Solving differential equations with eigenvalues, matrix exponentials and diagonalization: $u' = Au$

Kevin Jia

12 December 2020

1 Solving 1st and 2nd order linear differential equations using $u' = Au$

a) Take any differential equation that satisfies the requirements of the section title (our example will be $y'' + ay' + by = 0$).

b) Using your equation, you can substitute y for $e^{\lambda t}$. Below it is applied to our example. This substitution should allow you to isolate all λ terms and to solve for it in future steps.

$$y'' + ay' + by = 0 \tag{1}$$

$$(e^{\lambda t})'' + a(e^{\lambda t})' + b(e^{\lambda t}) = 0 \tag{2}$$

$$\lambda^2 e^{\lambda t} + a\lambda e^{\lambda t} + be^{\lambda t} = 0 \tag{3}$$

$$e^{\lambda t}(\lambda^2 + a\lambda + b) = 0 \tag{4}$$

$$\implies \lambda^2 + a\lambda + b = 0 \text{ } (e^{\lambda t} \text{ cannot be } 0) \tag{5}$$

c) Let $u\colon \mathbb{R} \longrightarrow \mathbb{R}^2$ and u be a function of t - $u(t)$. Set u to $\begin{bmatrix} y \\ y' \end{bmatrix}$.

d) We now have an eigenvalue equation that can solve for y. From this you can derive the $n \times n$ matrix A in $u' = Au$. Deriving A requires that you take into account the original differential equation itself, as well as some obviously common calculus knowledge. After you have found A and u, you have built the framework $u' = Au$ from which we will work off of to solve your original differential equation.

$$A = \begin{bmatrix} 0 & 1 \\ -b & -a \end{bmatrix} \tag{6}$$

$$u' = Au \tag{7}$$

$$\begin{bmatrix} y \\ y' \end{bmatrix}' = \begin{bmatrix} 0 & 1 \\ -b & -a \end{bmatrix} \begin{bmatrix} y \\ y' \end{bmatrix} \tag{8}$$

$$\begin{bmatrix} y' \\ y'' \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -b & -a \end{bmatrix} \begin{bmatrix} y \\ y' \end{bmatrix} \tag{9}$$

$$\begin{bmatrix} y' \\ y'' \end{bmatrix} = \begin{bmatrix} y' \\ -ay' - by \end{bmatrix} \tag{10}$$

$$\implies y' = y' \tag{11}$$

$$\implies y'' = -ay' - by \tag{12}$$

The two implied equations that are directly above can also be derived from the original differential equation itself, as said earlier in step **d** (the exception here is the first one, although the first one is pretty trivial).

$$y'' + ay' + by = 0 \tag{13}$$

$$\implies y'' = -ay' - by \tag{14}$$

$y' = y'$ is trivial though cannot be derived from the original differential equation itself.

e) We know that u is going to be of the form

$$\sum_{i=1}^{n} c_i e^{\lambda_i t} x_i \tag{15}$$

Because $Ax = \lambda x$, then $u = e^{\lambda t} x$ will solve $u' = Au$, where λ is an eigenvalue of A and where x is an eigenvector of A. Differentiation is linear, and therefore the form of u written in equation **15** is correct.

f) Now we can solve for all eigenvalues using the final equation you got from step **b**. Here, in our example we solve the quadratic $\lambda^2 + a\lambda + b = 0$ to get solutions λ_1 and λ_2.

g) Now that you have your solutions λ_1 through to λ_n, you can calculate the solution u to $u' = Au$ from equation **15**. Our solution to $u' = Au$ given our example differential equation is $c_1 e^{\lambda_1 t} x_1 + c_2 e^{\lambda_2 t} x_2$. With your eigenvalues from step **f**, you can work out your eigenvectors x_1 through to x_n using $(A - \lambda I)x = 0$ - these will be useful later on if you are given $u(0)$. For our example,

$$x_1 = \begin{bmatrix} 1 \\ \lambda_1 \end{bmatrix} \tag{16}$$

$$x_2 = \begin{bmatrix} 1 \\ \lambda_2 \end{bmatrix} \tag{17}$$

h) If you are given a $u(0)$, then you can work out the constants c_1 through to c_n in your final answer for u.

$$u(0) = \sum_{i=1}^{n} c_i x_i \tag{18}$$

Equation **18** is trivial, and using your given $u(0)$ vector and your eigenvectors from step **g**, you can work out numerical values for your constants c_1 through to c_n, and finalize your answer of u.

i) Now that we've messed around with $u' = Au$ for most of the document, it is time to return to our original aim - to solve our up-to-2nd-order linear differential equation. Using your sum notation solution for u and also what you originally set u to be in step **c**, you can find y by matching components. For our example,

$$u = c_1 e^{\lambda_1 t} \begin{bmatrix} 1 \\ \lambda_1 \end{bmatrix} + c_2 e^{\lambda_2 t} \begin{bmatrix} 1 \\ \lambda_2 \end{bmatrix} = \begin{bmatrix} c_1 e^{\lambda_1 t} + c_2 e^{\lambda_2 t} \\ c_1 \lambda_1 e^{\lambda_1 t} + c_2 \lambda_1 e^{\lambda_2 t} \end{bmatrix} = \begin{bmatrix} y \\ y' \end{bmatrix} \tag{19}$$

$$\implies y = c_1 e^{\lambda_1 t} + c_2 e^{\lambda_2 t} \tag{20}$$

We now have our solution to our differential equation $y'' + ay + by = 0$:
y = $c_1 e^{\lambda_1 t} + c_2 e^{\lambda_2 t}$, and you have your solution to your differential equation.

2 Applying matrix exponentials to $u' = Au$ (and diagonalization where possible)

a) We will repeat steps the section **1** steps **a** and **c**, and the part of deriving A from step **d**, but this time without an example differential equation.

b) We will now define the matrix exponential. The definition will use the Maclaurin expansion of e^x. For now we will assume diagonalizable matrices (we will deal with non-diagonalizable A matrices later in the next section), and recall from section **1** that $u\colon \mathbb{R} \longrightarrow \mathbb{R}^2$, and that u is also a function of t - $u(t)$.

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} ... \tag{21}$$

$$\implies e^A = I + A + \frac{A^2}{2!} + \frac{A^3}{3!} ... \text{ (recall that } A^0 = I) \tag{22}$$

$$e^x = \sum_{i=0}^{\infty} \frac{x^i}{i!} \text{ (note the 0-based indexing)} \tag{23}$$

$$\implies e^A = \sum_{i=0}^{\infty} \frac{A^i}{i!} \text{ (this is our final definition)} \tag{24}$$

$$\implies e^{At} = \sum_{i=0}^{\infty} \frac{(At)^i}{i!} \tag{25}$$

c) We can take the form of $u(0)$ from equation **18**. From this form of $u(0)$ we can derive that (let x_n be the nth eigen**vector** of A),

$$u(0) = \begin{bmatrix} x_1 \cdots x_n \end{bmatrix} \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} \tag{26}$$

$$\implies u(0) = Xc \tag{27}$$

d) Using equation **27** we can now "matricise" our original form for u that came from equation **15**,

$$u = \sum_{i=1}^{n} c_i e^{\lambda_i t} x_i \tag{28}$$

$$u = \begin{bmatrix} x_1 \cdots x_n \end{bmatrix} \left(\sum_{i=1}^{n} e^{\lambda_i t} \right) \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} \tag{29}$$

$$u = X \left(\sum_{i=1}^{n} e^{\lambda_i t} \right) c \tag{30}$$

We can replace $\sum_{i=1}^{n} e^{\lambda_i t}$ with a diagonal $e^{\Lambda t}$ matrix,

$$u = X \begin{bmatrix} e^{\lambda_1 t} & & \huge 0 \\ & \ddots & \\ \huge 0 & & e^{\lambda_n t} \end{bmatrix} c \tag{31}$$

$$u = X \begin{bmatrix} e^{\lambda_1 t} & & \huge 0 \\ & \ddots & \\ \huge 0 & & e^{\lambda_n t} \end{bmatrix} X^{-1} X c \tag{32}$$

Let us briefly leave this flow of equations and go to diagonalization in the next step.

e) We will on this brief sidetrack, apply diagonalization to e^{At},

$$e^{At} = \sum_{i=0}^{\infty} \frac{(At)^i}{i!} \text{ (recall this from equation } \mathbf{25}\text{)} \tag{33}$$

$$e^{At} = \sum_{i=0}^{\infty} \frac{(X \Lambda X^{-1} t)^i}{i!} \tag{34}$$

$$e^{At} = X \left(\sum_{i=0}^{\infty} \frac{(\Lambda t)^i}{i!} \right) X^{-1} \text{ (recall that } (X \Lambda X^{-1})^k = X \Lambda^k X^{-1}\text{)} \tag{35}$$

If we matricise the sum expression of equation **35**, we get

$$\sum_{i=0}^{\infty}\frac{(\Lambda t)^i}{i!}=\sum_{i=0}^{\infty}\frac{\left(\begin{bmatrix}\lambda_1 & & 0\\ & \ddots & \\ 0 & & \lambda_n\end{bmatrix}t\right)^i}{i!} \tag{36}$$

$$\sum_{i=0}^{\infty}\frac{(\Lambda t)^i}{i!}=\sum_{i=0}^{\infty}\frac{\begin{bmatrix}(\lambda_1 t)^i & & 0\\ & \ddots & \\ 0 & & (\lambda_n t)^i\end{bmatrix}}{i!} \tag{37}$$

$$\sum_{i=0}^{\infty}\frac{(\Lambda t)^i}{i!}=\sum_{i=0}^{\infty}\begin{bmatrix}\frac{(\lambda_1 t)^i}{i!} & & 0\\ & \ddots & \\ 0 & & \frac{(\lambda_n t)^i}{i!}\end{bmatrix} \tag{38}$$

$$\implies e^{\Lambda t}=\begin{bmatrix}e^{\lambda_1 t} & & 0\\ & \ddots & \\ 0 & & e^{\lambda_n t}\end{bmatrix} \tag{39}$$

Recall that for equation **37**, any diagonal matrix D has $D^k=\begin{bmatrix}d_{11}^k & & 0\\ & \ddots & \\ 0 & & d_{nn}^k\end{bmatrix}$.

f) Returning to where we branched off in step **d**, if we substitute equation **39** into equation **32**,

$$u=X\begin{bmatrix}e^{\lambda_1 t} & & 0\\ & \ddots & \\ 0 & & e^{\lambda_n t}\end{bmatrix}X^{-1}Xc \tag{40}$$

$$\implies u=Xe^{\Lambda t}X^{-1}Xc \tag{41}$$

$$\implies u=X\sum_{i=0}^{\infty}\frac{(\Lambda t)^i}{i!}X^{-1}u(0) \tag{42}$$

$$\implies u=\sum_{i=0}^{\infty}\frac{(X\Lambda X^{-1}t)^i}{i!}u(0) \tag{43}$$

$$\implies u = \sum_{i=0}^{\infty} \frac{(At)^i}{i!} u(0) \tag{44}$$

$$\implies u = e^{At} u(0) \tag{45}$$

If we return to equation **41**, and simplify it

$$u = Xe^{\Lambda t} X^{-1} X c \tag{46}$$

$$u = Xe^{\Lambda t} c \tag{47}$$

This is remarkably similar to the u given in equation **15**, but now this u is in a matricised form, with the sum being concealed in $e^{\Lambda t}$. Solving for the eigenvalues and eigenvectors of A (steps **f** and **g** of section **1**), as well as utilising a given $u(0)$ (step **h** of section **1**) will solve u in $u' = Au$ with u in the form of equation **35**, and then the final step of matching coefficients (step **i** of section **1**) will solve y in the original differential equation.

3 Dealing with non-diagonalizable A matrices

a) Solving for u with a non-diagonalizable A gives **two** solutions for u - of which you can combine both into one (differentiation is linear). The first arrives from the same method in section 1, while the second involves carefully manipulating the Maclaurin expansion of matrix exponentials when we cannot turn e^{At} into $Xe^{\Lambda t}X^{-1}$. We will be detailing the second method in this section.

b) We can rewrite e^{At} (here λ is any eigenvalue of A),

$$e^{At} = e^{\lambda I t} e^{(A-\lambda I)t} \tag{48}$$

We can then begin manipulate the Maclaurin expansion of $e^{(A-\lambda I)t}$,

$$e^{(A-\lambda I)t} = \sum_{i=0}^{n} \frac{((A-\lambda I)t)^i}{i!} \tag{49}$$

$$e^{(A-\lambda I)t} = I + (A-\lambda I)t + \frac{((A-\lambda I)t)^2}{2} + \dots \frac{((A-\lambda I)t)^n}{n!} \tag{50}$$

$$e^{(A-\lambda I)t} = I + (A-\lambda I)t \tag{51}$$

The reason we can erase the rest of the expansion is because $(A-\lambda I)^2 = 0$ by the Cayley-Hamilton theorem - which states that given a characteristic polynomial $p_A(\lambda) = \lambda^n + \sum_{i=0}^{n-1} c_i \lambda^i$ (where c_0 through to c_{n-1} are constants in $\mathbb{R}$) of a matrix A, then $p_A(A) = A^n + \sum i = 0^{n-1} c_i A^i = 0$. 2×2 matrices with repeated eigenvalues definitely cannot be diagonalized because of a lack of independent eigenvectors, and all of this along with the Cayley-Hamilton theorem yields the aforementioned $(A-\lambda I)^2 = 0$ expression. Therefore we now have an expression for e^{At},

$$e^{At} = e^{\lambda I t} e^{(A-\lambda I)t} \tag{52}$$

$$e^{At} = e^{\lambda t}(I + (A - \lambda I)t) \tag{53}$$

c) If we do another quick sidetrack to explain how we got from $e^{\lambda It}$ to $e^{\lambda t}$ when progressing from equations **52** to **53**,

$$e^{\lambda It} = \sum_{i=0}^{n} \frac{(\lambda It)^i}{i!} \tag{54}$$

$$\implies e^{\lambda It} = \sum_{i=0}^{n} \frac{\begin{bmatrix} \lambda t & & 0 \\ & \ddots & \\ 0 & & \lambda t \end{bmatrix}^i}{i!} \tag{55}$$

$$\implies e^{\lambda It} = \sum_{i=0}^{n} \frac{\begin{bmatrix} (\lambda t)^i & & 0 \\ & \ddots & \\ 0 & & (\lambda t)^i \end{bmatrix}}{i!} \tag{56}$$

$$\implies e^{\lambda It} = \sum_{i=0}^{n} \begin{bmatrix} \frac{(\lambda t)^i}{i!} & & 0 \\ & \ddots & \\ 0 & & \frac{(\lambda t)^i}{i!} \end{bmatrix} \tag{57}$$

$$\implies e^{\lambda It} = \begin{bmatrix} e^{\lambda t} & & 0 \\ & \ddots & \\ 0 & & e^{\lambda t} \end{bmatrix} \tag{58}$$

Multiplying any matrix by cI (where c is a constant in $\mathbb{R}$) is the same is multiplying any matrix by just c, and therefore in our $e^{\lambda It}$ in the expression $e^{\lambda It}(I + (A - \lambda I)t)$ can be substituted it for $e^{\lambda t}$.

d) The expression for e^{At} in equation **52** now should make solving for u with a non-diagonalizable A trivial enough given that the process for solving for eigenvalues, eigenvectors and eventually the solution y to your original differential equation is already detailed in the end section 1, from step f up until the end of the section.