Some Basic Matrix Theorems

Richard E. Quandt
Princeton University

Definition 1. Let A be a square matrix of order n and let λ be a scalar quantity. Then $\det(\mathrm{A}-\lambda\mathrm{I})$ is called the characteristic polynomial of A.

It is clear that the characteristic polynomial is an n^{th} degree polynomial in λ and $\det(\mathrm{A}-\lambda\mathrm{I}) = 0$ will have n (not necessarily distinct) solutions for λ.

Definition 2. The values of λ that satisfy $\det(\mathrm{A}-\lambda\mathrm{I}) = 0$ are the characteristic roots or eigenvalues of A.

It follows immediately that for each λ that is a solution of $\det(\mathrm{A}-\lambda\mathrm{I}) = 0$ there exists a nontrivial x (i.e., $x \neq 0$) such that

$$(A - \lambda I)x = 0. \tag{1}$$

Definition 3. The vectors x that satisfy Eq.(1) are the characteristic vectors or eigenvectors of A.

Now consider a particular eigenvalue λ and its corresponding eigenvector x, for which we have

$$\lambda x = Ax. \tag{2}$$

Premultiply (2) by an arbitrary nonsingular matrix P we obtain

$$\lambda Px = PAx = PAP^{-1}Px, \tag{3}$$

and defining $Px = y$,

$$\lambda y = PAP^{-1}y. \tag{4}$$

Hence λ is an eigenvalue and y is an eigenvector of the matrix PAP^{-1}.

Definition 4. The matrices A and PAP^{-1} are called similar matrices.

Exercise 1. We have shown above that any eigenvalue of A is also an eigenvalue of PAP^{-1}. Now show the converse, i.e., that any eigenvalue of PAP^{-1} is also an eigenvalue of A.

Definition 5. A matrix A is symmetric if $A = A'$.

Theorem 1. *The eigenvalues of symmetric matrices are real.*

Proof. A polynomial of n^{th} degree may, in general, have complex roots. Assume then, contrary to the assertion of the theorem, that λ is a complex number. The corresponding eigenvector x may have one or more complex elements, and for this λ and this x we have

$$Ax = \lambda x. \tag{5}$$

Both sides of Eq. (5) are, in general, complex, and since they are equal to one another, their complex conjugates are also equal. Denoting the conjugates of λ and x by $\overline{\lambda}$ and $\overline{x}$ respectively, we have

$$A\overline{x} = \overline{\lambda}\overline{x}, \tag{6}$$

since $\overline{(a+bi)(c+di)} = \overline{ac - bd + (ad+bc)i} = ac - bd - (ad+bc)i = (a-bi)(c-di)$. Premultiply (5) by $\overline{x}'$ and premultiply (6) by x' and subtract, which yields

$$\overline{x'}Ax - x'A\overline{x} = (\lambda - \overline{\lambda})\overline{x}'x. \tag{7}$$

Each term on the left hand side is a scalar and and since A is symmetric, the left hand side is equal to zero. But $\overline{x}'x$ is the sum of products of complex numbers times their conjugates, which can never be zero unless all the numbers themselves are zero. Hence λ equals its conjugate, which means that λ is real. ∎

Theorem 2. *The eigenvectors of a symmetric matrix A corresponding to different eigenvalues are orthogonal to each other.*

Proof. Let $\lambda_i \neq \lambda_j$. Substitute in Eq. (5) first λ_i and its corresponding eigenvector x_i, and premultiply it by x_j', which is the eigenvector corresponding to λ_j. Then reverse the procedure and substitute in (5) the j^{th} eigenvalue and eigenvector and premultiply by x_i'. Subtracting the two results from one another yields $(\lambda_i - \lambda_j)x_i'x_j = 0$, from which it follows that $x_i'x_j = 0$. ∎

Corollary 1. *If all the eigenvalues of a symmetric matrix A are distinct, the matrix X, which has as its columns the corresponding eigenvectors, has the property that $X'X = I$, i.e., X is an orthogonal matrix.*

Proof. To prove this we need merely observe that (1) since the eigenvectors are nontrivial (i.e., do not have all zero elements), we can replace each eigenvector by a corresponding vector which is obtained from the original one by dividing each of its elements by the squareroot of the sum of squares of its elements—thus insuring that each of these vectors has length 1; and (2) the n vectors are mutually orthogonal and hence form a orthonormal basis in n-space.

Theorem 3. *If λ_i is a repeated root with multiplicity $m \geqq 2$, then there exist m orthonormal eigenvectors corresponding to λ_i.*

Proof. First, we note that corresponding to λ_i there will be at least one eigenvector x_i. For any arbitrary nonzero vector x_i one can always find an additional $n-1$ vectors y_j, $j = 2, \ldots, n$, so that x_i, together with the $n-1$ y-vectors forms an orthonormal basis. Collect the y vectors in a matrix Y, i.e.,

$$Y = [y_2, \ldots, y_n],$$

and define

$$B = [\, x_i \quad Y \,]. \tag{8}$$

Then

$$B'AB = \begin{bmatrix} \lambda_i x_i' x_i & x_i' AY \\ \lambda_i Y' x_i & Y'AY \end{bmatrix} = \begin{bmatrix} \lambda_i & 0 \\ 0 & Y'AY \end{bmatrix} \tag{9}$$

since (1) the products in the first column under the $_{11}$ element are products of orthogonal vectors, and (2) replacing in the first row (other than in the $_{11}$ element) the terms $x_i'A$ by $\lambda_i x_i'$ also leads to products of orthogonal vectors. B is an orthogonal matrix, hence its transpose is also its inverse. Therefore A and $B'AB$ are similar matrices (see Definition 4) and they have the same eigenvalues. From (9), the characteristic polynomial of $B'AB$ can be written as

$$\det(B'AB - \lambda I_n) = (\lambda_i - \lambda)\det(Y'AY - \lambda I_{n-1}). \tag{10}$$

If a root, say λ_i, has multiplicity $m \geqq 2$, then in the factored form of the polynomial the term $(\lambda_i - \lambda)$ occurs m times; hence if $m \geqq 2$, $\det(Y'AY - \lambda_i I_{n-1}) = 0$, and the null space of $(B'AB - \lambda_i I_n)$ has dimension greater than or equal to 2. In particular, if $m = 2$, the null space has dimension 2, and there are two linearly independent and orthogonal eigenvectors in this nullspace.[1] If the multiplicity is greater, say 3, then there are at least two orthogonal eigenvectors x_{i_1} and x_{i_2} and we can find another $n-2$ vectors y_j such that $[x_{i_1}, x_{i_2}, y_3, \ldots, y_n]$ is an orthonormal basis and repeat the argument above.

It also follows that if a root has multiplicity m, there cannot be more than m orthogonal eigenvectors corresponding to that eigenvalue, for that would lead to the conclusion that we could find more than n orthoganl eigenvectors, which is not possible. ∎

[1] Note that any set of n linearly independent vectors in n-space can be transformed into an orthonormal basis the Schmidt orthogonalization process.