This document is a **transcript** of the lecture, with extra summary and vocabulary sections for your convenience. Due to time constraints, the notes may sometimes only contain limited illustrations, proofs, and examples; for a more thorough discussion of the course content, **consult the textbook**.

Summary

Quick summary of today's notes. Lecture starts on next page.

- Given real numbers $a, b \in \mathbb{R}$, define $a + bi = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$.

  In this notation, we think of 1 as the matrix $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ and i as the matrix $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$.

  The set of *complex numbers* is $\mathbb{C} = \{a + bi : a, b \in \mathbb{R}\} = \left\{ \begin{bmatrix} a & -b \\ b & a \end{bmatrix} : a, b \in \mathbb{R} \right\}$.

  We view $\mathbb{R}$ as a subset of $\mathbb{C}$ by setting $a = a + 0i = \begin{bmatrix} a & 0 \\ 0 & a \end{bmatrix}$.

- We can add, subtract, multiply, and invert complex numbers, since they are 2×2 matrices.

  The identity "$i^2 = -1$" holds in the sense that $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}^2 = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} = -\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$.

- Once we get used to these operations, another useful way to view the elements of $\mathbb{C}$ is as formal expressions $a + bi$ where $a, b \in \mathbb{R}$ and i is a symbol that satisfies $i^2 = -1$.

  Addition, subtraction, and multiplication work just like polynomials, but substituting -1 for i^2.

- Suppose $p(x) = a_n x^n + a_{n-1} x^{n-1} + \dots a_1 x + a_0$ is a polynomial with coefficients $a_0, a_1, \dots, a_n \in \mathbb{C}$.

  Assume $a_n \neq 0$ so that $p(x)$ has *degree* n.

  Then there are are n (not necessarily distinct) complex numbers $r_1, r_2, \dots, r_n \in \mathbb{C}$ such that

  $$p(x) = a_n(x - r_1)(x - r_2) \cdots (x - r_n).$$

  The numbers $r_1, r_2, \dots, r_n$ are the *roots* of $p(x)$.

1 Last time: methods to check diagonalizability

Let n be a positive integer and let A be an $n \times n$ matrix.

Remember that A is *diagonalizable* if $A = PDP^{-1}$ where P is an invertible $n \times n$ matrix and D is an $n \times n$ diagonal matrix. In other words, A is diagonalizable if A is similar to a diagonal matrix.

Suppose $v_1, v_2, \dots, v_n \in \mathbb{R}^n$ are linearly independent vectors and $\lambda_1, \lambda_2, \dots, \lambda_n$ are numbers. Define

$$P = \left[\begin{array}{cccc} v_1 & v_2 & \dots & v_n \end{array}\right] \qquad \text{and} \qquad D = \left[\begin{array}{cccc} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{array}\right].$$

If $A = PDP^{-1}$ then $Av_i = PDP^{-1}v_i = PDe_i = \lambda_i Pe_i = \lambda_i v_i$ for each $i = 1, 2, \dots, n$.

In other words, when $A = PDP^{-1}$, the columns of P are a basis for $\mathbb{R}^n$ made up of eigenvectors of A.

Matrices that are not diagonalizable.

Proposition. Let A be an $n \times n$ upper-triangular matrix with all entries on the diagonal equal to λ.

If A is not the diagonal matrix λI, then A is not diagonalizable.

Proof. Suppose $A = PDP^{-1}$ where D is diagonal. Every diagonal entry of D is an eigenvalue for A.

The only eigenvalue of A is λ so $D = \lambda I$ and $A = P(\lambda I)P^{-1} = \lambda PIP^{-1} = \lambda PP^{-1} = \lambda I$. □

The following result summarizes everything we need to know about diagonalizability: how to determine if a matrix A is diagonalizable, and then how to compute the decomposition $A = PDP^{-1}$ if it exists.

Theorem. Let A be an $n \times n$ matrix. Suppose $\lambda_1, \lambda_2, \dots, \lambda_p$ are the distinct eigenvalues of A.

Let $d_i = \dim \mathrm{Nul}(A - \lambda_i I)$ for $i = 1, 2, \dots, p$.

By the definition of an eigenvalue, we have $1 \leq d_i \leq n$ for each i. Moreover, the following holds:

1. We always have $d_1 + d_2 + \dots + d_p \leq n$.
2. The matrix A is diagonalizable if and only if $d_1 + d_2 + \dots + d_p = n$.
3. Suppose A is diagonalizable. Let $D_i = \lambda_i I_{d_i}$ and define D as the $n \times n$ diagonal matrix

$$D = \left[\begin{array}{cccc} D_1 & & & \\ & D_2 & & \\ & & \ddots & \\ & & & D_p \end{array}\right].$$

Choose n vectors $v_1, v_2, \dots, v_n \in \mathbb{R}^n$ such that the first d_1 vectors are a basis for $\mathrm{Nul}(A - \lambda_1 I)$, the next d_2 vectors are a basis for $\mathrm{Nul}(A - \lambda_2 I)$, the next d_3 vectors are a basis for $\mathrm{Nul}(A - \lambda_3 I)$, and so on, so that the last d_p vectors are basis for $\mathrm{Nul}(A - \lambda_p I)$. Then $A = PDP^{-1}$ for

$$P = \left[\begin{array}{cccc} v_1 & v_2 & \dots & v_n \end{array}\right].$$

2 Complex numbers

For the rest of this lecture, let $i = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$. Recall that $I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$.

Suppose $a, b \in \mathbb{R}$. Both i and I_2 are 2×2 matrices, so we can form the sum $aI_2 + bi$.

To simplify our notation, we will write 1 instead of I_2 and $a + bi$ instead of $aI_2 + bi$.

We consider $a = a + 0i$ and $bi = 0 + bi$ and $0 = 0 + 0i$. With this convention, we have

$$a + bi = a \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + b \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix} + \begin{bmatrix} 0 & -b \\ b & 0 \end{bmatrix} = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}.$$

Define $\mathbb{C} = \{a + bi : a, b \in \mathbb{R}\} = \left\{ \begin{bmatrix} a & -b \\ b & a \end{bmatrix} : a, b \in \mathbb{R} \right\}$. This is called the set of *complex numbers*.

According to our definition, each element of $\mathbb{C}$ is a 2×2 matrix, to be called a *complex number*.

Fact. We can add complex numbers together. If $a, b, c, d \in \mathbb{R}$ then

$$(a + bi) + (c + di) = \begin{bmatrix} a & -b \\ b & a \end{bmatrix} + \begin{bmatrix} c & -d \\ d & c \end{bmatrix} = \begin{bmatrix} a + c & -b - d \\ b + d & a + c \end{bmatrix} = (a + c) + (b + d)i \in \mathbb{C}.$$

Clearly $\boxed{(a + bi) + (c + di) = (c + di) + (a + bi) = (a + c) + (b + d)i}$.

Fact. We can subtract complex numbers. If $a, b, c, d \in \mathbb{R}$ then

$$(a + bi) - (c + di) = \begin{bmatrix} a & -b \\ b & a \end{bmatrix} - \begin{bmatrix} c & -d \\ d & c \end{bmatrix} = \begin{bmatrix} a - c & -b + d \\ b - d & a - c \end{bmatrix} = (a - c) + (b - d)i \in \mathbb{C}.$$

Fact. We can multiply complex numbers. If $a, b, c, d \in \mathbb{R}$ then

$$(a + bi)(c + di) = \begin{bmatrix} a & -b \\ b & a \end{bmatrix} \begin{bmatrix} c & -d \\ d & c \end{bmatrix} = \begin{bmatrix} ac - bd & -(ad + bc) \\ ad + bc & ac - bd \end{bmatrix} = (ac - bd) + (ad + bc)i \in \mathbb{C}.$$

Note that $\boxed{(a + bi)(c + di) = (c + di)(a + bi) = (ac - bd) + (ad + bc)i}$.

Fact. We can multiply complex numbers by real numbers. If $a, b, x \in \mathbb{R}$ then define

$$(a + bi)x = x(a + bi) = x \begin{bmatrix} a & -b \\ b & a \end{bmatrix} = \begin{bmatrix} ax & -bx \\ bx & ax \end{bmatrix} = (ax) + (bx)i \in \mathbb{C}.$$

Note that this is the same as the product $(a + bi)(x + 0i)$.

Fact. We can divide complex numbers by nonzero real numbers. If $a, b, x \in \mathbb{R}$ and $x \neq 0$ then define

$$(a + bi)/x = (a + bi)(1/x) = (a/x) + (b/x)i.$$

We sometimes write $\frac{p}{q}$ instead of p/q. Both expressions means the same thing.

A complex number $a + bi$ is *nonzero* if $a \neq 0$ or $b \neq 0$. Since

$$\det(a + bi) = \det \begin{bmatrix} a & -b \\ b & a \end{bmatrix} = a^2 + b^2,$$

which is only zero if $a = b = 0$, every nonzero complex number is invertible as a matrix.

Fact. This fact lets us divide complex numbers. If $a, b, c, d \in \mathbb{R}$ and $c + di \neq 0$ then define

$$(a+bi)/(c+di) = \left[\begin{array}{cc} a & -b \\ b & a \end{array}\right] \left[\begin{array}{cc} c & -d \\ d & c \end{array}\right]^{-1}.$$

We can write this more explicitly as

$$\begin{aligned} (a+bi)/(c+di) &= \left[\begin{array}{cc} a & -b \\ b & a \end{array}\right] \left[\begin{array}{cc} c & -d \\ d & c \end{array}\right]^{-1} \\ &= \frac{1}{c^2+d^2} \left[\begin{array}{cc} a & -b \\ b & a \end{array}\right] \left[\begin{array}{cc} c & d \\ -d & c \end{array}\right] \\ &= \frac{1}{c^2+d^2} \left[\begin{array}{cc} ac+bd & ad-bc \\ bc-ad & ac+bd \end{array}\right] = \frac{ac+bd}{c^2+d^2} + \frac{bc-ad}{c^2+d^2} i \in \mathbb{C}. \end{aligned}$$

The last formula is not so easy to remember.

It may be easier to divide complex numbers using the following method:

Example. We have $\frac{3-4i}{2+i} = \frac{(3-4i)(2-i)}{(2+i)(2-i)} = \frac{6-3i-8i+4i^2}{4-i^2} = \frac{6-11i-4}{5} = \frac{2-11i}{5} = \frac{2}{5} - \frac{11}{5} i.$

More generally, if $c + di \neq 0$ then we always have $\boxed{\frac{a+bi}{c+di} = \frac{(a+bi)(c-di)}{c^2+d^2}}$ since

$$\frac{a+bi}{c+di} = (a+bi)(c+di)^{-1} = \frac{1}{c^2+d^2}(a+bi)(c-di) = \frac{(a+bi)(c-di)}{c^2+d^2}.$$

The *complex conjugate* of $c + di$ is its matrix transpose, in other words, the complex number

$$\overline{c+di} = (c+di)^\top = c - di \in \mathbb{C}.$$

When $c + di$ is nonzero, the complex conjugate is related to the inverse by the identity

$$(c+di)^{-1} = \left[\begin{array}{cc} c & -d \\ d & c \end{array}\right]^{-1} = \frac{1}{c^2+d^2} \left[\begin{array}{cc} c & d \\ -d & c \end{array}\right] = \frac{1}{c^2+d^2} \cdot \overline{c+di}.$$

Since $x, y \in \mathbb{C}$ satisfy $xy = yx$ and $(xy)^\top = y^\top x^\top$ (since complex numbers are matrices), it follows that

$$\boxed{\overline{xy} = \overline{y} \cdot \overline{x} = \overline{x} \cdot \overline{y}.}$$

We can also add complex numbers $a + bi$ with real numbers c when $a, b, c \in \mathbb{R}$.

To do this, we set $c = c + 0i$ and define $(a + bi) + c = c + (a + bi) = (a + bi) + (c + 0i) = (a + c) + bi$.

Under this convention, we have

$$\begin{aligned} i^2 + 1 = (0+i)(0+i) + (1+0i) &= \left[\begin{array}{cc} 0 & -1 \\ 1 & 0 \end{array}\right] \left[\begin{array}{cc} 0 & -1 \\ 1 & 0 \end{array}\right] + \left[\begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array}\right] \\ &= \left[\begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array}\right] - \left[\begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array}\right] = \left[\begin{array}{cc} 0 & 0 \\ 0 & 0 \end{array}\right] = 0 + 0i = 0. \end{aligned}$$

Thus it makes sense to write $\boxed{i^2 = -1}$. In a similar way:

Theorem. Define the exponential function $\mathbb{C} \to \mathbb{C}$ by the convergent power series

$$e^x = 1 + \frac{1}{1}x + \frac{1}{1\cdot 2}x^2 + \frac{1}{1\cdot 2\cdot 3}x^3 + \frac{1}{1\cdot 2\cdot 3\cdot 4}x^4 + \dots$$

Then $e^1 = e = 2.71828\dots$ and $\boxed{e^{i\pi}+1=0}$.

Proof. We need two facts from calculus:

$$-1 = \cos\pi = 1 - \frac{1}{1\cdot 2}\pi^2 + \frac{1}{1\cdot 2\cdot 3\cdot 4}\pi^4 - \frac{1}{1\cdot 2\cdot 3\cdot 4\cdot 5\cdot 6}\pi^6 + \dots$$
$$0 = \sin\pi = \frac{1}{1}\pi - \frac{1}{1\cdot 2\cdot 3}\pi^3 + \frac{1}{1\cdot 2\cdot 3\cdot 4\cdot 5}\pi^5 - \frac{1}{1\cdot 2\cdot 3\cdot 4\cdot 5\cdot 6\cdot 7}\pi^7 + \dots$$

We have

$$i = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \quad i^2 = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}, \quad i^3 = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \quad \text{and} \quad i^0 = i^4 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Thus $i^{n+4} = i^n$ for all n.

Also, we have $(i\pi)^n = \pi^n i^n$. It follows that

$$e^{i\pi} = \begin{bmatrix} 1 - \frac{1}{1\cdot 2}\pi^2 + \frac{1}{1\cdot2\cdot3\cdot4}\pi^4 - \frac{1}{1\cdot2\cdot3\cdot4\cdot5\cdot6}\pi^6 + \dots & \frac{1}{1}\pi - \frac{1}{1\cdot2\cdot3}\pi^3 + \frac{1}{1\cdot2\cdot3\cdot4\cdot5}\pi^5 - \frac{1}{1\cdot2\cdot3\cdot4\cdot5\cdot6\cdot7}\pi^7 + \dots \\ \frac{1}{1}\pi - \frac{1}{1\cdot2\cdot3}\pi^3 + \frac{1}{1\cdot2\cdot3\cdot4\cdot5}\pi^5 - \frac{1}{1\cdot2\cdot3\cdot4\cdot5\cdot6\cdot7}\pi^7 + \dots & 1 - \frac{1}{1\cdot 2}\pi^2 + \frac{1}{1\cdot2\cdot3\cdot4}\pi^4 - \frac{1}{1\cdot2\cdot3\cdot4\cdot5\cdot6}\pi^6 + \dots \end{bmatrix}.$$

By our two facts, this is just $e^{i\pi} = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} = -1 + 0i$. Thus $e^{i\pi} + 1 = (-1+0i) + (1+0i) = 0$. $\square$

After a while, we tend to forget that complex numbers are 2×2 matrices and instead view the elements of $\mathbb{C}$ as formal expressions $a + bi$ where $a, b \in \mathbb{R}$ and i is a symbol that satisfies $i^2 = -1$.

We can add, subtract, and multiply such expressions just like polynomials, but substituting -1 for i^2. This convention gives the same operations as we saw above.

Moreover, this makes it clearer how to view $\mathbb{R}$ as a subset of $\mathbb{C}$, by setting $a = a + 0i$.

The *real part* of a complex number $a + bi \in \mathbb{C}$ is $\Re(a + bi) = a \in \mathbb{R}$.

The *imaginary part* of $a + bi \in \mathbb{C}$ is $\Im(a + bi) = b \in \mathbb{R}$.

Remark. It can be helpful to draw the complex number $a + bi \in \mathbb{C}$ as the vector $\begin{bmatrix} a \\ b \end{bmatrix} \in \mathbb{R}^2$.

The number $i(a + bi) = -b + ai \in \mathbb{C}$ then corresponds to the vector $\begin{bmatrix} -b \\ a \end{bmatrix} \in \mathbb{R}^2$, which is given by *rotating* $\begin{bmatrix} a \\ b \end{bmatrix}$ *ninety degrees counterclockwise.* (Try drawing this yourself.)

The main reason it is useful to work with complex numbers is the following theorem about polynomials.

Suppose $p(x) = a_n x^n + a_{n-1}x^{n-1} + \dots + a_1 x + a_0$ is a polynomial with coefficients $a_0, a_1, \dots, a_n \in \mathbb{C}$.

Assume $a_n \neq 0$ so that $p(x)$ has *degree* n.

Even though we think of complex numbers are 2×2 matrices, this expression for $p(x)$ still makes sense for $x \in \mathbb{C}$: if we plug in any complex number for x then $a_n x^n + a_{n-1}x^{n-1} + \dots + a_1 x + a_0 \in \mathbb{C}$.

Theorem (Fundamental theorem of algebra)**.** Define $p(x)$ as above. There are n (not necessarily distinct) complex numbers $r_1, r_2, \dots, r_n \in \mathbb{C}$ such that $p(x) = a_n(x - r_1)(x - r_2) \cdots (x - r_n)$.

One calls the numbers $r_1, r_2, \dots, r_n$ the *roots* of $p(x)$.

A root r has *multiplicity* m if exactly m of the numbers $r_1, r_2, \dots, r_n$ are equal to r.

The use of complex numbers in this theorem is essential. The statement fails if we use $\mathbb{R}$ instead of $\mathbb{C}$.

Example: if $p(x) = x^2 + 1$ then there **do not exist** real numbers $r_1, r_2 \in \mathbb{R}$ with $p(x) = (x - r_1)(x - r_2)$.

However, we do have $x^2 + 1 = (x - i)(x + i)$.

3 Vocabulary

Keywords from today's lecture:

1. **Complex number**.

   We define a complex number to be either

   - A matrix $a + bi = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$ where $a, b \in \mathbb{R}$ and $i = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$.
   - A formal expression "$a + bi$" where $a, b \in \mathbb{R}$ and i is a symbol that has $i^2 = -1$.

   The first definition makes it clear how to add, subtract, multiply, and divide complex numbers (use matrix operations). The second definition is secretly just a way of abbreviating the first definition.

   The set of complex numbers is denoted $\mathbb{C}$.

   Example:

   $1 + 2i$ corresponds to $\begin{bmatrix} 1 & -2 \\ 2 & 1 \end{bmatrix}$.

   $(1 + 2i) + (2 + 3i) = 3 + 5i$ corresponds to $\begin{bmatrix} 1 & -2 \\ 2 & 1 \end{bmatrix} + \begin{bmatrix} 2 & -3 \\ 3 & 2 \end{bmatrix} = \begin{bmatrix} 3 & -5 \\ 5 & 3 \end{bmatrix}$.

   $(1 + 2i)(2 + 3i) = -4 + 7i$ corresponds to $\begin{bmatrix} 1 & -2 \\ 2 & 1 \end{bmatrix}\begin{bmatrix} 2 & -3 \\ 3 & 2 \end{bmatrix} = \begin{bmatrix} -4 & -6 \\ 7 & -4 \end{bmatrix}$.

   $(1 + 2i)^{-1} = \frac{1}{5} - \frac{2}{5}i$ corresponds to $\begin{bmatrix} 1 & -2 \\ 2 & 1 \end{bmatrix}^{-1} = \frac{1}{5}\begin{bmatrix} 1 & 2 \\ -2 & 1 \end{bmatrix}$.

2. **Complex conjugation**.

   If $a, b \in \mathbb{R}$ then *complex conjugate* of $a + bi \in \mathbb{C}$ is $\overline{a + bi} = a - bi \in \mathbb{C}$.

   If $y, z \in \mathbb{C}$ then $\overline{y + z} = \overline{y} + \overline{z}$ and $\overline{yz} = \overline{y} \cdot \overline{z}$ and $\overline{y^{-1}} = \overline{y}^{-1}$.

3. **Fundamental theorem of algebra**.

   Any polynomial
   $$f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$$
   with coefficients $a_0, a_1, \dots, a_n \in \mathbb{C}$ and $a_n \neq 0$ can be factored as
   $$f(x) = a_n(x - \lambda_1)(x - \lambda_2)\cdots(x - \lambda_n)$$
   for some not necessarily distinct complex numbers $\lambda_1, \lambda_2, \dots, \lambda_n \in \mathbb{C}$.