THE MATRIX EXPONENTIAL FUNCTION

MICHAEL VANVALKENBURGH

This handout covers the material from class on November 10 (Wednesday), November 12 (Friday), and the beginning of November 15 (Monday).

1. Definition and Basic Properties of the Matrix Exponential

Recall that the exponential function $f(x) = e^x$ on the line has two equivalent definitions:

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

and

$$e^x = \lim_{N \to \infty} \left(1 + \frac{x}{N}\right)^N .$$

It is an amazing fact that these expressions also make sense when "plugging in matrices." That is, for any $A \in \mathbb{M}^{n,n}$,

$$e^A = \sum_{k=0}^{\infty} \frac{A^k}{k!} = \lim_{N \to \infty} \left(I + \frac{1}{N} A\right)^N \in \mathbb{M}^{n,n}.$$

(We adopt the convention that $A^0 = I$ for any matrix A, just as with the convention that $0^0 = 1$.)

Example. If the matrix A is *nilpotent* (that is, if A^k is the zero matrix for some positive integer k), then the above series expansion for e^A is actually a *finite sum.* Note in comparison that the only nilpotent scalar is the number 0; thus it is sometimes easier to compute the exponential of a matrix than it is to compute the exponential of a real number! For example, if

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix},$$

then we have $A^2 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$, and one can see from either expression above that $e^A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$.

Date: November 14, 2010.

Example. If $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, then for any $t \in \mathbb{R}$ we have

$$e^{tA} = \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix}.$$

In words, we say that $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ is the "infinitesimal generator" of rotation about the origin.

Now one may ask: Which properties of the scalar exponential are still true for the matrix exponential?

Theorem 1. *For any $A \in \mathbb{M}^{n,n}$ and any $t, s \in \mathbb{R}$, we have*

$$e^{(t+s)A} = e^{tA}e^{sA}.$$

Proof. We write out the right side:

$$e^{tA}e^{sA} = \left(\sum_{j=0}^{\infty} \frac{t^j A^j}{j!}\right)\left(\sum_{k=0}^{\infty} \frac{s^k A^k}{k!}\right).$$

The coefficient of A^N is

$$\sum_{m=0}^{N} \frac{t^m}{m!}\frac{s^{N-m}}{(N-m)!} = \frac{1}{N!}\sum_{m=0}^{N} \binom{N}{m} t^m s^{N-m} = \frac{1}{N!}(t+s)^N,$$

which proves the theorem. □

Corollary 1. *For any $A \in \mathbb{M}^{n,n}$, $t \in \mathbb{R}$, the matrix e^{tA} is invertible with inverse e^{-tA}.*

Theorem 2. *If $A, B \in \mathbb{M}^{n,n}$ commute (that is, if $AB = BA$), then for any $t \in \mathbb{R}$ we have*

$$e^{t(A+B)} = e^{tA}e^{tB}.$$

Proof. By the same type of calculation as before, the coefficient of t^N in $e^{tA}e^{tB}$ is

$$\frac{1}{N!}\sum_{k=0}^{N} \binom{N}{k} A^k B^{N-k}.$$

Since A and B commute by hypothesis, this is equal to

$$\frac{1}{N!}(A+B)^N,$$

which proves the theorem. □

However, not *all* properties of the scalar exponential hold for the matrix exponential, as the following example shows:

Note. If $A = \begin{pmatrix} 0 & -t \\ 0 & 0 \end{pmatrix}$ and $B = \begin{pmatrix} 0 & 0 \\ t & 0 \end{pmatrix}$, then

$$e^A = \begin{pmatrix} 1 & -t \\ 0 & 1 \end{pmatrix}, \qquad e^B = \begin{pmatrix} 1 & 0 \\ t & 1 \end{pmatrix}, \qquad e^{A+B} = \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix},$$

and

$$e^A e^B = \begin{pmatrix} 1 - t^2 & -t \\ t & 1 \end{pmatrix}.$$

So e^{A+B} and $e^A e^B$ are not necessarily equal.

2. A Concrete Expression for e^A

We can use the Jordan Canonical Form (JCF) of a matrix A to find a concrete expression for the matrix e^A. First of all, we state the Jordan Canonical Form Theorem without proof.

Theorem 3. *Every $A \in \mathbb{M}^{n,n}(\mathbb{C})$ is similar to a matrix of the form*

$$J = \begin{pmatrix} J_0 & 0 & 0 & \cdots & 0 \\ 0 & J_1 & 0 & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & 0 & \cdots & J_s \end{pmatrix},$$

with

$$J_0 = \begin{pmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & 0 & \cdots & \lambda_q \end{pmatrix}$$

and, for $i = 1, \ldots, s$,

$$J_i = \begin{pmatrix} \lambda_{q+i} & 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & \lambda_{q+i} & 1 & 0 & \cdots & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdots & \cdot & \cdot \\ 0 & 0 & 0 & 0 & \cdots & \lambda_{q+i} & 1 \\ 0 & 0 & 0 & 0 & \cdots & 0 & \lambda_{q+i} \end{pmatrix},$$

where the λ_j, $j = 1, \ldots, q+s$, are the (not necessarily distinct) eigenvalues of A.

From this very powerful theorem, we can easily derive the following facts (which we have already discussed, earlier in the semester):

Corollary 2. *The determinant of a matrix is the product of its eigenvalues, and the trace of a matrix is the sum of its eigenvalues. In both cases, the eigenvalues are repeated according to their algebraic multiplicities. In shorthand,* $\det A = \prod \lambda_j$ *and* $\mathrm{tr} A = \sum \lambda_j$.

To compute a concrete expression for e^A, say that A has Jordan Canonical Form J, as written above. So $A = PJP^{-1}$ for some invertible matrix P. Thus

$$\begin{aligned} e^{tA} &= e^{tPJP^{-1}} = Pe^{tJ}P^{-1} \\ &= P \begin{pmatrix} e^{tJ_0} & \cdots & 0 \\ 0 & \ddots & 0 \\ 0 & \cdots & e^{tJ_s} \end{pmatrix} P^{-1} \end{aligned}$$

Since J_0 is a diagonal matrix, it is easy to see that

$$e^{tJ_0} = \begin{pmatrix} e^{t\lambda_1} & \cdots & 0 \\ 0 & \ddots & 0 \\ 0 & \cdots & e^{t\lambda_q} \end{pmatrix}.$$

Hence it remains to find an expression for e^{tJ_i}, for $i = 1, \ldots, s$.

We write $J_i = \lambda_{q+i} I_{r_i} + Z_i$, where I_{r_i} is the $r_i \times r_i$ identity matrix and Z_i is the $r_i \times r_i$ matrix having zeros everywhere except for 1's on the superdiagonal. By Theorem 2 above we have

$$e^{tJ_i} = e^{t\lambda_{q+i}} e^{tZ_i},$$

and one can check that

$$e^{tZ_i} = \begin{pmatrix} 1 & t & \frac{t^2}{2!} & \cdots & \frac{t^{r_i-1}}{(r_i-1)!} \\ 0 & 1 & t & \cdots & \frac{t^{r_i-2}}{(r_i-2)!} \\ 0 & 0 & 1 & \cdots & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & 0 & \cdots & t \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}.$$

Putting this all together gives an explicit expression for e^{tA}. (See also Homework 9.)

3. The Determinant of a Matrix Exponential

In this section we compute $\det e^A$. Suppose that $A \in \mathbb{M}^{n,n}$ has eigenvalues $\{\lambda_1, \ldots, \lambda_n\}$, repeated according to algebraic multiplicity. Since the determinant is the product and the trace is the sum of the eigenvalues, and since the eigenvalues of e^A are e^{λ_j}, we should have

$$\det(e^A) = \prod e^{\lambda_j} = e^{\sum \lambda_j} = e^{\mathrm{tr} A}.$$

However, there is a subtle mistake in this argument: We need to make sure that the eigenvalues of e^A have the same algebraic multiplicities as the corresponding eigenvalues of A. To avoid this complication we give a somewhat different argument.

Theorem 4. *As $\epsilon \to 0$,*

$$\det(I + \epsilon A) = 1 + \epsilon \operatorname{tr}(A) + \mathcal{O}(\epsilon^2).$$

Proof. Let $\{\lambda_j\}$ be the eigenvalues of A. By considering the characteristic polynomials of A and $I + \epsilon\, A$, it is easy to see that the eigenvalues of $I + \epsilon\, A$ are $\{1 + \epsilon\lambda_j\}$, and that λ_j has the same algebraic multiplicity as $1 + \epsilon\lambda_j$. Thus

$$\begin{aligned} \det(I + \epsilon\, A) &= \prod_{j=1}^{n}(1 + \epsilon\lambda_j) \\ &= 1 + \epsilon\left(\sum_{j=1}^{n}\lambda_j\right) + \mathcal{O}(\epsilon^2) \\ &= 1 + \epsilon \operatorname{tr}(A) + \mathcal{O}(\epsilon^2). \end{aligned}$$

□

With this tool in hand, we come to the main theorem of this section:

Theorem 5. $\det(e^A) = e^{\operatorname{tr}(A)}$.

Proof.

$$\begin{aligned} \det(e^A) &= \det\left(\lim_{N\to\infty}\left(I + \frac{1}{N}A\right)^N\right) \\ &= \lim_{N\to\infty}\left(\det\left(I + \frac{1}{N}A\right)\right)^N \\ &= \lim_{N\to\infty}\left(1 + \frac{\operatorname{tr}(A)}{N} + \mathcal{O}\left(\frac{1}{N^2}\right)\right)^N \\ &= e^{\operatorname{tr}(A)}. \end{aligned}$$

□

Hence we have another proof of the following fact:

Corollary 3. *For any $A \in \mathbb{M}^{n,n}$, $e^A \in \mathbb{M}^{n,n}$ is an invertible matrix.*

4. An Interesting Question

What is the relationship between some family of matrices $\mathcal{F} \subset \mathbb{M}^{n,n}$ and the corresponding family of matrices $\mathcal{G} = \{e^A;\ A \in \mathcal{F}\}$?

5. Connection with Fundamental Matrices

Returning to systems of linear, first-order ordinary differential equations with constant coefficients, consider

$$\mathbf{x}'(t) = A\mathbf{x}(t). \tag{$*$}$$

(Here we use the boldface $\mathbf{x}$ to denote a vector.)

Similarly to the scalar case, this is solved by an exponential function; only now it is a *matrix* exponential function. Indeed, the important fact is that

$$\begin{aligned}\frac{d}{dt}\,e^{tA} &= \frac{d}{dt}\sum_{k=1}^{\infty}\frac{t^kA^k}{k!}\\ &= \sum_{k=1}^{\infty}\frac{t^{k-1}A^k}{(k-1)!}\\ &= Ae^{tA}.\end{aligned}$$

This shows that the solution of the initial value problem

$$\begin{cases}\mathbf{x}'(t) = A\mathbf{x}(t)\\ \mathbf{x}(t_0) = \mathbf{x}_0\end{cases}$$

is given by

$$\mathbf{x}(t) = e^{(t-t_0)A}\mathbf{x}_0.$$

Moreover, since e^{tA} has linearly independent columns (after all, it is *invertible*), the matrix

$$\mathbf{X}(t) = e^{tA}$$

is a fundamental matrix for the system $(*)$.

6. Variation of Parameters from this Point of View

To solve

$$\mathbf{x}'(t) = A\mathbf{x}(t) + \mathbf{f}(t),$$

we look for a solution of the form

$$\mathbf{x}_p(t) = e^{tA}\mathbf{v}(t),$$

where $\mathbf{v}(t)$ is to be determined. We simply plug this form into the equation and find that we want:

$$e^{tA}\mathbf{v}'(t) = \mathbf{f}(t).$$

Thus, by integrating,

$$\mathbf{x}_p(t) = e^{(t-t_0)A}\xi_0 + \int_{t_0}^{t} e^{(t-s)A}\mathbf{f}(s)\,ds$$

is the particular solution satisfying $\mathbf{x}_p(t_0) = \xi_0$.

7. An introduction to Dynamical Systems in the Plane

We now think of $e^{tA} : \mathbb{R}^n \to \mathbb{R}^n$ as a "propagator." It takes a vector $\mathbf{x}_0$ and moves it t units of time into the future. For any fixed $t \in \mathbb{R}$, this gives an invertible linear mapping from $\mathbb{R}^n$ to $\mathbb{R}^n$. It is interesting to note that this mapping preserves orientation (since $\det e^{tA} > 0$), and it preserves volume in $\mathbb{R}^n$ if and only if $\mathrm{tr}(A) = 0$ (or $t = 0$). As t changes, this gives an invertible "flow" in $\mathbb{R}^n$.

Given an initial point $\mathbf{x}_0$, we have a trajectory

$$\mathbb{R} \ni t \to e^{tA}\mathbf{x}_0 \in \mathbb{R}^n.$$

This is quite possibly a curved path as a function of t. Our problem now is to classify all possible "pictures" in $\mathbb{R}^2$. We call these pictures *phase portraits*.

it is easiest to see what happens when the initial point is an eigenvector of A. If $\mathbf{x}$ is an eigenvector of A with eigenvalue λ, then

$$e^{tA}\mathbf{x} = e^{t\lambda}\mathbf{x}.$$

Hence the point simply moves along a straight line as a function of t. Thus a first step is to find the eigenvalues and eigenvectors of the matrix A.

To simplify the situation, we consider the Jordan Canonical Form of A (which is easy to derive when $n = 2$). Then the flow given by e^{tA} is just the flow of e^{tJ}, but with respect to different coordinates.

For $A \in \mathbb{M}^{2,2}(\mathbb{R})$, $\det(A) \neq 0$, A is similar to one of the following:

(i) $\begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix}$, where $\mu < \lambda < 0$ or $0 < \mu < \lambda$.

(ii) $\begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix}$, where $\lambda > 0$ or $\lambda < 0$.

(iii) $\begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix}$, where $\mu < 0 < \lambda$.

(iv) $\begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$, where $\lambda > 0$ or $\lambda < 0$.

(v) $\begin{pmatrix} \sigma & \nu \\ -\nu & \sigma \end{pmatrix}$, where $\sigma, \nu \neq$ and ($\sigma > 0$ or $\sigma < 0$).

(vi) $\begin{pmatrix} 0 & \nu \\ -\nu & 0 \end{pmatrix}$, where $\nu \neq 0$.

In Homework 10 you will study the case (iv) in detail, and you will consider the case when $\det(A) = 0$ (not given here). I will sketch some of the phase portraits in class.

THE END.