Supplementary notes on matrix exponentials

This document is an addendum to [1], the notes *Introduction to Matrix Exponentials* posted on the Math 252 web page. The topics we discuss here are implicitly contained in those notes, but with a different emphasis.

Here is a brief guide to integrating the two documents. One should read [1] carefully up through the middle of page six; this will cover (i) the definition of matrix exponentials via power series, including several nice examples of how this definition works in practice, and (ii) the idea, whose treatment begins on the bottom of page six, of computing matrix exponentials using similarity to a matrix whose exponential may be easily computed:

$$A = S\Lambda S^{-1} \quad \Rightarrow \quad e^{At} = Se^{\Lambda t}S^{-1}. \tag{1}$$

Then one can profitably read the current document in parallel with portion of [1] from the middle of page six up to the bottom of page eleven (that is, up to the section labeled "Forcing Terms") as a brief summary of an alternative approach to applying (1) when A is either a diagonalizable matrix or a 2×2 non-diagonalizable matrix.

We consider then an $n\times n$ matrix A; we want to calculate the matrix exponential e^A or rather, since our purpose is the application to solutions of ordinary differential equations, the matrix e^{At}. We will write $\lambda_1,\ldots,\lambda_n$ for the n roots of the n^{th} order *characteristic polynomial* $\det(A-\lambda I)=0$; if this polynomial has repeated roots the any particular root will appear in the list $\lambda_1,\ldots,\lambda_n$ a number of times equal to its multiplicity as a root of the polynomial. Of course, the λ_k are the eigenvalues of A.

Case I. The first case we consider is that in which there exists a set of n linearly independent eigenvectors of A, say $\mathbf{v}^{(1)},\ldots,\mathbf{v}^{(n)}$ with $A\mathbf{v}^{(k)}=\lambda_k\mathbf{v}^{(k)}$; these will then automatically form a basis of the vector space $\mathbb{R}^n$. One can always find such a basis of eigenvectors when the eigenvalues $\lambda_1,\ldots,\lambda_n$ are all distinct; in that case there is an eigenvector $\mathbf{v}^{(k)}$ for each eigenvalue λ_k and these eigenvectors are automatically linearly independent.

Note: When there are repeated eigenvalues there may or may not be a basis of eigenvectors; the only way to find out is, for each eigenvalue λ, to solve for $\mathbf{v}$ in the equation $(A-\lambda I)\mathbf{v}$ and to see whether or not one finds m linearly independent solutions when λ is a root of the characteristic polynomial of multiplicity m (that is, whether or not $\operatorname{rank}(A-\lambda I)=n-m$).

Now given the eigenvectors $\mathbf{v}^{(1)},\ldots,\mathbf{v}^{(n)}$ we form the matrices $S=(\mathbf{v}^{(1)}\ \ldots\ \mathbf{v}^{(n)})$, whose columns are these eigenvectors, and Λ, the diagonal matrix with the eigenvalues $\lambda_1,\ldots,\lambda_k$ on the diagonal:

$$S = (\mathbf{v}^{(1)}\ \ldots\ \mathbf{v}^{(n)}) = \begin{pmatrix} v_1^{(1)} & \cdots & v_1^{(n)} \\ \vdots & \ddots & \vdots \\ v_n^{(1)} & \cdots & v_n^{(n)} \end{pmatrix}$$

$$\Lambda = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix} = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & \Large 0 & \\ & \Large 0 & \ddots & \\ & & & \lambda_n \end{pmatrix}$$

Since one can form a matrix product AB by applying A to each of the columns of B one has then that

$$AS = (A\mathbf{v}^{(1)} \ \dots \ A\mathbf{v}^{(n)}) = (\lambda_1\mathbf{v}^{(1)} \ \dots \ \lambda_1\mathbf{v}^{(n)}). \tag{2}$$

On the other hand, if we introduce the vectors $\mathbf{u}_1, \dots, \mathbf{u}_n$ which form the columns of the $n \times n$ identity matrix:

$$\mathbf{u}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \qquad \mathbf{u}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \qquad \mathbf{u}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \quad \dots, \quad \mathbf{u}_n = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix},$$

and use the fact that for any $n \times n$ matrix B, $B\mathbf{u}_k$ is the k^{th} column of B, we have

$$S\Lambda = S(\lambda_1\mathbf{u}_1 \ \dots \ \lambda_n\mathbf{u}_n) = (\lambda_1 S\mathbf{u}_1 \ \dots \ \lambda_n S\mathbf{u}_n) = (\lambda_1\mathbf{v}_1^{(1)} \ \dots \ \lambda_n\mathbf{v}^{(n)}). \tag{3}$$

Comparing (2) and (3) we finally have

$$AS = S\Lambda \quad \Rightarrow \quad A = S\Lambda S^{-1}. \tag{4}$$

This says that A is *diagonalizable*, that is, similar to a diagonal matrix. (One also sees easily that if a is diagonalizable then the columns of the diagonalizing matrix S must be eigenvectors of A, so that **digonalizability of A is equivalent to the existence of a basis consisting of of eigenvectors of A.**)

Once we have diagonalized A as in (4) we may calculate e^{At} easily. For as shown in [1] (see equation (9) there) we have from (4) that $e^{At} = Se^{\Lambda t}S^{-1}$. On the other hand, from the series definition of the matrix exponential we find that the exponential of any diagonal matrix D is just the diagonal matrix whose diagonal entries are the exponentials of the entries on the diagonal of D. Thus we have

$$e^{At} = Se^{\Lambda t}S^{-1} = S \begin{pmatrix} e^{\lambda_1 t} & & & \\ & e^{\lambda_2 t} & 0 & \\ & 0 & \ddots & \\ & & & e^{\lambda_n t} \end{pmatrix} S^{-1}.$$

Case II. When A is not diagonalizable one needs other methods to compute e^{At}. One may, for example, use the fact that every matrix A is similar to an upper triangular matrix: $A = SUS^{-1}$, with U upper triangular. This is the approach taken in [1]. Alternatively one may use that fact that every A is similar to a matrix in *Jordan canonical form*: $A = SKS^{-1}$ with K in Jordan form. Computation of the Jordan form can be rather complicated, but its exponentiation is easy.

When A is a 2×2 matrix the two methods are essentially equivalent; we will explain how they work, using notation similar to that we introduced in class. Suppose then that

A is a 2×2 matrix whose characteristic polynomial has λ_0 as a double root: $\det(A - \lambda I) = (\lambda - \lambda_0)^2$. Suppose further that $A - \lambda_0 I$ has rank one, so that when we solve $\det(A - \lambda_0 I) = \mathbf{v}$ for the eigenvector $\mathbf{v}$ we find only a one parameter family of solutions, all scalar multiples of some nonzero eigenvector $\mathbf{v}$. As we have discussed in class, one may then find a vector $\mathbf{w}$ such that $(A - \lambda_0 I)\mathbf{w} = \mathbf{v}$, i.e., $A\mathbf{w} = \lambda_0\mathbf{w} + \mathbf{v}$; $\mathbf{w}$ is a *generalized eigenvector*. Using the two vectors $\mathbf{v}$ and $\mathbf{w}$ we may now show that A is similar to a matrix K in Jordan form and thus construct e^{At}.

Let us introduce the matrix $S = (\mathbf{v}\,\mathbf{w})$ whose columns are the vectors $\mathbf{v}$ and $\mathbf{w}$. Then, just as in (2), we have

$$AS = (A\mathbf{v}\, A\mathbf{w}) = \big((\lambda_0\mathbf{v})\ (\lambda_0\mathbf{w} + \mathbf{v})\big). \tag{5}$$

On the other hand, let $K = \begin{pmatrix} \lambda_0 & 1 \\ 0 & \lambda_0 \end{pmatrix}$; then as in (3) we have

$$SK = S\big((\lambda_0\mathbf{u}_1)\ (\mathbf{u}_1 + \lambda_0\mathbf{u}_2)\big) = \big((\lambda_0 S\mathbf{u}_1)\ (S\mathbf{u}_1 + \lambda_0 S\mathbf{u}_2)\big) = \big((\lambda_0\mathbf{v})\ (\mathbf{v} + \lambda_0\mathbf{w})\big). \tag{6}$$

Comparing (5) and (6) we find that

$$AS = SK \quad \Rightarrow \quad A = SKS^{-1}. \tag{7}$$

On the other hand, from the definition of matrix exponentiation we find that $e^{Kt} = \begin{pmatrix} e^{\lambda_0 t} & te^{\lambda_0 t} \\ 0 & e^{\lambda_0 t} \end{pmatrix}$ (see the derivation of equation (12) on page nine of [1], taking $c = 1$ there). Thus from (7) we have

$$e^{At} = Se^{Kt}S^{-1} = S\begin{pmatrix} e^{\lambda_0 t} & te^{\lambda_0 t} \\ 0 & e^{\lambda_0 t} \end{pmatrix} S^{-1}.$$

This is the desired formula.

Note that the solution of the initial value problem $\mathbf{y}' = A\mathbf{y}$, $\mathbf{y}(0) = \mathbf{y}_0$, is $\mathbf{y}(t) = e^{At}\mathbf{y}_0$. But then if we write $\mathbf{c} = S^{-1}\mathbf{y}_0$ with $c = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}$ we have

$$\begin{aligned} e^{At}\mathbf{y}_0 = Se^{Kt}S^{-1}\mathbf{y}_0 &= S\begin{pmatrix} e^{\lambda_0 t} & te^{\lambda_0 t} \\ 0 & e^{\lambda_0 t} \end{pmatrix}\begin{pmatrix} c_1 \\ c_2 \end{pmatrix} \\ &= (\mathbf{v}\,\mathbf{w})\begin{pmatrix} c_1 e^{\lambda_0 t} + c_2 te^{\lambda_0 t} \\ c_2 e^{\lambda_0 t} \end{pmatrix} = (c_1 + c_2 t)e^{\lambda_0 t}\mathbf{v} + c_2 e^{\lambda_0 t}\mathbf{w}. \end{aligned}$$

This is the form of the solution that we have used earlier.

References

[1] Notes: *Introduction to Matrix Exponentials* posted on the Math 252 web page.