Exponentials and Rotations

Allan Yashinski

July 21, 2021

18.1   Introduction

We study how rotations arise from matrix exponentials.

18.2   An example

Suppose we wish to solve the differential equation

$$\bar{x}'(t) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \bar{x}(t), \qquad \bar{x}(0) = \begin{bmatrix} 2 \\ -5 \end{bmatrix}.$$

We know the solution is given by $\bar{x}(t) = e^{tA}\bar{x}_0$, where A is the given 2×2 matrix above. To compute the matrix exponential, we begin to diagonalize A. We first find the eigenvalues:

$$\det(A - \lambda I) = 0$$
$$\det \begin{bmatrix} -\lambda & -1 \\ 1 & \lambda \end{bmatrix} = 0$$
$$\lambda^2 + 1 = 0$$
$$\lambda^2 = -1$$
$$\lambda = \pm\sqrt{-1} = \pm i.$$

The eigenvalues of A are not real numbers, but we can continue as we would if they were. To diagonalize, we need to find eigenvectors. For $\lambda = i$, we solve $(A - iI)\bar{v} = \bar{0}$, which has augmented matrix

$$\begin{bmatrix} -i & -1 & 0 \\ 1 & -i & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & -i & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

The first equation says that $v_1 = iv_2$. Since v_2 is free, we can take $v_2 = 1$, which gives the eigenvector $\bar{v} = \begin{bmatrix} i \\ 1 \end{bmatrix}$. We can verify it is an eigenvector since

$$A\bar{v} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} i \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ i \end{bmatrix} \quad \text{and} \quad i\bar{v} = i\begin{bmatrix} i \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ i \end{bmatrix}.$$

In a similar way, one can find the eigenvector $\begin{bmatrix} -i \\ 1 \end{bmatrix}$ for the eigenvalue $\lambda = -i$. So we can write[1] $A = PDP^{-1}$ where

$$P = \begin{bmatrix} i & -i \\ 1 & 1 \end{bmatrix} \quad \text{and} \quad D = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}.$$

Using the usual 2×2 inversion formula, we get

$$P^{-1} = \frac{1}{2i} \begin{bmatrix} 1 & i \\ -1 & i \end{bmatrix} = \begin{bmatrix} (1/2)i^{-1} & 1/2 \\ -(1/2)i^{-1} & 1/2 \end{bmatrix} = \begin{bmatrix} -i/2 & 1/2 \\ i/2 & 1/2 \end{bmatrix}$$

---

[1]Diagonalization works the same way with complex eigenvalues/eigenvectors.

because $i^{-1} = -i$. So we have

$$A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} i & -i \\ 1 & 1 \end{bmatrix} \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix} \begin{bmatrix} -i/2 & 1/2 \\ i/2 & 1/2 \end{bmatrix}.$$

At this point, we may be somewhat concerned because the original differential equation involved only real numbers, and we now have complex numbers all over the place. Nonetheless, we persevere.

Using $e^{tA} = Pe^{tD}P^{-1}$, we have

$$e^{tA} = \begin{bmatrix} i & -i \\ 1 & 1 \end{bmatrix} \begin{bmatrix} e^{it} & 0 \\ 0 & e^{-it} \end{bmatrix} \begin{bmatrix} -i/2 & 1/2 \\ i/2 & 1/2 \end{bmatrix}.$$

At this point, we recall Euler's formula:

Theorem 18.2.0.1 (Euler's formula)**.** For any real number x, we have

$$e^{ix} = \cos x + i \sin x.$$

This directly tells us what the first complex exponential is. For the second, we have

$$e^{-it} = e^{i(-t)} = \cos(-t) + i \sin(-t) = \cos t - i \sin t,$$

where the last equality follows from trig identities. So we can simplify the complex exponentials and compute

$$\begin{aligned}
e^{tA} &= \begin{bmatrix} i & -i \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \cos t + i \sin t & 0 \\ 0 & \cos t - i \sin t \end{bmatrix} \begin{bmatrix} -i/2 & 1/2 \\ i/2 & 1/2 \end{bmatrix} \\
&= \begin{bmatrix} i \cos t - \sin t & -i \cos t - \sin t \\ \cos t + i \sin t & \cos t - i \sin t \end{bmatrix} \begin{bmatrix} -i/2 & 1/2 \\ i/2 & 1/2 \end{bmatrix} \\
&= \begin{bmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{bmatrix}.
\end{aligned}$$

Amazing! Not only have all the imaginary parts cancelled, the result is pleasantly simple (it is a rotation matrix.) We can finish solving the differential equation:

$$\bar{x}(t) = e^{tA}\bar{x}_0 = \begin{bmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{bmatrix} \begin{bmatrix} 2 \\ -5 \end{bmatrix} = \begin{bmatrix} 2\cos t + 5\sin t \\ 2\sin t - 5\cos t \end{bmatrix}.$$

It was pretty cool that all the imaginary parts cancelled, but we sort of knew this had to happen in advance. The matrix e^{tA} is defined to be the sum of the series

$$e^{tA} = \sum_{k=0}^{\infty} \frac{t^k}{k!} A^k.$$

Since A is real, every term of the series is real, and therefore the sum has to be real.

Looking at the original system from a different perspective, it is not surprising that $\sin t$ and $\cos t$ appeared. The system is

$$\begin{cases} x_1'(t) = - x_2(t) \\ x_2'(t) = x_1(t). \end{cases}$$

If we take the derivative of the first equation and substitute in the second, we obtain

$$x_1''(t) = -x_2'(t) = -x_1(t),$$

and similarly $x_2''(t) = -x_2(t)$. The functions $\sin t$ and $\cos t$ are two functions whose second derivative is equal to its negative. Any function that has this property must actually be a linear combination of $\sin t$ and $\cos t$, just like we saw in the solution above.

18.3   More properties of matrix exponentials

We'd like to understand better why the matrices e^{tA} were rotation matrices in the previous example. Our goal will be to understand which properties of A imply that e^{tA} are rotation matrices.

Recall that an *orthogonal* matrix is an $n \times n$ matrix Q such that $Q^T Q = I_n$. In other words, $Q^T = Q^{-1}$. Equivalently, Q is orthogonal if and only if its columns are an orthonormal basis for $\mathbb{R}^n$. It follows from the definition of orthogonal matrix that $\det Q = \pm 1$. An orthogonal matrix with determinant 1 is a *rotation*, and an orthogonal matrix with determinant -1 is a *reflection*.

So we'd like to know under what circumstances e^{tA} is an orthogonal matrix with determinant 1. We'll investigate more properties of matrix exponentials.

Suppose A and B are $n \times n$ matrices. What happens when we multiply $e^A e^B$? For numbers, a and b, we know that $e^a e^b = e^{a+b}$. Unfortunately, this property does not carry over to matrices.

Exercise 18.1. Give an example of two 2×2 matrices A, B for which $e^A e^B \neq e^{A+B}$.

The reason why the result fails for matrices is because matrix multiplication is noncommutative. So there is a partial result in the case where A and B commute.

Theorem 18.3.0.1. If $AB = BA$, then $e^A e^B = e^{A+B}$.

4

It might seem rare to have commuting matrices, but it can be applied in situations where the second matrix has a strong relationship to the first, as in the following results.

Corollary 18.3.0.1. Let A be an $n \times n$ matrix.

1. For any real numbers t and s, we have $e^{tA}e^{sA} = e^{(t+s)A}$.

2. The matrix e^A is invertible, and $(e^A)^{-1} = e^{-A}$.

For the first, note that tA and sA commute. For the second, note that A and $-A$ commute. Consequently,

$$e^A e^{-A} = e^{A+(-A)} = e^{\bar{0}_{n \times n}} = I_n.$$

Since we are interested in orthogonal matrices, we will need to take the transpose of a matrix exponential. This is easy enough.

Theorem 18.3.0.2. For any $n \times n$ matrix A, $(e^A)^T = e^{(A^T)}$.

This follows because the transpose operation is linear and it commutes with powers of a matrix, meaning $(A^k)^T = (A^T)^k$. This justifies the following:

$$(e^A)^T = \left(\sum_{k=0}^{\infty} \frac{1}{k!} A^k \right)^T = \sum_{k=0}^{\infty} \frac{1}{k!} (A^k)^T = \sum_{k=0}^{\infty} \frac{1}{k!} (A^T)^k = e^{(A^T)}.$$

The last thing we will need is some information about the eigenvalues and eigenvectors of e^A.

Theorem 18.3.0.3. Let A be an $n \times n$ matrix. Then

1. If λ is an eigenvalue for A, then e^λ is an eigenvalue for e^A.

2. More precisely, if $\bar{v}$ is an eigenvector for A with eigenvalue λ, then $\bar{v}$ is an eigenvector for e^A with eigenvalue e^λ.

So A and e^A have the same eigenvectors, but the eigenvalue for e^A is the exponential of the eigenvalue for A. This follows from the definition of e^A. Suppose $A\bar{v} = \lambda\bar{v}$. Then

$$e^A \bar{v} = \sum_{k=0}^{\infty} \frac{1}{k!} A^k \bar{v} = \sum_{k=0}^{\infty} \frac{1}{k!} \lambda^k \bar{v} = \left(\sum_{k=0}^{\infty} \frac{\lambda^k}{k!} \right) \bar{v} = e^\lambda \bar{v}.$$

18.4 Rotations as matrix exponentials

Now we determine conditions on A that will make e^A an orthogonal matrix. Orthogonal means $(e^A)^T = (e^A)^{-1}$. Note that the former equals $e^{(A^T)}$ and the latter equals e^{-A}. So this would work if A satisfied the condition $A^T = -A$.

Definition 18.4.0.1. An $n \times n$ matrix is called *skew-symmetric* if $A^T = -A$.

Theorem 18.4.0.1. If A is a skew-symmetric $n \times n$ matrix, then e^A is orthogonal. Additionally, $\det(e^A) = 1$, so e^A is a rotation matrix.

So what does a skew-symmetric matrix look like? For a 2×2 matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, we see A is skew-symmetric if and only if

$$A^T = -A$$
$$\begin{bmatrix} a & c \\ b & d \end{bmatrix} = \begin{bmatrix} -a & -b \\ -c & -d \end{bmatrix}.$$

This gives the system

$$\begin{cases} a = -a \\ b = -c \\ d = -d \end{cases}$$

The first equation implies $a = 0$, and similarly $d = 0$ from the last equation. Hence a 2×2 skew-symmetric matrix has the form $\begin{bmatrix} 0 & -b \\ b & 0 \end{bmatrix}$. This is precisely the family of matrices we took the exponential of when we computed

$$\exp\left(\begin{bmatrix} 0 & -t \\ t & 0 \end{bmatrix} \right) = \begin{bmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{bmatrix}.$$

Note that if A is a skew-symmetric matrix, then any scalar multiple tA is also skew-symmetric. The result below follows.

Theorem 18.4.0.2. If A is a skew-symmetric $n \times n$ matrix, then e^{tA} is a rotation matrix for each real number t.

Let's try to better understand what happens in the case where A is a 3×3 skew-symmetric matrix. Such a matrix necessarily has the form

$$A = \begin{bmatrix} 0 & -a & -b \\ a & 0 & -c \\ b & c & 0 \end{bmatrix}.$$

For reasons that will become clearer later, it will be convenient to consider a skew-symmetric matrix in the form

$$A = \begin{bmatrix} 0 & -z & y \\ z & 0 & -x \\ -y & x & 0 \end{bmatrix}.$$

We know that e^A, and more generally e^{tA}, are rotations in $\mathbb{R}^3$. What kind of rotations are they? More specifically, what is the axis of rotation, and what is the angle of rotation?

Let's compute the eigenvalues of this matrix:

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

$$\det \begin{bmatrix} -\lambda & -z & y \\ z & -\lambda & -x \\ -y & x & -\lambda \end{bmatrix} = 0$$

$$-\lambda(\lambda^2 + x^2) + z(-\lambda z - xy) + y(xz - \lambda y) = 0$$

$$-\lambda^3 - \lambda x^2 - \lambda z^2 - \lambda y^2 = 0$$

$$-\lambda^3 - \lambda||\bar{v}||^2 = 0$$

$$-\lambda(\lambda^2 + ||\bar{v}||^2) = 0,$$

where $\bar{v} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$. It follows that the eigenvalues are $\lambda = 0$ and $\lambda = \pm i||\bar{v}||$. Let's focus on the $\lambda = 0$ eigenvalue now. As it turns out, $\bar{v}$ is an eigenvector for $\lambda = 0$. We can confirm this:

$$A\bar{v} = \begin{bmatrix} 0 & -z & y \\ z & 0 & -x \\ -y & x & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = 0\bar{v}.$$

The importance of this is that we know that the same vector $\bar{v}$ is an eigenvector for e^A with eigenvalue $e^0 = 1$. That is, $e^A\bar{v} = \bar{v}$. A vector that is fixed by a rotation in $\mathbb{R}^3$ must point along the axis of rotation! Note that we will also have $e^{tA}\bar{v} = \bar{v}$ for any t. It turns out that the norm of $\bar{v}$ is determining the angle of rotation. The following theorem describes the whole story.

Theorem 18.4.0.3. Let $\bar{u} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$ be a unit vector and let $A = \begin{bmatrix} 0 & -z & y \\ z & 0 & -x \\ -y & x & 0 \end{bmatrix}.$
Then rotation about the line through the origin in the direction of $\bar{u}$ by θ radians is given by the matrix $e^{\theta A}$.

Note that if we consider any of the unit vectors $u = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$, we can

7

recover the formulas for rotations about the x-, y-, or z-axes. For example, rotation about the x-axis by θ radians is given by

$$\exp\left(\theta \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}\right) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{bmatrix}.$$

This could be worked out by hand in a manner similar to the first example given in this document.

We can give any rotation matrix in the case where the axis goes through the origin. For example, suppose we want to rotate about the axis determined by the vector $\bar{v} = \begin{bmatrix} 3 \\ 2 \\ -6 \end{bmatrix}$ by 27 degrees. The angle is $\theta = \dfrac{2\pi}{360}(27) = \dfrac{3\pi}{20}$ radians.
We need the unit vector

$$\bar{u} = \frac{1}{||\bar{v}||}\bar{v} = \frac{1}{\sqrt{9+4+36}} \begin{bmatrix} 3 \\ 2 \\ -6 \end{bmatrix} = \begin{bmatrix} 3/7 \\ 2/7 \\ -6/7 \end{bmatrix}.$$

We can then use the `expm` command in MATLAB to compute our rotation matrix

$$\exp\left(\frac{3\pi}{20} \begin{bmatrix} 0 & 6/7 & 2/7 \\ -6/7 & 0 & -3/7 \\ -2/7 & 3/7 & 0 \end{bmatrix}\right) \approx \begin{bmatrix} 0.9110 & 0.4025 & 0.0897 \\ -0.3758 & 0.8999 & -0.2213 \\ -0.1697 & 0.1679 & 0.9711 \end{bmatrix}.$$

More generally, we can do rotations about any line in $\mathbb{R}^3$ (not necessarily through the origin), if we use homogeneous coordinates. Say we want to rotate by 27 degrees about the line which goes through the point $(8, -4, 5)$ and points in the direction of $\bar{v} = \begin{bmatrix} 3 \\ 2 \\ -6 \end{bmatrix}$. We use the translation matrix

$$T(8, -4, 5) = \begin{bmatrix} 1 & 0 & 0 & 8 \\ 0 & 1 & 0 & -4 \\ 0 & 0 & 1 & 5 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

and its inverse, and compute

$$\begin{bmatrix} 1 & 0 & 0 & 8 \\ 0 & 1 & 0 & -4 \\ 0 & 0 & 1 & 5 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0.9110 & 0.4025 & 0.0897 & 0 \\ -0.3758 & 0.8999 & -0.2213 & 0 \\ -0.1697 & 0.1679 & 0.9711 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & -8 \\ 0 & 1 & 0 & 4 \\ 0 & 0 & 1 & -5 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 0.9110 & 0.4025 & 0.0897 & 1.8734 \\ -0.3758 & 0.8999 & -0.2213 & 3.7122 \\ -0.1697 & 0.1679 & 0.9711 & 2.1741 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

For example, if we want to apply this rotation to the point $(10, -11, -2)$, we compute

$$\begin{bmatrix} 0.9110 & 0.4025 & 0.0897 & 1.8734 \\ -0.3758 & 0.8999 & -0.2213 & 3.7122 \\ -0.1697 & 0.1679 & 0.9711 & 2.1741 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 10 \\ -11 \\ -2 \\ 1 \end{bmatrix} = \begin{bmatrix} 6.3770 \\ -9.5021 \\ -3.3122 \\ 1 \end{bmatrix}.$$

The rotated point is $(6.3770, -9.5021, -3.3122)$.