MATH 1030 Chapter 16

The lecture is based on Beezer, A first course in Linear algebra. Ver 3.5 Downloadable at http://linear.ups.edu/download.html .

The print version can be downloaded at http://linear.ups.edu/download/fcla-3.50-print.pdf .

Reference.

Beezer, Ver 3.5 Subsection EEM (print version p283-285), Subsection CEE and ECEE (print version p289-297)

Exercise.

Exercises with solutions can be downloaded at http://linear.ups.edu/download/fcla-3.50-solution-manual.pdf

(Replace $\mathbb{C}$ by $\mathbb{R}$)Section EE, p103-108, all except C22, M60. Note that for the questions regarding diagonalizability, use our method instead of the method in the solution manual.

16.1 Eigenvalues and Eigenvectors of a Matrix

Definition 16.1 (Eigenvalues and Eigenvectors of a Matrix)**.** Suppose that A is a square matrix of size n, $\mathbf{x}$ a non-zero vector in $\mathbb{R}^n$, and λ a scalar in $\mathbb{R}$. We say $\mathbf{x}$ is an **eigenvector** of A with **eigenvalue** λ if

$$A\mathbf{x} = \lambda\mathbf{x}.$$

Example 16.2. Let

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

and let

$$\mathbf{u} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \ \mathbf{v} = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \ \mathbf{w} = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}.$$

Then

$$A\mathbf{u} = \begin{bmatrix}4\\4\\4\end{bmatrix} = 4\mathbf{u}, \;\; A\mathbf{v} = \begin{bmatrix}1\\-1\\0\end{bmatrix} = 1\mathbf{v}, \;\; A\mathbf{w} = \begin{bmatrix}1\\0\\-1\end{bmatrix} = 1\mathbf{w}.$$

So $\mathbf{u}$ is an eigenvector of A with eigenvalue 4, $\mathbf{v}$ is an eigenvector of A with eigenvalue 1, and $\mathbf{w}$ is an eigenvector A with eigenvalue 1. Now let $\mathbf{x} = 100\mathbf{u}$. Then

$$A\mathbf{x} = 100A\mathbf{u} = 400\mathbf{u} = 4\mathbf{x}.$$

So $\mathbf{x}$ is an eigenvector of A with eigenvalue 4.Next let $\mathbf{y} = \mathbf{v} + \mathbf{w}$, then

$$A\mathbf{y} = A\mathbf{v} + A\mathbf{w} = \mathbf{v} + \mathbf{w} = 1\mathbf{y}.$$

So $\mathbf{y}$ is an eigenvector of A with eigenvalue 1. Finally, let $\mathbf{z} = \mathbf{u} + \mathbf{v} = \begin{bmatrix}2\\0\\1\end{bmatrix}$.

Then

$$A\mathbf{z} = 4\mathbf{u} + \mathbf{v} = \begin{bmatrix}5\\3\\4\end{bmatrix}$$

is not a scalar multiple of $\mathbf{z}$. So $\mathbf{z}$ is not an eigenvector. This shows that sum of eigenvectors need not be an eigenvector.

Example 16.3. Consider the matrix

$$A = \begin{bmatrix}204 & 98 & -26 & -10\\ -280 & -134 & 36 & 14\\ 716 & 348 & -90 & -36\\ -472 & -232 & 60 & 28\end{bmatrix}$$

and the vectors

$$\mathbf{x} = \begin{bmatrix}1\\-1\\2\\5\end{bmatrix} \mathbf{y} = \begin{bmatrix}-3\\4\\-10\\4\end{bmatrix} \qquad \mathbf{z} = \begin{bmatrix}-3\\7\\0\\8\end{bmatrix} \mathbf{w} = \begin{bmatrix}1\\-1\\4\\0\end{bmatrix}.$$

Then

$$A\mathbf{x} = \begin{bmatrix}204 & 98 & -26 & -10\\ -280 & -134 & 36 & 14\\ 716 & 348 & -90 & -36\\ -472 & -232 & 60 & 28\end{bmatrix} \begin{bmatrix}1\\-1\\2\\5\end{bmatrix} = \begin{bmatrix}4\\-4\\8\\20\end{bmatrix} = 4\begin{bmatrix}1\\-1\\2\\5\end{bmatrix} = 4\mathbf{x}$$

so that $\mathbf{x}$ is an eigenvector of A with eigenvalue $\lambda = 4$.

Also,

$$A\mathbf{y} = \begin{bmatrix} 204 & 98 & -26 & -10 \\ -280 & -134 & 36 & 14 \\ 716 & 348 & -90 & -36 \\ -472 & -232 & 60 & 28 \end{bmatrix} \begin{bmatrix} -3 \\ 4 \\ -10 \\ 4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} = 0 \begin{bmatrix} -3 \\ 4 \\ -10 \\ 4 \end{bmatrix} = 0\mathbf{y}$$

so that $\mathbf{y}$ is an eigenvector of A with eigenvalue $\lambda = 0$.

Also,

$$A\mathbf{z} = \begin{bmatrix} 204 & 98 & -26 & -10 \\ -280 & -134 & 36 & 14 \\ 716 & 348 & -90 & -36 \\ -472 & -232 & 60 & 28 \end{bmatrix} \begin{bmatrix} -3 \\ 7 \\ 0 \\ 8 \end{bmatrix} = \begin{bmatrix} -6 \\ 14 \\ 0 \\ 16 \end{bmatrix} = 2 \begin{bmatrix} -3 \\ 7 \\ 0 \\ 8 \end{bmatrix} = 2\mathbf{z}$$

so that $\mathbf{z}$ is an eigenvector of A with eigenvalue $\lambda = 2$.

Finally,

$$A\mathbf{w} = \begin{bmatrix} 204 & 98 & -26 & -10 \\ -280 & -134 & 36 & 14 \\ 716 & 348 & -90 & -36 \\ -472 & -232 & 60 & 28 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \\ 4 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ -2 \\ 8 \\ 0 \end{bmatrix} = 2 \begin{bmatrix} 1 \\ -1 \\ 4 \\ 0 \end{bmatrix} = 2\mathbf{w}$$

so that $\mathbf{w}$ is an eigenvector of A with eigenvalue $\lambda = 2$.

We have demonstrated four eigenvectors of A. Are there more? Yes, any nonzero scalar multiple of an eigenvector is again an eigenvector. In this example, setting $\mathbf{u} = 30\mathbf{x}$, we have

$$A\mathbf{u} = A(30\mathbf{x}) = 30A\mathbf{x} = 30(4\mathbf{x}) = 4(30\mathbf{x}) = 4\mathbf{u}$$

so that $\mathbf{u}$ is also an eigenvector of A with the same eigenvalue, $\lambda = 4$.

The vectors $\mathbf{z}$ and $\mathbf{w}$ are both eigenvectors of A for the same eigenvalue $\lambda = 2$, yet this is not as simple as the two vectors just being scalar multiples of each other (they are not). Look what happens when we add them together, forming $\mathbf{v} = \mathbf{z} + \mathbf{w}$, which we then multiply by A:

$$A\mathbf{v} = A(\mathbf{z} + \mathbf{w}) = A\mathbf{z} + A\mathbf{w}$$

$$= 2\mathbf{z} + 2\mathbf{w} = 2(\mathbf{z} + \mathbf{w}) = 2\mathbf{v}.$$

Hence, $\mathbf{v}$ is also an eigenvector of A with eigenvalue $\lambda = 2$. It would appear that the set of eigenvectors that are associated with a fixed eigenvalue is closed under the vector space operations of $\mathbb{R}^n$.

The vector $\mathbf{y}$ is an eigenvector of A for the eigenvalue $\lambda = 0$, so $A\mathbf{y} = 0\mathbf{y} = \mathbf{0}$. But this also means that $\mathbf{y} \in \mathcal{N}(A)$. There would appear to be a connection here also.

Definition 16.4 (Eigenspace of a Matrix)**.** Suppose A is a square matrix and λ is an eigenvalue of A. Then the **eigenspace** of A for λ, denoted by $\mathcal{E}_A(\lambda)$, is the set of all eigenvectors of A with eigenvalue λ, together with the zero vector.

Theorem 16.5 (Eigenspace of a Matrix is a Null Space)**.** *Suppose that A is a square matrix of size n and λ is an eigenvalue of A. Then*

$$\mathcal{E}_A(\lambda) = \mathcal{N}(A - \lambda I_n).$$

In particular, $\mathcal{E}_A(\lambda)$ is a subspace of $\mathbb{R}^n$.

Proof of Eigenspace of a Matrix is a Null Space. First, notice that $\mathbf{0} \in \mathcal{E}_A(\lambda)$ (by definition) and $\mathbf{0} \in \mathcal{N}(A - \lambda I_n)$.

For any nonzero vector $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{x} \in \mathcal{E}_A(\lambda) \iff A\mathbf{x} = \lambda\mathbf{x}$
$\iff A\mathbf{x} - \lambda\mathbf{x} = \mathbf{0}$
$\iff A\mathbf{x} - \lambda I_n\mathbf{x} = \mathbf{0}$
$\iff (A - \lambda I_n)\,\mathbf{x} = \mathbf{0}$
$\iff \mathbf{x} \in \mathcal{N}(A - \lambda I_n)$.
This completes the proof. □

16.2 Existence of Eigenvalues and Eigenvectors

Definition 16.6 (Characteristic Polynomial)**.** Suppose that A is a square matrix of size n. Then the **characteristic polynomial** of A is the polynomial $p_A(x)$ defined by:

$$p_A(x) = \det(A - xI_n)$$

Theorem 16.7 (Degree of the Characteristic Polynomial)**.** *Suppose that A is a square matrix of size n. Then the characteristic polynomial $p_A(x)$ has degree n.*

Proof of Degree of the Characteristic Polynomial. You can skip the proof. The following briefly explains why the theorem is true. It is not a rigorous proof.We have

$$p_A(x) = \begin{vmatrix} a_{11} - x & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - x & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} - x \end{vmatrix}.$$

The determinant is a sum of products of entries of $A - xI_n$, and all such products have degree at most $n - 1$, except the product of the diagonal entries,

$$(a_{11} - x)(a_{22} - x)\cdots(a_{nn} - x),$$

which has degree n.

Remark: We can also see that the leading coefficient is $(-1)^n$. $\square$

Theorem 16.8 (Eigenvalues of a Matrix are Roots of Characteristic Polynomials)**.** *Suppose that A is a square matrix. Then λ is an eigenvalue of A if and only if $p_A(\lambda) = 0$.*

Proof of Eigenvalues of a Matrix are Roots of Characteristic Polynomials. A number λ is an eigenvalue of A if and only if $\mathcal{E}_A(\lambda) = \mathcal{N}(A - \lambda I_n)$ is not the zero vector space $\{\mathbf{0}\}$.

$\iff A - \lambda I_n$ is singular

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

$\square$

Example 16.9. Consider

$$F = \begin{bmatrix} -13 & -8 & -4 \\ 12 & 7 & 4 \\ 24 & 16 & 7 \end{bmatrix}.$$

We compute

$$p_F(x) = \det(F - xI_3)$$

$$= \begin{vmatrix} -13-x & -8 & -4 \\ 12 & 7-x & 4 \\ 24 & 16 & 7-x \end{vmatrix}$$

$$= (-13-x)\begin{vmatrix} 7-x & 4 \\ 16 & 7-x \end{vmatrix} + (-8)(-1)\begin{vmatrix} 12 & 4 \\ 24 & 7-x \end{vmatrix} + (-4)\begin{vmatrix} 12 & 7-x \\ 24 & 16 \end{vmatrix}$$

$$= (-13-x)((7-x)(7-x) - 4(16))$$

$$+(-8)(-1)(12(7-x) - 4(24)) + (-4)(12(16) - (7-x)(24))$$

$$= 3 + 5x + x^2 - x^3$$

$$= -(x-3)(x+1)^2.$$

Example 16.10. In Example 16.9, we found the characteristic polynomial of

$$F = \begin{bmatrix} -13 & -8 & -4 \\ 12 & 7 & 4 \\ 24 & 16 & 7 \end{bmatrix}$$

to be $p_F\left(x\right) = -(x-3)(x+1)^2$. Being written in factored form, we can simply read off its roots; they are $x = 3$ and $x = -1$. By the previous theorem, $\lambda = 3$ and $\lambda = -1$ are both eigenvalues of F. Moreover, these are the only eigenvalues of F.

Example 16.11. Example 16.9 and Example 16.10 describe the characteristic polynomial and eigenvalues of the 3×3 matrix

$$F = \begin{bmatrix} -13 & -8 & -4 \\ 12 & 7 & 4 \\ 24 & 16 & 7 \end{bmatrix}.$$

We will now take each eigenvalue in turn and compute its eigenspace. To do this, we row-reduce the matrix $F - \lambda I_3$ in order to find all solutions to the homogeneous system $F - \lambda I_3\mathbf{x} = \mathbf{0}$. We then express the eigenspace $\mathcal{E}_F\left(\lambda\right)$ as the nullspace of $F - \lambda I_3$. Then we can write the nullspace as the span of a basis.

$$\lambda = 3: \qquad F - 3I_3 = \begin{bmatrix} -16 & -8 & -4 \\ 12 & 4 & 4 \\ 24 & 16 & 4 \end{bmatrix} \xrightarrow{\text{RREF}} \begin{bmatrix} \boxed{1} & 0 & \frac{1}{2} \\ 0 & \boxed{1} & -\frac{1}{2} \\ 0 & 0 & 0 \end{bmatrix}$$

$$\mathcal{E}_F\left(3\right) = \mathcal{N}\left(F - 3I_3\right) = \text{Span}\left\{\begin{bmatrix} -\frac{1}{2} \\ \frac{1}{2} \\ 1 \end{bmatrix}\right\} = \text{Span}\left\{\begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix}\right\}$$

$$\lambda = -1: \qquad F + 1I_3 = \begin{bmatrix} -12 & -8 & -4 \\ 12 & 8 & 4 \\ 24 & 16 & 8 \end{bmatrix} \xrightarrow{\text{RREF}} \begin{bmatrix} \boxed{1} & \frac{2}{3} & \frac{1}{3} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\mathcal{E}_F\left(-1\right) = \mathcal{N}\left(F + 1I_3\right) = \text{Span}\left\{\begin{bmatrix} -\frac{2}{3} \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -\frac{1}{3} \\ 0 \\ 1 \end{bmatrix}\right\} = \text{Span}\left\{\begin{bmatrix} -2 \\ 3 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 3 \end{bmatrix}\right\}$$

Eigenspaces in hand, we can easily compute eigenvectors by forming nontrivial linear combinations of the basis vectors describing each eigenspace. In particular, notice that we can pretty up our basis vectors by using scalar multiples to clear out fractions.

16.3 Examples of Computing Eigenvalues and Eigenvectors

Example 16.12. Consider the matrix

$$B = \begin{bmatrix} -2 & 1 & -2 & -4 \\ 12 & 1 & 4 & 9 \\ 6 & 5 & -2 & -4 \\ 3 & -4 & 5 & 10 \end{bmatrix}.$$

Then

$$p_B(x) = 8 - 20x + 18x^2 - 7x^3 + x^4 = (x-1)(x-2)^3.$$

So the eigenvalues are $\lambda = 1,\, 2$. Computing eigenvectors, we find

$$\lambda = 1: \qquad B - 1I_4 = \begin{bmatrix} -3 & 1 & -2 & -4 \\ 12 & 0 & 4 & 9 \\ 6 & 5 & -3 & -4 \\ 3 & -4 & 5 & 9 \end{bmatrix} \xrightarrow{\text{RREF}} \begin{bmatrix} \boxed{1} & 0 & \frac{1}{3} & 0 \\ 0 & \boxed{1} & -1 & 0 \\ 0 & 0 & 0 & \boxed{1} \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\mathcal{E}_B(1) = \mathcal{N}(B - 1I_4) = \text{Span}\left\{ \begin{bmatrix} -\frac{1}{3} \\ 1 \\ 1 \\ 0 \end{bmatrix} \right\} = \text{Span}\left\{ \begin{bmatrix} -1 \\ 3 \\ 3 \\ 0 \end{bmatrix} \right\}$$

$$\lambda = 2: \qquad B - 2I_4 = \begin{bmatrix} -4 & 1 & -2 & -4 \\ 12 & -1 & 4 & 9 \\ 6 & 5 & -4 & -4 \\ 3 & -4 & 5 & 8 \end{bmatrix} \xrightarrow{\text{RREF}} \begin{bmatrix} \boxed{1} & 0 & 0 & 1/2 \\ 0 & \boxed{1} & 0 & -1 \\ 0 & 0 & \boxed{1} & 1/2 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\mathcal{E}_B(2) = \mathcal{N}(B - 2I_4) = \text{Span}\left\{ \begin{bmatrix} -\frac{1}{2} \\ 1 \\ -\frac{1}{2} \\ 1 \end{bmatrix} \right\} = \text{Span}\left\{ \begin{bmatrix} -1 \\ 2 \\ -1 \\ 2 \end{bmatrix} \right\}$$

Example 16.13. Consider the matrix

$$C = \begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{bmatrix}.$$

Then

$$p_C(x) = -3 + 4x + 2x^2 - 4x^3 + x^4 = (x-3)(x-1)^2(x+1).$$

So the eigenvalues are $\lambda = 3,\ 1,\ -1$. Computing eigenvectors, we find

$$\lambda = 3: \qquad C - 3I_4 = \begin{bmatrix} -2 & 0 & 1 & 1 \\ 0 & -2 & 1 & 1 \\ 1 & 1 & -2 & 0 \\ 1 & 1 & 0 & -2 \end{bmatrix} \xrightarrow{\text{RREF}} \begin{bmatrix} \boxed{1} & 0 & 0 & -1 \\ 0 & \boxed{1} & 0 & -1 \\ 0 & 0 & \boxed{1} & -1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\mathcal{E}_C(3) = \mathcal{N}(C - 3I_4) = \operatorname{Span}\left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \right\}$$

$$\lambda = 1: \qquad C - 1I_4 = \begin{bmatrix} 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{bmatrix} \xrightarrow{\text{RREF}} \begin{bmatrix} \boxed{1} & 1 & 0 & 0 \\ 0 & 0 & \boxed{1} & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\mathcal{E}_C(1) = \mathcal{N}(C - 1I_4) = \operatorname{Span}\left\{ \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ -1 \\ 1 \end{bmatrix} \right\}$$

$$\lambda = -1: \qquad C + 1I_4 = \begin{bmatrix} 2 & 0 & 1 & 1 \\ 0 & 2 & 1 & 1 \\ 1 & 1 & 2 & 0 \\ 1 & 1 & 0 & 2 \end{bmatrix} \xrightarrow{\text{RREF}} \begin{bmatrix} \boxed{1} & 0 & 0 & 1 \\ 0 & \boxed{1} & 0 & 1 \\ 0 & 0 & \boxed{1} & -1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\mathcal{E}_C(-1) = \mathcal{N}(C + 1I_4) = \operatorname{Span}\left\{ \begin{bmatrix} -1 \\ -1 \\ 1 \\ 1 \end{bmatrix} \right\}$$

Example 16.14. Consider the matrix

$$E = \begin{bmatrix} 29 & 14 & 2 & 6 & -9 \\ -47 & -22 & -1 & -11 & 13 \\ 19 & 10 & 5 & 4 & -8 \\ -19 & -10 & -3 & -2 & 8 \\ 7 & 4 & 3 & 1 & -3 \end{bmatrix}.$$

Then

$$p_E\left(x\right) = -16 + 16x + 8x^2 - 16x^3 + 7x^4 - x^5 = -(x-2)^4(x+1).$$

So the eigenvalues are $\lambda = 2,\ -1$. Computing eigenvectors, we find

$$\lambda = 2:$$

$$E - 2I_5 = \begin{bmatrix} 27 & 14 & 2 & 6 & -9 \\ -47 & -24 & -1 & -11 & 13 \\ 19 & 10 & 3 & 4 & -8 \\ -19 & -10 & -3 & -4 & 8 \\ 7 & 4 & 3 & 1 & -5 \end{bmatrix} \xrightarrow{\text{RREF}} \begin{bmatrix} \boxed{1} & 0 & 0 & 1 & 0 \\ 0 & \boxed{1} & 0 & -\frac{3}{2} & -\frac{1}{2} \\ 0 & 0 & \boxed{1} & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\mathcal{E}_E\left(2\right) = \mathcal{N}\left(E - 2I_5\right) = \text{Span}\left\{ \begin{bmatrix} -1 \\ \frac{3}{2} \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ \frac{1}{2} \\ 1 \\ 0 \\ 1 \end{bmatrix} \right\} = \text{Span}\left\{ \begin{bmatrix} -2 \\ 3 \\ 0 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 2 \\ 0 \\ 2 \end{bmatrix} \right\}$$

$$\lambda = -1:$$

$$E + 1I_5 = \begin{bmatrix} 30 & 14 & 2 & 6 & -9 \\ -47 & -21 & -1 & -11 & 13 \\ 19 & 10 & 6 & 4 & -8 \\ -19 & -10 & -3 & -1 & 8 \\ 7 & 4 & 3 & 1 & -2 \end{bmatrix} \xrightarrow{\text{RREF}} \begin{bmatrix} \boxed{1} & 0 & 0 & 2 & 0 \\ 0 & \boxed{1} & 0 & -4 & 0 \\ 0 & 0 & \boxed{1} & 1 & 0 \\ 0 & 0 & 0 & 0 & \boxed{1} \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\mathcal{E}_E\left(-1\right) = \mathcal{N}\left(E + 1I_5\right) = \text{Span}\left\{ \begin{bmatrix} -2 \\ 4 \\ -1 \\ 1 \\ 0 \end{bmatrix} \right\}$$

Example 16.15. Consider the matrix

$$H = \begin{bmatrix} 15 & 18 & -8 & 6 & -5 \\ 5 & 3 & 1 & -1 & -3 \\ 0 & -4 & 5 & -4 & -2 \\ -43 & -46 & 17 & -14 & 15 \\ 26 & 30 & -12 & 8 & -10 \end{bmatrix}.$$

Then

$$p_H\left(x\right) = -6x + x^2 + 7x^3 - x^4 - x^5 = x(x-2)(x-1)(x+1)(x+3).$$

So the eigenvalues are $\lambda = 2,\ 1,\ 0,\ -1,\ -3$.

Computing eigenvectors, we find

$\lambda = 2:$

$$H - 2I_5 = \begin{bmatrix} 13 & 18 & -8 & 6 & -5 \\ 5 & 1 & 1 & -1 & -3 \\ 0 & -4 & 3 & -4 & -2 \\ -43 & -46 & 17 & -16 & 15 \\ 26 & 30 & -12 & 8 & -12 \end{bmatrix} \xrightarrow{\text{RREF}} \begin{bmatrix} \boxed{1} & 0 & 0 & 0 & -1 \\ 0 & \boxed{1} & 0 & 0 & 1 \\ 0 & 0 & \boxed{1} & 0 & 2 \\ 0 & 0 & 0 & \boxed{1} & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\mathcal{E}_H\left(2\right) = \mathcal{N}\left(H - 2I_5\right) = \operatorname{Span}\left\{\begin{bmatrix} 1 \\ -1 \\ -2 \\ -1 \\ 1 \end{bmatrix}\right\}$$

$\lambda = 1:$

$$H - 1I_5 = \begin{bmatrix} 14 & 18 & -8 & 6 & -5 \\ 5 & 2 & 1 & -1 & -3 \\ 0 & -4 & 4 & -4 & -2 \\ -43 & -46 & 17 & -15 & 15 \\ 26 & 30 & -12 & 8 & -11 \end{bmatrix} \xrightarrow{\text{RREF}} \begin{bmatrix} \boxed{1} & 0 & 0 & 0 & -\frac{1}{2} \\ 0 & \boxed{1} & 0 & 0 & 0 \\ 0 & 0 & \boxed{1} & 0 & \frac{1}{2} \\ 0 & 0 & 0 & \boxed{1} & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\mathcal{E}_H\left(1\right) = \mathcal{N}\left(H - 1I_5\right) = \operatorname{Span}\left\{\begin{bmatrix} \frac{1}{2} \\ 0 \\ -\frac{1}{2} \\ -1 \\ 1 \end{bmatrix}\right\} = \operatorname{Span}\left\{\begin{bmatrix} 1 \\ 0 \\ -1 \\ -2 \\ 2 \end{bmatrix}\right\}$$

$\lambda = 0:$

$$H - 0I_5 = \begin{bmatrix} 15 & 18 & -8 & 6 & -5 \\ 5 & 3 & 1 & -1 & -3 \\ 0 & -4 & 5 & -4 & -2 \\ -43 & -46 & 17 & -14 & 15 \\ 26 & 30 & -12 & 8 & -10 \end{bmatrix} \xrightarrow{\text{RREF}} \begin{bmatrix} \boxed{1} & 0 & 0 & 0 & 1 \\ 0 & \boxed{1} & 0 & 0 & -2 \\ 0 & 0 & \boxed{1} & 0 & -2 \\ 0 & 0 & 0 & \boxed{1} & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\mathcal{E}_H(0) = \mathcal{N}(H - 0I_5) = \text{Span}\left\{ \begin{bmatrix} -1 \\ 2 \\ 2 \\ 0 \\ 1 \end{bmatrix} \right\}$$

$\lambda = -1:$

$$H + 1I_5 = \begin{bmatrix} 16 & 18 & -8 & 6 & -5 \\ 5 & 4 & 1 & -1 & -3 \\ 0 & -4 & 6 & -4 & -2 \\ -43 & -46 & 17 & -13 & 15 \\ 26 & 30 & -12 & 8 & -9 \end{bmatrix} \xrightarrow{\text{RREF}} \begin{bmatrix} \boxed{1} & 0 & 0 & 0 & -1/2 \\ 0 & \boxed{1} & 0 & 0 & 0 \\ 0 & 0 & \boxed{1} & 0 & 0 \\ 0 & 0 & 0 & \boxed{1} & 1/2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\mathcal{E}_H(-1) = \mathcal{N}(H + 1I_5) = \text{Span}\left\{ \begin{bmatrix} \frac{1}{2} \\ 0 \\ 0 \\ -\frac{1}{2} \\ 1 \end{bmatrix} \right\} = \text{Span}\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \\ -1 \\ 2 \end{bmatrix} \right\}$$

$\lambda = -3:$

$$H + 3I_5 = \begin{bmatrix} 18 & 18 & -8 & 6 & -5 \\ 5 & 6 & 1 & -1 & -3 \\ 0 & -4 & 8 & -4 & -2 \\ -43 & -46 & 17 & -11 & 15 \\ 26 & 30 & -12 & 8 & -7 \end{bmatrix} \xrightarrow{\text{RREF}} \begin{bmatrix} \boxed{1} & 0 & 0 & 0 & -1 \\ 0 & \boxed{1} & 0 & 0 & \frac{1}{2} \\ 0 & 0 & \boxed{1} & 0 & 1 \\ 0 & 0 & 0 & \boxed{1} & 2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\mathcal{E}_H(-3) = \mathcal{N}(H + 3I_5) = \text{Span}\left\{ \begin{bmatrix} 1 \\ -\frac{1}{2} \\ -1 \\ -2 \\ 1 \end{bmatrix} \right\} = \text{Span}\left\{ \begin{bmatrix} -2 \\ 1 \\ 2 \\ 4 \\ -2 \end{bmatrix} \right\}$$

16.4 Similar Matrices

Definition 16.16 (Similar Matrices)**.** Suppose that A and B are square matrices of size n. Then A and B are **similar** if there exists a nonsingular matrix of size n, S, such that $A = S^{-1}BS$. We will also say A is similar to B via S. Finally, we will refer to $S^{-1}BS$ as a **similarity transformation** when we want to emphasize the way that S changes B.

Example 16.17. Define

$$B = \begin{bmatrix} -5 & -7 \\ 4 & 6 \end{bmatrix}$$
$$S = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}.$$

Check that S is nonsingular and then compute

$$\begin{aligned} A &= S^{-1}BS \\ &= \begin{bmatrix} -3 & 2 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} -5 & -7 \\ 4 & 6 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix} \\ &= \begin{bmatrix} 89 & 145 \\ -54 & -88 \end{bmatrix}. \end{aligned}$$

It follows that A and B are similar.

Example 16.18. Define

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

Check that S is nonsingular and then compute

$$
\begin{aligned}
A &= S^{-1}BS \\
&= \begin{bmatrix} 10 & 1 & 0 & 2 & -5 \\ -1 & 0 & 1 & 0 & 0 \\ 3 & 0 & 2 & 1 & -3 \\ 0 & 0 & -1 & 0 & 1 \\ -4 & -1 & 1 & -1 & 1 \end{bmatrix} \begin{bmatrix} -4 & 1 & -3 & -2 & 2 \\ 1 & 2 & -1 & 3 & -2 \\ -4 & 1 & 3 & 2 & 2 \\ -3 & 4 & -2 & -1 & -3 \\ 3 & 1 & -1 & 1 & -4 \end{bmatrix} \begin{bmatrix} 1 & 2 & -1 & 1 & 1 \\ 0 & 1 & -1 & -2 & -1 \\ 1 & 3 & -1 & 1 & 1 \\ -2 & -3 & 3 & 1 & -2 \\ 1 & 3 & -1 & 2 & 1 \end{bmatrix} \\
&= \begin{bmatrix} -10 & -27 & -29 & -80 & -25 \\ -2 & 6 & 6 & 10 & -2 \\ -3 & 11 & -9 & -14 & -9 \\ -1 & -13 & 0 & -10 & -1 \\ 11 & 35 & 6 & 49 & 19 \end{bmatrix}.
\end{aligned}
$$

This shows that A and B are similar.

Example 16.19. Define

$$
B = \begin{bmatrix} -13 & -8 & -4 \\ 12 & 7 & 4 \\ 24 & 16 & 7 \end{bmatrix}
\qquad
S = \begin{bmatrix} 1 & 1 & 2 \\ -2 & -1 & -3 \\ 1 & -2 & 0 \end{bmatrix}.
$$

Check that S is nonsingular and then compute

$$
\begin{aligned}
A &= S^{-1}BS \\
&= \begin{bmatrix} -6 & -4 & -1 \\ -3 & -2 & -1 \\ 5 & 3 & 1 \end{bmatrix} \begin{bmatrix} -13 & -8 & -4 \\ 12 & 7 & 4 \\ 24 & 16 & 7 \end{bmatrix} \begin{bmatrix} 1 & 1 & 2 \\ -2 & -1 & -3 \\ 1 & -2 & 0 \end{bmatrix} \\
&= \begin{bmatrix} -1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & -1 \end{bmatrix}.
\end{aligned}
$$

Theorem 16.20 (Similarity is an Equivalence Relation)**.** *Suppose that A, B and C are square matrices of size n. Then*

1. *(Reflexive) A is similar to A.*

2. *(Symmetric) If A is similar to B, then B is similar to A.*

3. *(Transitive) If A is similar to B and B is similar to C, then A is similar to C.*

Proof of Similarity is an Equivalence Relation. To see that A is similar to A, we need only demonstrate a nonsingular matrix that effects a similarity transformation of A to A. We can take I_n, which is nonsingular and satisfies $I_n^{-1}AI_n = I_nAI_n = A$.

If we assume that A is similar to B, then we know there exists is a nonsingular matrix S so that $A = S^{-1}BS$. But then S^{-1} is invertible and therefore nonsingular. So

$$(S^{-1})^{-1}A(S^{-1}) = SAS^{-1} = SS^{-1}BSS^{-1}$$

$$= \left(SS^{-1}\right) B \left(SS^{-1}\right) = I_nBI_n = B$$

and we see that B is similar to A.

Assume that A is similar to B and that B is similar to C. This gives us the existence of nonsingular matrices, S and R, such that $A = S^{-1}BS$ and $B = R^{-1}CR$. Since S and R are invertible, so too is RS, which has inverse $S^{-1}R^{-1}$. Then we compute

$$(RS)^{-1}C(RS) = S^{-1}R^{-1}CRS = S^{-1}\left(R^{-1}CR\right)S$$

$$= S^{-1}BS = A$$

so A is similar to C via the nonsingular matrix RS. □

Theorem 16.21 (Similar Matrices have Equal Eigenvalues)**.** *Suppose that A and B are similar matrices. Then the characteristic polynomials of A and B are equal, that is, $p_A(x) = p_B(x)$.*

Proof of Similar Matrices have Equal Eigenvalues. Let n denote the size of A and B. Since A and B are similar, there exists a nonsingular matrix S, such that

$A = S^{-1}BS$. Then

$$\begin{aligned}
p_A(x) &= \det(A - xI_n) \\
&= \det\left(S^{-1}BS - xI_n\right) \\
&= \det\left(S^{-1}BS - xS^{-1}I_nS\right) \\
&= \det\left(S^{-1}BS - S^{-1}xI_nS\right) \\
&= \det\left(S^{-1}(B - xI_n)S\right) \\
&= \det\left(S^{-1}\right)\det(B - xI_n)\det(S) \\
&= \det\left(S^{-1}\right)\det(S)\det(B - xI_n) \\
&= \det\left(S^{-1}S\right)\det(B - xI_n) \\
&= \det(I_n)\det(B - xI_n) \\
&= 1\det(B - xI_n) \\
&= p_B(x).
\end{aligned}$$

□

Example 16.22. We claim that the matrices

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

are not similar.

To show this, we compute

$$p_A(x) = \begin{vmatrix} 1-x & 2 \\ 3 & 4-x \end{vmatrix} = (1-x)(4-x) - 6 = x^2 - 5x - 2$$

and

$$p_B(x) = \begin{vmatrix} 1-x & 2 \\ 0 & 4-x \end{vmatrix} = (1-x)(4-x) = x^2 - 5x + 4.$$

Because $p_A(x) \neq p_B(x)$, we conclude that A and B are not similar.

Example 16.23. Same characteristic polynomial, but not similar

Define

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \qquad B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

We have

$$p_A(x) = p_B(x) = 1 - 2x + x^2 = (x-1)^2,$$

so that A and B have equal characteristic polynomials. If the converse of the above theorem were true, then A and B would be similar. Suppose this is the case. More precisely, suppose there exists is a nonsingular matrix S so that $A = S^{-1}BS$. Then

$$A = S^{-1}BS = S^{-1}I_2S = S^{-1}S = I_2$$

Clearly $A \neq I_2$. This contradiction tells us that the converse of the above theorem is false.

16.5 Diagonalizability

Good things happen when a matrix is similar to a diagonal matrix. For example, the eigenvalues of the matrix are the entries on the diagonal of the diagonal matrix. It is also much simpler matter to compute high powers of the matrix. Diagonalizable matrices are also of interest in more abstract settings. Here are the relevant definitions, then our main theorem for this section.

Definition 16.24 (Diagonal Matrix)**.** Suppose that A is a square matrix of size n. Then A is a **diagonal matrix** if $[A]_{ij} = 0$ whenever $i \neq j$, i.e.

$$A = \begin{bmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ 0 & 0 & \lambda_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & \cdots & \lambda_n \end{bmatrix}.$$

We will often denote such a matrix A by $\operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$.

Definition 16.25 (Diagonalizable Matrix)**.** Suppose that A is a square matrix. Then A is **diagonalizable** if A is similar to a diagonal matrix, i.e, there exists an invertible matrix S and real numbers $\lambda_1, \lambda_2, \ldots, \lambda_n$ such that

$$S^{-1}AS = \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n).$$

Example 16.26. Let

$$B = \begin{bmatrix} -7 & -6 & -12 \\ 5 & 5 & 7 \\ 1 & 0 & 4 \end{bmatrix}.$$

This matrix is similar to a diagonal matrix, as can be seen by the following computation with the nonsingular matrix S:

$$S^{-1}BS = \begin{bmatrix} -5 & -3 & -2 \\ 3 & 2 & 1 \\ 1 & 1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} -7 & -6 & -12 \\ 5 & 5 & 7 \\ 1 & 0 & 4 \end{bmatrix} \begin{bmatrix} -5 & -3 & -2 \\ 3 & 2 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} -1 & -1 & -1 \\ 2 & 3 & 1 \\ -1 & -2 & 1 \end{bmatrix} \begin{bmatrix} -7 & -6 & -12 \\ 5 & 5 & 7 \\ 1 & 0 & 4 \end{bmatrix} \begin{bmatrix} -5 & -3 & -2 \\ 3 & 2 & 1 \\ 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}.$$

Theorem 16.27 (Diagonalization Characterization)**.** *Suppose that A is a square matrix of size n. Then A is diagonalizable if and only if there exists a linearly independent set T that contains n eigenvectors of A.*

Proof of Diagonalization Characterization. ($\Rightarrow$) Suppose that A is diagonalizable. Then there exists an invertible matrix S and real numbers $\lambda_1, \ldots, \lambda_n$ such that

$$S^{-1}AS = \operatorname{diag}(\lambda_1, \ldots, \lambda_n).$$

Let $\mathbf{S}_i$ be column i of S. Let T be the columns of S. Because S is invertible (nonsingular), the columns of S are linearly independent. Also

$$S^{-1}AS = \operatorname{diag}(\lambda_1, \ldots, \lambda_n).$$

So

$$AS = S \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$$

or equivalently

$$[A\mathbf{S}_1 | A\mathbf{S}_2 | \cdots | A\mathbf{S}_n] = [\lambda_1\mathbf{S}_1 | \lambda_2\mathbf{S}_2 | \cdots | \lambda_n\mathbf{S}_n].$$

Hence, for $1 \leq i \leq n$, we have:

$$A\mathbf{S}_i = \lambda_i\mathbf{S}_i.$$

Obviously $\mathbf{S}_i \neq \mathbf{0}$, because S is nonsingular. So $\mathbf{S}_i$ is an eigenvector with eigenvalue λ_i. Hence T is a linearly independent set consisting of eigenvectors of A.

($\Leftarrow$) Suppose that $T = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is a linearly independent set consisting of eigenvectors of A with eigenvalues $\lambda_1, \ldots, \lambda_n$, i.e. $A\mathbf{v}_i = \lambda_i\mathbf{v}_i$ for $i = 1, \ldots, n$. Let $D = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$. Let

$$S = [\mathbf{v}_1 | \mathbf{v}_2 | \cdots | \mathbf{v}_n].$$

Because T is linearly independent, S is invertible. Similarly to the above computation, we compute:

$$AS = [A\mathbf{v}_1|A\mathbf{v}_2|\cdots|A\mathbf{v}_n] = [\lambda_1\mathbf{v}_1|\lambda_2\mathbf{v}_2|\cdots|\lambda_n\mathbf{v}_n] = SD.$$

So

$$S^{-1}AS = S^{-1}SD = D.$$

Therefore A is diagonalizable. □

Remark. Notice that the proof is constructive. To diagonalize a matrix, we need only locate n linearly independent eigenvectors. Then we can construct a nonsingular matrix S, using the eigenvectors as columns, with the property that $S^{-1}AS$ is a diagonal matrix (D). The entries on the diagonal of D will be the eigenvalues of the eigenvectors used to create S, in the same order as the eigenvectors appear in S. We illustrate this by **diagonalizing** some matrices.

Example 16.28. Consider the matrix

$$F = \begin{bmatrix} -13 & -8 & -4 \\ 12 & 7 & 4 \\ 24 & 16 & 7 \end{bmatrix}$$

from previous examples. The eigenvalues and eigenspaces of F's are

$$\lambda = 3 \qquad \mathcal{E}_F(3) = \text{Span}\left\{\begin{bmatrix} -\frac{1}{2} \\ \frac{1}{2} \\ 1 \end{bmatrix}\right\}$$

$$\lambda = -1 \qquad \mathcal{E}_F(-1) = \text{Span}\left\{\begin{bmatrix} -\frac{2}{3} \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -\frac{1}{3} \\ 0 \\ 1 \end{bmatrix}\right\}$$

Define the matrix S to be the 3×3 matrix whose columns are the three basis vectors in the eigenspaces for F:

$$S = \begin{bmatrix} -\frac{1}{2} & -\frac{2}{3} & -\frac{1}{3} \\ \frac{1}{2} & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$$

Check that S is nonsingular (row-reduces to the identity matrix, or has a nonzero determinant).

Remark : After we introduce Theorem Theorem 16.29, you don't need to check that S is nonsingular. See examples below).

The three columns of S are a linearly independent set. By Theorem Theorem 16.27 (Diagonalization Characterization) we now know that F is diagonalizable. Furthermore, the construction in the proof of Theorem Theorem 16.27 (Diagonalization Characterization) tells us that $S^{-1}FS = \operatorname{diag}(3, -1, -1)$. Let us check this directly:

$$\begin{aligned}
S^{-1}FS &= \begin{bmatrix} -\frac{1}{2} & -\frac{2}{3} & -\frac{1}{3} \\ \frac{1}{2} & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} -13 & -8 & -4 \\ 12 & 7 & 4 \\ 24 & 16 & 7 \end{bmatrix} \begin{bmatrix} -\frac{1}{2} & -\frac{2}{3} & -\frac{1}{3} \\ \frac{1}{2} & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \\
&= \begin{bmatrix} 6 & 4 & 2 \\ -3 & -1 & -1 \\ -6 & -4 & -1 \end{bmatrix} \begin{bmatrix} -13 & -8 & -4 \\ 12 & 7 & 4 \\ 24 & 16 & 7 \end{bmatrix} \begin{bmatrix} -\frac{1}{2} & -\frac{2}{3} & -\frac{1}{3} \\ \frac{1}{2} & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \\
&= \begin{bmatrix} 3 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}.
\end{aligned}$$

Theorem 16.29. *Suppose that A is a square matrix of size n. Suppose that $\lambda_1, \ldots, \lambda_k$ are all of the distinct eigenvalues of A. Then A is diagonalizable if and only if*

$$\sum_{i=1}^{k} \dim \mathcal{E}_A(\lambda_i) = \dim \mathcal{E}_A(\lambda_1) + \cdots + \dim \mathcal{E}_A(\lambda_k) = n. \tag{16.1}$$

Suppose that the above condition is satisfied by A and let $T_i = \{\mathbf{v}_{i1}, \mathbf{v}_{i2}, \mathbf{v}_{i3}, \ldots, \mathbf{v}_{id_i}\}$ be a basis for the eigenspace of λ_i, $\mathcal{E}_A(\lambda_i)$, for each $1 \le i \le k$ and let $d_i = \dim \mathcal{E}_A(\lambda_i)$. Then

$$T = T_1 \cup T_2 \cup T_3 \cup \cdots \cup T_k$$

is a set of linearly independent eigenvectors for A with size n. By Theorem Theorem 16.27 (Diagonalization Characterization), let S be a square matrix whose i-th column is the i-th vector of the set T, i.e.

$$S = [\mathbf{v}_{11}|\cdots|\mathbf{v}_{1d_1}|\mathbf{v}_{21}|\cdots|\mathbf{v}_{2d_2}|\cdots|\mathbf{v}_{k1}|\cdots|\mathbf{v}_{kd_k}]$$

Then

$$S^{-1}AS = \operatorname{diag}(\underbrace{\lambda_1, \ldots, \lambda_1}_{d_1}, \underbrace{\lambda_2, \ldots, \lambda_2}_{d_2}, \ldots, \underbrace{\lambda_k, \ldots, \lambda_k}_{d_k}).$$

Proof of Theorem 16.29. See Beezer's textbook. □

Remark. Equation (16.1) may be rewritten as:

$$\sum_{i=1}^{k} n\left(A-\lambda_i I_n\right)=n\left(A-\lambda_1 I_n\right)+\cdots+n\left(A-\lambda_k I_n\right)=n,$$

where on the left-hand side $n(M)$ denotes the nullity of a matrix M. or

$$\sum_{i=1}^{k}\left(n-r\left(A-\lambda_i I_n\right)\right)=\left(n-r\left(A-\lambda_1 I_n\right)\right)+\cdots+\left(n-r\left(A-\lambda_k I_n\right)\right)=n,$$

where $r(M)$ denotes the rank of a matrix M.

Corollary 16.30 (Distinct Eigenvalues implies Diagonalizable)**.** *Suppose that A is a square matrix of size n with n distinct eigenvalues. Then A is diagonalizable.*

Proof of Distinct Eigenvalues implies Diagonalizable. You can skip the proof. See the textbook. □

Example 16.31. Determine if the matrix B in Example 16.12 is diagonalizable. The characteristic polynomial is

$$p_B\left(x\right)=\det\left(B-xI_4\right)=(x-1)(x-2)^3.$$

We conclude that $\lambda_1=1$ and $\lambda_2=2$ are all of the distinct eigenvalues of B.

In Example 16.12, we compute the RREF of $B-I_4$ and $B-2I_4$. By the RREFs, we have $r\left(B-I_4\right)=3$, $r\left(B-2I_4\right)=3$. Therefore

$$\dim\mathcal{E}_B\left(1\right)=n\left(B-I_4\right)=4-r\left(B-I_4\right)=1$$

and

$$\dim\mathcal{E}_B\left(2\right)=n\left(B-2I_4\right)=4-r\left(B-2I_4\right)=4-3=1.$$

Now

$$\dim\mathcal{E}_B\left(1\right)+\dim\mathcal{E}_B\left(2\right)=1+1=2\neq 4.$$

By Theorem Theorem 16.29, B is not diagonalizable.

Example 16.32. Determine if the matrix C in Example 16.13 is diagonalizable.Because

$$p_C\left(x\right)=-3+4x+2x^2-4x^3+x^4=(x-3)(x-1)^2(x+1),$$

all the distinct eigenvalues are $\lambda_1=3$, $\lambda_2=1$ and $\lambda_3=-1$. We have

$$\sum_{i=1}^{3}\dim\mathcal{E}_C\left(\lambda_i\right)=\sum_{i=1}^{3}(4-r\left(C-\lambda_i I_4\right))$$

$$=(4-3)+(4-2)+(4-3)=4.$$

By Theorem Theorem 16.29, C is diagonalizable.

Example 16.33. Determine if the matrix E in Example 16.14 is diagonalizable.The characteristic polynomial is $p_E(x) = -(x-2)^4(x+1)$. The eigenvalues are $\lambda_1 = 2$ and $\lambda_2 = -1$ and we have

$$\dim \mathcal{E}_E(\lambda_1) + \dim \mathcal{E}_E(\lambda_2) = (5 - r(E - 2I_4)) + (5 - r(E + I_5))$$

$$= (5-3) + (5-4) = 2 + 1 = 3 \neq 5.$$

So E is not diagonalizable.

Example 16.34. Determine if the matrix H in Example 16.15 is diagonalizable. Because $p_H(x) = x(x-2)(x-1)(x+1)(x+3)$, has 5 distinct eigenvalues, Theorem Corollary 16.30 (Distinct Eigenvalues implies Diagonalizable) implies that H is diagonalizable.

Example 16.35. Diagonalize C in Example 16.13 (see also Example 16.32). By the computation in Example 16.32, C is diagonalizable. By the computation in Example 16.32, we have that $\left\{\begin{bmatrix}1\\1\\1\\1\end{bmatrix}\right\}$ is a basis for $\mathcal{E}_C(3)$, $\left\{\begin{bmatrix}-1\\1\\0\\0\end{bmatrix}, \begin{bmatrix}0\\0\\-1\\1\end{bmatrix}\right\}$ is a basis for $\mathcal{E}_C(1)$, $\left\{\begin{bmatrix}-1\\-1\\1\\1\end{bmatrix}\right\}$ is a basis for $\mathcal{E}_C(-1)$. By Theorem Theorem 16.29,

$$S = \begin{bmatrix}1 & -1 & 0 & -1\\ 1 & 1 & 0 & -1\\ 1 & 0 & -1 & 1\\ 1 & 0 & 1 & 1\end{bmatrix}.$$

Then $S^{-1}CS = \mathrm{diag}(3, 1, 1, -1)$.

Remark : Note that the invertibility of S is guaranteed by Theorem Theorem 16.29.

Example 16.36. Diagonalize H in Example 16.15 (see also Example 16.33). By the discussion of Example 16.33, H is diagonalizable. By the computation in Example 16.15, $\begin{bmatrix}1\\-1\\-2\\-1\\1\end{bmatrix}, \begin{bmatrix}1\\0\\-1\\-2\\2\end{bmatrix}, \begin{bmatrix}-1\\2\\2\\0\\1\end{bmatrix}, \begin{bmatrix}1\\0\\0\\-1\\2\end{bmatrix}$ and $\begin{bmatrix}-2\\1\\2\\4\\-2\end{bmatrix}$ are bases for

$\mathcal{E}_H(2), \mathcal{E}_H(1), \mathcal{E}_H(0), \mathcal{E}_H(-1)$ and $\mathcal{E}_H(-3)$ respectively. By Theorem Theorem 16.29, let

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

Then

$$S^{-1}HS = \operatorname{diag}(2, 1, 0, -1, -3).$$

Example 16.37. Determine if

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

is diagonalizable. If it is diagonalizable, find S such that $S^{-1}JS$ is diagonal.

Step 1: $p_J(x) = -x^3 + 6x^2 - 9x + 4 = -(x-4)(x-1)^2$. All the distinct eigenvalues of J is $\lambda_1 = 4$, $\lambda_2 = 1$.

Step 2 :

$$J - 4I_3 \xrightarrow{\text{RREF}} \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}.$$

$$\dim \mathcal{E}_J(4) = 3 - r(J - 4I_3) = 3 - 2 = 1.$$

$$J - I_3 \xrightarrow{\text{RREF}} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\dim \mathcal{E}_J(1) = 3 - r(J - I_3) = 3 - 1 = 2.$$

Now

$$\dim \mathcal{E}_J(4) + \dim \mathcal{E}_J(1) = 1 + 2 = 3.$$

By Theorem Theorem 16.29, J is diagonalizable.

Step 3 : $\left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\}$ is a basis for $\mathcal{E}_J\left(4\right)$. $\left\{ \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \right\}$ is a basis for $\mathcal{E}_J\left(1\right)$.
By Theorem Theorem 16.29, we can take

$$S = \begin{bmatrix} 1 & -1 & -1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}.$$

Then

$$S^{-1}JS = \operatorname{diag}(4, 1, 1).$$

Example 16.38. Determine if

$$K = \begin{bmatrix} -4 & -4 & 0 & -11 & -2 & -5 \\ 138 & 87 & -6 & 248 & 44 & 122 \\ -24 & -16 & 2 & -44 & -8 & -20 \\ -62 & -39 & 2 & -110 & -20 & -54 \\ -63 & -39 & 3 & -114 & -19 & -57 \\ 56 & 35 & -2 & 101 & 18 & 51 \end{bmatrix}$$

is diagonalizable and if it is diagonalizable, find S such that $S^{-1}KS$ is diagonal.

Step 1: The characteristic polynomial is

$$p_K\left(x\right) = \det\left(K - xI_6\right) = (x+1)(x-1)^2(x-2)^3.$$

The eigenvalues are $\lambda_1 = -1, \lambda_2 = 1$ and $\lambda_3 = 2$.

Step 2 :

$$K + I_4 \xrightarrow{\text{RREF}} \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & \frac{1}{9} \\ 0 & 1 & 0 & 0 & 0 & -\frac{22}{9} \\ 0 & 0 & 1 & 0 & 0 & \frac{4}{9} \\ 0 & 0 & 0 & 1 & 0 & \frac{10}{9} \\ 0 & 0 & 0 & 0 & 1 & \frac{10}{9} \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

$$\dim \mathcal{E}_K\left(-1\right) = n\left(K + I_4\right) = 6 - r\left(K + I_4\right) = 6 - 5 = 1.$$

$$A - I_6 \xrightarrow{\text{RREF}} \begin{bmatrix} 1 & 0 & 0 & 0 & \frac{1}{18} & \frac{1}{6} \\ 0 & 1 & 0 & 0 & \frac{5}{18} & -\frac{13}{6} \\ 0 & 0 & 1 & 0 & \frac{2}{9} & \frac{2}{3} \\ 0 & 0 & 0 & 1 & \frac{1}{18} & \frac{7}{6} \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

$$\dim \mathcal{E}_K(1) = n(K - I_6) = 6 - r(K - I_6) = 6 - 4 = 2.$$

$$K - 2I_6 \xrightarrow{\text{RREF}} \begin{bmatrix} 1 & 0 & 0 & \frac{1}{2} & -1 & -\frac{1}{2} \\ 0 & 1 & 0 & 2 & 2 & 2 \\ 0 & 0 & 1 & -\frac{3}{2} & -2 & -\frac{7}{2} \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\dim \mathcal{E}_K(2) = n(K - 2I_6) = 6 - r(K - 2I_6) = 6 - 3 = 3.$$

Because

$$\dim \mathcal{E}_K(-1) + \dim \mathcal{E}_K(1) + \dim \mathcal{E}_K(2) = 1 + 2 + 3 = 6.$$

By Theorem Theorem 16.29, K is diagonalizable.

Step 3 : A basis for $\mathcal{E}_K(-1) = \mathcal{N}(K + I_6)$ is

$$\left\{ \begin{bmatrix} -1 \\ 22 \\ -4 \\ -10 \\ -10 \\ 9 \end{bmatrix} \right\}$$

(we use the method in Lecture 8 Example 16.12 and multiply the result by 9 to clear the denominator.) A basis for $\mathcal{E}_K(1) = \mathcal{N}(K - I_6)$ is

$$\left\{ \begin{bmatrix} -1 \\ -5 \\ -4 \\ -1 \\ 18 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 13 \\ -4 \\ -7 \\ 0 \\ 6 \end{bmatrix} \right\}$$

(Again, we use the method in Lecture 8 Example 16.12 and multiply the first vector by 18 and the second vector by 6 to clear the denominators.)

$$\left\{ \begin{bmatrix} -1 \\ -4 \\ 3 \\ 2 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ -2 \\ 2 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ -4 \\ 7 \\ 0 \\ 0 \\ 2 \end{bmatrix} \right\}.$$

(Again, we use the method in Lecture 8 Example 16.12 and multiply the first and the third vector by 2 to clear the denominators.) So we can take

$$S = \begin{bmatrix} -1 & -1 & -1 & 1 & 1 & -1 \\ 22 & 13 & -5 & -4 & -2 & -4 \\ -4 & -4 & -4 & 7 & 2 & 3 \\ -10 & -7 & -1 & 0 & 0 & 2 \\ -10 & 0 & 18 & 0 & 1 & 0 \\ 9 & 6 & 0 & 2 & 0 & 0 \end{bmatrix}.$$

Then

$$S^{-1}AS = \mathrm{diag}(-1, 1, 1, 2, 2, 2).$$

Example 16.39. Determine if

$$J = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}, \ L = \begin{bmatrix} -8 & 6 & 6 \\ -9 & 7 & 6 \\ -9 & 6 & 7 \end{bmatrix}$$

are similar. If they are similar, Find R such that $R^{-1}JR = L$.

The characteristic polynomials

$$p_J(x) = -(-4+x)(-1+x)^2 = p_L(x).$$

(If the characteristic polynomials are different, J and L are not similar, end of the story.)

In Example 16.37, we know that J is diagonalizable. Following the same procedure as before, we can show that L is diagonalizable (**Exercise**).

We have:

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

$$Q^{-1}LQ = \mathrm{diag}(4, 1, 1).$$

So L is similar to $\mathrm{diag}(4, 1, 1)$ which in turns similar to J. So J is similar to L (Theorem Theorem 16.20 (Similarity is an Equivalence Relation)). In fact:

$$S^{-1}JS = Q^{-1}LQ$$

$$(SQ^{-1})^{-1}J(SQ^{-1}) = L.$$

So we can take

$$R = SQ^{-1} = \begin{bmatrix} -5 & 3 & 3 \\ -2 & \frac{5}{3} & \frac{4}{3} \\ -2 & \frac{4}{3} & \frac{5}{3} \end{bmatrix}.$$

16.6 Powers of Matrices

Suppose s is a positive integer. Recall the notation

$$A^s = \underbrace{A \cdots A}_{s}$$

Powers of a diagonal matrix are easy to compute. The case of a diagonalizable matrix is only slightly more difficult. Suppose that A is similar to a diagonal matrix $D = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$. Let S be an invertible matrix such that

$$S^{-1}AS = D.$$

Then

$$A = SDS^{-1}.$$

$$A^s = \underbrace{SDS^{-1}SDS^{-1} \cdots SDS^{-1}}_{s} = S\underbrace{D \cdots D}_{s}S^{-1} = SD^sS^{-1}$$

$$= S\operatorname{diag}(\lambda_1^s, \ldots, \lambda_n^s)S^{-1}.$$

Example 16.40. Let s be a positive integer and

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

We want to find a closed formula for A^s. The characteristic polynomial of A is

$$p_A(x) = \det(A - xI_2) = (1-x)(2-x) - 12 = x^2 - 3x - 10 = (x+2)(x-5).$$

For $\lambda = -2$

$$A + 2I_2 = \begin{bmatrix} 3 & 3 \\ 4 & 4 \end{bmatrix} \xrightarrow{\text{RREF}} \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}.$$

So

$$\left\{ \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right\}$$

is a basis for $\mathcal{E}_A(-2)$. For $\lambda = 5$

$$A - 5I_2 = \begin{bmatrix} -4 & 3 \\ 4 & -3 \end{bmatrix} \xrightarrow{\text{RREF}} \begin{bmatrix} 1 & -\frac{3}{4} \\ 0 & 0 \end{bmatrix}.$$

So

$$\left\{ \begin{bmatrix} 3 \\ 4 \end{bmatrix} \right\}$$

is a basis for $\mathcal{E}_A(5)$. Let

$$S = \begin{bmatrix} 1 & 3 \\ -1 & 4 \end{bmatrix}.$$

Then

$$S^{-1}AS = \operatorname{diag}(-2, 5)$$

so that

$$A^s = S \operatorname{diag}((-2)^s, 5^s) S^{-1}$$

$$= \begin{bmatrix} 1 & 3 \\ -1 & 4 \end{bmatrix} \begin{bmatrix} (-2)^s & 0 \\ 0 & 5^s \end{bmatrix} \begin{bmatrix} \frac{4}{7} & -\frac{3}{7} \\ \frac{1}{7} & \frac{1}{7} \end{bmatrix} = \begin{bmatrix} \frac{1}{7}(-1)^s 2^{s+2} + \frac{3\times 5^s}{7} & \frac{1}{7}(-3)(-2)^s + \frac{3\times 5^s}{7} \\ -\frac{1}{7}(-1)^s 2^{s+2} + \frac{4\times 5^s}{7} & \frac{3(-2)^s}{7} + \frac{4\times 5^s}{7} \end{bmatrix}.$$

Example 16.41. High power of a diagonalizable matrix

Suppose that

$$A = \begin{bmatrix} 19 & 0 & 6 & 13 \\ -33 & -1 & -9 & -21 \\ 21 & -4 & 12 & 21 \\ -36 & 2 & -14 & -28 \end{bmatrix}.$$

We wish to compute A^{20}. Normally this would require 19 matrix multiplications. But since A is diagonalizable, we can simplify the computations substantially.

First, we diagonalize A. With

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

we find

$$\begin{aligned} D &= S^{-1}AS \\ &= \begin{bmatrix} -6 & 1 & -3 & -6 \\ 0 & 2 & -2 & -3 \\ 3 & 0 & 1 & 2 \\ -1 & -1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 19 & 0 & 6 & 13 \\ -33 & -1 & -9 & -21 \\ 21 & -4 & 12 & 21 \\ -36 & 2 & -14 & -28 \end{bmatrix} \begin{bmatrix} 1 & -1 & 2 & -1 \\ -2 & 3 & -3 & 3 \\ 1 & 1 & 3 & 3 \\ -2 & 1 & -4 & 0 \end{bmatrix} \\ &= \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \end{aligned}$$

Using this, we compute

$$
\begin{aligned}
A^{20} = SD^{20}S^{-1} &= S \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}^{20} S^{-1} \\
&= S \begin{bmatrix} (-1)^{20} & 0 & 0 & 0 \\ 0 & (0)^{20} & 0 & 0 \\ 0 & 0 & (2)^{20} & 0 \\ 0 & 0 & 0 & (1)^{20} \end{bmatrix} S^{-1} \\
&= \begin{bmatrix} 1 & -1 & 2 & -1 \\ -2 & 3 & -3 & 3 \\ 1 & 1 & 3 & 3 \\ -2 & 1 & -4 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1048576 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} -6 & 1 & -3 & -6 \\ 0 & 2 & -2 & -3 \\ 3 & 0 & 1 & 2 \\ -1 & -1 & 1 & 1 \end{bmatrix} \\
&= \begin{bmatrix} 6291451 & 2 & 2097148 & 4194297 \\ -9437175 & -5 & -3145719 & -6291441 \\ 9437175 & -2 & 3145728 & 6291453 \\ -12582900 & -2 & -4194298 & -8388596 \end{bmatrix}.
\end{aligned}
$$

Generally

$$
A^s = SD^sS^{-1} = \begin{bmatrix} 1 - 6(-1)^s + 3\,2^{s+1} & 1 + (-1)^s & -1 - 3(-1)^s + 2^{s+1} & -1 - 6(-1)^s + 2^{s+2} \\ -3 + 12(-1)^s - 9\,2^s & -3 - 2(-1)^s & 3 + 6(-1)^s - 3\,2^s & 3 + 12(-1)^s - 3\,2^{s+} \\ -3 - 6(-1)^s + 9\,2^s & -3 + (-1)^s & 3 - 3(-1)^s + 3\,2^s & 3 - 6(-1)^s + 3\,2^{s+1} \\ 12(-1)^s - 3\,2^{s+2} & -2(-1)^s & 6(-1)^s - 2^{s+2} & 12(-1)^s - 2^{s+3} \end{bmatrix}
$$

16.7 Summary

In below A always denote a square matrix of size n

1. If $\mathbf{x} \neq 0$ and $A\mathbf{x} = \lambda x$, then $\mathbf{x}$ is called an **eigenvector** of A with eigenvalue λ.
2. $p_A(x) = \det(A - xI_n)$ is called the **characteristic function** of A.
   (a) It is a polynomial of degree n with leading coefficient $(-1)^n$.
   (b) λ is an eigenvalue if and only if $p_A(\lambda) = 0$, i.e., λ is a root of $p_A(x)$.
3. $\mathcal{E}_A(\lambda)$: **eigenspace** of A for an eigenvalue λ.

(a) It is the set of of all eigenvectors of A for λ, together with the zero vector, i.e.

$$\mathcal{E}_{\lambda}(A) = \{\mathbf{x} \in \mathbb{R}^n \mid A\mathbf{x} = \lambda\mathbf{x}\}.$$

(b) $\mathcal{E}_{\lambda}(A) = \mathcal{N}(A - \lambda I_n)$.

(c) $\mathcal{E}_{\lambda}(A)$ is a subspace of $\mathbb{R}^n$.

4. $\alpha_A(\lambda)$: **algebraic multiplicity**. The power of $(x - \lambda)$ in the factorization of $p_A(x)$.

5. $\gamma_A(\lambda)$: **geometric multiplicity**. It is $\dim \mathcal{E}_A(\lambda) = n(A - \lambda I_n)$.

6. Basic properties. Suppose λ is an eigenvalue of A and $\mathbf{x}$ is an eigenvector of A with eigenvalue λ.

   (a) A is invertible if and only if $\lambda = 0$ is not an eigenvalue.

   (b) For positive integer s, $\mathbf{x}$ is an eigenvector of A^s with eigenvalue λ^s.

   (c) If $\lambda \neq 0$, $\mathbf{x}$ is an eigenvector of A^{-1} with eigenvalue λ^{-1}.

   (d) λ is an eigenvector of A^t (but $\mathbf{x}$ may **not** be an eigenvector of A)

7. Computational questions

   (a) Find all the eigenvalues of A: find all the roots of $p_A(x)$.

   (b) Find $\mathcal{E}_A(\lambda)$: find $\mathcal{N}(A - \lambda I_n)$, this can be done by finding the RREF of $A - \lambda I_n$.

   (c) Find $\alpha_A(\lambda)$: Find the power of $x - \lambda$ in the factorization of $p_A(x)$.

   (d) Find a basis for $\mathcal{E}_A(\lambda)$: again, this is same as finding basis of $\mathcal{N}(A - \lambda I_n)$. This can be done by $A - \lambda I_n \xrightarrow{\text{RREF}} B$ and use the standard method in finding basis (see Lecture 8 Theorem 4, Example 16.12).

   (e) Find $\gamma_A(\lambda)$: same as finding $n(A - \lambda I_n) = n - r(A - \lambda I_n)$. Suppose $A - \lambda I_n \xrightarrow{\text{RREF}} B$. Then $\gamma_A(\lambda) = n - r(B) = n-$ number of pivot columns of B.