Power Series of Square Matrices

Before I dive into things, let me remind you of what's known as Euler's formula. For any real number θ, Euler's formula says

$$e^{i\theta} = \cos(\theta) + i\sin(\theta).$$

It's derived as follows. Recall the Taylor series from calculus

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}$$

which has an infinite radius of convergence. Let $x = i\theta$ and rearrange the sum into even and odd terms

$$e^{i\theta} = \sum_{n=0}^{\infty} \frac{(i\theta)^n}{n!} = \sum_{k=0}^{\infty} \frac{(i\theta)^{2k}}{(2k)!} + \sum_{k=0}^{\infty} \frac{(i\theta)^{2k+1}}{(2k+1)!}.$$

Note, $n = 2k$ are the even terms and $n = 2k+1$ are the odd terms. Since $i^2 = -1$

$$(i)^{2k} = (-1)^k, \quad (i)^{2k+1} = i(-1)^k.$$

Plug these into the sums above to get

$$e^{i\theta} = \sum_{k=0}^{\infty} \frac{(-1)^k \theta^{2k}}{(2k)!} + i\sum_{k=0}^{\infty} \frac{(-1)^k \theta^{2k+1}}{(2k+1)!}.$$

Finally, recognize the left sum is the Taylor series for $\cos(\theta)$ and the right sum is the Taylor series for $\sin(\theta)$. There you go.

Let $\theta \to -\theta$ in Euler's formula and use $\cos(-\theta) = \cos(\theta)$ and $\sin(-\theta) = -\sin(\theta)$ to see

$$e^{-i\theta} = \cos(\theta) - i\sin(\theta).$$

Therefore

$$\frac{e^{i\theta} + e^{-i\theta}}{2} = \cos(\theta) \quad \text{and} \quad \frac{e^{i\theta} - e^{-i\theta}}{2i} = \sin(\theta).$$

Every mathematics student should know Euler's formula and the implied complex exponential formulae given above for $\cos(\theta)$ and $\sin(\theta)$.

Now, suppose a given power series

$$f(x) = \sum_{n=0}^{\infty} f_n\, x^n$$

has positive radius of convergence R, i.e. it converges on the open disk $\{x : |x| < R\}$ but diverges (when $R < \infty$) on $\{x : |x| > R\}$. $1/R$ can be obtained by the root test. For example, from calculus, you should all remember

$$e^x = \sum_{n=0}^{\infty} \frac{1}{n!}\, x^n, \quad \cos(x) = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!}\, x^{2n}, \quad \sin(x) = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!}\, x^{2n+1},$$

each of which has an infinite radius of convergence.

Power series make a natural vehicle to define functions of square matrices, and they often come up in applications. Notice that if we define

$$f(A) \equiv \sum_{n=0}^{\infty} f_n A^n = f_0 I + f_1 A + f_2 A^2 + \cdots$$

for the given sequence of numbers f_n which define $f(x)$, when A is an $m \times m$ matrix, so is A^n, and therefore $f(A)$ is an $m \times m$ matrix. Now consider

$$f(At) \equiv \sum_{n=0}^{\infty} f_n (At)^n = f_0 I + f_1 tA + f_2 t^2 A^2 + \cdots,$$

where t is a scalar variable.

Without getting into convergence issues, let's play around with $f(At)$. I'm going to differentiate the matrix power series term-by-term.

$$\frac{d}{dt} f(At) = \sum_{n=1}^{\infty} f_n \, n t^{n-1} A^n.$$

Notice the sum starts at $n = 1$ because the $n = 0$ term is constant. Rearrange to write

$$\frac{d}{dt} f(At) = A \sum_{n=1}^{\infty} n f_n \, (At)^{n-1} = A \, f'(At)$$

where $f'(x)$ is the derivative of $f(x)$, i.e.

$$f(x) = f_0 + f_1 x + f_2 x^2 + f_3 x^3 + \cdots \quad \Rightarrow \quad f'(x) = f_1 + 2 f_2 x + 3 f_3 x^2 + \cdots.$$

For example, given $f(x) = e^x \;\Rightarrow\; f'(x) = e^x$, and so we have

$$\frac{d}{dt} e^{At} = A \, e^{At},$$

or given $f(x) = \sin(x) \;\Rightarrow\; f'(x) = \cos(x)$

$$\frac{d}{dt} \sin(At) = A \, \cos(At).$$

What I'm really interested in showing you here is how to evaluate such power series in closed (non infinite series) form. It's quite easy when A is diagonalizable, and we'll do it immediately below. It's also not too hard, but can be complicated, when A is not diagonalizable. Perhaps you'll see these somewhat exceptional cases treated when you learn about the *Jordan canonical form*.

Now, suppose for a given $m \times m$ matrix A there is an invertible $m \times m$ matrix R such that

$$R^{-1} A R = \Lambda \equiv \mathrm{diag}(\lambda_1, \ldots, \lambda_m).$$

This is possible if and only if the columns of R are independent eigenvectors of A, and the diagonal entries of Λ are the associated eigenvalues. Now

$$R^{-1}AR = \Lambda \quad \Rightarrow \quad A = R\Lambda R^{-1},$$

and plug this in

$$f(At) = \sum_{n=0}^{\infty} f_n t^n A^n \quad \Rightarrow \quad f(At) = \sum_{n=0}^{\infty} f_n t^n (R\Lambda R^{-1})^n.$$

But

$$(R\Lambda R^{-1})^n = R\Lambda R^{-1} \cdots R\Lambda R^{-1} = R\Lambda^n R^{-1}.$$

Therefore

$$f(At) = R\left(\sum_{n=0}^{\infty} f_n t^n \Lambda^n\right) R^{-1} = R\, f(\Lambda t)\, R^{-1}.$$

Since Λ is diagonal, $f(\Lambda t)$ is particularly easy to evaluate. In fact

$$f(\Lambda t) = \operatorname{diag}(f(\lambda_1 t), \ldots, f(\lambda_m t)),$$

where $f(\lambda_i t)$ is the scalar valued power series function. (When A can not be diagonalized, this step is slightly more complicated.)

Here're a couple of examples.

First a 2×2 example with complex eigenvalues. Compute that

$$A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \quad \Rightarrow \quad \lambda = -i,\ \mathbf{r} = \begin{pmatrix} i \\ 1 \end{pmatrix}, \quad \lambda = i,\ \mathbf{r} = \begin{pmatrix} -i \\ 1 \end{pmatrix},$$

and from this

$$R = \begin{pmatrix} i & -i \\ 1 & 1 \end{pmatrix}, \quad \Lambda = \begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix}, \quad R^{-1} = \frac{1}{2i}\begin{pmatrix} 1 & i \\ -1 & i \end{pmatrix}.$$

From our work above we find

$$e^{At} = R\, e^{\Lambda t}\, R^{-1} = \begin{pmatrix} i & -i \\ 1 & 1 \end{pmatrix} \begin{pmatrix} e^{-it} & 0 \\ 0 & e^{it} \end{pmatrix} \frac{1}{2i} \begin{pmatrix} 1 & i \\ -1 & i \end{pmatrix}$$

and multiply this out and use Euler's formula to get

$$\frac{1}{2i}\begin{pmatrix} ie^{-it} + ie^{it} & -e^{-it} + e^{it} \\ e^{-it} - e^{it} & ie^{-it} + ie^{it} \end{pmatrix} = \begin{pmatrix} \cos(t) & \sin(t) \\ -\sin(t) & \cos(t) \end{pmatrix}.$$

That's it. Pretty simple. Just for the fun of it, let's confirm that $\dfrac{d}{dt}e^{At} = Ae^{At}$.

$$\frac{d}{dt}e^{At} = \begin{pmatrix} -\sin(t) & \cos(t) \\ -\cos(t) & -\sin(t) \end{pmatrix}$$

$$Ae^{At} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} \cos(t) & \sin(t) \\ -\sin(t) & \cos(t) \end{pmatrix} = \begin{pmatrix} -\sin(t) & \cos(t) \\ -\cos(t) & -\sin(t) \end{pmatrix},$$

which agrees with what was shown in the general case above.

For the second example, consider

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \quad \Rightarrow \quad \lambda = 0,\ \mathbf{r} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \quad \lambda = 2,\ \mathbf{r} = \begin{pmatrix} 1 \\ 1 \end{pmatrix},$$

and from this get

$$R = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}, \quad \Lambda = \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix}, \quad R^{-1} = \frac{1}{2}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}.$$

This time let's calculate both $\sin(At)$ and $\cos(At)$.

$$\begin{aligned}
\sin(At) &= \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} \sin(0t) & 0 \\ 0 & \sin(2t) \end{pmatrix} \frac{1}{2}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \\
&= \frac{1}{2}\begin{pmatrix} \sin(2t) & \sin(2t) \\ \sin(2t) & \sin(2t) \end{pmatrix}, \\
\cos(At) &= \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} \cos(0t) & 0 \\ 0 & \cos(2t) \end{pmatrix} \frac{1}{2}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \\
&= \frac{1}{2}\begin{pmatrix} 1+\cos(2t) & -1+\cos(2t) \\ -1+\cos(2t) & 1+\cos(2t) \end{pmatrix}.
\end{aligned}$$

Again, let's confirm the general result that $\dfrac{d}{dt} f(At) = Af'(At)$ for this example when $f(x) = \sin(x)$, $f'(x) = \cos(x)$.

$$\begin{aligned}
\frac{d}{dt}\sin(At) &= \frac{1}{2}\begin{pmatrix} 2\cos(2t) & 2\cos(2t) \\ 2\cos(2t) & 2\cos(2t) \end{pmatrix} \\
A\cos(At) &= \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \frac{1}{2}\begin{pmatrix} 1+\cos(2t) & -1+\cos(2t) \\ -1+\cos(2t) & 1+\cos(2t) \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 2\cos(2t) & 2\cos(2t) \\ 2\cos(2t) & 2\cos(2t) \end{pmatrix},
\end{aligned}$$

which is the expected result.

Here are a few exercises. Many of these would make very good exam questions.

---

1. The following matrix A has the given eigenvalues and corresponding eigenvectors.

$$A = \begin{pmatrix} -1 & 6 \\ -1 & 4 \end{pmatrix}, \quad \lambda = 1,\ \mathbf{r} = \begin{pmatrix} 3 \\ 1 \end{pmatrix}, \quad \lambda = 2,\ \mathbf{r} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}.$$

(a) Compute $\cos(At)$. (b) Compute $\sin(At)$.

(c) Confirm that $\dfrac{d}{dt}\sin(At) = A\cos(At)$.

2. Let A be the matrix given in exercise 1.

(a) Compute e^{At}. (b) Confirm that $\dfrac{d}{dt}e^{At} = A\,e^{At}$.

3. The following matrix A has the given eigenvalues and corresponding eigenvectors.

$$A = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}, \quad \lambda = 1 - i, \ \mathbf{r} = \begin{pmatrix} 1 \\ i \end{pmatrix}, \quad \lambda = 1 + i, \ \mathbf{r} = \begin{pmatrix} 1 \\ -i \end{pmatrix}.$$

(a) Compute e^{At}. (b) Confirm that $\dfrac{d}{dt} e^{At} = A\, e^{At}$.

4. The following 3×3 matrix A has the given eigenvalues and corresponding eigenvectors.

$$A = \begin{pmatrix} 3 & -1 & 1 \\ -2 & 4 & 2 \\ -1 & 1 & 5 \end{pmatrix}, \quad \lambda = 2, \ \mathbf{r} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad \lambda = 4, \ \mathbf{r} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \quad \lambda = 6, \ \mathbf{r} = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}.$$

(a) Compute e^{At}. (b) Confirm that $\dfrac{d}{dt} e^{At} = A\, e^{At}$.

5. Recall the Taylor series from calculus: $\log(1 - x) = -\sum_{n=1}^{\infty} x^n/n$ has interval of convergence $\{x : |x| < 1\}$. Let A be the 2×2 matrix from exercise 1. For t restricted to $|t| < 1/2$ do the following.

(a) Compute $\log(I - At)$. (b) Confirm that $\dfrac{d}{dt} \log(I - At) = -A\,(I - At)^{-1}$.

6. The matrix A has eigenvalues and corresponding eigenvectors.

$$A = \begin{pmatrix} -5 & 18 \\ -3 & 10 \end{pmatrix}, \quad \lambda = 1, \ \mathbf{r} = \begin{pmatrix} 3 \\ 1 \end{pmatrix}, \quad \lambda = 4, \ \mathbf{r} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}.$$

Determine a matrix, say $\sqrt{A}$, that satisfies $\left(\sqrt{A}\right)^2 = A$.

Hint: Write $A = R\Lambda R^{-1} = R\sqrt{\Lambda}\sqrt{\Lambda}R^{-1} = (R\sqrt{\Lambda}R^{-1})^2$.