Chapter 8

Canonical Forms

8.1 Eigenvalues and Eigenvectors

Definition 8.1.1. *Let V be a vector space over the field F and let T be a linear operator on V. An* ***eigenvalue*** *of T is a scalar $\lambda \in F$ such that there exists a non-zero vector $\underline{v} \in V$ with $T\underline{v} = \lambda\underline{v}$. Any vector $\underline{v}$ such that $T\underline{v} = \lambda\underline{v}$ is called an* ***eigenvector*** *of T associated with the eigenvalue value λ.*

Definition 8.1.2. *The* ***spectrum*** *$\sigma(T)$ of a linear operator $T\colon V \to V$ is the set of all scalars such that the operator $(T - \lambda I)$ is not invertible.*

Example 8.1.3. *Let $V = \ell_2$ be the Hilbert space of infinite square-summable sequences and $T : V \to V$ be the right-shift operator defined by*

$$T(v_1, v_2, \ldots) = (0, v_1, v_2, \ldots).$$

Since T is not invertible, it follows that the scalar 0 is in the spectrum of T. But, it is not an eigenvalue because $T\underline{v} = \underline{0}$ implies $\underline{v} = 0$ and an eigenvector must be a non-zero vector. In fact, this operator does not have any eigenvalues.

For finite-dimensional spaces, things are quite a bit simpler.

Theorem 8.1.4. *Let A be the matrix representation of a linear operator on a finite-dimensional vector space V, and let λ be a scalar. The following are equivalent:*

1. *λ is an eigenvalue of A*

2. *the operator $(A - \lambda I)$ is singular*

3. $\det(A - \lambda I) = 0$.

Proof. First, we show the first and third are equivalent. If λ is an eigenvalue of A, then there exists a vector $\underline{v} \in V$ such that $A\underline{v} = \lambda\underline{v}$. Therefore, $(A - \lambda I)\underline{v} = 0$ and $(A - \lambda I)$ is singular. Likewise, if $(A - \lambda I)\underline{v} = 0$ for some $\underline{v} \in V$ and $\lambda \in F$, then $A\underline{v} = \lambda\underline{v}$. To show the second and third are equivalent, we note that the determinant of a matrix is zero iff it is singular. □

The last criterion is important. It implies that every eigenvalue λ is a root of the polynomial

$$\chi_A(\lambda) \triangleq \det(\lambda I - A)$$

called the **characteristic polynomial** of A. The equation $\det(A - \lambda I) = 0$ is called the characteristic equation of A. The spectrum $\sigma(A)$ is given by the roots of the characteristic polynomial $\chi_A(\lambda)$.

Let A be a matrix over the field of real or complex numbers. A nonzero vector $\underline{v}$ is called a **right eigenvector** for the eigenvalue λ if $A\underline{v} = \lambda\underline{v}$. It is called a **left eigenvector** if $\underline{v}^H A = \lambda\underline{v}^H$.

Definition 8.1.5. *Let λ be an eigenvalue of the matrix A. The **eigenspace** associated with λ is the set $E_\lambda = \{\underline{v} \in V | A\underline{v} = \lambda\underline{v}\}$. The **algebraic multiplicity** of λ is the multiplicity of the zero at $t = \lambda$ in the characteristic polynomial $\chi_A(t)$. The **geometric multiplicity** of an eigenvalue λ is equal to dimension of the eigenspace E_λ or nullity$(A - tI)$.*

Theorem 8.1.6. *If the eigenvalues of an $n \times n$ matrix are all distinct, then the eigenvectors of A are linearly independent.*

Proof. We will prove the slightly stronger statement: if $\lambda_1, \lambda_2, \ldots, \lambda_k$ are distinct eigenvalues with eigenvectors $\underline{v}_1, \underline{v}_2, \ldots, \underline{v}_k$, then the eigenvectors are linearly independent. Suppose that

$$\sum_{i=1}^{k} c_i \underline{v}_i = \underline{0}$$

for scalars $c_1, c_2, \ldots, c_k$. Notice that one can annihilate $\underline{v}_j$ from this equation by multiplying both sides by $(A - \lambda_j I)$. So, multiplying both sides by a product of these matrices gives

$$\prod_{j=1,j\neq m}^{k}(A-\lambda_j I)\sum_{i=1}^{k} c_j \underline{v}_i = \left(\prod_{j=1,j\neq m}^{k}(A-\lambda_j I)\right) c_m \underline{v}_m$$
$$= c_m \prod_{j=1,j\neq m}^{k}(\lambda_m - \lambda_j) = \underline{0}.$$

Since all eigenvalues are distinct, we must conclude that $c_m = 0$. Since the choice of m was arbitrary, it follows that $c_1, c_2, \ldots, c_k$ are all zero. Therefore, the vectors $\underline{v}_1, \underline{v}_2, \ldots, \underline{v}_k$ are linearly independent. □

Definition 8.1.7. *Let T be a linear operator on a finite-dimensional vector space V. The operator T is* ***diagonalizable*** *if there exists a basis $\mathcal{B}$ for V such that each basis vector is an eigenvector of T,*

$$[T]_{\mathcal{B}} = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}$$

Similarly, a matrix A is diagonalizable if there exists an invertible matrix S such that

$$A = S\Lambda S^{-1}$$

where Λ is a diagonal matrix.

Theorem 8.1.8. *If an $n \times n$ matrix has n linearly independent eigenvectors, then it is diagonalizable.*

Proof. Suppose that the $n \times n$ matrix A has n linearly independent eigenvectors, which we denote by $\underline{v}_1, \ldots, \underline{v}_n$. Let the eigenvalue of $\underline{v}_i$ be denoted by λ_i so that

$$A\underline{v}_j = \lambda_j \underline{v}_j, \quad j = 1, \ldots, n.$$

In matrix form, we have

$$A\begin{bmatrix} \underline{v}_1 & \cdots & \underline{v}_n \end{bmatrix} = \begin{bmatrix} A\underline{v}_1 & \cdots & A\underline{v}_n \end{bmatrix}$$
$$= \begin{bmatrix} \lambda_1\underline{v}_1 & \cdots & \lambda_n\underline{v}_n \end{bmatrix}.$$

We can rewrite the last matrix on the right as

$$\begin{bmatrix} \lambda_1 \underline{v}_1 & \cdots & \lambda_n \underline{v}_n \end{bmatrix} = \begin{bmatrix} \underline{v}_1 & \cdots & \underline{v}_n \end{bmatrix} \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix} = S\Lambda.$$

where

$$S = \begin{bmatrix} \underline{v}_1 & \cdots & \underline{v}_n \end{bmatrix} \quad \text{and} \quad \Lambda = \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix},$$

Combining these two equations, we obtain the equality

$$AS = S\Lambda.$$

Since the eigenvectors are linearly independent, the matrix S is full rank and hence invertible. We can therefore write

$$A = S\Lambda S^{-1}$$
$$\Lambda = S^{-1}AS.$$

That is, the matrix A is diagonalizable. □

The type of the transformation from A to Λ arises in a variety of contexts.

Definition 8.1.9. *If there exists an invertible matrix T such that*

$$A = TBT^{-1},$$

then matrices A and B are said to be ***similar****.*

If A and B are similar, then they have the same eigenvalues. Similar matrices can be considered representations of the same linear operator using different bases.

Lemma 8.1.10. *Let A be an $n \times n$ Hermitian matrix (i.e., $A^H = A$). Then, the eigenvalues of A are real and the eigenvectors associated with distinct eigenvalues are orthogonal.*

Proof. First, we notice that $A = A^H$ implies $\underline{v}^H A \underline{v}$ is real because

$$\overline{s} = \left(\underline{v}^H A \underline{v}\right)^H = \underline{v}^H A^H \underline{v} = \underline{v}^H A \underline{v} = s.$$

If $A\underline{v} = \lambda_1 \underline{v}$, left multiplication by $\underline{v}^H$ shows that

$$\underline{v}^H A\underline{v} = \lambda_1 \underline{v}^H \underline{v} = \lambda_1 \|\underline{v}\|.$$

Therefore, λ_1 is real. Next, assume that $A\underline{w} = \lambda_2 \underline{w}$ and $\lambda_2 \neq \lambda_1$. Then, we have

$$\lambda_1 \lambda_2 \underline{w}^H \underline{v} = \underline{w}^H A^H A\underline{v} = \underline{w} A^2 \underline{v} = \lambda_1^2 \underline{w}^H \underline{v}.$$

We also assume, without loss of generality, that $\lambda_1 \neq 0$. Therefore, if $\lambda_2 \neq \lambda_1$, then $\underline{w}^H \underline{v} = 0$ and the eigenvectors are orthogonal. □

8.2 Applications of Eigenvalues

8.2.1 Differential Equations

It is well known that the solution of the 1st-order linear differential equation

$$\frac{d}{dt} x(t) = a x(t)$$

is given by

$$x(t) = e^{at} x(0).$$

It turns out that this formula can be extended to coupled differential equations. Let A be a diagonalizable matrix and consider the the set of 1st order linear differential equations defined by

$$\frac{d}{dt} \underline{x}(t) = A\underline{x}(t).$$

Using the decomposition $A = S\Lambda S^{-1}$ and the substitution $\underline{x}(t) = S\underline{y}(t)$, we find that

$$\begin{aligned} \frac{d}{dt} \underline{x}(t) &= \frac{d}{dt} S\underline{y}(t) \\ &= S \frac{d}{dt} \underline{y}(t). \end{aligned}$$

and

$$\begin{aligned} \frac{d}{dt} \underline{x}(t) &= A\underline{x}(t) \\ &= AS\underline{y}(t). \end{aligned}$$

This implies that

$$\frac{d}{dt}\underline{y}(t) = S^{-1}AS\underline{y}(t) = \Lambda\underline{y}(t).$$

Solving each individual equation gives

$$y_j(t) = e^{\lambda_j t} y_j(0)$$

and we can group them together in matrix form with

$$\underline{y}(t) = e^{\Lambda t}\underline{y}(0).$$

In terms of $\underline{x}(t)$, this gives

$$\underline{x}(t) = Se^{\Lambda t}S^{-1}\underline{x}(0).$$

In the next section, we will see this is equal to $\underline{x}(t) = e^{At}\underline{x}(0)$.

8.2.2 Functions of a Matrix

The diagonal form of a diagonalizable matrix can be used in a number of applications. One such application is the computation of matrix exponentials. If $A = S\Lambda S^{-1}$ then

$$A^2 = S\Lambda S^{-1}S\Lambda S^{-1} = S\Lambda^2 S^{-1}$$

and, more generally,

$$A^n = S\Lambda^n S^{-1}.$$

Note that Λ^n is obtained in a straightforward manner as

$$\Lambda^n = \begin{bmatrix} \lambda_1^n & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n^n \end{bmatrix}.$$

This observation drastically simplifies the computation of the matrix exponential e^A,

$$e^A = \sum_{i=0}^{\infty} \frac{A^i}{i!} = S\left(\sum_{i=0}^{\infty} \frac{\Lambda^i}{i!}\right)S^{-1} = Se^{\Lambda}S^{-1},$$

where

$$e^{\Lambda} = \begin{bmatrix} e^{\lambda_1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & e^{\lambda_n} \end{bmatrix}.$$

Theorem 8.2.1. *Let $p(\cdot)$ be a given polynomial. If λ is an eigenvalue of A, while $\underline{v}$ is an associated eigenvector, then $p(\lambda)$ is an eigenvalue of the matrix $p(A)$ and $\underline{v}$ is an eigenvector of $p(A)$ associated with $p(\lambda)$.*

Proof. Consider $p(A)\underline{v}$. Then,

$$p(A)\underline{v} = \sum_{k=0}^{l} p_k A^k \underline{v} = \sum_{k=0}^{l} p_k \lambda^k \underline{v} = p(\lambda)\underline{v}.$$

That is $p(A)\underline{v} = p(\lambda)\underline{v}$. □

A matrix A is singular if and only if 0 is an eigenvalue of A.

8.3 The Jordan Form

Not all matrices are diagonalizable. In particular, if A has an eigenvalue whose algebraic multiplicity is larger than its geometric multiplicity, then that eigenvalue is called **defective**. A matrix with a defective eigenvalue is not diagonalizable.

Theorem 8.3.1. *Let A be an $n \times n$ matrix. Then A is diagonalizable if and only if there is a set of n linearly independent vectors, each of which is an eigenvector of A.*

Proof. If A has n linearly independent eigenvectors $\underline{v}_1, \ldots, \underline{v}_n$, then let S be an invertible matrix whose columns are there n vectors. Consider

$$\begin{aligned} S^{-1}AS &= S^{-1}\begin{bmatrix} A\underline{v}_1 & \cdots & A\underline{v}_n \end{bmatrix} \\ &= S^{-1}\begin{bmatrix} \lambda_1\underline{v}_1 & \cdots & \lambda_n\underline{v}_n \end{bmatrix} \\ &= S^{-1}S\Lambda = \Lambda. \end{aligned}$$

Conversely, suppose that there is a similarity matrix S such that $S^{-1}AS = \Lambda$ is a diagonal matrix. Then $AS = S\Lambda$. This implies that A times the ith column of S is the ith diagonal entry of Λ times the ith column of S. That is, the ith column of S is an eigenvector of A associated with the ith diagonal entry of Λ. Since S is nonsingular, there are exactly n linearly independent eigenvectors. □

Definition 8.3.2. *The **Jordan normal form** of any matrix $A \in \mathbb{C}^{n\times n}$ with $l \leq n$ linearly independent eigenvectors can be written as*

$$A = TJT^{-1},$$

where T is an invertible matrix and J is the block-diagonal matrix

$$J = \begin{bmatrix} J_{m_1}(\lambda_1) & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & J_{m_l}(\lambda_l) \end{bmatrix}.$$

The $J_m(\lambda)$ are $m \times m$ matrices called Jordan blocks, and they have the form

$$J_m(\lambda) = \begin{bmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda \end{bmatrix}.$$

It is important to note that the eigenvalues $\lambda_1, \ldots, \lambda_l$ are not necessarily distinct (i.e., multiple Jordan blocks may have the same eigenvalue). The Jordan matrix J associated with any matrix A is unique up to the order of the Jordan blocks. Moreover, two matrices are similar iff they are both similar to the same Jordan matrix J.

Since every matrix is similar to a Jordan block matrix, one can gain some insight by studying Jordan blocks. In fact, Jordan blocks exemplify the way that matrices can be degenerate. For example, $J_m(\lambda)$ has the single eigenvector $\underline{e}_1$ (i.e., the standard basis vector) and satisfies

$$J_m(0)\underline{e}_{j+1} = \underline{e}_j \text{ for } j = 1, 2, \ldots, m-1.$$

So, the reason this matrix has only one eigenvector is that left-multiplication by this matrix shifts all elements in a vector up element.

Computing the Jordan normal form of a matrix can be broken into two parts. First, one can identify, for each distinct eigenvalue λ, the **generalized eigenspace**

$$G_\lambda = \left\{ \underline{v} \in \mathbb{C}^n \middle| (A - \lambda I)^n \underline{v} = \underline{0} \right\}.$$

Let $\lambda_1, \ldots, \lambda_k$ be the distinct eigenvalues of A ordered by decreasing magnitude. Let d_j be the dimension of G_{λ_j}, which is equal to the sum of the sizes of the Jordan blocks associated with λ, then $\sum_{j=1}^{k} d_j = n$. Let T be a matrix whose first d_1 columns for a basis for G_{λ_1}, next d_2 columns form a basis for G_{λ_2}, and so on. In this case, the matrix $T^{-1}AT$ is block diagonal and the j-th block B_j is associated with the eigenvalue λ_j.

To put A in Jordan normal form, we now need to transform each block matrix B into Jordan normal form. One can do this by identifying the subspace V_j that is not mapped to $\underline{0}$ by $(B-\lambda I)^{j-1}$ (i.e., $\mathcal{N}\left(\left((B-\lambda I)^{j-1}\right)^{\perp}\right)$). This gives the sequence $V_1, \ldots, V_J$ of non-empty subspaces (e.g., V_j is empty for $j > J$). Now, we can form a sequence of bases $W_J, W_{J-1}, \ldots, W_1$ recursively starting from W_J with

$$W_j = W_{j+1} \cup \{(B-\lambda I)\underline{w} | \underline{w} \in W_{j+1}\} \cup \text{basis}(V_j - V_{j-1}),$$

where basis$(V_j - V_{j-1})$ is some set basis vectors that extends V_{j-1} to V_j. Each vector in W_j gives rise to a length j **Jordan chain** of vectors $\underline{v}_{i-1} = (B-\lambda I)\underline{v}_i \in W_{i-1}$ starting from any $\underline{v}_j \in W_j$. Each vector $\underline{v}_j$ defined in this way is called a **generalized eigenvector** of order j. By correctly ordering the basis W_1 as columns of T, one finds that $T^{-1}BT$ is a Jordan matrix.

Example 8.3.3. *Consider the matrix*

$$\begin{bmatrix} 4 & 0 & 1 & 0 \\ 2 & 2 & 3 & 0 \\ -1 & 0 & 2 & 0 \\ 4 & 0 & 1 & 2 \end{bmatrix}.$$

First, we find the characteristic polynomial

$$\chi_A(t) = \det(tI - A) = t^4 - 10t^3 + 37t^2 - 60t + 36 = (t-2)^2(t-3)^2.$$

Next, we find the eigenvectors associated with the eigenvalues $\lambda_1 = 3$ and $\lambda_2 = 2$. This is done by finding a basis $\underline{v}_1^{(i)}, \underline{v}_2^{(i)}, \ldots$ for the nullspace of $A - \lambda_i I$ and gives

$$\begin{aligned} \underline{v}_1^{(1)} &= [1\ -1\ -1\ 3]^T \\ \underline{v}_1^{(2)} &= [0\ 1\ 0\ 0]^T \\ \underline{v}_2^{(2)} &= [0\ 0\ 0\ 1]^T. \end{aligned}$$

Since the eigenvalue λ_1 has algebraic multiplicity 2 and geometric multiplicity 1, we still need to find another generalized eigenvector associated with this eigenspace.

In particular, we need a vector $\underline{w}$ which satisfies $(A-\lambda_1 I)\underline{w} = \underline{v}_1^{(1)}$. This gives

$$\begin{bmatrix} 1 & 0 & 1 & 0 \\ 2 & -1 & 3 & 0 \\ -1 & 0 & -1 & 0 \\ 4 & 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ w_3 \\ w_4 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ -1 \\ 3 \end{bmatrix}.$$

Using the pseudoinverse of $(A-\lambda_1 I)$, one finds that $\underline{w} = \left[\frac{11}{12}\ \frac{37}{12}\ \frac{1}{12}\ \frac{9}{12}\right]$. Using this, we construct the Jordan normal form by noting that

$$\begin{bmatrix} 4 & 0 & 1 & 0 \\ 2 & 2 & 3 & 0 \\ -1 & 0 & 2 & 0 \\ 4 & 0 & 1 & 2 \end{bmatrix} \left[\underline{v}_1^{(1)} \quad \underline{w} \quad \underline{v}_1^{(2)} \quad \underline{v}_2^{(2)} \right] = \left[3\underline{v}_1^{(1)} \quad \underline{v}_1^{(1)} + 3\underline{w} \quad 2\underline{v}_1^{(2)} \quad 2\underline{v}_2^{(2)} \right]$$

$$= \left[\underline{v}_1^{(1)} \quad \underline{w} \quad \underline{v}_1^{(2)} \quad \underline{v}_2^{(2)} \right] \begin{bmatrix} 3 & 1 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}.$$

This implies that $A = TJT^{-1}$ with

$$T = \left[\underline{v}_1^{(1)} \quad \underline{w} \quad \underline{v}_1^{(2)} \quad \underline{v}_2^{(2)} \right] = \begin{bmatrix} 1 & \frac{11}{12} & 0 & 0 \\ -1 & \frac{37}{12} & 1 & 0 \\ -1 & \frac{1}{12} & 0 & 0 \\ 3 & \frac{9}{12} & 0 & 1 \end{bmatrix}.$$

8.4 Applications of Jordan Normal Form

Jordan normal form often allows one to extend to all matrices results that are easy to prove for diagonalizable matrices.

8.4.1 Convergent Matrices

Definition 8.4.1. *An $n \times n$ matrix A is **convergent** if $\|A^k\| \to 0$ for any norm.*

Of course, this is equivalent to the statement "A^k converges to the all zero matrix". Since all finite-dimensional vector norms are equivalent, it also follows that this condition does not depend on the norm chosen.

Recall that the spectral radius $\rho(A)$ of a matrix A is the magnitude of the largest eigenvalue. If A is diagonalizable, then $A^k = T\Lambda^k T^{-1}$ and it is easy to see that

$$\|A^k\| \leq \|T\|\|\Lambda^k\|\|T^{-1}\|.$$

Since all finite-dimensional vector norms are equivalent, we know that $\|\Lambda^k\| \leq M\|\Lambda^k\|_1 = M\rho(A)^k$. Therefore, A is convergent if $\rho(A) < 1$. If $\rho(A) \geq 1$, then it is easy to show that $\|\Lambda^k\| > 0$ and therefore that $\|A^k\| > 0$. For general matrices, we can instead use the Jordan normal form and the following lemma.

Lemma 8.4.2. *The Jordan block $J_m(\lambda)$ is convergent iff $|\lambda| < 1$.*

Proof. This follows from the fact that $J_m(\lambda) = \lambda I + N$, where $[N]_{i,j} = \delta_{i+1,j}$. Using the Binomial formula, we write

$$\begin{aligned}\|(\lambda I + N)^k\| &= \left\| \sum_{i=0}^{k} \binom{k}{i} N^i \lambda^{k-i} \right\| \\ &\leq \sum_{i=0}^{m-1} \binom{k}{i} |\lambda|^{k-i},\end{aligned}$$

where the second step follows from the fact that $\|N^i\|$ is 1 for $i = 1, \ldots, m-1$ and zero for $i \geq m$. Notice that $|\binom{k}{i}\lambda^{k-i}| \leq k^{m-1}|\lambda|^{k-m+1}$ for $0 \leq i \leq m-1$. Since $k^{m-1}|\lambda|^{k-m+1} \to 0$ as $k \to \infty$ iff $|\lambda| < 1$, we see that each term in the sum converges to zero under the same condition. On the other hand, if $|\lambda| \geq 1$, then $|\left[(\lambda I + N)^k\right]_{1,1}| \geq 1$ for all $k \geq 0$. □

Theorem 8.4.3. *A matrix $A \in \mathbb{C}^{n\times n}$ is convergent iff $\rho(A) < 1$.*

Proof. Using the Jordan normal form, we can write $A = TJT^{-1}$, where J is a block diagonal with k Jordan blocks $J_1, \ldots, J_k$. Since J is block diagonal, we also have that $\|J^k\| \leq \sum_{i=1}^{k} \|J_i^k\|$. If $\rho(A) < 1$, then the eigenvalue λ associated with each Jordan block satisfies $\|\lambda\| < 1$. In this case, the lemma shows that $\|J_i^k\| \to 0$ which implies that $\|J^k\| \to 0$. Therefore, $\|A^k\| \to 0$ and A is convergent. On the other hand, if $\rho(A) \geq 1$, then there is a Jordan block J_i with $|\lambda| \geq 1$ and $|[J_i^k]_{1,1}| \geq 1$ for all $k \geq 0$. □

In some cases, one can make stronger statements about large powers of a matrix.

Definition 8.4.4. *A matrix A has a* ***unique eigenvalue of maximum modulus*** *if the Jordan block associated with that eigenvalue is 1×1 and all other Jordan blocks are associated with eigenvalues of smaller magnitude.*

The following theorem shows that a properly normalized matrix of this type converges to a non-zero limit.

Theorem 8.4.5. *If A has a unique eigenvalue λ_1 of maximum modulus, then*

$$\lim_{k\to\infty} \frac{1}{\lambda_1^k} A^k = \underline{u}\underline{v}^H,$$

where $A\underline{u} = \lambda_1 \underline{u}$, $\underline{v}^H A = \lambda_1 \underline{v}^H$, and $\underline{v}^H \underline{u} = 1$.

Proof. Let $B = \frac{1}{\lambda_1} A$ so that maximum modulus eigenvalue is now 1. Next, choose the Jordan normal form $B = TJT^{-1}$ so that the Jordan block associated with the eigenvalue 1 is in the top left corner of J. In this case, it follows from the lemma that J^n converges to $\underline{e}_1 \underline{e}_1^H$ as $n \to \infty$. This implies that $B^n = TJ^nT^{-1}$ converges to $T\underline{e}_1\underline{e}_1^H T^{-1} = \underline{u}\underline{v}^H$ where $\underline{u}$ is the first column of T and $\underline{v}^H$ is the first row of T^{-1}.

By construction, the first column of T is the right eigenvector $\underline{u}$ and satisfies $A\underline{u} = \lambda_1 \underline{u}$. Likewise, the first row of T^{-1} is the left eigenvector $\underline{v}^H$ associated with the eigenvalue 1 because $B^H = T^{-H} J^H T^H$ and the first column of T^{-H} (i.e., Hermitian conjugate of first row of T^{-1}) is the right eigenvector of A^H associated with λ_1. Therefore, $\underline{v}^H A = \lambda_1 \underline{v}^H$. Finally, the fact that $\underline{u} = B^n \underline{u} \to \underline{u}\underline{v}^H \underline{u}$ implies that $\underline{v}^H \underline{u} = 1$. □