> **Problem 5.** Consider the data set of $n = 20$ data points with $q = 2$, shown in Figure 2:
>
> $(1, -1 + 2j/9), (-1, -1 + 2j/9),$
>
> for $j = 0, \ldots, 9$. Run the k-means algorithm with $k = 2, 3, 4$. Initialize the centers to the first k points in the list
>
> $(-1, -1), (1, 1), (-1, 1), (1, -1).$
>
> Display the clustered data. Discuss the effects of choosing the "wrong" value for k. Then, repeat the experiment, initializing the centers to $(0, -1 + 2j/(k-1)), j = 0, \ldots, k-1$. Note that although the answer is different, it is also a local minimizer of the (nonconvex) function R. Compare with the first set of answers and discuss the difficulty it illustrates with this kind of clustering.

Sensitivities of the clustering to the initial choice of centers and the number of clusters are serious pitfalls. As we see in Problem 6, another serious pitfall arises from the sensitivity of the clustering to variable transformations.

> **Problem 6.** Consider the data set from Problem 5, but multiply the second component of each data point by 100. Repeat the clustering experiments, applying the same transformation to the initial centers. Discuss why coordinate scaling is important in clustering algorithms.

Through our investigations, we see that despite its pitfalls, clustering is an important tool for data classification, noise reduction, and storage savings. Check the Web site for data for the problems and, later, for sample solutions (http://computer.org/cise/homework/v5n5.htm).

Acknowledgments
Simon P. Schurr provided helpful comments on this project.

Partial Solution to "Robot Control: Swinging Like a Pendulum"

By Dianne P. O'Leary and Yalin E. Sagduyu

> **Problem 1.** Consider the *undriven damped pendulum* modeled by
>
> $$m\ell \frac{d^2\theta(t)}{dt^2} + c\frac{d\theta(t)}{dt} + mg\sin(\theta(t)) = u(t), \quad (1)$$
>
> when $u(t) = 0$ and $c > 0$. Linearize the second-order nonlinear differential equation using the approximation $\sin(\theta(t)) \approx \theta(t)$. Transform this equation into a first-order system of ODEs of the form $\mathbf{y}' = A\mathbf{y}$, where A is a 2×2 matrix, and the two components of the vector $\mathbf{y}(t)$ represent $y_1(t) = \theta(t)$ and $y_2(t) = d\theta(t)/dt$. Determine the eigenvalues of A. Show that the damped system is *stable*—that the real part of each eigenvalue is negative—and that the undamped system is not. Use the eigenvalue information to show how the solutions behave in the damped and undamped systems.

Answer: Under the transformation, Equation 1 becomes

$$\begin{bmatrix} 1 & 0 \\ c & m\ell \end{bmatrix}\begin{bmatrix} y'_1(t) \\ y'_2(t) \end{bmatrix} = \begin{bmatrix} y_2(t) \\ -mg\sin(y_1(t)) \end{bmatrix}$$

or

$$\begin{bmatrix} y'_1(t) \\ y'_2(t) \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ -c/(m\ell) & 1/(m\ell) \end{bmatrix}\begin{bmatrix} y_2(t) \\ -mg\sin(y_1(t)) \end{bmatrix}.$$

Replacing $\sin(y_1(t))$ with $y_1(t)$ gives the system

$$\mathbf{y}' = \begin{bmatrix} y'_1(t) \\ y'_2(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -g/\ell & -c/(m\ell) \end{bmatrix}\begin{bmatrix} y_1(t) \\ y_2(t) \end{bmatrix} = \mathbf{A}\mathbf{y}.$$

The eigenvalues of the matrix A are the roots of $\det(A - \lambda I) = 0$, or the roots of $\lambda^2 + \lambda c/(m\ell) + g/\ell = 0$:

$$\lambda_{1,2} = -\frac{c}{2m\ell} \pm \sqrt{\frac{c^2}{4m^2\ell^2} - \frac{g}{\ell}}.$$