Math 252 — Spring 2000

Matrix Exponentials for Complex Eigenvalues

Introduction. Most of the examples arising in this course involve 2 by 2 matrices. There are special tricks in this case that allow easy computation of eigenvalues and eigenvectors, so such systems can be solved without excessive computation. It is also customary in the *Never-Never-Land* of textbook exercises and examination questions to arrange for real eigenvalues to be integers, so that eigenvalues can be found by recognizing the factors of the characteristic polynomial. However, it is also important to allow exercises leading to complex eigenvalues. In the simplest case, the eigenvalues will turn out to be $a+bi$ with a and b integers, but this turns out not to be simple enough to get all answers easily if you use methods developed for the real case. In order to appreciate the discussion that follows, you should work through *one* problem with complex eigenvalues. (Exercises 9 through 14 of section 3.4 of the textbook may be considered representative examples.)

The characteristic polynomial. The characteristic polynomial of a matrix A is defined in general to be $\det(A-\lambda I)$. In the 2-by-2 case, this has a simple form that can be derived by introducing individual variables for the matrix entries. Thus,

$$\det\begin{pmatrix} a-\lambda & b \\ c & d-\lambda \end{pmatrix} = (a-\lambda)(d-\lambda) - bc = \lambda^2 - (a+d)\lambda + (ad-bc).$$

The coefficient of λ^2 is 1, the constant term is $\det(A)$ — since that is what the definition reduces to when $\lambda = 0$ — and the only other coefficient is the negative of the sum of the diagonal entries. The sum of the diagonal entries of any matrix is called the *trace* of the matrix, and is surprisingly important.

The roots r_0 and r_1 of a polynomial $\lambda^2 - A\lambda + B$ are characterized by the polynomial being a multiple of $(\lambda - r_0)(\lambda - r_1)$. Expanding the product and comparing coefficients of λ^2 shows that the multiplier can only be 1. Then, comparing the other coefficients gives $r_0 + r_1 = A$ and $r_0 r_1 = B$. In the case in which B is an integer with only a few factors, one can find all ways of writing B as a product of two integers. If one of the sums is equal to A, these factors give the roots. This is not really useful, except for recognizing the equations that have small integer roots. (Although factoring plays an important role in elementary mathematics, its value depends on having memorized multiplication tables. In particular, no fast ways to factor large integers are known, a fact that forms the basis of some modern encryption methods.)

One method that always works is *Completing the square*, or the *Quadratic formula* obtained by applying the method to a general equation whose coefficients are independent parameters. The *method* is often better than the *formula* because it can exploit special properties of the coefficients to simplify calculations. This shows that every quadratic equation can be solved by extracting the square root of a single quantity computed from the coefficients. If this quantity is negative, there are no real roots, but a system of *complex numbers* has been created that gives a interpretation of these quantities that is consist with the familiar rules of algebra. Every complex number has a unique representation as $a + bi$ where a and b are real numbers and i is a square root of -1.

For 2-by-2 matrices, it is easily seen, by direct calculation, that

$$A^2 = \operatorname{tr}(A)A - \det(A),$$

which is a special case of a result known as the Cayley-Hamilton theorem.

Complex exponentials. Assuming that familiar rules of algebra and calculus extend to complex numbers, the expression

$$y = \cos t + i \sin t$$

has derivative

$$\frac{dy}{dt} = -\sin t + i \cos t = iy.$$

Thus, y should be Ce^{it} for some C. Evaluating at $t = 0$ gives $C = 1$. Combining this formula with the addition formula for exponentials leads to

$$e^{(a+bi)t} = a^{at}\big(\cos bt + i \sin bt\big).$$

If we get a solution of a differential equation in terms of complex exponentials, we can use this formula to get solutions in terms of familiar real functions with complex numerical coefficients. This general solution gives unique coefficients to satisfy any given complex initial conditions. If the initial conditions are real, the solution will turn out to be real. Any of the results that we obtain by exploiting this notational device can be checked in the given equation to show that the solution is correct. As long as results are checked, such speculative methods are safe.

One way to find the solution is to copy the process used for real eigenvalues. An eigenvector for the eigenvalue λ is any *nonzero* solution of

$$\begin{pmatrix} a-\lambda & b \\ c & d-\lambda \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

For example, one could take

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} b \\ \lambda - a \end{pmatrix},$$

which *obviously* makes the first entry of the product zero. The value of second entry is just the negative of the characteristic polynomial, so it will be zero precisely when λ is an eigenvalue. If $\lambda = r + si$, this method gives

$$\begin{pmatrix} b \\ \lambda - a \end{pmatrix} e^{\lambda t} = \begin{pmatrix} be^{rt}\cos st + ibe^{rt}\sin st \\ (r-a+si)e^{rt}\cos st + (-s+(r-a)i)e^{rt}\sin st \end{pmatrix}.$$

as one solution of the differential equation. (**Exercise.** Check that this satisfies the given equation.)

The complex conjugate of this solution is a solution corresponding to the eigenvalue that is the complex conjugate of λ. The general solution is a linear combination of the two special solutions we have. Finding the coefficient in this linear combination corresponding to given initial conditions leads to solving equations whose coefficients are the entries of the eigenvectors.

When matrices have complex entries, every step requires additional care, so even familiar calculations may become tedious. Following the steps of the method with the numbers used in exercises doesn't seem to introduce the simplifications that one finds when dealing with integer eigenvalues, so the best approach would seem to be to carry a symbolic calculation through to a conclusion that would give a formula to solve general problems. We will do better! A few interesting properties of the solution will be recorded that allow the solution to be guessed. Then, the verification that the guess is correct will be routine.

If a linear problem has solutions that are a complex expression and its conjugate, then the solution will also include the sum and difference of those expressions. The sum is twice the real part of the expression and the difference is $2i$ times the coefficient of i (this coefficient is called "the imaginary part" of the expression although it is a real expression) in the expression. Thus, the real and imaginary parts of a solution will be solutions, and the original solution is a combination with complex coefficient of these solutions. This gives a description in terms of real numbers or functions of the solutions of the original problem, and the real linear combinations of these solutions give all real solutions.

Applying this to our given differential equation, we see can write a basis for the space of real solutions. (**Exercise:** Do this.) We see that all solutions can be expressed as combinations of vector multiples of the functions $e^{rt}\cos st$ and $e^{rt}\sin st$ (though only some linear combinations are solutions).

All solutions can be found from e^{At}, which is a matrix whose first column solves the initial value problem $x(0) = 1$, $y(0) = 0$, and whose second column satisfies $x(0) = 0$, $y(0) = 1$. Alternatively, the *whole* matrix $X(t) = e^{At}$ is characterized by $dX(t)/dt = AX(t)$, $X(0) = I$. If we solve this problem, any initial value problem for this equation is solved by multiplying this matrix by the column of initial values (in that order).

The formula. At this point, we know that

$$e^{At} = Me^{rt}\cos st + Ne^{rt}\sin st,$$

where M and N are 2-by-2 matrices of real numbers. Putting $t = 0$ in this expression gives (immediately!) that $M = I$. We now need only find N.

The product rule of differential calculus builds the derivative of e^{At} from two parts which are essentially the cases corresponding to $r = 0$ and to $s = 0$, so these simpler cases will give parts of the formula we seek that can be assembled later. If $s = 0$, the only possibility is $N = 0$. That is, the real part of the action of A should be represented by a rI. This means that the part of A corresponding to si should be $A - rI$, which is a matrix of trace zero. The part of the Cayley-Hamilton Theorem that we have verified by direct computation shows that the square of such a matrix is a multiple of the identity. In the case of imaginary roots, the determinant of the matrix will be positive so that the square of the matrix will be a negative multiple of the identity. In particular, any matrix J of trace zero and determinant 1 seems to be trying to play the role of the special number i.

For such a matrix, we have

$$e^{Jt} = I\cos t + N\sin t$$

for some matrix N. Differentiating this gives

$$Je^{Jt} = -I\sin t + N\cos t$$

and evaluating this at $t = 0$ shows that the only possibility is $N = J$.

Any matrix of trace zero and positive determinant is a real multiple of such a matrix J. In particular, for our given matrix A with eigenvalues $r \pm si$,

$$A = rI + sJ$$

for some matrix J with $J^2 = -I$. In this case, we expect

$$e^{At} = Ie^{rt}\cos st + Je^{rt}\sin st. \qquad (E)$$

Having guessed the form of the answer, we can prove that (E) is correct by showing that the quantity on the right satisfies $dX(t)/dt = AX(t)$ and $X(0) = I$. (**Exercise:** Do this.) Since formula (E) looks exactly like Euler's equation $e^{ix} = \cos x + i\sin x$, it is easy to remember. To use it, we need only discover that the eigenvalues of A are $r + si$, and use this to compute $J = (A - rI)/s$

Direction of rotation. The equation

$$\begin{pmatrix} x \\ y \end{pmatrix} = (I \cos\theta + J \sin\theta) \begin{pmatrix} x_0 \\ y_0 \end{pmatrix},$$

for any given x_0 and y_0, gives parametric equations of an ellipse. Thus,

$$e^{At} \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$$

describes a path that combines multiplication by the scale factor e^{rt} and rotation-like motion around an ellipse. We would like to know whether this rotation-like motion is clockwise or counterclockwise.

Geometry suggests that the behavior of all vectors will be the same, so let us look at the the positive x-axis by setting $x_0 = 1$ and $y_0 = 0$. The positive (i.e., counterclockwise) direction is characterized by $y > 0$ for small t, but the value of y given by (E) is the positive quantity $e^{rt} \sin st$ times the lower left entry of the matrix J. If this entry of J is positive, the motion is counterclockwise; if negative, the motion is clockwise.

Anticlimax. This method was limited to 2-by-2 matrices since the pair of conjugate eigenvalues gave a unique way to construct J from A. To apply this to an n-by-n matrix A, one should first identify a plane spanned by the real and imaginary parts of the eigenvector corresponding to a complex eigenvalue of A. This plane is taken into itself by A. Hence, if S is a n-by-2 matrix whose columns are a basis for this plane, then there is a 2-by-2 matrix A_2 such that $AS = SA_2$. The action of A on this plane is given by the matrix A_2 and the method just described applies to A_2. This calculation is easy for any basis of the plane — no special basis is required. If $n = 3$, and there is one real eigenvalue λ_0, then the eignvector corresponding to λ_0 is the nullspace of the matrix $A - \lambda_0 I$, so it is found by solving a system of equations with this matrix of coefficients (there are three equations, but one will be revealed to be redundant in the course of solving the equations). The plane corresponding to the other eigenvalues (which we are assuming to be a complex conjugate pair) is the column space of $A - \lambda_0 I$, and we can take any two linearly independent columns as our basis.

The weak link in this is the computation of the characteristic polynomial. Accurate computation of a 3-by-3 determinant with some polynomial entries requires special care. There are robust ways to find characteristic polynomials that avoid the evaluation of determinates, but it would be too much af a digression to discuss them here.

Exercise. Solve the following initial value problem using formula (E). In more detail: (**1**) write the differential equation as $X' = AX$; (**2**) find the characteristic polynomial of A; (**3**) solve to obtain eigenvalues $r \pm si$; (**4**) write $A = rI + sJ$ and check that $J^2 = -I$; (**5**) write e^{At} from formula (E); (**6**) multiply by the column of initial values; (7) check that you have a solution.

$$\frac{dx}{dt} = 7x - 13y \qquad x(0) = 3$$
$$\frac{dy}{dt} = 2x - 3y \qquad y(0) = 1$$