## Math you need to know

## 1 Linear algebra

Linear algebra is mainly concerned with solving equations of the form

$$\mathbf{A} \cdot \mathbf{x} = \mathbf{y} \,, \tag{1} \quad \{\texttt{lin\_eq}\}$$

which is written in terms of components as

$$\sum_{j} A_{ij} x_j = y_j \,. \tag{2}$$

Generally,  $\mathbf{y}$  is known and we want to find  $\mathbf{x}$ . For that, we need the inverse of  $\mathbf{A}$ . That inverse, denoted  $\mathbf{A}^{-1}$ , is the solution to the equation

$$\mathbf{A}^{-1} \cdot \mathbf{A} = \mathbf{I} \tag{3} \quad \{\texttt{soln\_lin\_e}\}$$

where  $\mathbf{I}$  is the identity matrix; it has 1's along the diagonal and 0's in all the off diagonal elements. In components, this is written

$$\sum_{ij} A_{ij}^{-1} A_{jk} = \delta_{ik} \tag{4}$$

where  $\delta_{ik}$  is the Kronecker delta,

$$\delta_{ik} = \begin{cases} 1 & i = k \\ 0 & i \neq k \end{cases}$$
(5) {kronecker}

If we know the inverse, then we can write down the solution to Eq. (1),

$$\mathbf{x} = \mathbf{A} \cdot \mathbf{y} \,. \tag{6}$$

That all sounds reasonable, but what really just happened is that we traded one problem (Eq. (1)) for another (Eq. (3)). To understand why that's a good trade, we need to understand linear algebra – which really means we need to understand the properties of matrices. So that's what the rest of this section is about.

Probably the most important thing we need to know about matrices is that they have eigenvectors and eigenvalues, defined via

$$\mathbf{A} \cdot \mathbf{v}_k = \lambda_k \mathbf{v}_k \,. \tag{7} \quad \{\texttt{eigen}\}$$

Note that  $\lambda_k$  is a scalar (it's just a number). If **A** is  $n \times n$ , then there are *n* distinct eigenvectors (except in *very* degenerate cases, which we typically don't worry about), each with its own eigenvalue.

To find the eigenvalues and eigenvectors, note that Eq. (7) can be written (dropping the subscript k, for reasons that will become clear shortly),

$$\left[\mathbf{A} - \lambda \mathbf{I}\right] \cdot \mathbf{v} = 0. \tag{8} \quad \{\texttt{eigen0}\}$$

For most values of  $\lambda$ , this corresponds to n equations and n unknowns, which means that **v** is uniquely determined. Unfortunately, it's uniquely determined to be 0. So that's not very useful. However, for particular values of  $\lambda$ , some of the n equations are redundant – meaning, more technically, some of the rows of the matrix  $\mathbf{A} - \lambda \mathbf{I}$  are linearly dependent. In that case, there is a vector, **v**, that is nonzero and solves Eq. (8). That's an eigenvector, and the corresponding value of  $\lambda$  is its eigenvalue.

To see how this works in practice, consider the following  $2 \times 2$  matrix,

$$\mathbf{A} = \left(\begin{array}{cc} 5 & 2\\ 4 & 3 \end{array}\right) \,. \tag{9}$$

For this matrix, Eq. (8) can be written

$$\begin{pmatrix} 5-\lambda & 2\\ 4 & 3-\lambda \end{pmatrix} \begin{pmatrix} v_1\\ v_2 \end{pmatrix} = \begin{pmatrix} 0\\ 0 \end{pmatrix}.$$
(10) {eigen1}

As is easy to verify, for most values of  $\lambda$  (for instance,  $\lambda = 0$ ), the only solution is  $v_1 = v_2 = 0$ . However, for two special values of  $\lambda$ , 1 and 7, there are nonzero values of  $v_1$  and  $v_2$  that solve Eq. (10) (as is easy to verify). Note also that when  $\lambda$  takes on either of these values, the determinant of  $\mathbf{A} - \lambda \mathbf{I}$  is zero (see Eq. (30a) for the definition of the determinant of a  $2 \times 2$  matrix).

In fact, this is general: the eigenvalues associated with matrix  $\mathbf{A}$  are found by solving the so-called characteristic equation,

$$Det[\mathbf{A} - \lambda \mathbf{I}] = 0 \tag{11} \qquad (11) \quad \{\texttt{characteri}\)$$

where Det stands for determinant (more on that shortly). If **A** is  $n \times n$ , then this is an  $n^{\text{th}}$  order polynomial; that polynomial has n solutions. Those solutions correspond to the n eigenvalues. For each eigenvalue, one must then solve Eq. (7) to find the eigenvectors.

So what's a determinant? There's a formula for computing it, but it's so complicated that it's rarely used (you can look it up on Wikipedia if you want). However, you should know about its properties, three of the most important being

$$Det[\mathbf{A} \cdot \mathbf{B}] = Det[\mathbf{A}]Det[\mathbf{B}]$$
(12a)

$$\operatorname{Det}[\mathbf{A}^T] = \operatorname{Det}[\mathbf{A}] \tag{12b}$$

$$Det[\mathbf{A}] = \prod_{k} \lambda_k \,. \tag{12c} \quad \{\texttt{det\_lambdas}\}$$

Here superscript T denotes transpose, which is pretty much what it sounds like,

$$A_{ij}^T = A_{ji} \,. \tag{13}$$

Matrices also have adjoint, or left, eigenvectors, for which we usually use a dagger,

$$\mathbf{v}_k^{\dagger} \cdot \mathbf{A} = \lambda_k \mathbf{v}_k^{\dagger}$$
. (14) {adjoint}

Note that the eigenvalues are the same. To see why, write Eq. (14) as

$$\mathbf{A}^T \cdot \mathbf{v}_k^{\dagger} = \lambda_k \mathbf{v}_k^{\dagger} \,. \tag{15}$$

These are found through the characteristic equation,

$$Det[\mathbf{A}^T - \lambda \mathbf{I}] = 0.$$
(16)

Because  $\text{Det}[\mathbf{A}] = \text{Det}[\mathbf{A}^T]$ , and  $\mathbf{I}^T = \mathbf{I}$ , this is the same as Eq. (11).

This analysis tells us that  $\mathbf{v}_k$  and  $\mathbf{v}_k^{\dagger}$  share the same eigenvalue. They also share something else: an orthogonality condition. That condition is written

$$\mathbf{v}_k^{\dagger} \cdot \mathbf{v}_l = \delta_{kl} \tag{17} \quad \{\texttt{ortho}\}$$

where, recall,  $\delta_{kl}$  is the Kronecker delta (defined in Eq. (5)). The  $k \neq l$  part of this equation is easy to show. Write

$$\mathbf{v}_{k}^{\dagger} \cdot \mathbf{A} \cdot \mathbf{v}_{l} = \lambda_{l} \mathbf{v}_{k}^{\dagger} \cdot \mathbf{v}_{l} = \lambda_{k} \mathbf{v}_{k}^{\dagger} \cdot \mathbf{v}_{l} .$$
<sup>(18)</sup>

The first equality came from Eq. (7); the second from Eq. (14). Consequently,

$$(\lambda_k - \lambda_l) \mathbf{v}_k^{\dagger} \cdot \mathbf{v}_l = 0.$$
<sup>(19)</sup>

If all the eigenvalues are different, then  $\mathbf{v}_k^{\dagger} \cdot \mathbf{v}_l = 0$  whenever  $k \neq l$ . If some of the eigenvalues are the same, it turns out that one can still choose the eigenvectors so that  $\mathbf{v}_k^{\dagger} \cdot \mathbf{v}_l = 0$  whenever  $k \neq l$ . (That's reasonably straightforward to show; I'll leave it as an exercise for the reader.) So why set  $\mathbf{v}_k^{\dagger} \cdot \mathbf{v}_k$  to 1? It's a convention, but it will become clear, when we work actual problems, that it's a good convention.

Note that Eq. (17) doesn't pin down the magnitudes of the eigenvectors or their adjoints; it pins down only the magnitude of their product: the eigenvalues can be scaled by any factor, so long as the associated adjoint eigenvector is scaled by the inverse of that factor. As far as I know, there's no generally agreed upon convention for setting the scale factor; what one chooses depends on the problem. Fortunately, all quantities of interest involve products of  $\mathbf{v}_k$  and  $\mathbf{v}_k^{\dagger}$ , so the scale factor doesn't matter.

As a (rather important) aside, if  $\mathbf{A}$  is symmetric, then the eigenvectors and adjoint eigenvectors are the same (as is easy to show from Eqs. (7) and (14)). Moreover, the orthogonality conditions implies that

$$\mathbf{v}_k \cdot \mathbf{v}_l = \delta_{kl} \,. \tag{20}$$

Thus, for symmetric matrices (unlike non-symmetric ones), the magnitude of  $\mathbf{v}_k$  is fully determined by the orthogonality conditions: all eigenvectors have a Euclidean length of 1.

There are several reasons to know about eigenvectors and eigenvalues. Most of them hinge on the fact that a matrix can be written in terms of its eigenvectors, adjoint eigenvectors, and eigenvalues as

$$\mathbf{A} = \sum_{k} \lambda_k \mathbf{v}_k \mathbf{v}_k^{\dagger} \,. \tag{21} \quad \{\texttt{eigen\_expan}\}$$

To show that this equality holds, all we need to do is show that it holds for any vector,  $\mathbf{u}$ . In other words, if Eq. (21) holds, then we must have

$$\mathbf{A} \cdot \mathbf{u} = \sum_{k} \lambda_k \mathbf{v}_k \mathbf{v}_k^{\dagger} \cdot \mathbf{u} \,. \tag{22} \quad \{\texttt{adotu}\}$$

(This is actually an if an only if, which we won't prove.) Because eigenvectors are generally complete, any vector  $\mathbf{u}$  can be written uniquely as the sum of eigenvectors,

$$\mathbf{u} = \sum_{k} a_k \mathbf{v}_k \,. \tag{23}$$

(And, of course, the same is true of adjoint eigenvectors.) Using Eq. (7), along with the orthogonality condition, Eq. (17), we see that Eq. (22) is indeed satisfied. The only time this doesn't work is when the eigenvectors aren't complete, but that almost never happens. When it does, though, one must be careful.

One of the reasons Eq. (21) is important is that it gives us an expression for the inverse of a matrix,

$$\mathbf{A}^{-1} = \sum_{k} \lambda_k^{-1} \mathbf{v}_k \mathbf{v}_k^{\dagger} \,. \tag{24} \quad \{\texttt{inverse}\}$$

That this really is the inverse is easy to show using the orthogonality condition, Eq. (17) (it's the main reasons we used  $\delta_{kl}$  on the right side, rather than some constant times  $\delta_{kl}$ ). As an aside, this generalizes to

$$f(\mathbf{A}) = \sum_{k} f(\lambda_k) \mathbf{v}_k \mathbf{v}_k^{\dagger}$$
(25)

where f is any function that has a Taylor series expansion. We will rarely use this, but it does come up.

One of the important things about Eq. (24) is that it can be used to solve our original problem, Eq. (1). Using Eq. (3), we see that

$$\mathbf{x} = \sum_{k} \lambda_k^{-1} \mathbf{v}_k \mathbf{v}_k^{\dagger} \cdot \mathbf{y} \,. \tag{26} \quad \{\text{expansion}\}$$

So, once we know the eigenvectors, adjoint eigenvalues, and eigenvectors, finding  $\mathbf{x}$  amounts to computing a bunch of dot products.

The following is a bit of an aside, but it will come up later when we solve differential equations. If one of the eigenvalues of  $\mathbf{A}$  is zero then, technically, its inverse does not exist – the expression in Eq. (24) is infinity. However, it's still possible for  $\mathbf{A}^{-1} \cdot \mathbf{y}$  to exist; all we need is for  $\mathbf{y}$  to be orthogonal to any adjoint eigenvector whose corresponding eigenvalue is zero. But this isn't quite the end of the story. Suppose we want to solve Eq. (1) when  $\lambda_1 = 0$  and  $\mathbf{v}_1^{\dagger} \cdot \mathbf{y} = 1$ . In that case, the solution is

$$\mathbf{x} = \sum_{k=2}^{n} \lambda_k^{-1} \mathbf{v}_k \mathbf{v}_k^{\dagger} \cdot \mathbf{y} + c_1 \mathbf{v}_1$$
(27)

where  $c_1$  is any constant. Because  $\mathbf{A} \cdot \mathbf{v}_1 = 0$ , this satisfies Eq. (1). So if  $\mathbf{A}$  isn't invertible, we can have a continuum of solutions! We'll actually use this fact when we solve linear differential equations.

There is one more definition we need: the trace of a matrix, often denoted Tr. The trace is the sum of the diagonal elements,

$$\operatorname{Tr}[\mathbf{A}] \equiv \sum_{i} A_{ii} \,. \tag{28} \quad \{\texttt{trace}\}$$

Using Eq. (21), along with the orthogonality conditions, Eq. (17), it is reasonably easy to show that

$$\operatorname{Tr}[\mathbf{A}] = \sum_{k} \lambda_k \,, \tag{29} \quad \{\texttt{trace\_lambda}\}$$

which is why it's an important quantity.

We end this section by applying some of these ideas to a  $2 \times 2$  matrix. For that, the determinant, which we'll call D, and the inverse are given by

$$D = A_{11}A_{22} - A_{12}A_{22} \tag{30a} \{\texttt{det}\}$$

$$\mathbf{A}^{-1} = \frac{1}{D} \begin{pmatrix} A_{22} & -A_{12} \\ -A_{21} & A_{11} \end{pmatrix}.$$
 (30b)

The second expression is easy to verify. The characteristic equation, Eq. (11), also has a simple form; using T for the trace (Eq. (28)), that equation is given by

$$\lambda^2 - T\lambda + D = 0. \tag{31}$$

This has two solutions, corresponding to the two eigenvalues, which we'll call  $\lambda_{\pm}$ ,

$$\lambda_{\pm} = \frac{T \pm (T^2 - 4D)^{1/2}}{2} \,. \tag{32}$$

(You should verify that  $\lambda_+ + \lambda_- = D$  and  $\lambda_+\lambda_- = D$ . Thus, at least for  $2 \times 2$  matrices, Eqs. (29) and (12c) are correct.)

Finally, note that

$$T^{2} - 4D = (A_{11} - A_{22})^{2} + 4A_{12}A_{21}.$$
(33)

For arbitrary matrices, this quantity can be negative, and the eigenvalues can be complex (for real matrices, they always come in complex conjugate pairs). For symmetric matrices, on the other hand, this quantity is non-negative. Consequently, real symmetric  $2 \times 2$  matrices have real eigenvalues. This turns out to generalize: all real symmetric matrices have real eigenvalues. This isn't that hard to show; we'll leave it as an exercise for the reader.

To be continued ...

- 2 Fourier transforms (and other infinite dimensional linear operators)
- **3** Ordinary differential equations (ODEs)
- 3.1 Linear ODEs
- 3.2 Nonlinear ODEs
- 3.3 Bifurcation theory
- 3.4 Stochastic differential equations
- 4 The central limit theorem
- 5 Taylor expansions
- 6 Integrals
- 7 Distributions