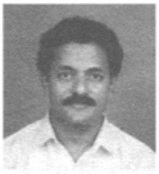
Multiplicities of an Eigenvalue: Some Observations

M Thamban Nair



M Thamban Nair is an Associate Professor of Mathematics at Indian Institute of Technology Madras, Chennai. His research interest is in the area of numerical functional analysis; in particular, spectral approximation of operators, operator equations, and inverse and ill-posed problems.

Another title for this article could be '*What makes a matrix non-diagonalizable?*' Of course, the question can be posed not only for matrices, but also for linear transformations on any linear space.

Recall that an $n \times n$ matrix A is said to be *diagonalizable* if there exists an $n \times n$ diagonal matrix D and an $n \times n$ invertible matrix U such that

$$A = UDU^{-1}.$$

We may observe that, if D is a diagonal matrix, say with diagonal entries $\lambda_1, \ldots, \lambda_n$, and U is an invertible matrix with columns $u_1, \ldots, u_n$, then the relation $A = UDU^{-1}$, i.e., $AU = UD$, is same as

$$Au_j = \lambda_j u_j, \qquad j \in \{1, \ldots, n\}.$$

Thus, if A is diagonalizable, then A has n linearly independent eigenvectors. Conversely, if A has n linearly independent eigenvectors $u_1, \ldots, u_n$ and if $\lambda_1, \ldots, \lambda_n$ are the associated eigenvalues, possibly repeated, then taking U with columns $u_1, \ldots, u_n$, and D as the diagonal matrix with diagonal entries $\lambda_1, \ldots, \lambda_n$, we have $AU = UD$, i.e., $A = UDU^{-1}$. In fact, D is the matrix representation of A with respect to the basis consisting of columns of U.

From the above discussion we have the following characterization for diagonalizability of A:

An $n \times n$ matrix A is diagonalizable if and only if A has n linearly independent eigenvectors.

Keywords
Eigenvalues, algebraic multiplicity, geometric multiplicity, ascent, generalized eigenspace.

A matrix is diagonalizable if and only if ascent of every eigenvalue is one, that is, if and only if for each eigenvalue, the geometric multiplicity and the algebraic multiplicity are equal.

In view of the above characterization, the question of diagonalizability can be posed for any linear transformation on a finite dimensional vector space.

Now let us look at the issue in a slightly different manner. Suppose $\lambda_1, \ldots, \lambda_k$ are the distinct eigenvalues of an $n \times n$ matrix A. By the above discussion, we can infer that A is diagonalizable if and only if for each $j = 1, \ldots, k$, there are linearly independent eigenvectors $u_{j,1}, \ldots, u_{j,n_j}$ of A associated with λ_j such that

$$n_1 + \ldots + n_k = n,$$

and in that case, columns of U are the vectors

$$u_{1,1}, \ldots, u_{1,n_1}, \ldots, u_{k,1}, \ldots, u_{k,n_k},$$

and diagonal elements of D are $\lambda_1, \ldots, \lambda_k$ with each λ_j repeated n_j times.

Such a nice situation does not prevail if one of the eigenvalues, say λ_j, is *defective*, in the sense that n_j is not large enough so as to satisfy the relation $n_1 + \ldots + n_k = n$. A simple example of a matrix which is not diagonalizable is the 2×2 matrix

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

It is easily seen that $\lambda = 1$ is the only eigenvalue of A and there is only one linearly independent eigenvector associated with this eigenvalue. In this example, $\lambda = 1$ is a defective eigenvalue of A.

In this article we define certain quantities called *geometric multiplicity, algebraic multiplicity* and *ascent* of an eigenvalue, and then prove a generalized form of the *diagonalization theorem* which, in particular, show that a matrix is diagonalizable if and only if ascent of every eigenvalue is one, that is, if and only if for each eigenvalue, the geometric multiplicity and the algebraic multiplicity are equal. We shall do this in the context of a linear transformation on a finite dimensional vector space.

First some definitions:

Let X be a finite dimensional vector space over the field $\mathbb{C}$ of complex numbers, and let $T : X \to X$ be a linear transformation, i.e.,

$$T(\alpha x + \beta y) = \alpha T(x) + \beta T(y)$$

for every $x, y \in X$ and for every $\alpha, \beta \in \mathbb{C}$. An element $\lambda \in \mathbb{C}$ is said to be an *eigenvalue* of T if there exists a nonzero $x \in X$ such that

$$Tx = \lambda x,$$

and in that case x is called an *eigenvector* of T. (Here, and below, for $x \in X$, we may also use the notation Tx for the element $T(x)$ in X.)

Two of the important subspaces associated with $T: X \to X$ are its *kernel* and its *range*, defined by

$$\ker(T) := \{x \in X : Tx = 0\},$$

$$\text{range}(T) := \{Tx : x \in X\},$$

respectively.

A subspace M of X is said to be *invariant* under T if $Tx \in M$ for every $x \in M$.

Clearly, kernel and range of T are invariant under T.

The following definition is motivated by our discussion on diagonalizablity of a matrix: The linear transformation $T : X \to X$ is said to be *diagonalizable* if X has a basis consisting of eigenvectors of T.

Clearly, $\lambda \in \mathbb{C}$ is an eigenvalue of T if and only if the transformation $T - \lambda I$ is not injective, if and only if the kernel of $T - \lambda I$, is not the zero space. Thus, eigenvectors of T corresponding to an eigenvalue λ are the nonzero elements of $\ker(T - \lambda I)$.

The subspace $\ker(T - \lambda I)$ is called the *eigenspace* of T corresponding to the eigenvalue λ, and the dimension of the eigenspace $\ker(T - \lambda I)$ is called the *geometric multiplicity* of λ.

If $\lambda_1, \ldots, \lambda_k$ are the distinct eigenvalues of T, then it can be easily seen that T is diagonalizable if and only if

$$\dim X = \dim \ker(T - \lambda_1 I) + \ldots + \dim \ker(T - \lambda_k I).$$

This is equivalent to saying that

$$X = \ker(T - \lambda_1 I) + \ldots + \ker(T - \lambda_k I)$$

with

$$\ker(T - \lambda_i I) \cap \ker(T - \lambda_j I) = \{0\} \quad \text{for} \quad i \neq j.$$

For subsets M, N of X, we used the notation $M + N$ for the set

$$\{x + y : x \in M,\ y \in N\}.$$

Now suppose that $\lambda \in \mathbb{C}$ is an eigenvalue of T, and for $j = 1, 2, \ldots,$ let us denote

$$K_j = \ker(T - \lambda I)^j.$$

Then we observe the following :

(i) $\{0\} \subseteq K_1 \subseteq K_2 \ldots \subseteq X$.

(ii) There exists j such that $K_j = K_{j+1}$

(iii) If $K_j = K_{j+1}$ for some j, then $K_j = K_{j+i}$ for every $i = 1, 2, \ldots$.

Clearly (i) follows from the definition of K_j's, and (ii) is a consequence of the finite dimensionality of X. To see (iii), suppose $K_j = K_{j+1}$ for some j. Now if $x \in K_{j+2}$, i.e., if $(T - \lambda I)^{j+2}x = 0$, then we have $(T - \lambda I)x \in K_{j+1} = K_j$, so that $(T - \lambda I)^{j+1}x = 0$, i.e., $x \in K_{j+1}$. Thus we have proved that $K_{j+2} \subseteq K_{j+1}$. Now using (i), we have $K_{j+2} = K_{j+1}$. Continuing the above argument, we obtain $K_j = K_{j+i}$ for every $i \in \{1, 2, \ldots\}$.

From the above observations, it is obvious that there exists a positive integer ℓ such that

$$K_{\ell-1} \neq K_\ell = K_{\ell+i} \text{ for every } \; i = 1, 2, \ldots.$$

This positive integer ℓ is called the *ascent* of λ, the subspace K_ℓ is called the *generalized eigenspace* and elements of K_ℓ are called the *generalized eigenvectors* of T corresponding to the eigenvalue λ. The dimension of K_ℓ is called the *algebraic multiplicity* of λ.

Let g_λ, ℓ_λ and m_λ, respectively, be the geometric multiplicity, ascent and the algebraic multiplicity of an eigenvalue λ of T. Then, from the observations (i)–(iii) above, it is clear that

$$g_\lambda \leq m_\lambda \quad \text{and} \quad \ell_\lambda \leq m_\lambda.$$

Also, we observe that

$$g_\lambda = m_\lambda \quad \text{if and only if} \quad \ell_\lambda = 1.$$

In fact, we have the following result ([1]):

Theorem 1 *Suppose λ is an eigenvalue of a linear transformation $T : X \to X$ with geometric multiplicity g, algebraic multiplicity m, and ascent ℓ. Then*

$$\ell + g - 1 \leq m \leq \ell g.$$

Proof. For each $i \in \mathbb{N}$, let $K_i = \ker(T - \lambda I)^i$ and $g_i = \dim K_i$. Then we know that

$$K_1 \subset K_2 \subset \ldots \subset K_\ell = K_{\ell+j} \qquad \forall j \in \mathbb{N}.$$

From this, it is clear that

$$g_1 + i - 1 \leq g_i \qquad \forall i \in \{1, \ldots, \ell\}.$$

In particular, since $g_1 = g$ and $g_\ell = m$, we have $g + \ell + 1 \leq m$. Next, we prove

$$g_i \leq g_{i-1} + g_1. \tag{1}$$

Once, this is proved, it follows that $g_2 \leq 2g_1 = 2g$, $g_3 \leq g_2 + g_1 \leq 3g$, etc., $m = g_\ell \leq \ell g$.

For the proof of (1), write K_i as

$$K_i = K_{i-1} + Y_i \quad \text{with} \quad K_{i-1} \cap Y_i = \{0\}.$$

Hence, it is enough to show that $\dim Y_i \leq g$. So, suppose that $u_1, \ldots, u_k$ be a basis of Y_i. Then we notice that $(T - \lambda I)^{i-1}u_1, \ldots, (T - \lambda I)^{i-1}u_k$ are linearly independent. Indeed, if $\alpha_1, \ldots, \alpha_k$ are scalars such that $\sum_{j=1}^k \alpha_j (T - \lambda I)^{i-1}u_j = 0$, then, we have $\sum_{j=1}^k \alpha_j u_j \in K_{i-1}$. This happens only if $\sum_{j=1}^k \alpha_j u_j = 0$, since $\sum_{j=1}^k \alpha_j u_j \in Y_i$. Thus, it follows that $\alpha_j = 0$ for $j = 1, \ldots, k$. Thus, $(T - \lambda I)^{i-1}u_1, \ldots, (T - \lambda I)^{i-1}u_k$ are linearly independent. Also, since $u_j \in K_i$, $(T - \lambda I)^{i-1}u_j \in K_1$ for each $j = 1, \ldots, k$. Thus, we have proved that $k \leq g$.
This completes the proof. □

Our attempt is to prove that, if $\lambda_1, \ldots, \lambda_k$ are the distinct eigenvalues of T with algebraic multiplicities $m_1, \ldots, m_k$ and ascents $\ell_1, \ldots, \ell_k$ respectively, then

$$m_1 + \ldots + m_k = \dim X, \tag{2}$$

so that T is diagonalizable if and only if $\ell_j = 1$ for every $j = 1, \ldots, k$. For showing this, it is enough to prove that

$$X = \ker(T - \lambda_1)^{\ell_1} + \ldots + \ker(T - \lambda_k)^{\ell_k}$$

with

$$\ker(T - \lambda_i I)^{\ell_i} \cap \ker(T - \lambda_j I)^{\ell_j} = \{0\} \quad \text{for} \quad i \neq j.$$

All this discussion would be in vain if T has no eigenvalues. So first let us prove

One can prove the existence of an eigenvalue without using the concept of the determinant.

Theorem 2 *Every linear transformation $T : X \to X$ has at least one eigenvalue.*

Usually this result is proved by invoking the concept of *determinant* of a linear transformation or matrix. The following simple and elegant proof which does not depend on the concept of determinant is due to Sheldon Axler [2].

Proof of Theorem 2. Suppose $\dim X = n$, and $x \in X$ is a nonzero element. Since the set $\{x, T(x), \ldots, T^n(x)\}$ consisting of $n+1$ elements is linearly dependent, there exists complex numbers $\alpha_0, \alpha_1, \ldots, \alpha_n$ with at least one of them nonzero, such that

$$\alpha_0 x + \alpha_1 T(x) + \ldots + \alpha_n T^n(x) = 0.$$

Writing

$$p(t) = \alpha_0 + \alpha_1 t + \ldots + \alpha_n t^n, p(T) = \alpha_0 I + \alpha_1 T + \ldots + \alpha_n T^n,$$

we have

$$p(T)(x) = 0.$$

Suppose $k = \max\{i : \alpha_i \neq 0\}$. Then, by Fundamental Theorem of Algebra, there exists $\lambda_1, \ldots, \lambda_k$ in $\mathbb{C}$ such that

$$p(t) = \alpha_k (t - \lambda_1) \ldots (t - \lambda_k).$$

Thus we see that

$$(T - \lambda_1 I) \ldots (T - \lambda_k I)(x) = p(T)(x) = 0.$$

The above relation shows that at least one of $T - \lambda_1 I, \ldots,$ $T - \lambda_k I$ is not injective, so that at least one of $\lambda_1, \ldots, \lambda_k$ is an eigenvalue of T. □

As a first step towards the proof of (2) we prove:

Theorem 3 *Let λ be an eigenvalue of T with ascent ℓ. Let $K = \ker(T - \lambda I)^\ell$ and $R = \text{range}(T - \lambda I)^\ell$. Then*

(i) *K and R are invariant under T, and*

(ii) *$X = K + R$ with $K \cap R = \{0\}$.*

(iii) *X is spanned by generalized eigenvectors of T.*

(iv) *λ is the only eigenvlaue of $T|_K$.*

Proof. It is easy to see that

$$T(T - \lambda I)^j = (T - \lambda I)^j T \quad \forall j \in \mathbb{N}.$$

Hence, if $x \in K$, then

$$(T - \lambda I)^\ell (Tx) = T(T - \lambda I)^\ell (x) = 0,$$

so that $Tx \in K$ showing that K is invariant under T. To see that R is invariant under T, let $y \in R$, and let $x \in X$ such that $y = (T - \lambda I)^\ell x$. Then we have

$$Ty = T(T - \lambda I)^\ell (x) = (T - \lambda I)^\ell (Tx) \in R,$$

showing that R is invariant under T. Thus, proofs of (i) and (ii) are over.

Next suppose that $x \in K \cap R$. Then we have

$$(T - \lambda I)^\ell x = 0 \quad \text{and} \quad x = (T - \lambda I)^\ell u$$

for some $u \in X$, so that

$$0 = (T - \lambda I)^\ell x = (T - \lambda I)^{2\ell} u.$$

Thus $u \in \ker(T - \lambda I)^{2\ell} = \ker(T - \lambda I)^\ell$. Hence $x = (T - \lambda I)^\ell u = 0$. Thus we have proved that $K \cap R = \{0\}$. This, together with the relation $\dim X = \dim K + \dim R$ implies that $X = K + R$, proving (ii).

We show (iii) by induction on the dimension of X. Clearly the result is true if $n = 1$. Next suppose $n > 1$ and that the result is true for all vector spaces of dimension less than n. Since $K \neq \{0\}$, we see from (ii) that $R \neq X$, so that $\dim R < n$, and by induction assumption R is spanned by the generalized eigenvectors of $T|_R$. Since generalized eigenvectors of $T|_R$ are generalized eigenvectors of T as well, and elements of K are already generalized eigenvectors of T, it follows that X is spanned by the generalized eigenvectors of T.

To see (iv), first observe by Theorem 2 that $T|_K$ has at least one eigenvalue. Suppose μ is an eigenvalue of $T|_K$ and $x \neq 0$ is a corresponding eigenvector. Then we have $(T - \lambda I)x = (\mu - \lambda)x$ so that

$$0 = (T - \lambda I)^{\ell} x = (\mu - \lambda)^{\ell} x.$$

This shows that $\mu = \lambda$, as desired. □

Theorem 4 *If λ and μ are distinct eigenvalues of T with corresponding ascents ℓ_λ and ℓ_μ respectively, then*

$$\ker(T - \lambda I)^{\ell_\lambda} \cap \ker(T - \mu I)^{\ell_\mu} = \{0\}.$$

Proof. Suppose λ and μ are distinct eigenvalues of T with corresponding ascents ℓ_λ and ℓ_μ, respectively. Let us denote ℓ_λ by ℓ, and

$$K := \ker(T - \lambda I)^{\ell}, \qquad R := \mathrm{range}(T - \lambda I)^{\ell}.$$

Since, by Theorem 3 (ii), $K \cap R = \{0\}$, it is enough to show that

$$\ker(T - \mu I)^{\ell_\mu} \subseteq R.$$

We shall, in fact, show by induction that

$$\ker(T - \mu I)^{j} \subseteq R, \quad j = 1, 2, \ldots \qquad (3).$$

For proving (3), first let $j = 1$ and $x \in \ker(T - \mu I)$. Since $X = K + R$, by Theorem 3 (ii), there exists $x_1 \in K$ and

$x_2 \in R$ such that $x = x_1 + x_2$. Now, by Theorem 3(i), K and R are invariant under T, so that $(T - \mu I)x = 0$ implies $(T - \mu I)x_1 = 0$ and $(T - \mu I)x_2 = 0$. In particular, if $x_1 \neq 0$, then μ is an eigenvalue of $T|_K$, a contradiction to the fact (ref. Theorem 3(iv)) that λ is the only eigenvalue of $T|_K$. Hence $x_1 = 0$, so that $x = x_2 \in R$.

Next suppose that the result (3) is true for some $j \geq 1$. Let $x \in \ker(T - \mu I)^{j+1}$, and $x_1 \in K$ and $x_2 \in R$ are such that $x = x_1 + x_2$. Again, using the invariance of K and R under T, $(T - \mu I)^{j+1} x_1 = 0$ and $(T - \mu I)^{j+1} x_2 = 0$. In particular, using the induction hypothesis, we have

$$(T - \mu I)x_1 \subseteq K \cap \ker(T - \mu I)^j \subseteq K \cap R = \{0\}.$$

Since λ is the only eigenvalue of $T|_K$, it then follows that $x_1 = 0$, and hence $x = x_2 \in R$. Thus (3) is proved, and the proof of the theorem is complete. □

Now the following theorem is obvious from Theorem 3 (iii) and Theorem 4.

Theorem 5 *Let $\lambda_1, \ldots, \lambda_k$ be distinct eigenvalues of T with ascents $\ell_1, \ldots, \ell_k$ respectively. Then*

$$X = \ker(T - \lambda_1 I)^{\ell_1} + \ldots + \ker(T - \lambda_k I)^{\ell_k}$$

with

$$\ker(T - \lambda_i I)^{\ell_i} \cap \ker(T - \lambda_j I)^{\ell_j} = \{0\} \quad \textit{for} \quad i \neq j.$$

In particular,

$$\dim X = \dim[\ker(T - \lambda_1 I)^{\ell_1}] + \ldots + \dim[\ker(T - \lambda_k I)^{\ell_k}].$$

Note that the above theorem includes the relation (2).

As a corollary to the above theorem we have

Theorem 6 *Let $\lambda_1, \ldots, \lambda_k$ be distinct eigenvalues of T with ascents $\ell_1, \ldots, \ell_k$ respectively. Then T is diagonalizable if and only if $\ell_j = 1$ for every $j \in \{1, \ldots, k\}$.*

Finally we would like to mention a class of linear transformations on finite dimensional inner product spaces, the so called self-adjoint operators, which can be diagonalized.

Suppose that X is a finite dimensional inner product space with (positive definite, hermitian) inner product $\langle \cdot, \cdot \rangle$. Recall that a linear transformation transformation $T : X \to X$ is called a *self-adjoint operator* if

$$\langle Tx, y \rangle = \langle x, Ty \rangle \quad \forall x, y \in X.$$

Suppose $T : X \to X$ is a self-adjoint operator. If $x \in X$ and $\lambda \in \mathbb{C}$ are such that $Tx = \lambda x$, then we have

$$\lambda \langle x, x \rangle = \langle \lambda x, x \rangle = \langle Tx, x \rangle = \langle x, Tx \rangle = \langle x, \lambda x \rangle = \bar{\lambda} \langle x, x \rangle,$$

where $\bar{\lambda}$ denotes the complex conjugate of λ. From this it follows that every eigenvalue of T is a real number. We also obtain that if $x \in \ker(T - \lambda I)^2$, then

$$\langle (T - \lambda I)x, (T - \lambda I)x \rangle = \langle (T - \lambda I)^2 x, x \rangle = 0.$$

so that $\ker(T - \lambda I)^2 \subseteq \ker(T - \lambda I)$. We already know that $\ker(T - \lambda I) \subseteq \ker(T - \lambda I)^2$. Hence, $\ker(T - \lambda I)^2 = \ker(T - \lambda I)$, and consequently, ascent of every eigenvalue of T is 1. Thus we have proved the following *diagonalization theorem* for self-adjoint operators.

Theorem 7 *Suppose $T : X \to X$ is a self-adjoint operator on a finite dimensional inner product space X. Then every eigenvalue of T is real and has ascent 1. In particular, if $\lambda_1, \ldots, \lambda_k$ are distinct eigenvalues of T, then*

$$X = \ker(T - \lambda_1 I) + \ldots + \ker(T - \lambda_k I).$$

Suggested Reading

[1] B V Limaye and M T Nair, On multiplicities and ascent of an eigenvalue of a linear operator, *Mathematics Student*, Vol.64, pp.162-166, 1995.

[2] S Axler, Down with determinants!, *American Mathematical Monthly*, pp.139-154, Feb.1995.

Address for Correspondence
M Thamban Nair
Associate Professor
Department of Mathematics
Indian Institute of Technology
Chennai 600 036, India.
Email: mtnair@iitm.ac.in