Part I

Basic Formulation

Models are to be used, not believed.
H. Theil (Principles of Econometrics)

1

Getting Concepts and Gathering Tools

1.1 Vectors and Scalars

In science and engineering we come across many quantities which require both magnitude and direction for their complete specification, e.g., velocity, acceleration, momentum, force, angular momentum, torque, electrical current density, electric and magnetic fields, pressure and temperature gradients, heat flow and so on. To deal with such quantities, we need laws to represent, combine and manipulate them. Instead of creating these laws separately for each of these quantities, it makes good sense to create a mathematical model to set up common laws for all quantities requiring both magnitude and direction to be specified. This idea is neither new nor alien: right from our childhood we deal with real numbers and integers which are the mathematical objects representing a value of 'something'. This 'something' is anything which can be quantified or measured and whose value is specified as a single entity: length, mass, time, energy, area, volume, curvature, cash in your pocket, the size of the memory and the speed of your computer, bank interest rates $\cdots$. The combination and manipulation of these values is effected by combining and manipulating the corresponding real numbers. Similarly, the values of the quantities specified by magnitude and direction are represented by vectors. A vector is *completely specified by its magnitude and direction*. Note that the magnitude of a vector is specified by a single real number ≥ 0, so if we wish to change only the magnitude of a vector, we must have the facility to multiply a vector by a real number, which we call a *scalar* in this context. Henceforth, in this book, by a scalar we mean a real number. Thus, in order to develop an algebra on the set of vectors, we need to associate with it the set of scalars and define the laws for multiplying a vector by a scalar. If we multiply a vector by -1 we get the vector with same magnitude but opposite in direction, which, when added to the original vector gives the zero vector, that is, a vector with zero magnitude and no direction. Two vectors are equal if *they have equal magnitudes and the same direction*.

In this book we are using boldfaced letters for vectors. A symbol which is not bold, may represent the magnitude of the corresponding vector, or a scalar.

1.2 Space and Direction

We have not attempted to formally define 'space' or 'direction' as these are the integral parts of our experience right from birth. By space we mean the space we live in and move around. We experience direction by our motion as well as by observing other moving objects. We call our space three dimensional, (3-D) because given any two different directions, we can always choose a third direction such that going through any sequence of displacements along any two of them, we will never move along the third and also because given any set of four different directions we can always find a sequence of displacements through any three of them, which will take us along the fourth. *In this book, any n-dimensional object is denoted* n-D. We also assume that space is a continuum, that is, any region of space can be divided arbitrarily and indefinitely into smaller and smaller regions. Further, we assume that space is an inert vacuum, whose sole purpose is to make room for different physical phenomena to occur in it. We denote this space by a symbol $\mathbb{R}^3$. You may wonder about this weird symbol. However, we will understand it in due course. For the time being we just view this symbol as a short name for our space with the above properties.

In order to incorporate the concept of direction in our model, we note that any straight line in space specifies two directions, each by the sense in which the line is traversed. In order to pick one of these two directions, we may put an arrow-head on the line, pointing in the direction we want to indicate. Thus, a straight line with an arrow is our first model for specifying direction in space (see Fig. 1.1(a)). We will refine it shortly. Note that if we parallelly transport a line with an arrow, (that is, the transported line is always parallel to the original one), it indicates the same direction. Thus, two different directions in space correspond to two intersecting straight lines with arrows appropriately placed on them. One of these directions (which we call 'reference direction') can be reached from the other by rotating the other direction about the line normal to the plane containing the two intersecting lines and passing through the point of intersection, until both, the lines and the arrows, coincide (see Fig. 1.1(b)). The angular advance made by the rotating line is simply the angle between the two directions. This angle can be measured by drawing a circle of radius r in the plane of two intersecting lines with its center at the point of intersection and measuring the length of the arc of this circle, say S, swept by the rotating line. The angle θ swept by the rotating line is then given by

$$S = r\theta.$$

Any arbitrary circle drawn in the specified plane can be used to get the value of angle θ via the above equation $(\theta = S/r)$. In other words, the radius r is arbitrary. It is convenient to choose a unit circle, that is, a circle with radius unity, $(r = 1)$, so that the arc-length and the angle swept by the rotating line are numerically equal (see Fig. 1.1(c)). Such a arc-length measure of angle is called 'radian measure'. Since the length of the circumference of a unit circle is 2π, the angle corresponding to one complete rotation is 2π. The angle corresponding to half the circumference is π and so on.

This procedure still leaves an ambiguity in defining the angle between two directions. We can rotate one of the directions (so as to coincide with the other direction) in two ways. The sense of one rotation is reverse to that of the other. Each of these rotations correspond to different angles, say θ and $2\pi - \theta$ (see Fig. 1.1(b)). Which of these rotations do we choose? We place a clock with its center at the point of intersection of the two lines so as to view it from the top. We then choose the rotation in the sense opposite to that of the hands of the clock. This is called counterclockwise rotation.

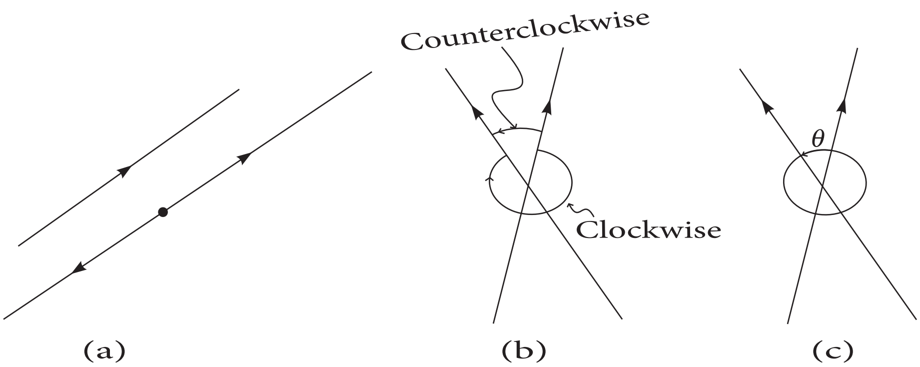

Fig. 1.1 (a) A line indicates two possible directions. A line with an arrow specifies a unique direction. (b) The angle between two directions is the amount by which one line is to be rotated so as to coincide with the other along with the arrows. Note the counterclockwise and clockwise rotations. (c) The angle between two directions is measured by the arc of the unit circle swept by the rotating direction.

The angle swept by a counterclockwise rotation is taken to be positive, while the angle swept by a clockwise rotation is negative. Note that we can always choose the angle between two directions to be $\leq \pi$ by choosing which direction is to be rotated counterclockwise towards which (see Fig. 1.2).

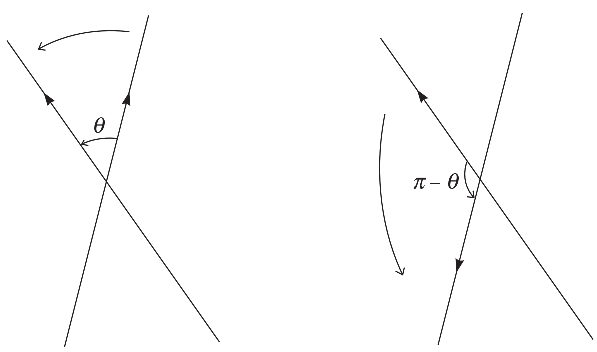

Fig. 1.2 We can choose the angle between directions $\leq \pi$ by choosing which direction is to be rotated (counterclockwise) towards which.

The angle between two directions is used to specify one direction relative to the other. If you reflect on your experience, you will realize that the only way to specify a direction is to specify it relative to some other reference direction which you can determine by observing something like a magnetic needle. To appreciate this, imagine that you are on a ship sailing in the mid-pacific. Suppose that you have no device like a magnetic compass or a gyroscope on the ship (I do not recommend this!) and that clouds block your vision of the pole star and the other stars. Then it is impossible to tell in which direction your ship is moving.

Exercise Consider three different *non-coplanar lines*[1] intersecting at a point O. Take a point P which is not on any of these three lines. Put arrows on these three lines to specify three directions (Draw a figure). Construct a path starting at O and ending at P on which you are moving either in or opposite to one of the three directions you have specified by putting arrows on the three lines. Convince yourself that this is always possible. In the light of the statements made in the first para of this section, this exercise demonstrates that our space is three dimensional. □

1.3 Representing Vectors in Space

Let us now consider a physical quantity, say electric field, whose 'values' are vectors. We call such a quantity, a 'vector quantity'. Each value is a specific vector, with given magnitude and direction. For example, magnitude of earth's magnetic field can be specified as say, 0.37 gauss and the direction can be given relative to that implied by earth's polar axis. Any such vector can be represented in space as follows. Given the magnitude and the direction of the vector, we draw a line in space in the direction of the vector. Then, we mark out a segment of this line whose length is proportional to the magnitude of the vector and then put an arrow at one of the ends of this segment to indicate the direction of the vector. For example, to represent a vector specifying a value of the electric field, we may choose a length of 1 cm to correspond to the magnitude of 1 volt/meter. An electric field vector of magnitude x volts/meter is then represented by a segment of length x cm. Once chosen, the same constant of proportionality must be used to represent all vectors corresponding to the electric field. Every vector giving a possible value of a vector quantity is completely represented in space by the corresponding segment with an arrow at one of its ends. Of course, the arrow can be placed anywhere on the line segment, not necessarily at one of its ends.

The end opposite to the arrow on the vector (drawn in space) is called its base point. Since a vector is completely specified by its magnitude and direction, it can be represented in space at any point as its base point, because changing the base point does not change the length or the direction of the vector. Two or more representations of the same vector based at different points in space are to be taken as the same vector (see Fig. 1.3).

[1] Any number of lines *all* of which fall on the same plane are called *coplanar*. A collection of lines which are not coplanar is called non-coplanar. A pair of intersecting lines is coplanar.

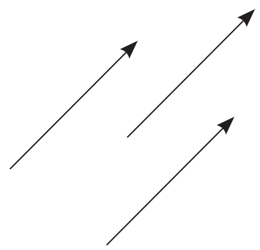

Fig. 1.3 Different representations of the same vector in space

Henceforth, by a vector, we will mean the representation of a value of a vector quantity in space, which is simply proportional to the actual value of the vector quantity it represents. *This enables us to specify every vector by its length and direction, without any reference to the physical quantity it represents. This gives us the freedom to set up the laws of combining two or more vectors in the same sense as we set up the laws for combining real or complex numbers without reference to the quantities they correspond to.* Thus, we can develop the theory of vectors independent of which physical quantity they represent and common to all applications of vectors. The vectors giving the possible positions of a point particle in space (relative to some origin) are called the *position vectors.* The set of all vectors is in one to one correspondence with the set of points in space.

In some applications, a vector has to be localized in space, that is, it has to be based at a particular point in space and cannot be parallel transported. A typical example is – the forces applied at a given set of points on a body which is in mechanical equilibrium, so that the net force on the body is zero, as well as the net torque about any point of the body is zero. Here, the set of applied forces are vectors fixed at the points of application. Such a localization of vectors can be effected by assigning them to the points in space or to the corresponding position vectors. If the number of vectors we are dealing with is finite and small, we can assign this set of vectors to the corresponding set of position vectors by giving an explicit table of assignment. If the vectors and the corresponding position vectors form a continuum, then the assignment takes the form of a vector valued function of the position vector variable, say $\mathbf{f}(\mathbf{x})$, which is called a vector field (see section 1.15).

Apart from the vectors representing the values of vector quantities in space, we need to draw another kind of vectors in space. These are called unit vectors whose length is always unity. Thus, two unit vectors differ only in direction. A unit vector replaces the 'line with an arrow' model to specify a direction in space. The sole purpose of a unit vector is to specify a direction in space. In particular, the length of a unit vector does not correspond to the magnitude of any physical quantity. We shall always denote a unit vector by a hat over it, so that you can recognize it as a unit vector even if that is not explicitly stated. Given a vector $\mathbf{a}$, $\hat{\mathbf{a}}$ will denote the unit vector in the direction of $\mathbf{a}$. Thus, every vector $\mathbf{a} \neq \mathbf{0}$ can be written as

$$\mathbf{a} = |\mathbf{a}|\hat{\mathbf{a}},$$

where $|\mathbf{a}|$ denotes the magnitude of $\mathbf{a}$.

The geometric interpretation of the set of real numbers is a straight line, that is, the set of real numbers is in one to one correspondence with the points on the line. Similarly, the set of vectors is in one to one correspondence with the points in the three dimensional space $\mathbb{R}^3$. To see this one to one correspondence, consider the set of vectors comprising all possible values of some vector quantity. We can construct the set containing the representatives of these vectors in space. One to one correspondence between these two sets is obvious by construction. To transfer this correspondence to the points in $\mathbb{R}^3$ we take an arbitrary point in space say O, called origin and represent every vector with O as the base point. Since the vectors have all possible magnitudes and directions, every point in space is at the tip of some vector based at O, representing a possible value of the vector quantity. In this way, a unique magnitude and direction is assigned to every point in space, establishing the one to one correspondence between the set of vectors and the set of points in space. We could have chosen any other point, say O' as the origin and base all vectors at O'. This gives a new representation for each vector in the set of vectors obtained by parallelly transporting each vector based at O to that based at O'. These two are the representations of the same set of vectors (values of a vector quantity). However, they generate two different one to one correspondences with the points in $\mathbb{R}^3$ as can be seen from Fig. 1.4. We see that changing the origin from O to O' makes a vector correspond to two different points in space (or, makes a point in space correspond to two different vectors) as we assign a vector (based at O or O') to a point in space. Thus, changing the origin changes the one to one correspondence between the set of vectors and the points in space. Later, we will have a closer look at the one to one correspondence between $\mathbb{R}^3$ and the set of vectors (values of a vector quantity).

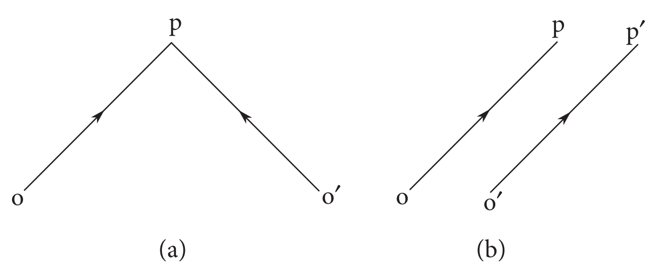

Fig. 1.4 Shifting origin makes (a) two different vectors correspond to the same point and (b) two different points correspond to the same vector

1.4 Addition and its Properties

Let us now see how to add two vectors. We will define the addition of vectors using the representatives of the values of a vector quantity in space. This frees vector addition from the corresponding vector quantities.

To add $\mathbf{a}$ and $\mathbf{b}$, base the vector $\mathbf{b}$ at the tip of $\mathbf{a}$. Then, the vector joining the base point of $\mathbf{a}$ to the tip of $\mathbf{b}$, in that direction, is the vector $\mathbf{a}+\mathbf{b}$. You can check that $\mathbf{a}+\mathbf{b}=\mathbf{b}+\mathbf{a}$ (see Fig. 1.5). Notice that the vectors $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{a}+\mathbf{b}$ form a (planar) triangle and hence are coplanar.

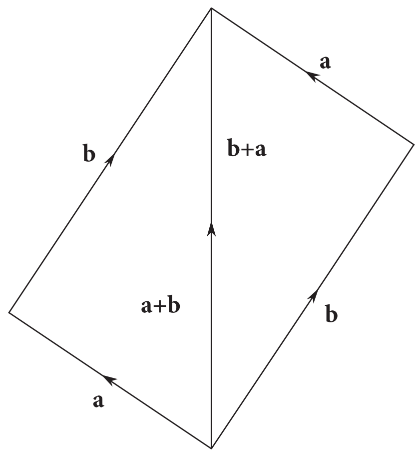

Fig. 1.5 Vector addition is commutative

The vector $\mathbf{a}+\mathbf{b}$ is sometimes called the resultant of $\mathbf{a}$ and $\mathbf{b}$. The rule of adding two or more vectors is motivated by the net displacement of an object in space, resulting due to many successive displacements. Thus, if we go from A to B by travelling 10 km NE (vector $\mathbf{a}$) and then from B to C by travelling 6 km W (vector $\mathbf{b}$) the net displacement, 8 km due North from A to C (vector $\mathbf{c}$), is obtained as depicted in Fig. 1.6(a), which is the same as that given by $\mathbf{c}=\mathbf{a}+\mathbf{b}$. Figure 1.6(b) shows the net displacement ($\mathbf{f}$) after four successive displacements ($\mathbf{a},\mathbf{b},\mathbf{c},\mathbf{d}$) which is consistent with $\mathbf{f}=\mathbf{a}+\mathbf{b}+\mathbf{c}+\mathbf{d}$.

We can now list the properties of vector addition and multiplication by a scalar.

(1) *Closure* If $\mathbf{a},\mathbf{b}$ are in $\mathbb{R}^3$ then $\mathbf{a}+\mathbf{b}$ is also in $\mathbb{R}^3$. That is, addition of two vectors results in a vector.

(2) *Commutativity* $\mathbf{a}+\mathbf{b}=\mathbf{b}+\mathbf{a}$ (see Fig. 1.5).

(3) *Associativity* For all vectors $\mathbf{a},\mathbf{b},\mathbf{c}$ in $\mathbb{R}^3$, $\mathbf{a}+(\mathbf{b}+\mathbf{c})=(\mathbf{a}+\mathbf{b})+\mathbf{c}$. Thus, while adding three or more vectors, it does not matter which two you add first, which two next etc, that is, the order in which you add does not matter (see Fig. 1.6(b)).

(4) *Identity* There is a unique vector $\mathbf{0}$ such that for every vector $\mathbf{a}$ in $\mathbb{R}^3$, $\mathbf{a}+\mathbf{0} = \mathbf{a}$.

(5) *Inverse* For every vector $\mathbf{a}\neq\mathbf{0}$ in $\mathbb{R}^3$, there is a unique vector $-\mathbf{a}$ such that $\mathbf{a}+(-\mathbf{a})=\mathbf{0}$ and $\mathbf{0}\pm\mathbf{0}=\mathbf{0}$.

To every pair α and $\mathbf{a}$ where α is a scalar (i.e., a real number) and $\mathbf{a}$ in $\mathbb{R}^3$ there is a vector $\alpha\mathbf{a}$ in $\mathbb{R}^3$. If we denote by $|\mathbf{a}|$ the magnitude of $\mathbf{a}$, then the magnitude of $\alpha\mathbf{a}$ is $|\alpha|\,|\mathbf{a}|$. If $\alpha>0$, the direction of $\alpha\mathbf{a}$ is the same as that of $\mathbf{a}$, while if $\alpha<0$ then the direction of $\alpha\mathbf{a}$ is opposite to that of $\mathbf{a}$. If $\alpha>0$, then $\alpha\mathbf{a}$ is said to be the *scaling* of $\mathbf{a}$ by α. Note that $\alpha=1/|\mathbf{a}|$ produces unit vector $\hat{\mathbf{a}}$ in the direction of $\mathbf{a}$. We have, for the scalar multiplication,

(1) *Associativity* $\alpha(\beta\mathbf{a})=(\alpha\beta)\mathbf{a}$.

(2) *Identity* $1\mathbf{a}=\mathbf{a}$.

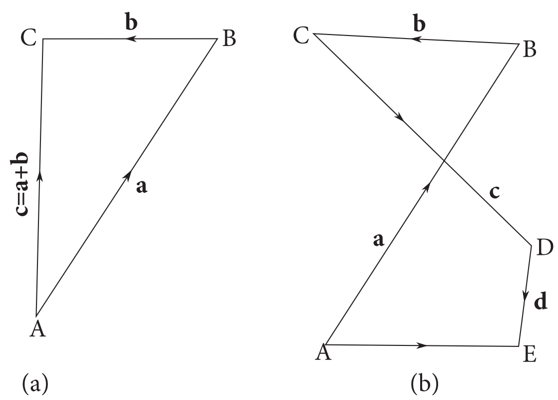

Fig. 1.6 (a) Addition of two vectors (see text). (b) Vector $\overline{AE}$ equals $\mathbf{a}+\mathbf{b}+\mathbf{c}+\mathbf{d}$. Draw different figures, adding $\mathbf{a},\mathbf{b},\mathbf{c},\mathbf{d}$ in different orders to check that this vector addition is associative.

Multiplication by scalars is distributive, namely,

(3) $\alpha(\mathbf{a}+\mathbf{b}) = \alpha\mathbf{a}+\alpha\mathbf{b}$.

(4) $(\alpha+\beta)\mathbf{a} = \alpha\mathbf{a}+\beta\mathbf{a}$.

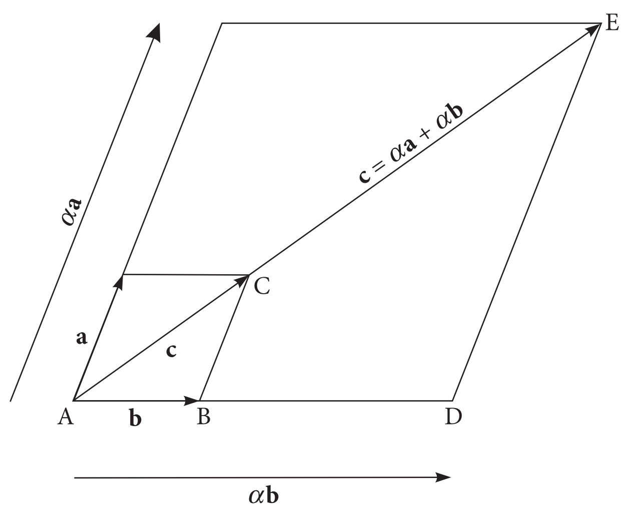

Fig. 1.7 $\alpha\mathbf{a}+\alpha\mathbf{b} = \alpha(\mathbf{a}+\mathbf{b})$

Note that these properties are shared by *all vectors independent of the context in which they are used and independent of which vector quantity they correspond to.* As explained in section 1.3, this is true of all the algebra of vectors and operations on vectors we develop in this book and will not be stated explicitly again.