

TOPICS IN ALGEBRA: EIGENVALUES AND EIGENVECTORS

PHIL MAYER

1. VECTOR SPACES AND BASES

Suppose V is a vector space. Given some vectors in V , we can add them with every nice property we might expect: associativity, an identity, inverses, and commutativity. So $(V, +)$ is a commutative group. We can also multiply vectors by scalars taken from a field, typically \mathbb{R} or \mathbb{C} . The standard examples of vector spaces that we will discuss are therefore \mathbb{R}^n and \mathbb{C}^n . Recall the following definitions about vector spaces:

Definition 1.1. A **basis** is a set $B = \{\vec{b}_1, \dots, \vec{b}_n\}$ of linearly independent vectors that generate the vector space. Any vector $\vec{v} \in V$ can be written as $\vec{v} = c_1\vec{b}_1 + \dots + c_n\vec{b}_n$, a linear combination of the basis vectors, in at least one way.

Definition 1.2. We say that a set of vectors is **linearly independent** if and only if any vector in V can be written as a linear combination of the basis vectors in at most one way.

So given some basis B , any vector \vec{v} in the vector space can be written as a combination of the basis vectors $\vec{b}_i \in B$ in *exactly* one way. Though we can chose to work in any basis we want, we typically work in the standard basis.

Definition 1.3. The **standard basis** for \mathbb{R}^n , denoted E or E_n , is the set of n unique vectors having $e_i = 1$ for exactly one index i and $e_j = 0$ for all other indices j . For example, in \mathbb{R}^2 , the standard basis is $E_2 = \{[1, 0], [0, 1]\}$.

Given an arbitrary vector $\vec{v} \in \mathbb{R}^n$, we can write \vec{v} as a linear combination of its basis vectors. For example, given $\vec{v} = [a, b]$ in \mathbb{R}^2 , we can write it as a linear combination of standard basis vectors in E_2 as follows:

$$[\vec{v}]_{E_2} = a[1, 0] + b[0, 1]$$

We can also write \vec{v} in terms of other bases. Suppose we want to express \vec{v} in terms of $B = \{[1, 1], [1, -1]\}$. In order to keep track of this change in representation, we denote \vec{v} with respect to B as $[\vec{v}]_B$. To compute \vec{v} 's new coordinates with respect to B , we solve the following linear system for (x, y) :

$$[\vec{v}]_B = x[1, 1] + y[1, -1]$$

We can quickly solve the system by encapsulating it into the following matrix and row-reducing.

$$\left[\begin{array}{cc|c} 1 & 1 & a \\ 1 & -1 & b \end{array} \right] \xrightarrow{rref} \left[\begin{array}{cc|c} 1 & 0 & \frac{a+b}{2} \\ 0 & 1 & \frac{a-b}{2} \end{array} \right]$$

Here \vec{v} 's new coordinates with respect to B are given on the right side of row the augmented matrix. Overall, we have:

$$\vec{v} = [a, b]_{E_2} = \left[\frac{a+b}{2}, \frac{a-b}{2} \right]_B$$

For example, $[3, 7]_{E_2} = [5, -2]_B$. We can also apply a similar process to change the representation of an arbitrary vector in B back to E_2 . First, we write $[\vec{v}]_B$ as a linear combination of its basis vectors:

$$\begin{aligned} \left[\frac{a+b}{2}, \frac{a-b}{2} \right]_B &= \frac{a+b}{2} \cdot [1, 1] + \frac{a-b}{2} \cdot [1, -1] \\ &= \left[\frac{a+b}{2} + \frac{a-b}{2}, \frac{a+b}{2} - \frac{a-b}{2} \right] \\ &= \left[\frac{2a}{2}, \frac{2b}{2} \right] \\ \left[\frac{a+b}{2}, \frac{a-b}{2} \right]_B &= [a, b]_{E_2} \end{aligned}$$

We will see another technique for quickly changing basis in the next section.

2. LINEAR TRANSFORMATIONS AND CHANGE OF BASIS

Suppose we have some vector space $V \subset \mathbb{R}^n$ and a linear transformation $T : V \rightarrow V$. Then $\forall \vec{u}, \vec{v} \in V$, $T(\vec{u} + \vec{v}) = T(\vec{u}) + T(\vec{v})$ and $\forall c \in \mathbb{R}$, $T(c\vec{v}) = cT(\vec{v})$. Linear transformations can also be represented as matrices. For example, consider the transformation:

$$T([x, y]) = [x + 2y, 3y]$$

This definition implicitly describes T 's action on vectors with respect to the standard basis. Generating its matrix representation (with respect to E_2), denoted $M_E(T)$, is therefore easy: we take a 2×2 matrix and place $T(\vec{e}_1)$ and $T(\vec{e}_2)$ down the columns:

$$\begin{aligned} M_E(T) &= [T(\vec{e}_1), T(\vec{e}_2)] \\ &= \left[T\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right), T\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) \right] \\ M_E(T) &= \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix} \end{aligned}$$

$M_E(T)$ provides an equivalent way to apply T to vectors in the standard basis. We can also express T as a matrix with respect to other bases. For example, consider the bases:

$$B_1 = \{[1, 0, 0], [1, 1, 0], [1, 1, 1]\} \text{ and } B_2 = \{[1, 1, 0], [1, 0, 1], [0, 1, 0]\}$$

and the linear transformation:

$$T([x, y, z]) = [2x - y, y + z, z - 3x]$$

In order to find the matrices $A = M_B(T)$ and $A' = M_{B'}(T)$, we first need to represent T 's action on vectors in the two bases. For each basis vector in B , we see that:

$$\begin{aligned} T(\vec{b}_1) &= T\left(\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 2 \\ 0 \\ -3 \end{bmatrix} \\ T(\vec{b}_2) &= T\left(\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ 1 \\ -3 \end{bmatrix} \\ T(\vec{b}_3) &= T\left(\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ 2 \\ -2 \end{bmatrix} \end{aligned}$$

But each of these vectors is still in the standard basis. We can quickly bring each of these vectors into B by multiplying by a change of basis, or change of coordinates, matrix.

Definition 2.1. The **change of basis** matrix $C_{B_1 \rightarrow B_2}$ brings vectors from B_1 to B_2 . For example, given $[\vec{v}]_{B_1}$:

$$[\vec{v}]_{B_2} = C_{B_1 \rightarrow B_2} \cdot [\vec{v}]_{B_1}$$

We derive the change of basis matrix from B_1 to B_2 by augmenting the matrix representation of B_2 (the new basis) with the matrix representation of B_1 (the old basis) and row-reducing.

$$[M_{B_2} \mid M_{B_1}] \xrightarrow{\text{rref}} [I_n \mid C_{B_1 \rightarrow B_2}]$$

Continuing our example, we find the change of basis matrix from the standard basis E_3 to B by:

$$[M_B \mid M_{E_3}] = \left[\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array} \right] \xrightarrow{\text{rref}} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & -1 & 0 \\ 0 & 1 & 0 & 0 & 1 & -1 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array} \right]$$

So our change of basis matrix is:

$$C_{E \rightarrow B} = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix}$$

We then multiply it by each of our previous results $T(\vec{b}_i)$ to get $[T(\vec{b}_i)]_B$. The matrix $A = M_B(T)$ is then given by putting each product down the columns of a new matrix. The same process can be applied to get $A' = M_{B'}(T)$. The two matrices are:

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

$$A' = \begin{bmatrix} 4 & 4 & -1 \\ -3 & -2 & 0 \\ -3 & -3 & 2 \end{bmatrix}$$

Given a vector $[\vec{v}]_B$, the product $A \cdot [\vec{v}]_B$ represents T 's action on \vec{v} with respect to B . Thus the product is a vector with respect B . There is a direct relationship between the matrices A and A' that allows for faster computation of either matrix. Though the process of generating the matrix for T with respect to either basis is fairly straightforward, it is cumbersome. The following theorem describes the relationship between the two matrices, giving a formula for quickly computing one matrix given the other.

Theorem 2.2. *Given two bases B and B' and the change of basis matrix $C_{B' \rightarrow B}$, the matrix representations of a linear transformation T with respect to both bases, $A = M_B(T)$ and $A' = M_{B'}(T)$, are related by:*

$$A' = (C_{B' \rightarrow B})^{-1} \cdot A \cdot C_{B' \rightarrow B}$$

It can also be shown that for a change of basis matrix $C_{B' \rightarrow B}$, its inverse acts in the opposite direction. So $(C_{B' \rightarrow B})^{-1} = C_{B \rightarrow B'}$. This gives us the following corollary.

Corollary 2.3. Under the same conditions as the previous theorem, A and A' are also related by:

$$A' = C_{B \rightarrow B'} \cdot A \cdot C_{B' \rightarrow B}$$

To finish our example, it can be seen that the two matrices are indeed related by both formulas.

3. EIGENVALUES AND EIGENVECTORS

Definition 3.1. Consider a linear transformation $T: \mathbb{R}^n \rightarrow \mathbb{R}^n$. Then λ is an **eigenvalue** for T if and only if $\exists \vec{v} \neq \vec{0}$ with $T(\vec{v}) = \lambda\vec{v}$ where \vec{v} is an **eigenvector** for T associated with λ . If A is a matrix with $A\vec{v} = \lambda\vec{v}$ and $\vec{v} \neq \vec{0}$, then \vec{v} is an eigenvector for A associated with eigenvalue λ .

We can find eigenvalues and eigenvectors associated with a matrix A (and thus its corresponding linear transformations as well) by solving the equation for λ :

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

The derivation is as follows:

$$\begin{aligned} A\vec{v} &= \lambda\vec{v} \\ A\vec{v} - \lambda\vec{v} &= \vec{0} \\ \vec{v} \cdot (A - \lambda I) &= \vec{0}, \text{ but since } \vec{v} \neq \vec{0}: \\ A - \lambda I &= \vec{0} \\ \det(A - \lambda I) &= \det(\vec{0}) \\ \det(A - \lambda I) &= 0 \end{aligned}$$

While $\vec{v} = \vec{0}$ is always a solution to $A\vec{v} = \lambda\vec{v}$, nontrivial eigenvalue-eigenvector solutions exist if and only if $\det(A - \lambda I) \neq 0$. Eigenvalues and eigenvectors also have the following properties:

Theorem 3.2. *If λ is an eigenvalue for a matrix A with eigenvector \vec{v} , then λ^k is an eigenvalue for A^k with eigenvector \vec{v} for some $k \in \mathbb{N}$ with $k > 0$. Symbolically:*

$$A^k \vec{v} = \lambda^k \vec{v}$$

Theorem 3.3. *If A is an invertible matrix and $\lambda \neq 0$ is an eigenvalue for A with eigenvector \vec{v} , then $\frac{1}{\lambda}$ is an eigenvalue for A^{-1} with eigenvector \vec{v} .*

Let's apply some of this information and find some eigenvalues and eigenvectors. Consider the matrix:

$$A = \begin{bmatrix} 2 & 3 \\ 2 & 1 \end{bmatrix}$$

Then the eigenvalues of A are given by expanding out $\det(A - \lambda I)$, factoring, and finding its roots.

$$\begin{aligned}\det(A - \lambda I) &= \begin{vmatrix} 2 - \lambda & 3 \\ 2 & 1 - \lambda \end{vmatrix} \\ &= (2 - \lambda)(1 - \lambda) - 2 \cdot 3 \\ &= \lambda^2 - 3\lambda - 4 \\ \det(A - \lambda I) &= (\lambda - 4)(\lambda + 1)\end{aligned}$$

We call this factorization the characteristic polynomial for A . It tells us that the eigenvalues for A are $\lambda = -1, 4$.

Definition 3.4. The **characteristic polynomial** for a matrix A , given by $p(\lambda) = \det(A - \lambda I)$, has the eigenvalues for A as its roots.

Notice that A was 2×2 , $p(\lambda)$ is of degree 2, and there are two eigenvalues for A . To find a basis for the space of eigenvectors associated with each λ , we solve the equation $(A - \lambda I_2)\vec{x} = \vec{0}$, plugging in λ . For instance, for $\lambda = 4$ we have:

$$\begin{bmatrix} -2 & 3 \\ 2 & -3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

If we augment $A - \lambda I_2$ with $\vec{0}$ and row-reduce, we only find the trivial solution $\vec{x} = \vec{0}$. To find nontrivial solutions, we solve the following system by hand:

$$\begin{aligned}-2x + 3y &= 0 \\ 2x - 3y &= 0\end{aligned}$$

Since the first equation is simply a scalar multiple of the second, we only need to find (x, y) solutions to $2x - 3y = 0$. One nontrivial solution that works is $\vec{v} = [3, 2]$. This solution \vec{v} is therefore an eigenvector for A corresponding to $\lambda = 4$. Now if we repeat the process for $\lambda = -1$, we have:

$$\begin{bmatrix} 3 & 3 \\ 2 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

and since $x + y = 0$, $\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$

So $\vec{v} = [1, -1]$ is an eigenvector corresponding to $\lambda = -1$. As we will see, having found the eigenvalues (and a basis for the eigenvectors associated with each eigenvalue) will be useful to us in the next section. In the meantime, we can also prove the following theorem relevant to the eigenvectors we found.

Definition 3.5. The **eigenspace** of an eigenvalue λ , denoted E_λ , is the set of eigenvectors corresponding to λ unioned with $\vec{0}$.

Theorem 3.6. *If λ is an eigenvalue for A , then the eigenspace E_λ is a vector space.*

Proof. Let λ be an eigenvalue for A .

Let $\vec{v}_1, \vec{v}_2 \in E_\lambda$.

Then $A \cdot (\vec{v}_1 + \vec{v}_2) = A\vec{v}_1 + A\vec{v}_2 = \lambda\vec{v}_1 + \lambda\vec{v}_2 = \lambda(\vec{v}_1 + \vec{v}_2)$, so $\vec{v}_1 + \vec{v}_2 \in E_\lambda$.

Now let $c \in \mathbb{R}$.

Then $A \cdot (c\vec{v}) = cA\vec{v} = c\lambda\vec{v} = \lambda(c\vec{v})$, so $c\vec{v} \in E_\lambda$.

Finally, $\vec{0} \in E_\lambda$ by definition.

So E_λ is a vector space. □

4. DIAGONALIZATION

Recall that matrix multiplication is a computationally expensive process. By hand, even small matrices are tedious to multiply. On a computer, even the best algorithms are trumped by the raw number of calculations required to multiply large matrices: even the best algorithms can't achieve an efficiency of $O(n^2)$. If we diagonalize a matrix, however, we can at least improve the process of raising it to positive powers, achieving $O(n)$ time complexity.

Theorem 4.1. *Suppose $\lambda_1, \dots, \lambda_n$ are scalars ($\lambda_i \in \mathbb{C}$), $\vec{v}_1, \dots, \vec{v}_n$ are nonzero vectors, and A is an $n \times n$ matrix. Then let $C = [\vec{v}_1, \dots, \vec{v}_n]$ be a matrix with the vectors down the columns. Let D be a matrix whose off-diagonal elements are 0 and its diagonal elements are $\lambda_1, \dots, \lambda_n$. Then $AC = CD$ if and only if $\lambda_1, \dots, \lambda_n$ are the eigenvalues of A with \vec{v}_i being an eigenvector associated with λ_i .*

Proof. Suppose $\lambda_1, \dots, \lambda_n$ are eigenvalues for A , and C, D are defined as above.

Then $AC = [A\vec{v}_1, \dots, A\vec{v}_n]$ and $CD = [\lambda_1\vec{v}_1, \dots, \lambda_n\vec{v}_n]$.

Now $AC = CD$ if and only if $\forall i$ with $0 < i \leq n$, $\lambda_i\vec{v}_i = A\vec{v}_i$.

But each λ_i is an eigenvalue for A , so $AC = CD$. □

Definition 4.2. Given two $n \times n$ matrices P and Q , we say that P is **similar** to Q if and only if $\exists C$ invertible such that $P = C^{-1}QC$. It can be shown that similarity is an equivalence relation. Furthermore, P and Q are similar if and only if they represent the same linear transformation relative to different bases.

We can now begin discussing diagonalization, starting with the following definition. We then present an important corollary to the previous theorem.

Definition 4.3. A is **diagonalizable** if and only if A is similar to a diagonal matrix.

Corollary 4.4. A is diagonalizable if and only if there exists a basis for \mathbb{R}^n (or \mathbb{C}^n) consisting of eigenvectors corresponding to eigenvalues of A .

For example, in the previous section we found that A had eigenvalues $\lambda = -1, 4$ with respective eigenvectors $[3, 2], [1, -1]$. We can then choose D, C such that:

$$D = \begin{bmatrix} -1 & 0 \\ 0 & 4 \end{bmatrix}, \quad C = \begin{bmatrix} 3 & 1 \\ 2 & -1 \end{bmatrix}$$

In this case, we can then invert C and observe that $A = C^{-1}AC$, so A is similar to D . In some cases, however, we can only find a matrix C where $AC = CD$. This case is typically caused by the eigenvectors being dependent: C becomes non-invertible.

As mentioned previously, the main benefit of diagonalization is that we can raise matrices to positive powers much more efficiently.

Theorem 4.5. *Suppose $A = C^{-1}DC$. Then for some $k \in \mathbb{N}$, $A^k = C^{-1}D^kC$.*

Proof.

$$\begin{aligned}
 A^k &= (C^{-1}DC)^k \\
 &= (C^{-1}DC)(C^{-1}DC)\cdots(C^{-1}DC) \\
 &= C^{-1} \cdot D \cdot (CC^{-1}) \cdot D \cdots D \cdot (CC^{-1}) \cdot D \cdot C \\
 &= C^{-1} \cdot D \cdots D \cdot C \\
 A^k &= C^{-1} \cdot D^k \cdot C
 \end{aligned}$$

□

We can quickly determine if a matrix is diagonalizable by employing some of the following theorems.

Theorem 4.6. *Suppose A has n distinct eigenvalues λ_i . Then $\{\vec{v}_1, \dots, \vec{v}_n\}$ are independent, so A is diagonalizable.*

To understand the next theorem about diagonalizability, we first define some additional properties of eigenvalues.

Definition 4.7. The **algebraic multiplicity** of an eigenvalue λ for a matrix A is the number of times λ appears as a root of the characteristic equation $p(\lambda)$ for A .

Definition 4.8. The **geometric multiplicity** of an eigenvalue λ for A is the dimension of λ 's eigenspace: $\dim(E_\lambda)$.

Theorem 4.9. *If A is a matrix with eigenvalue λ , then the geometric multiplicity of λ is less than or equal to its algebraic multiplicity.*

Theorem 4.10. *A matrix A is diagonalizable if and only if for all eigenvalues λ_i , the geometric multiplicity of λ_i equals its algebraic multiplicity.*

5. THE JORDAN-CANONICAL FORM

Though the previous section discussed conditions for a given matrix A to be similar to a diagonal matrix, sometimes A simply cannot be diagonalized. Fortunately, we will see that every matrix is similar to something almost as good: a Jordan canonical form. To begin investigating Jordan canonical forms, we first need to define a Jordan block.

Definition 5.1. A **Jordan block**, typically denoted J , is a matrix in which the diagonal elements are equal, the elements immediately above the diagonal are 1, and all other elements are 0.

Some examples of Jordan blocks include:

$$J_1 = [\lambda] \quad J_2 = \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix} \quad J_3 = \begin{bmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{bmatrix}$$

Definition 5.2. A **Jordan canonical form** is a matrix consisting of Jordan blocks down the diagonal, placed corner-to-corner. A 4×4 example of a Jordan canonical form (with two identical 2×2 blocks) would resemble:

$$J = \begin{bmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 2 \end{bmatrix}$$

Now given some $n \times n$ matrix A , we construct its similar Jordan canonical form by setting the diagonal elements of a new $n \times n$ matrix J equal to the eigenvalues of A (repeating some if necessary based on their algebraic multiplicities). We then examine the properties of A to determine the number of Jordan blocks in J and their respective sizes. We can also examine a Jordan canonical form and deduce facts about matrices that it might be similar to. For example, in the matrix above, we know that matrices similar to J would have $\lambda = 4$ as an eigenvalue repeated four times in its characteristic polynomial.

To further investigate how Jordan canonical forms behave, let's find a Jordan canonical form similar to the following matrix:

$$A = \begin{bmatrix} 3 & 0 & 1 & 1 \\ 0 & 3 & 0 & 2 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix}$$

It is clear that the eigenvalues for A are $\lambda = 3$, repeated four times. Now if we define $B = A - 3I$, then we have:

$$B = \begin{bmatrix} 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

We can now ask: what does B do to vectors in E_4 , the standard basis? If we multiply B by each vector, we see that:

$$B\vec{e}_1 = \vec{0}, \quad B\vec{e}_2 = \vec{0}, \quad B\vec{e}_3 = \vec{e}_1, \quad B\vec{e}_4 = \vec{e}_1 + 2\vec{e}_2$$

We can now begin assembling “strings” of vectors that more succinctly describe this behavior. At this point, the following strings provide the most useful information for us:

$$\vec{e}_3 \rightarrow \vec{e}_1 \rightarrow \vec{0}, \quad \vec{e}_2 \rightarrow \vec{0}$$

Since we can describe B 's action on the standard basis vectors by two strings, we can now claim that A 's similar Jordan canonical form J has two Jordan blocks. Interestingly, $\text{nullity}(B) = 2$ as well. We can almost fully characterize J ; we just need to know if its blocks are of sizes 3×3 and 1×1 , or 2×2 and 2×2 . To find out, we square B . In this case, it turns out that $B^2 = 0$, which means that $\text{nullity}(B^2) = 4$. In addition, B^2 's action on vectors in E_4 can be described by the following strings:

$$\vec{e}_3 \rightarrow \vec{e}_1 \rightarrow \vec{0}, \quad \vec{e}_4 \rightarrow \vec{e}_2 \rightarrow \vec{0}$$

So the Jordan canonical form J similar to A has two Jordan blocks, each 2×2 and corresponding to $\lambda = 3$:

$$J = \begin{bmatrix} 3 & 1 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 3 \end{bmatrix}$$

Overall, our findings are supported by the following theorem.

Theorem 5.3. *Let A be a matrix with eigenvalue λ . Let J be its corresponding Jordan canonical form. Then the number of Jordan blocks in J is equal to the number of strings, the geometric multiplicity of λ , and $\text{nullity}(J - \lambda I)$. Furthermore, the size of each block is equal to the length of its corresponding string.*

Furthermore, we can also state the major theorem summarizing our exploration of Jordan canonical forms.

Theorem 5.4. *Let A be a square matrix. Then there exists an invertible matrix C such that the matrix $J = C^{-1}AC$ is a Jordan canonical form. This Jordan canonical form is unique, except for the order of the Jordan blocks of which it is composed.*