LINEAR SYSTEM OF ORDINARY DIFFERENTIAL EQUATIONS

ELEFTHERIOS GKIOULEKAS

The purpose of this handout is to summarize the theory of the matrix exponential. You may note that this is a different approach from the one proposed in your textbook. The main motivation for preferring the matrix exponential is that it provides a unified theory that can solve *all* problems, including all *forced* problems. For examples, see Bronson [1]. A detailed exposition of the theory is given by Apostol [2].

We want to solve a linear system of ordinary differential equations of the form

$$\frac{dx_a}{dt} = \sum_{b=1}^{n} A_{ab} x_b + f_b, \tag{1}$$

where f_b is a known forcing function, and $\mathbf{A} = [A_{ab}]$ is a matrix of constant coefficients. The essential result of the matrix exponentials theory is that the solution is given by

$$\mathbf{x}(t) = \exp(t\mathbf{A})\mathbf{x}(0) + \exp(t\mathbf{A}) \int_0^t \exp(-s\mathbf{A})\mathbf{f}(s), \tag{2}$$

where $\mathbf{x}(0)$ is the initial condition, which we know. The first term represents the homogeneous solution, whereas the second term represents the particular solution. The matrix exponential is given by

$$e^{\mathbf{A}} = \exp(\mathbf{A}) = \sum_{n=0}^{+\infty} \frac{\mathbf{A}^n}{n!}. \tag{3}$$

Note that this is a generalization of the Taylor series expansion of the standard exponential function. Naturally, $e^{\mathbf{A}}$ is a $n \times n$ matrix. The problem at hand is to evaluate the matrix exponential without having to evaluate an infinite sum, which is not practical.

1. Eigenvalues and eigenvectors

The evaluation of the matrix exponential requires the use of the concept of eigenvalues and eigenvectors. We say that a matrix $\mathbf{A}$ has eigenvalue λ with eigenvector $\mathbf{v}$, if the equation $\mathbf{Av} = \lambda\mathbf{v}$ is satisfied. To find the eigenvalues, you solve the following equation

$$\det(\mathbf{A} - \lambda\mathbf{I}) = 0. \tag{4}$$

Then, you substitute the numbers that you find to $\mathbf{Av} = \lambda\mathbf{v}$ and solve that to find the corresponding eigenvector.

The geometric interpretation is that the eigenvector represents a direction such that any vector that points towards that direction (or in the opposite direction) will not be *rotated* when you apply the matrix on it. It will only stretch or shrink, and the corresponding eigenvalue will tell you how much. If all the eigenvalues are distinct, then you're dealing with a very nice matrix, because you will also have corresponding eigenvectors that are orthogonal to each other. This means that every other vector can be decomposed as

$$\mathbf{x} = \sum_{a=1}^{n} c_a \mathbf{v}_a, \tag{5}$$

where the numbers c_a are the coordinates of the vector $\mathbf{x}$ using the eigenvectors $\mathbf{v}_a$ of the matrix $\mathbf{A}$ as a frame of reference. It follows that when you apply the matrix to an arbitrary vector, all that happens is that you're multiplying its coordinates with certain constant numbers, the eigenvalues λ_a , regardless of who this vector is. This only holds however *if you use the eigenvectors to define the coordinate system.* Furthermore, this may fail if you have repeated eigenvalues.

2. Matrix exponential

The most convenient technique for evaluating the matrix exponential was invented by Putzer. It is based on the following two claims. First, there will always exist numbers a_k such that the matrix exponential $e^{\mathbf{A}}$ can be written as a finite sum given by

$$e^{\mathbf{A}} = \sum_{k=0}^{n-1} a_k \mathbf{A}^k. \tag{6}$$

These numbers can be calculated by using the following claim: Suppose that we define a polynomial given by

$$p(x) = \sum_{k=0}^{n-1} a_k x^k. \tag{7}$$

Then, suppose that you know all the eigenvalues of the matrix $\mathbf{A}$. Then, for each eigenvalue λ you may obtain one equation as follows.

$$\begin{aligned} \lambda \text{ simple eigenvalue} &\Longrightarrow e^{\lambda} = p(\lambda) \\ \lambda \text{ double eigenvalue} &\Longrightarrow e^{\lambda} = p(\lambda) = p'(\lambda) \\ \lambda \text{ triple eigenvalue} &\Longrightarrow e^{\lambda} = p(\lambda) = p'(\lambda) = p''(\lambda) \end{aligned} \tag{8}$$

etc. Solving these equations together as a system, will give you the numbers a_k.

This is a straightforward calculation that always works. However, that doesn't mean that you will always have to do it. In certain cases, we can work it out in general and derive special results that will allow you to circumvent the need to solve the system of linear equations.

For example, if your matrix is nice enough to have n distinct eigenvalues, then the matrix exponential can be calculated from the following expressions.

$$\begin{aligned} \exp(t\mathbf{A}) &= \sum_{k=1}^{n} e^{\lambda_k t} \mathbf{L}_k \\ \mathbf{L}_k &= \prod_{j \in [n]-\{k\}} \frac{1}{\lambda_k - \lambda_j} (\mathbf{A} - \lambda_j \mathbf{I}) \end{aligned} \tag{9}$$

This result by itself covers most circumstances.

Most useful is the 2×2 case. In this case, we distinguish between two possibilities: that the eigenvalues are distinct, or that they are equal. Suppose that you have found the eigenvalues, and they are λ_1, λ_2. Then, if $\lambda_1 = \lambda_2 = \lambda$ then

$$\exp(t\mathbf{A}) = e^{\lambda t}(\mathbf{I} + t(\mathbf{A} - \lambda\mathbf{I})) \tag{10}$$

If $\lambda_1 \neq \lambda_2$ then

$$\begin{aligned} \exp(t\mathbf{A}) &= \frac{\mathbf{A} - \lambda_2\mathbf{I}}{\lambda_1 - \lambda_2}e^{\lambda_1 t} + \frac{\mathbf{A} - \lambda_1\mathbf{I}}{\lambda_2 - \lambda_1}e^{\lambda_2 t} \\ &= \frac{\lambda_1 e^{\lambda_2 t} - \lambda_2 e^{\lambda_1 t}}{\lambda_1 - \lambda_2}\mathbf{I} + \frac{e^{\lambda_1 t} - e^{\lambda_2 t}}{\lambda_1 - \lambda_2}\mathbf{A} \end{aligned} \tag{11}$$

This result is sufficient to solve all 2×2 problems. Keep in mind, that when you know the matrix exponential, that allows you to write down solutions for any kind of forcing functions.

3. The eigenvector method

Now let us consider the method used by the textbook. This method is expedient only when all the eigenvalues are distinct *and* you do not have forcing in your problem. It is also conceptually interesting.

Suppose that $\mathbf{v}$ is an eigenvector of the matrix $\mathbf{A}$ with eigenvalue λ. What happens if you apply the matrix $e^{t\mathbf{A}}$ on the eigenvector $\mathbf{v}$ instead? Using the definition of the matrix exponential, you may show that

$$\exp(t\mathbf{A})\mathbf{v} = e^{\lambda t}\mathbf{v}. \tag{12}$$

Suppose that you want to solve the unforced problem

$$\frac{dx_a}{dt} = \sum_{b=1}^{n} A_{ab}x_b. \tag{13}$$

The key is to write the initial condition as a linear combination of the eigenvectors

$$\mathbf{x}(0) = \sum_{a=1}^{n} c_a\mathbf{v}_a. \tag{14}$$

Using this, we may calculate the solution as follows:

$$\begin{aligned} \mathbf{x}(t) &= \exp(t\mathbf{A})\mathbf{x}(0) = \exp(t\mathbf{A})\sum_{a=1}^{n} c_a\mathbf{v}_a \\ &= \sum_{a=1}^{n} c_a \exp(t\mathbf{A})\mathbf{v}_a = \sum_{a=1}^{n} c_a \exp(\lambda_a t)\mathbf{v}_a \end{aligned} \tag{15}$$

What you get is the method of the textbook. Did you see how the matrix exponential disappeared? The idea is that *if* you don't have forcing, and *if* the eigenvalues *are* distinct, then you can circumvent the computation of the matrix exponential by evaluating the *eigenvectors* instead, which is an easier computation.

Don't forget that it is easy to diagnose whether your solutions are correct by substituting them back to the original equations.

References

[1] R. Bronson. *Differential Equations.* Scaum Outline Series McGrew–Hill, New York, 1994.
[2] T. M. Apostol. *Calculus.* John Wiley & Sons, New York, 1969.