4 Matrix methods

In this week we will learn about how linear algebra is used to solve problems in physics. We will also use this chance to introduce Mathematica, which is a powerful programming language when it comes to symbolic manipulation.

It is useful for our purposes to think of matrices as transformations: they simply act on a vector, transforming it to another vector.

$$\vec{v_2} = \mathbf{M}\vec{v_1} \tag{30}$$

This can be made quite visual and geometric, as will be seen in the Mathematica scripts.

4.1 Linear equations in physics

It is worth simply recalling the fact that linear algebra has the word linear built right into the name. This word gets thrown around a lot[2], and means different things in different contexts. For us in this context it is a system with evolution only depending on first powers of its variables, and constants. Contrast this to what we discussed in the general ODE solvers in Sect. 2, where any rate of change could be solved in an iterative way.

This apparent limitation is in fact not as problematic as it may seem. By this point in your physics careers you have likely noticed that harmonic oscillators are ubiquitous. Lecturers of mine (TK) have gone as far as saying that $F = -kx$ is the most important equation in all of physics. This simple statement of a restoring force which acts against displacement in the systems is extremely profound. To see why this occurs so frequently it is sufficient to study a 1D physical system close to a minimum of potential $V(x)$. In this case we can study the force as

$$m\ddot{x} = F = -\frac{\partial V}{\partial x} = -V'(x). \tag{31}$$

If the system is close to a minimum at x_0 can be taylor expanded

$$V(x) \approx V(x_0) + V'(x_0)(x - x_0) + V''(x_0)\frac{(x - x_0)^2}{2} \tag{32}$$

$$= V(x_0) + V''(x_0)\frac{(x - x_0)^2}{2}, \tag{33}$$

where in the second line we used $V'(x_0) = 0$ by the definition of a potential minimum. Combining these two equations we can then find

$$\ddot{x} \approx -\frac{1}{m}V''(x_0)(x - x_0). \tag{34}$$

As an exmple we can plug in the usual spring potential $kx^2/2$ with minimum $x_0 = 0$, $V''(x_0) = k$ and get the expected $\ddot{x} = -kx/m$.

To summarise: any system with a stable equilibrium will exhibit harmonic oscillations around that minimum for sufficiently small perturbations. This implies they're susceptible to the linear algebra techniques learned here.

[2]The most fundamental definition I know of is that a map between vector spaces, $f : \mathbb{V} \to \mathbb{W}$, is linear if the vector space structure is preserved in the mapping, e.g. $f(\alpha\vec{v}_1 + \beta\vec{v}_2) = \alpha f(\vec{v}_1) + \beta f(\vec{v}_2)$. This usually requires some distilling before it is directly helpful in physics -TK

4.2 Eigenanalysis

Many problems in physics can be solved by studying eigenvectors and eigenvalues of problems. These represent interesting states that are *stable* under some transformation. By this we mean states that become simple rescaled versions of themselves under the influence of the matrix. The fundamental equation is

$$\mathbf{M}\vec{v} = \lambda\vec{v}, \tag{35}$$

which upon rearranging gives

$$(\mathbf{M} - \lambda\mathbf{I})\vec{v} = \vec{0}. \tag{36}$$

This means the transformation offered by $\mathbf{M} - \lambda\mathbf{I}$ must be singular, and thus its determinant must be 0. This gives the eigenvalue equation

$$\left|\mathbf{M} - \lambda\mathbf{I}\right| = 0. \tag{37}$$

which in practice requires solving an N dimensional equation for an $N \times N$ dimensional matrix. This provides the N distinct eigenvalues $\{\lambda_i\}$ and corresponding vectors $\{\vec{v}_i\}$.

One way of justifying our interest in eigenanalysis is to remember that one can choose a set of basis vectors from the vector space, and use these to represent any other vector

$$\vec{v} = \Sigma_{i=0}^{N} v_i \vec{e_i}. \tag{38}$$

Taking the basis to be made up of the eigenvectors we then find

$$\mathbf{M}\vec{v} = \Sigma_{i=0}^{N} \lambda_i v_i \vec{e_i}, \tag{39}$$

which emphasises the real change in the vector occurring because the different λ_i stretch it different amounts in different directions. If all λ_i were the same coefficient then no angle change would occur, and all vectors are eigenvectors. Recall however that in the case of degenerate eigenvalues we have an ambiguity in choosing eigenvectors to begin with – it works out consistent.

The generality and importance of all this can't be understated: we can understand the actions of an arbitrary transformation by understanding its actions on the eigenbasis.

4.2.1 Eigenanalysis example

Continuing with the above example of harmonic oscillators we will consider the case of three coupled oscillators. Specifically we have three masses (m_1, m_2, m_3) with positions (x_1, x_2, x_3). These are connected pairwise with two strings such that masses 1 and 2 are connected with a coupling constant k_{12} and equilibrium length a_{12}. The equations of motion for such a system are

$$\begin{aligned} m_1\ddot{x_1} &= k_{12}(x_2 - x_1 - a_{12}) \\ m_2\ddot{x_2} &= -k_{12}(x_2 - x_1 - a_{12}) + k_{23}(x_3 - x_2 - a_{23}) \\ m_3\ddot{x_3} &= -k_{23}(x_3 - x_2 - a_{23}). \end{aligned} \tag{40}$$

This has the typical appearance of ODEs we have solved in Sect. 2: a set of derivatives specifying the evolution of a system in terms of the state of the system itself, including couplings between terms. One notable feature however is that no higher powers of x_i appear. This hints towards our linear algebra application, and as such the system can be written as a matrix equation

$$\begin{bmatrix} \ddot{x_1} \\ \ddot{x_2} \\ \ddot{x_3} \end{bmatrix} = \begin{bmatrix} -\frac{k_{12}}{m_1} & \frac{k_{12}}{m_1} & 0 \\ \frac{k_{12}}{m_2} & -\frac{k_{12}}{m_2} - \frac{k_{23}}{m_2} & \frac{k_{23}}{m_2} \\ 0 & \frac{k_{23}}{m_3} & -\frac{k_{23}}{m_3} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} -k_{12}a_{12} \\ k_{12}a_{12} - k_{23}a_{23} \\ k_{23}a_{23} \end{bmatrix}, \tag{41}$$

which we will write as

$$\ddot{\vec{X}} = \mathbf{Q}\vec{X} + \vec{A}. \tag{42}$$

The approach we will take to solving this problem is to choose some new coordinates where the equations decouple. We will propose that there is some transformation matrix $\mathbf{T}$ that maps from the old basis to the new one. We thus define $\mathbf{T}\vec{X} = \vec{Y}$ where $\vec{Y}$ is the uncoupled coordinates. We now note

$$\begin{aligned} &\ddot{\vec{X}} = \mathbf{Q}\vec{X} + \vec{A} \\ \implies &\mathbf{T}^{-1}\ddot{\vec{Y}} = \mathbf{Q}\mathbf{T}^{-1}\vec{Y} + \vec{A} \\ \implies &\ddot{\vec{Y}} = \mathbf{T}\mathbf{Q}\mathbf{T}^{-1}\vec{Y} + \mathbf{T}\vec{A}. \end{aligned} \tag{43}$$

This combination of a matrix *sandwiched* by a transformation and its inverse is very common, and receives the name of a similarity transform. It should be read from right to left as "convert back to old coordinates, where this matrix usually applies. Apply the matrix. Map back to new coordinates".

Now comes the key step: we will demand that this new matrix $\mathbf{T}\mathbf{Q}\mathbf{T}^{-1}$ be diagonal, since we want uncoupled equations. Which new basis is it that has a purely diagonal matrix? It is precisely the eigenbasis. This means the state vector $\vec{Y}$ now has 3 components specifying the amount of each eigenstate, and these do not mix under action of $\mathbf{Q}$ by definition - they can only be rescaled.

To be more specific the tranformed matrix will be diagonal with the eigenvalues $\mathbf{T}\mathbf{Q}\mathbf{T}^{-1} = \mathrm{diag}(\lambda_1, \lambda_2, \lambda_3)$ while the inverse transformation matrix will have eigenvectors forming the three columns $\mathbf{T}^{-1} = [\vec{v_1}|\vec{v_2}|\vec{v_3}]$, such that $\vec{Y} = (1, 0, 0)^T$ simply maps to the first eigenvector, as expected.

Let's yield results from all this mathematical manipulation. Using the above assertions we can write

$$\ddot{y}_i = \lambda_i y_i + b_i, \tag{44}$$

where we define b_i as components of $\vec{B} = \mathbf{T}\vec{A}$. This gives us familiar ODEs for the much anticipated harmonic motion. Specifically this means we have the solutions

$$y_i = \begin{cases} \alpha_i e^{\sqrt{\lambda_i}t} + \beta_i e^{-\sqrt{\lambda_i}t} - \frac{b_i}{\lambda_i} & \text{if } \lambda_i \neq 0 \\ \alpha_i t + \beta_i + \frac{1}{2} b_i t^2 & \text{if } \lambda_i = 0 \end{cases} \tag{45}$$

If $\lambda_i < 0$, this is SHM and we can write $y_i = A_i \cos\omega_i t + B_i \sin\omega_i t - b_i/\lambda_i$ with $\omega_i = \sqrt{-\lambda_i}$ if we prefer.

4.3 Abstract vector spaces

A subtle notion which will become clearer in more advanced subjects of physics is that the mathematics of matrices and vectors is more general than the *boxes of numbers* we have been used to.[3] This can be emphasised firstly by realising that the vectors and matrices discussed so far seem to have a geometric reality of their own (they can be drawn and studied as transformations). A much more compelling case however is that continuous functions can be cast into exactly the same language. Consider functions as vectors (they can be added,

[3]Or graveyards of numbers as Frederic Schuller likes to call them in his fantastic lecture series on all aspects of Theoretical Physics. -TK

subtracted, scaled, and all the other operations vectors follow) with differential operators as the transformative *matrices* (they take in a function and return a transformed function). Consider for example the *eigenfunctions* of $\mathrm{d}/\mathrm{d}x$:

$$\frac{\mathrm{d}f}{\mathrm{d}x} = \lambda f(x) \implies f(x) = \mathrm{e}^{\lambda x}, \tag{46}$$

where λ isn't quantised if we consider the space of all differentiable functions on $(-\infty, \infty)$, because this vector space has an uncountable infinity of dimensions.

Let's take this one step further and consider what we are doing in the typical quantum mechanics solution. Essentially we solve the eigenequation of the Hamiltonian, giving wavefunctions which only receive a rescaling under the action of the Hamiltonian (the coefficient corresponds to energy). We then say that these functions can be a basis for other functions and understand that a general wavefunction changes under the Hamiltonian because its underlying eigencomposition is getting stretched by different amounts. Another example is Fourier analysis and the use of integral transforms to solve differential equations.

4.4 Abstract vector space example

We will demonstrate the notions above through an example. We will consider the quantum harmonic oscillator

$$\hat{H} = \frac{1}{2m}\hat{p}^2 + \frac{1}{2}m\omega^2\hat{x}^2. \tag{47}$$

We simplify the algebra by choosing to use units in which $m = 1$, $\omega = 1$ and $\hbar = 1$, so $\hat{p} = -i\frac{\partial}{\partial x}$ and

$$\hat{H} = \frac{-1}{2}\frac{\partial^2}{\partial x^2} + \frac{1}{2}x^2, \tag{48}$$

which we know has eigensolutions

$$\hat{H}\psi_n(x) = \left(n + \frac{1}{2}\right)\psi_n(x) = E_n\psi_n(x) \tag{49}$$

where $\psi_n(x)$ is given in terms of Hermite polynomials[4] H_n:

$$\psi_n(x) = \frac{1}{\sqrt{2^n n!}}\left(\frac{1}{\pi}\right)^{1/4} e^{-\frac{x^2}{2}} H_n(x). \tag{50}$$

The assertion is then that any arbitrary wavefunction $\phi(x)$ can be written as some superposition of these:

$$\phi(x) = \sum_n \alpha_n \psi_n(x) \tag{51}$$

Now comes the **key point**: If we simply provide these coefficients α_n (potentially an infinite number of them) we have in essence specified the entire function itself. This means the function has a perfectly valid *vector* representation $\phi(x) \to (\alpha_1, \alpha_2, \alpha_3, ...)^T$. Additionally, we know the action of the Hamiltonian on this vector – it rescales the n^{th} entry according to the energy E_n from Eq. 49. This means

$$\hat{H} \to \begin{bmatrix} \frac{1}{2} & 0 & 0 & \cdots \\ 0 & \frac{3}{2} & 0 & \cdots \\ 0 & 0 & \frac{5}{2} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}. \tag{52}$$

[4]Be careful with the definition of these, as physicists use a different convention to mathematicians.

Now we have fully mapped the differential operator and function to matrix and vector respectively. Compare the diagonal matrix in Eq. 52 to the expression in Eq. 43 where we stated $\mathbf{TQT}^{-1}$ was the diagonal matrix of eigenvalues. The mathematics in each case is exactly the same, except here didn't need to transform into this nicer decoupled basis from some other more physically intuitive basis like we did before, because we already set up the problem in the nicer basis.

Now comes the real example. We are going to add a quartic perturbation to harmonic motion

$$\hat{H} = \frac{1}{2}\hat{p}^2 + \frac{1}{2}\hat{x}^2 + \lambda\hat{x}^4, \tag{53}$$

so now we don't know what the eigenfunctions are. How do we proceed?

One key point is that λ is considered to be a small perturbation. And in the limit of $\lambda \to 0$ we expect to recover the usual eigenfunctions in Eq. 50. We will therefore decide to cast the equation into matrix and vector exactly as before, understanding that any off-diagonal non-zero values will be small, in proportion to λ. We can then numerically find the eigenbasis rather than analytically, similar to the classical problem considered before.

In general, the casting of a function to its n^{th} vector component follows

$$\phi_n = \int \mathrm{d}x\, \psi_n^*(x)\, \phi(x) \tag{54}$$

and the matrix is mapped with

$$H_{mn} = \int \mathrm{d}x\, \psi_m^*(x)\, \hat{H}\, \psi_n(x). \tag{55}$$

Try to stop for a second and justify the above vector and matrix using these full expressions. Punchline in the footnote[5].

Constructing the matrix corresponding to δH can be done in a brute force way (we will show this in the Mathematica notebook), but a more elegant solution can be found via the raising and lowering operators, which have the effect

$$\hat{a}^\dagger \psi_n = \sqrt{n+1}\, \psi_{n+1}, \tag{56}$$

$$\hat{a}\, \psi_n = \sqrt{n}\, \psi_{n-1}. \tag{57}$$

Looking back at our Quantum Mechanics notes, we see that we can write the $\hat{x}$ operator in terms of $\hat{a}^\dagger$ and $\hat{a}$, as

$$\hat{x} = \frac{1}{\sqrt{2}}\left(\hat{a}^\dagger + \hat{a}\right). \tag{58}$$

We can find the action of $\hat{x}$ on ψ_n in this way:

$$\hat{x}\, \psi_n = \frac{1}{\sqrt{2}}\left(\sqrt{n+1}\, \psi_{n+1} + \sqrt{n}\, \psi_{n-1}\right). \tag{59}$$

Iterating this formula four times, we obtain

$$\begin{aligned}
\hat{x}^4\psi_n &= \frac{1}{4}\left(\hat{a}^\dagger + \hat{a}\right)^4 \psi_n \\
&= \frac{1}{4}\Big(\sqrt{n-3}\sqrt{n-2}\sqrt{n-1}\sqrt{n}\, \psi_{n-4} + 2\sqrt{n-1}\sqrt{n}(2n-1)\psi_{n-2} + 3(2n(n+1)+1)\psi_n \\
&\quad + 2\sqrt{n+1}\sqrt{n+2}(2n+3)\psi_{n+2} + \sqrt{n+1}\sqrt{n+2}\sqrt{n+3}\sqrt{n+4}\, \psi_{n+4}\Big).
\end{aligned} \tag{60}$$

In this way, we can calculate the elements of the matrix H_{mn} without having to perform any integrals.

[5] One of the keys to the simple expressions earlier was orthogonality of the eigenmodes, as well as $\hat{H}$ not mixing them. $\hat{x}^4$ will not have similar behaviour.