Sect. 5.1: (Cont...)

Previously, we defined eigenvalues (e-values) and eigenvectors (e-vectors).

- e-values are scalars associated $A \in \mathbb{R}^{n \times n}$
- e-vectors are vectors associated $A \in \mathbb{R}^{n \times n}$
- They satisfy

$$\underbrace{A\underline{x}}_{\text{eigenvector}} = \overbrace{\lambda}^{\text{eigenvalue}}\underline{x}$$

Triangular Matrices

The e-values for a triangular matrix are the entries on its diagonal

Ex

$$A = \begin{vmatrix} 3 & 6 & -8 \\ 0 & 0 & 6 \\ 0 & 0 & 2 \end{vmatrix} \qquad B = \begin{vmatrix} 4 & 0 & 0 \\ -2 & 1 & 0 \\ 5 & 3 & 4 \end{vmatrix}$$

$$\lambda = 3, 0, 2 \qquad \lambda = 4, 1$$

To see this, we know that for λ to be an e-value, the equation

$$A\underline{x} = \lambda \underline{x} \iff (A - \lambda I)\underline{x} = \underline{0}$$

has a nontrivial solution if and only if it has at least one free variable. For this to be the case for a triangular matrix, at least one diagonal of $A-\lambda I$ must be 0, so the diagonal entries of A must be e-values.

<u>Ex</u> $A = \begin{vmatrix} 2 & 3 & 2 \\ 0 & -4 & 0 \\ 0 & 0 & 1 \end{vmatrix}$, then for $\lambda = -4$

$$A - \lambda I = A + 4I = \begin{vmatrix} 2+4 & 3 & 2 \\ 0 & -4+4 & 0 \\ 0 & 0 & 1+4 \end{vmatrix} = \begin{vmatrix} 6 & 3 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 5 \end{vmatrix}$$

So,

$$(A - \lambda I)\underline{x} = \underline{0} \implies \begin{vmatrix} 6 & 3 & 2 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 5 & 0 \end{vmatrix}$$

free variable for x_2 where we plugged in $\lambda = -4$

<u>Theorem</u>

If $\underline{v}_1 \cdots \underline{v}_r$ are e-vectors that correspond to distinct e-values $\lambda_1 \cdots \lambda_r$ of an $n \times n$ matrix A, then $\{\underline{v}_1 \cdots \underline{v}_r\}$ is linearly independent

<u>Pf.</u> (by contradiction)

Suppose the contrary, i.e, $\{\underline{v}_1 \cdots \underline{v}_r\}$ are linearly dependent. Since $\underline{v}_1$ is an e-vector, we know that $\underline{v}_1 \neq \underline{0}$ (by definition), so by a previous theorem (Sect. 1.7) we know that some vector $\underline{v}_{p+1}$ is a linear combination of $\underline{v}_1 \cdots \underline{v}_p$ if for $p+1 \le r$ if $\{\underline{v}_1 \cdots \underline{v}_r\}$ is lin. dep. So, there exist scalars $c_1 \cdots c_p$ such that

$$c_1 \underline{v}_1 + c_2 \underline{v}_2 + \cdots + c_p \underline{v}_p = \underline{v}_{p+1} \quad (*)$$

We can multiply both sides of $(*)$ by the matrix A. This gives

$$A(c_1 \underline{v}_1 + c_2 \underline{v}_2 + \cdots + c_p \underline{v}_p) = A\underline{v}_{p+1}$$

$$c_1 A\underline{v}_1 + c_2 A\underline{v}_2 + \cdots + c_p A\underline{v}_p = A\underline{v}_{p+1}$$

Since $\underline{v}_1 \cdots \underline{v}_p$ are e-vectors, we know that $A\underline{v}_j = \lambda_j \underline{v}_j$ for $j = 1:p+1$, so the equation becomes

a.) $$c_1 \lambda_1 \underline{v}_1 + c_2 \lambda_2 \underline{v}_2 + \cdots + c_p \lambda_p \underline{v}_p = \lambda_{p+1} \underline{v}_{p+1}$$

We can also multiply $(*)$ by λ_{p+1}

$$\lambda_{p+1}(c_1 \underline{v}_1 + c_2 \underline{v}_2 + \cdots + c_p \underline{v}_p) = \lambda_{p+1} \underline{v}_{p+1}$$

b.) $$c_1 \lambda_{p+1} \underline{v}_1 + c_2 \lambda_{p+1} \underline{v}_2 + \cdots + c_p \lambda_{p+1} \underline{v}_p = \lambda_{p+1} \underline{v}_{p+1}$$

Subtracting a.) − b.) gives

$$(c_1 \lambda_1 \underline{v}_1 + \cdots + c_p \lambda_p \underline{v}_p) - (c_1 \lambda_{p+1} \underline{v}_1 + \cdots + c_p \lambda_{p+1} \underline{v}_p) = \underline{0}$$

$$(c_1 \lambda_1 \underline{v}_1 - c_1 \lambda_{p+1} \underline{v}_1) + \cdots + (c_p \lambda_p \underline{v}_p + \cdots c_p \lambda_{p+1} \underline{v}_p) = \underline{0}$$

$$c_1(\lambda_1 - \lambda_{p+1})\underline{v}_1 + \cdots + c_p(\lambda_p - \lambda_{p+1})\underline{v}_p = \underline{0}$$

We know $\underline{v}_1 \cdots \underline{v}_p \neq \underline{0}$ because they're e-vectors.
We know $\lambda_j - \lambda_{p+1} \neq \underline{0}$ because $\lambda_1 \cdots \lambda_p$ are distinct.
This implies for the equation to hold, we must $c_1 = c_2 = \cdots = c_p = 0$. This produces a contradiction $(*)$, because

$$c_1\underline{v}_1 + c_2\underline{v}_2 + \cdots + c_p\underline{v}_p = \underline{v}_{p+1} \quad (*)$$

$$0\,\underline{v}_1 + 0\,\underline{v}_2 + \cdots + 0\,\underline{v}_p = \underline{v}_{p+1}$$

$$\underline{0} = \underline{v}_{p+1}$$

But, $\underline{v}_{p+1}$ is an e-vector and cannot be $\underline{0}$ by definition. So, the assumption that $\underline{v}_{p+1}$ is a linear combination of $\underline{v}_1 \cdots \underline{v}_p$ must be false $\Rightarrow$ $\underline{v}_1 \cdots \underline{v}_r$ must be linearly independent.

Sect 5.2: The Characteristic Equation

In Sect 5.1, we defined e-values and e-vectors as solutions to

$$A\underline{x} = \lambda x$$

$\hookrightarrow$ $\underline{x} \in \mathbb{R}^n$ vector & $\lambda \in \mathbb{R}$ scalar

We learned how to determine an e-vector or find a basis for the e-space corresponding to a known e-value

↳ To do this, we found the null space of $A - \lambda I$ by solving

$$(A - \lambda I)\underline{x} = \underline{0}$$

This begs the question

Q: How do we find all the e-values of a matrix $A \in \mathbb{R}^{n \times n}$?

A: The characteristic equation.

Def

The characteristic equation is the polynomial defined as:

$$\det(A - \lambda I) = 0 \leftarrow \text{scalar}$$

Ex

Let $A = \begin{vmatrix} 2 & 3 \\ 3 & -6 \end{vmatrix}$ and $A - \lambda I = \begin{vmatrix} 2-\lambda & 3 \\ 3 & -6-\lambda \end{vmatrix}$

Then the char. equation is

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

$$(2-\lambda)(-6-\lambda) - 3 \cdot 3 = 0$$

$$\lambda^2 + 4\lambda - 21 = 0 \quad \swarrow$$

a polynomial in the variable λ

Ex

Let $A = \begin{vmatrix} 5 & -2 & 6 & -1 \\ 0 & 3 & -8 & 0 \\ 0 & 0 & 5 & 4 \\ 0 & 0 & 0 & 1 \end{vmatrix}$ and $A - \lambda I = \begin{vmatrix} 5-\lambda & -2 & 6 & -1 \\ 0 & 3-\lambda & -8 & 0 \\ 0 & 0 & 5-\lambda & 4 \\ 0 & 0 & 0 & 1-\lambda \end{vmatrix}$

Then, the char. equation is

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

$$(5-\lambda)(3-\lambda)(5-\lambda)(1-\lambda) = 0$$

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

Observations

- For $A \in \mathbb{R}^{n \times n}$, the characteristic equation is a polynomial of degree n.
- For triangular matrices, the characterist equation is the product of the diagonal entries of $A - \lambda I$. This polynomial is already factored, so it's easy to find its roots.
- Theorem: A scalar λ is an e-value of an $n \times n$ matrix A if and only if λ is a root of the characteristic equation, i.e.

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

- Since the degree of the characteristic equation is n, it has at most n roots, i.e, there are at most n distinct e-values for a matrix $A \in \mathbb{R}^{n \times n}$

- If $\lambda = 0$ is an e-value, then it satisfies

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

$\Rightarrow$ A is not invertible by the invertible matrix theorem.

<u>Theorem</u> (IMT Cont...)

Let A be an $n \times n$ matrix, then A is invertible if and only if

- 0 is **not** an eigenvalue of A