The structure of an endomorphism

Similarity of $n \times n$ matrices

Similarity of $n \times n$ matrices

In mathematics, we study mathematical objects.

Similarity of $n \times n$ matrices

In mathematics, we study mathematical objects. A profitable way to approach this is to study how different objects differ.

Similarity of $n \times n$ matrices

In mathematics, we study mathematical objects. A profitable way to approach this is to study how different objects differ. We do this by studying structure-preserving ("comparison") maps $f : X \to Y$.

Similarity of $n \times n$ matrices

In mathematics, we study mathematical objects. A profitable way to approach this is to study how different objects differ. We do this by studying structure-preserving ("comparison") maps $f \colon X \to Y$. In linear algebra, this leads to the study of linear transformations $T \colon \mathbb{V} \to \mathbb{W}$,

Similarity of $n \times n$ matrices

In mathematics, we study mathematical objects. A profitable way to approach this is to study how different objects differ. We do this by studying structure-preserving (“comparison”) maps $f\colon X \to Y$. In linear algebra, this leads to the study of linear transformations $T\colon \mathbb{V} \to \mathbb{W}$, and, as a special case, endomorphisms $T\colon \mathbb{V} \to \mathbb{V}$.

Similarity of $n \times n$ matrices

In mathematics, we study mathematical objects. A profitable way to approach this is to study how different objects differ. We do this by studying structure-preserving ("comparison") maps $f\colon X \to Y$. In linear algebra, this leads to the study of linear transformations $T\colon \mathbb{V} \to \mathbb{W}$, and, as a special case, endomorphisms $T\colon \mathbb{V} \to \mathbb{V}$. We try to classify these endomorphisms up to "change of basis" in order to understand the essentially different possibilities.

Similarity of $n \times n$ matrices

In mathematics, we study mathematical objects. A profitable way to approach this is to study how different objects differ. We do this by studying structure-preserving ("comparison") maps $f: X \to Y$. In linear algebra, this leads to the study of linear transformations $T: \mathbb{V} \to \mathbb{W}$, and, as a special case, endomorphisms $T: \mathbb{V} \to \mathbb{V}$. We try to classify these endomorphisms up to "change of basis" in order to understand the essentially different possibilities. (E.g., "rotations through angle θ", "reflections", "scalings", "shears", etc.)

Similarity of $n \times n$ matrices

In mathematics, we study mathematical objects. A profitable way to approach this is to study how different objects differ. We do this by studying structure-preserving ("comparison") maps $f\colon X \to Y$. In linear algebra, this leads to the study of linear transformations $T\colon \mathbb{V} \to \mathbb{W}$, and, as a special case, endomorphisms $T\colon \mathbb{V} \to \mathbb{V}$. We try to classify these endomorphisms up to "change of basis" in order to understand the essentially different possibilities. (E.g., "rotations through angle θ", "reflections", "scalings", "shears", etc.) We write $A \sim B$ is A and B are matrices for the same endomorphism but with respect to different bases.

Similarity of $n \times n$ matrices

In mathematics, we study mathematical objects. A profitable way to approach this is to study how different objects differ. We do this by studying structure-preserving ("comparison") maps $f : X \to Y$. In linear algebra, this leads to the study of linear transformations $T : \mathbb{V} \to \mathbb{W}$, and, as a special case, endomorphisms $T : \mathbb{V} \to \mathbb{V}$. We try to classify these endomorphisms up to "change of basis" in order to understand the essentially different possibilities. (E.g., "rotations through angle θ", "reflections", "scalings", "shears", etc.) We write $A \sim B$ is A and B are matrices for the same endomorphism but with respect to different bases. (A is **similar** to B.)

Similarity of $n \times n$ matrices

In mathematics, we study mathematical objects. A profitable way to approach this is to study how different objects differ. We do this by studying structure-preserving ("comparison") maps $f\colon X \to Y$. In linear algebra, this leads to the study of linear transformations $T\colon \mathbb{V} \to \mathbb{W}$, and, as a special case, endomorphisms $T\colon \mathbb{V} \to \mathbb{V}$. We try to classify these endomorphisms up to "change of basis" in order to understand the essentially different possibilities. (E.g., "rotations through angle θ", "reflections", "scalings", "shears", etc.) We write $A \sim B$ is A and B are matrices for the same endomorphism but with respect to different bases. (A is **similar** to B.)

$A \sim B$ iff there is an invertible matrix C so that $B = C^{-1}AC$.

Similarity of $n \times n$ matrices

In mathematics, we study mathematical objects. A profitable way to approach this is to study how different objects differ. We do this by studying structure-preserving ("comparison") maps $f\colon X \to Y$. In linear algebra, this leads to the study of linear transformations $T\colon \mathbb{V} \to \mathbb{W}$, and, as a special case, endomorphisms $T\colon \mathbb{V} \to \mathbb{V}$. We try to classify these endomorphisms up to "change of basis" in order to understand the essentially different possibilities. (E.g., "rotations through angle θ", "reflections", "scalings", "shears", etc.) We write $A \sim B$ is A and B are matrices for the same endomorphism but with respect to different bases. (A is **similar** to B.)

$A \sim B$ iff there is an invertible matrix C so that $B = C^{-1}AC$. Another way to say this is "B is a conjugate of A", or "B is the conjugate of A by C".

Similarity of $n \times n$ matrices

In mathematics, we study mathematical objects. A profitable way to approach this is to study how different objects differ. We do this by studying structure-preserving ("comparison") maps $f \colon X \to Y$. In linear algebra, this leads to the study of linear transformations $T \colon \mathbb{V} \to \mathbb{W}$, and, as a special case, endomorphisms $T \colon \mathbb{V} \to \mathbb{V}$. We try to classify these endomorphisms up to "change of basis" in order to understand the essentially different possibilities. (E.g., "rotations through angle θ", "reflections", "scalings", "shears", etc.) We write $A \sim B$ is A and B are matrices for the same endomorphism but with respect to different bases. (A is **similar** to B.)

$A \sim B$ iff there is an invertible matrix C so that $B = C^{-1}AC$. Another way to say this is "B is a conjugate of A", or "B is the conjugate of A by C". (Any invertible matrix can be a change-of-basis matrix:

Similarity of $n \times n$ matrices

In mathematics, we study mathematical objects. A profitable way to approach this is to study how different objects differ. We do this by studying structure-preserving ("comparison") maps $f : X \to Y$. In linear algebra, this leads to the study of linear transformations $T : \mathbb{V} \to \mathbb{W}$, and, as a special case, endomorphisms $T : \mathbb{V} \to \mathbb{V}$. We try to classify these endomorphisms up to "change of basis" in order to understand the essentially different possibilities. (E.g., "rotations through angle θ", "reflections", "scalings", "shears", etc.) We write $A \sim B$ is A and B are matrices for the same endomorphism but with respect to different bases. (A is **similar** to B.)

$A \sim B$ iff there is an invertible matrix C so that $B = C^{-1}AC$. Another way to say this is "B is a conjugate of A", or "B is the conjugate of A by C". (Any invertible matrix can be a change-of-basis matrix: $C = [\mathbf{c}_1 \cdots \mathbf{c}_n]$ is invertible iff it is the C-o-B matrix from $(\mathbf{c}_1, \ldots, \mathbf{c}_n)$ to $(\mathbf{e}_1, \ldots, \mathbf{e}_n)$.)

Similarity of $n \times n$ matrices

In mathematics, we study mathematical objects. A profitable way to approach this is to study how different objects differ. We do this by studying structure-preserving ("comparison") maps $f\colon X \to Y$. In linear algebra, this leads to the study of linear transformations $T\colon \mathbb{V} \to \mathbb{W}$, and, as a special case, endomorphisms $T\colon \mathbb{V} \to \mathbb{V}$. We try to classify these endomorphisms up to "change of basis" in order to understand the essentially different possibilities. (E.g., "rotations through angle θ", "reflections", "scalings", "shears", etc.) We write $A \sim B$ is A and B are matrices for the same endomorphism but with respect to different bases. (A is **similar** to B.)

$A \sim B$ iff there is an invertible matrix C so that $B = C^{-1}AC$. Another way to say this is "B is a conjugate of A", or "B is the conjugate of A by C". (Any invertible matrix can be a change-of-basis matrix: $C = [\mathbf{c}_1 \cdots \mathbf{c}_n]$ is invertible iff it is the C-o-B matrix from $(\mathbf{c}_1, \ldots, \mathbf{c}_n)$ to $(\mathbf{e}_1, \ldots, \mathbf{e}_n)$.)

Note that "similarity" is an equivalence relation on the set of $n \times n$-matrices over $\mathbb{F}$.

Example

Example

It can be shown that rotations of $\mathbb{R}^3$ are linear, hence are representable by matrices.

Example

It can be shown that rotations of $\mathbb{R}^3$ are linear, hence are representable by matrices. By choosing a basis where the first basis vector is the axis of rotation, the other basis vectors lie in the equatorial plane, and the three basis vectors form a right handed system, any rotation of $\mathbb{R}^3$ will have a matrix of the form

Example

It can be shown that rotations of $\mathbb{R}^3$ are linear, hence are representable by matrices. By choosing a basis where the first basis vector is the axis of rotation, the other basis vectors lie in the equatorial plane, and the three basis vectors form a right handed system, any rotation of $\mathbb{R}^3$ will have a matrix of the form

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta) & -\sin(\theta) \\ 0 & \sin(\theta) & \cos(\theta) \end{bmatrix}.$$

Example

It can be shown that rotations of $\mathbb{R}^3$ are linear, hence are representable by matrices. By choosing a basis where the first basis vector is the axis of rotation, the other basis vectors lie in the equatorial plane, and the three basis vectors form a right handed system, any rotation of $\mathbb{R}^3$ will have a matrix of the form

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta) & -\sin(\theta) \\ 0 & \sin(\theta) & \cos(\theta) \end{bmatrix}.$$

E-values are $1, \cos(\theta) + i\sin(\theta), \cos(\theta) - i\sin(\theta)$.

Example

It can be shown that rotations of $\mathbb{R}^3$ are linear, hence are representable by matrices. By choosing a basis where the first basis vector is the axis of rotation, the other basis vectors lie in the equatorial plane, and the three basis vectors form a right handed system, any rotation of $\mathbb{R}^3$ will have a matrix of the form

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta) & -\sin(\theta) \\ 0 & \sin(\theta) & \cos(\theta) \end{bmatrix}.$$

E-values are $1, \cos(\theta) + i\sin(\theta), \cos(\theta) - i\sin(\theta)$. So, every rotation matrix in $\mathbb{R}^3$ is similar to a matrix of this kind, and no two different matrices of this kind are similar to each other.

Example

It can be shown that rotations of $\mathbb{R}^3$ are linear, hence are representable by matrices. By choosing a basis where the first basis vector is the axis of rotation, the other basis vectors lie in the equatorial plane, and the three basis vectors form a right handed system, any rotation of $\mathbb{R}^3$ will have a matrix of the form

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta) & -\sin(\theta) \\ 0 & \sin(\theta) & \cos(\theta) \end{bmatrix}.$$

E-values are $1, \cos(\theta) + i\sin(\theta), \cos(\theta) - i\sin(\theta)$. So, every rotation matrix in $\mathbb{R}^3$ is similar to a matrix of this kind, and no two different matrices of this kind are similar to each other. This shows that the "structure" of a rotation is classified by the angle of rotation.

The simplest case: $\mathbb{F}$ is algebraically closed

The simplest case: $\mathbb{F}$ is algebraically closed

A field is **algebraically closed** if every polynomial $p(x)$ with degree > 0 and leading coefficient 1 factors into linear factors which all lie in $\mathbb{F}$.

The simplest case: $\mathbb{F}$ is algebraically closed

A field is **algebraically closed** if every polynomial $p(x)$ with degree > 0 and leading coefficient 1 factors into linear factors which all lie in $\mathbb{F}$.

- $\mathbb{Q}$ is not algebraically closed, since $p(x) = x^2 - 2$ does not factor into linear factors with coefficients in $\mathbb{Q}$.

The simplest case: $\mathbb{F}$ is algebraically closed

A field is **algebraically closed** if every polynomial $p(x)$ with degree > 0 and leading coefficient 1 factors into linear factors which all lie in $\mathbb{F}$.

- $\mathbb{Q}$ is not algebraically closed, since $p(x) = x^2 - 2$ does not factor into linear factors with coefficients in $\mathbb{Q}$.

The simplest case: $\mathbb{F}$ is algebraically closed

A field is **algebraically closed** if every polynomial $p(x)$ with degree > 0 and leading coefficient 1 factors into linear factors which all lie in $\mathbb{F}$.

- $\mathbb{Q}$ is not algebraically closed, since $p(x) = x^2 - 2$ does not factor into linear factors with coefficients in $\mathbb{Q}$.
- $\mathbb{R}$ is not algebraically closed, since $p(x) = x^2 + 1$ does not factor into linear factors with coefficients in $\mathbb{R}$.

The simplest case: $\mathbb{F}$ is algebraically closed

A field is **algebraically closed** if every polynomial $p(x)$ with degree > 0 and leading coefficient 1 factors into linear factors which all lie in $\mathbb{F}$.

- $\mathbb{Q}$ is not algebraically closed, since $p(x) = x^2 - 2$ does not factor into linear factors with coefficients in $\mathbb{Q}$.
- $\mathbb{R}$ is not algebraically closed, since $p(x) = x^2 + 1$ does not factor into linear factors with coefficients in $\mathbb{R}$.

The simplest case: $\mathbb{F}$ is algebraically closed

A field is **algebraically closed** if every polynomial $p(x)$ with degree > 0 and leading coefficient 1 factors into linear factors which all lie in $\mathbb{F}$.

- $\mathbb{Q}$ is not algebraically closed, since $p(x) = x^2 - 2$ does not factor into linear factors with coefficients in $\mathbb{Q}$.
- $\mathbb{R}$ is not algebraically closed, since $p(x) = x^2 + 1$ does not factor into linear factors with coefficients in $\mathbb{R}$.

Fundamental Theorem of Algebra.

The simplest case: $\mathbb{F}$ is algebraically closed

A field is **algebraically closed** if every polynomial $p(x)$ with degree > 0 and leading coefficient 1 factors into linear factors which all lie in $\mathbb{F}$.

- $\mathbb{Q}$ is not algebraically closed, since $p(x) = x^2 - 2$ does not factor into linear factors with coefficients in $\mathbb{Q}$.
- $\mathbb{R}$ is not algebraically closed, since $p(x) = x^2 + 1$ does not factor into linear factors with coefficients in $\mathbb{R}$.

Fundamental Theorem of Algebra. $\mathbb{C}$ is algebraically closed.

The simplest case: $\mathbb{F}$ is algebraically closed

A field is **algebraically closed** if every polynomial $p(x)$ with degree > 0 and leading coefficient 1 factors into linear factors which all lie in $\mathbb{F}$.

- $\mathbb{Q}$ is not algebraically closed, since $p(x) = x^2 - 2$ does not factor into linear factors with coefficients in $\mathbb{Q}$.
- $\mathbb{R}$ is not algebraically closed, since $p(x) = x^2 + 1$ does not factor into linear factors with coefficients in $\mathbb{R}$.

Fundamental Theorem of Algebra. $\mathbb{C}$ is algebraically closed.

More fundamental theorem without a name. If $\mathbb{F}$ is any field, then $\mathbb{F}$ is contained in a larger field $\overline{\mathbb{F}}$ such that

1. $\overline{\mathbb{F}}$ is algebraically closed, and

The field $\overline{\mathbb{F}}$ is unique up to isomorphism.

The simplest case: $\mathbb{F}$ is algebraically closed

A field is **algebraically closed** if every polynomial $p(x)$ with degree > 0 and leading coefficient 1 factors into linear factors which all lie in $\mathbb{F}$.

- $\mathbb{Q}$ is not algebraically closed, since $p(x) = x^2 - 2$ does not factor into linear factors with coefficients in $\mathbb{Q}$.
- $\mathbb{R}$ is not algebraically closed, since $p(x) = x^2 + 1$ does not factor into linear factors with coefficients in $\mathbb{R}$.

Fundamental Theorem of Algebra. $\mathbb{C}$ is algebraically closed.

More fundamental theorem without a name. If $\mathbb{F}$ is any field, then $\mathbb{F}$ is contained in a larger field $\overline{\mathbb{F}}$ such that

1. $\overline{\mathbb{F}}$ is algebraically closed, and
2. every element of $\overline{\mathbb{F}}$ is a root of some polynomial with coefficients in $\mathbb{F}$.

The field $\overline{\mathbb{F}}$ is unique up to isomorphism.

Similarity over an algebraically closed field (ACF)

Similarity over an algebraically closed field (ACF)

Let A be a matrix with entries in some ACF, like $\mathbb{C}$.

Similarity over an algebraically closed field (ACF)

Let A be a matrix with entries in some ACF, like $\mathbb{C}$.
$\chi_A(\lambda) = (\lambda - r_1)^{m_1} \cdots (\lambda - r_k)^{m_k}$

Similarity over an algebraically closed field (ACF)

Let A be a matrix with entries in some ACF, like $\mathbb{C}$.
$\chi_A(\lambda) = (\lambda - r_1)^{m_1} \cdots (\lambda - r_k)^{m_k}$
The elements $r_1, \ldots, r_k$ are the e-values of A and r_i has algebraic multiplicity m_i.

Similarity over an algebraically closed field (ACF)

Let A be a matrix with entries in some ACF, like $\mathbb{C}$.
$\chi_A(\lambda) = (\lambda - r_1)^{m_1} \cdots (\lambda - r_k)^{m_k}$
The elements $r_1, \ldots, r_k$ are the e-values of A and r_i has algebraic multiplicity m_i.

Claim.

Similarity over an algebraically closed field (ACF)

Let A be a matrix with entries in some ACF, like $\mathbb{C}$.
$\chi_A(\lambda) = (\lambda - r_1)^{m_1} \cdots (\lambda - r_k)^{m_k}$
The elements $r_1, \ldots, r_k$ are the e-values of A and r_i has algebraic multiplicity m_i.

Claim. Similar matrices have the same characteristic polynomials.

Similarity over an algebraically closed field (ACF)

Let A be a matrix with entries in some ACF, like $\mathbb{C}$.
$\chi_A(\lambda) = (\lambda - r_1)^{m_1} \cdots (\lambda - r_k)^{m_k}$
The elements $r_1, \ldots, r_k$ are the e-values of A and r_i has algebraic multiplicity m_i.

Claim. Similar matrices have the same characteristic polynomials. (Why do we care?)

Similarity over an algebraically closed field (ACF)

Let A be a matrix with entries in some ACF, like $\mathbb{C}$.
$\chi_A(\lambda) = (\lambda - r_1)^{m_1} \cdots (\lambda - r_k)^{m_k}$
The elements $r_1, \ldots, r_k$ are the e-values of A and r_i has algebraic multiplicity m_i.

Claim. Similar matrices have the same characteristic polynomials. (Why do we care?)

Proof.

Similarity over an algebraically closed field (ACF)

Let A be a matrix with entries in some ACF, like $\mathbb{C}$.
$\chi_A(\lambda) = (\lambda - r_1)^{m_1} \cdots (\lambda - r_k)^{m_k}$
The elements $r_1, \ldots, r_k$ are the e-values of A and r_i has algebraic multiplicity m_i.

Claim. Similar matrices have the same characteristic polynomials. (Why do we care?)

Proof.
Assume $B = C^{-1}AC$.

Similarity over an algebraically closed field (ACF)

Let A be a matrix with entries in some ACF, like $\mathbb{C}$.
$\chi_A(\lambda) = (\lambda - r_1)^{m_1} \cdots (\lambda - r_k)^{m_k}$
The elements $r_1, \ldots, r_k$ are the e-values of A and r_i has algebraic multiplicity m_i.

Claim. Similar matrices have the same characteristic polynomials. (Why do we care?)

Proof.
Assume $B = C^{-1}AC$.

$$\chi_B(\lambda) \quad = \det(\lambda I - C^{-1}AC) = \det(C^{-1}(\lambda I - A)C)$$

Similarity over an algebraically closed field (ACF)

Let A be a matrix with entries in some ACF, like $\mathbb{C}$.
$\chi_A(\lambda) = (\lambda - r_1)^{m_1} \cdots (\lambda - r_k)^{m_k}$
The elements $r_1, \ldots, r_k$ are the e-values of A and r_i has algebraic multiplicity m_i.

Claim. Similar matrices have the same characteristic polynomials. (Why do we care?)

Proof.
Assume $B = C^{-1}AC$.

$$\begin{aligned}\chi_B(\lambda) &= \det(\lambda I - C^{-1}AC) = \det(C^{-1}(\lambda I - A)C) \\ &= \underline{\det(C^{-1})}\det(\lambda I - A)\underline{\det(C)} = \det(\lambda I - A) = \chi_A(\lambda)\end{aligned}$$

Similarity over an algebraically closed field (ACF)

Let A be a matrix with entries in some ACF, like $\mathbb{C}$.
$\chi_A(\lambda) = (\lambda - r_1)^{m_1} \cdots (\lambda - r_k)^{m_k}$
The elements $r_1, \ldots, r_k$ are the e-values of A and r_i has algebraic multiplicity m_i.

Claim. Similar matrices have the same characteristic polynomials. (Why do we care?)

Proof.
Assume $B = C^{-1}AC$.

$$\begin{aligned}\chi_B(\lambda) &= \det(\lambda I - C^{-1}AC) = \det(C^{-1}(\lambda I - A)C) \\ &= \underline{\det(C^{-1})}\det(\lambda I - A)\underline{\det(C)} = \det(\lambda I - A) = \chi_A(\lambda)\end{aligned}$$

It follows that if $A \sim B$, then A and B have the same e-values and the e-values have the same algebraic multiplicities.

Algebraic multiplicity versus geometric multiplicity

Claim.

Algebraic multiplicity versus geometric multiplicity

Claim. If $m = \dim(V_r)$, then $\lambda - r$ has multiplicity at least m as a factor of $\chi_A(\lambda)$.

Algebraic multiplicity versus geometric multiplicity

Claim. If $m = \dim(V_r)$, then $\lambda - r$ has multiplicity at least m as a factor of $\chi_A(\lambda)$.

Reason.

Algebraic multiplicity versus geometric multiplicity

Claim. If $m = \dim(V_r)$, then $\lambda - r$ has multiplicity at least m as a factor of $\chi_A(\lambda)$.

Reason. Choose a basis $\mathcal{B}$ for the subspace V_r and extend it to a basis $(\mathcal{B}, \mathcal{C})$ for the whole space.

Algebraic multiplicity versus geometric multiplicity

Claim. If $m = \dim(V_r)$, then $\lambda - r$ has multiplicity at least m as a factor of $\chi_A(\lambda)$.

Reason. Choose a basis $\mathcal{B}$ for the subspace V_r and extend it to a basis $(\mathcal{B}, \mathcal{C})$ for the whole space. With respect to this basis, A has a matrix $\begin{bmatrix} rI_m & X \\ \mathbf{0} & Y \end{bmatrix}$

Claim. If $m = \dim(V_r)$, then $\lambda - r$ has multiplicity at least m as a factor of $\chi_A(\lambda)$.

Reason. Choose a basis $\mathcal{B}$ for the subspace V_r and extend it to a basis $(\mathcal{B}, \mathcal{C})$ for the whole space. With respect to this basis, A has a matrix $\begin{bmatrix} rI_m & X \\ \mathbf{0} & Y \end{bmatrix}$ so

$$\begin{aligned}\chi_A(\lambda) &= \det\left(\lambda I - \begin{bmatrix} rI_m & X \\ \mathbf{0} & Y \end{bmatrix}\right) = \det\left(\begin{bmatrix} (\lambda - r)I_m & -X \\ \mathbf{0} & \lambda I - Y \end{bmatrix}\right) \\ &= \det\left(\left[(\lambda - r)I_m\right]\right) \cdot \det\left(\left[\lambda I - Y\right]\right) = (\lambda - r)^m \chi_Y(\lambda),\end{aligned}$$

Algebraic multiplicity versus geometric multiplicity

Claim. If $m = \dim(V_r)$, then $\lambda - r$ has multiplicity at least m as a factor of $\chi_A(\lambda)$.

Reason. Choose a basis $\mathcal{B}$ for the subspace V_r and extend it to a basis $(\mathcal{B}, \mathcal{C})$ for the whole space. With respect to this basis, A has a matrix $\begin{bmatrix} rI_m & X \\ \mathbf{0} & Y \end{bmatrix}$ so

$$\begin{aligned} \chi_A(\lambda) &= \det\left(\lambda I - \begin{bmatrix} rI_m & X \\ \mathbf{0} & Y \end{bmatrix}\right) = \det\left(\begin{bmatrix} (\lambda - r)I_m & -X \\ \mathbf{0} & \lambda I - Y \end{bmatrix}\right) \\ &= \det(\left[(\lambda - r)I_m\right]) \cdot \det(\left[\lambda I - Y\right]) = (\lambda - r)^m \chi_Y(\lambda), \end{aligned}$$

so $(\lambda - r)^m$ is a factor of $\chi_A(\lambda)$. This shows that the algebraic multiplicity of r is at least as large as the geometric multiplicity.

Distinct eigenspaces are independent

Distinct eigenspaces are independent

Claim.

Distinct eigenspaces are independent

Claim. Suppose that the e-values of A are $r_1, \ldots, r_k$.

Distinct eigenspaces are independent

Claim. Suppose that the e-values of A are $r_1, \ldots, r_k$. If $\mathcal{B}_i$ is a basis for V_{r_i}, then $(\mathcal{B}_1, \ldots, \mathcal{B}_k)$ is independent.

Distinct eigenspaces are independent

Claim. Suppose that the e-values of A are $r_1, \ldots, r_k$. If $\mathcal{B}_i$ is a basis for V_{r_i}, then $(\mathcal{B}_1, \ldots, \mathcal{B}_k)$ is independent.

Reason.

Distinct eigenspaces are independent

Claim. Suppose that the e-values of A are $r_1, \ldots, r_k$. If $\mathcal{B}_i$ is a basis for V_{r_i}, then $(\mathcal{B}_1, \ldots, \mathcal{B}_k)$ is independent.

Reason.
Assume the claim is false, and write down a dependence relation of shortest length:

Distinct eigenspaces are independent

Claim. Suppose that the e-values of A are $r_1, \ldots, r_k$. If $\mathcal{B}_i$ is a basis for V_{r_i}, then $(\mathcal{B}_1, \ldots, \mathcal{B}_k)$ is independent.

Reason.
Assume the claim is false, and write down a dependence relation of shortest length:

$$c_1\mathbf{b}_1 + \cdots + c_u\mathbf{b}_u = \mathbf{0}.$$

Distinct eigenspaces are independent

Claim. Suppose that the e-values of A are $r_1, \ldots, r_k$. If $\mathcal{B}_i$ is a basis for V_{r_i}, then $(\mathcal{B}_1, \ldots, \mathcal{B}_k)$ is independent.

Reason.
Assume the claim is false, and write down a dependence relation of shortest length:

$$c_1\mathbf{b}_1 + \cdots + c_u\mathbf{b}_u = \mathbf{0}.$$

Multiply A and obtain a new dependence relation

Distinct eigenspaces are independent

Claim. Suppose that the e-values of A are $r_1, \ldots, r_k$. If $\mathcal{B}_i$ is a basis for V_{r_i}, then $(\mathcal{B}_1, \ldots, \mathcal{B}_k)$ is independent.

Reason.
Assume the claim is false, and write down a dependence relation of shortest length:

$$c_1 \mathbf{b}_1 + \cdots + c_u \mathbf{b}_u = \mathbf{0}.$$

Multiply A and obtain a new dependence relation

$$c_1 r_{i_1} \mathbf{b}_1 + \cdots + c_u r_{i_u} \mathbf{b}_u = \mathbf{0}.$$

Distinct eigenspaces are independent

Claim. Suppose that the e-values of A are $r_1, \ldots, r_k$. If $\mathcal{B}_i$ is a basis for V_{r_i}, then $(\mathcal{B}_1, \ldots, \mathcal{B}_k)$ is independent.

Reason.
Assume the claim is false, and write down a dependence relation of shortest length:

$$c_1\mathbf{b}_1 + \cdots + c_u\mathbf{b}_u = \mathbf{0}.$$

Multiply A and obtain a new dependence relation

$$c_1 r_{i_1}\mathbf{b}_1 + \cdots + c_u r_{i_u}\mathbf{b}_u = \mathbf{0}.$$

The r's cannot all be equal, so this is an essentially new dependence relation.

Distinct eigenspaces are independent

Claim. Suppose that the e-values of A are $r_1, \ldots, r_k$. If $\mathcal{B}_i$ is a basis for V_{r_i}, then $(\mathcal{B}_1, \ldots, \mathcal{B}_k)$ is independent.

Reason.
Assume the claim is false, and write down a dependence relation of shortest length:

$$c_1\mathbf{b}_1 + \cdots + c_u\mathbf{b}_u = \mathbf{0}.$$

Multiply A and obtain a new dependence relation

$$c_1 r_{i_1}\mathbf{b}_1 + \cdots + c_u r_{i_u}\mathbf{b}_u = \mathbf{0}.$$

The r's cannot all be equal, so this is an essentially new dependence relation. Use the dependence relations together to find a shorter one.

Distinct eigenspaces are independent

Claim. Suppose that the e-values of A are $r_1, \ldots, r_k$. If $\mathcal{B}_i$ is a basis for V_{r_i}, then $(\mathcal{B}_1, \ldots, \mathcal{B}_k)$ is independent.

Reason.
Assume the claim is false, and write down a dependence relation of shortest length:

$$c_1\mathbf{b}_1 + \cdots + c_u\mathbf{b}_u = \mathbf{0}.$$

Multiply A and obtain a new dependence relation

$$c_1 r_{i_1}\mathbf{b}_1 + \cdots + c_u r_{i_u}\mathbf{b}_u = \mathbf{0}.$$

The r's cannot all be equal, so this is an essentially new dependence relation. Use the dependence relations together to find a shorter one. Contradiction!

Distinct eigenspaces are independent

Claim. Suppose that the e-values of A are $r_1, \ldots, r_k$. If $\mathcal{B}_i$ is a basis for V_{r_i}, then $(\mathcal{B}_1, \ldots, \mathcal{B}_k)$ is independent.

Reason.
Assume the claim is false, and write down a dependence relation of shortest length:

$$c_1\mathbf{b}_1 + \cdots + c_u\mathbf{b}_u = \mathbf{0}.$$

Multiply A and obtain a new dependence relation

$$c_1 r_{i_1}\mathbf{b}_1 + \cdots + c_u r_{i_u}\mathbf{b}_u = \mathbf{0}.$$

The r's cannot all be equal, so this is an essentially new dependence relation. Use the dependence relations together to find a shorter one. Contradiction!

Consequence.

Distinct eigenspaces are independent

Claim. Suppose that the e-values of A are $r_1, \ldots, r_k$. If $\mathcal{B}_i$ is a basis for V_{r_i}, then $(\mathcal{B}_1, \ldots, \mathcal{B}_k)$ is independent.

Reason.
Assume the claim is false, and write down a dependence relation of shortest length:

$$c_1\mathbf{b}_1 + \cdots + c_u\mathbf{b}_u = \mathbf{0}.$$

Multiply A and obtain a new dependence relation

$$c_1 r_{i_1}\mathbf{b}_1 + \cdots + c_u r_{i_u}\mathbf{b}_u = \mathbf{0}.$$

The r's cannot all be equal, so this is an essentially new dependence relation. Use the dependence relations together to find a shorter one. Contradiction!

Consequence. A is diagonalizable

Distinct eigenspaces are independent

Claim. Suppose that the e-values of A are $r_1, \ldots, r_k$. If $\mathcal{B}_i$ is a basis for V_{r_i}, then $(\mathcal{B}_1, \ldots, \mathcal{B}_k)$ is independent.

Reason.
Assume the claim is false, and write down a dependence relation of shortest length:

$$c_1\mathbf{b}_1 + \cdots + c_u\mathbf{b}_u = \mathbf{0}.$$

Multiply A and obtain a new dependence relation

$$c_1 r_{i_1}\mathbf{b}_1 + \cdots + c_u r_{i_u}\mathbf{b}_u = \mathbf{0}.$$

The r's cannot all be equal, so this is an essentially new dependence relation. Use the dependence relations together to find a shorter one. Contradiction!

Consequence. A is diagonalizable iff $\mathbb{V}$ has a basis of e-vectors for A

Distinct eigenspaces are independent

Claim. Suppose that the e-values of A are $r_1, \ldots, r_k$. If $\mathcal{B}_i$ is a basis for V_{r_i}, then $(\mathcal{B}_1, \ldots, \mathcal{B}_k)$ is independent.

Reason.
Assume the claim is false, and write down a dependence relation of shortest length:

$$c_1\mathbf{b}_1 + \cdots + c_u\mathbf{b}_u = \mathbf{0}.$$

Multiply A and obtain a new dependence relation

$$c_1 r_{i_1}\mathbf{b}_1 + \cdots + c_u r_{i_u}\mathbf{b}_u = \mathbf{0}.$$

The r's cannot all be equal, so this is an essentially new dependence relation. Use the dependence relations together to find a shorter one. Contradiction!

Consequence. A is diagonalizable iff $\mathbb{V}$ has a basis of e-vectors for A iff geometric multiplicity = algebraic multiplicity for each e-value of A.

Generalized e-vectors

Generalized e-vectors

A vector v is a e-vector e-value 0 for A iff $Av = 0v = \mathbf{0}$

Generalized e-vectors

A vector v is a e-vector e-value 0 for A iff $Av = 0v = \mathbf{0}$ iff $v \in \text{null}(A)$.

Generalized e-vectors

A vector v is a e-vector e-value 0 for A iff $Av = 0v = \mathbf{0}$ iff $v \in \mathrm{null}(A)$. We call v a “generalized e-vector” for e-value 0 if $v \in \mathrm{null}(A^k)$ for some k.

Generalized e-vectors

A vector v is a e-vector e-value 0 for A iff $Av = 0v = \mathbf{0}$ iff $v \in \text{null}(A)$. We call v a "generalized e-vector" for e-value 0 if $v \in \text{null}(A^k)$ for some k. More generally,

Generalized e-vectors

A vector v is a e-vector e-value 0 for A iff $Av = 0v = \mathbf{0}$ iff $v \in \text{null}(A)$. We call v a "generalized e-vector" for e-value 0 if $v \in \text{null}(A^k)$ for some k. More generally,

Definition. v is a **generalized e-vector** for e-value λ if

$v \in \text{null}((A - \lambda I)^k) = \text{null}((A - \lambda I)^k)$ for some k.

Generalized e-vectors

A vector v is a e-vector e-value 0 for A iff $Av = 0v = \mathbf{0}$ iff $v \in \text{null}(A)$. We call v a "generalized e-vector" for e-value 0 if $v \in \text{null}(A^k)$ for some k. More generally,

Definition. v is a **generalized e-vector** for e-value λ if

$v \in \text{null}((A - \lambda I)^k) = \text{null}((A - \lambda I)^k)$ for some k.

Facts.

Generalized e-vectors

A vector v is a e-vector e-value 0 for A iff $Av = 0v = \mathbf{0}$ iff $v \in \text{null}(A)$. We call v a "generalized e-vector" for e-value 0 if $v \in \text{null}(A^k)$ for some k. More generally,

Definition. v is a **generalized e-vector** for e-value λ if

$v \in \text{null}((A - \lambda I)^k) = \text{null}((A - \lambda I)^k)$ for some k.

Facts.

1. If A is an $n \times n$ matrix over an algebraically closed field $\mathbb{F}$, then $\mathbb{V} = \mathbb{F}^n$ has a basis of generalized e-vectors.

Generalized e-vectors

A vector v is a e-vector e-value 0 for A iff $Av = 0v = \mathbf{0}$ iff $v \in \text{null}(A)$. We call v a "generalized e-vector" for e-value 0 if $v \in \text{null}(A^k)$ for some k. More generally,

Definition. v is a **generalized e-vector** for e-value λ if

$v \in \text{null}((A - \lambda I)^k) = \text{null}((A - \lambda I)^k)$ for some k.

Facts.

1. If A is an $n \times n$ matrix over an algebraically closed field $\mathbb{F}$, then $\mathbb{V} = \mathbb{F}^n$ has a basis of generalized e-vectors.

Generalized e-vectors

A vector v is a e-vector e-value 0 for A iff $Av = 0v = \mathbf{0}$ iff $v \in \text{null}(A)$. We call v a "generalized e-vector" for e-value 0 if $v \in \text{null}(A^k)$ for some k. More generally,

Definition. v is a **generalized e-vector** for e-value λ if

$v \in \text{null}((A - \lambda I)^k) = \text{null}((A - \lambda I)^k)$ for some k.

Facts.

1. If A is an $n \times n$ matrix over an algebraically closed field $\mathbb{F}$, then $\mathbb{V} = \mathbb{F}^n$ has a basis of generalized e-vectors.
2. With respect to the basis in the previous item, A has an almost-diagonal form, called the **Jordan form** for A.

Generalized e-vectors

A vector v is a e-vector e-value 0 for A iff $Av = 0v = \mathbf{0}$ iff $v \in \text{null}(A)$. We call v a "generalized e-vector" for e-value 0 if $v \in \text{null}(A^k)$ for some k. More generally,

Definition. v is a **generalized e-vector** for e-value λ if

$v \in \text{null}((A - \lambda I)^k) = \text{null}((A - \lambda I)^k)$ for some k.

Facts.

1. If A is an $n \times n$ matrix over an algebraically closed field $\mathbb{F}$, then $\mathbb{V} = \mathbb{F}^n$ has a basis of generalized e-vectors.
2. With respect to the basis in the previous item, A has an almost-diagonal form, called the **Jordan form** for A.

Generalized e-vectors

A vector v is a e-vector e-value 0 for A iff $Av = 0v = \mathbf{0}$ iff $v \in \text{null}(A)$. We call v a "generalized e-vector" for e-value 0 if $v \in \text{null}(A^k)$ for some k. More generally,

Definition. v is a **generalized e-vector** for e-value λ if

$v \in \text{null}((A - \lambda I)^k) = \text{null}((A - \lambda I)^k)$ for some k.

Facts.

1. If A is an $n \times n$ matrix over an algebraically closed field $\mathbb{F}$, then $\mathbb{V} = \mathbb{F}^n$ has a basis of generalized e-vectors.
2. With respect to the basis in the previous item, A has an almost-diagonal form, called the **Jordan form** for A.

Example.

Generalized e-vectors

A vector v is a e-vector e-value 0 for A iff $Av = 0v = \mathbf{0}$ iff $v \in \text{null}(A)$. We call v a "generalized e-vector" for e-value 0 if $v \in \text{null}(A^k)$ for some k. More generally,

Definition. v is a **generalized e-vector** for e-value λ if

$v \in \text{null}((A - \lambda I)^k) = \text{null}((A - \lambda I)^k)$ for some k.

Facts.

1. If A is an $n \times n$ matrix over an algebraically closed field $\mathbb{F}$, then $\mathbb{V} = \mathbb{F}^n$ has a basis of generalized e-vectors.
2. With respect to the basis in the previous item, A has an almost-diagonal form, called the **Jordan form** for A.

Example. $A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$ is nondiagonalizable, but in Jordan form.