectors in the set. This will leave us with a linearly independent collection of vectors comprised of $a^1, \ldots, a^{\dim(S)}$ plus any remaining b^i's. This set of vectors still spans S, and therefore, they form a basis of S. From Theorem 2.2.1 this means it has $\dim(S)$ vectors. It follows that the remaining collection of vectors cannot include any b^i's. Hence, $a^1, \ldots, a^{\dim(S)}$ is a basis for S. $\square$

We began this section describing how a basis can be constructed by repeatedly removing vectors from a collection B that spans the vector space. Let us close describing how a basis can be constructed by repeatedly appending vectors to a collection. We start with a collection of vectors in a vector space S. If the collection does not span S, we append an element of $x \in S$ that is not a linear combination of vectors already in the collection. Each vector appended maintains the linear independence of the collection. By Theorem 2.2.2, the collection spans S once it includes $\dim(S)$ vectors. Hence, we end up with a basis.

Note that we could apply the procedure we just described even if S were not a vector space but rather an arbitrary subset of $\Re^M$. Since there can not be more than M linearly independent vectors in S, the procedure must terminate with a collection of no more than M vectors. The set S would be a subset of the span of these vectors. In the event that S is equal to the span it is a vector space, otherwise it is not. This observation leads to the following theorem:

Theorem 2.2.3. *A set $S \in \Re^M$ is a vector space if and only if any linear combination of vectors in S is in S.*

Note that this theorem provides an alternative definition of a *vector space*.

2.2.2 Orthogonality

Two vectors $x, y \in \Re^M$ are said to be *orthogonal* (or *perpendicular*) if $x^T y = 0$. Figure 2.5 presents four vectors in $\Re^2$. The vectors $[1\ 0]^T$ and $[0\ 1]^T$ are orthogonal, as are $[1\ 1]^T$ and $[-1\ 1]^T$. On the other hand, $[1\ 0]^T$ and $[1\ 1]^T$ are not orthogonal.

Any collection of nonzero vectors $a^1, \ldots, a^N$ that are orthogonal to one another are linearly independent. To establish this, suppose that they are orthogonal, and that

$$\sum_{j=1}^{N} \alpha_j a^j = 0.$$

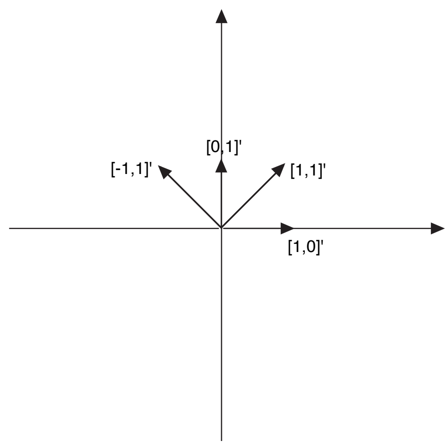


Figure 2.5: The vectors $[1\ 0]^T$ and $[0\ 1]^T$ are orthogonal, as are $[1\ 1]^T$ and $[-1\ 1]^T$. On the other hand, $[1\ 0]^T$ and $[1\ 1]^T$ are not orthogonal.

Multiplying both sides by $(a^k)^T$, we obtain

$$0 = (a^k)^T \sum_{j=1}^{N} \alpha_j a^j = \alpha_k (a^k)^T a^k.$$

Because $(a^k)^T a^k > 0$, it must be that $\alpha_k = 0$. This is true for all k, and so the a^i's are linearly independent. An *orthogonal basis* is one in which all vectors in the basis are orthogonal to one another.

2.2.3 Orthogonal Subspaces

Two subspaces S and T of $\Re^M$ are said to be orthogonal if $x^T y = 0$ for all $x \in S$ and $y \in T$. To establish that two subspaces S and T of $\Re^M$ are orthogonal, it suffices to show that a collection of vectors $u^1, \ldots, u^N$ that span S are orthogonal to a collection $v^1, \ldots, v^K$ that span T. To see why, consider arbitrary vectors $x \in S$ and $y \in T$. For some numbers $c_1, \ldots, c_N$ and $d_1, \ldots, d_K$, we have

$$x^T y = \left(\sum_{j=1}^{N} c_j u^j\right)^T \left(\sum_{k=1}^{K} d_k v^k\right) = \sum_{j=1}^{N} \sum_{k=1}^{K} c_j d_k (u^j)^T v^k = 0,$$

since each u^j is orthogonal to each v^k. We capture this in terms of a theorem.

Theorem 2.2.4. *Let S and T be the subspaces spanned by $u^1, \ldots, u^N \in \Re^M$ and $v^1, \ldots, v^K \in \Re^M$, respectively. S and T are orthogonal subspaces if and only if $(u^i)^T v^j = 0$ for all $i \in \{1, \ldots, N\}$ and $j \in \{1, \ldots, K\}$.*

The notion of orthogonality offers an alternative way to characterize a subspace. Given a vector space S, we can define its *orthogonal complement*, which is the set $S^\perp \subseteq \Re^M$ of vectors that are orthogonal to all vectors in S. But is $S^\perp$ a vector space? We will now establish that it is.

From Theorem 2.2.3, it suffices to show that any linear combination of M vectors in $S^\perp$ is in $S^\perp$. Suppose $t^1, \ldots t^M$ are vectors in $S^\perp$ and that $x = \sum_{i=1}^M \alpha_i t^i$ for some numbers $\alpha_1, \ldots, \alpha_M$. Then, for any $s \in S$

$$s^T x = \sum \alpha_i s^T t^i = \sum \alpha_i \cdot 0 = 0.$$

This establishes the following theorem.

Theorem 2.2.5. *The orthogonal complement $S^\perp \subseteq \Re^M$ of a vector space $S \subseteq \Re^M$ is also a vector space.*

The following theorem establishes an interesting property of orthogonal subspaces.

Theorem 2.2.6. *The dimensions of a vector space $S \subseteq \Re^M$ and its orthogonal complement $S^\perp \subseteq \Re^M$ satisfy $\dim(S) + \dim(S^\perp) = M$.*

Proof: Suppose $s^1, \ldots, s^N$ form a basis for S, and $t^1, \ldots, t^L$ form a basis for $S^\perp$. It is certainly the case that $N + L \leq M$ because there can not be more than M linearly independent vectors in $\Re^M$.

We will assume that $N + L < M$, and derive a contradiction. If $N + L < M$, there must be a vector $x \in \Re^M$ that is not a linear combination of vectors in S and $S^\perp$. Let $\alpha_1^*, \ldots, \alpha_N^*$ be a set of numbers that minimizes the function

$$f(\alpha_1, \ldots, \alpha_N) = (x - \sum_{i=1}^N \alpha_i s^i)^T (x - \sum_{i=1}^N \alpha_i s^i).$$

Such numbers exist because the function is quadratic with a nonnegative range. Let $y = x - \sum_{i=1}^N \alpha_i^* s^i$. Note that $y \notin S^\perp$, because if $y \in S^\perp$ then x would be a linear combination of elements of S and $S^\perp$. It follows that there is a vector $u \in S$ such that $y^T u \neq 0$. Let $\beta = -\frac{u^T y}{u^T u}$. Then,

$$\begin{aligned}(y + \beta u)^T (y + \beta u) &= y^T y + 2\beta u^T y + \beta^2 u^T u \\ &= y^T y - 2\frac{u^T y}{u^T u} u^T y + \left(-\frac{u^T y}{u^T u}\right)^2 u^T u\end{aligned}$$

$$
\begin{aligned}
&= y^T y - 2\frac{(u^T y)^2}{u^T u} + \frac{(u^T y)^2}{u^T u} \\
&= y^T y - \frac{(u^T y)^2}{u^T u} \\
&< y^T y
\end{aligned}
$$

which contradicts that fact that $y^T y$ is the minimum of $f(\alpha_1, \ldots, \alpha_N)$. It follows that $N + L = M$. $\square$

2.2.4 Vector Spaces Associated with Matrices

Given a matrix $A \in \Re^{M \times N}$, there are several subspaces of $\Re^M$ and $\Re^N$ worth studying. An understanding of these subspaces facilitates intuition about linear systems of equations. The first of these is the subspace of $\Re^M$ spanned by the columns of A. This subspace is called the *column space* of A, and we denote it by $\mathcal{C}(A)$. Note that, for any vector $x \in \Re^N$, the product Ax is in $\mathcal{C}(A)$, since it is a linear combination of columns of A:

$$Ax = \sum_{j=1}^{N} x_j A_{*j}.$$

The converse is also true. If $b \in \Re^M$ is in $\mathcal{C}(A)$ then there is a vector $x \in \Re^N$ such that $Ax = b$.

The *row space* of A is the subspace of $\Re^N$ spanned by the rows of A. It is equivalent to the column space $\mathcal{C}(A^T)$ of A^T. If a vector $c \in \Re^N$ is in the row space of A then there is a vector $y \in \Re^M$ such that $A^T y = c$ or, written in another way, $y^T A = c^T$.

Another interesting subspace of $\Re^N$ is the *null space*. This is the set of vectors $x \in \Re^N$ such that $Ax = 0$. Note that the null space is the set of vectors that are orthogonal to every row of A. Hence, by Theorem 2.2.4, the null space is the orthogonal complement of the row space. We denote the null space of a matrix A by $\mathcal{N}(A)$.

The *left null space* is analogously defined as the set of vectors $y \in \Re^M$ such that $y^T A = 0$, or written in another way, $A^T y = 0$. Again, by Theorem 2.2.4, the left null space is the set of vectors that are orthogonal to every column of A and is equivalent to $\mathcal{N}(A^T)$, the null space of A^T. The following theorem summarizes our observations relating null spaces to row and column spaces.

Theorem 2.2.7. *The left null space is the orthogonal complement of the column space. The null space is the orthogonal complement of the row space.*

Applying Theorems 2.2.6 and 2.2.7, we deduce the following relationships among the dimensionality of spaces associated with a matrix.

Theorem 2.2.8. *For any matrix* $A \in \Re^{M \times N}$, $\dim(\mathcal{C}(A)) + \dim(\mathcal{N}(A^T)) = M$ *and* $\dim(\mathcal{C}(A^T)) + \dim(\mathcal{N}(A)) = N$.

Recall from Section 2.2.1 that a basis for a vector space can be obtained by starting with a collection of vectors that span the space and repeatedly removing vectors that are linear combinations of others until all remaining columns are linearly independent. When the dimension $\dim(\mathcal{C}(A))$ of the column space of a matrix $A \in \Re^{M \times N}$ is less than N, the columns are linearly dependent. We can therefore repeatedly remove columns of a matrix A until we arrive at a matrix $B \in \Re^{M \times K}$ with K linearly independent columns that span $\mathcal{C}(A)$. Hence, $\dim(\mathcal{C}(A)) = \dim(\mathcal{C}(B)) = K \leq N$. Perhaps more surprising is the fact that the process of removing columns that are linear combinations of others does not change the dimension of the row space. In particular, $\dim(\mathcal{C}(B^T)) = \dim(\mathcal{C}(A^T))$, as we will now establish.

Theorem 2.2.9. *Let A be a matrix. If we remove a column of A that is a linear combination of the remaining columns, the row space of the resulting matrix has the same dimension as the row space of A.*

Proof: Suppose that the last column of A is a linear combination of the other columns. Partition A as follows:

$$A = \left[\begin{array}{cc} B & c \\ r^T & \alpha \end{array} \right]$$

where $B \in \Re^{(M-1)\times(N-1)}, c \in \Re^{M-1}, r^T \in \Re^{N-1}$, and $\alpha \in \Re$. Note that $[c^T \ \alpha]^T = A_{*N}^T$ and $[r^T \ \alpha] = A_{M*}$.

Because the last column of A is a linear combination of the other columns, there is a vector $x \in \Re^{N-1}$ such that

$$\left[\begin{array}{c} c \\ \alpha \end{array} \right] = \left[\begin{array}{c} B \\ r^T \end{array} \right] x.$$

In other words, $\alpha = r^T x$ and $c = Bx$.

We will show that the last row of A is a linear combination of the other rows of A if and only if the last row of

$$\left[\begin{array}{c} B \\ r^T \end{array} \right]$$

is a linear combination of the previous rows. That is, if and only if r^T is a linear combination of the rows of B.

Suppose the last row of A is a linear combination of the previous rows of A. Then $[r^T \ \alpha] = y^T[B \ c]$ for some $y \in \Re^{M-1}$. This means $r^T = y^T B$, and therefore, that r^T is a linear combination of the rows of B.

Conversely, suppose r^T is a linear combination of the rows of B. This means for some $y \in \Re^{M-1}$ that $r^T = y^T B$. But then $\alpha = r^T x = y^T Bx = y^T c$. Thus $[r^T \ \alpha] = y^T[B \ c]$, and therefore, the last row of A is a linear combination of the previous rows of A.

The same argument can be applied to any column that is a linear combination of other columns and any row. It follows that a row is linearly independent of other rows after removing a column that is linearly dependent on other columns if and only if it was linearly independent of other rows prior to removing the column. This implies the result that we set out to prove. $\square$

The procedure for removing columns can also be applied to eliminating rows until all rows are linearly independent. This will change neither the row space nor the dimension of the column space. If we apply the procedure to columns and then to rows, the resulting matrix will have linearly independent rows and linearly independent columns.

Suppose we start with a matrix $A \in \Re^{M\times N}$ and eliminate columns and rows until we arrive at a matrix $C \in \Re^{K\times L}$ with linearly independent columns and linearly independent rows. Because the columns are in $\Re^K$ and linearly independent, there can be no more than K of them. This means that $L \leq K$. Similarly, since the rows are in $\Re^L$ and linearly independent, there can be no more than L of them. Therefore, $K \leq L$. This gives us the following theorem.

Theorem 2.2.10. *For any matrix A,* $\dim(\mathcal{C}(A)) = \dim(\mathcal{C}(A^T))$.

The dimension of the column space, or equivalently, the row space, of $A \in \Re^{M\times N}$ is referred to as the *rank* of A. We say that a matrix has *full column rank* if $\dim(\mathcal{C}(A)) = N$ and *full row rank* if $\dim(\mathcal{C}(A)) = M$. A matrix is said to have *full rank* if it has either *full column rank* or *full row rank*.

2.3 Linear Systems of Equations

Not all lines on a plane go through the origin. In fact, most do not. They are however all translations of lines that do go through the origin. Using the terminology of linear algebra we say that lines going through the origin are subspaces, while the translated lines are affine subspaces. An *affine subspace* is a translated subspace. That is, an affine subspace $D \subseteq \Re^N$ is described by a subspace $S \subseteq \Re^N$ and a vector $d \in \Re^N$, with $D = \{x|x = d + s, s \in S\}$. We define $\dim(D) = \dim(S)$.

Note that if $d \in S$ then $D = S$. An example of this would be translating a line in the direction of the line. On the other hand, if $d \notin S$ then $D \neq S$; in fact, the intersection of D and S is empty in this case. There are also a number of different vectors d that can be used to describe a given affine subspace D. Also, the difference between any two vectors in D is in S, in fact $S = \{x - y|x, y \in D\}$.

Of particular interest to us is the case when the dimension of $D \subseteq \Re^N$ is $N - 1$, in which case D is called a *hyperplane*. If $N = 2$, a hyperplane D is a line. If $N = 3$, a hyperplane D is a plane in a three-dimensional space.

Hyperplanes arise in several situations of interest to us. For example, the set of solutions to a single linear equation $a^T x = b$, where $a, x \in \Re^N$, $b \in \Re$, and $a \neq 0$, is a hyperplane. To see why this set should be a hyperplane, consider first the equation $a^T x = 0$. The set of solutions is the set of vectors orthogonal to a, which is an $(N-1)$-dimensional subspace. Let d be a particular solution to $a^T x = b$. Then, given any vector $\overline{x}$ that is orthogonal to a, its translation $\overline{x} + d$ satisfies $a^T(\overline{x} + d) = b$. The converse is also true: if $\overline{x}$ solves $a^T \overline{x} = b$ then $a^T(\overline{x} - d) = 0$. It follows that the solutions of $a^T x = b$ comprise a hyperplane.

Consider the matrix equation $Ax = b$ where x is unknown. One way to think of this is by partitioning A row by row to obtain

$$Ax = \begin{bmatrix} (A_{1*})^T \\ (A_{2*})^T \\ \vdots \\ (A_{M*})^T \end{bmatrix} x = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_M \end{bmatrix}$$

or,

$$Ax = \begin{bmatrix} (A_{1*})^T x \\ (A_{2*})^T x \\ \vdots \\ (A_{M*})^T x \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_M \end{bmatrix}$$

What this means is that $(A_{1*})^T x = b_1$, $(A_{2*})^T x = b_2$, and so on. Thus $Ax = b$, can be interpreted as M equations of the form $A_{i*}x = b_i$. Each of these equations describes a hyperplane, and for x to satisfy the equation, x must lie in the hyperplane. For x to satisfy all of the equations simultaneously, and hence satisfy $Ax = b$, x must be in the intersection of the M hyperplanes defined by the rows of $Ax = b$.

Another way to think about the equation $Ax = b$ involves partitioning A column by column. We then have

$$\begin{bmatrix} A_{*1} & A_{*2} & \dots & A_{*N} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix} = b$$

or $A_{*1}x_1 + A_{*2}x_2 + \ldots + A_{*N}x_N = b$. This means b is a linear combination of the columns, for which each x_j is the coefficient of the jth column A_{*j}. Geometrically, the columns of A are vectors in $\Re^M$ and we are trying to express b as a linear combination of them.

The two interpretations we have described for the equation $Ax = b$ are both helpful. Concepts we will study are sometimes more easily visualized with one interpretation or the other.

2.3.1 Solution Sets

In this section we will study some properties of solutions to $Ax = b$. Suppose we have two solutions x and y. Then $A(x - y) = Ax - Ay = b - b = 0$ so that $x - y$ is in the null space of A. The converse is also true, suppose x is a solution, and z is in the null space of A. Then $A(x+z) = Ax+Az = b+0 = b$. This means the set of all solutions is a translation of the null space of A. Because the null space of A is a subspace of $\Re^N$, this means the following:

Theorem 2.3.1. *For any $A \in \Re^{M\times N}$ and $b \in \Re^M$, the set S of solutions to $Ax = b$ is an affine subspace of $\Re^N$, with* $\dim(S) = \dim(\mathcal{N}(A))$.

We make some further observations. If the rank of A is M, the columns of A span $\Re^M$. This means that for any b, there will be at least one solution to $Ax = b$. Now, suppose the rank of A is N. Then, by Theorem 2.2.8, $\dim(\mathcal{N}(A)) = 0$. It follows from Theorem 2.3.1 that there can not be more than one solution to $Ax = b$. We summarize these observations in the following theorem.

Theorem 2.3.2. *For any $A \in \Re^{M\times N}$ and $b \in \Re^M$, the set S of solutions to $Ax = b$ satisfies*

1. $|S| \geq 1$ *if* $\dim(\mathcal{C}(A)) = M$;

2. $|S| \leq 1$ *if* $\dim(\mathcal{C}(A)) = N$.

Given a matrix $A \in \Re^{M \times N}$, we can define a function $f(x) = Ax$, mapping $\Re^N$ to $\Re^M$. The range of this function is the column space of A. If we restrict the domain to the row space of A, the function is a one-to-one mapping from row space to column space, as captured by the following theorem.

Theorem 2.3.3. *For any matrix $A \in \Re^{M \times N}$, the function $f(x) = Ax$ is a one-to-one mapping from the row space of A to the column space of A.*

Proof: Since the columns of A span the column space, the range of f is the column space. Furthermore, each vector in the column space is given by $f(x) \in \Re^M$ for some $x \in \mathcal{C}(A^T)$. This is because any such $y \in \Re^N$ can be written as $y = x + z$ for some $x \in \mathcal{C}(A^T)$ and $z \in \mathcal{N}(A)$, so $f(y) = A(x + z) = Ax = f(x)$. To complete the proof, we need to establish that only one element of the row space can map to any particular element of the column space.

Suppose that for two vectors $x^1, x^2 \in \mathcal{C}(A^T)$, both in the row space of A, we have $Ax^1 = Ax^2$. Because x^1 and x^2 are in the row space of A, there are vectors $y^1, y^2 \in \Re^M$ such that $x^1 = A^T y^1$ and $x^2 = A^T y^2$. It follows that $AA^T y^1 = AA^T y^2$, or $AA^T(y^1 - y^2) = 0$. This implies that $(y^1 - y^2)^T AA^T(y^1 - y^2) = 0$. This is the square of the norm of $A^T(y^1 - y^2)$, and the fact that it is equal to zero implies that $A^T(y^1 - y^2) = 0$. Hence, $x^1 - x^2 = 0$, or $x^1 = x^2$. The result follows $\square$

2.3.2 Matrix Inversion

Given a matrix $A \in \Re^{M \times N}$, a matrix $B \in \Re^{N \times M}$ is said to be a *left inverse* of A if $BA = I$. Analogously, a matrix $B \in \Re^{N \times M}$ is said to be a *right inverse* of A if $AB = I$. A matrix cannot have both a left inverse and a right inverse unless both are equal. To see this, suppose that the left inverse of A is B and that the right inverse is C. We would then have

$$B = B(AC) = (BA)C = C.$$

Note that if $N \neq M$, A cannot have both a left inverse and a right inverse. This is because the left inverse is in $\Re^{N \times M}$, while the right inverse is in $\Re^{M \times N}$, so they cannot possibly be equal. Hence, only square matrices can have both left and right inverses. If A is a square matrix, any left inverse must be equal to any right inverse. So when a left and right inverse exists, there is a single matrix that is simultaneously the unique left inverse and the unique right inverse. This matrix is simply referred to as the *inverse*, and denoted by A^{-1}.

How do we find inverses of a matrix $A \in \Re^{M \times N}$? To find a right inverse B, so that $AB = I$, we can partition the matrix B into columns $[B_{*1} \ \cdots \ B_{*M}]$ and I into columns $[e^1 \ \cdots \ e^M]$, and solve $AB_{*j} = e^j$ to find each jth column of B. Therefore, the matrix A has a right inverse if and only if each e^j is in its column space. Since the vectors $e^1, \ldots, e^N$ span $\Re^M$, the column space of A must be $\Re^M$. Hence, A has a right inverse if and only if it has rank M – in other words, full row rank.

For left inverses, the picture is similar. We first transpose the equation $BA = I$ to get $A^T B^T = I^T = I$, then partition by column as we did to find a right inverse. Now we see that A has a left inverse if and only if its row space is $\Re^N$. Hence, A has a left inverse if and only if it has full column rank.

We summarize our observations in the following theorem.

Theorem 2.3.4. *For any matrix A,*
(a) *A has a right inverse if and only if it has full row rank;*
(b) *A has a left inverse if and only if it has full column rank;*
(c) *if A has full rank, then it has at least one left inverse or right inverse;*
(d) *if A is square and has full rank, then A has an inverse, which is unique.*

Note, that in Theorem 2.3.3, we described a function defined by A, and saw that it was a one-to-one mapping from the row space to the column space. If A is has an inverse, then the column space and the row space will both be $\Re^N$ (remember that A must be square). This gives us the following.

Theorem 2.3.5. *Let $f(x) = Ax$ for a matrix $A \in \Re^{M \times N}$. Then, f is a one-to-one mapping from $\Re^N$ to $\Re^M$ if and only if A has an inverse. Furthermore, this can only be the case if $N = M$.*

One final note on inverses, if A has an inverse then so does A^T, and in fact, $(A^T)^{-1} = (A^{-1})^T$. This is verified by simply multiplying the two together $(A^T)(A^{-1})^T = (A^{-1}A)^T = I^T = I$, and the check for left inverse is similar. To simplify notation, the inverse of a transpose is often denoted as A^{-T}.

2.4 Contingent Claims

Prices of assets traded in public exchanges fluctuate over time. If we make an investment today, our payoff is subject to future market prices. Many times we consider assets whose payoffs are completely determined by other assets. We refer to these as *derivative* assets.

For example, we might consider corporate stock, as being fundamental – a stock's value is derived from the state of a company and market expectations about its future. Other assets – known as *derivatives* – are contracts that promise payoffs that are functions of the future price of other assets. In a sense, these assets are synthetic.

Whether fundamental or synthetic, assets traded in a market can be thought of abstractly as *contingent claims*. This term refers to the fact that the future payoff from an asset is contingent on the future state of variables dictated by the market. We present some common examples of contingent claims.

Example 2.4.1. (Stocks, Bonds, and Options) *We describe four types of contingent claims. A share of corporate stock entitles its holder to a fraction of the company's value. This value is determined by the market. A zero-coupon bond is a contract that offers the holder a* \$1 *payoff on some future date, referred to as the maturity date, specified by terms of the contract. A European put option is a contract that offers its holder the option to sell a share of stock to the grantor at a particular price, referred to as the strike price, on an particular future date, referred to as the expiration date. The strike price and expiration date are specified by terms of the contract. A European call option, on the other hand, offers its holder the option to buy a share of stock from the grantor at a strike price on an expiration date.*

Consider four assets: a share of stock, a zero-coupon bond maturing in one year, and European put and call options on the stock, both expiring in one year, with strike prices of \$40 *and* \$60*. We could purchase any combination of these assets today and liquidate in one year to receive a payoff. The payoff of each asset will only depend on the price of the stock one year from now. Hence, we can visualize the payoff in terms of a function mapping the stock price one year from now to the associated payoff from the asset. Figure 2.6 illustrates the payoff functions associated with each asset.*

Each payoff function provides the value of one unit of an asset one year from now, as a function of the stock price. In the case of the stock, the value is the stock price itself. For the put, we would exercise our option to sell the stock only if the the strike price exceeds the stock price. In this event, we would purchase a share of stock at the prevailing stock price and sell it at the strike price of \$40*, keeping the difference as a payoff. If the stock price exceeds the strike price, we would discard the contract without exercising. The story is similar for the call, except that we would only exercise our option to buy if the stock price exceeds the strike price. In this event, we would purchase one share of stock at the strike price of* \$60 *and sell it at the prevailing stock price, keeping the difference as a payoff.*

Suppose that the price of the stock in the preceding example can only take on values in $\{1, \ldots, 100\}$ a year from now. Then, payoff functions are conveniently represented in terms of *payoff vectors*. For example, the payoff

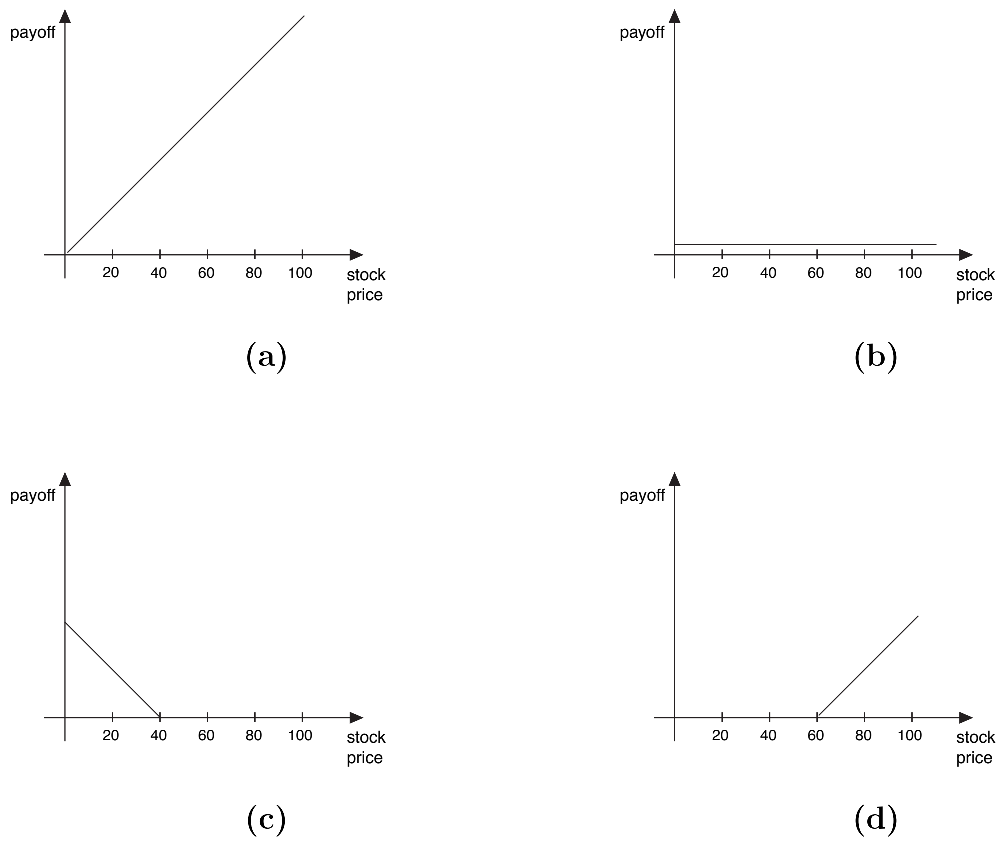


Figure 2.6: Payoff functions of a share of stock (a), a zero-coupon bond (b), a European put option (c), and a European call option (d).

vector $a^1 \in \Re^{100}$ for the stock would be defined by $a^1_i = i$. Similarly, the payoff vectors for the zero-coupon bond, European put option, and European call option would be $a^2_i = 1$, $a^3_i = \max(40 - i, 0)$, and $a^4_i = \max(i - 60, 0)$, respectively.

More generally, we may be concerned with large numbers of assets driven by many market prices – perhaps even the entire stock market. Even in such cases, the vector representation applies, so long as we enumerate all possible outcomes of interest. In particular, given a collection of N assets and M possible outcomes for relevant market prices, the payoff function associated with any asset $j \in \{1, \ldots, N\}$ can be thought of as a payoff vector $a^j \in \Re^M$, where each ith component is the payoff of the asset in outcome $i \in \{1, \ldots, M\}$.

It is sometimes convenient to represent payoff vectors for all assets in terms of a single matrix. Each column of this *payoff matrix* P is the payoff vector for an asset. In particular,

$$P = \left[\begin{array}{ccc} a^1 & \ldots & a^N \end{array} \right].$$

2.4.1 Structured Products and Market Completeness

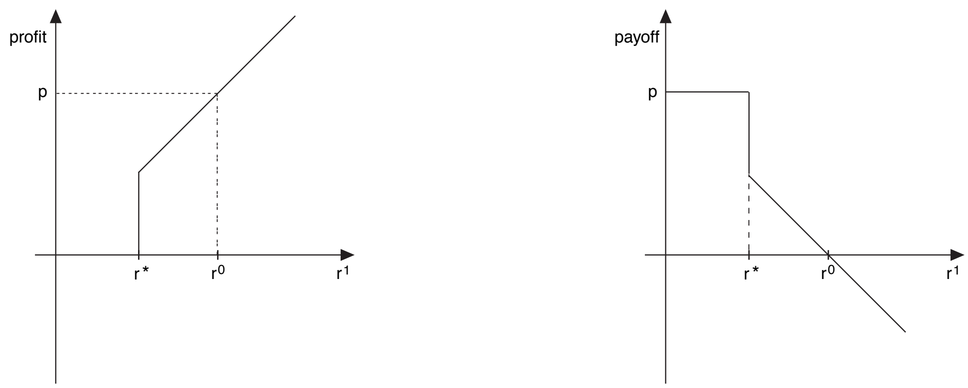
Investment banks serve investors with specialized needs. One service offered by some banks involves structuring and selling assets that accommodate a customer's demands. Such assets are called *structured products*. We provide a simple example.

Example 2.4.2. (Currency Hedge) *A company is planning to set up marketing operations in a foreign country. It is motivated by favorable sales projections. However, the company faces significant risks that are unrelated to its products or operation. If the Dollar value of the foreign currency depreciates, it may operate at a loss even if sales projections are met. In addition, if the economy of the country is unstable, there may be a dramatic devaluation arising from poor economic conditions. In this event, the sales projections would become infeasible and the company would pull out of the country entirely.*

The anticipated profit over the coming year from operating in this foreign country is p Dollars, so long as the Dollar value of the currency remains at its current level of r_0. However, if the Dollar value were to change to r_1, the company's anticipated profit would become pr_1/r_0. If the exchange rate drops to a critical value of $r_1 = r^*$, the company would shut down operations and pull out of the country entirely. The anticipated profit over the coming year as a function of the exchange rate is illustrated in Figure 2.7(a).

The company consults with an investment bank, expressing a desire to meet its profit projection during its first year of operation by focusing on product sales, without having to face risks associated with the country's currency and economic conditions. The bank designs a structured product that offers a payoff in one year that is contingent on the prevailing exchange rate r_1. The payoff function is illustrated in Figure 2.7(b). By purchasing this structured product, the company can rest assured that its profit over the coming year from this foreign operation will be p Dollars, so long as its sales projections are met.

When selling a structured product, an investment bank may be taking on significant risks. For example, if the country discussed in the above example turns out to have an economic crisis, the bank will owe a large sum of money to the company. This risk can be avoided, however, if the bank is able to replicate the structured product. It is said that a structured product with a



(a) **(b)**

Figure 2.7: (a) A company's profit as a function of the exchange rate. (b) The payoff function for a protective structured product.

payoff vector $b \in \Re^M$ can be *replicated* with assets available in the market if

$$Px = b,$$

for some $x \in \Re^N$. In financial terms, x represents a portfolio that can be acquired by trading in the market, and Px is the payoff vector associated with this portfolio. Each nonnegative x_j represents a number of units of the jth asset held in the portfolio. Each negative x_j represents a quantity sold short. The term *short-sell* refers to a process whereby one borrows units of an asset from a broker, sells them, and returns them to the broker at a later time after buying an equal number of units of the same asset from the market.

If $Px = b$, the payoff from the portfolio equals that of the structured product in every possible future outcome. Hence, by acquiring the portfolio when selling the structured product, the bank mitigates all risk. It is simply intermediating between the market and its customer. To do this right, the bank must solve the equation $Px = b$.

Can every possible contingent claim be replicated by assets available in the market? To answer this question, recall that, by definition, a contingent claim with payoff function $b \in \Re^M$ can be replicated if and only if $Px = b$ for some x. Hence, to replicate every possible contingent claim, $Px = b$ must have a solution for every $b \in \Re^M$. This is true only if $\mathcal{C}(P) = \Re^M$. In this event, the market is said to be *complete* – that is, any new asset that might be introduced to the market can be replicated by existing assets.

2.4.2 Pricing and Arbitrage

Until now, we have focussed on payoffs – what we can make from investing in assets. We now bring attention to prices – what we have to pay for assets. Let us represent initial prices of assets in the market by a vector $\rho \in \Re^N$. Hence, the unit cost of asset j is ρ_j and a portfolio $x \in \Re^N$ requires an investment of $\rho^T x$.

How should an investment bank price a structured product? The answer seems clear: since the investment bank is replicating the product using a portfolio, the bank should think of the price as being equal that of the portfolio. Of course, the bank will also tack on a premium to generate some profit.

If an asset in the market can be replicated by others, this should also translate to a price relationship. In particular, the price of a replicating portfolio should equal that of the replicated asset. If it were greater, one could short-sell the portfolio and purchase the asset to generate immediate profits without incurring cost or risk. Similarly, if the price of the replicating portfolio were lower than that of the asset, one would short sell the asset and purchase the portfolio. In either case, the initial investment is negative (the individual carrying out the transactions receives money) and the future payoff is zero, regardless of the outcome.

What we have just described is an arbitrage opportunity. More generally, an *arbitrage opportunity* is an investment strategy that involves a negative initial investment and guarantees a nonnegative payoff. In mathematical terms, an arbitrage opportunity is represented by a portfolio $x \in \Re^N$ such that $\rho^T x < 0$ and $Px \geq 0$. Under the assumption that arbitrage opportunities do not exist, it often possible to derive relationships among asset prices. We provide a simple example

Example 2.4.3. (Put-Call Parity) *Consider four assets:*
(a) *a stock currently priced at s_0 that will take on a price $s_1 \in \{1, \ldots, 100\}$ one month from now;*
(b) *a zero-coupon bond priced at β_0, maturing one month from now;*
(c) *a European put option currently priced at p_0 with a strike price $\kappa > 0$, expiring one month from now;*
(d) *a European call option currently priced at c_0 with the same strike price κ, expiring one month from now.*

The payoff vectors $a^1, a^2, a^3, a^4 \in \Re^M$ are given by

$$a_i^1 = \mu_i, \quad a_i^2 = 1, \quad a_i^3 = \max(\kappa - \mu_i, 0), \quad a_i^4 = \max(\mu_i - \kappa, 0),$$

for $i \in \{1, \ldots, M\}$. Note that if we purchase one share of the stock and one put option and short-sell one call option and κ units of the bond, we are guaranteed zero payoff; i.e.,

$$a^1 - \kappa a^2 + a^3 - a^4 = 0.$$

The initial investment in this portfolio would be

$$s_0 - \kappa\beta_0 + p_0 - c_0.$$

If this initial investment is nonzero, there would be an arbitrage opportunity. Hence, in the absence of arbitrage opportunities, we have the pricing relationship

$$s_0 - \kappa\beta_0 + p_0 - c_0 = 0,$$

which is known as the put-call parity.

Because arbitrage opportunities are lucrative, one might wish to determine whether they exist. In fact, one might consider writing a computer program that automatically detects such opportunities whenever they are available. We have not yet developed the tools to do this – but we will. In particular, linear programming offers a solution to this problem. Linear programming will also offer an approach to mitigating risk and pricing when selling structured products that can not be replicated by assets in the market. Duality theory will also offer interesting insights about pricing contingent claims in the absence of arbitrage opportunities. So the story of contingent claims and arbitrage is not over but to be continued in subsequent chapters.

2.5 Exercises

Question 1

Let $a = [1, 2]^T$ and $b = [2, 1]^T$. On the same graph, draw each of the following

1. The set of all points $x \in \Re^2$ that satisfy $a^T x = 0$.
2. The set of all points $x \in \Re^2$ that satisfy $a^T x = 1$.
3. The set of all points $x \in \Re^2$ that satisfy $b^T x = 0$.
4. The set of all points $x \in \Re^2$ that satisfy $b^T x = 1$.
5. The set all points $x \in \Re^2$ that satisfy $[a\ b]x = [0, 1]^T$.
6. The set all points $x \in \Re^2$ that satisfy $[a\ b]x = [0, 2]^T$.

In addition, shade the region that consists of all $x \in \Re^2$ that satisfy $a^T x \leq 0$.

Question 2

Consider trying to solve $Ax = b$ where

$$A = \begin{bmatrix} 1 & 2 \\ 0 & 3 \\ 2 & -1 \end{bmatrix}$$

1. Find a b so that $Ax = b$ has no solution.

2. Find a non-zero b so that $Ax = b$ has a solution.

Question 3

Find two $x \in \Re^4$ that solve all of the following equations.

$$\begin{aligned} [0,2,0,0]\,x &= 4 \\ [1,0,0,0]\,x &= 3 \\ [2,-1,-2,1]\,x &= 0 \end{aligned}$$

Write $[4,3,0]^T$ as a linear combination of $[0,1,2]^T, [2,0,-1]^T, [0,0,-2]^T$ and $[0,0,1]^T$ in two different ways. Note: When we say write x as a linear combination of a, b and c, what we mean is find the coefficients of the a, b and c. For example $x = 5a - 2b + c$.

Question 4

Suppose $A, B \in \Re^{3\times 3}$ are defined by $A_{ij} = i + j$ and $B_{ij} = (-1)^{ij}$ for each i and j. What is A^T? AB?

Suppose we now change some elements of A so that $A_{1j} = e_j$[1]. What is A now?

Question 5

Suppose U and V are both subspaces of S. Are $U \cap V$ and $U \cup V$ subspaces of S? Why, or why not? Hint: Think about whether or not linear combinations of vectors are still in the set.

Question 6

$$a = \begin{bmatrix} 1 \\ 4 \\ -1 \\ 2 \\ 3 \end{bmatrix}, b = \begin{bmatrix} 1 \\ 2 \\ 0 \\ -3 \\ 1 \end{bmatrix}, c = \begin{bmatrix} 0 \\ 2 \\ -1 \\ 5 \\ 2 \end{bmatrix}, d = \begin{bmatrix} 1 \\ 0 \\ 2 \\ -8 \\ -1 \end{bmatrix}, e = \begin{bmatrix} 1 \\ 10 \\ -1 \\ 17 \\ 9 \end{bmatrix}, f = \begin{bmatrix} -4 \\ -2 \\ 1.5 \\ 27 \\ 2 \end{bmatrix}$$

1. The span of $\{a, b, c\}$ is an N dimensional subspace of $\Re^M$. What are M and N?
2. What is $6a + 5b - 3e + 2f$?
3. Write a as a linear combination of b, e, f.
4. write f as a linear combination of a, b, e.
5. The span of $\{a, b, c, d, e, f\}$ is an N dimensional subspace of $\Re^M$. What are M and N?

Question 7

Suppose I have 3 vectors x, y and z. I know $x^T y = 0$ and that y is not a multiple of z. Is it possible for $\{x, y, z\}$ to be linearly dependent? If so, give an example. If not, why not?.

Question 8

Find 2 matrices $A, B \in \Re^{2\times2}$ so that none of the entries in A or B are zero, but AB is the zero matrix. Hint: Orthogonal vectors.

Question 9

Suppose the only solution to $Ax = 0$ is $x = 0$. If $A \in \Re^{M\times N}$ what is its rank, and why?

Question 10

The system of equations

$$\begin{aligned} 3x + ay &= 0 \\ ax + 3y &= 0 \end{aligned}$$

always has as a solution $x = y = 0$. For some a the equations have more than one solution. Find two such a.

Question 11

In $\Re^3$, describe the 4 subspaces (column, row, null, left-null) of the matrix

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

Question 12

In real life there are many more options than those described in class, and they can have differing expiry dates, strike prices, and terms.

1. Suppose there are 2 call options with strike prices \$ 1 and \$ 2. Which of the two will have the higher price?

2. Suppose the options were put options. Which would now have the higher price?

3. If an investor thinks that the market will be particularly volatile in the coming weeks, but does not know whether the market will go up or down, they may choose to buy an option that pays $|S - K|$ where S will be the value of a particular share in one months time, and K is the *current* price of the share. If the price of a zero coupon bond maturing in one month is 1, then what is the difference between a call option and a put option both having strike price K. (Report the answer in terms of the strike price K.)

4. Suppose I am considering buying an option on a company. I am considering options with a strike price of \$ 100, but there are different expiry dates. The call option expiring in one month costs \$ 10 while the put option expiring in one month costs \$ 5. If the call option expiring in two months costs \$ 12 and zero coupon bonds for both time frames cost \$ 0.8, then how much should a put option expiring in two months cost?

Question 13

Find a matrix whose row space contains $[1\ 2\ 1]^T$ and whose null space contains $[1\ -2\ 1]^T$ or show that there is no such matrix.

Question 14

True or False:

1. If U is orthogonal to V then $U^{\perp}$ is orthogonal to $V^{\perp}$.
2. If U is orthogonal to V and V is orthogonal to W, then it is always true that U us orthogonal to W.
3. If U is orthogonal to V and V is orthogonal to W, then it is never true that U us orthogonal to W.

Question 15

Can the null space of AB be bigger than the null space of B? If so, give an example. If not, why not.

Question 16

Find two different matrices that have the same null, row, column, and left-null spaces.