Lecture 14

Solutions to in-class exercises:

1. Determine if the following matrices are diagonalizable.

a) $\begin{pmatrix} 1 & 7 & 8 & 9 & 10 & 11 \\ & 2 & 12 & 13 & 14 & 15 \\ & & 3 & 16 & 17 & 18 \\ & & & 4 & 19 & 20 \\ & 0 & & & 5 & 21 \\ & & & & & 6 \end{pmatrix}$

char poly $= (1-\lambda)(2-\lambda)(3-\lambda)(4-\lambda)(5-\lambda)(6-\lambda)$. 6 different eigenvalues for a 6×6 matrix

$\Rightarrow$ Diagonalizable.

b) $\begin{pmatrix} 3 & 0 & 1 \\ 0 & -2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$ char poly $= (3-\lambda)^2(-2-\lambda) \Rightarrow a_3 = 2$

$g_3 = \dim(E_3) = \dim\left(\operatorname{Ker}\begin{pmatrix} 0 & 0 & 1 \\ 0 & -5 & 0 \\ 0 & 0 & 0 \end{pmatrix}\right) = 3 - \operatorname{rank}\begin{pmatrix} 0 & 0 & 1 \\ 0 & -5 & 0 \\ 0 & 0 & 0 \end{pmatrix} = 1 < 2$

$\Rightarrow$ Not diagonalizable.

2. a) Find a 5×5 matrix with eigenvalues -1 and 0, $a_{-1} = 3$, $a_0 = 2$, $g_{-1} = 1$, $g_0 = 2$.

$\begin{pmatrix} -1 & 1 & & & \\ & -1 & 1 & & \\ & & -1 & & \\ & & & 0 & \\ & & & & 0 \end{pmatrix}$: char poly $= (-1-\lambda)^3(-\lambda)^2 \Rightarrow a_{-1} = 3$ ✓

$a_0 = 2$ ✓

$g_{-1} = \dim\left(\operatorname{Ker}\begin{pmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & 0 & & \\ & & & 1 & \\ & & & & 1 \end{pmatrix}\right) = 5 - \operatorname{rank}\begin{pmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & 0 & & \\ & & & 1 & \\ & & & & 1 \end{pmatrix} = 5 - 4 = 1.$ ✓

$g_0 = \dim\left(\operatorname{Ker}\begin{pmatrix} -1 & 1 & & & \\ & -1 & 1 & & \\ & & -1 & & \\ & & & 0 & \\ & & & & 0 \end{pmatrix}\right) = 5 - \operatorname{rank} = 5 - 3 = 2$ ✓

b) Find a 7×7 matrix with eigenvalues -1 and 0, $a_{-1} = 3$, $a_0 = 2$, $g_{-1} = 1$, $g_0 = 2$.

$\begin{pmatrix} -1 & 1 & & & & & \\ & -1 & 1 & & & & \\ & & -1 & & & & \\ & & & 0 & & & \\ & & & & 0 & & \\ & & & & & 0 & -1 \\ & & & & & 1 & 0 \end{pmatrix}$ char poly: $(-1-\lambda)^3(-\lambda)^2(\lambda^2+1)$

(The rest is similar to a)).

3. Write $\frac{1}{2-3i}$ in the form $a+bi$. $\quad \frac{1}{2-3i} = \frac{(2+3i)}{(2-3i)(2+3i)} = \frac{2+3i}{4+9} = \frac{2}{13} + \frac{3}{13}i.$

Recap: • Diagonalizability $\Leftrightarrow$ $\sum g_\lambda = n$

• Problems that can arise: • $\sum a_\lambda < n$ Example: $\begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 2 \end{pmatrix}$ char poly $= (\lambda^2+1)(\lambda-2)$.

$\sum_{\lambda \text{ eig}} a_\lambda = a_2 = 1 < 3$.

• $g_\lambda < a_\lambda$ Example: $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ char poly $= (1-\lambda)^2$

$g_1 = 1 < 2 = g_2$.

Today: "We get rid of the first problem".

Discussion: Complex numbers are of the form $a+bi$, with $a, b \in \mathbb{R}$, and i a number s.t. $i^2 = -1$.

They can be added, subtracted, multiplied, and divided

$(a+bi)(c+di) = (ac-bd) + i(ad+bc)$. $\quad \frac{1}{a+bi} = \frac{a-bi}{a^2+b^2} = \frac{a}{a^2+b^2} - i \cdot \frac{b}{a^2+b^2}$

Notation: $\mathbb{C}$.

Complex numbers were invented in order to solve polynomial equations. An example could be $x^2+x+1=0$. The solutions are $x = \frac{-1 \pm \sqrt{1-4}}{2} = \frac{-1 \pm \sqrt{-3}}{2}$

Now the introduction of $i = \sqrt{-1}$ allows us to write $\sqrt{-3} = \sqrt{-1} \cdot \sqrt{3} = i \cdot \sqrt{3}$.

So our complex solutions are $\frac{-1+i\sqrt{3}}{2}$ and $\frac{-1-i\sqrt{3}}{2}$.

When finding eigenvalues of a matrix, sometimes we get factors such as (λ^2+9) or $(\lambda^2+\lambda+1)$, which cannot be factored further using real numbers. However, using complex numbers, both of these factor: $\lambda^2+9 = (\lambda+3i)(\lambda-3i)$ and $\lambda^2+\lambda+1 = \left(\lambda - \frac{1+i\sqrt{3}}{2}\right)\left(\lambda - \frac{1-i\sqrt{3}}{2}\right)$.

One may wonder if using complex numbers, one could factor <u>any</u> polynomial into linear factors.

Theorem 1 ("Fundamental" theorem of algebra). Let $p(x)$ be a polynomial with coefficients in $\mathbb{C}$ of degree n.

Then, $p(x) = (x-z_1)(x-z_2)\cdots(x-z_n)$ for some complex numbers z_i, possibly repeated.

Proof: omitted.

Corollary: Let A be an $n \times n$ matrix. If we allow complex eigenvalues, $\sum_{\lambda \text{ eig.}} a_\lambda = n$.

Example 1: $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \Rightarrow$ char poly $(A) = \lambda^2 + 1 = (\lambda + i)(\lambda - i) \Rightarrow \begin{cases} a_i = 1 \\ a_{-i} = 1 \end{cases}$ } these add up to $n = 2$

Since it has 2 different eigenvalues, A is diagonalizable "over $\mathbb{C}$".

Now $E_i = \operatorname{Ker} \begin{pmatrix} -i & -1 \\ 1 & -i \end{pmatrix} =$

$$\left(\begin{array}{cc|c} -i & -1 & 0 \\ 1 & -i & 0 \end{array}\right) \xrightarrow{I \to \frac{1}{-i} I} \left(\begin{array}{cc|c} 1 & \underbrace{\tfrac{1}{i}}_{=-i} & 0 \\ 1 & -i & 0 \end{array}\right) \xrightarrow{II \to II - I} \left(\begin{array}{cc|c} 1 & -i & 0 \\ 0 & 0 & 0 \end{array}\right) \Rightarrow E_i = \operatorname{Span}\left(\begin{pmatrix} i \\ 1 \end{pmatrix}\right)$$

$E_{-i} = \operatorname{Ker} \begin{pmatrix} i & -1 \\ 1 & i \end{pmatrix}$

$$\left(\begin{array}{cc|c} i & -1 & 0 \\ 1 & i & 0 \end{array}\right) \xrightarrow{I \to \frac{1}{i} I} \left(\begin{array}{cc|c} 1 & \underbrace{-\tfrac{1}{i}}_{=i} & 0 \\ 1 & i & 0 \end{array}\right) \xrightarrow{II \to II - I} \left(\begin{array}{cc|c} 1 & i & 0 \\ 0 & 0 & 0 \end{array}\right) \Rightarrow E_{-i} = \operatorname{Span}\left(\begin{pmatrix} -i \\ 1 \end{pmatrix}\right)$$

So set $S = \begin{pmatrix} i & -i \\ 1 & 1 \end{pmatrix}$. Then $S^{-1} = \begin{pmatrix} -\frac{i}{2} & \frac{1}{2} \\ \frac{i}{2} & \frac{1}{2} \end{pmatrix} \Rightarrow \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} i & -i \\ 1 & 1 \end{pmatrix} \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} \begin{pmatrix} -\frac{i}{2} & \frac{1}{2} \\ \frac{i}{2} & \frac{1}{2} \end{pmatrix}$

Definition 1 ("over $\mathbb{C}$"): A matrix is diagonalizable over $\mathbb{C}$ if it can be diagonalized using complex numbers. (See Example 1)

Remark: We will say "over $\mathbb{R}$" or "over $\mathbb{C}$" when we need to specify which numbers we are using. By default, you should assume everything is over $\mathbb{C}$.

Discussion: The use of complex allows us to "almost diagonalize" every matrix, the only problem being when the geometric and algebraic multiplicities do not coincide.

Example 2: The matrix $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ is still not diagonalizable, not even over $\mathbb{C}$.

After diagonal form, the best next thing is the Jordan normal form.

Definition 2: A Jordan matrix is a matrix of the form

$$\begin{pmatrix} \boxed{\begin{matrix} \lambda_1 & 1 & & \\ & \lambda_1 & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_1 \end{matrix}} & & & \\ & \boxed{\begin{matrix} \lambda_2 & 1 & & \\ & \lambda_2 & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_2 \end{matrix}} & & \\ & & \ddots & \\ & & & \boxed{\begin{matrix} \lambda_k & 1 & & \\ & \lambda_k & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_k \end{matrix}} \end{pmatrix}$$

where $\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_k$.

(blank entries are zeroes)

Theorem 2: Any $n \times n$ matrix (with real or complex entries) is similar over $\mathbb{C}$ to a Jordan matrix.

Remark: In other words, for any $n \times n$ matrix A, there exists an invertible matrix S such that $A = S^{-1} J S$, and J is a Jordan matrix.

Proof: omitted.

Definition 3: The matrix J in Theorem 2 is called the Jordan normal form of A.

Theorem 3: Two matrices are similar if and only if they have the same Jordan normal form.

Proof: omitted.

Example 3: $\begin{pmatrix} 2 & 1 & & \\ & 2 & 1 & \\ & & 2 & 0 \\ & & & 2 \end{pmatrix}$ and $\begin{pmatrix} 2 & 1 & & \\ & 2 & 0 & \\ & & 2 & 1 \\ & & & 2 \end{pmatrix}$ are not similar by Theorem 3.

Remark: we won't learn how to find/compute the Jordan normal form of a matrix, although it's not very far from the "algorithm" to diagonalize a matrix.

Discussion: we have seen linear algebra "over $\mathbb{R}$" and "over $\mathbb{C}$", and I'm telling you that everything "works out the same" in the complex case. To be more rigorous, one would have to repeat the whole course replacing $\mathbb{R}$ by $\mathbb{C}$. Better yet, one could create an abstract notion of which $\mathbb{R}$ and $\mathbb{C}$ are special cases, that way we only have to do the work once. Moreover, this could be applied to "numbers" that aren't real or complex.

Note: the rest of today's lecture is not examinable.

Definition 4:

Classic definition [edit]

Formally, a field is a set F together with two binary operations on F called *addition* and *multiplication*.[1] A binary operation on F is a mapping $F \times F \to F$, that is, a correspondence that associates with each ordered pair of elements of F a uniquely determined element of F.[2][3] The result of the addition of a and b is called the sum of a and b, and is denoted $a + b$. Similarly, the result of the multiplication of a and b is called the product of a and b, and is denoted ab or $a \cdot b$. These operations are required to satisfy the following properties, referred to as *field axioms* (in these axioms, a, b, and c are arbitrary elements of the field F):

- Associativity of addition and multiplication: $a + (b + c) = (a + b) + c$, and $a \cdot (b \cdot c) = (a \cdot b) \cdot c$.
- Commutativity of addition and multiplication: $a + b = b + a$, and $a \cdot b = b \cdot a$.
- Additive and multiplicative identity: there exist two different elements 0 and 1 in F such that $a + 0 = a$ and $a \cdot 1 = a$.
- Additive inverses: for every a in F, there exists an element in F, denoted $-a$, called the *additive inverse* of a, such that $a + (-a) = 0$.
- Multiplicative inverses: for every $a \neq 0$ in F, there exists an element in F, denoted by a^{-1} or $1/a$, called the *multiplicative inverse* of a, such that $a \cdot a^{-1} = 1$.
- Distributivity of multiplication over addition: $a \cdot (b + c) = (a \cdot b) + (a \cdot c)$.

This may be summarized by saying: a field has two operations, called addition and multiplication; it is an abelian group under addition with 0 as the additive identity; the nonzero elements are an abelian group under multiplication with 1 as the multiplicative identity; and multiplication distributes over addition.

Even more summarized: a field is a commutative ring where $0 \neq 1$ and all nonzero elements are invertible.

Example 4: $\mathbb{Q}, \mathbb{R}, \mathbb{C}$.

Example 5: Consider the set $F = \{0,1\}$, with addition table
$$\begin{array}{c|cc} + & 0 & 1 \\ \hline 0 & 0 & 1 \\ 1 & 1 & 0 \end{array}$$
and multiplication table
$$\begin{array}{c|cc} \cdot & 0 & 1 \\ \hline 0 & 0 & 0 \\ 1 & 0 & 1 \end{array}$$

This field is denoted $\mathbb{F}_2$, "the field of two elements".

Definition 5: A vector space over a field F is a set V with a binary operation $+$ called sum as well as a scalar multiplication sending $(\lambda, v) \mapsto \lambda v \in V$ for each $\lambda \in F$, satisfying:

Axiom	Meaning
Associativity of vector addition	$\mathbf{u} + (\mathbf{v} + \mathbf{w}) = (\mathbf{u} + \mathbf{v}) + \mathbf{w}$
Commutativity of vector addition	$\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$
Identity element of vector addition	There exists an element $\mathbf{0} \in V$, called the *zero vector*, such that $\mathbf{v} + \mathbf{0} = \mathbf{v}$ for all $\mathbf{v} \in V$.
Inverse elements of vector addition	For every $\mathbf{v} \in V$, there exists an element $-\mathbf{v} \in V$, called the *additive inverse* of $\mathbf{v}$, such that $\mathbf{v} + (-\mathbf{v}) = \mathbf{0}$.
Compatibility of scalar multiplication with field multiplication	$a(b\mathbf{v}) = (ab)\mathbf{v}$ [nb 3]
Identity element of scalar multiplication	$1\mathbf{v} = \mathbf{v}$, where 1 denotes the multiplicative identity in F.
Distributivity of scalar multiplication with respect to vector addition	$a(\mathbf{u} + \mathbf{v}) = a\mathbf{u} + a\mathbf{v}$
Distributivity of scalar multiplication with respect to field addition	$(a + b)\mathbf{v} = a\mathbf{v} + b\mathbf{v}$

Fact: the notions we have seen in this course can all be carried out in this abstract setting.

Application: lights out

Lights out is a game where squares on a grid are on/off and the goal is to turn all of them off.

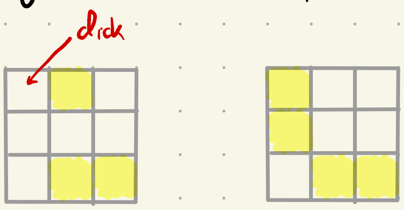
The catch is that clicking on a square also changes the adjacent squares:

We can win this game using linear algebra!

Consider $(\mathbb{F}_2)^9 = \left\{ \begin{pmatrix} x_{11} \\ x_{12} \\ x_{13} \\ x_{21} \\ x_{22} \\ x_{23} \\ x_{31} \\ x_{32} \\ x_{33} \end{pmatrix} : x_{ij} \in \mathbb{F}_2 \right\}$, canonical basis $e_{11}, \dots, e_{33}$.

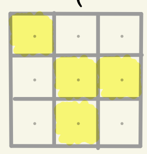
Then lights configurations $\xleftrightarrow{1:1}$ $\mathbb{F}_2^9$

Clicking the top left square $\longleftrightarrow$ adding the vector $e_{11}+e_{12}+e_{21}$. For instance:



$e_{12}+e_{32}+e_{33} \longrightarrow e_{11} + 2e_{12} + e_{21} + e_{32} + e_{33}$

This holds similarly for the other squares. Define $v_{11}, v_{12}, v_{13}, v_{21}, \dots, v_{33}$ similarly.

Let $y \in (\mathbb{F}_2)^9$ describe our configuration:  $\rightarrow y = e_{11}+e_{22}+e_{23}+e_{32} = \begin{pmatrix} 1\\0\\0\\0\\1\\1\\0\\1\\0 \end{pmatrix}$

The question "which squares should I click?"

$\longleftrightarrow$ what scalars $\lambda_{11}, \dots, \lambda_{33} \in \mathbb{F}_2$ should I choose so that $\lambda_{11}v_{11}+\dots+\lambda_{33}v_{33} = y$?

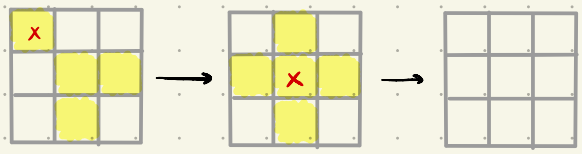
This is a system of linear equations! Namely, $A = \begin{pmatrix} | & & | \\ v_{11} & \cdots & v_{33} \\ | & & | \end{pmatrix}$, $x = \begin{pmatrix} \lambda_{11} \\ \vdots \\ \lambda_{33} \end{pmatrix}$, $\Rightarrow Ax = y$.

To solve this, take $A^{-1}y$.

$$A = \begin{pmatrix} 1&1&0&1&0&0&0&0&0\\ 1&1&1&0&1&0&0&0&0\\ 0&1&1&0&0&1&0&0&0\\ 1&0&0&1&1&0&1&0&0\\ 0&1&0&1&1&1&0&1&0\\ 0&0&1&0&1&1&0&0&1\\ 0&0&0&1&0&0&1&1&0\\ 0&0&0&0&1&0&1&1&1\\ 0&0&0&0&0&1&0&1&1 \end{pmatrix} \Rightarrow A^{-1} = \begin{pmatrix} 1&0&1&0&0&1&1&1&0\\ 0&0&0&0&1&0&1&1&1\\ 1&0&1&1&0&0&0&1&1\\ 0&0&1&0&1&1&0&0&1\\ 0&1&0&1&1&1&0&1&0\\ 1&0&0&1&1&0&1&0&0\\ 1&1&0&0&0&1&1&0&1\\ 1&1&1&0&1&0&0&0&0\\ 0&1&1&1&0&0&1&0&1 \end{pmatrix}$$

$$A^{-1}y = \begin{pmatrix} 1&0&1&0&0&1&1&1&0\\ 0&0&0&0&1&0&1&1&1\\ 1&0&1&1&0&0&0&1&1\\ 0&0&1&0&1&1&0&0&1\\ 0&1&0&1&1&1&0&1&0\\ 1&0&0&1&1&0&1&0&0\\ 1&1&0&0&0&1&1&0&1\\ 1&1&1&0&1&0&0&0&0\\ 0&1&1&1&0&0&1&0&1 \end{pmatrix} \begin{pmatrix} 1\\0\\0\\0\\1\\1\\0\\1\\0 \end{pmatrix} = \begin{pmatrix} 1\\0\\0\\0\\1\\0\\0\\0\\0 \end{pmatrix} = e_{11} + e_{22}.$$

And indeed:



done!

In-class exercise: diagonalize the matrix $\begin{pmatrix} & 1 & \\ & & 1 \\ 1 & & \end{pmatrix}$ over $\mathbb{C}$.