# **Section 3.8.** Eigenanalysis; Canonical Factorizations

**Note.** In this lengthy section, we define eigenvalues and eigenvectors for a square matrix, and consider some of their properties. We consider geometric and algebraic multiplicity of eigenvalues, similar matrices, diagonalization of matrices, and orthogonal diagonalizations of matrices. We define the spectrum, spectral radius, and spectral decomposition.

**Definition.** Let $A$ be an $n \times n$ matrix. If $v$ is a nonzero $n$-vector and $c$ is a scalar satisfying $Av = cv$ then $v$ is an *eigenvector* of $A$ and $c$ is an *eigenvalue* (or *characteristic value*) of $A$. Such $c$ and $v$ together form an *eigenpair*.

**Note.** Although we do not allow the 0 vector to be an eigenvector, the scalar 0 can be an eigenvalue (which would have associated nonzero eigenvectors). In this section, Gentle continues to assume the **entries of matrices are real** but will **allow the eigenvalues and eigenvectors to be complex or have complex entries**, respectively.

**Definition.** Let $A$ be an $n \times n$ matrix. If $w$ is an $n$-vector, $w \neq 0$, and $c$ is a scalar satisfying $w^T A = cw^T$ then $w$ is a *left eigenvector* of $A$.

**Note.** If $w$ is a left eigenvector of $A$ and $A$ is symmetric, then $(w^T A)^T = (cw^T)^T$ or $A^T w = cw$ or $Aw = cw$ and $w$ is an eigenvector of $A^T = A$.

**Theorem 3.8.1.** If $v$ is an eigenvector of $A$ and $w$ is a left eigenvector of $A$ with a different associated eigenvalue, then $v \perp w$.

**Theorem 3.8.2. Basic Properties of Eigenvalues and Eigenvectors.**

Let $A$ be an $n \times n$ real matrix with eigenpair $c$ and $v$.

**(1)** For any nonzero $b \in \mathbb{C}$, $bv$ is an eigenvector of $A$.

**(2)** For any nonzero $b \in \mathbb{C}$, $bc$ is an eigenvalue of $bA$.

**(3)** If $A$ is nonsingular then $1/c$ and $v$ are an eigenpair for $A^{-1}$.

**(5)** If $A$ is diagonal or triangular with diagonal entries $a_{ii}$, then the eigenvalues of $A$ are $a_{ii}$. For $A$ diagonal, the corresponding eigenvectors are $e_i$ (the $i$th unit vector in $\mathbb{R}^n$).

**(6)** $c^k$ and $v$ are an eigenpair for $A^k$ for $k \in \mathbb{N}$.

**(7)** If $A$ is an $n \times m$ matrix and $B$ is an $m \times n$ matrix then the nonzero eigenvalues of $AB$ are the same as the nonzero eigenvalues of $BA$. If $m = n$ then the eigenvalues of $AB$ and $BA$ are the same (the "nonzero" restriction is removed here).

**(8)** If $A$ and $B$ are $n \times n$ matrices and $B^{-1}$ exists, then the eigenvalues of $BAB^{-1}$ are the same as the eigenvalues of $A$.

**Note.** Gentle states as part (4) of Theorem 3.8.2: If $A$ is square and $c \neq 0$ then $1/c$ and $v$ are an eigenpair for $A^+$. This is false, though, as shown by $A = \begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix}$, which has eigenvalues $\lambda_1 = 0$ and $\lambda_2 = 0$, and $A^+ = \begin{bmatrix} 1/10 & 1/10 \\ 2/10 & 2/10 \end{bmatrix}$, which has eigenvalues $\lambda_1 = 0$ and $\lambda_2 = 3/10$ (so $c = 3$ is an eigenvalue of $A$ but $1/c = 1/3$

is not an eigenvalue of $A^+$). The proof of Theorem 3.8.2 is to be given in Exercise 3.16.

**Corollary 3.8.3.** The set of eigenvectors of a $n \times n$ matrix $A$ associated with given eigenvalue $c$, along with the 0 vector, form a subspace of $\mathbb{C}^n$ (or of $\mathbb{R}^n$ if we restrict ourselves to real numbers). The subspace is the *eigenspace* of $A$ associated with eigenvalue $c$.

**Note.** If $c$ and $v$ are an eigenpair for square matrix $A$, then $Av = cv$ for nonzero $v$, or $Av - cv = (A - c\mathcal{I})v = 0$. If $A - c\mathcal{I}$ is nonsingular (and so invertible) then $v = (A - c\mathcal{I})^{-1}0 = 0$ and this is the only solution. Since $v$ is nonzero by the definition of eigenvectors, we must have $(A - c\mathcal{I})$ singular; that is, $\det(A - c\mathcal{I}) = 0$.

**Definition.** For $n \times n$ matrix $A$, the $n$ degree polynomial $p_A(c) = \det(A - c\mathcal{I})$ is the *characteristic polynomial*. The equation $p_A(c) = 0$ is the *characteristic equation*.

**Note.** Some texts define the characteristic equation as $q_A(c) = \det(c\mathcal{I} - A)$ (see, for example, Definition 1.2.3 on page 49 of R. A. Horn and C. R. Johnson's *Matrix Analysis*, 2nd Edition, Cambridge University Press, 2013). This does not affect the eigenvalues of $A$, but will have an effect when we define the companion matrix of a polynomial below. Since $c\mathcal{I} - A = -(A - c\mathcal{I})$, then by Theorem 3.1.B the relationship between $p_A$ and $q_A$ is $p_A(c) = (-1)^n q_A(c)$.

**Note 3.8.A.** By definition, $\det(A - c\mathcal{I}) = \sum_{\pi \in S_n} \sigma(\pi) \prod_{i=1}^n b_{i\,\pi(i)}$ where $b_{i\,\pi(i)}$ is the $(i, \pi(i))$ entry of $A - cI$. For $\pi$ the identity in $S_n$ (and so $\sigma(\pi) = 1$ since the identity is even), $\sigma(\pi) \prod_{i=1}^n b_{i\,\pi(i)} = +\prod_{i=1}^n (a_{ii} - c)$ includes a term of the form $(-c)^n = (-1)^n c^n$. For $\pi$ not the identity in $S_n$, $\prod_{i=1}^n b_{i\,\pi(i)}$ only has terms involving powers of $c$ of order at most $n-1$ (actually, of degree at most $n-2$ since a permutation producing $n-1$ diagonal elements in $\prod_{i=1}^n b_{i\,\pi(i)}$ would necessarily have to include the final diagonal element). So $\det(A - c\mathcal{I})$ is an $n$ degree polynomial with the coefficient of $c^n$ as $(-1)^n$, where $A$ is $n \times n$.

**Note.** Every eigenvalue of $A$ is a root of the characteristic polynomial and a solution to the characteristic equation. By the Fundamental Theorem of Algebra, every $n \times n$ matrix (with real or complex entries) has $n$ eigenvalues counting multiplicity.

**Theorem 3.8.4. The Cayley-Hamilton Theorem.**
For $n \times n$ matrix $A$ with characteristic polynomial $p_A$ we have $p_A(A) = 0$.

**Theorem 3.8.5.** Let $q(c) = s_0 + s_1 c + s_2 c^2 + \cdots + s_{n-1} c^{n-1} + c^n$ be a monic polynomial. Then $q(c) = \det(c\mathcal{I} - A)$ for some $n \times n$ matrix $A$. In particular, $q(c) = \det(c\mathcal{I} - A)$ for

$$A = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -s_0 & -s_1 & -s_2 & \cdots & -s_{n-1} \end{bmatrix}.$$

Matrix $A$ is called a *companion matrix* for polynomial $q$.

**Note.** Notice that Gentle states that $q$ is the characteristic polynomial of $A$, but this is not the case (see page 109). However, as commented above, $p(c) = (-1)^n q(c)$ for characteristic polynomial $p$ of $n \times n$ matrix $A$. The confusion arises from the two different definitions of characteristic polynomial ($\det(A - c\mathcal{I})$ versus $\det(c\mathcal{I} - A)$ in some other sources, such as Horn and Johnson's text). To further muddle things, sometimes the companion matrix of $q$ is defined as the transpose of what is given in Theorem 3.8.5 (see Horn and Johnson's Definition 3.3.13 on pages 194 and 195); this has no real effect, since $\det(A) = \det(A^T)$ for all square $A$ by Theorem 3.1.A.

**Theorem 3.8.6.** Let $A$ be an $n \times n$ matrix with eigenvalues $c_1, c_2, \ldots, c_n$. Then $\det(A) = \prod_{i=1}^n c_i$ and $\mathrm{tr}(A) = \sum_{i=1}^n c_i$.

**Theorem 3.8.7.** Let $A$ be a real square matrix and $(c, v)$ an eigenpair (possibly complex) for $A$.

**(1)** $c$ is an eigenvalue of $A^T$. The eigenvectors of $A^T$ (which are by definition left eigenvectors of $A$) are not necessarily the same as the eigenvectors of $A$.

**(2)** There is a left eigenvector such that $c$ is the associated eigenvalue.

**(3)** $(\overline{c}, \overline{v})$ is an eigenpair of $A$, where $\overline{c}$ and $\overline{v}$ denote the complex conjugates of $c$ and $v$, respectively.

**(4)** $c\overline{c}$ is an eigenvalue of $A^T A$.

**(5)** $c$ is real if $A$ is symmetric.

**Note.** The proof of Theorem 3.8.7 is to be given in Exercise 3.18.

**Note.** We now address some properties of the set of eigenvalues of a matrix. Throughout, we consider the real and complex eigenvalues. Recall that the modulus of complex $z = a + ib$ is $|z| = \sqrt{a^2 + b^2}$.

**Definition.** The set of all distinct eigenvalues of a matrix $A$ is the *spectrum* of the matrix, denoted $\sigma(A)$.

**Definition.** Let $A$ be an $n \times n$ matrix with eigenvalues $c_1, c_2, \ldots, c_n$ where we index the eigenvalues such that $|c_1| \geq |c_2| \geq \cdots \geq |c_n|$. The *spectral radius* of $A$, denoted $\rho(A)$, is the maximum modulus of an eigenvalue: $\rho(A) = \max_{1 \leq i \leq n} |c_i|$. The set of complex numbers $\{z \mid |z| = \rho(A)\}$ is the *spectral circle* of $A$. An eigenvalue with modulus $\rho(A)$ (such as $c_1$) is a *dominant eigenvalue*.

**Note.** By Theorem 3.8.2(2), if $c$ is an eigenvalue of $A$ and $b \in \mathbb{C}$, then $bc$ is an eigenvalue of $bA$. So if $A$ has a nonzero eigenvalue then the matrix $(1/\rho(A))A$ is a scaled matrix with spectral radius 1.

**Theorem 3.8.8.** Let $A$ be an $n \times n$ matrix with distinct eigenvalues $\{c_1, c_2, \ldots, c_k\}$ and corresponding eigenvectors $\{x_1, x_2, \ldots, x_k\}$ where $(c_i, x_i)$ is an eigenpair for $A$. Then $\{x_1, x_2, \ldots, x_k\}$ is a set of linearly independent vectors. That is, eigenvectors associated with distinct eigenvalues are linearly independent.

**Definition.** Let $c$ be an eigenvalue of square matrix $A$. The dimension of the eigenspace of $A$ associated with $c$ (see Corollary 3.8.3) is the *geometric multiplicity* of $c$.

**Note.** Since the eigenspace of $A$ associated with eigenvalue $c$ is mapped to the 0 vector by $A - c\mathcal{I}$, the geometric multiplicity of $c$ is the nullity of $A - c\mathcal{I}$.

**Definition.** Let $A$ be a square matrix with characteristic polynomial $p_A$ and eigenvalue $c$. If $c$ is a root of $p_A$ of multiplicity $m$ then $c$ is an eigenvalue of *algebraic multiplicity* $m$. If the algebraic multiplicity of $c$ is 1 then $c$ is a *simple eigenvalue*. If the geometric and algebraic multiplicities of $c$ are the same then $c$ is a *semisimple eigenvalue*.

**Definition.** Let $A$ and $B$ be $n \times n$ matrices. If there exists nonsingular $n \times n$ matrix $P$ such that $B = P^{-1}AP$ then $A$ and $B$ are *similar*. If $A$ and $B$ are similar and $B$ is diagonal then $A$ is *diagonalizable*. If there is $n \times n$ orthogonal matrix $Q$ such that $B = Q^T AQ$ then $A$ and $B$ are *orthogonally similar*. If $B$ is diagonal and $A$ and $B$ are orthogonally similar then $A$ is *orthogonally diagonalizable* and $QBQ^T$ is an *orthogonally diagonal factorization* of $A$. If $A$ and $B$ are orthogonally similar and $B$ is upper triangular then $QBQ^T$ is a *Schur factorization* of $A$.

**Theorem 3.8.9.** For any square matrix $A$, a Schur factorization exists.

**Note.** If $B = Q^T AQ$ where $Q$ is orthogonal and $B$ is diagonal, then by Theorem 3.8.2(5 and 8) (and the fact that $Q^{-1} = Q^T$ by Theorem 3.7.1) the eigenvalues of $A$ are the same as the diagonal elements of $B$. Similarly, if $A = QBQ^T$ is a Schur factorization of $A$ (so that $B$ is upper triangular) then the eigenvalues of $A$ are the diagonal entries of $B$. Though the eigenvalues of these similar matrices are the same, they do not necessarily have the same eigenvectors, as is to be shown in Exercise 3.19(b).

**Note.** Let $A$ be an $n \times n$ matrix. Let $C$ be a diagonal matrix with diagonal entries as the eigenvalues of $A$ repeated according to multiplicity. Let $V$ be a matrix with $i$th column an eigenvector corresponding to eigenvalue $c_{ii}$. Then $AV = VC$, as we prove in the next theorem. *If* $V$ is invertible (which would require $\text{rank}(V) = n$ and so would require the linear independence of the eigenvectors in $V$) then $A = VCV^{-1}$.

**Theorem 3.8.10.** Let $A$ be an $n \times n$ matrix, let $c_1, c_2, \ldots, c_n$ be (possibly complex) scalars, and let $v_1, v_2, \ldots, v_n$ be nonzero $n$-vectors. Let $V$ be an $n \times n$ matrix with $i$th column $v_i$ for $1 \leq i \leq n$ and let $C = \text{diag}(c_1, c_2, \ldots, c_n)$. Then $AV = VC$ if and only if $c_1, c_2, \ldots, c_n$ are eigenvalues of $A$ and $v_j$ is an eigenvector of $A$ corresponding to $c_j$ for $j = 1, 2, \ldots, n$.

**Definition.** If $A$ is an $n \times n$ matrix and $A = VCV^{-1}$ for some invertible matrix $V$ and some diagonal matrix $C$, then $A$ is *diagonalizable* (or *simple*) and $A = VCV^{-1}$ is a *diagonal factorization* of $A$. If $A$ has eigenvalues $c_1, c_2, \ldots, c_n$ and $C = \text{diag}(c_1, c_2, \ldots, c_n)$ then $A = VCV^{-1}$ (if this exists) is a *similar canonical form* of $A$. A matrix which is not diagonalizable is a *deficient matrix*.

**Note.** If the eigenvalues of $A$ are distinct then by Theorem 3.8.8 $\text{rank}(V) = n$ (where the columns of $V$ are eigenvectors of $A$, as described above) and so $V^{-1}$ exists and $A$ is diagonalizable.

**Theorem 3.8.11. Diagonalizability Theorem.**

Let $A$ be an $n \times n$ matrix with distinct eigenvalues $c_1, c_2, \ldots, c_k$ with algebraic multiplicities $m_1, m_2, \ldots, m_k$, respectively. Then $A$ is diagonalizable if and only if $\text{rank}(A - c_i\mathcal{I}) = n - m_i$ for $i = 1, 2, \ldots, k$.

**Note.** By the rank nullity equation (Theorem 3.5.4), $\dim(\mathcal{N}(A - c_i\mathcal{I})) = n - \text{rank}(A - c_i\mathcal{I})$, or $\text{rank}(A - c_i\mathcal{I}) = n - \dim(\mathcal{N}(A - c_i\mathcal{I}))$. So $\text{rank}(A - c_i\mathcal{I}) = n - m_i$ if and only if $m_i = \dim(\mathcal{N}(A - c_i\mathcal{I}))$. Now the geometric multiplicity of $c_i$ is $\dim(\mathcal{N}(A - c_i\mathcal{I}))$ so, by Theorem 3.8.11, $A$ is diagonalizable if and only if the geometric multiplicity equals the algebraic multiplicity for each eigenvalue of $A$.

**Example.** We can use the Diagonalizability Theorem to find a nondiagonalizable matrix. Consider $A = \begin{bmatrix} 0 & 1 & 2 \\ 2 & 3 & 0 \\ 0 & 4 & 5 \end{bmatrix}$. We find that the characteristic polynomial equation is $c^3 - 8c^2 + 13c - 6 = (c - 6)(c - 1)^2 = 0$. So $c = 1$ is an eigenvalue of algebraic multiplicity 2. But

$$\text{rank}(A - 1I) = \text{rank}\left(\begin{bmatrix} -1 & 1 & 2 \\ 2 & 2 & 0 \\ 0 & 4 & 4 \end{bmatrix}\right) = 2 \neq n - m_i = 3 - 2 = 1.$$

So by the Diagonalizability Theorem, $A$ is not diagonalizable.

**Note.** On page 117, Gentle argues that a symmetric matrix is diagonalizable. Something more general in fact holds, so we go to another source to prove the general result. The following is based on Professor Ron Freiwald's (of Washington University in St. Louis) website on Orthogonally Diagonalizable Matrices (accessed 4/25/2020).

**Theorem 3.8.A.** A (real) $n \times n$ matrix $A$ is orthogonally diagonalizable if and only if $A$ is symmetric.

**Note.** Of course, a non-symmetric matrix can be diagonalizable, just not *orthogonally* diagonalizable.

**Theorem 3.8.12.** If $A$ is an $n \times n$ diagonalizable matrix where $A = VCV^{-1}$ for diagonal $C$, then

**(1)** there are $n$ linearly independent eigenvectors of $A$,

**(2)** the number of nonzero eigenvalues of $A$ is equal to $\mathrm{rank}(A)$.

**Note.** Gentle states that for diagonalizable $A$ with $A = VCV^{-1}$, and for "function" $f$ of a scalar, we can define

$$f(A) = V\mathrm{diag}(f(c_1), f(c_2), \ldots, f(c_n))V^{-1}$$

where $c_1, c_2, \ldots, c_n$ are the eigenvalues of $A$. In the current setting, the domain of $f$ is presumably $\mathbb{C}$. This definition is fine if $f$ is a polynomial. For other functions ($f(x) = \exp(x)$ is particularly useful), we postpone the discussion until we introduce matrix norms.

**Note 3.8.B.** Since every (real) symmetric matrix is orthogonally diagonalizable, say $A = QCQ^T = QCQ^{-1}$, then $AQ = QC$. By Theorem 3.8.10, the columns of $Q$ are eigenvectors of $A$. So the eigenvectors of a (real) symmetric matrix can be chosen to be orthonormal.

**Note.** If $v_1, v_2, \ldots, v_n$ are orthonormal $n$-vectors then $\mathcal{I} = \sum_{i=1}^n v_i v_i^T$ by Exercise 3.8.A. Therefore, for $A$ symmetric with orthonormal eigenvectors $v_1, v_2, \ldots, v_n$ we have

$$A = A\mathcal{I} = A\sum_{i=1}^n v_i v_i^T = \sum_{i=1}^n A v_i v_i^T = \sum_{i=1}^n c_i v_i v_i^T$$

where $c_i$ is the eigenvalue of $A$ corresponding to eigenvector $v_1$.

**Definition.** If $A$ is a symmetric matrix with orthonormal eigenvectors $v_i$ and corresponding eigenvalues $c_i$ with $i = 1, 2, \ldots, n$, then $A = \sum_{i=1}^n c_i v_i v_i^T$ is the *spectral decomposition* of $A$. Let $P_i = v_i v_i^T$. Then the $P_i$ are *spectral projectors.*

**Example.** We give a quick example of spectral decomposition using an example from Fraleigh and Beauregard (Example 8.4.3). Consider symmetric $A = \begin{bmatrix} 3 & -2 \\ -2 & 0 \end{bmatrix}$. We find $c_1 = 4$ and $c_2 = -1$, $v_1 = \begin{bmatrix} 2/\sqrt{5} \\ -1/\sqrt{5} \end{bmatrix}$ and $v_2 = \begin{bmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \end{bmatrix}$. Notice

$$\sum_{i=1}^n v_i v_i^T = \begin{bmatrix} 2/\sqrt{5} \\ -1/\sqrt{5} \end{bmatrix} \begin{bmatrix} \frac{2}{\sqrt{5}} & -\frac{1}{\sqrt{5}} \end{bmatrix} + \begin{bmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{bmatrix}$$

$$= \begin{bmatrix} 4/5 & -2/5 \\ -2/5 & 1/5 \end{bmatrix} + \begin{bmatrix} 1/5 & 2/5 \\ 2/5 & 4/5 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Also

$$\sum_{i=1}^n c_i v_i v_i^T = 4 \begin{bmatrix} 4/5 & -2/5 \\ -2/5 & 1/5 \end{bmatrix} - 1 \begin{bmatrix} 1/5 & 2/5 \\ 2/5 & 4/5 \end{bmatrix} = \begin{bmatrix} 3 & -2 \\ -2 & 0 \end{bmatrix} = A.$$

**Note.** Since the spectral projectors satisfy $P_i = v_i v_i^T$ where $\mathcal{I} = \sum_{i=1}^n v_i v_i^T$ (by Exercise 3.8.A), then $\sum_{i=1}^n P_i = \mathcal{I}$. Also, $P_i^T = (v_i v_i^T)^T = v_i v_i^T = P_i$ so the $P_i$ are symmetric. Next,

$$P_i P_i = (v_i v_i^T)(v_i v_i^T) = v_i(v_i^T v_i)v_i^T = v_i \|v_i\|^2 v_i^T = v_i(1)v_i^T = v_i v_i^T = P_i$$

and for $i \neq j$,

$$P_i P_j = (v_i v_i^T)(v_j v_j^T) = v_i(v_i^T v_j)v_j^T = v_i\langle v_i, v_j\rangle v_j^T = v_i(0)v_j^T = 0.$$

That is, $P_i P_j = \begin{cases} P_i & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$ where "0" is the $n \times n$ zero matrix. Of course the spectral decomposition in terms of the spectral projectors is $A = \sum_{i=1}^n c_i P_i$. Since, by Theorem 3.8.2(6), the eigenvalues of $A^k$ are $c_i^k$ with eigenvector $v_i$. so the spectral decomposition of $A^k$ is $A^k = \sum_{i=1}^n c_i^k v_i v_i^T = \sum_{i=1}^n c_i^k P_i$.

**Note.** For $A$ a symmetric matrix and $A = \sum_{i=1}^n c_i v_i v_i^T$ as the spectral decomposition of $A$ with the columns of matrix $V$ as the $v_i$, we have that $V$ is invertible. So for any $x \in \mathbb{R}^n$, we have $x = Vb$ for some $b \in \mathbb{R}^n$ (namely, $b = V^{-1}x$). Recall from Section 3.2 that a quadratic form is of the form $x^T Ax$ for $A$ an $n \times n$ matrix and $x \in \mathbb{R}^n$. So with $x = Vb$ we have the quadratic form

$$x^T Ax = x^T \left(\sum_{i=1}^n c_i v_i v_i^T\right) x = (Vb)^T \left(\sum_{i=1}^n c_i v_i v_i^T\right) Vb$$

$$= \sum_{i=1}^n b^T V^T c_i v_i v_i^T V b = \sum_{i=1}^n b^T V^T v_i v_i^T V b c_i$$

since $c_i$ is a scalar. Now $v_i^T V$ is the $1 \times n$ vector $[\langle v_i, v_1\rangle \ \langle v_i, v_2\rangle \ \cdots \ \langle v_i, v_n\rangle] = e_i$ where $e_i$ is the $i$th standard basis (row) vector of $\mathbb{R}^n$. Therefore $v_i^T V b = b_i$ (here, $b_i$ is the $i$th component of vector $b$). Similarly, $b^T V^T v_i = b^T e_i^T = b_i$. Therefore,

$$x^T A x = \sum_{i=1}^n b^T V^T v_i v_i^T V b c_i = \sum_{i=1}^n b_i^2 c_i.$$

Since each $c_i$ is real by Theorem 3.8.7(5), we can consider $\max\{c_i\}$. So we have

$$x^T A x = \sum_{i=1}^n b_i^2 c_i \le \max\{c_i\} \sum_{i=1}^n b_i^2 = \|b\|^2 \max\{c_i\}.$$

Since $x = Vb$ then $x^T x = (Vb)^T (Vb) = b^T V^T V b = b^T V^{-1} V b = b^T b$ (since $V$ is orthogonal and so $V^T = V^{-1}$). Therefore for $x \ne 0$, since $\|x\| = \|Vb\| = \|b\|$ by Exercise 3.7.C, we have

$$\frac{x^T A x}{x^T x} = \frac{x^T A x}{\langle x, x\rangle} \le \max\{c_i\}.$$

**Definition.** For symmetric $n \times n$ real matrix $A$, the function $R_A : \mathbb{R}^n \to \mathbb{R}$ defined as

$$R_A(x) = \frac{x^T A x}{x^T x} = \frac{\langle x, Ax\rangle}{\langle x, x\rangle} \text{ (where } x \ne 0)$$

is the *Rayleigh quotient* of $A$.

**Note.** The Rayleigh quotient is used in a numerical technique that approximates the eigenvalues of maximum modulus and the corresponding eigenvectors (see Fraleigh and Beauregard's *Linear Algebra*, 3rd Editon, Section 8.4 "Computing Eigenvalues and Eigenvectors"). It is named for John William Strut, the third

Baron Rayleigh (1842–1919). He was a physicist and studied sound and optics. He is most famous for explaining why the sky is blue (it is due to the physical process called Rayleigh scattering); he also is a codiscoverer of the element argon for which he won the 1904 Nobel prize in physics.

**Note.** Recall from Section 3.2 that the inner product of $n \times m$ matrices $A$ and $B$ where the columns of $A$ are $a_1, a_2, \ldots, a_m$ and the columns of $B$ are $b_1, b_2, \ldots, b_m$, is $\langle A, B\rangle = \sum_{j=1}^m a_j^T b_j$. So for orthonormal $v_1, v_2, \ldots, v_m$ we have that the $j$th column of $v_i v_i^T$ is $v_i^j v_i$ where $v_i^j$ is the $j$th entry of $v_i$. So

$$\langle v_i v_i^T, v_i v_i^T\rangle = \sum_{j=1}^n (v_i^j v_i)^T (v_i^j v_i) = \sum_{j=1}^n (v_i^j)^2 v_i^T v_i = \|v_i\|^2 \langle v_i, v_i\rangle = \|v_i\|^4 = 1.$$

For $k \neq i$,

$$\langle v_i v_i^T, v_k v_k^T\rangle = \sum_{j=1}^n (v_i^j v_i)^T (v_k^j v_k) = \sum_{j=1}^n v_i^j v_k^j v_i^T v_k = \sum_{j=1}^n v_i^j v_k^j \langle v_i, v_k\rangle = 0.$$

So the $n \times n$ matrices $P_i = v_i v_i^T$ form an orthonormal system of matrices. So by Corollary 2.2.2 (the Fourier expansion of matrix $A$), for any $n \times n$ symmetric matrix $A$,

$$A = \langle A, v_1 v_1^T\rangle v_1 v_1^T + \langle A, v_2 v_2^T\rangle v_2 v_2^T + \cdots + \langle A, v_n v_n^T\rangle v_n v_n^T.$$

But the spectral decomposition of $A$ is $A = \sum_{i=1}^n c_i v_i v_i^T$ and since representations with respect to a given basis are unique, then $c_i = \langle A, v_i v_i^T\rangle$ for $i = 1, 2, \ldots, n$.

**Theorem 3.8.13.** If $A$ is a symmetric matrix where $(c, v)$ is an eigenpair for $A$ with $v^T v = \|v\|^2 = 1$, then for any $k \in \mathbb{N}$ we have $(A - cvv^T)^k = A^k - c^k vv^T$.

**Note.** A result related to Theorem 3.8.13 holds for nonsymmetric square matrices. Exercise 3.8.B states: "Let $A$ be an $n \times n$ (not necessarily symmetric) matrix. Let $w$ be a left eigenvector for eigenvalue $c$ and let $v$ be a right eigenvector for eigenvalue $c$, where $w^T v = 1$. Prove that for $k \in \mathbb{N}$, $(A - cvw^T)^k = A^k - c^k vw^T$."

**Theorem 3.8.14.** Any real symmetric matrix is positive definite if and only if all of its eigenvalues are positive. Any real symmetric matrix is nonnegative definite if and only if all of its eigenvalues are nonnegative.

**Note.** If square matrix $A$ is positive definite and orthogonally diagonalizable then $A = VCV^T$ for orthogonal $V$ and $A^{-1} = (VCV^T)^{-1} = VC^{-1}V^{-1} = VC^{-1}V^T$. Since the eigenvalues (i.e., the diagonal entries of $C$) are positive by Theorem 3.8.14, then the diagonal entries of $C^{-1}$ are positive (i.e. the eigenvalues of $C^{-1}$) and so by Theorem 3.8.14, $C^{-1}$ is positive definite.

**Theorem 3.8.15.**

**(1)** If symmetric matrix $A$ is positive definite then there is nonsingular $P$ such that $P^T AP = \mathcal{I}$.

**(2)** Suppose symmetric matrix $A$ is nonnegative definite and $A = VCV^T$ where $V$ is orthogonal (such $V$ exists by Theorem 3.8.A) and $C = \text{diag}(c_1, c_2, \ldots, c_n)$ where the eigenvalues of $A$ are $c_1, c_2, \ldots, c_n$. Then there is diagonal nonnegative definite matrix $S$ such that $(VSV^T)^2 = A$.

**Definition.** If symmetric matrix $A$ is nonnegative definite then the matrix $VSV^T$ of Theorem 3.8.15(2) where $(VSV^T)^2 = A$ is the *square root* of $A$, denoted $A^{1/2}$.

**Note.** For $r \in \mathbb{N}$ we can similarly define $A^{1/r}$ by letting $S = \text{diag}(\sqrt[r]{c_1}, \sqrt[r]{c_2}, \ldots, \sqrt[r]{c_n})$. If symmetric $A$ is positive definite then all eigenvalues are positive by Theorem 3.8.14 and by Theorem 3.8.6 $\det(A)$ is the product of the eigenvalues and so $A$ is invertible by Theorem 3.3.16. If $c$ is an eigenvalue of $A$ then $1/c$ is an eigenvalue of $A^{-1}$ by Theorem 3.8.2(4) and so $A^{-1}$ is positive definite by Theorem 3.8.14 ($A^{-1}$ is symmetric since $(A^{-1})^T = (A^T)^{-1}$ by Theorem 3.3.7). So we can define the square root of $A^{-1}$, denoted $A^{-1/2}$. Similarly we can define $A^{-1/r}$ for $r \in \mathbb{N}$.

**Definition.** Let $A$ and $B$ be $n \times n$ matrices. A value $c \in \mathbb{C}$ such that $\det(A - cB) = 0$ is a *generalized eigenvalue* of $A$ with respect to $B$. If $v \in \mathbb{R}^n$ satisfies $Av = cBv$ then $v$ is a *generalized eigenvector* of $A$ with respect to $B$ for $c$.

**Note.** Gentle claims without proof that every $n \times m$ matrix $A$ has a singular value decomposition, which we define next. We give a proof of the existence of such a decomposition from another source.

**Definition.** For an $n \times m$ matrix $A$, a factorization $A = UDV^T$, where $U$ is an $n \times n$ orthogonal matrix, $V$ is an $m \times m$ orthogonal matrix, and $D$ is an $n \times m$ diagonal matrix with nonnegative entries is a *singular value decomposition* of $A$. (An $n \times m$ diagonal matrix has $\min\{n, m\}$ elements on the diagonal and all other entries are zero.) The nonzero entries of $D$ are the *singular values* of $A$.

**Note.** The following proof in the notation of Gentle is based on Harville's *Matrix Algebra From a Statistician's Perspective* (Springer, 1997; see pages 550-51).

**Theorem 3.8.16.** Let $A$ be an $n \times m$ matrix. Then there exists a singular value decomposition of $A$.

**Definition.** Let $A$ be an $n \times m$ matrix of rank $r$ with singular value decomposition $A = UDV^T$ where $D = \mathrm{diag}(d_1, d_2, \ldots, d_n)$ (here, $D$ is $n \times m$) and $d_1 \geq d_2 \geq \cdots \geq d_n \geq 0$ (so $d_{r+1} = d_{r+2} = \cdots = d_n = 0$). Let the columns of $U$ be $u_i$ and the columns of $V$ be $v_i$. Then we can express $A$ as $A = UDV^T = \sum_{i=1}^r d_i u_i v_i^T$. This is a *spectral decomposition* of $A$.

**Note.** If $A$ is $n \times n$ and symmetric then $A$ is orthogonally diagonalizable (Theorem 3.8.A) and the previous definition reduces to the definition of spectral decomposition given previously for symmetric matrices.

**Theorem 3.8.17.** Let $A$ be an $n \times m$ matrix with spectral decomposition $A = UDV^T = \sum_{i=1}^r d_i u_i v_i^T$. Then $\langle u_i v_i^T, u_j v_j^T \rangle = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$ and $d_i = \langle A, u_i v_i^T \rangle$. That is, the spectral decomposition is a Fourier expansion of $A$.

*Revised: 7/11/2020*