15.4 Conditioning of Eigenvalue Problems

The following $n \times n$ matrix

$$A = \begin{pmatrix} 0 & & & & & \\ 1 & 0 & & & & \\ & 1 & 0 & & & \\ & & \ddots & \ddots & & \\ & & & 1 & 0 & \\ & & & & 1 & 0 \end{pmatrix}$$

has the eigenvalue 0 with multiplicity n. However, if we perturb the top rightmost entry of A by ϵ, it is easy to see that the characteristic polynomial of the matrix

$$A(\epsilon) = \begin{pmatrix} 0 & & & & & \epsilon \\ 1 & 0 & & & & \\ & 1 & 0 & & & \\ & & \ddots & \ddots & & \\ & & & 1 & 0 & \\ & & & & 1 & 0 \end{pmatrix}$$

is $X^n - \epsilon$. It follows that if $n = 40$ and $\epsilon = 10^{-40}$, $A(10^{-40})$ has the eigenvalues $10^{-1}e^{k2\pi i/40}$ with $k = 1, \ldots, 40$. Thus, we see that a very small change ($\epsilon = 10^{-40}$) to the matrix A causes a significant change to the eigenvalues of A (from 0 to $10^{-1}e^{k2\pi i/40}$). Indeed, the relative error is 10^{-39}. Worse, due to machine precision, since very small numbers are treated as 0, the error on the computation of eigenvalues (for example, of the matrix $A(10^{-40})$) can be very large.

This phenomenon is similar to the phenomenon discussed in Section 9.5 where we studied the effect of a small perturbation of the coefficients of a linear system $Ax = b$ on its solution. In Section 9.5, we saw that the behavior of a linear system under small perturbations is governed by the condition number $\mathrm{cond}(A)$ of the matrix A. In the case of the eigenvalue problem (finding the eigenvalues of a matrix), we will see that the conditioning of the problem depends on the condition number of the change of basis matrix P used in reducing the matrix A to its diagonal form $D = P^{-1}AP$, rather than on the condition number of A itself. The following proposition in which we assume that A is diagonalizable and that the matrix norm $\| \ \|$ satisfies a special condition (satisfied by the operator norms $\| \ \|_p$ for $p = 1, 2, \infty$), is due to Bauer and Fike (1960).

Proposition 15.11. *Let $A \in \mathrm{M}_n(\mathbb{C})$ be a diagonalizable matrix, P be an invertible matrix, and D be a diagonal matrix $D = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ such that*

$$A = PDP^{-1},$$