Matrix diagonalisation and its applications in chemistry

Last time, we saw how to find eigenvectors and eigenvalues of a matrix. There are some complications that arise for certain matrices, and we will need to come back to those in a later lecture. For now, though, let's look at a couple of examples where eigenvectors are useful in chemistry.

Diagonalisation of a matrix

The main purpose of finding eigenvectors is for *matrix diagonalisation.* This section explains what that is. We will consider a 3×3 matrix in the following, to simplify the discussion, but the results are easily extendable to any square matrix.

Suppose that the 3×3 matrix A has three eigenvectors, $\mathbf{v}_i$, with corresponding eigenvalues λ_i (with $i = 1, 2$ or 3). That is,

$$\mathsf{A}\mathbf{v}_i = \lambda_i \mathbf{v}_i.$$

From these three eigenvectors, we construct a 3×3 matrix S, using the eigenvectors as the columns.

$$\mathsf{S} = \begin{pmatrix} \uparrow & \uparrow & \uparrow \\ \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 \\ \downarrow & \downarrow & \downarrow \end{pmatrix}. \tag{5.1}$$

In fact, S is really just a row vector of the column vectors, i.e.

$$\mathsf{S} = \begin{pmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 \end{pmatrix}.$$

We will deliberately abuse the notation, by introducing the arrows in (5.1), to make it clear that the vectors are the columns of the matrix.

Then, it follows that

$$\mathsf{AS} = \mathsf{SD} \qquad \text{where} \qquad \mathsf{D} = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}. \tag{5.2}$$

This will take some explaining. First, let's convince ourselves it works, using an example.

Example 1. Let
$$A = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}.$$

The eigenvectors and eigenvalues are

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} \qquad \mathbf{v}_2 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \qquad \mathbf{v}_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

and

$$\lambda_1 = -1 \qquad \lambda_2 = -1 \qquad \lambda_3 = 2.$$

Construct the matrix S as defined in (5.1), and confirm that (5.2) holds.

We construct S by putting the three column vectors into a row vector:

$$S = \begin{pmatrix} 1 & 1 & 1 \\ -1 & 0 & 1 \\ 0 & -1 & 1 \end{pmatrix}.$$

Then, by matrix multiplication, we find that

$$AS = \begin{pmatrix} -1 & -1 & 2 \\ 1 & 0 & 2 \\ 0 & 1 & 2 \end{pmatrix}.$$

On the other hand, using the definition of D in (5.2), we have that

$$D = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{pmatrix},$$

from which it is easy to show that

$$SD = \begin{pmatrix} -1 & -1 & 2 \\ 1 & 0 & 2 \\ 0 & 1 & 2 \end{pmatrix}.$$

Therefore, (5.2) holds in this case.

The question is: why does it work in general?

Think about how the first column of AS was calculated in the example above. It was the dot products of rows of A with the first column of S.

Given how S was defined, the first column of AS is therefore dot products of the rows of A with the vector $\mathbf{v}_1$, i.e. exactly what we do when calculating $A\mathbf{v}_1$.

Now, given that $\mathbf{v}_1$ is an eigenvector of A, we know that $A\mathbf{v}_1 = \lambda_1 \mathbf{v}_1$. Putting this all together, the first column of AS must be $\lambda_1 \mathbf{v}_1$. Look back to the example above, and you will see that this is the case.

You will probably need to read those last three paragraphs a few times to understand what is going on. Read them slowly, a sentence at a time, and look back at the example above for a concrete example.

By extending the argument, we find that

$$\mathsf{AS} = \begin{pmatrix} \uparrow & \uparrow & \uparrow \\ \lambda_1 \mathbf{v}_1 & \lambda_2 \mathbf{v}_2 & \lambda_3 \mathbf{v}_3 \\ \downarrow & \downarrow & \downarrow \end{pmatrix},$$

that is, the 3×3 matrix where each column is $\lambda_i \mathbf{v}_i$. Again, look back at the example above, and check that you understand this.

That's the hardest part. Basically, when S is premultiplied by A, its columns get multiplied by the corresponding eigenvalues. But another way to multiply the columns of S by the eigenvalues is to postmultiply by D as defined in (5.2). This is just a simple matter of algebra to prove, because it's easy to verify that

$$\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} \begin{pmatrix} \alpha & 0 & 0 \\ 0 & \beta & 0 \\ 0 & 0 & \gamma \end{pmatrix} = \begin{pmatrix} \alpha a & \beta b & \gamma c \\ \alpha d & \beta e & \gamma f \\ \alpha g & \beta h & \gamma i \end{pmatrix}$$

Hence, we have confirmed (5.2) as required.

Now, **if S happens to be invertible**, we can premultiply both sides of (5.2) to get

$$\mathsf{S}^{-1}\mathsf{AS} = \mathsf{D} \tag{5.3}$$

This is the crucial, crucial result. We say that 'A has been *diagonalised* by a similarity transformation'. Remember that S has to be the matrix of eigenvectors, and also has to be invertible, for this formula to work.

If we can construct an invertible, square matrix of eigenvectors, S, we say that A is *diagonalisable*. Not all matrices are diagonalisable. Indeed, we saw a 2×2 matrix in the last lecture that had *no* eigenvectors!

Using diagonalisation to solve coupled differential equations

This example illustrates the power of matrix diagonalisation. Suppose we had to solve the following differential equations:

$$\frac{\mathrm{d}y_1}{\mathrm{d}t} = 3y_1$$

$$\frac{\mathrm{d}y_2}{\mathrm{d}t} = 2y_2$$

$$\frac{\mathrm{d}y_3}{\mathrm{d}t} = 6y_3.$$

These are straightforward, because they can be solved individually, and each is just a separable equation. Clearly,

$$y_1 = c_1 \exp(3t)$$
$$y_2 = c_2 \exp(2t)$$
$$y_3 = c_3 \exp(6t).$$

where the c_is are the integration constants. But what if we had something like the following?

$$\frac{dy_1}{dt} = -k_1 y_1 \tag{5.4}$$

$$\frac{dy_2}{dt} = k_1 y_1 - k_2 y_2 \tag{5.5}$$

$$\frac{dy_3}{dt} = k_2 y_2. \tag{5.6}$$

These equations are *coupled*, which makes them much more difficult to solve as they stand. Matrix diagonalisation provides a clever way to decouple them, so that they can be solved independently.

N.B. The example here has been chosen specifically because it's the set of differential equations for the kinetics of $A \rightarrow B \rightarrow C$. You solved it last Term on one of the problem sheets. Now we'll see the easy method.

We start by writing the equations in matrix form:

$$\begin{pmatrix} y_1' \\ y_2' \\ y_3' \end{pmatrix} = \begin{pmatrix} -k_1 & 0 & 0 \\ k_1 & -k_2 & 0 \\ 0 & k_2 & 0 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}.$$

That is,

$$\mathbf{y}' = \mathsf{A}\mathbf{y}. \tag{5.7}$$

Now, let's suppose the three $y_i(t)$ functions can be expressed in terms of three new functions, $u_i(t)$, by relationships of the form:

$$y_1 = s_{11}u_1 + s_{12}u_2 + s_{13}u_3 \tag{5.8}$$
$$y_2 = s_{21}u_1 + s_{22}u_2 + s_{23}u_3 \tag{5.9}$$
$$y_3 = s_{31}u_1 + s_{32}u_2 + s_{33}u_3 \tag{5.10}$$

where each s_{ij} is a constant. In other words,

$$\mathbf{y} = \mathsf{S}\mathbf{u}. \tag{5.11}$$

By differentiating everything in (5.8) with respect to x, we can similarly write

$$\mathbf{y}' = \mathsf{S}\mathbf{u}'$$

with exactly the same matrix S. Subsituting these last two equations into (5.7), we get

$$\mathsf{S}\mathbf{u}' = \mathsf{A}\mathsf{S}\mathbf{u}.$$

If we assume that S is invertible, we can premultiply both sides by S^{-1} to get

$$\mathbf{u}' = S^{-1}AS\mathbf{u}.$$

That was all just algebra. *Now comes the key idea.* We've not yet specified the coefficients of S. Let's *choose* S to be the matrix of eigenvectors, so that (5.3) is satisfied. Then the equation above becomes

$$\mathbf{u}' = D\mathbf{u},$$

i.e.

$$\begin{pmatrix} u_1' \\ u_2' \\ u_3' \end{pmatrix} = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix}.$$

(with the eigenvalues along the diagonal of D). But this is just

$$\frac{du_i}{dx} = \lambda_i u_i$$

for all i; the equations are now decoupled. Each just solves to give $u_i = c_i \exp(\lambda_i x)$, and therefore

$$\mathbf{u} = \begin{pmatrix} c_1 \exp(\lambda_1 t) \\ c_2 \exp(\lambda_2 t) \\ c_3 \exp(\lambda_3 t) \end{pmatrix}.$$

Then, using (5.11), we can obtain the solution for $\mathbf{y}$:

$$\mathbf{y} = S \begin{pmatrix} c_1 \exp(\lambda_1 t) \\ c_2 \exp(\lambda_2 t) \\ c_3 \exp(\lambda_3 t) \end{pmatrix} \tag{5.12}$$

$$= \begin{pmatrix} \uparrow & \uparrow & \uparrow \\ \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 \\ \downarrow & \downarrow & \downarrow \end{pmatrix} \begin{pmatrix} c_1 \exp(\lambda_1 t) \\ c_2 \exp(\lambda_2 t) \\ c_3 \exp(\lambda_3 t) \end{pmatrix}. \tag{5.13}$$

In short: the solution of the coupled differential equations has been expressed simply in terms of the eigenvalues, and eigenvector matrix, of A.

The individual functions, y_i, that solve (5.4) etc. are then found by multiplying this out and taking the components of the vector $\mathbf{y}$.

Notice the key thing that matrix diagonalisation does. It decouples complicated sets of equations, so that they can be solved separately. As a side-effect, it also makes it much easier to keep track of all the solutions, because all the separate parts are lumped together into vectors and matrices.

Let's see how this works for the specific equations in (5.4).

Example 2. Solve
$$\mathbf{y'} = \mathsf{A}\mathbf{y}$$
where
$$\mathsf{A} = \begin{pmatrix} -k_1 & 0 & 0 \\ k_1 & -k_2 & 0 \\ 0 & k_2 & 0 \end{pmatrix}.$$

The eigenvalues of A are easily shown to be

$$\lambda_1 = -k_1 \qquad \lambda_2 = -k_2 \qquad \lambda_3 = 0.$$

The corresponding eigenvectors are

$$\mathbf{v}_1 = \begin{pmatrix} k_1 - k_2 \\ -k_1 \\ 1 \end{pmatrix} \qquad \mathbf{v}_2 = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} \qquad \mathbf{v}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

Hence, from (5.13),

$$\mathbf{y} = \begin{pmatrix} k_1 - k_2 & 0 & 0 \\ -k_1 & 1 & 0 \\ 1 & -1 & 1 \end{pmatrix} \begin{pmatrix} c_1 \exp(-k_1 t) \\ c_2 \exp(-k_2 t) \\ c_3 \end{pmatrix}.$$

On multiplying this out, we find that

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} c_1(k_1 - k_2)\exp(-k_1 t) \\ -c_1 k_1 \exp(-k_1 t) + c_2 \exp(-k_2 t) \\ c_1 \exp(-k_1 t) - c_2 \exp(-k_2 t) + c_3 \end{pmatrix}.$$

It looks grim, but this is the general solution for the kinetics of $A \to B \to C$. The initial conditions (i.e. initial concentrations in this case) would then be used to find c_i.

Second application: normal modes

Another useful application of matrices to chemistry is in figuring out the *normal modes* of vibration of a molecule. You've seen the idea already in the vibrational modes of CO_2. We will consider a simplified version of the problem.

Let's assume that carbon and oxygen atoms have the same mass. (Trust me, I'm a theorist.) Let's also restrict ourselves to motion along one axis; that is, we will not allow the molecule to bend. Finally, let's assume that the atoms are held together by classical mechanical springs. Then, our 'toy' model of CO_2 looks like three equal masses connected by two springs:
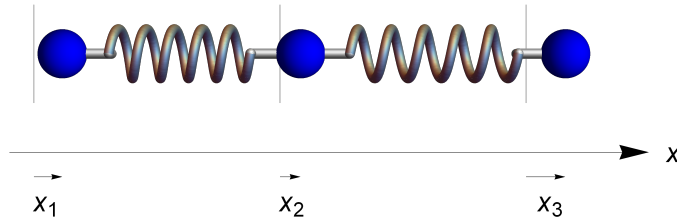
According to Newton's second law,

$$F = ma = m\frac{d^2x}{dt^2}.$$

So we can write down an equation of motion for each of the masses. They look like the following:

$$m\frac{d^2x_1}{dt^2} = k(x_2 - x_1)$$

$$m\frac{d^2x_2}{dt^2} = -k(x_2 - x_1) + k(x_3 - x_2)$$

$$m\frac{d^2x_3}{dt^2} = -k(x_3 - x_2).$$

The variables x_i are the displacements of the particles from their equilibrium locations. To understand the signs in the equations, note that positive forces cause acceleration to the right, and negative forces cause acceleration to the left.

In matrix form, as in (5.7), this becomes

$$\mathbf{x}'' = \mathsf{A}\mathbf{x}$$

where

$$\mathsf{A} = \omega^2 \begin{pmatrix} -1 & 1 & 0 \\ 1 & -2 & 1 \\ 0 & 1 & -1 \end{pmatrix}$$

and $\omega^2 = k/m$.

Using exactly the same arguments as before, we can decouple the three differential equations by transforming to a new set of coordinates. We let

$$\mathbf{x} = \mathsf{S}\mathbf{u} \tag{5.14}$$

[in direct analogy to (5.11)] which leads to

$$\mathsf{S}\mathbf{u}'' = \mathsf{A}\mathsf{S}\mathbf{u}.$$

Again, taking S to be the matrix of eigenvectors of A, and assuming it is invertible, we will find that

$$\mathbf{u}'' = \mathsf{D}\mathbf{u}. \tag{5.15}$$

with $\mathsf{D} = \mathsf{S}^{-1}\mathsf{A}\mathsf{S}$.

Let's figure out the eigenvectors and eigenvalues of A. It's straightforward to show that they are

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \qquad \mathbf{v}_2 = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \qquad \mathbf{v}_3 = \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}$$

with

$$\lambda_1 = 0 \qquad \lambda_2 = -\omega^2 \qquad \lambda_3 = -3\omega^2.$$

Hence, (5.15) gives us

$$\frac{d^2 u_1}{dt^2} = 0$$
$$\frac{d^2 u_2}{dt^2} = -\omega^2 u_2$$
$$\frac{d^2 u_3}{dt^2} = -3\omega^2 u_3.$$

Notice that

$$\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

is an eigenvector, with eigenvalue 0. This is always the case in these calculations, if there are no external forces on the molecule, because all the forces acting on each mass must add to zero. (That is, the rows of A must all add to zero.) This is why there is always a translational normal mode, in each of the x, y and z directions.

Compare these simple equations to the horrible ones we started with. This is why diagonalisation is such an important method.

The equation for u_1 is Newton's second law for a particle without any forces acting on it. It integrates to something of the form

$$u_1 = A + Bt \tag{5.16}$$

which is simply the 'SUVAT' equation for a particle experiencing no acceleration.

For u_2, we have

$$\frac{d^2 u_2}{dt^2} + \omega^2 u_2 = 0,$$

which you learned how to solve last Term. Its auxiliary equation is $\lambda^2 + \omega^2 = 0$, which leads to two complex roots ($\lambda = \pm i\omega$) and hence

$$u_2 = C\cos(\omega t) + D\sin(\omega t). \tag{5.17}$$

Similarly,

$$u_3 = E\cos(\sqrt{3}\omega t) + F\sin(\sqrt{3}\omega t). \tag{5.18}$$

Hence our toy molecule obeys the same laws of motion as three 'particles' moving *independently*; one is a free particle, and the other two obey simple harmonic motion.

The actual atoms in the molecule don't move independently, of course. To understand how they move, we need to express the x_i coordinates back in terms of u_i using (5.14). In written-out form, using the eigenvectors of A, this looks like

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 & -1 & 1 \\ 1 & 0 & -2 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix}.$$

I'm going to write this as

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = u_1 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + u_2 \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} + u_3 \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}.$$

You can check for yourself that this is exactly the same as the equation above.

We see that the motion of the atoms is a linear combination of three *normal modes*. Each mode involves collective motion of all the atoms. The time-dependence of the motion enters via (5.16), (5.17) and (5.18).

For example, the first term above is

$$u_1 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = (A + Bt) \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} A + Bt \\ A + Bt \\ A + Bt \end{pmatrix}.$$

this is saying that all atoms move in exactly the same way, with the time dependence of a free particle. This corresponds to translational motion of the whole molecule, without any changes in bond length.

The second term above is

$$u_2 \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} = [C \cos(\omega t) + D \sin(\omega t)] \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}.$$

This is atom 2 staying still, but atoms 1 and 3 moving in opposite directions with a harmonic oscillator time dependence. This is the symmetric stretching mode.

The third term is the antisymmetric stretch. It too has a harmonic oscillator time dependence, but now the outer atoms move in the opposite direction from the middle one.

Finally, notice that the frequency of the antisymmetric stretch is $\sqrt{3}\omega$, compared to ω for the symmetric stretch. The ratio of the vibrational frequencies should therefore be approximately $\sqrt{3} = 1.732\ldots$ according to our model. This agrees pretty well with the experimental values for CO_2, even given the assumptions of classical mechanics and equal masses:

Molecule	Sym. stretch (cm^{-1})	Antisym. stretch (cm^{-1})	Ratio
$^{12}C^{16}O_2$	1333	2349	1.762
$^{13}C^{16}O_2$	1334	2283	1.711

They are called normal modes because the eigenvectors that describe them are mutually perpendicular. This means that there can be no transfer of energy/motion between the modes, i.e. motion in each mode is separate from the others.