Another perspective on the matrix exponential

Review. We achieved the milestone to introduce a **matrix exponential** in such a way that we can treat a system of DEs, say $\boldsymbol{y}' = M\boldsymbol{y}$ with $\boldsymbol{y}(0) = \boldsymbol{c}$, just as if the matrix M was a number: namely, the unique solution is simply $\boldsymbol{y} = e^{Mx}\boldsymbol{c}$.

The price to pay is that the matrix e^{Mx} requires some work to actually compute (and proceeds by first determining a different matrix solution $\Phi(x)$ using eigenvectors and eigenvalues). We offer below another way to think about e^{Mx} (using Taylor series).

(exponential function) e^x is the unique solution to $y' = y$, $y(0) = 1$.

From here, it follows that $e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \ldots$.

The latter is the Taylor series for e^x at $x = 0$ that we have seen in Calculus II.

Important note. We can actually construct this infinite sum directly from $y' = y$ and $y(0) = 1$.

Indeed, observe how each term, when differentiated, produces the term before it. For instance, $\frac{\mathrm{d}}{\mathrm{d}x}\frac{x^3}{3!} = \frac{x^2}{2!}$.

Review. We defined the **matrix exponential** e^{Mx} as the unique matrix solution to the IVP

$$\boldsymbol{y}' = M\boldsymbol{y}, \quad \boldsymbol{y}(0) = I.$$

We next observe that we can also make sense of the matrix exponential e^{Mx} as a power series.

Theorem 74. Let M be $n \times n$. Then the **matrix exponential** satisfies

$$e^M = I + M + \frac{1}{2!}M^2 + \frac{1}{3!}M^3 + \ldots$$

Proof. Define $\Phi(x) = I + Mx + \frac{1}{2!}M^2x^2 + \frac{1}{3!}M^3x^3 + \ldots$

$$\begin{aligned}\Phi'(x) &= \frac{\mathrm{d}}{\mathrm{d}x}\left[I + Mx + \frac{1}{2!}M^2x^2 + \frac{1}{3!}M^3x^3 + \ldots\right]\\ &= 0 + M + M^2x + \frac{1}{2!}M^3x^2 + \ldots = M\Phi(x).\end{aligned}$$

Clearly, $\Phi(0) = I$. Therefore, $\Phi(x)$ is the fundamental matrix solution to $\boldsymbol{y}' = M\boldsymbol{y}$, $\boldsymbol{y}(0) = I$.

But that's precisely how we defined e^{Mx} earlier. It follows that $\Phi(x) = e^{Mx}$. Now set $x = 1$. □

Example 75. If $A = \begin{bmatrix} 2 & 0 \\ 0 & 5 \end{bmatrix}$, then $A^{100} = \begin{bmatrix} 2^{100} & 0 \\ 0 & 5^{100} \end{bmatrix}$.

Example 76. If $A = \begin{bmatrix} 2 & 0 \\ 0 & 5 \end{bmatrix}$, then $e^A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 2 & 0 \\ 0 & 5 \end{bmatrix} + \frac{1}{2!}\begin{bmatrix} 2^2 & 0 \\ 0 & 5^2 \end{bmatrix} + \cdots = \begin{bmatrix} e^2 & 0 \\ 0 & e^5 \end{bmatrix}$.

Clearly, this works to obtain e^D for every diagonal matrix D.

In particular, for $Ax = \begin{bmatrix} 2x & 0 \\ 0 & 5x \end{bmatrix}$, $e^{Ax} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 2x & 0 \\ 0 & 5x \end{bmatrix} + \frac{1}{2!}\begin{bmatrix} (2x)^2 & 0 \\ 0 & (5x)^2 \end{bmatrix} + \cdots = \begin{bmatrix} e^{2x} & 0 \\ 0 & e^{5x} \end{bmatrix}$.

The following is a preview of how the matrix exponential deals with repeated characteristic roots.

Example 77. Determine e^{Ax} for $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$.

Solution. If we compute eigenvalues, we find that we get $\lambda = 0, 0$ (multiplicity 2) but there is only one 0-eigenvector (up to multiples). This means we are stuck with this approach—however, see next extra section how we could still proceed.

The key here is to observe that $A^2 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$. It follows that $e^{Ax} = I + Ax = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 0 & x \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & x \\ 0 & 1 \end{bmatrix}$.

Armin Straub
straub@southalabama.edu

Extra: The case of repeated eigenvalues with too few eigenvectors

Review. To construct a fundamental matrix solution $\Phi(x)$ to $\boldsymbol{y}' = M\boldsymbol{y}$, we compute eigenvectors: Given a λ-eigenvector $\boldsymbol{v}$, we have the corresponding solution $\boldsymbol{y}(x) = \boldsymbol{v}e^{\lambda x}$.
If there are enough eigenvectors, we can collect these as columns to obtain $\Phi(x)$.
The next example illustrates how to proceed if there are not enough eigenvectors.

In that case, instead of looking only for solutions of the type $\boldsymbol{y}(x) = \boldsymbol{v}e^{\lambda x}$, we also need to look for solutions of the type $\boldsymbol{y}(x) = (\boldsymbol{v}x + \boldsymbol{w})e^{\lambda x}$. This can only happen if an eigenvalue is a repeated root of the characteristic polynomial.

Example 78. Let $M = \begin{bmatrix} 8 & 4 \\ -1 & 4 \end{bmatrix}$.

(a) Determine the general solution to $\boldsymbol{y}' = M\boldsymbol{y}$.

(b) Determine a fundamental matrix solution to $\boldsymbol{y}' = M\boldsymbol{y}$.

(c) Compute e^{Mx}.

(d) Solve the initial value problem $\boldsymbol{y}' = M\boldsymbol{y}$ with $\boldsymbol{y}(0) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$.

Solution.

(a) We determine the eigenvectors of M. The characteristic polynomial is:

$\det(M - \lambda I) = \det\left(\begin{bmatrix} 8-\lambda & 4 \\ -1 & 4-\lambda \end{bmatrix}\right) = (8-\lambda)(4-\lambda) + 4 = \lambda^2 - 12\lambda + 36 = (\lambda - 6)(\lambda - 6)$

Hence, the eigenvalues are $\lambda = 6, 6$ (meaning that 6 has multiplicity 2).

- To find eigenvectors $\boldsymbol{v}$ for $\lambda = 6$, we need to solve $\begin{bmatrix} 2 & 4 \\ -1 & -2 \end{bmatrix}\boldsymbol{v} = \boldsymbol{0}$.
  Hence, $\boldsymbol{v} = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$ is an eigenvector for $\lambda = 6$. There is no independent second eigenvector.

- We therefore search for a solution of the form $\boldsymbol{y}(x) = (\boldsymbol{v}x + \boldsymbol{w})e^{\lambda x}$ with $\lambda = 6$.
  $\boldsymbol{y}'(x) = (\lambda\boldsymbol{v}x + \lambda\boldsymbol{w} + \boldsymbol{v})e^{\lambda x} \overset{!}{=} M\boldsymbol{y} = (M\boldsymbol{v}x + M\boldsymbol{w})e^{\lambda x}$
  Equating coefficients of x, we need $\lambda\boldsymbol{v} = M\boldsymbol{v}$ and $\lambda\boldsymbol{w} + \boldsymbol{v} = M\boldsymbol{w}$.
  Hence, $\boldsymbol{v}$ must be an eigenvector (which we already computed); we choose $\boldsymbol{v} = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$.
  [Note that any multiple of $\boldsymbol{y}(x)$ will be another solution, so it doesn't matter which multiple of $\begin{bmatrix} -2 \\ 1 \end{bmatrix}$ we choose.]
  $\lambda\boldsymbol{w} + \boldsymbol{v} = M\boldsymbol{w}$ or $(M - \lambda)\boldsymbol{w} = \boldsymbol{v}$ then becomes $\begin{bmatrix} 2 & 4 \\ -1 & -2 \end{bmatrix}\boldsymbol{w} = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$.
  One solution is $\boldsymbol{w} = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$. [We only need one.]

Hence, the general solution is $C_1\begin{bmatrix} -2 \\ 1 \end{bmatrix}e^{6x} + C_2\left(\begin{bmatrix} -2 \\ 1 \end{bmatrix}x + \begin{bmatrix} -1 \\ 0 \end{bmatrix}\right)e^{6x}$.

(b) The corresponding fundamental matrix solution is $\Phi = \begin{bmatrix} -2e^{6x} & -(2x+1)e^{6x} \\ e^{6x} & x e^{6x} \end{bmatrix}$.

(c) Note that $\Phi(0) = \begin{bmatrix} -2 & -1 \\ 1 & 0 \end{bmatrix}$, so that $\Phi(0)^{-1} = \begin{bmatrix} 0 & 1 \\ -1 & -2 \end{bmatrix}$. It follows that

$$e^{Mx} = \Phi(x)\Phi(0)^{-1} = \begin{bmatrix} -2e^{6x} & -(2x+1)e^{6x} \\ e^{6x} & x e^{6x} \end{bmatrix}\begin{bmatrix} 0 & 1 \\ -1 & -2 \end{bmatrix} = \begin{bmatrix} (2x+1)e^{6x} & 4x e^{6x} \\ -x e^{6x} & -(2x-1)\, e^{6x} \end{bmatrix}.$$

(d) The solution to the IVP is $\boldsymbol{y}(x) = e^{Mx}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} (2x+1)e^{6x} & 4x e^{6x} \\ -x e^{6x} & -(2x-1)\, e^{6x} \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} (2x+1)e^{6x} \\ -x e^{6x} \end{bmatrix}$.

Phase portraits and phase plane analysis

Our goal is to visualize the solutions to systems of equations. This works particularly well in the case of systems of two differential equations.

A system of two differential equations that can be written as

$$\frac{\mathrm{d}x}{\mathrm{d}t} = f(x,y)$$
$$\frac{\mathrm{d}y}{\mathrm{d}t} = g(x,y)$$

is called **autonomous** because it doesn't depend on the independent variable t.

Comment. Can you show that if $x(t)$ and $y(t)$ are a pair of solutions, then so is the pair $x(t+t_0)$ and $y(t+t_0)$?

We can visualize solutions to such a system by plotting the points $(x(t), y(t))$ for increasing values of t so that we get a curve (and we can attach an arrow to indicate the direction we're flowing along that curve). Each such curve is called the **trajectory** of a solution.

Even better, we can do such a **phase portrait** without solving to get a formula for $(x(t), y(t))$! That's because we can combine the two equations to get

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{g(x,y)}{f(x,y)},$$

which allows us to make a slope field! If a trajectory passes through a point (x,y), then we know that the slope at that point must be $\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{g(x,y)}{f(x,y)}$.

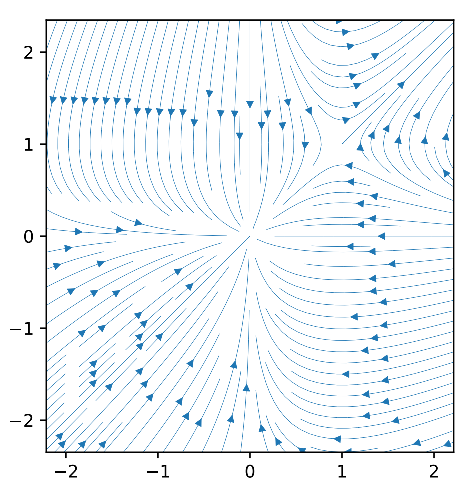
This allows us to sketch trajectories. However, it does not tell us everything about the corresponding solution $(x(t), y(t))$ because we don't know at which times t the solution passes through the points on the curve.

However, we can visualize the speed with which a solution passes through the trajectory by attaching to a point (x,y) not only the slope $\frac{g(x,y)}{f(x,y)}$ but the vector $\begin{bmatrix} f(x,y) \\ g(x,y) \end{bmatrix}$. That vector has the same direction as the slope but it also tells us in which direction we are moving and how fast (by its magnitude).

Example 79. Sketch some trajectories for the system $\frac{\mathrm{d}x}{\mathrm{d}t} = x\cdot(y-1)$, $\frac{\mathrm{d}y}{\mathrm{d}t} = y\cdot(x-1)$.

Solution. Let's look at the point $(x,y) = (2,-1)$, for instance. Then the DEs tell us that $\frac{\mathrm{d}x}{\mathrm{d}t} = x\cdot(y-1) = -4$ and $\frac{\mathrm{d}y}{\mathrm{d}t} = y\cdot(x-1) = -1$. We therefore attach the vector $\left(\frac{\mathrm{d}x}{\mathrm{d}t}, \frac{\mathrm{d}y}{\mathrm{d}t}\right) = (-4,-1)$ to $(x,y) = (2,-1)$.

Note that if we use $\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{y\cdot(x-1)}{x\cdot(y-1)}$ directly, we find the slope $\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{-1}{-4} = \frac{1}{4}$. This is slightly less information because it doesn't tell us that we are moving "left and down" as the arrows in the following plot indicate:



Armin Straub
straub@southalabama.edu

Comment. In this example, we can solve the slope-field equation $\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{y(x-1)}{x(y-1)}$ using separation of variables. Do it! We end up with the implicit solutions $y - \ln|y| = x - \ln|x| + C$.

If we plot these curves for various values of C, we get trajectories in the plot above. However, note that none of this solving is needed for plotting by itself.

Sage. We can make Sage create such phase portraits for us!

```
>>> x,y = var('x y')
>>> streamline_plot((x*(y-1),y*(x-1)), (x,-3,3), (y,-3,3))
```

Equilibrium solutions

(x_0, y_0) is an **equilibrium point** of the system $\frac{\mathrm{d}x}{\mathrm{d}t} = f(x,y)$, $\frac{\mathrm{d}y}{\mathrm{d}t} = g(x,y)$ if

$$f(x_0, y_0) = 0 \quad \text{and} \quad g(x_0, y_0) = 0.$$

In that case, we have the **constant (equilibrium) solution** $x(t) = x_0$, $y(t) = y_0$.

Comment. Equilibrium points are also called **critical points** (or stationary points or rest points).

In a phase portrait, the equilibrium solutions are just a single point.

Recall that every other solution $(x(t), y(t))$ corresponds to a curve (parametrized by t), called the **trajectory** of the solution (and we can adorn it with an arrow that indicates the direction of the "flow" of the solution).

We can learn a lot from how solutions behave near equilibrium points.

An equilibrium point is called:

- **stable** if all nearby solutions remain close to the equilibrium point;
- **asymptotically stable** if all nearby solutions remain close and "flow into" the equilibrium;
- **unstable** if it is not stable (some nearby solutions "flow away" from the equilibrium).

Comment. Note that asymptotically stable is a stronger condition than stable. A typical example of a stable, but not asymptotically stable, equilibrium point is one where nearby solutions loop around the equilibrium point without coming closer to it.

Advanced comment. For asymptotically stable, we kept the condition that nearby solutions remain close because there are "weird" instances where trajectories come arbitrarily close to the equilibrium, then "flow away" but eventually "flow into" (this would constitute an unstable equilibrium point).

Example 80. (cont'd) Consider again the system $\frac{\mathrm{d}x}{\mathrm{d}t} = x \cdot (y-1)$, $\frac{\mathrm{d}y}{\mathrm{d}t} = y \cdot (x-1)$.

(a) Determine the equilibrium points.

(b) Using the phase portrait from Example 79, classify the stability of each equilibrium point.

Solution.

(a) We solve $x(y-1) = 0$ (that is, $x = 0$ or $y = 1$) and $y(x-1) = 0$ (that is, $x = 1$ or $y = 0$).
We conclude that the equilibrium points are $(0,0)$ and $(1,1)$.

(b) $(0,0)$ is asymptotically stable (because all nearby solutions "flow into" $(0,0)$).
$(1,1)$ is unstable (because some nearby solutions "flow away" from $(1,1)$).
Comment. We will soon learn how to determine stability without the need for a plot.
Comment. If you look carefully at the phase portrait near $(1,1)$, you can see that certain solutions get attracted at first to $(1,1)$ and then "flow away" at the last moment. This suggests that there is a single trajectory which actually "flows into" $(1,1)$. This constellation is typical and is called a **saddle point**.