Let us continue our list from last time with two more types of matrices:

Session 23
May 4, 2020

c) Unitary matrices

Recall that normal matrices A can be diagonalized with orthonormal eigenvectors $\vec{v}_1, \dots, \vec{v}_n$, i.e.,

$$\langle \vec{v}_i, \vec{v}_j \rangle = \underbrace{\delta_{ij} := \begin{cases} 0 & \text{for } i \neq j \\ 1 & \text{for } i = j \end{cases}}_{\text{this is called "Kronecker delta"}}.$$

Diagonalization means $\Lambda = V^{-1} A V$, where

$$\Lambda = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \text{ with } \lambda_1, \dots, \lambda_n \text{ the eigenvalues, and } V = \overbrace{\left(\vec{v}_1 \,\middle|\, \vec{v}_2 \,\middle|\, \dots \,\middle|\, \vec{v}_n \right)}^{\text{matrix with eigenvectors as columns}}.$$

Let us compute $V^{\dagger}$: We find

$$V^{\dagger} = \left.\begin{pmatrix} \overline{\vec{v}_1}^T \\ \overline{\vec{v}_2}^T \\ \vdots \\ \overline{\vec{v}_n}^T \end{pmatrix}\right\} \text{matrix with row vectors } \vec{v}_1^T = ((\vec{v}_1)_1, \dots, (\vec{v}_1)_n) \text{ that are also complex conjugated}$$

Then we find

$$V^{\dagger} V = \begin{pmatrix} \overline{\vec{v}_1}^T \\ \vdots \\ \overline{\vec{v}_n}^T \end{pmatrix} \left(\vec{v}_1 \,\middle|\, \dots \,\middle|\, \vec{v}_n \right) = \begin{pmatrix} \langle \vec{v}_1, \vec{v}_1 \rangle & \langle \vec{v}_1, \vec{v}_2 \rangle & \dots & \langle \vec{v}_1, \vec{v}_n \rangle \\ \langle \vec{v}_2, \vec{v}_1 \rangle & \langle \vec{v}_2, \vec{v}_2 \rangle & & \vdots \\ \vdots & & \ddots & \vdots \\ \langle \vec{v}_n, \vec{v}_1 \rangle & \dots & & \langle \vec{v}_n, \vec{v}_n \rangle \end{pmatrix}$$

because $\vec{v}_1, \dots, \vec{v}_n$ are orthonormal!

$$= \begin{pmatrix} 1 & & 0 \\ & 1 & \\ 0 & & \ddots \\ & & & 1 \end{pmatrix}$$

$\Rightarrow V^{\dagger} V = \overbrace{I}^{\text{identity matrix}}$. In the same way we find $V V^{\dagger} = I$. So $V^{\dagger} = V^{-1}$.

So this is a very special type of matrix, where the inverse is just the Hermitian adjoint, which can be very easily computed.

This makes this class of matrices so interesting that it has a name.

$\boxed{\text{Definition: An } n\times n \text{ matrix } \underbrace{U}_{} \text{ (possibly with complex entries) is called unitary if } U^{-1}=U^{+}.}$

Let us always try to call unitary matrices U

Important properties of unitary matrices U:

- As we saw above, unitary matrices are exactly those that diagonalize normal matrices A: $\Lambda = U^{+}AU$.
- They preserve lengths (such maps are also called isometries):

$$\boxed{|U\vec{x}|^2 = \langle U\vec{x}, U\vec{x}\rangle = \langle \vec{x}, \underbrace{U^{+}U}_{=I}\vec{x}\rangle = \langle \vec{x},\vec{x}\rangle = |\vec{x}|^2.}$$

And they preserve angles: $\langle U\vec{x}, U\vec{y}\rangle = \langle \vec{x}, U^{+}U\vec{y}\rangle = \langle \vec{x},\vec{y}\rangle$

This is in fact equivalent to $U^{-1}=U^{+}$, and often taken as the definition of unitary.

- U is in particular normal, since $U^{+}U = I = UU^{+}$. Thus, U itself is also diagonalizable.
- $1 = \det I = \det(U^{+}U) = (\det U^{+})(\det U) \overset{\det U^{T} = \det U}{=} \overline{(\det U)}(\det U) = |\det U|^2$, so $\boxed{|\det U| = 1}$.
- What about the eigenvalues of U? If $\vec{x}$ is an eigenvector with eigenvalue λ, we find

$$|\lambda|^2|\vec{x}|^2 = |\lambda\vec{x}|^2 = \langle \lambda\vec{x}, \lambda\vec{x}\rangle \overset{U\vec{x}=\lambda\vec{x}}{=} \langle U\vec{x}, U\vec{x}\rangle = \langle \vec{x}, \underbrace{U^{+}U}_{=I}\vec{x}\rangle \overset{U\text{ unitary}}{=} \langle \vec{x},\vec{x}\rangle = |\vec{x}|^2 \neq 0.$$

because eigenvectors are non-zero

Thus, $|\lambda|^2 = 1$, i.e., $\boxed{|\lambda| = 1}$. (Thus, if you remember complex numbers and Euler's formula, $\lambda = e^{i\varphi}$ for some $\varphi \in [0, 2\pi)$.)

Again, an application from my research field, quantum mechanics. There, particles are described by a so-called wave function $\vec{\Psi}$ that evolves in time, i.e., we have $\vec{\Psi}(t)$, where t = time. $\vec{\Psi}$ is an element of some vector space (which is usually infinite-dimensional). Now $|\vec{\Psi}|^2$ is the total probability to find a particle somewhere, so it should always be one. Thus, the time evolution of $\vec{\Psi}(t)$ is described by a unitary matrix: $\vec{\Psi}(t) = U(t)\vec{\Psi}(0)$, where $\vec{\Psi}(0)$

is the initial condition. But $\vec{\psi}(t)$ should also be solution to a linear first-order equation. (Linearity is the reason for Schrödinger's cat!) So $i\frac{d}{dt}\vec{\psi}(t) = H\vec{\psi}(t)$, where H is some matrix; this is called the Schrödinger equation. The solution should be $\underbrace{\vec{\psi}(t)}_{\text{vector}} = \underbrace{e^{-iHt}}_{\text{matrix}}\underbrace{\vec{\psi}(0)}_{\text{vector}}$.
So we should find that $U(t) = e^{-iHt}$. The question is then what type of matrix H makes the matrix e^{-iHt} unitary? The answer is that these are the Hermitian matrices:

A matrix U is unitary if and only if it can be written as e^{-iH} with Hermitian H.

More examples in the exercises.

D) Orthogonal matrices

Again, we might want to specialize to matrices with real entries, i.e., when $A^\dagger = A^T$.

Definition: A real $n\times n$ matrix Q is called orthogonal if $Q^{-1} = Q^T$.

So orthogonal matrices are special cases of unitary matrices (similar to how symmetric matrices were special cases of Hermitian ones). So the remarks from C) still apply.

In particular we have that $\det Q = \pm 1$ and all eigenvalues equal ± 1.

Note:

- One can show that orthogonal matrices exactly describe rotations and reflections. This is one of the reasons that make them so interesting.

- They will also be important for matrix decompositions, as we will see shortly.
- Any real symmetric matrix A is diagonalized by an orthogonal matrix, i.e., $\underbrace{\Lambda}_{\text{diagonal matrix}} = Q^T A Q$.

This concludes our discussion of eigenvalues/eigenvectors. Let me make one final remark.

We saw that diagonalizable matrices are very nice, because $A = V \Lambda V^{-1}$. One question that we left open is what to do about matrices that cannot be diagonalized.

Is there some transformation that makes such matrices "almost diagonal"? There is, and it is called Jordan normal form.

The general result is that any matrix A can be written as $A = VJV^{-1}$, where J is of the form

$$J = \begin{pmatrix} \boxed{\begin{matrix} \lambda_1 & 1 & 0 \\ & \ddots & 1 \\ 0 & & \lambda_1 \end{matrix}} & & & 0 \\ & \boxed{\begin{matrix} \lambda_2 & 1 & 0 \\ & \ddots & 1 \\ 0 & & \lambda_2 \end{matrix}} & & \\ & & \ddots & \\ 0 & & & \boxed{\begin{matrix} \lambda_k & 1 & 0 \\ & \ddots & 1 \\ 0 & & \lambda_k \end{matrix}} \end{pmatrix}$$

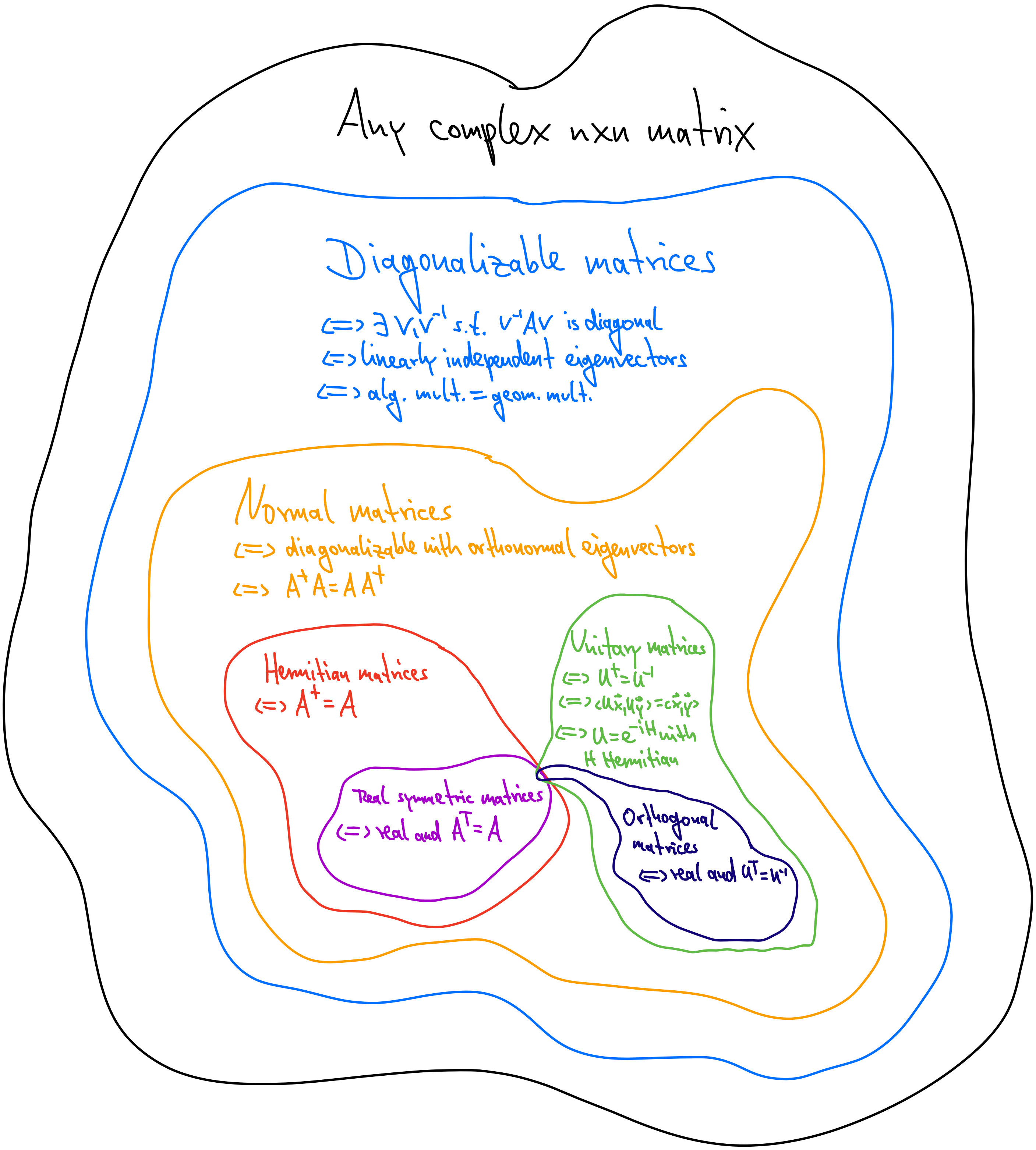
, i.e., with all eigenvalues with their multiplicity on the diagonal and possibly 1's on the diagonal right above.

With such J's one can still do a lot, e.g., powers of A can still be computed relatively easily.

But a more thorough discussion of the Jordan normal form is beyond the scope of this class.

Summary of classes of matrices:

(Note: any $n \times n$ matrix has a Jordan normal form.)



Note: The overlap of Hermitian and unitary matrices are exactly the Hermitian matrices with eigenvalues ± 1.

4.7 Matrix Decompositions

In this chapter we consider useful ways how a matrix can be written as a product of other, simpler matrices. (We already saw one example: $A = V^{-1} \Lambda V$, with Λ diagonal.) These are called "matrix decompositions", or "matrix factorizations".

4.7.1 LU Decomposition

Let us consider a system of n linear equations with n unknowns:

$$A\vec{x} = \vec{b} \quad \text{, where } A \text{ is an } n{\times}n \text{ matrix and } \vec{b} \in \mathbb{R}^n.$$

Let us assume that A is invertible ($\det A \neq 0$) i.e., there is a unique solution $\vec{x} = A^{-1}\vec{b}$.

Last semester you saw how to find solutions using Gaussian elimination. This could be formulated using type I, II, III row operations, which could be represented by matrix multiplication with invertible matrices T_1, T_2, T_3.

Recall:

- $$T_1 = \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & 0 & \cdots & 1 & & \\ & & & \vdots & 1 \ddots 1 & \vdots & & \\ & & & 1 & \cdots & 0 & & \\ & & & & & & 1 & \\ & & & & & & & \ddots \\ & & & & & & & & 1 \end{pmatrix} = \text{interchange two rows,}$$

- $$T_2 = \begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 & & \\ & & & \lambda & \\ & & & & 1 \\ & & & & & \ddots \\ & & & & & & 1 \end{pmatrix} = \text{multiply a row by a constant } \lambda,$$

  - λ is on the lower left when a row above is added to a lower row
  - λ is on the upper right if a row is added to a higher one

- $$T_3 = \begin{array}{c} \begin{pmatrix} 1 & & & \\ & \ddots & & \\ & \lambda & \ddots & \\ & & & 1 \end{pmatrix} \leftarrow \text{row } j \\ \begin{array}{l} \quad\ \ \uparrow \\ \ \text{column } i \end{array} \end{array} = \text{add a multiple of one row to another } (\lambda \cdot \text{row } i + \text{row } j)$$

e.g., $\begin{pmatrix} 1 & 0 \\ \lambda & 1 \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a & b \\ \lambda a + c & \lambda b + d \end{pmatrix}$, $\rightarrow$ $\lambda \cdot$ first row added to second row

$\begin{pmatrix} 1 & \lambda \\ 0 & 1 \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a + \lambda c & b + \lambda d \\ c & d \end{pmatrix}$. $\rightarrow$ $\lambda \cdot$ second row added to first row

Then bringing the system $A\vec{x} = \vec{b}$ into $\overbrace{\text{upper triangular form}}^{\begin{pmatrix} \ddots & \cdots \\ 0 & \ddots \end{pmatrix}\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \vdots \end{pmatrix}}$ ("forward elimination") can be written as: $\underbrace{M_N M_{N-1} \cdots M_1}_{=:M} A \vec{x} = M_N M_{N-1} \cdots M_1 \vec{b}$, where $\underbrace{\text{each } M_i \in \{T_1, T_2, T_3\}}_{\text{each } M_i \text{ is one of the operations above}}$.

Writing $M = M_N M_{N-1} \cdots M_1$, we have $\overbrace{\underbrace{MA}_{\text{in components, upper triangular means } (MA)_{ij} = 0 \text{ for } i > j}}^{\text{upper triangular}} \vec{x} = M\vec{b}$

If we forget about $\vec{x}$ and $\vec{b}$ for a moment, we have shown

$$MA = \text{upper triangular} = U = \begin{pmatrix} u_{11} & u_{12} & \cdots & \cdots & u_{1n} \\ 0 & u_{22} & \ddots & & \vdots \\ 0 & 0 & u_{33} & \ddots & \vdots \\ \vdots & & \ddots & \ddots & \\ 0 & \cdots & \cdots & 0 & u_{nn} \end{pmatrix}.$$

If A is non-singular, i.e., $\det A \neq 0$, then all diagonal entries $u_{ii} \neq 0$.

But all type I, II, III operations can be reversed, meaning $M_N, \ldots, M_1$ and thus M are invertible. So:

$$U = M_N \cdots M_1 A \iff \overbrace{M_1^{-1} \cdots M_N^{-1}}^{M^{-1}} U = A \quad (\iff M^{-1} U = A)$$

Question: Can we say anything about the form of M^{-1}?

Yes, but there is one subtlety concerning interchanging rows. Suppose the rows of A are already ordered in such a way that only type III operations are needed to bring A into upper triangular form.

E.g., for $A = \begin{pmatrix} 0 & a & b \\ 0 & c & d \\ e & 0 & 0 \end{pmatrix}$, $a,b,c,d,e \neq 0$, we first need to interchange rows 1 and 3; then we can bring the matrix into upper triangular form.

Recall that type III matrices are of the form $T_3 = \begin{pmatrix} 1 & & \\ & \ddots & \\ \lambda & & 1 \end{pmatrix}$.

1's on the diagonal, λ in row j and column i, all other entries 0.

Now, $T_3^{-1} = \begin{pmatrix} 1 & & \\ & \ddots & \\ -\lambda & & 1 \end{pmatrix}$ (can be easily checked).

So T_3^{-1} is lower triangular. And the product of lower triangular matrices remains lower triangular.

So M^{-1} from above is lower triangular!

This is the LU decomposition: $A = LU$, with $U =$ upper triangular, $L =$ lower triangular

In order to formulate this as a theorem, we might have to reorder the rows first.

Theorem (LU Decomposition): Let A be an invertible $n \times n$ matrix. Then we can decompose $PA = LU$, where L is a lower triangular matrix, U is an upper triangular matrix, and P is a matrix that permutes the rows of A.

Applications and examples: next time