Linear Algebra (end)

We now investigate some general ideas about when diagonalization is possible. In the generic case, eigenvalues will not be repeated. Is this condition sufficient to ensure that the matrix is diagonalizable? The answer is yes and results from the following

Theorem: Assume all eigenvalues λ_1, λ_2,.., λ_n of the linear map T are distinct. Then the associated eigenvectors v_1, v_2,.., v_n are linearly independent.

Proof: (by contradiction) assume that the eigenvectors are linearly dependent. Then, there is a smallest k ($1 \leq k \leq n$) such that the list $\{v_1,.., v_k\}$ is linearly independent but the list $\{v_1,.., v_k, v_{k+1}\}$ is linearly dependent. Therefore (see homework problem #1) there must exist a_1,.., a_k, not all zero, such that

$$v_{k+1} = \mathsf{a}_1 v_1 + .. + \mathsf{a}_k v_k \tag{1}$$

Apply T on both sides of (1) to obtain

$$T v_{k+1} = \mathsf{a}_1 T v_1 + .. + \mathsf{a}_k T v_k$$

or

$$\lambda_{k+1} v_{k+1} = \mathsf{a}_1 \lambda_1 v_1 + .. + \mathsf{a}_k \lambda_k v_k \tag{2}$$

Now, multiply (1) through by λ_{k+1} and subtract from (2) to obtain

$$0 = \mathsf{a}_1(\lambda_{k+1} - \lambda_1) v_1 + .. + \mathsf{a}_k(\lambda_{k+1} - \lambda_k) v_k \tag{3}$$

Since all $(\lambda_{k+1} - \lambda_i)$ in (3) are non-zero, the $\mathsf{a}_i(\lambda_{k+1} - \lambda_i)$ are not all zero and the first k eigenvectors would be linearly dependent, a contradiction! So all n eigenvectors are linearly independent. *QED*

Clearly, if the n eigenvectors are linearly independent, they form a basis for the n-dimensional vector space **V**. So, the matrix B (in the original basis $\mathcal{E}$) made up of the n eigenvectors (as columns) is nonsingular and we can write $\mathrm{B}^{-1}\mathrm{AB} = \Lambda$ (with diagonal elements λ_i).

If eigenvalues are repeated, the situation is murky. Often, the matrix A won't be diagonalizable, but this is not always true. For instance, the $n \times n$ identity I has the sole eigenvalue 1 (with multiplicity n) and it is evidently diagonalizable.

More is known on eigenvalues and eigenvectors for specific classes of matrices. We now briefly study four classes: (1) orthogonal, (2) real symmetric, (3) unitary, and (4) hermitian.

The matrix Q is orthogonal if it has real coefficients and its column (or row) vectors are orthonormal. In other words $Q^T Q = I$. Suppose λ is an eigenvalue of Q with corresponding eigenvector v, both possibly complex although Q is real. We can write $Qv = \lambda v$ and therefore, by conjugating and transposing this equality:

$$Q\overline{v} = \overline{\lambda}\,\overline{v} \quad \text{and} \quad v^T Q^T = \lambda v^T \tag{4}$$

Multiplying term by term in (4) yields: $v^T Q^T Q\overline{v} = \lambda\overline{\lambda}\, v^T\overline{v}$ or, since $Q^T Q = I$ and $v^T\overline{v} = \|v\|^2 \neq 0$, $\lambda\overline{\lambda} = 1$. This means that $\lambda = e^{i\theta}$. So, an orthogonal matrix has complex eigenvalues on the unit circle in **C**. A typical orthogonal matrix reads (check it):

$$Q = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

Its eigenvalues are $\lambda = e^{\pm i\theta}$.

The matrix S is real symmetric if it has real coefficients and $S^T = S$. Suppose λ is eigenvalue with eigenvector v. We can write $Sv = \lambda v$. Conjugating and transposing this equality yields

$$S\overline{v} = \overline{\lambda}\,\overline{v} \quad \text{and} \quad v^T S^T = v^T S = \lambda v^T \tag{5}$$

Multiplying the first equality in (5) by v^T and substituting the second equality yields $v^T S\overline{v} = \lambda v^T\overline{v} = \overline{\lambda}\, v^T\overline{v}$. It follows that $\lambda = \overline{\lambda}$. So, all eigenvalues of a real symmetric matrix are real. But there is more: suppose that we have two distinct eigenvalues $\lambda_i \neq \lambda_j$ of the same symmetric matrix S and two associated eigenvectors v_i and v_j. We can write $Sv_i = \lambda_i v_i$ and $Sv_j = \lambda_j v_j$. Multiplying the first equality by v_j^T and the second by v_i^T and transposing the latter yields

$$v_j^T S v_i = \lambda_i v_j^T v_i = (v_i^T S v_j)^T = \lambda_j v_j^T v_i$$

Since $\lambda_i \neq \lambda_j$ we must have $v_j^T v_i$ and the eigenvectors are orthogonal! In fact, we may even choose normal eigenvectors. So, if all eigenvalues are distinct, the eigenvectors form an orthonormal basis and S is similar to a real diagonal matrix with a similarity matrix Q (made up of the orthonormal eigenvectors) being an *orthogonal* matrix: $Q^{-1}SQ = \Lambda$. In fact, this results holds *without* the restriction that the eigenvalues are all distinct. As an example, consider

$$S = \begin{pmatrix} 1 & \sqrt{2} \\ \sqrt{2} & 0 \end{pmatrix}$$

The characteristic polynomial reads $p(\lambda) = \lambda(\lambda - 1) - 2 = (\lambda + 1)(\lambda - 2)$. The matrix of normalized eigenvectors reads (with $Q^T Q = I$ and $Q^T SQ = \Lambda$):

$$Q = \begin{pmatrix} \frac{-1}{\sqrt{3}} & \sqrt{\frac{2}{3}} \\ \sqrt{\frac{2}{3}} & \frac{1}{\sqrt{3}} \end{pmatrix}$$

The matrix U is unitary if it has complex coefficients and $U^*U = I$ with notation $U^* = (\overline{U})^T$, where $\overline{U}$ is the complex conjugate matrix of U (the matrix U^* is often called the *transconjugate* of U). This is a complex generalization of the orthogonal case and the same result holds: all eigenvalues are of the form $\lambda = e^{i\theta}$ (see homework).

The matrix H is hermitian if it has complex coefficients and $H^* = H$. This is a generalization of the real symmetric case with a similar result: all eigenvalues are real and H is unitarily similar to a real diagonal matrix: $U^*HU = \Lambda$ (see homework).

Homework

1. Prove equality (1) under the given assumptions.

2. Prove that the trace of a diagonalizable matrix is the sum of its eigenvalues (note: this is true even if it not diagonalizable).

3. Show that a triangular matrix has its eigenvalues on the diagonal.

4. Show that all eigenvalues of a unitary matrix have magnitude 1. Hint: follow steps similar to the orthogonal case.

5. Show that if all eigenvalues of a hermitian matrix are distinct, then it is unitarily equivalent to a real diagonal matrix. Hint: follow steps similar to the symmetric case.