Eigenvalues & Eigenvectors Part II

Summary: $i := \sqrt{-1}$, or even better,

$$(i)^2 := -1$$

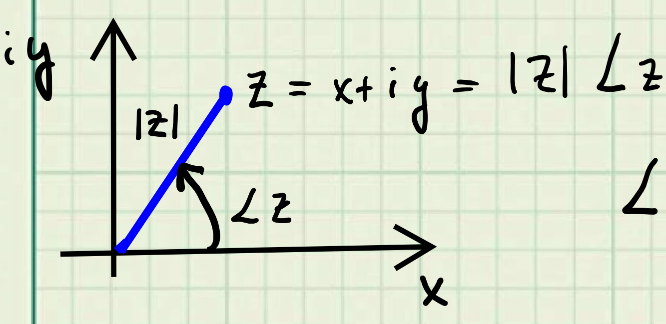
EE's (like Grizzle) write $j = \sqrt{-1}$

Imaginary Numbers: $i\beta,\ \beta \in \mathbb{R}$

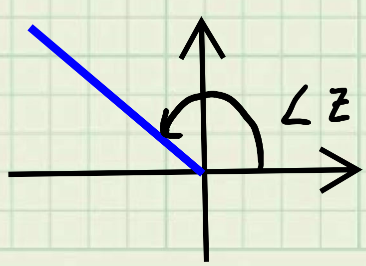
Complex Numbers: $\alpha + i\beta,\ \alpha, \beta \in \mathbb{R}$

$\mathbb{C} := \{x + iy \mid x \in \mathbb{R}, y \in \mathbb{R}\}$ $\quad \mathbb{R} \subset \mathbb{C}$

Pro Tip: multiplying Complex Numbers



$$\angle z = \begin{cases} \operatorname{atan}\left(\frac{y}{x}\right) & x > 0 \\ \pi - \operatorname{atan}\left(\frac{y}{|x|}\right) & x < 0 \end{cases}$$



$\angle z$ = angle of z

$z = |z| \cdot \angle z$ phasor notation

Fact: $z_1 * z_2 = |z_1| \cdot |z_2| \angle \angle z_1 + \angle z_2$

$$\frac{z_1}{z_2} = \frac{|z_1|}{|z_2|} \angle \angle z_1 - \angle z_2$$

Magnitudes multiply and angles add!

$$\therefore \boxed{z^k = |z|^k \angle k\angle z}$$

End Pro Tip

Why? $z = x + iy$

$$\text{Let } \theta = \angle z = \begin{cases} \operatorname{atan}\left(\frac{y}{x}\right) & x > 0 \\ \pi - \operatorname{atan}\left(\frac{y}{|x|}\right) & x < 0 \end{cases}$$

$$e^x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \cdots + \frac{x^k}{k!} + \cdots$$ Exponential function

Euler: $e^{i\theta} = \cos(\theta) + i\sin(\theta)$

$$z = |z| \cdot e^{i\theta}$$

$$\therefore z_1 * z_2 = |z_1| \cdot e^{i\theta_1} \cdot |z_2| \cdot e^{i\theta_2} = |z_1||z_2| e^{i(\theta_1 + \theta_2)}$$

Linear Difference Equations

We start with

$$z[k+1] = a\, z[k], \quad a \in \mathbb{C}, \quad z_0 \in \mathbb{C}$$

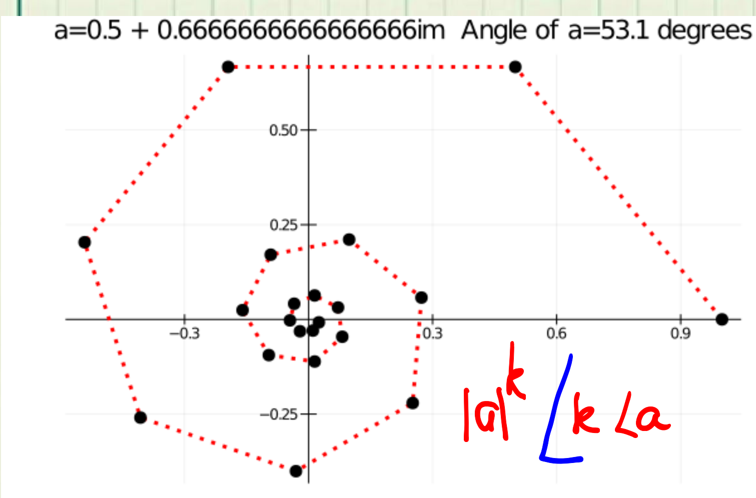
$$z[1] = a\, z[0]$$
$$z[2] = a\, z[1] = a^2 z[0]$$
$$z[3] = a\, z[2] = a^3 z[0]$$
$$\vdots$$

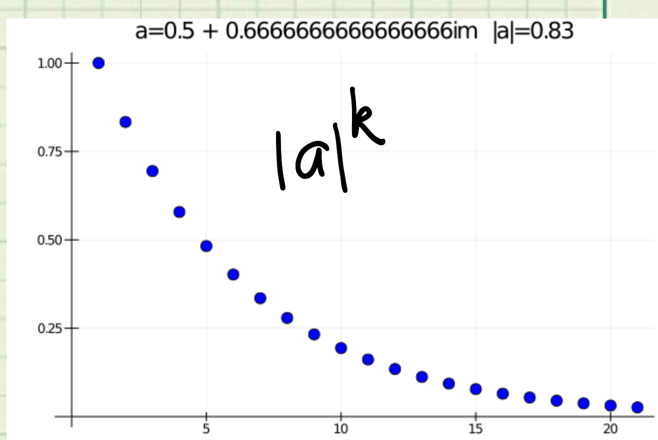
$$\boxed{z[k] = a^k z[0]}$$

We write $a = |a| \angle a = |a| e^{i \angle a}$

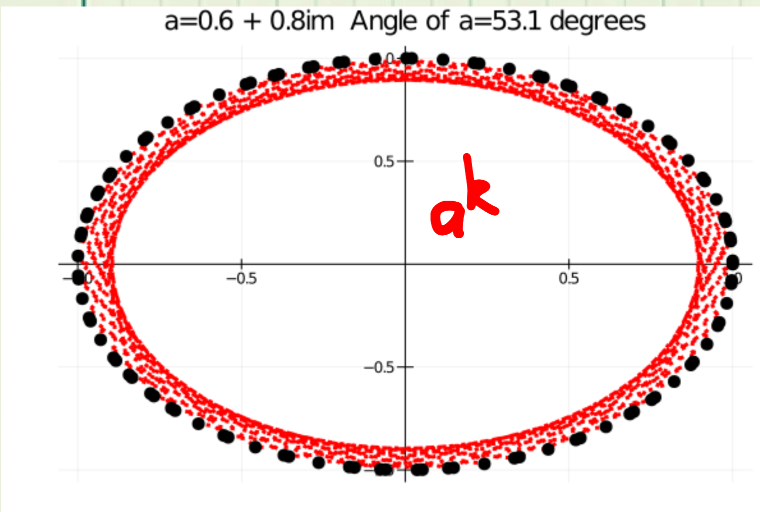
$$a^k = |a| \angle k \angle a = |a|^k e^{i k \angle a}$$

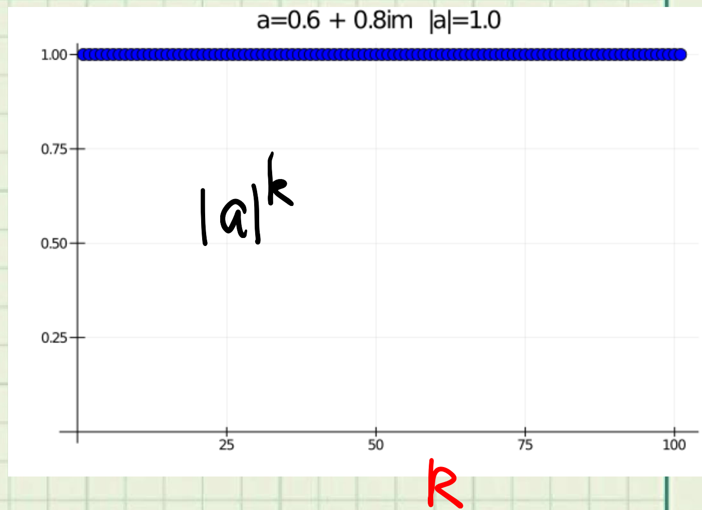
Case 1 $|a| < 1$ $\quad |z_0| \neq 0 \Rightarrow |z[k]| \searrow 0$



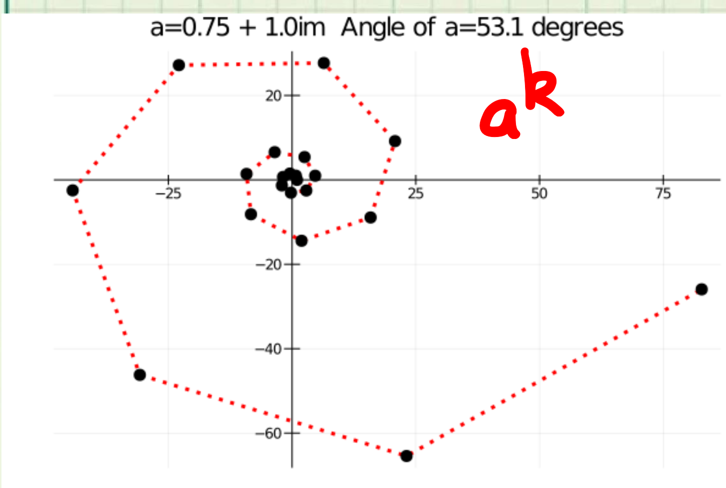


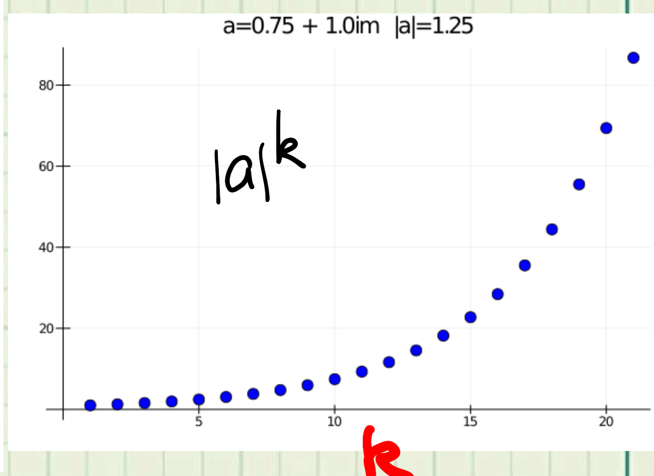
Case 2 $|a| = 1$ $\quad |z[k]| = |z_0| \quad k \to \infty$





Case 3 $|a| > 1$ $\quad |z_0| \neq 0 \Rightarrow |z[k]| \nearrow \infty$





$\therefore$ Scalar case is doable!

Even for "complex" numbers

Matrix Case?????

$X[k+1] = A\, X[k]$, $\quad A = n \times n$, x_0 given

$x[1] = A\, x[0]$

$x[2] = A \cdot x[1] = A \cdot A\, x[0] = A^2\, x[0]$

$x[3] = A \cdot x[2] = A \cdot A^2\, x[0] = A^3\, x[0]$

$\vdots$

$x[n] = A^n\, x[0]$

Challenge: Give conditions on A such that $\|x[k]\|$ contracts, blows up, stays bounded as $k \to \infty$

Even better: For a given initial condition x_0, predict the evolution of $x[k]$!

Start with A diagonal and 2×2

$$A = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$$

Exercise: $A^2 = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}\begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} = \begin{bmatrix} (\lambda_1)^2 & 0 \\ 0 & (\lambda_2)^2 \end{bmatrix}$

$$A^k = \begin{bmatrix} (\lambda_1)^k & 0 \\ 0 & (\lambda_2)^k \end{bmatrix}$$

$$\begin{bmatrix} x_1[k] \\ x_2[k] \end{bmatrix} = \begin{bmatrix} (\lambda_1)^k & 0 \\ 0 & (\lambda_2)^k \end{bmatrix}\begin{bmatrix} x_1[0] \\ x_2[0] \end{bmatrix} = \begin{bmatrix} (\lambda_1)^k x_1[0] \\ (\lambda_2)^k x_2[0] \end{bmatrix}$$

In general, we would have that

$$x_j[k] = (\lambda_j)^k \; x_j[0] \qquad 1 \leq j \leq n$$

$$x_j[0] \neq 0 \text{ and } \begin{cases} |\lambda_j| < 1 & \Rightarrow |x_j[k]| \xrightarrow[k\to\infty]{} 0 \\ |\lambda_j| = 1 & \Rightarrow |x_j[k]| = |x_j[0]| \\ |\lambda_j| > 1 & \Rightarrow |x_j[k]| \xrightarrow[k\to\infty]{} \infty \end{cases}$$

The elements on the diagonal of a diagonal matrix will soon be called eigenvalues, denoted by $\{\lambda_1, \lambda_2, \ldots, \lambda_n\}$

Key feature of a Diagonal Matrix

$$A = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{bmatrix}, \quad \text{let } v_j = e_j \quad \left\{ \text{Canonical basis vectors} \right.$$

THEN

$$\boxed{Av_j = \lambda_j v_j}$$

Multiplication of a special set of vectors $\{v_1, v_2, \ldots, v_n\}$ by A reduces to scaling the vector! A is acting like a scalar. Seems crazy!

Consider $x[k+1] = A\,x[k]$, and

suppose $\boxed{Av = \lambda v}$ We set $x_0 = v$

$$x[1] = Ax_0 = Av = \lambda v$$
$$x[2] = Ax[1] = A(\lambda v) = \lambda Av = \lambda^2 v$$
$$x[3] = A\,x[2] = A(\lambda^2 v) = \lambda^2 Av = \lambda^3 v$$
$$\vdots$$
$$x[k] = \lambda^k v$$

$\therefore$ Our previous analysis applies

$$|\lambda| < 1 \Rightarrow \|x[k]\| = \|\lambda^k v\| = |\lambda|^k \|v\| \xrightarrow[k\to\infty]{} 0$$
$$|\lambda| = 1 \Rightarrow \|x[k]\| = \|v\|$$
$$|\lambda| > 1 \Rightarrow \|x[k]\| = |\lambda|^k \|v\| \xrightarrow[k\to\infty]{} \infty$$

Question: Given an $n\times n$ matrix A, when does there exist a **basis** $\{v_1, v_2, \ldots, v_n\}$ for $\mathbb{R}^n$ such that $\boxed{Av_j = \lambda_j v_j}$?

Cool Things to Learn on Your Own

ROB 101 Handout: Grizzle & Ghaffari

December 4, 2020

Notes for Computational Linear Algebra by Jessy Grizzle, Director of Michigan Robotics

https://umich.instructure.com/courses/403066/files/folder/Booklet%3A%20Notes%20for%20Computational%20Linear%20Algebra

Learning Objectives

- Introduce material that is commonly included in a second or third year Linear Algebra Course
- Provide a resource for use after you leave ROB 101.

Outcomes

- Complex numbers obey the same rules of arithmetic as the real numbers, if you really understand the real numbers!
- Eigenvalues and eigenvectors of square matrices
- Symmetric matrices have real eigenvalues and admit orthonormal eigenvectors
- Positive definite matrices allow one to generalize the Euclidean norm
- The Singular Value Decomposition (SVD) allows one to quantify the degree to which vectors are linearly independent. This is super useful in engineering practice.
- Matrices are good for other things than representing systems of equations: they also allow one to transform vectors in interesting ways, giving rise to the concept of *linear transformations*.
- Many more facts about basis vectors.

1 Complex Numbers and Complex Vectors

Also introduce $\mathbb{C}^n$.

`https://youtu.be/T647CGsuOVU`

`https://youtu.be/2HrSG0fdxLY`

`https://youtu.be/N9QOLrfcKNc`

`https://youtu.be/DThAoT3q2V4`

`https://youtu.be/65wYmy8Pf-Y`

$$\mathbb{C}^n := \left\{ \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{bmatrix} \middle| \; z_i \in \mathbb{C} \right\}$$

Also introduce $\mathbb{C}^n$.

2 Eigenvalues and Eigenvectors

The study of eigenvalues and eigenvectors is very traditional in Linear Algebra courses. We skipped them for two reasons: (1) they get complicated really fast and (2) their most important application is probably the Singular Value Decomposition (SVD), a topic that we also skipped. The usual application of eigenvalues and eigenvectors is to "diagonalize" a matrix, and quite frankly, it is not a compelling application in serious engineering problems. Nevertheless, you may need this material in a future course, so we will give you the highlights.

2.1 General Square Matrices

Temporary Def. Let A be an $n \times n$ matrix with real coefficients. A scalar $\lambda \in \mathbb{R}$ is an **eigenvalue** (e-value) of A, if there exists a non-zero vector $v \in \mathbb{R}^n$ such that $A \cdot v = \lambda v$. Any such vector v is called an **eigenvector** (e-vector) associated with λ.

We note that if v is an e-vector, then so is αv for any $\alpha \neq 0$, and therefore, e-vectors are not unique. To find eigenvalues, we need to have conditions under which there exists $v \in \mathbb{R}^n$, $v \neq 0$, such that $A \cdot v = \lambda v$. Here they are,

$$A \cdot v = \lambda v \iff (\lambda I - A) \cdot v = 0 \overset{v \neq 0}{\iff} \det(\lambda I - A) = 0.$$

Example 2.1 *Let A be the 2×2 real matrix $A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$. Determine, if any, its e-values and e-vectors.*

Solution: To find e-values, we need to solve

$$\det(\lambda I - A) = \begin{vmatrix} \lambda & -1 \\ 1 & \lambda \end{vmatrix} = \lambda^2 + 1 = 0.$$

We compute the discriminant of this quadratic equation and we find

$$b^2 - 4ac = -4 < 0,$$

and therefore there are no real solutions. Hence, by our *temporary definition*, this 2×2 real matrix does not have any e-values, and hence, neither does it have any e-vectors.

If we were to allow e-values to be complex numbers, then we'd have two e-values corresponding to the two complex solutions of the quadratic equation $\lambda^2 + 1 = 0$, namely, $\lambda_1 = j$ and $\lambda_2 = -j$.

We'll see shortly that we'll also need to allow the e-vectors to have complex entries. Hence, we need to generalize our temporary definition. ■

Permanent Definition of Eigenvalues and Eigenvectors

Let A be an $n \times n$ matrix with real or complex coefficients. A scalar $\lambda \in \mathbb{C}$ is an **eigenvalue** (e-value) of A, if there exists a non-zero vector $v \in \mathbb{C}^n$ such that $Av = \lambda v$. Any such vector v is called an **eigenvector** (e-vector) associated with λ.

Eigenvectors are not unique.

- To find e-values, we solve $\det(\lambda I - A) = 0$ because

$$A \cdot v = \lambda v \iff (\lambda I - A) \cdot v = 0 \overset{v \neq 0}{\iff} \det(\lambda I - A) = 0. \tag{1}$$

- To find e-vectors, we find any non-zero $v \in \mathbb{C}^n$ such that

$$(\lambda I - A) \cdot v = 0. \tag{2}$$

Of course, if you prefer, you can solve $(A - \lambda I)v = 0$ when seeking e-vectors.

$$\det(\lambda I - A) = \lambda^n + \alpha_{n-1}\lambda^{n-1} + \cdots + \alpha_1 \lambda + \alpha_0$$

$$= (\lambda - \lambda_1)(\lambda - \lambda_2) \cdots (\lambda - \lambda_n)$$

Note: A real matrix $\Rightarrow$ $\alpha_{n-1}, \ldots, \alpha_0$ are real

$\{\lambda_1, \lambda_2, \ldots, \lambda_n\}$

Distinct e-values: $\lambda_k \neq \lambda_\ell \quad k \neq \ell$

Repeated e-values: $\lambda_k = \lambda_\ell$ some $k \neq \ell$

For each λ_k, we have $v_k \in \mathbb{C}^n$ such that $\boxed{A v_k = \lambda_k v_k}$

IF e-values are DISTINCT

$\{v_1, v_2, \ldots, v_n\}$ form a basis of $\mathbb{C}^n$ (hence, are linearly indep.)

λ_k real $\Rightarrow$ v_k real

λ_k complex $\Rightarrow$ v_k complex

$$v_k = v_k^{real} + j\, v_k^{imag}$$

{DEMO TIME}

Fundamental Theorem of Algebra (and a bit More)

Let A be an $n \times n$ matrix with real or complex coefficients. Then the following statements are true

- $\det(\lambda I - A) = \lambda^n + \alpha_{n-1}\lambda^{n-1} + \cdots \alpha_1\lambda + \alpha_0$, and if A is real, so are the coefficients $\alpha_{n-1}, \ldots, \alpha_0$.
- The degree n polynomial $\lambda^n + \alpha_{n-1}\lambda^{n-1} + \cdots \alpha_1\lambda + \alpha_0$ has n roots $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$ such that

$$\det(\lambda I - A) = (\lambda - \lambda_1)(\lambda - \lambda_2)\cdots(\lambda - \lambda_n).$$

  Each of the roots λ_i, $1 \leq i \leq n$, is an e-value of A.

- The e-values $\{\lambda_1, \ldots, \lambda_n\}$ are **distinct** if $\lambda_i \neq \lambda_k$ for all $i \neq k$.
- If $\lambda_i = \lambda_k$ for some $i \neq k$, then λ_i is a **repeated e-value**. The e-values can then be grouped into $1 \leq p < n$ sets of distinct roots $\{\lambda_1, \ldots, \lambda_p\}$ such that

$$\det(\lambda I - A) = (\lambda - \lambda_1)^{m_1}(\lambda - \lambda_2)^{m_2}\cdots(\lambda - \lambda_p)^{m_p}.$$

  The integer m_i is called the **algebraic multiplicity** of λ_i and their sum satisfies $m_1 + m_2 + \cdots + m_p = n$.

- An e-vector associated with λ_i is computed by finding non-zero solutions to (2).
- If the matrix A is real, then the e-values occur in **complex conjugate pairs**, that is, if λ_i is an e-value then so is λ_i^*.
- If the matrix A is real and the e-value λ_i is real, then the e-vector v_i can always be chosen to be real, that is, $v_i \in \mathbb{R}^n$ instead of $v_i \in \mathbb{C}^n$.
- There will always be at least one non-zero solution to (2), and because any non-zero multiple of a solution is also a solution, there will always be an infinite number of solutions to (2).
- If λ_i is a repeated e-value with algebraic multiplicity m_i, then the number of linearly independent e-vectors associated with λ_i is upper bounded by m_i. Another way to say this is, $1 \leq \dim\left(\mathrm{Null}(A)\right) \leq m_i$.
- **In Julia**, after `using LinearAlgebra`, the commands are $\Lambda =$ `eigvals(A)` and $V =$ `eigvecs(A)`

Example 2.2 *Let A be the 2×2 real matrix that we treated in Example 2.1, namely, $A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$. Determine its e-values and e-vectors in the sense of our "permanent" definition.*

Solution: As in Example 2.1, to find e-values, we solve

$$\det(\lambda I - A) = \begin{vmatrix} \lambda & -1 \\ 1 & \lambda \end{vmatrix} = \lambda^2 + 1 = 0.$$

We apply the quadratic equation and determine $\lambda_1 = j$ and $\lambda_2 = -j$. To find the eigenvectors, we solve

$$(A - \lambda_i I)v_i = 0.$$

The eigenvectors are

$$v_1 = \begin{bmatrix} 1 \\ j \end{bmatrix}, v_2 = \begin{bmatrix} 1 \\ -j \end{bmatrix}.$$

Note that the eigenvalues and eigenvectors each form complex conjugate pairs. Indeed,

$$\lambda_2 = \lambda_1^* \text{ and } v_2 = v_1^*.$$

■

Example 2.3 *Let A be the $n \times n$ identity matrix. Determine its e-values and e-vectors.*

Solution: $\det(\lambda I - I) = \det((\lambda - 1)I = 0 \iff \lambda = 1$. Alternatively, you can compute that $\det(\lambda I - I) = (\lambda - 1)^n$. Hence, the e-value $\lambda = 1$ is repeated n times, that is, $m_1 = n$. What are the e-vectors? We seek to solve

$$(A - \lambda I) \cdot v = 0 \iff (I - 1 \cdot I) \cdot v = 0 \iff 0_n \cdot v = 0,$$

where 0_n is the $n \times n$ matrix of all zeros! Hence, any non-zero vector $v \in \mathbb{R}^n$ is an e-vector. Moreover, if $\{v_1, \ldots, v_n\}$ is a basis for $\mathbb{R}^n$, then $\{v_1, \ldots, v_n\}$ is a set of n linearly independent e-vectors associated with $\lambda_1 = 1$. ■

Example 2.4 *Let $a \in \mathbb{R}$ be a constant and let A be the 4×4 matrix below. Determine its e-values and e-vectors.*

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

Solution: To find the e-values, we solve

$$\det(\lambda I - A) = \det\left(\begin{bmatrix} (\lambda - a) & -1 & 0 & 0 \\ 0 & (\lambda - a) & -1 & 0 \\ 0 & 0 & (\lambda - a) & -1 \\ 0 & 0 & 0 & (\lambda - a) \end{bmatrix}\right) = (\lambda - a)^4 = 0,$$

and hence there is one distinct e-value $\lambda_1 = a$. To solve for e-vector(s) we consider

$$0 = (A - aI) \cdot v = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \cdot v$$

and we find that the only solutions are multiples of

$$v = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

■

We've seen the extremes! A matrix with a single distinct e-value and a complete set of e-vectors (there were n linearly independent e-vectors associated with the e-value), and another matrix with a single distinct e-value, but only one linearly independent e-vector associated with it.

When the e-values are Distinct, the e-vectors form a Basis

Let A be an $n \times n$ matrix with coefficients in $\mathbb{R}$ or $\mathbb{C}$. If the e-values $\{\lambda_1, \ldots, \lambda_n\}$ are distinct, that is, $\lambda_i \neq \lambda_j$ for all $1 \leq i \neq j \leq n$, then the e-vectors $\{v_1, \ldots, v_n\}$ are linearly independent in ($\mathbb{C}^n$,$\mathbb{C}$).

Restatement of the result: If $\{\lambda_1, \ldots, \lambda_n\}$ are distinct, then $\{v_1, \ldots, v_n\}$ is a basis for ($\mathbb{C}^n$,$\mathbb{C}$).

2.2 Real Symmetric Matrices

We recall that a real $n \times n$ matrix A is **symmetric** if $A^\top = A$. E-values and e-vectors of symmetric matrices have nicer properties than those of general matrices.

E-values and E-vectors of Symmetric Matrices

- The e-values of a symmetric matrix are real. Because the e-values are real and the matrix is real, we can always chose the e-vectors to be real. Moreover, we can always normalize the e-vectors to have **norm one**.
- Just as with general matrices, the e-values of a symmetric matrix may be distinct or repeated. However, even when an e-value λ_i is repeated m_i times, there are always m_i linearly independent e-vectors associated with it. By applying Gram-Schmidt, we can always chose these e-vectors to be **orthonormal**.
- E-vectors associated with distinct e-values are automatically orthogonal. To be clear,

$$\left(A^\top = A,\ Av_i = \lambda_i v_i, Av_k = \lambda_k v_k, \text{ and } \lambda_i \neq \lambda_k\right) \implies v_i \perp v_k.$$

  Since we can assume they have length one, we have that the e-vectors are **orthonormal**.
- **In summary**, when A is symmetric, there is always an orthonormal basis $\{v_1, v_2, \ldots, v_n\}$ for $\mathbb{R}^n$ consisting of e-vectors of A. In other words, for all $1 \leq i \leq n$, $Av_i = \lambda_i v_i$, $||v_i|| = 1$, and for $k \neq i$, $v_k \perp v_i$.