On the equality of algebraic and geometric multiplicities of matrix eigenvalues

J. Ding\textsuperscript{a}, Noah H. Rhee\textsuperscript{b,*}

\textsuperscript{a} Department of Mathematics, The University of Southern Mississippi, Hattiesburg, MS 39406-5045, USA
\textsuperscript{b} Department of Mathematics and Statistics, University of Missouri-Kansas City, Kansas City, MO 64110-2499, USA

\textbf{A B S T R A C T}

We summarize seventeen equivalent conditions for the equality of algebraic and geometric multiplicities of an eigenvalue for a complex square matrix. As applications, we give new proofs of some important results related to mean ergodic and positive matrices.

© 2011 Elsevier Ltd. All rights reserved.

\textbf{1. Introduction}

Let $C^{n \times n}$ denote the space of all complex $n \times n$ matrices. For a matrix $A \in C^{n \times n}$, $N(A)$ and $R(A)$ denote the null space and the range space of $A$, respectively. We denote the set of all eigenvalues of $A$ by $\sigma(A)$. The notation $\rho(A)$ stands for the spectral radius of $A$, which is the maximum of the absolute value of all the eigenvalues. Suppose $\lambda \in \sigma(A)$. Then the \textit{algebraic multiplicity} of $\lambda$ is the number of times it is repeated as a zero of the characteristic polynomial. In other words, $\text{alg mult}_A(\lambda_i) = m_i$ if and only if $p_A(x) = (x - \lambda_1)^{m_1} \cdots (x - \lambda_s)^{m_s}$ is the characteristic polynomial for $A$. The \textit{geometric multiplicity} of $\lambda$ is $\text{dim} N(A - \lambda I)$. In other words, $\text{geom mult}_A(\lambda)$ is the maximal number of linearly independent eigenvectors associated with $\lambda$. In general we have

$$ \text{geom mult}_A(\lambda) \leq \text{alg mult}_A(\lambda).$$

The following three results are well known. For example, see [1].

1. A matrix $A \in C^{n \times n}$ is diagonalizable if and only if $\text{geom mult}_A(\lambda) = \text{alg mult}_A(\lambda)$ for each $\lambda \in \sigma(A)$.

2. For $A \in C^{n \times n}$, $\lim_{k \to \infty} A^k$ exists if and only if $\rho(A) < 1$ or else $\rho(A) = 1$, where $\lambda = 1$ is the only eigenvalue on the unit circle, and the algebraic multiplicity of $\lambda = 1$ equals its geometric multiplicity.

3. For $A \in C^{n \times n}$, the Cesáro limit

$$ \lim_{k \to \infty} \frac{I + A + \cdots + A^{k-1}}{k}$$

exists if and only if $\rho(A) < 1$ or else $\rho(A) = 1$ with the algebraic multiplicity equal to the geometric multiplicity for each eigenvalue on the unit circle.

\textsuperscript{*} Corresponding author.
\textit{E-mail address:} rheen@umkc.edu (N.H. Rhee).

0893-9659/$-$ see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.aml.2011.06.030
So it is important to know necessary and sufficient conditions under which the algebraic multiplicity equals the geometric multiplicity of a given eigenvalue.

In Section 2 we summarize seventeen equivalent conditions for the equality of algebraic and geometric multiplicities of an eigenvalue for any complex square matrix. Then in Section 3 we give some applications of the results in Section 2.

2. Equality of algebraic and geometric multiplicities

The following theorem may not be unknown. But it appears that it has not been summarized in the way we present here.

Theorem 1. Let $A \in \mathbb{C}^{n \times n}$ and let $\lambda$ be an eigenvalue of $A$. Then the following four statements are equivalent:

(a) There exist bi-orthonormal bases $\{x_1, \ldots, x_m\}$ of $N(A - \lambda I)$ and $\{y_1, \ldots, y_m\}$ of $(A^H - \bar{\lambda} I)^\perp$ in the sense that $y_j^H x_k = \delta_{jk}$, \forall $j, k = 1, \ldots, m$. Note that $m$ is the geometric multiplicity of $\lambda$. Here the superscript $H$ denotes the conjugate transpose.

(b) $\mathbb{C}^n = N(A - \lambda I) \oplus R(A - \lambda I)$.

(c) $N(A - \lambda I) \cap R(A - \lambda I) = \{0\}$.

(d) The geometric multiplicity and the algebraic multiplicity of $\lambda$ are equal.

Proof. We first prove that (a) implies (b). Suppose that there exist bi-orthonormal bases $\{x_1, \ldots, x_m\}$ of $N(A - \lambda I)$ and $\{y_1, \ldots, y_m\}$ of $(A^H - \bar{\lambda} I)^\perp$. Let $P = \sum_{j=1}^m x_j y_j^H$. Then using the bi-orthogonal property $y_j^H x_k = \delta_{jk}$, \forall $j, k = 1, \ldots, m$, we have

$$P^2 = \sum_{j=1}^m x_j y_j^H \sum_{k=1}^m x_k y_k^H = \sum_{j=1}^m \sum_{k=1}^m x_j y_j^H x_k y_k^H = \sum_{j=1}^m x_j y_j^H = P,$$

which shows that $P$ is a projection. Note that $R(P) = \text{Span}\{x_1, \ldots, x_m\} = N(A - \lambda I)$ and $N(P) = \{y_1, \ldots, y_m\} = (N(A^H - \bar{\lambda} I))^\perp = R(A - \lambda I)$, where $\perp$ denotes the orthogonal complement. Since $P$ is a projection, we have

$$\mathbb{C}^n = P \oplus N(P) = N(A - \lambda I) \oplus R(A - \lambda I).$$

Note that (b) immediately implies (c) by the definition of the direct sum.

Now we prove that (c) implies (d). Let $J$ be the Jordan canonical form of $A$. Then $A = JSJ^{-1}$ for some invertible matrix $S$. Suppose that the algebraic multiplicity of $\lambda$ is greater than the geometric multiplicity of $\lambda$. Then there is at least one Jordan block in $J$ associated with $\lambda$ which is at least $2 \times 2$. We may assume that this Jordan block is the first Jordan block in $J$. Then $A_{s1} = \lambda s_1$ and $A_{s2} = s_1 + \lambda s_2$. In other words, $(A - \lambda I)s_1 = 0$ and $(A - \lambda I)s_2 = s_1$ Here $s_1$ and $s_2$ denote the first and second columns of the matrix $S$. So

$$0 \neq s_1 \in N(A - \lambda I) \cap R(A - \lambda I),$$
a contradiction to the assumption (c). Thus we conclude that the geometric multiplicity and the algebraic multiplicity of $\lambda$ are equal.

Finally, we show that (d) implies (a). Suppose that

$$\dim N(A - \lambda I) = \dim N(A^H - \bar{\lambda} I) = m.$$ 

Let $\{x_1, \ldots, x_m\}$ be a basis of $N(A - \lambda I)$. Since the geometric multiplicity of $\lambda$ equals the algebraic multiplicity of $\lambda$, the set $\{x_1, \ldots, x_m\}$ can be extended to a Jordan basis for $A$. Let $S$ be the matrix whose columns form this Jordan basis for $A$. Then $A = JSJ^{-1}$, where $J$ is the Jordan canonical form of $A$ and the top $m \times m$ submatrix of $J$ is diag($\lambda, \ldots, \lambda$). Let $T = (S^{-1})^H$ and we denote the first $m$ columns of $T$ by $y_1, \ldots, y_m$. Then $T^H S = I$ directly implies that $y_j^H x_k = \delta_{jk}$, \forall $j, k = 1, \ldots, m$. Furthermore,

$$A^H T = (S^{-1})^H J^H S^H T = T^H S^H (S^{-1})^H = T^H,$$

which immediately implies that $A^H y_j = \bar{\lambda} y_j$ for $j = 1, 2, \ldots, m$. So $\{y_1, \ldots, y_m\}$ is a basis for $N(A^H - \bar{\lambda} I)$. \qed

In the remaining part of this section we provide several more equivalent conditions for the equality of algebraic multiplicity and geometric multiplicity of an eigenvalue. To do so, we introduce several necessary concepts. Let $A$ be a square matrix. The minimal polynomial of $A$ is the unique monic polynomial $g$ of minimal degree such that $g(A) = 0$. By the Cayley–Hamilton theorem, the minimal polynomial is a factor of the characteristic polynomial of $A$. Moreover, the zeros of the minimal polynomial are exactly the eigenvalues of $A$.

Let $z$ be a complex variable. The function $R(z, A) = (A - zI)^{-1}$ of $z$ is called the resolvent of $A$ and is analytic in the complement of $\sigma(A)$. Every eigenvalue $\lambda$ of $A$ is a pole of $R(z, A)$.

From the theory of Jordan normal forms, every square matrix is similar to its Jordan canonical form which is composed of Jordan blocks. Each Jordan block is related to an eigenvalue of $A$ as its diagonal elements and $1$s along the super-diagonal, and all other entries are zero. The size of the largest Jordan block with respect to eigenvalue $\lambda$ is exactly the multiplicity of $\lambda$ in the factorization of the minimal polynomial of $A$. 
The index \( \nu(\lambda) \) of an eigenvalue \( \lambda \) of \( A \) is the smallest nonnegative integer \( k \) such that
\[
N((A - \lambda I)^{k+1}) = N((A - \lambda I)^k).
\]

A generalized eigenvector associated with eigenvalue \( \lambda \) of \( A \) is a nonzero vector in \( N((A - \lambda I))^{\nu(\lambda)} \) that is not an eigenvector of \( A \).

We summarize more equivalent conditions in the following theorem.

**Theorem 2.** Let \( A \in \mathbb{C}^{n\times n} \) and let \( \lambda \) be an eigenvalue of \( A \). Then the following statements are equivalent to any of the statements (a) through (d) of **Theorem 1**:

(e) \( \lambda \) is a simple root of the minimal polynomial of \( A \).

(f) The derivative of the minimal polynomial of \( A \) at \( \lambda \) is nonzero.

(g) \( \lambda \) is a stable eigenvalue. That is, \( \lambda \) remains a simple root of the minimal polynomial under a small perturbation to \( A \).

(h) The order of \( \mathbb{R}(z, A) \) at the pole \( \lambda \) is 1.

(i) There are one by one Jordan blocks associated with \( \lambda \) in the Jordan canonical form of \( A \), where \( m \) is the algebraic multiplicity of \( \lambda \).

(j) The maximal size of the Jordan blocks associated with \( \lambda \) is 1.

(k) All the Jordan blocks associated with \( \lambda \) are \( 1 \times 1 \).

(l) The index of \( \lambda \) equals 1.

(m) \( N((A - \lambda I)^{k+1}) = N((A - \lambda I)^k) \) for all positive integers \( k \).

(n) \( \mathbb{R}(A - \lambda I)^{k+1} = \mathbb{R}(A - \lambda I)^k \) for all positive integers \( k \).

(o) For any eigenvector \( x \) associated with \( \lambda \), the equation \( (A - \lambda I)y = x \) has no solution.

(p) There are no generalized eigenvectors with respect to \( \lambda \).

(q) The algebraic multiplicity of \( \lambda \) is 1.

### 3. Applications

We first give an application of our theorems to the ergodic theory of general matrices. A square matrix \( A \in \mathbb{C}^{n\times n} \) of spectral radius 1 is said to be **mean ergodic** if the sequence of its Cesàro averages
\[
A_k = \frac{1}{k} \sum_{j=0}^{k-1} A^j
\]
converges to a matrix \( E \). From the identity
\[
AA_k = \frac{k+1}{k} A_{k+1} - \frac{1}{k} I = A_k A
\]
we see that the limit matrix \( E \) satisfies
\[
AE = E \quad \text{and} \quad EA = E.
\]

Moreover, from \( AE = E \) we have \( A_k E = E \). It follows by taking limit \( k \to \infty \) that \( E^2 = E \), so \( E \) is a projection matrix onto \( R(E) \) along \( N(E) \) and
\[
\mathbb{C}^n = R(E) \oplus N(E).
\]

Now we show that \( R(E) = N(A - I) \) and \( N(E) = R(A - I) \). From the first equation in (1) we have \( R(E) \subset N(A - I) \). On the other hand, for any \( x \in N(A - I) \) we have \( Ax = x \) from which \( A_k x = x \) for all \( k \). Thus \( Ex = x \) and so \( x \in R(E) \). Hence \( R(E) = N(A - I) \). From the second equation in (1) we have \( R(A - I) \subset N(E) \). Furthermore, we have
\[
\dim N(E) = n - \dim R(E) = n - \dim N(A - I) = \dim R(A - I),
\]
and hence it follows that \( N(E) = R(A - I) \). So (2) becomes
\[
\mathbb{C}^n = N(A - I) \oplus R(A - I).
\]

It may be that \( N(A - I) = \{0\} \). To prevent that from happening, we need to assume that 1 is an eigenvalue of \( A \). Now, from **Theorem 1**, we have

**Corollary 1.** If \( A \in \mathbb{C}^{n\times n} \) is mean ergodic and if \( 1 \in \sigma(A) \), then the eigenvalue 1 has the same algebraic and geometric multiplicity.

**Remark.** It is known [1, page 633] that a matrix \( A \in \mathbb{C}^{n\times n} \) is mean ergodic if and only if each eigenvalue on the unit circle has the same algebraic and geometric multiplicity. Here we offered another proof of a part of this result.
A sufficient condition for the mean ergodicity of $A$ is that the power sequence $\{\|A^k\|\}$ is uniformly bounded for some matrix norm $\| \cdot \|$ (Lemma 2.5 of [2]). For completeness we provide a proof of this result.

**Lemma 1.** Let $A \in \mathbb{R}^{n \times n}$ be such that $\rho(A) = 1$. If the sequence $\{\|A^k\|\}$ is uniformly bounded for some matrix norm $\| \cdot \|$, then $A$ is mean ergodic.

**Proof.** In this proof we need the identity

\[ AA_k - A_k = \frac{1}{k} (A^k - I), \tag{3} \]

which can be verified by a direct computation. Since the sequence $\{\|A^k\|\}$ is uniformly bounded, so is the sequence $\{\|A_k\|\}$. The Balzano-Weierstrass theorem implies that there exists a subsequence $\{A_{k_j}\}$ of $\{A_k\}$ and a matrix $E \in \mathbb{C}^{n \times n}$ such that

\[ \lim_{j \to \infty} A_{k_j} = E. \]

We now show that $\lim_{k \to \infty} A_k = E$. Suppose that there is a subsequence $\{A_{k'_j}\}$ of $\{A_k\}$ with $\lim_{j \to \infty} A_{k'_j} = E'$. Then (3) implies that $AE = E$ and $AE' = E'$, so $A_k E = E$ and $A_k E' = E'$ for all positive integers $k$. Since

\[ \|E - E'\| = \|A_{k'_j}(E - A_{k'_j}) + A_{k'_j}(A_{k'_j} - E')\| \leq \|A_{k'_j}\| \|E - A_{k'_j}\| + \|A_{k'_j}\| \|A_{k'_j} - E'\| \to 0 \]

as $j \to \infty$, we see that $E = E'$. Since all convergent subsequences of $\{A_k\}$ converge to $E$, we have $\lim_{k \to \infty} A_k = E$. \qed

A nonnegative square matrix is called a stochastic matrix if its every row sums to be 1. We have the following result.

**Corollary 2.** The algebraic multiplicity and geometric multiplicity of all the eigenvalues on the unit circle for any stochastic matrix are equal.

**Proof.** Since the product of any two stochastic matrices is a stochastic matrix, the sequence $\{\|A^k\|\}$ of a stochastic matrix $A$ is uniformly bounded by 1, where $\| \cdot \|$ is the matrix $\infty$-norm. It is a standard result that $\rho(A) = 1$. So by Lemma 1 $A$ is mean ergodic. Since $A$ is nonnegative, its spectral radius 1 is an eigenvalue, and by Corollary 1 the algebraic multiplicity and geometric multiplicity of the eigenvalue 1 are the same. Let $\lambda$ be any eigenvalue of $A$ such that $|\lambda| = 1$ and $\lambda \neq 1$. Let $B = \lambda A$. Then clearly $\rho(B) = 1$ and $\|B^k\|_\infty$ is uniformly bounded by 1. So by Lemma 1 again $B$ is mean ergodic. Furthermore $1 \in \sigma(B)$ because if $x$ is an eigenvector of $A$ corresponding to $\lambda$ then

\[ Bx = \lambda Ax = \lambda (\lambda x) = |\lambda|^2 x = x. \]

So by Corollary 1 the algebraic multiplicity and geometric multiplicity of the eigenvalue 1 of $B$ are the same. Since the Jordan structure of $A$ corresponding to $\lambda$ is exactly the same as that of $B$ corresponding to 1, the algebraic multiplicity and geometric multiplicity of the eigenvalue $\lambda$ of $A$ are the same. Since $\lambda$ is an arbitrary eigenvalue of $A$ on the unit circle, the proof is completed. \qed

**Remark.** Corollary 2 is known [1, page 696]. Here we offered a different proof.

Now we give another application of Theorem 1. Let $A \in \mathbb{R}^{n \times n}$ be a positive matrix. Then we have the following facts.

(i) The spectral radius $\rho(A)$ is an eigenvalue of $A$, and there exists a positive eigenvector of $A$ associated with $\rho(A)$.
(ii) The geometric multiplicity of $\rho(A)$ is 1.
(iii) The algebraic multiplicity of $\rho(A)$ is 1.

Here we provide a different proof for (iii). For completeness we provide proofs for (i) and (ii). For (i) we present a proof found in [1]. In the following $|A|$ is the absolute value matrix of $A$ obtained by taking absolute value for each entry of $A$, and the absolute value vector $|x|$ of $x$ is defined similarly. The vector inequalities $|x| \leq |y|$ and $|x| < |y|$ are understood componentwise.

Suppose that $(\lambda, x)$ is any eigenpair of $A$ with $|\lambda| = \rho(A)$. We assert that

\[ A|x| = \rho(A)|x| \quad \text{and} \quad |x| > 0. \]

which establishes (i). Note that

\[ \rho(A)|x| = |\lambda x| = |Ax| \leq |A||x| = A|x|. \]

So we have $\rho(A)|x| \leq |A||x|$. Now we show that equality holds. Let $y = A|x| > 0$ and $z = y - \rho(A)|x|$. We know that $z \geq 0$. Suppose that $z_i > 0$ for some $i$. Then $Az > 0$. So there exists a number $\varepsilon > 0$ such that $Az > \varepsilon y$, or, equivalently,

\[ \frac{A - \rho(A)}{\rho(A) + \varepsilon} y > y. \]
Writing this $By > y$, where $B = A/\rho(A) + \varepsilon > 0$, one can easily see that

$$0 < y < By < \cdots < B^ky < \cdots.$$  

Since $\rho(B) < 1$, $\lim_{k \to \infty} B^ky = 0$. So we arrive at $0 < y < 0$, which is impossible. This establishes the fact that $z = 0$. So $A|x| = \rho(A)|x|$. The fact that $|x| > 0$ follows from $|x| = A|x|/\rho(A)$. Here we used the fact that $\rho(A) > 0$ because $A > 0$.

To establish (ii) we present the following standard argument. Suppose that $x$ and $y$ are a pair of linearly independent eigenvectors of $A$ associated with $\rho(A)$. Select a nonzero component from $y$, say $y_i$, and set $z = x - (x_i/y_i)y$. Since $Az = \rho(A)z$ and $z \neq 0$, we know from the argument in proving (i) that $A|z| = \rho(A)|z|$ and that $|z| > 0$. But this contradicts the fact that $z_i = 0$. Thus the geometric multiplicity of $\rho(A)$ must be 1.

In existing textbooks on nonnegative matrices, there are various proofs of (iii). For example, in book [3], the proof of this fact for irreducible nonnegative square matrices (Theorem 1.4.3), which includes positive matrices as a special case, is based on a calculus argument applied to the derivative of the characteristic polynomial of $A$ with respect to $\lambda$ and is almost two pages long. Basically the same analytic proof was adopted in another book [4] (Theorem 1.4.4(v)). An algebraic proof which is shorter was given in [5], based on the classic Schur triangulation theorem and the matrix powers limit theorem. A similar proof using the Jordan canonical form and the powers limit theorem appears in [1]. A quite different and more analytical approach based on the theory of dynamical systems appears in the book [6]. Most recently, the authors of [7] presented another proof of this famous fact, as a corollary of their eigenvalue perturbation result for specially rank-one updated matrices. This last proof is contained in the textbook [2].

Now we present yet another proof of (iii), using Theorem 1

**Corollary 3.** The algebraic multiplicity of $\rho(A)$ of a positive matrix is 1.

**Proof.** Using (i) and (ii) we see that there are two vectors $x$ and $y$ in $\mathbb{R}^n$ such that

$$Ax = \rho(A)x, \quad A^Ty = \rho(A)y, \quad x > 0, \quad y > 0$$

and that the geometric multiplicity of $\rho(A)$ equals 1. Then $\{x\}$ and $\{y\}$ are bases for $N(A - \rho(A)I)$ and $N(A^T - \rho(A)I)$, respectively. Since $x$ and $y$ are positive vectors, they can be rescaled so that $y^T x = 1$, Hence the condition (a) of the Theorem 1 in Section 2 is fulfilled with $\lambda = \rho(A)$ and $m = 1$. Thus the equivalent condition (d) guarantees that the algebraic multiplicity of $\rho(A)$ must be 1. \qed

**Remark.** (i), (ii) and (iii) are also true when $A$ is an irreducible and nonnegative matrix. They follow from the fact that for an irreducible nonnegative matrix $A \in \mathbb{R}^{n \times n}$ we have $(I + A)^{n-1} > 0$. We note that (i), (ii) and (iii) for irreducible nonnegative matrices are important parts of the Perron–Frobenius theory of nonnegative matrices.

**Acknowledgment**

The authors are thankful for the helpful suggestions of an anonymous referee which improved the paper.

**References**