

Math you need to know

1 Linear algebra

Linear algebra is mainly concerned with solving equations of the form

$$\mathbf{A} \cdot \mathbf{x} = \mathbf{y}, \tag{1} \quad \{\text{lin_eq}\}$$

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} . The inverse, denoted \mathbf{A}^{-1} , is the solution to the equation

$$\mathbf{A}^{-1} \cdot \mathbf{A} = \mathbf{I} \tag{3} \quad \{\text{soln_lin_eq}\}$$

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 δ_{ik} is the Kronecker delta,

$$\delta_{ik} = \begin{cases} 1 & i = k \\ 0 & i \neq k. \end{cases} \tag{5} \quad \{\text{kroncker}\}$$

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

$$\mathbf{x} = \mathbf{A}^{-1} \cdot \mathbf{y}. \tag{6}$$

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

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

Note that λ_k is a scalar (it's just a number). If \mathbf{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),

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

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

To see how this works in practice, consider the following 2×2 matrix,

$$\mathbf{A} = \begin{pmatrix} 5 & 2 \\ 4 & 3 \end{pmatrix}. \quad (9)$$

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

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

As is easy to verify, for most values of λ (for instance, $\lambda = 0$), the only solution is $v_1 = v_2 = 0$. However, for two special values of λ , 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 λ takes on either of these values, the determinant of $\mathbf{A} - \lambda\mathbf{I}$ is zero (see Eq. (31a) for the definition of the determinant of a 2×2 matrix).

The fact that the determinant vanishes is general: the eigenvalues associated with matrix \mathbf{A} are found by solving the so-called characteristic equation,

$$\text{Det}[\mathbf{A} - \lambda\mathbf{I}] = 0 \quad (11) \quad \{\text{characteristic}\}$$

where Det stands for determinant (more on that shortly). If \mathbf{A} is $n \times n$, then this is an n^{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 the properties of determinants, three of the most important being

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

$$\text{Det}[\mathbf{A}] = \prod_k \lambda_k \quad (12b) \quad \{\text{det_lambdas}\}$$

$$\text{Det}[\mathbf{A}^T] = \text{Det}[\mathbf{A}] \quad (12c) \quad \{\text{det_transpose}\}$$

$$\text{Det}[\mathbf{A}^{-1}] = \text{Det}[\mathbf{A}]^{-1}. \quad (12d) \quad \{\text{det_inverse}\}$$

Note that Eq. (12d) follows from Eq. (12a), so we don't really need it. Superscript T denotes transpose, which is pretty much what it sounds like,

$$A_{ij}^T = A_{ji}. \quad (13)$$

Matrices also have adjoint, or left, eigenvectors, for which one often uses a dagger,

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

Note that we have taken the eigenvalues associated with the adjoint eigenvectors to be the same as the ones associated with the eigenvectors. To see why this is correct, write Eq. (14) as

$$\mathbf{A}^T \cdot \mathbf{v}_k^\dagger = \lambda_k \mathbf{v}_k^\dagger. \quad (15)$$

These are found through the characteristic equation,

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

Because $\text{Det}[\mathbf{A}] = \text{Det}[\mathbf{A}^T]$ (Eq. (12c)), and $\mathbf{I}^T = \mathbf{I}$ (true for all diagonal matrices), 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

$$\mathbf{v}_k^\dagger \cdot \mathbf{v}_l = \delta_{kl} \quad (17) \quad \{\text{ortho}\}$$

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

$$\mathbf{v}_k^\dagger \cdot \mathbf{A} \cdot \mathbf{v}_l = \lambda_l \mathbf{v}_k^\dagger \cdot \mathbf{v}_l = \lambda_k \mathbf{v}_k^\dagger \cdot \mathbf{v}_l. \quad (18)$$

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

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

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 eigenvectors 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)). In this case, the orthogonality conditions implies that

$$\mathbf{v}_k \cdot \mathbf{v}_l = \delta_{kl}. \quad (20)$$

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

$$\mathbf{A} = \sum_k \lambda_k \mathbf{v}_k \mathbf{v}_k^\dagger. \quad (21) \quad \{\text{eigen_expansi}\}$$

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

$$\mathbf{A} \cdot \mathbf{u} = \sum_k \lambda_k \mathbf{v}_k \mathbf{v}_k^\dagger \cdot \mathbf{u}. \quad (22) \quad \{\text{adotu}\}$$

(This is actually an if and 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,

$$\mathbf{u} = \sum_k a_k \mathbf{v}_k. \quad (23)$$

(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,

$$\mathbf{A}^{-1} = \sum_k \lambda_k^{-1} \mathbf{v}_k \mathbf{v}_k^\dagger. \quad (24) \quad \{\text{inverse}\}$$

To see that this really is the inverse, use orthogonality condition, Eq. (17) to write

$$\mathbf{A}^{-1} \cdot \mathbf{A} = \sum_{kl} \lambda_k^{-1} \mathbf{v}_k \mathbf{v}_k^\dagger \cdot \lambda_l \mathbf{v}_l \mathbf{v}_l^\dagger = \sum_k \mathbf{v}_k \mathbf{v}_k^\dagger. \quad (25)$$

It's very important to realize that the right hand side is the identity matrix. The reasoning is the same as that used to show that Eq. (21) is correct; you should convince yourself of this!

As an aside, this generalizes to

$$f(\mathbf{A}) = \sum_k f(\lambda_k) \mathbf{v}_k \mathbf{v}_k^\dagger \quad (26)$$

where f is any function that has a Taylor series expansion. We will rarely use this, but it does occasionally come up, and it's a good thing to know.

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

$$\mathbf{x} = \sum_k \lambda_k^{-1} \mathbf{v}_k \mathbf{v}_k^\dagger \cdot \mathbf{y}. \quad (27) \quad \{\text{expansion_solv}\}$$

So, once we know the eigenvectors, adjoint eigenvalues, and eigenvectors (and we have the machinery to do that: we just have to solve the characteristic equation for the eigenvalues, and then solve some linear equations for the eigenvectors and their adjoints), 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

$$\mathbf{x} = \sum_{k=2}^n \lambda_k^{-1} \mathbf{v}_k \mathbf{v}_k^\dagger \cdot \mathbf{y} + c_1 \mathbf{v}_1 \quad (28)$$

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,

$$\text{Tr}[\mathbf{A}] \equiv \sum_i A_{ii}. \quad (29) \quad \{\text{trace}\}$$

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

$$\text{Tr}[\mathbf{A}] = \sum_k \lambda_k, \quad (30) \quad \{\text{trace_lambdas}\}$$

which is why it's an important quantity.

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

$$D = A_{11}A_{22} - A_{12}A_{21} \quad (31a) \quad \{\text{det}\}$$

$$\mathbf{A}^{-1} = \frac{1}{D} \begin{pmatrix} A_{22} & -A_{12} \\ -A_{21} & A_{11} \end{pmatrix}. \quad (31b)$$

The first expression is a definition; the second is easy to verify. The characteristic equation, Eq. (11), also has a simple form; using T for the trace (Eq. (29)), that equation is given by

$$\lambda^2 - T\lambda + D = 0. \quad (32)$$

This has two solutions, corresponding to the two eigenvalues, which we'll call λ_\pm ,

$$\lambda_\pm = \frac{T \pm (T^2 - 4D)^{1/2}}{2}. \quad (33)$$

(You should verify that $\lambda_+ + \lambda_- = T$ and $\lambda_+ \lambda_- = D$. Thus, at least for 2×2 matrices, Eqs. (30) and (12b) are correct.)

Finally, note that

$$T^2 - 4D = (A_{11} - A_{22})^2 + 4A_{12}A_{21}. \quad (34)$$

For arbitrary matrices, this quantity can be negative, and the eigenvalues can be complex (for real matrices – matrices whose elements are all real – complex eigenvalues always come in complex conjugate pairs). For symmetric matrices, on the other hand, this quantity is non-negative. Consequently, real symmetric 2×2 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.

1.1 Matrix identities

inverse proportional to cofactors

det prod = prod det

log det = trace log

derivative of log det

2 Fourier transforms (and other infinite dimensional linear operators)

We often want to express a function in terms of other functions. For instance, we might want to write a function $f(x)$ as the sum of sines and cosines. If x has infinite range, that sum turns into integrals, and we end up with an inverse Fourier transform,

$$f(x) = \int_{-\infty}^{\infty} \frac{dk}{2\pi} e^{ikx} \hat{f}(k) \quad (35) \quad \{\text{ft}\}$$

where $i = \sqrt{-1}$. Why we would want to do this is discussed below.

Not surprisingly given its name, the inverse Fourier transform has an inverse – that’s the Fourier transform,

$$\hat{f}(k) = \int_{-\infty}^{\infty} dx e^{-ikx} f(x) \quad (36) \quad \{\text{ift}\}$$

Be aware that there are various conventions for signs and factors of 2π . Sometimes i is replaced by $-i$ in the exponent; sometimes k is replaced by $2\pi k$ in the exponent, in which case the factor of 2π in the denominator goes away; and sometimes the Fourier transform has $(2\pi)^{1/2}$ in the denominator, or even no denominator at all. And engineers have an annoying habit of using j , which they set to $-i$. In each case, the inverse Fourier transform is adjusted accordingly, so that it really is the inverse.

The first observation is that the Fourier transform can be thought of as a glorified dot product between the matrix whose k^{th} element is e^{ikx} . The matrix has an uncountably infinite number of elements, but that’s kind of a detail. And if we wanted to actually compute the Fourier transform, we would first discretize x and k , giving us a matrix with a countably infinite number of elements, and then truncate, giving us a standard matrix. Thus, if you understand linear algebra, you understand Fourier transforms.

That said, the fact that the matrix e^{ikx} has an uncountably infinite number of elements does introduce some issue. In particular, the relevant identity matrix (see Eq. (3)) is now a Dirac delta function,

$$\int \frac{dx}{2\pi} e^{ikx} e^{-ik'x} = \delta(k - k'). \quad (37) \quad \{\text{delta}\}$$

The term on the right – the Dirac delta function – is not a true function, but we’ll call it a function anyway. And it’s the best function in the world. It’s discussed in Sec. 7.1 below, but briefly: it is infinitely narrow and infinitely high, and the narrowness and height are just

right so that the area under it is 1. So it really is the continuum equivalent of the identity matrix. As an aside, Eq. (37) is a really useful representation of the delta function; it comes up all the time in probability theory, Bayesian inference, statistical physics, and probably lots of other places.

So why would we ever want to Fourier transform a function? There are actually lots of reasons, most of which have to do with the kind of eigenvector/eigenvalue analysis we did in the linear algebra section. We'll give just one: the convolution theorem. Suppose that a function $g(x)$ is given in terms of a convolution,

$$g(x) = \int dx' K(x - x')f(x'). \quad (38)$$

This comes up all the time, especially in neuroscience. Suppose we Fourier transformed both sides. In that case, as is easy to show using Eq. (36),

$$\hat{g}(k) = \hat{K}(k)\hat{f}(k) \quad (39) \quad \{\text{conv_theorem}\}$$

where

$$\hat{K}(k) \equiv \int_{-\infty}^{\infty} dx e^{-ikx} K(x). \quad (40)$$

Equation (39) is the convolution theorem. What it says is that once we Fourier transform, convolution reduces to multiplication. This makes convolutions easier in Fourier transformed space than in the original space simply because multiplication is easier than integration. It also allows us to find $f(x)$ in terms of $g(x)$; it's given by the inverse Fourier transform of $\hat{K}(k)/\hat{f}(k)$. This method for finding $f(x)$ is equivalent to using Eq. (27) to solve Eq. (1). (Like I said, Fourier transforms are just linear algebra.) We will see other uses for Fourier transforms when we discuss linear differential equations.

Fourier transforms are just one way of expressing a function as a sum (or integral) of other functions. There are basically an infinite number of ways to do this, and there's a whole theory behind it. Which we won't go into. Instead, we'll just write down two common alternatives. One is the Fourier sum. Suppose you have some function $f(x)$ that has period L , meaning $f(x + L) = f(x)$. In that case, $f(x)$ can be written

$$f(x) = \sum_{k=-\infty}^{\infty} e^{2\pi ik/L} \hat{f}(k) \quad (41)$$

where the sum is over integer k . This too has an inverse; as is easy to verify,

$$\hat{f}(k) = \int \frac{dx}{L} e^{-2\pi ik/L} f(x). \quad (42)$$

Note that this is a mixed discrete/continuous transform: the Fourier sum is discrete; its inverse is continuous. That happens all the time.

Finally, we end with the Laplace transform of the function $f(t)$,

$$\hat{f}(s) = \int_0^{\infty} dt e^{-st} f(t). \quad (43)$$

Its inverse is sort of complicated,

$$f(t) = \int ds 2\pi i e^{st} \hat{f}(s) \quad (44)$$

where the integration is along the imaginary s -axis. Complications come in because the path of integration can be deformed in complex s -space, but we won't go into that. We will point out, though, that if you make the transformation $s \rightarrow is$, this looks much like an inverse Fourier transform. Laplace transforms are sometimes used to solve linear ordinary differential equations, and they are important for computing averages of Dirichlet distributions.

3 Ordinary differential equations (ODEs)

Ordinary differential equations are equations of the form

$$\frac{dx_i}{dt} = f_i(\mathbf{x}). \quad (45) \quad \{\text{general_ode}\}$$

If $f_i(\mathbf{x})$ is a linear function of \mathbf{x} , it's a linear ODE; those we can solve. If, on the other hand, $f_i(\mathbf{x})$ is a nonlinear function of \mathbf{x} , there is generally no solution. In one and two dimensions we can draw pictures that tell us a lot; beyond that life becomes very difficult, mainly because the equations can admit chaotic dynamics.

Here we'll focus mainly on problems we can solve: linear ODEs, and nonlinear ODEs in one and two dimensions. We'll also briefly consider stochastic ODEs, for which a noise term is added to the right hand side of Eq. (45).

3.1 Linear ODEs

Linear ODEs have the form

$$\frac{dx_i}{dt} = \sum_j A_{ij} x_j. \quad (46)$$

Or, in vector notation,

$$\frac{d\mathbf{x}}{dt} = \mathbf{A} \cdot \mathbf{x}. \quad (47) \quad \{\text{lin_ode}\}$$

Linearity refers to the fact that the equations are linear in \mathbf{x} . This has an important consequence: if $\mathbf{y}(t)$ and $\mathbf{z}(t)$ are both solutions to Eq. (47), then so is their sum. The way we solve these equations, then, is to find all possible solutions, and add them together.

As is well known, mainly because somebody figured it out a long time ago, the solutions to Eq. (47) are, generally, exponentials: $\mathbf{x} \propto e^{\lambda t}$. (There are exceptions, which we'll talk about briefly below.) The problem is to find which values of λ are the relevant ones. To do that, assume $\mathbf{x} = \mathbf{v}e^{\lambda t}$, and insert that into Eq. (47). That yields

$$\lambda \mathbf{v} = \mathbf{A} \cdot \mathbf{v}. \quad (48)$$

This is an equation we've seen before: it's the eigenvalue equation given in Eq. (7). (Not really a big coincidence; all linear equations are essentially the same.) We know, therefore, that if \mathbf{A} is $n \times n$, then there will be n eigenvalues, n eigenvectors, and n adjoint eigenvectors. As usual, we denote these λ_k , \mathbf{v}_k and \mathbf{v}_k^\dagger . Then, the most general solution to Eq. (47) is

$$\mathbf{x}(t) = \sum_k a_k \mathbf{v}_k e^{\lambda_k t}. \quad (49) \quad \{\text{soln1_lin_ode}\}$$

Note that the solution depends on n numbers – the a_k . So it's really a family of solutions. But for a particular problem, they have particular values. Usually (but not always), one is given an initial value problem, for which $\mathbf{x}(t=0)$, commonly known as $\mathbf{x}(0)$, is known. Inserting this into the above solution gives

$$\mathbf{x}(0) = \sum_k a_k \mathbf{v}_k. \quad (50)$$

Using the orthogonality condition, Eq. (17), we can solve for the a_k ,

$$a_k = \mathbf{v}_k^\dagger \cdot \mathbf{x}(0). \quad (51)$$

Inserting this into Eq. (49), we have

$$\mathbf{x}(t) = \sum_k e^{\lambda_k t} \mathbf{v}_k \mathbf{v}_k^\dagger \cdot \mathbf{x}(0). \quad (52) \quad \{\text{soln2_lin_ode}\}$$

This is nice, compact expression. And it becomes especially useful if we're interested in the long time limit, because in that case

$$\mathbf{x}(t) \approx e^{\lambda_{k_0} t} \mathbf{v}_{k_0} \mathbf{v}_{k_0}^\dagger \cdot \mathbf{x}(0) \quad (53)$$

where λ_{k_0} is the largest eigenvalue. Of course, if there are two eigenvalues that tie for largest, both have to be included.

Eq. (47) is known as a homogeneous ODE. Often we have to solve homogeneous ones, which have the form

$$\frac{d\mathbf{x}}{dt} - \mathbf{A} \cdot \mathbf{x} = \mathbf{g}(t). \quad (54) \quad \{\text{inhomo_ode}\}$$

We have sort of seen this before: the left hand side is a linear operator (meaning it's linear in \mathbf{x}), so all we have to do is invert it. We know a lot about inverting matrices (which are also linear operators; they're just particularly simple ones). Inverting linear operators is a little trickier, but the same principles apply: find the eigenvalues and eigenvectors, and use those.

So how do we proceed? There are lot of ways, but one of my favorites is to write

$$\mathbf{x}(t) = \sum_k a_k(t) \mathbf{v}_k \quad (55)$$

where the \mathbf{v}_k are the eigenvectors of \mathbf{A} . Inserting this into Eq. (54) gives us

$$\sum_k \frac{da_k(t)}{dt} \mathbf{v}_k - \mathbf{A} \cdot \sum_k a_k \mathbf{v}_k = \mathbf{g}(t). \quad (56)$$

Using the orthogonality relationships, and the fact that $\mathbf{A} \cdot \mathbf{v}_k = \lambda_k \mathbf{v}_k$, we find that equations for the individual a_k ,

$$\frac{da_k(t)}{dt} - \lambda_k a_k = \mathbf{v}_k^\dagger \cdot \mathbf{g}(t). \quad (57)$$

So we have reduced the problem: now we have to solve n one dimensional ODEs, and they all have the same form. So once we solve one, we've solved all of them!

Because we essentially have one equation to solve, to ease notation we'll consider the equation

$$\frac{da}{dt} - \lambda a = g(t). \quad (58) \quad \{\text{inhomo_ode_1d}\}$$

We're going to use the Green function approach, because it's powerful and intuitive. The idea is to solve the equation

$$\frac{dG(t, t')}{dt} - \lambda G(t, t') = \delta(t - t') \quad (59) \quad \{\text{green}\}$$

where the delta function is discussed in Sec. 7.1. Briefly, it's zero when $t \neq t'$, infinite when $t = t'$, and it integrates to one. Basically, it's the continuous analog of the identity matrix. Thus, the Green function is, essentially, the inverse of the linear operator $d/dt - \lambda$. As such, that inverse operates on $g(t)$. And, in fact, that's (mainly) correct,

$$\left[\frac{d}{dt} - \lambda \right] \int dt' G(t - t') g(t') = \int dt' \delta(t - t') g(t') = g(t). \quad (60)$$

(You may want to consult Sec. 7.1 on the delta function.) Thus,

$$a(t) = \int dt' G(t, t') g(t'). \quad (61)$$

There are, though, a couple of twists. We haven't specified the limits of integration, and that turns out to matter. Also, the most general solution includes the homogeneous solution,

$$a(t) = \int dt' G(t, t') g(t') + a_0 e^{\lambda t}. \quad (62)$$

The homogeneous solution is used to satisfy initial, or boundary, conditions. (Boundary conditions are typically, although not always, conditions on $a(t)$ at $t = \pm\infty$.)

So how do we solve Eq. (59)? For most values of t , the right hand side is zero, and the solution is just $e^{\lambda t}$. However, there can be a discontinuity at $t = t'$. We thus write

$$G(t, t') = e^{\lambda(t-t')} \begin{cases} c_1 & t < t' \\ c_2 & t > t' \end{cases} \quad (63)$$

To be continued ...

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

7.1 Delta function

{sec:delta}

8 Lagrange multipliers