opposite directions. If the angle θ between the vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ is a right angle, its cosine is 0, and so, from (2.07), the scalar product $\langle \boldsymbol{x}, \boldsymbol{y} \rangle$ is 0. Conversely, if $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = 0$, then $\cos\theta = 0$ unless $\boldsymbol{x}$ or $\boldsymbol{y}$ is a zero vector. If $\cos\theta = 0$, it follows that $\theta = \pi/2$. Thus, if two nonzero vectors have a zero scalar product, they are at right angles. Such vectors are often said to be **orthogonal**, or, less commonly, **perpendicular**. This definition implies that the zero vector is orthogonal to everything.

Since the cosine function can take on values only between -1 and 1, a consequence of (2.07) is that

$$|\boldsymbol{x}^\top \boldsymbol{y}| \leq \|\boldsymbol{x}\| \, \|\boldsymbol{y}\|. \tag{2.08}$$

This result, which is called the **Cauchy-Schwartz inequality**, says that the absolute value of the inner product of $\boldsymbol{x}$ and $\boldsymbol{y}$ can never be greater than the length of the vector $\boldsymbol{x}$ times the length of the vector $\boldsymbol{y}$. Only if $\boldsymbol{x}$ and $\boldsymbol{y}$ are parallel does the inequality in (2.08) become the equality (2.05). Readers are asked to prove this result in Exercise 2.2.

Subspaces of Euclidean Space

For arbitrary positive integers n, the elements of an n–vector can be thought of as the coordinates of a point in E^n. In particular, in the regression model (2.01), the regressand $\boldsymbol{y}$ and each column of the matrix of regressors $\boldsymbol{X}$ can be thought of as vectors in E^n. This makes it possible to represent a relationship like (2.01) geometrically.

It is obviously impossible to represent all n dimensions of E^n physically when $n > 3$. For the pages of a book, even three dimensions can be too many, although a proper use of perspective drawings can allow three dimensions to be shown. Fortunately, we can represent (2.01) without needing to draw in n dimensions. The key to this is that there are only three vectors in (2.01): $\boldsymbol{y}$, $\boldsymbol{X\beta}$, and $\boldsymbol{u}$. Since only two vectors, $\boldsymbol{X\beta}$ and $\boldsymbol{u}$, appear on the right-hand side of (2.01), only two dimensions are needed to represent it. Because $\boldsymbol{y}$ is equal to $\boldsymbol{X\beta} + \boldsymbol{u}$, these two dimensions suffice for $\boldsymbol{y}$ as well.

To see how this works, we need the concept of a **subspace** of a Euclidean space E^n. Normally, such a subspace has a dimension lower than n. The easiest way to define a subspace of E^n is in terms of a set of **basis vectors**. A subspace that is of particular interest to us is the one for which the columns of $\boldsymbol{X}$ provide the basis vectors. We may denote the k columns of $\boldsymbol{X}$ as $\boldsymbol{x}_1$, $\boldsymbol{x}_2, \ldots \boldsymbol{x}_k$. Then the subspace associated with these k basis vectors is denoted by $\mathcal{S}(\boldsymbol{X})$ or $\mathcal{S}(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k)$. The basis vectors are said to **span** this subspace, which in general is a k–dimensional subspace.

The subspace $\mathcal{S}(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k)$ consists of every vector that can be formed as a **linear combination** of the $\boldsymbol{x}_i$, $i = 1, \ldots, k$. Formally, it is defined as

$$\mathcal{S}(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k) \equiv \Big\{ \boldsymbol{z} \in E^n \;\Big|\; \boldsymbol{z} = \sum_{i=1}^{k} b_i \boldsymbol{x}_i, \quad b_i \in \mathbb{R} \Big\}. \tag{2.09}$$