Jim Lambers
MAT 461/561
Spring Semester 2009-10
Lecture 17 Notes

These notes correspond to Section 7.2 in the text.

Eigenvalues and Eigenvectors

We have learned what it means for a sequence of vectors to converge to a limit. However, using the definition alone, it may still be difficult to determine, conclusively, whether a given sequence of vectors converges. For example, suppose a sequence of vectors is defined as follows: we choose the initial vector $\mathbf{x}^{(0)}$ arbitrarily, and then define the rest of the sequence by

$$\mathbf{x}^{(k+1)} = A\mathbf{x}^{(k)}, \quad k = 0, 1, 2, \ldots$$

for some matrix A. Such a sequence will actually arise when we discuss the convergence of various iterative methods for solving systems of linear equations.

An important question will be whether a sequence of this form converges to the zero vector. This will be the case if

$$\lim_{k\to\infty} \|\mathbf{x}^{(k)}\| = 0$$

in some vector norm. From the definition of $\mathbf{x}^{(k)}$, we must have

$$\lim_{k\to\infty} \|A^k\mathbf{x}^{(0)}\| = 0.$$

From the submultiplicative property of matrix norms,

$$\|A^k\mathbf{x}^{(0)}\| \leq \|A\|^k\|\mathbf{x}^{(0)}\|,$$

from which it follows that the sequence will converge to the zero vector if $\|A\| < 1$. However, this is only a *sufficient* condition; it is not *necessary*.

To obtain a sufficient *and* necessary condition, it is necessary to achieve a better understanding of the effect of matrix-vector multiplication on the magnitude of a vector. However, because matrix-vector multiplication is a complicated operation, this understanding can be difficult to acquire. Therefore, it is helpful to identify circumstances under which this operation can be simply described.

To that end, we say that a nonzero vector $\mathbf{x}$ is an *eigenvector* of an $n \times n$ matrix A if there exists a scalar λ such that

$$A\mathbf{x} = \lambda\mathbf{x}.$$

The scalar λ is called an *eigenvalue* of A corresponding to $\mathbf{x}$. Note that although $\mathbf{x}$ is required to be nonzero, it is possible that λ can be zero. It can also be complex, even if A is a real matrix.

If we rearrange the above equation, we have

$$(A - \lambda I)\mathbf{x} = \mathbf{0}.$$

That is, if λ is an eigenvalue of A, then $A - \lambda I$ is a singular matrix, and therefore $\det(A - \lambda I) = 0$. This equation is actually a polynomial in λ, which is called the *characteristic polynomial* of A. If A is an $n \times n$ matrix, then the characteristic polynomial is of degree n, which means that A has n eigenvalues, which may repeat.

The following properties of eigenvalues and eigenvectors are helpful to know:

- If λ is an eigenvalue of A, then there is at least one eigenvector of A corresponding to λ
- If there exists an invertible matrix P such that $B = PAP^{-1}$, then A and B have the same eigenvalues. We say that A and B are *similar*, and the transformation PAP^{-1} is called a *similarity transformation*.
- If A is a symmetric matrix, then its eigenvalues are real.
- If A is a *skew-symmetric* matrix, meaning that $A^T = -A$, then its eigenvalues are either equal to zero, or are purely imaginary.
- If A is a real matrix, and $\lambda = u + iv$ is a complex eigenvalue of A, then $\bar{\lambda} = u - iv$ is also an eigenvalue of A.
- If A is a triangular matrix, then its diagonal entries are the eigenvalues of A.
- $\det(A)$ is equal to the product of the eigenvalues of A.
- $\text{tr}(A)$, the sum of the diagonal entries of A, is also equal to the sum of the eigenvalues of A.

It follows that any method for computing the roots of a polynomial can be used to obtain the eigenvalues of a matrix A. However, in practice, eigenvalues are normally computed using iterative methods that employ orthogonal similarity transformations to reduce A to upper triangular form, thus revealing the eigenvalues of A. In practice, such methods for computing eigenvalues are used to compute roots of polynomials, rather than using polynomial root-finding methods to compute eigenvalues, because they are much more robust with respect to roundoff error.

It can be shown that if each eigenvalue λ of a matrix A satisfies $|\lambda| < 1$, then, for any vector $\mathbf{x}$,

$$\lim_{k\to\infty} A^k\mathbf{x} = \mathbf{0}.$$

Furthermore, the converse of this statement is also true: if there exists a vector $\mathbf{x}$ such that $A^k\mathbf{x}$ does not approach $\mathbf{0}$ as $k \to \infty$, then at least one eigenvalue λ of A must satisfy $|\lambda| \geq 1$.

Therefore, it is through the eigenvalues of A that we can describe a necessary and sufficient condition for a sequence of vectors of the form $\mathbf{x}^{(k)} = A^k\mathbf{x}^{(0)}$ to converge to the zero vector. Specifically, we need only check if the magnitude of the largest eigenvalue is less than 1. For convenience, we define the *spectral radius* of A, denoted by $\rho(A)$, to be $\max|\lambda|$, where λ is an eigenvalue of A. We can then conclude that the sequence $\mathbf{x}^{(k)} = A^k\mathbf{x}^{(0)}$ converges to the zero vector if and only if $\rho(A) < 1$.

The spectral radius is closely related to natural (induced) matrix norms. Let λ be the largest eigenvalue of A, with $\mathbf{x}$ being a corresponding eigenvector. Then, for any natural matrix norm $\|\cdot\|$, we have

$$\rho(A)\|\mathbf{x}\| = |\lambda|\|\mathbf{x}\| = \|\lambda\mathbf{x}\| = \|A\mathbf{x}\| \leq \|A\|\|\mathbf{x}\|.$$

Therefore, we have $\rho(A) \leq \|A\|$. When A is symmetric, we also have

$$\|A\|_2 = \rho(A).$$

For a general matrix A, we have

$$\|A\|_2 = [\rho(A^TA)]^{1/2},$$

which can be seen to reduce to $\rho(A)$ when $A^T = A$, since, in general, $\rho(A^k) = \rho(A)^k$.

Because the condition $\rho(A) < 1$ is necessary and sufficient to ensure that $\lim_{k\to\infty} A^k\mathbf{x} = \mathbf{0}$, it is possible that such convergence may occur even if $\|A\| \geq 1$ for some natural norm $\|\cdot\|$. However, if $\rho(A) < 1$, we can conclude that

$$\lim_{k\to\infty} \|A^k\| = 0,$$

even though $\lim_{k\to\infty} \|A\|^k$ may not even exist.

In view of the definition of a matrix norm, that $\|A\| = 0$ if and only if $A = 0$, we can conclude that if $\rho(A) < 1$, then A^k converges to the zero matrix as $k \to \infty$. In summary, the following statements are all equivalent:

1. $\rho(A) < 1$
2. $\lim_{k\to\infty} \|A^k\| = 0$, for any natural norm $\|\cdot\|$
3. $\lim_{k\to\infty} (A^k)_{ij} = 0$, $i, j = 1, 2, \ldots, n$
4. $\lim_{k\to\infty} A^k\mathbf{x} = \mathbf{0}$

We will see that these results are very useful for analyzing the convergence behavior of various iterative methods for solving systems of linear equations.