Example 134. Consider the following system of (second-order) initial value problems:

$$\begin{array}{l} y_1'' = 2y_1' - 3y_2' + 7y_2 \\ y_2'' = 4y_1' + y_2' - 5y_1 \end{array} \qquad y_1(0) = 2,\ \ y_1'(0) = 3,\ \ y_2(0) = -1,\ \ y_2'(0) = 1$$

Write it as a first-order initial value problem in the form $\boldsymbol{y}' = A\boldsymbol{y}$, $\boldsymbol{y}(0) = \boldsymbol{y}_0$.

Solution. Introduce $y_3 = y_1'$ and $y_4 = y_2'$. Then, the given system translates into

$$\boldsymbol{y}' = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 7 & 2 & -3 \\ -5 & 0 & 4 & 1 \end{bmatrix} \boldsymbol{y}, \quad \boldsymbol{y}(0) = \begin{bmatrix} 2 \\ -1 \\ 3 \\ 1 \end{bmatrix}.$$

Review. Jordan normal form

Example 135.

(a) What are the possible Jordan normal forms of a 3×3 matrix with eigenvalues $3, 3, 3$?

(b) What are the possible Jordan normal forms of a 4×4 matrix with eigenvalues $3, 3, 3, 3$?

(c) What if the matrix is 5×5 and has eigenvalues $4, 4, 3, 3, 3$?

Solution.

(a) $\begin{bmatrix} 3 & & \\ & 3 & \\ & & 3 \end{bmatrix}, \begin{bmatrix} 3 & & \\ & 3 & 1 \\ & & 3 \end{bmatrix}, \begin{bmatrix} 3 & 1 & \\ & 3 & 1 \\ & & 3 \end{bmatrix}$

The dimension of the 3-eigenspace equals the number of Jordan blocks: 3, 2, 1, respectively.

Comment. Note that, say, $\begin{bmatrix} 3 & 1 & \\ & 3 & \\ & & 3 \end{bmatrix}$ is equivalent to $\begin{bmatrix} 3 & & \\ & 3 & 1 \\ & & 3 \end{bmatrix}$ because the ordering of the diagonal blocks does not matter (as you known from diagonalization).

(b) Now, there are 5 possibilities:

$\begin{bmatrix} 3 & & & \\ & 3 & & \\ & & 3 & \\ & & & 3 \end{bmatrix}, \begin{bmatrix} 3 & & & \\ & 3 & & \\ & & 3 & 1 \\ & & & 3 \end{bmatrix}, \begin{bmatrix} 3 & 1 & & \\ & 3 & & \\ & & 3 & 1 \\ & & & 3 \end{bmatrix}, \begin{bmatrix} 3 & & & \\ & 3 & 1 & \\ & & 3 & 1 \\ & & & 3 \end{bmatrix}, \begin{bmatrix} 3 & 1 & & \\ & 3 & 1 & \\ & & 3 & 1 \\ & & & 3 \end{bmatrix}$

The dimension of the 3-eigenspace equals the number of Jordan blocks: 4, 3, 2, 2, 1, respectively.

(c) $\begin{bmatrix} 3 & & & & \\ & 3 & & & \\ & & 3 & & \\ & & & 4 & \\ & & & & 4 \end{bmatrix}, \begin{bmatrix} 3 & & & & \\ & 3 & & & \\ & & 3 & & \\ & & & 4 & 1 \\ & & & & 4 \end{bmatrix}, \begin{bmatrix} 3 & & & & \\ & 3 & 1 & & \\ & & 3 & & \\ & & & 4 & \\ & & & & 4 \end{bmatrix}, \begin{bmatrix} 3 & & & & \\ & 3 & 1 & & \\ & & 3 & & \\ & & & 4 & 1 \\ & & & & 4 \end{bmatrix}, \begin{bmatrix} 3 & 1 & & & \\ & 3 & 1 & & \\ & & 3 & & \\ & & & 4 & \\ & & & & 4 \end{bmatrix}, \begin{bmatrix} 3 & 1 & & & \\ & 3 & 1 & & \\ & & 3 & & \\ & & & 4 & 1 \\ & & & & 4 \end{bmatrix}$

Note that this is just all possible (namely, 3) Jordan normal forms of a 3×3 matrix with eigenvalues $3, 3, 3$ combined with all possible (namely, 2) Jordan normal forms of a 2×2 matrix with eigenvalues $4, 4$. In total, that makes $3 \cdot 2 = 6$ possibilities.

Comment. Let $p(n)$ be the number of inequivalent Jordan normal forms of an $n \times n$ matrix with a single eigenvalue, n times repeated. We have seen that $p(2) = 2$, $p(3) = 3$, $p(4) = 5$. Note that $p(n)$ is equal to the number of ways of writing n as an ordered sum of positive integers: for instance, $p(4) = 5$ because $4 = 3 + 1 = 2 + 2 = 2 + 1 + 1 = 1 + 1 + 1 + 1$.

$p(n)$ is referred to as the **partition function** and, surprisingly, is a remarkably interesting mathematical object.

https://en.wikipedia.org/wiki/Partition_function_(number_theory)

Armin Straub
straub@southalabama.edu

Example 136. (summary of small cases)

(a) There are 2 possible Jordan normal forms of a 2×2 matrix with eigenvalues λ, λ.

Namely. $\begin{bmatrix} \lambda & \\ & \lambda \end{bmatrix}, \begin{bmatrix} \lambda & 1 \\ & \lambda \end{bmatrix}$

(b) There are 3 possible Jordan normal forms of a 3×3 matrix with eigenvalues $\lambda, \lambda, \lambda$.

Namely. $\begin{bmatrix} \lambda & & \\ & \lambda & \\ & & \lambda \end{bmatrix}, \begin{bmatrix} \lambda & & \\ & \lambda & 1 \\ & & \lambda \end{bmatrix}, \begin{bmatrix} \lambda & 1 & \\ & \lambda & 1 \\ & & \lambda \end{bmatrix}$

(c) There are 5 possible Jordan normal forms of a 4×4 matrix with eigenvalues $\lambda, \lambda, \lambda, \lambda$.

Namely. $\begin{bmatrix} \lambda & & & \\ & \lambda & & \\ & & \lambda & \\ & & & \lambda \end{bmatrix}, \begin{bmatrix} \lambda & & & \\ & \lambda & & \\ & & \lambda & 1 \\ & & & \lambda \end{bmatrix}, \begin{bmatrix} \lambda & 1 & & \\ & \lambda & & \\ & & \lambda & 1 \\ & & & \lambda \end{bmatrix}, \begin{bmatrix} \lambda & & & \\ & \lambda & 1 & \\ & & \lambda & 1 \\ & & & \lambda \end{bmatrix}, \begin{bmatrix} \lambda & 1 & & \\ & \lambda & 1 & \\ & & \lambda & 1 \\ & & & \lambda \end{bmatrix}$

Example 137. What are the possible Jordan normal forms of a 6×6 matrix with eigenvalues $3, 3, 7, 7, 7, 7$?

Solution. There are $2\cdot 5 = 10$ possible Jordan normal forms for such a matrix:

$$\begin{bmatrix} 3 &&&&& \\ & 3 &&&& \\ && 7 &&& \\ &&& 7 && \\ &&&& 7 & \\ &&&&& 7 \end{bmatrix}, \begin{bmatrix} 3 &&&&& \\ & 3 &&&& \\ && 7 & 1 && \\ &&& 7 && \\ &&&& 7 & \\ &&&&& 7 \end{bmatrix}, \begin{bmatrix} 3 &&&&& \\ & 3 &&&& \\ && 7 & 1 && \\ &&& 7 && \\ &&&& 7 & 1 \\ &&&&& 7 \end{bmatrix}, \begin{bmatrix} 3 &&&&& \\ & 3 &&&& \\ && 7 & 1 && \\ &&& 7 & 1 & \\ &&&& 7 & \\ &&&&& 7 \end{bmatrix}, \begin{bmatrix} 3 &&&&& \\ & 3 &&&& \\ && 7 & 1 && \\ &&& 7 & 1 & \\ &&&& 7 & 1 \\ &&&&& 7 \end{bmatrix},$$

$$\begin{bmatrix} 3 & 1 &&&& \\ & 3 &&&& \\ && 7 &&& \\ &&& 7 && \\ &&&& 7 & \\ &&&&& 7 \end{bmatrix}, \begin{bmatrix} 3 & 1 &&&& \\ & 3 &&&& \\ && 7 & 1 && \\ &&& 7 && \\ &&&& 7 & \\ &&&&& 7 \end{bmatrix}, \begin{bmatrix} 3 & 1 &&&& \\ & 3 &&&& \\ && 7 & 1 && \\ &&& 7 && \\ &&&& 7 & 1 \\ &&&&& 7 \end{bmatrix}, \begin{bmatrix} 3 & 1 &&&& \\ & 3 &&&& \\ && 7 & 1 && \\ &&& 7 & 1 & \\ &&&& 7 & \\ &&&&& 7 \end{bmatrix}, \begin{bmatrix} 3 & 1 &&&& \\ & 3 &&&& \\ && 7 & 1 && \\ &&& 7 & 1 & \\ &&&& 7 & 1 \\ &&&&& 7 \end{bmatrix}$$

Example 138. How many different Jordan normal forms are there in the following cases?

(a) A 8×8 matrix with eigenvalues $1, 1, 2, 2, 2, 4, 4, 4$?

(b) A 11×11 matrix with eigenvalues $1, 1, 1, 2, 2, 2, 2, 4, 4, 4, 4$?

Solution.

(a) $2\cdot 3\cdot 3 = 18$ possible Jordan normal forms

(b) $3\cdot 5\cdot 5 = 75$ possible Jordan normal forms

Review.

- Let A be $n \times n$. The matrix exponential is

$$e^A = I + A + \frac{1}{2!}A^2 + \frac{1}{3!}A^3 + \cdots$$

  Then, $\frac{\mathrm{d}}{\mathrm{d}t}e^{At} = Ae^{At}$.

  Why? $\frac{\mathrm{d}}{\mathrm{d}t}e^{At} = \frac{\mathrm{d}}{\mathrm{d}t}\left(I + At + \frac{1}{2!}A^2t^2 + \frac{1}{3!}A^3t^3 + \cdots\right) = A + \frac{1}{1!}A^2t + \frac{1}{2!}A^3t^2 + \cdots = Ae^{At}$

- If $A = PDP^{-1}$, then $e^A = Pe^DP^{-1}$.

- The solution to $\boldsymbol{y}' = A\boldsymbol{y}$, $\boldsymbol{y}(0) = \boldsymbol{y}_0$ is $\boldsymbol{y}(t) = e^{At}\boldsymbol{y}_0$.

  Why? Because $\boldsymbol{y}'(t) = Ae^{At}\boldsymbol{y}_0 = A\boldsymbol{y}(t)$ and $\boldsymbol{y}(0) = e^{0A}\boldsymbol{y}_0 = \boldsymbol{y}_0$.

Example 139. The matrix exponential shares many other properties of the usual exponential:

- $e^Ae^B = e^{A+B} = e^Be^A$ if $AB = BA$

  Why the condition $\boldsymbol{AB} = \boldsymbol{BA}$? By the Taylor series, $e^{A+B} = I + (A+B) + \frac{(A+B)^2}{2!} + \ldots$ In order to simplify that to

$$e^Ae^B = \left(I + A + \frac{A^2}{2!} + \ldots\right)\left(I + B + \frac{B^2}{2!} + \ldots\right),$$

  we need that $(A+B)^2 = A^2 + AB + BA + B^2$ is the same as $A^2 + 2AB + B^2$. That's only the case if $AB = BA$.

- e^A is invertible and $(e^A)^{-1} = e^{-A}$

  Why? That actually follows from the previous property.

Example 140. Compute e^{At} for $A = \begin{bmatrix} 2 & 1 \\ & 2 \end{bmatrix}$.

Solution.

- Write $A = \begin{bmatrix} 2 & 1 \\ & 2 \end{bmatrix} = 2I + N$ with $N = \begin{bmatrix} 0 & 1 \\ & 0 \end{bmatrix}$. Note that $2I$ and N commute.

  Hence, $e^{At} = e^{2It + Nt} = e^{2It}e^{Nt}$.

- Note that $N^2 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$. Hence, $e^{Nt} = I + Nt + \frac{t^2}{2!}N^2 + \ldots = I + Nt = \begin{bmatrix} 1 & t \\ & 1 \end{bmatrix}$.

- Combined, $e^{At} = e^{2It + Nt} = e^{2It}e^{Nt} = \begin{bmatrix} e^{2t} & \\ & e^{2t} \end{bmatrix}\begin{bmatrix} 1 & t \\ & 1 \end{bmatrix} = \begin{bmatrix} e^{2t} & te^{2t} \\ & e^{2t} \end{bmatrix}$.

Advanced. Can you show that $A^n = \begin{bmatrix} 2^n & n2^{n-1} \\ & 2^n \end{bmatrix}$?

Armin Straub
straub@southalabama.edu

Example 141. Solve the differential equation

$$\boldsymbol{y}' = \underbrace{\begin{bmatrix} 2 & 1 \\ & 2 \end{bmatrix}}_{A} \boldsymbol{y}, \qquad \boldsymbol{y}(0) = \underbrace{\begin{bmatrix} -1 \\ 1 \end{bmatrix}}_{\boldsymbol{y}_0}$$

Solution. Repeating the work in the previous example, the solution to the differential equation is

$$\begin{aligned} \boldsymbol{y}(t) &= e^{At}\boldsymbol{y}_0 \\ &= e^{2It+Nt}\boldsymbol{y}_0 \qquad \text{with } N = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \\ &= e^{2It}e^{Nt}\boldsymbol{y}_0 \qquad \text{(because } 2It \text{ and } Nt \text{ commute)} \\ &= \begin{bmatrix} e^{2t} & \\ & e^{2t} \end{bmatrix}\left(1 + Nt + \frac{1}{2}(Nt)^2 + \frac{1}{3!}(Nt)^3 + \ldots\right)\boldsymbol{y}_0 \\ &= \begin{bmatrix} e^{2t} & \\ & e^{2t} \end{bmatrix}(1 + Nt)\boldsymbol{y}_0 \qquad \text{(because } N^2 = \mathbf{0}) \\ &= \begin{bmatrix} e^{2t} & \\ & e^{2t} \end{bmatrix}\begin{bmatrix} 1 & t \\ & 1 \end{bmatrix}\begin{bmatrix} -1 \\ 1 \end{bmatrix} \\ &= \begin{bmatrix} e^{2t} & \\ & e^{2t} \end{bmatrix}\begin{bmatrix} t-1 \\ 1 \end{bmatrix} = \begin{bmatrix} (t-1)e^{2t} \\ e^{2t} \end{bmatrix}. \end{aligned}$$

Check. We should verify that $y_1 = (t-1)e^{2t}$ and $y_2 = e^{2t}$ satisfy $y_1' = 2y_1 + y_2$ and $y_2' = 2y_2$.
Indeed, $y_1' = e^{2t} + (t-1)2e^{2t}$ equals $2y_1 + y_2 = 2(t-1)e^{2t} + e^{2t}$.

Comment. For applications, having solutions like $te^{\lambda t}$ or $t\cos(\lambda t)$ (when the eigenvalues are imaginary) is connected to the phenomenon of **resonance**, which you may have already seen.

Important comment. Note that we can immediately see from the solution that the original matrix A is not diagonalizable: there is a term te^{2t}, whereas in the diagonalizable case we would only see exponentials like e^{2t} by themselves.

In our upcoming discussion of complex numbers we will see that e^{2it} (here, $2i$ would be the eigenvalue) can be rewritten in terms of $\cos(2t)$ and $\sin(2t)$. Both of these are periodic and bounded, so that the same is true for every linear combination.

In that case, if the eigenvalue $2i$ was repeated in such a way that the matrix A is not diagonalizable, then we would get the functions $t\cos(2t)$ and $t\sin(2t)$ in our solutions. These, however, are not bounded! This phenomenon (getting solutions that are unbounded under the right/wrong circumstances) is called **resonance**.

https://en.wikipedia.org/wiki/Resonance

Understanding when resonance occurs is of crucial importance for practical applications.

straub@southalabama.edu

Example 142. Solve the IVP $\boldsymbol{y}' = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \boldsymbol{y}$ with $\boldsymbol{y}(0) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$.

Solution. Recall that the solution to $\boldsymbol{y}' = A\boldsymbol{y}$, $\boldsymbol{y}(0) = \boldsymbol{y}_0$ is $\boldsymbol{y} = e^{At}\boldsymbol{y}_0$.

- We first diagonalize $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$.
  - $\begin{vmatrix} -\lambda & 1 \\ 1 & -\lambda \end{vmatrix} = \lambda^2 - 1$, so the eigenvalues are ± 1.
  - The 1-eigenspace $\operatorname{null}\left(\begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}\right)$ has basis $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$.
  - The -1-eigenspace $\operatorname{null}\left(\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\right)$ has basis $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$.
  - Hence, $A = PDP^{-1}$ with $P = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$ and $D = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$.
- Compute the solution $\boldsymbol{y} = e^{At}\boldsymbol{y}_0$:

$$\begin{aligned} \boldsymbol{y} = e^{At}\boldsymbol{y}_0 &= Pe^{Dt}P^{-1}\boldsymbol{y}_0 \\ &= \underbrace{\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} e^t & 0 \\ 0 & e^{-t} \end{bmatrix}}_{= \begin{bmatrix} e^t & -e^{-t} \\ e^t & e^{-t} \end{bmatrix}} \underbrace{\frac{1}{2}\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix}}_{= \frac{1}{2}\begin{bmatrix} 1 \\ -1 \end{bmatrix}} = \frac{1}{2}\begin{bmatrix} e^t + e^{-t} \\ e^t - e^{-t} \end{bmatrix} \end{aligned}$$

Check. Indeed, $y_1 = \frac{1}{2}(e^t + e^{-t})$ and $y_2 = \frac{1}{2}(e^t - e^{-t})$ satisfy the system of differential equations $y_1' = y_2$ and $y_2' = y_1$ as well as the initial conditions $y_1(0) = 1$, $y_2(0) = 0$.

Comment. You have actually met these functions in Calculus! $y_1 = \cosh(t)$ and $y_2 = \sinh(t)$. Check out the next example for the connection to $\cos(t)$ and $\sin(t)$.

Example 143.

(a) Solve the IVP $\boldsymbol{y}' = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \boldsymbol{y}$ with $\boldsymbol{y}(0) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$.

(b) Show that $\boldsymbol{y} = \begin{bmatrix} \cos(t) \\ \sin(t) \end{bmatrix}$ solves the same IVP. What do you conclude?

Solution.

(a) $A = PDP^{-1}$ with $P = \begin{bmatrix} i & -i \\ 1 & 1 \end{bmatrix}$, $D = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}$.

The system is therefore solved by:

$$\begin{aligned} \boldsymbol{y}(t) &= Pe^{Dt}P^{-1}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} i & -i \\ 1 & 1 \end{bmatrix} \begin{bmatrix} e^{it} & \\ & e^{-it} \end{bmatrix} \frac{1}{2i}\begin{bmatrix} 1 & i \\ -1 & i \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \\ &= \frac{1}{2i}\begin{bmatrix} i & -i \\ 1 & 1 \end{bmatrix} \begin{bmatrix} e^{it} & \\ & e^{-it} \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \frac{1}{2i}\begin{bmatrix} i & -i \\ 1 & 1 \end{bmatrix} \begin{bmatrix} e^{it} \\ -e^{-it} \end{bmatrix} = \frac{1}{2i}\begin{bmatrix} ie^{it} + ie^{-it} \\ e^{it} - e^{-it} \end{bmatrix} \\ &= \frac{1}{2}\begin{bmatrix} e^{it} + e^{-it} \\ -ie^{it} + ie^{-it} \end{bmatrix} \end{aligned}$$

(b) Clearly, $\boldsymbol{y}(0) = \begin{bmatrix} \cos(0) \\ \sin(0) \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$. On the other hand, $y_1' = -\sin(t) = -y_2$ and $y_2' = \cos(t) = y_1$, so that $\boldsymbol{y}' = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \boldsymbol{y}$. Since the solution to the IVP is unique, it follows that $\begin{bmatrix} \cos(t) \\ \sin(t) \end{bmatrix} = \frac{1}{2}\begin{bmatrix} e^{it} + e^{-it} \\ -ie^{it} + ie^{-it} \end{bmatrix}$.

We have just discovered **Euler's identity**!

Theorem 144. (Euler's identity) $e^{i\theta} = \cos(\theta) + i\sin(\theta)$

Armin Straub
straub@southalabama.edu

Another short proof. Observe that both sides are the (unique) solution to the IVP $y' = iy$, $y(0) = 1$.

On lots of T-shirts. In particular, with $x = \pi$, we get $\boxed{e^{\pi i} = -1}$ or $e^{i\pi} + 1 = 0$ (which connects the five fundamental constants).

Rotation matrices

Example 145. Write down a 2×2 matrix Q for rotation by angle θ in the plane.

Comment. Why should we even be able to represent something like rotation by a matrix? Meaning that $Q\boldsymbol{x}$ should be the vector $\boldsymbol{x}$ rotated by θ. Recall from Linear Algebra I that every **linear map** can be represented by a matrix. Then think about why rotation is a linear map.

Solution. We can determine Q by figuring out $Q\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ (the first column of Q) and $Q\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ (the second column of Q).

Since $Q\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}$ and $Q\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -\sin\theta \\ \cos\theta \end{bmatrix}$, we conclude that $Q = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$.

Comment. Note that we don't need previous knowledge of $\cos$ and $\sin$. We could have introduced these trig functions on the spot.

Comment. Note that it is geometrically obvious that Q is orthogonal. (Why?)

It is clear that $\left\| \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix} \right\|^2 = 1$. Noting that $\left\| \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix} \right\|^2 = \cos^2\theta + \sin^2\theta$, we have rediscovered Pythagoras.

Advanced comment. Actually, every orthogonal 2×2 matrix Q with $\det(Q) = 1$ is a rotation by some angle θ. Orthogonal matrices with $\det(Q) = -1$ are reflections.

Example 146. As in the previous example, let Q_θ be the 2×2 matrix for rotation by angle θ in the plane. What is $Q_\alpha Q_\beta$?

Solution. Note that $Q_\alpha Q_\beta \boldsymbol{x}$ first rotates $\boldsymbol{x}$ by angle β and then by angle α. For geometric reasons, it is obvious that this is the same as if we rotated $\boldsymbol{x}$ by $\alpha + \beta$. It follows that $Q_\alpha Q_\beta = Q_{\alpha+\beta}$.

Comment. This allows us to derive interesting trig identities:

$$\begin{aligned} Q_\alpha Q_\beta &= \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix} \begin{bmatrix} \cos\beta & -\sin\beta \\ \sin\beta & \cos\beta \end{bmatrix} = \begin{bmatrix} \cos\alpha\cos\beta - \sin\alpha\sin\beta & \ldots \\ \ldots & \ldots \end{bmatrix} \\ Q_{\alpha+\beta} &= \begin{bmatrix} \cos(\alpha+\beta) & -\sin(\alpha+\beta) \\ \sin(\alpha+\beta) & \cos(\alpha+\beta) \end{bmatrix} \end{aligned}$$

It follows that $\cos(\alpha+\beta) = \cos\alpha\cos\beta - \sin\alpha\sin\beta$.

Comment. If we set $\beta = \alpha$, this simplifies to $\cos(2\alpha) = \cos^2\alpha - \sin^2\alpha = 2\cos^2\alpha - 1$, the double angle formula that you have probably used countless times in Calculus.

Comment. Similarly, we find an identity for $\sin(\alpha+\beta)$. Spell it out!

More on complex numbers

Let's recall some very basic facts about **complex numbers**:

- Every complex number can be written as $z = x + iy$ with real x, y.
- Here, the imaginary unit i is characterized by solving $x^2 = -1$.

  Important observation. The same equation is solved by $-i$. This means that, algebraically, we cannot distinguish between $+i$ and $-i$.
- The **conjugate** of $z = x + iy$ is $\bar{z} = x - iy$.

  Important comment. Since we cannot algebraically distinguish between $\pm i$, we also cannot distinguish between z and $\bar{z}$. That's the reason why, in problems involving only real numbers, if a complex number $z = x + iy$ shows up, then its **conjugate** $\bar{z} = x - iy$ has to show up in the same manner. With that in mind, have another look at Example 82.

- The **absolute value** of the complex number $z = x + iy$ is $|z| = \sqrt{x^2 + y^2} = \sqrt{\bar{z}z}$.
- The **norm** of the complex vector $\boldsymbol{z} = \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}$ is $\|\boldsymbol{z}\| = \sqrt{|z_1|^2 + |z_2|^2}$.
  Note that $\|\boldsymbol{z}\|^2 = \bar{z}_1 z_1 + \bar{z}_2 z_2 = \bar{\boldsymbol{z}}^T \boldsymbol{z}$.

Definition 147.

- For every matrix A, its **conjugate transpose** is $A^* = (\bar{A})^T$.
- The **dot product** (inner product) of complex vectors is $\boldsymbol{v} \cdot \boldsymbol{w} = \boldsymbol{v}^* \boldsymbol{w}$.
- A complex $n \times n$ matrix A is **unitary** if $A^* A = I$.

Comment. A^* is also written A^H (or $A^\dagger$ in quantum mechanics) and called the Hermitian conjugate.

Comment. For real matrices and vectors, the conjugate transpose is just the ordinary transpose. In particular, the dot product is the same.

Comment. Unitary matrices are the complex version of orthogonal matrices. (A real matrix if unitary if and only if it is orthogonal.)

Example 148. What is the norm of the vector $\begin{bmatrix} 1-i \\ 2+3i \end{bmatrix}$?

Solution. $\left\| \begin{bmatrix} 1-i \\ 2+3i \end{bmatrix} \right\|^2 = \begin{bmatrix} 1+i & 2-3i \end{bmatrix} \begin{bmatrix} 1-i \\ 2+3i \end{bmatrix} = |1-i|^2 + |2+3i|^2 = 2 + 13$. Hence, $\left\| \begin{bmatrix} 1-i \\ 2+3i \end{bmatrix} \right\| = \sqrt{15}$.

Example 149. Determine A^* if $A = \begin{bmatrix} 2 & 1-i \\ 3+2i & i \end{bmatrix}$.

Solution. $A^* = \begin{bmatrix} 2 & 3-2i \\ 1+i & -i \end{bmatrix}$

Example 150. What is $\frac{1}{2+3i}$?

Solution. $\frac{1}{2+3i} = \frac{2-3i}{(2+3i)(2-3i)} = \frac{2-3i}{13}$.

In general. $\frac{1}{z} = \frac{\bar{z}}{z\,\bar{z}} = \frac{\bar{z}}{|z|^2}$

Example 151. (extra) We can identify complex numbers $x + iy$ with vectors $\begin{bmatrix} x \\ y \end{bmatrix}$ in $\mathbb{R}^2$. Then, what is the geometric effect of multiplying with i?

Solution. Algebraically, the effect of multiplying $x + iy$ with i obviously is $i(x + iy) = -y + ix$.

Since multiplication with i is obviously linear, we can represent it using a 2×2 matrix J acting on vectors $\begin{bmatrix} x \\ y \end{bmatrix}$.

$J \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ (this is the same as saying $i \cdot 1 = i$) and $J \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$ (this is the same as saying $i \cdot i = -1$).

Hence, $J = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$. This is precisely the rotation matrix for a rotation by $90°$.

In other words, multiplication with i has the geometric effect of rotating complex numbers by $90°$.

Comment. The relation $i^2 = -1$ translates to $J^2 = -I$.

Complex numbers as 2×2 matrices. In light of the above, we can express complex numbers $x + iy$ as the 2×2 matrix $xI + yJ = \begin{bmatrix} x & -y \\ y & x \end{bmatrix}$. Adding and multiplying these matrices behaves exactly the same way as adding or multiplying the complex numbers directly.

For instance, $(2 + 3i)(4 - i) = 8 + 10i - 3i^2 = 11 + 10i$ versus $\begin{bmatrix} 2 & -3 \\ 3 & 2 \end{bmatrix} \begin{bmatrix} 4 & 1 \\ -1 & 4 \end{bmatrix} = \begin{bmatrix} 11 & -10 \\ 10 & 11 \end{bmatrix}$.

Likewise for inverses: $\frac{1}{2+3i} = \frac{2-3i}{(2+3i)(2-3i)} = \frac{2-3i}{13}$ versus $\begin{bmatrix} 2 & -3 \\ 3 & 2 \end{bmatrix}^{-1} = \frac{1}{13} \begin{bmatrix} 2 & 3 \\ -3 & 2 \end{bmatrix}$

Extra: More details on the spectral theorem

Let us add $\langle \boldsymbol{v}, \boldsymbol{w} \rangle$ to our notations for the dot product: $\langle \boldsymbol{v}, \boldsymbol{w} \rangle = \boldsymbol{v}^T \boldsymbol{w} = \boldsymbol{v} \cdot \boldsymbol{w}$.

- In our story of orthogonality, the important player has been the dot product. However, one could argue that the fundamental quantity is actually the norm:
  $\langle \boldsymbol{v}, \boldsymbol{w} \rangle = \frac{1}{4}(\|\boldsymbol{v} + \boldsymbol{w}\|^2 - \|\boldsymbol{v} - \boldsymbol{w}\|^2)$. See Example 24.

- Accepting the dot product as immensely important, we see that symmetric matrices (i.e. matrices A such that $A = A^T$) are of interest.
  For every matrix A, $\langle A\boldsymbol{v}, \boldsymbol{w} \rangle = \langle \boldsymbol{v}, A^T \boldsymbol{w} \rangle$.
  It follows that, a matrix A is symmetric if and only if $\langle A\boldsymbol{v}, \boldsymbol{w} \rangle = \langle \boldsymbol{v}, A\boldsymbol{w} \rangle$ for all vectors $\boldsymbol{v}, \boldsymbol{w}$.

- Similarly, let Q be an orthogonal matrix (i.e. Q is a square matrix with $Q^T Q = I$).
  Then, $\langle Q\boldsymbol{v}, Q\boldsymbol{w} \rangle = \langle \boldsymbol{v}, \boldsymbol{w} \rangle$.
  In fact, a matrix A is orthogonal if and only if $\langle A\boldsymbol{v}, A\boldsymbol{w} \rangle = \langle \boldsymbol{v}, \boldsymbol{w} \rangle$ for all vectors $\boldsymbol{v}, \boldsymbol{w}$.
  Comment. We observed in Example 145 that orthogonal matrices Q correspond to rotations ($\det Q = 1$) or reflections ($\det Q = -1$) [or products thereof]. The equality $\langle Q\boldsymbol{v}, Q\boldsymbol{w} \rangle = \langle \boldsymbol{v}, \boldsymbol{w} \rangle$ encodes the fact that these types (and only these!) of geometric transformations preserve angles and lengths.

(Spectral theorem)

A $n \times n$ matrix A is symmetric if and only if it can be decomposed as $A = PDP^T$, where

- D is a diagonal matrix, $(n \times n)$

The diagonal entries λ_i are the **eigenvalues** of A.

- P is orthogonal. $(n \times n)$

The columns of P are **eigenvectors** of A.

Note that, in particular, A is always diagonalizable, the eigenvalues (and hence, the eigenvectors) are all real, and, most importantly, the eigenspaces of A are orthogonal.
The "only if" part says that, if A is symmetric, then we get a diagonalization $A = PDP^T$. The "if" part says that, if $A = PDP^T$, then A is symmetric (which follows from $A^T = (PDP^T)^T = (P^T)^T D^T P^T = PDP^T = A$).

Let us prove the following important parts of the spectral theorem.

Theorem 152.

(a) If A is symmetric, then the eigenspaces of A are orthogonal.

(b) If A is real and symmetric, then the eigenvalues of A are real.

Proof.

(a) We need to show that, if $\boldsymbol{v}$ and $\boldsymbol{w}$ are eigenvectors of A with different eigenvalues, then $\langle \boldsymbol{v}, \boldsymbol{w} \rangle = \boldsymbol{0}$.
Suppose that $A\boldsymbol{v} = \lambda \boldsymbol{v}$ and $A\boldsymbol{w} = \mu \boldsymbol{w}$ with $\lambda \neq \mu$.
Then, $\lambda \langle \boldsymbol{v}, \boldsymbol{w} \rangle = \langle \lambda \boldsymbol{v}, \boldsymbol{w} \rangle = \langle A\boldsymbol{v}, \boldsymbol{w} \rangle = \langle \boldsymbol{v}, A^T \boldsymbol{w} \rangle = \langle \boldsymbol{v}, A\boldsymbol{w} \rangle = \langle \boldsymbol{v}, \mu \boldsymbol{w} \rangle = \mu \langle \boldsymbol{v}, \boldsymbol{w} \rangle$.
However, since $\lambda \neq \mu$, $\lambda \langle \boldsymbol{v}, \boldsymbol{w} \rangle = \mu \langle \boldsymbol{v}, \boldsymbol{w} \rangle$ is only possible if $\langle \boldsymbol{v}, \boldsymbol{w} \rangle = 0$.

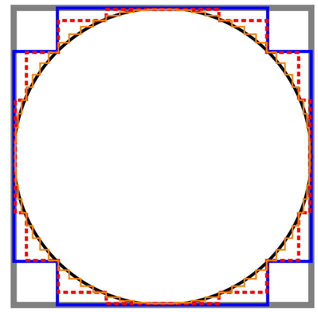
(b) Suppose λ is a nonreal eigenvalue with nonzero eigenvector $\boldsymbol{v}$. Then, $\bar{\boldsymbol{v}}$ is a $\bar{\lambda}$-eigenvector and, since $\lambda \neq \bar{\lambda}$, we have two eigenvectors with different eigenvalues. By the first part, these two eigenvectors must be orthogonal in the sense that $\bar{\boldsymbol{v}}^T \boldsymbol{v} = 0$. But $\bar{\boldsymbol{v}}^T \boldsymbol{v} = \boldsymbol{v}^* \boldsymbol{v} = \|\boldsymbol{v}\|^2 \neq 0$. This shows that it is impossible to have a nonzero eigenvector for a nonreal eigenvalue. □

Let us highlight the following point we used in our proof:

Let A be a real matrix. If $\boldsymbol{v}$ is a λ-eigenvector, then $\bar{\boldsymbol{v}}$ is a $\bar{\lambda}$-eigenvector.

See, for instance, Example 82. This is just a consequence of the basic fact that we cannot algebraically distinguish between $+i$ and $-i$.

Remark 153. (April Fools' Day!) π is the perimeter of a circle enclosed in a square with edge length 1. The perimeter of the square is 4, which approximates π. To get a better approximation, we "fold" the vertices of the square towards the circle (and get the blue polygon). This construction can be repeated for even better approximations and, in the limit, our shape will converge to the true circle. At each step, the perimeter is 4, so we conclude that $\pi = 4$, contrary to popular belief.



Can you pin-point the fallacy in this argument?

Comment. We'll actually come back to this. It's related to linear algebra in infinite dimensions.

Armin Straub
straub@southalabama.edu