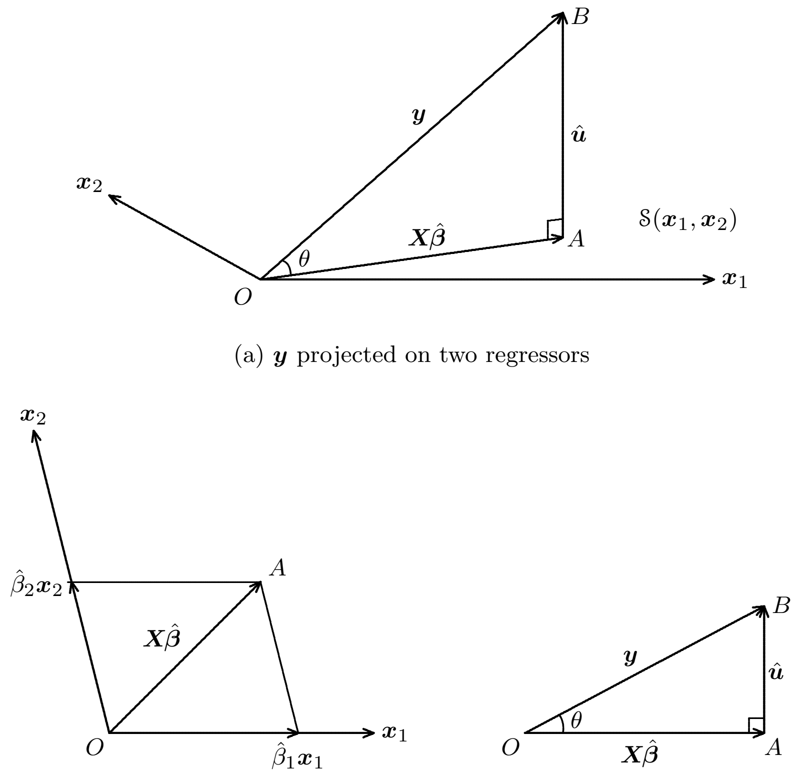


(a) $\boldsymbol{y}$ projected on two regressors

(b) The span $\mathcal{S}(\boldsymbol{x}_1, \boldsymbol{x}_2)$ of the regressors

(c) The vertical plane through $\boldsymbol{y}$

**Figure 2.11** Linear regression in three dimensions

a perpendicular" from $\boldsymbol{y}$ to the horizontal plane. The least-squares interpretation of the MM estimator $\hat{\boldsymbol{\beta}}$ can now be seen to be a consequence of simple geometry. The shortest distance from $\boldsymbol{y}$ to the horizontal plane is obtained by descending vertically on to it, and the point in the horizontal plane vertically below $\boldsymbol{y}$, labeled $A$ in the figure, is the closest point in the plane to $\boldsymbol{y}$. Thus $\|\hat{\boldsymbol{u}}\|$ minimizes $\|\boldsymbol{u}(\boldsymbol{\beta})\|$, the norm of $\boldsymbol{u}(\boldsymbol{\beta})$, with respect to $\boldsymbol{\beta}$. The squared norm, $\|\boldsymbol{u}(\boldsymbol{\beta})\|^2$, is just the sum of squared residuals, $\text{SSR}(\boldsymbol{\beta})$; see (1.49). Since minimizing the norm of $\boldsymbol{u}(\boldsymbol{\beta})$ is the same thing as minimizing the squared norm, it follows that $\hat{\boldsymbol{\beta}}$ is the OLS estimator.

Panel (b) of the figure shows the horizontal plane $\mathcal{S}(\boldsymbol{x}_1, \boldsymbol{x}_2)$ as a straightforward 2–dimensional picture, seen from directly above. The point $A$ is the point directly underneath $\boldsymbol{y}$, and so, since $\boldsymbol{y} = \boldsymbol{X}\hat{\boldsymbol{\beta}} + \hat{\boldsymbol{u}}$ by definition, the vector represented by the line segment $OA$ is the vector of fitted values, $\boldsymbol{X}\hat{\boldsymbol{\beta}}$. Geometrically, it is much simpler to represent $\boldsymbol{X}\hat{\boldsymbol{\beta}}$ than to represent just the