Again, for this lecture all our vector spaces are over the field $\mathbb{R}$.

Consider the $n \times n$ matrix A and a scalar $\lambda \in \mathbb{R}$.

$$\lambda \text{ is an eigenvalue of } A \iff \begin{matrix} A\vec{v} = \lambda\vec{v} \\ \exists \vec{v} \neq 0 \end{matrix} \iff A\vec{v} = \lambda I_n \vec{v} \iff$$

$$\iff (A - \lambda I_n)\vec{v} = 0 \iff \vec{v} \in \ker(A - \lambda I_n) \iff$$

$$\iff A - \lambda I_n \text{ not invertible} \iff \det(A - \lambda I_n) = 0$$

Eigenvalues and determinants; characteristic equation

Consider an $n \times n$ matrix A and a scalar λ. Then λ is an eigenvalue of A if (and only if)

$$\det(A - \lambda I_n) = 0.$$

This is called the *characteristic equation* (or the *secular equation*) of matrix A.

Example: Find the eigenvalues of the matrix

$$A = \begin{pmatrix} 1 & 2 \\ 4 & 3 \end{pmatrix}$$

$$\det(A - \lambda I_2) = \det\left(\begin{pmatrix} 1 & 2 \\ 4 & 3 \end{pmatrix} - \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix}\right) = \det\left(\begin{pmatrix} 1-\lambda & 2 \\ 4 & 3-\lambda \end{pmatrix}\right)$$

$$= (1-\lambda)(3-\lambda) - 2\cdot 4 = \lambda^2 - 4\lambda - 5 = (\lambda - 5)(\lambda + 1) = 0$$

So, $\det(A - \lambda I_2) = 0$ holds for $\lambda_1 = 5$ and $\lambda = -1$.

These two numbers are eigenvalues of A.

These two numbers are eigenvalues of A.

In general, for $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, $\det(A - \lambda I_2) =$

$$= \det \begin{pmatrix} a-\lambda & b \\ c & d-\lambda \end{pmatrix} = (a-\lambda)(d-\lambda) - bc = \lambda^2 - (a+d)\lambda + (ad-bc) = 0$$

Now,

Trace

The sum of the diagonal entries of a square matrix A is called the *trace* of A, denoted by tr A.

So, for 2×2 matrix A, the characteristic equation has the form

$$\lambda^2 - \operatorname{tr} A \cdot \lambda + \det A = 0$$

This can be generalized:

Characteristic polynomial

If A is an $n \times n$ matrix, then $\det(A - \lambda I_n)$ is a polynomial of degree n, of the form

$$\begin{aligned}&(-\lambda)^n + (\operatorname{tr} A)(-\lambda)^{n-1} + \cdots + \det A \\ &\quad = (-1)^n \lambda^n + (-1)^{n-1}(\operatorname{tr} A)\lambda^{n-1} + \cdots + \det A.\end{aligned}$$

This is called the *characteristic polynomial* of A, denoted by $f_A(\lambda)$.

Proof:

$$f_A(\lambda) = \det(A - \lambda I_n) = \det \begin{pmatrix} a_{11}-\lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22}-\lambda & \cdots & a_{2n} \\ & & & \\ a_{n1} & a_{n2} & \cdots & a_{nn}-\lambda \end{pmatrix}$$

Now any term in the expansion $f_A(\lambda) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} a_{i\sigma(i)}$, Now any term in the expansion $f_A(\lambda) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{i=1} a_{i\sigma(i)}$ is a polynomial of degree $\leq n$, so $f_A(\lambda)$ is itself a polynomial of degree $\leq n$. Now, consider the term in the sum corresponding to the diagonal

$$(a_{11}-\lambda)(a_{22}-\lambda)\cdots(a_{nn}-\lambda) = (-\lambda)^n + (a_{11}+a_{22}+\cdots+a_{nn})(-\lambda)^{n-1} +$$

$$+(\text{polynomial of degree } \leq n-2) =$$

$$= (-1)^n \lambda^n + (-1)^{n-1} \lambda^{n-1} \operatorname{tr} A + (\text{polyn. of degree } \leq n-2)$$

Any other term in the sum for determinant involves at least two scalars off the diagonal why? and its product is therefore a polynomial of degree less then equal to $n-2$. This implies that

$$f_A(\lambda) = (-1)^n \lambda^n + \operatorname{tr} A \, (-1)^{n-1} \lambda^{n-1} + (\text{pol. of degree } \leq n-2).$$

Of course, the constant term of the polynomial is

$$f_A(0) = \det(A - 0 \cdot I_n) = \det(A). \qquad \square$$

Corollary: $n \times n$ matrix has at most n eigenvalues.
If n is odd, $n \times n$ matrix has at least one eigenvalue.

Example: Find all eigenvalues of $A = \begin{pmatrix} 5 & 4 & 3 \\ 0 & 5 & 3 \\ 0 & 0 & 4 \end{pmatrix}$. Now,

$$\det(A - \lambda I_3) = \det \begin{pmatrix} 5-\lambda & 4 & 3 \\ 0 & 5-\lambda & 3 \end{pmatrix} = (5-\lambda)^2(4-\lambda)$$

$$\det(A-\lambda I_3) = \det\begin{pmatrix} 5-\lambda & 4 & 3 \\ 0 & 5-\lambda & 3 \\ 0 & 0 & 4-\lambda \end{pmatrix} = (5-\lambda)^2(4-\lambda)$$
$\lambda_0 = 5, \quad \lambda_1 = 4.$

Algebraic multiplicity of an eigenvalue

We say that an eigenvalue λ_0 of a square matrix A has *algebraic multiplicity* k if λ_0 is a root of multiplicity k of the characteristic polynomial $f_A(\lambda)$, meaning that we can write

$$f_A(\lambda) = (\lambda_0 - \lambda)^k g(\lambda)$$

for some polynomial $g(\lambda)$ with $g(\lambda_0) \neq 0$. We write $\mathrm{almu}(\lambda_0) = k$.

So, in our example $\mathrm{almu}(5) = 2$ and $\mathrm{almu}(4) = 1$.

Since $f_A(\lambda) = (\underset{\uparrow \lambda_0}{5}-\lambda)^{\overset{k}{2}}\underbrace{(4-\lambda)}_{g(\lambda)}$.

Of course, it follows that $n \times n$ matrix has at most n eigenvalues, even if they are counted with their algebraic multiplicities.

Eigenvalues, determinant, and trace

If an $n \times n$ matrix A has the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$, listed with their algebraic multiplicities, then

$$\det A = \lambda_1 \lambda_2 \cdots \lambda_n, \quad \text{the product of the eigenvalues}$$

and

$$\operatorname{tr} A = \lambda_1 + \lambda_2 + \cdots + \lambda_n, \quad \text{the sum of the eigenvalues.}$$

Proof: In this case, characteristic polynomial factors completely as

$$f(\lambda) = \det(A - \lambda I_n) = (\lambda_1 - \lambda)(\lambda_2 - \lambda)\cdots(\lambda_n - \lambda)$$

$$f_A(\lambda) = \det(A - \lambda I_n) = (\lambda_1 - \lambda)(\lambda_2 - \lambda) \cdots (\lambda_n - \lambda)$$

Now, $f_A(0) = \det(A) = \lambda_1 \cdots \lambda_n$ as claimed. The case with trace is an exercise. (CE).

Q: Having found the eigenvalue λ of $n \times n$ matrix A, how do we find the corresponding eigenvectors?

We want to find vectors $v \in \mathbb{R}^n$ such that $A\vec{v} = \lambda\vec{v}$, or $(A - \lambda I_n)\vec{v} = \vec{0}$. In other words we need to find the kernel of the matrix $A - \lambda I_n$.

Eigenspaces

Consider an eigenvalue λ of an $n \times n$ matrix A. Then the kernel of the matrix $A - \lambda I_n$ is called the *eigenspace* associated with λ, denoted by E_λ:

$$E_\lambda = \ker(A - \lambda I_n) = \{\vec{v} \text{ in } \mathbb{R}^n : A\vec{v} = \lambda\vec{v}\}.$$

Remark: Eigenvectors with eigenvalue λ are the nonzero vectors in the eigenspace E_λ.

Example: Find the eigenspaces of the matrix $A = \begin{pmatrix} 1 & 2 \\ 4 & 3 \end{pmatrix}$.
We already saw that eigenvalues of this matrix are -5 and -1. Now,

$$E_5 = \ker\left(\begin{pmatrix} 1 & 2 \\ 4 & 3 \end{pmatrix} - \begin{pmatrix} 5 & 0 \\ 0 & 5 \end{pmatrix}\right) = \ker\begin{pmatrix} -4 & 2 \\ 4 & -2 \end{pmatrix}$$

$\begin{pmatrix} x \\ y \end{pmatrix} \in E_5 \iff \begin{pmatrix} -4 & 2 \\ \end{pmatrix}\begin{pmatrix} x \\ \end{pmatrix} = \begin{pmatrix} -4x + 2y \\ \end{pmatrix} = 0$

$$\begin{pmatrix} x \\ y \end{pmatrix} \in E_5 \iff \begin{pmatrix} -4 & 2 \\ 4 & -2 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -4x+2y \\ 4x-2y \end{pmatrix} = 0$$

$$\iff \begin{cases} -4x+2y = 0 \\ 4x - 2y = 0 \end{cases} \quad \left| \begin{array}{l} 2x = y \\ 0 = 0 \end{array} \right. \qquad \begin{pmatrix} x \\ y \end{pmatrix} = t\begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

So, $E_5 = \operatorname{span}\begin{pmatrix} 1 \\ 2 \end{pmatrix}$. Similarly,

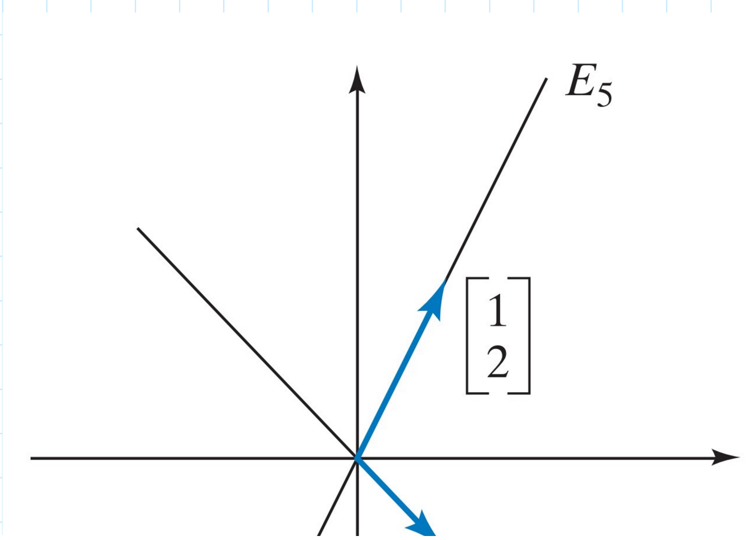
$$E_{-1} = \ker\begin{pmatrix} 2 & 2 \\ 4 & 4 \end{pmatrix}, \quad \begin{pmatrix} x \\ y \end{pmatrix} \in \ker\begin{pmatrix} 2 & 2 \\ 4 & 4 \end{pmatrix}$$

$$\iff \begin{cases} 2x + 2y = 0 \\ 4x + 4y = 0 \end{cases} \iff \begin{pmatrix} x \\ y \end{pmatrix} = t\begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

So, $E_{-1} = \operatorname{span}\begin{pmatrix} 1 \\ 1 \end{pmatrix}$. Indeed checking:

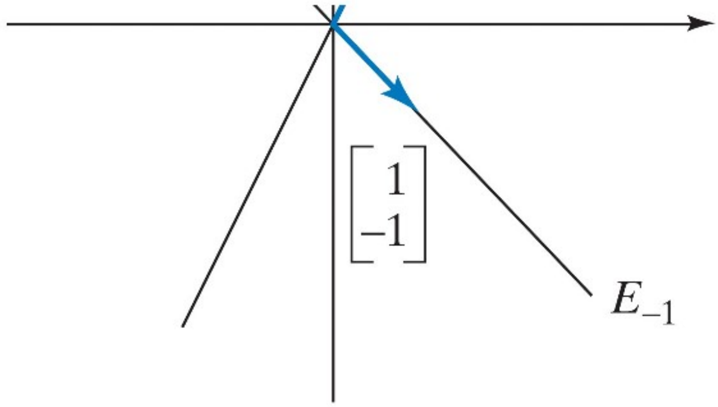
$$\begin{pmatrix} 1 & 2 \\ 4 & 3 \end{pmatrix}\begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 5 \\ 10 \end{pmatrix} = 5\begin{pmatrix} 1 \\ 2 \end{pmatrix} \quad \text{and}$$

$$\begin{pmatrix} 1 & 2 \\ 4 & 3 \end{pmatrix}\begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \end{pmatrix} = (-1)\begin{pmatrix} 1 \\ -1 \end{pmatrix}$$



The vectors $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$ and $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ form an eigenbasis of A, so A is diagonalizable with

$S = \begin{pmatrix} 1 & 1 \\ 2 & -1 \end{pmatrix}$ and



$S = \begin{pmatrix} 1 & 1 \\ 2 & -1 \end{pmatrix}$ and

$B = \begin{pmatrix} 5 & 0 \\ 0 & -1 \end{pmatrix}$. That

is $A = S B S^{-1}$.

Geomerically, the matrix A represents a scaling by factor 5 along the line spanned by vector $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$, while the line spanned by $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ is flipped over the origin.

Example: Find the eigenspaces of the matrix describing the orthogonal projection in 3 dimensions onto x-y plane.

Denote this matrix by A. $Ae_1 = e_1$, $Ae_2 = e_2$ and $Ae_3 = 0$. Therefore

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$ and A is diagonal with

eigenvalues $\lambda_1 = 1$ and $\lambda_2 = 0$.

$$E_1 = \ker(A - I_3) = \ker \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \ker(A - I_3) \iff \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = 0 \iff \begin{pmatrix} x \\ y \\ z \end{pmatrix} = t \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + s \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

So, E_1 is an x-y plane ($\text{span}(e_1, e_2)$).

$E_0 = \ker(A - 0 \cdot I_3) = \ker(A)$

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \ker(A) \iff \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = 0 \iff \begin{pmatrix} x \\ y \\ z \end{pmatrix} = t\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

So $E_0 = \mathrm{span}(e_3)$.

Example: Find eigenspaces of $A = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix}$ and diagonalize A if possible.

$$\det(A - \lambda I_3) = \det\begin{pmatrix} 1-\lambda & 1 & 1 \\ 0 & -\lambda & 1 \\ 0 & 0 & 1-\lambda \end{pmatrix} = -\lambda(1-\lambda)^2 \stackrel{!}{=} 0$$

$\lambda_1 = 0$ and $\lambda_2 = 1$ with $\mathrm{almu}(0) = 1$, $\mathrm{almu}(1) = 2$.

$E_0 = \ker(A - 0 \cdot I_3) = \ker A$.

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \ker A \iff \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = 0 \iff$$

$$\iff \begin{cases} x = -y \\ z = 0 \end{cases} \iff \begin{pmatrix} x \\ y \\ z \end{pmatrix} = t\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}$$

So, $E_0 = \mathrm{span}\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}$.

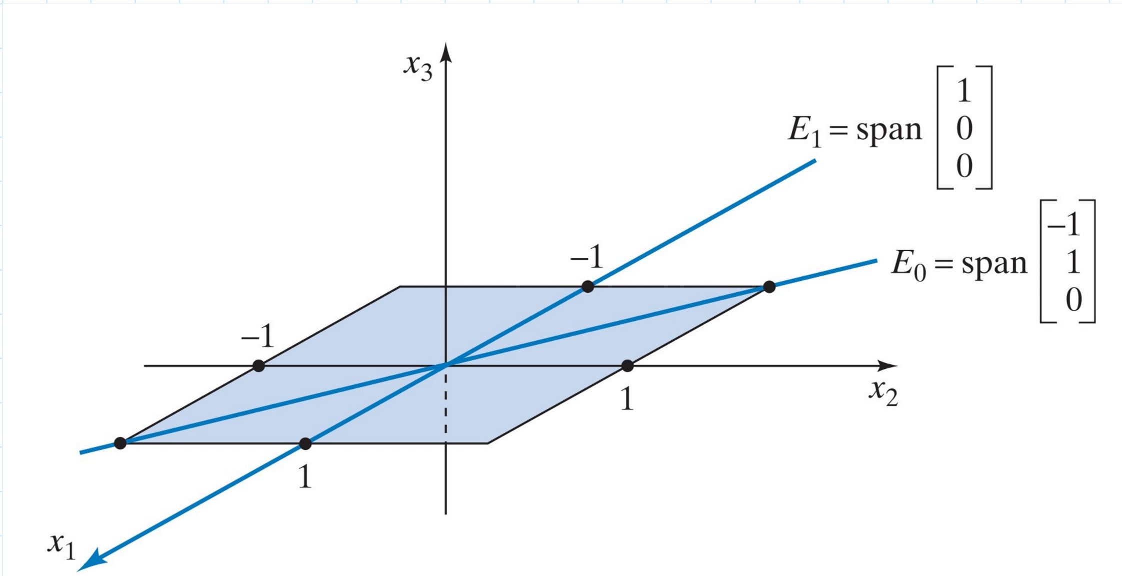
$$E_1 = \ker(A - 1 \cdot I_3) = \ker\begin{pmatrix} 0 & 1 & 1 \\ 0 & -1 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \ker\begin{pmatrix} 0 & 1 & 1 \\ 0 & -1 & 1 \\ 0 & 0 & 0 \end{pmatrix} \iff \begin{cases} y + z = 0 \\ -y + z = 0 \\ 0 = 0 \end{cases} \iff \begin{pmatrix} x \\ y \\ z \end{pmatrix} = t\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

$$\begin{pmatrix} z \end{pmatrix} \quad \begin{pmatrix} 0 & 0 & 0 \end{pmatrix} \quad 0=0 \quad \begin{pmatrix} z \end{pmatrix} \quad \begin{pmatrix} 0 \end{pmatrix}$$

So, $E_1 = \operatorname{span}\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$.

Both eigenspaces are in x-y plane. We can only find two linearly independent eigenvectors, one in each of the eigenspaces E_0 and E_1, so we are not able to construct the eigenbasis of A. Thus matrix A fails to be diagonalizable.



Now in an example about the projection on x-y plane we had two eigenvalues 0 and 1 and the matrix was diagonal. In the last example, again we had eigenvalues 0 and 1, but matrix was not diagonalizable at all. To discuss such cases, it is useful to introduce:

Geometric multiplicity

Consider an eigenvalue λ of an $n \times n$ matrix A. The dimension of eigenspace $E_\lambda = \ker(A - \lambda I_n)$ is called the *geometric multiplicity* of eigenvalue λ, denoted $\operatorname{gemu}(\lambda)$. Thus,

$$\operatorname{gemu}(\lambda) = \operatorname{nullity}(A - \lambda I_n) = n - \operatorname{rank}(A - \lambda I_n).$$

Geometric multiplicity

Consider an eigenvalue λ of an $n \times n$ matrix A. The dimension of eigenspace $E_\lambda = \ker(A - \lambda I_n)$ is called the *geometric multiplicity* of eigenvalue λ, denoted $\text{gemu}(\lambda)$. Thus,

$$\text{gemu}(\lambda) = \text{nullity}(A - \lambda I_n) = n - \text{rank}(A - \lambda I_n)$$

In general, as our last example shows, algebraic and geometric multiplicities differ.

Q1: Which square matrices are diagonalizable? That is, when does the eigenbasis of A exist?

Q2: If eigenbasis exists, how can we find one?

Eigenbases and geometric multiplicities

a. Consider an $n \times n$ matrix A. If we find a basis of each eigenspace of A and concatenate all these bases, then the resulting eigenvectors $\vec{v}_1, \ldots, \vec{v}_s$ will be linearly independent. (Note that s is the sum of the geometric multiplicities of the eigenvalues of A.) This result implies that $s \leq n$.

b. Matrix A is diagonalizable if (and only if) the geometric multiplicities of the eigenvalues add up to n (meaning that $s = n$ in part a).

Proof: a) Assume the set $V = \{\vec{v}_1, \ldots, \vec{v}_s\}$ is linearly dependent.
Since each $\vec{v}_i \neq 0$, every dependent subset of $V = \{\vec{v}_1, \ldots, \vec{v}_s\}$ must contain at least two eigenvectors.
If there is such a dependent pair, choose it. If not, ask if there is a dependent set of three vectors in V. If yes choose it. If no, ask if there is a dependent

V. If yes choose it. If no, ask if there is a dependent set of 4 vectors in V and so on.

Eventually, we will arrive at the set of j eigenvectors which is dependent and such that no $j-1$ vector subset of this set is dependent. By renumering we may assume that the set is $\{\vec{v}_1, \dots, \vec{v}_j\} \subset V$.

Since they are dependent we have

$$a_1\vec{v}_1 + \cdots + a_j\vec{v}_j = 0$$ for some constant

a_i such that not all vanish. By renumbering we may assume $a_j \neq 0$. Then with $b_i = -\frac{a_i}{a_j}$, we have a

(1) $\vec{v}_j = b_1\vec{v}_1 + \cdots + b_{j-1}\vec{v}_{j-1}$. Since $\vec{v}_j \neq 0$, not all b_i are zero. Multiplieng (1) by a matrix A on both sides and denoting $A\vec{v}_i = \lambda_i\vec{v}_i$ we get

$$\lambda_j\vec{v}_j = \lambda_1 b_1\vec{v}_1 + \cdots + \lambda_{j-1}b_{j-1}\vec{v}_{j-1} \qquad (2)$$

Multiplieng (1) by λ_j on both sides, we get

$$\lambda_j\vec{v}_j = \lambda_j b_1\vec{v}_1 + \cdots + \lambda_j b_{j-1}\vec{v}_{j-1} \qquad (3)$$

Substracting (2) from (3) to fing

$$0 = (\lambda_1 - \lambda_j)b_1\vec{v}_1 + \cdots + (\lambda_{j-1}-\lambda_j)b_j\vec{v}_{j-1} \qquad (4)$$

Now, there must be at least one nonzero

Now, there must be at least one nonzero b_k, such that $\lambda_j \neq \lambda_k$ because $\vec{v}_j$ can not be expressed as the linear combination of vectors $\vec{v}_i$ that are all in the same eigenspace E_{λ_j}.
Therefore (4) means $\{\vec{v}_1, \dots, \vec{v}_{j-1}\}$ are linearly dependent. A contradiction.

b) Follows directly from part a). There exists an eigenbasis if and only if $s = n$ in part a)

Corollary:

An $n \times n$ matrix with n distinct eigenvalues

If an $n \times n$ matrix A has n distinct eigenvalues, then A is diagonalizable. We can construct an eigenbasis by finding an eigenvector for each eigenvalue.