Eigenvectors and eigenvalues

You can find approximations to solutions of differential equations by numerical methods, but you can't understand their true behaviour without some more theoretical approach. For differential equations with constant coefficients, this means looking at eigenvalues and eigenvectors of matrices. The link between the two is the exponential function applied to matrices.

1. The exponential of a matrix

If X is a matrix, then

$$e^X = I + X + \frac{X^2}{2} + \frac{X^3}{3!} + \cdots$$

This is a series that always converges, and at about the same rate as that for e^x if x is one of the entries in X.

Example. Say

$$X = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$$

Then

$$e^X = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} + (1/2!) \begin{bmatrix} 1 & 0 \\ 0 & 2^2 \end{bmatrix} + (1/3!) \begin{bmatrix} 1 & 0 \\ 0 & 2^3 \end{bmatrix} + (1/4!) \begin{bmatrix} 1 & 0 \\ 0 & 2^4 \end{bmatrix} + \cdots$$

$$= \begin{bmatrix} 1 + 1 + 1/2! + 1/3! + \cdots & 0 \\ 0 & 1 + 2 + 2^2/2! + 2^3/3! + \cdots \end{bmatrix}$$

$$= \begin{bmatrix} e & 0 \\ 0 & e^2 \end{bmatrix}$$

And in general if X is a diagonal matrix

$$X = \begin{bmatrix} x_1 & 0 & & \\ 0 & x_2 & 0 & \\ & & \cdots & \\ 0 & \cdots & 0 & x_n \end{bmatrix}$$

then

$$e^X = \begin{bmatrix} e^{x_1} & 0 & & \\ 0 & e^{x_2} & 0 & \\ & & \cdots & \\ 0 & \cdots & 0 & e^{x_n} \end{bmatrix}$$

This formula for the exponential of a diagonal matrix is itself rather special, but it can be combined with another one. Matrices can often be related to a diagonal matrix by means of its **eigenvectors**. I recall that an eignevector of a matrix is a non-zero vector v such that vA is a scalar multiple of v. In other words, $vA = cv$ for some constant c. The constant is called the corresponding **eigenvalue**. If A is an $n \times n$ matrix then usually we can find n linearly independent eigenvectors v_i. Let V be the matrix whose rows are the vectors v_i. Then the equations $v_i A = c_i v_i$ translate to

$$VA = CV$$

where C is a diagonal matrix whose entries are the c_i. Then

$$VAV^{-1} = C, \quad A = V^{-1}CV .$$

Here's where the next rule for exponentials comes in:

- If $A = V^{-1}CV$ then

$$e^A = V^{-1}e^C V .$$

This technique applies in particular to symmetric matrices, for which the eigenvalues are always real, and n linearly independent eigenvectors always exist.

2. Matrix exponentials and differential equations

If t is a single real variable and A a matrix, then the function

$$E(t) = e^{tA}$$

will give for each t a matrix whose entries are functions of t. Its derivative is the matrix-valued function we get just be taking the derivative of each entry. From the definition of e^{tA} in terms of the series it is easy to see that

$$\frac{d}{dt} e^{tA} = e^{At} A$$

jus as if A were a single number. The consequence of this is a method for solving systems of differential equations

$$Y' = YA, \quad Y(0) = Y_0 .$$

The solution is

$$Y(t) = Y_0 e^{At} .$$

This is the extension to systems of the solution of the single differential equation

$$y' = ay, \quad y(0) = y_0$$

by

$$y = y_0 e^{at} .$$

3. Finding an eigenvector for the largest eigenvalue

Suppose A to be an $n \times n$ symmetric matrix. There exist n linearly distinctly different eigenvectors. How do we find them? In these circumstances any vector v may be expressed as a linear combination of eigenvectors v_i with $v_i A = c_i v_i$):

$$v = a_1 v_1 + \cdots + a_n v_n$$

where say

$$|a_1| \geq |a_2| \geq \ldots |a_n|$$

so that a_1 is among the largest in magnitude. Suppose for simplicity in a first explanation that it is strictly the largest, that it is positive and that a_1 happens not to be 0. Then if A is applied to v we get

$$vA = a_1 c_1 v_1 + \cdots + a_n c_n v_n$$

and

$$vA^m = a_1 c_1^m v_1 + \cdots + a_n c_n^m v_n$$

and as $m \to \infty$ the size of the first vector grows in size compared to the others. Another way to say this is that the **direction** of vA^m becomes closer and closer to the direction of v_1 as m gets larger.

Two vectors u and v have the same direction if their normalizations $u/\|u\|$ and $v/\|v\|$ are the same. So *the vectors $vA^m/\|vA^m\|$ will converge to $v_1/\|v_1\|$.* In practice this works pretty well, and as we shall see it can be improved upon to work very well indeed. Here is a record of the calculation for

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}, \quad v = \begin{bmatrix} 1 & 0 \end{bmatrix} .$$

$$
\begin{array}{ll}
1.0000000 & 0.0000000 \\
0.0000000 & 1.0000000 \\
0.7071068 & 0.7071068 \\
0.4472136 & 0.8944272 \\
0.5547002 & 0.8320503 \\
0.5144957 & 0.8574929 \\
0.5299990 & 0.8479983 \\
0.5240974 & 0.8516583 \\
0.5263546 & 0.8502652 \\
0.5254929 & 0.8507980 \\
0.5258221 & 0.8505946 \\
0.5256963 & 0.8506723 \\
0.5257444 & 0.8506426 \\
0.5257260 & 0.8506540 \\
0.5257331 & 0.8506496 \\
0.5257304 & 0.8506513 \\
0.5257314 & 0.8506506 \\
0.5257310 & 0.8506508 \\
0.5257311 & 0.8506508 \\
0.5257311 & 0.8506508 \\
\end{array}
$$

This doesn't tell us what the eigenvalue is, but we can find it as the ratio of vA to v if v is an eigenvector.
If $c_1 < 0$? The directions flip sign every other iteartion.

4. Other eigenvectors

If v is an eigenvector for A then $vA = cv$, so if A isinvertible then $vA^{-1} = c^{-1}v$, and v is also an eigenvector for A^{-1}. If c is the smallest among the eigenvalues d then c^{-1} will be the largest among the d^{-1} so c^{-1} will be the largest among the eigenvalues for A^{-1}. Therefore if we calculate normalized vectors $vA^{-n}/\|vA^{-n}\|$ we can find an eigenvector for the smallest eigenvalue of A.

But what about all the ones in between? Suppose c_* to be a scalar. The eigenvectors of $A - c_*I$ are the same as those of A since if $vA = cv$ then

$$
v(A - c_*I) = vA - c_*v = (c - c_*)v
$$

so that $c - c_*$ is an eigenvalue of $A - c_*I$ for the same eigenvector. And then $(c - c_*)^{-1}$ is an eigenvector of $(A - c_*I)^{-1}$. Now if c_* is close to c then $c - c_*$ will be small and $(c - c_*)^{-1}$ will be large, and if it is close enough then it will be the largest among the eigenvalues of $A - c_*I$. So by making a lucky guess we can find the eigenvectors for $(A - c_*I)^{-1}$ by calculating and normalizing the $v(A - c_*I)^{-m}$ for large m. There are clever ways to make guesses by skill rather than by luck, but I won't describe those here. What I will point out is that the possible eigenvalues are sandwiched in bewteen the smallest and the largest, so you don't have an infinite range to work in. Very occasionally it will turn out that $A - c_*I$ has no inverse. This is luck indeeed, since then c_* must itself be an eigenvalue. But you then need to choose a slightly different c_* to find an eigenvector.

5. An example

Let

$$A = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 2 \end{bmatrix}.$$

If we apply A over again and normalize, we get the limit vector $[0.5, 0.70710677, 0.5]$. In fact, the exact value of the eigenvector is $[1/2, \sqrt{2}/2, 1/2]$.

The next step is to calculate A^{-1}. For large matrices the Gauss-Jordan procedure is the best way to do this, but for 3×3 matrices a simple formula is best. If $C = [a_{i,j}]$ is any $n \times n$ matrix, then its **adjoint** C^{adj} is the $n \times n$ matrix whose entry at position i, j is the determinant of the complementary $(n-1) \times (n-1)$ matrix, multiplied by $(-1)^{i+j}$. Thus the adjoint of

$$\begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{bmatrix}$$

is

$$\begin{bmatrix} \begin{vmatrix} a_{2,2} & a_{2,3} \\ a_{3,2} & a_{3,3} \end{vmatrix} & -\begin{vmatrix} a_{2,1} & a_{2,3} \\ a_{3,1} & a_{3,3} \end{vmatrix} & \begin{vmatrix} a_{2,1} & a_{2,2} \\ a_{3,1} & a_{3,2} \end{vmatrix} \\ -\begin{vmatrix} a_{1,2} & a_{1,3} \\ a_{3,2} & a_{3,3} \end{vmatrix} & \begin{vmatrix} a_{1,1} & a_{1,3} \\ a_{3,1} & a_{3,3} \end{vmatrix} & -\begin{vmatrix} a_{1,1} & a_{1,2} \\ a_{3,1} & a_{3,2} \end{vmatrix} \\ \begin{vmatrix} a_{1,2} & a_{1,3} \\ a_{2,2} & a_{2,3} \end{vmatrix} & -\begin{vmatrix} a_{1,1} & a_{1,3} \\ a_{2,1} & a_{2,3} \end{vmatrix} & \begin{vmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{vmatrix} \end{bmatrix}.$$

The inverse of A is the **transpose** of the adjoint, divided by the determinant of A. The transpose is what you get by flipping A with respect to its NW-SE diagonal.

For a 3×3 matrix the determinant is

$$a_{1,1}a_{2,2}a_{3,3} + a_{1,2}a_{2,3}a_{3,1} + a_{1,3}a_{2,1}a_{3,2} - a_{1,3}a_{2,2}a_{3,1} - a_{1,2}a_{2,1}a_{3,3} - a_{1,1}a_{3,2}a_{2,3}.$$

At any rate, when we apply the iteration and normalization to A^{-1} we get the eigenvector is $[1/2, -\sqrt{2}/2, 1/2]$. The maximum and minumum eigenvaluyes can be calculated to be $2 + \sqrt{2} \sim 3.4142$ and $2/(2+\sqrt{2}) \sim 0.5958$. What about the middle eigenvalue? If we try the technique mentioned above with $c = 2$ we get an error in inverting $A - cI$, which means that $c = 2$ is an eigenvalue. We were lucky! This will not find the eigenvector; to do this we must set c to something near 2, say 1.99. We then get $[\sqrt{2}/2, 0, -\sqrt{2}/2]$.