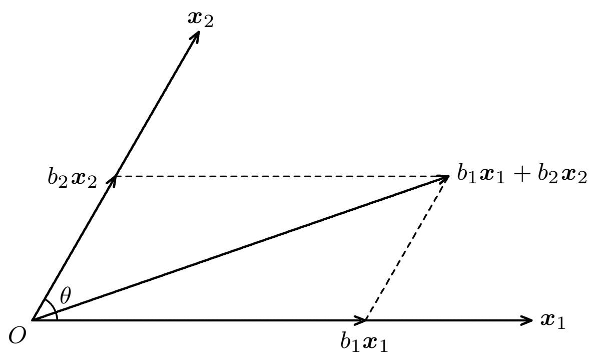

Figure 2.8 A 2-dimensional subspace

as our second dimension the vertical direction in the plane of the page, with the result that we can draw an arrow for $\boldsymbol{x}_2$, as shown.

Any vector in $\mathbb{S}(\boldsymbol{x}_1, \boldsymbol{x}_2)$ can be drawn in the plane of Figure 2.8. Consider, for instance, the linear combination of $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ given by the expression $\boldsymbol{z} \equiv b_1\boldsymbol{x}_1 + b_2\boldsymbol{x}_2$. We could draw the vector $\boldsymbol{z}$ by computing its length and the angle that it makes with $\boldsymbol{x}_1$. Alternatively, we could apply the rules for adding vectors geometrically that were illustrated in Figure 2.4 to the vectors $b_1\boldsymbol{x}_1$ and $b_2\boldsymbol{x}_2$. This is illustrated in the figure for the case in which $b_1 = 2/3$ and $b_2 = 1/2$.

In precisely the same way, we can represent any three vectors by arrows in 3–dimensional space, but we leave this task to the reader. It will be easier to appreciate the renderings of vectors in three dimensions in perspective that appear later on if one has already tried to draw 3–dimensional pictures, or even to model relationships in three dimensions with the help of a computer.

We can finally represent the regression model (2.01) geometrically. This is done in Figure 2.9. The horizontal direction is chosen for the vector $\boldsymbol{X\beta}$, and then the other two vectors $\boldsymbol{y}$ and $\boldsymbol{u}$ are shown in the plane of the page. It is clear that, by construction, $\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{u}$. Notice that $\boldsymbol{u}$, the error vector, is not orthogonal to $\boldsymbol{X\beta}$. The figure contains no reference to any system of axes, because there would be n of them, and we would not be able to avoid needing n dimensions to treat them all.

Linear Independence

In order to define the OLS estimator by the formula (1.46), it is necessary to assume that the $k \times k$ square matrix $\boldsymbol{X}^\top\boldsymbol{X}$ is invertible, or nonsingular. Equivalently, as we saw in Section 1.4, we may say that $\boldsymbol{X}^\top\boldsymbol{X}$ has full rank. This condition is equivalent to the condition that the columns of $\boldsymbol{X}$ should be **linearly independent**. This is a very important concept for econometrics. Note that the meaning of linear independence is quite different from the meaning

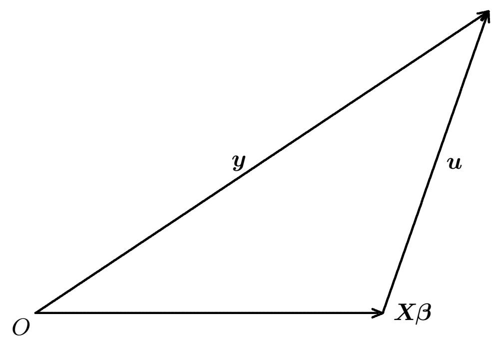

Figure 2.9 The geometry of the linear regression model

of statistical independence, which we discussed in Section 1.2. It is important not to confuse these two concepts.

The vectors $\boldsymbol{x}_1$ through $\boldsymbol{x}_k$ are said to be **linearly dependent** if we can write one of them as a linear combination of the others. In other words, there is a vector $\boldsymbol{x}_j$, $1 \leq j \leq k$, and coefficients c_i such that

$$\boldsymbol{x}_j = \sum_{i \neq j} c_i \boldsymbol{x}_i. \tag{2.10}$$

Another, equivalent, definition is that there exist coefficients b_i, at least one of which is nonzero, such that

$$\sum_{i=1}^{k} b_i \boldsymbol{x}_i = \boldsymbol{0}. \tag{2.11}$$

Recall that $\boldsymbol{0}$ denotes the **zero vector**, every component of which is 0. It is clear from the definition (2.11) that, if any of the $\boldsymbol{x}_i$ is itself equal to the zero vector, then the $\boldsymbol{x}_i$ are linearly dependent. If $\boldsymbol{x}_j = \boldsymbol{0}$, for example, then equation (2.11) is satisfied if we make b_j nonzero and set $b_i = 0$ for all $i \neq j$.

If the vectors $\boldsymbol{x}_i$, $i = 1, \ldots, k$, are the columns of an $n \times k$ matrix $\boldsymbol{X}$, then another way of writing (2.11) is

$$\boldsymbol{Xb} = \boldsymbol{0}, \tag{2.12}$$

where $\boldsymbol{b}$ is a k–vector with typical element b_i. In order to see that (2.11) and (2.12) are equivalent, it is enough to check that the typical elements of the two left-hand sides are the same; see Exercise 2.5. The set of vectors $\boldsymbol{x}_i$, $i = 1, \ldots, k$, is linearly independent if it is not linearly dependent, that is, if there are no coefficients c_i such that (2.10) is true, or (equivalently) no coefficients b_i such that (2.11) is true, or (equivalently, once more) no vector $\boldsymbol{b}$ such that (2.12) is true.