More about Eigen Structure and Diagonalization

Exercise: Let A be an $n \times n$ involutory matrix (i.e. $A^2 = I$), what are the eigenvalues of A?

Exercise: Let $f(\lambda)$ be the characteristic polynomial of an $n \times n$ matrix A. Find $f(\lambda)$.

Remarks: Let A and B be $n \times n$ matrices.

(1) We said that if A and B are similar, then they have the same eigenvalues. In particular, if A is diagonalizable, then A is similar to a diagonal matrix D. Thus, A and D have the same eigenvalues which means the diagonal elements of D are the eigenvalues of A. Now if A is diagonalizable, then there exists a nonsingular matrix P such that $P^{-1}AP = D$. In this case, we say P diagonalizes A or A is diagonalizable via P. Now after explaining the relationship between A and D (they have the same eigenvalues), the questions that arise are:

(a) What is the relationship between A and P?

(b) Are P and D unique?

The answer to this first is that the columns of P are eigenvectors of A. The answer to the second is P is not unique and D is not necessarily unique. In fact, if you multiply P by any nonzero number, then the resulting matrix will diagonalize A. Also, if you change the order of the columns of P, then the resulting matrix will diagonalize A. If you do so, D also may change (will have the same elements as the original, but the location of these elements may change), which means D is not necessarily unique.

Remark: To diagonalize an $n \times n$ matrix A that has n linearly independent eigenvectors, find n linearly independent eigenvectors $x_1, x_2, \cdots, x_n$, of A. Then form the matrix $P = [x_1 \ x_2 \ \cdots \ x_n]$. Now $P^{-1}AP = D$, where D is a diagonal matrix whose diagonal elements are the eigenvalues of A corresponding to the eigenvalues of $x_1, x_2, \cdots, x_n$.

(2) Now recall that the eigenvalues of Hermitian matrices and the diagonal elements are real and the eigenvalues of skew-Hermitian matrices and the diagonal elements have zero real parts. Also, recall that if M is unitary, then $|\det(M)| = 1$ (note $\det(M)$ can be complex), M is normal, $\|Mx\|_2 = \|x\|_2$, $\forall x \in \mathbb{C}^n$, and if λ is an eigenvalue of M, then $|\lambda| = 1$. Moreover, any two distinct columns of M are orthogonal (i.e. $M(*, i) \cdot M(*, j) = M(*, i)^H M(*, j) = 0$, when $i \neq j$, and each one of them is a unit vector). Here are more facts about these matrices:

(a) (Schur's Theorem) If A is an $n \times n$ matrix, then there exists a unitary matrix M such that $M^H AM = U$, where U is upper-triangular. Moreover, A and U have the same eigenvalues.

(b) From the above, note that we can write $A = MUM^H$ (*proof: exercise*). This is called the Schur decomposition of A or Schur normal form.

(c) From Schur's theorem: If A is an $n \times n$ hermitian or a skew-Hermitian matrix, then there exists a unitary matrix M such that $M^H AM = D$, where D is diagonal. Thus, A is diagonalizable, and we say in this case A is unitarily diagonalizable.

Proof of the Hermitian Case: By Schur's Theorem, there exists a unitary matrix M and an upper-triangular matrix U such that $M^H AM = U$. Now take the Hermitian transpose of both sides to get $M^H A^H (M^H)^H = U^H$. Thus, $M^H AM = U^H$. But, $M^H AM = U$ also. Therefore, $U^H = U$, which implies U is diagonal. The skew-Herimitian case is similar.

(d) U in Schur's decomposition is diagonal iff A is normal. Moreover, when U is normal, the rows of M are eigenvetors of A. Thus, A is unitarily diagonalizable iff A is normal. Moerover, A is normal iff A has a complete orthonormal set of eigenvectors.

(e) From the previous part, if you take A to be real symmetric, then there exists an orthogonal matrix M such that $M^T AM = D$, where D is diagonal. Thus, A is diagonalizable, and we say in this case A is orthogonally diagonalizable (*proof: exercise*). The eigenvalues of a real symmetric matrix are real and eigenvectors are real. Also, eigenvectors corresponding to different eigenvalues are orthogonal.

(f) If A is Herimitian, then eigenvectors that correspond to different eigenvalues are orthogonal.

Proof: Let (λ_1, x) and (λ_2, y) be two eigenpairs of A where $\lambda_1 \neq \lambda_2$. We have to prove that $x^H y = 0$. Now consider

$$(Ax)^H y = x^H A^H y = x^H Ay = \lambda_2 x^H y.$$

$$(Ax)^H y = (y^H Ax)^H = (\lambda_1 y^H x)^H = \lambda_1 x^H y.$$

Therefore, $\lambda_1 x^H y = \lambda_2 x^H y$, which implies $\lambda_1 x^H y - \lambda_2 x^H y = (\lambda_1 - \lambda_2)x^H y = 0,$. Since $\lambda_1 \neq \lambda_2$, then $x^H y = 0$.

(3) A is orthogonaly diagonalizable iff A has n orthonormal eigenvectors iff A is real symmetric.

(4) (Cayley-Hamilton Theorem) Every matrix satisfies its characteristic equation.

Transforming Complex Hermitian Eigenvalue Problems to Real Ones

Let $C = A + iB$ be a complex Hermitian matrix, where A and B are real $n \times n$ matrices and let $(\lambda, z = x + iy)$ be an eigenpair of C, where x and y are in $\mathbb{R}^n$ (recall that λ is real because C is Hermitian). Now, $(A+iB)(x+iy) = \lambda(x+iy)$ if and only if

$$\begin{bmatrix} A & -B \\ B & A \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \lambda \begin{bmatrix} x \\ y \end{bmatrix}.$$

Note that since C is Hermitian, then A is symmetric and B is skew-symmetric. Hence, the matrix $\begin{bmatrix} A & -B \\ B & A \end{bmatrix}$ is real symmetric. Thus, we managed to reduce a complex Hermitian eigenvalue problem of order n to a real symmetric eigenvalue problem of order $2n$.

The companion Matrix

Let A be the $n \times n$ matrix such that $a(k, k+1) = 1$, $k = 1, 2, \cdots, n-1$, and $a_(n, k) = -c_{k-1}$, $k = 1, 2, \cdots, n$. By expanding the determinant of $A - \lambda I$ across the last row, you'll find out that the characterestic polynomial of A is

$$p(\lambda) = (-1)^n \left(x^n + \sum_{k=1}^{n} c_{k-1} x^{k-1} \right).$$

The matrix A is called the *companion* matrix of the polynomial $p(\lambda)$. The companion matrix is sometimes defined to be the traspose of the matrix above and it satisfies the following: $Ae_k = e_{k+1}$, $k = 1, 2, \cdots, n-1$, and $Ae_n = [-c_0 \ \ -c_1 \ \ \cdots \ \ -c_{n-1}]^T$.

Definition: The *spectral radius* of an $n \times n$ matrix A, denoted $\rho(A)$, is the maximum eigenvalue of A in magnitude; i.e. if the eigenvalues of A are $\lambda_1, \lambda_2, \cdots, \lambda_n$, then $\rho(A) = \max_i |\lambda_i|$.

Definition: Let A be an $m \times n$ matrix and let B be the matrix such that $b_{ij} = |a_{ij}|$.

(1) The one-norm of A, denoted $\|A\|_1$, is the maximum column sum of B.
(2) The ∞-norm, denoted $\|A\|_\infty$, of A is the maximum rwo sum of B.
(3) The two-norm (or spectral norm) of A, denoted $\|A\|_2$, is the non-negative square root of $\rho(A^T A)$. Note that $A^T A$ is square and symmetric.
(4) The Frobenius norm of A, denoted $\|A\|_F$, is $\sqrt{\sum_{i=1}^{m} \sum_{j=1}^{n} |a_{ij}|^2}$.

Remark: There are properties and equivalent definitions for matrix norms, but we don't have time to go over that.

Theorem (Gerschgorin): Let A be an $n \times n$ matrix and define the disks

$$D_k = \{z \in \mathbb{C} \mid |z - a_{kk}| \leq \sum_{j \neq k} |a_{kj}|, \ k = 1, \ 2, \ \cdots, \ n.$$

If λ is an eigenvalue of A, then λ is located in $\bigcup_{k=1}^{n} D_k$.