



FIGURE 3.9. *Principal components of some input data points. The largest principal component is the direction that maximizes the variance of the projected data, and the smallest principal component minimizes that variance. Ridge regression projects \mathbf{y} onto these components, and then shrinks the coefficients of the low-variance components more than the high-variance components.*