University of Canterbury

EMTH210

ENGINEERING MATHEMATICS

Topic 4: Eigen problems

4.1 REVIEW OF FINDING DETERMINANTS

This section should all be revision for you. The determinant $|A|$ of an $n \times n$ matrix A can be defined recursively by

$$|A| = \sum_{i=1}^{n} a_{ij}(-1)^{i+j}|M_{ij}| \qquad \text{for any fixed } j\text{, or}$$

$$= \sum_{j=1}^{n} a_{ij}(-1)^{i+j}|M_{ij}| \qquad \text{for any fixed } i,$$

(ie we can calculate along rows or columns) where a_{ij} is the element of A in row i and column j

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & & \ddots & \vdots \\ a_{n1} & & & a_{nn} \end{pmatrix},$$

and where $(-1)^{i+j}$ is the so-called "checkerboard" sign

		j		
		1	2	3
	1	$+$	$-$	$+$
i	2	$-$	$+$	$-$
	3	$+$	$-$	$+$
	4	$-$	$+$	$-$

and where M_{ij} is the ij^{th} minor of A, that is A with the i^{th} row and j^{th} column deleted, such that for example

$$A = \begin{pmatrix} a_{11} & a_{22} & a_{33} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \quad \rightarrowtail \quad M_{12} = \begin{pmatrix} \dots & \dots & \dots \\ a_{21} & \vdots & a_{23} \\ a_{31} & \vdots & a_{33} \end{pmatrix} = \begin{pmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{pmatrix}.$$

So $|A|$ = sum along any row or column, element by element, of "element $\times$ element's checkerboard sign $\times$ determinant of element's minor".

Example 4.1.1

Find the determinant of

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} .$$

Solution.

♦

Example 4.1.2

Find the determinant of

$$A = \begin{pmatrix} 1 & 0 & 3 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 2 & 1 & 2 \\ 1 & 0 & 0 & 3 \end{pmatrix} .$$

Solution.

4.2 EIGENVALUES and EIGENVECTORS

The *eigenvectors* of a square matrix A are the non-zero vectors $\boldsymbol{x}$ that remain proportional to $\boldsymbol{x}$ when they are multiplied by A. In other words, they are the $\boldsymbol{x} \neq \mathbf{0}$ for which

$$A\boldsymbol{x} = \lambda \boldsymbol{x} \tag{1}$$

where λ is called the *eigenvalue* associated with the eigenvector $\boldsymbol{x}$.

You will note that the eigenvectors and eigenvalues "belong" to the square matrix A — we speak of "the eigenvalues and eigenvectors *of* A". The German word for "characteristic" or "own" is "eigen".

In (1) both λ and $\boldsymbol{x}$ are unknowns. Finding these unknowns for a given matrix A is called *solving an eigen problem*[1]. Other non-matrix problems in which the unknowns also depend on an unknown parameter are sometimes referred to as eigenproblems.

Example 4.2.1

Show that $\boldsymbol{x} = (1,1)^{\mathrm{T}}$ is an eigenvector of the following matrix, and find its associated eigenvalue:

$$A = \begin{pmatrix} 1 & 4 \\ 3 & 2 \end{pmatrix} .$$

Solution.

♦

[1] Or *eigenvalue problem*, etc.

So we see that by definition an eigenvector $\boldsymbol{x}$ of a matrix A is a vector whose direction is unchanged by the transformation A; it is only stretched or compressed, and the eigenvalue gives the stretch.

The general method to solve $A\boldsymbol{x} = \lambda\boldsymbol{x}$ for $\boldsymbol{x}$ and λ is as follows.

$$A\boldsymbol{x} = \lambda\boldsymbol{x} \iff A\boldsymbol{x} - \lambda\boldsymbol{x} = (A - \lambda I)\boldsymbol{x} = \mathbf{0} \ ,$$

where I is the identity matrix. Now we have already said that $\boldsymbol{x} \neq 0$. Imagine that our new matrix $A - \lambda I$ had an inverse, call it B. Then if we multiplied both sides of our equation by B (on the left) then on the LHS we would get $B(A - \lambda I)\boldsymbol{x} = \boldsymbol{x}$ while on the RHS we would get $B\mathbf{0} = \mathbf{0}$. Since the LHS equals the RHS this tells us that $\boldsymbol{x} = \mathbf{0}$, which we know it can't be. So we have found that *if* $(A - \lambda I)$ has an inverse *then* we get a contradiction. Therefore $(A - \lambda I)$ can't have an inverse — it must be *singular*. Recall that a singular matrix is one whose determinant is zero. Hence the strategy to find the eigenvalues and eigenvectors of A is

1. Solve $\det(A - \lambda I) = 0$ for the eigenvalues λ;
2. then solve $(A - \lambda I)\boldsymbol{x} = 0$ for the eigenvectors $\boldsymbol{x}$.

Example 4.2.2

Find the eigenvalues and eigenvectors of

$$A = \begin{pmatrix} 1 & 4 \\ 3 & 2 \end{pmatrix} \ .$$

Solution.

♦

Eigenvalues and eigenvectors are a crucial part of engineering mathematics. Over the next four lectures, we will see them applied to solving systems of differential equations, and to describing the strain of materials placed under shear stress.

Example 4.2.3

Find the eigenvalues and eigenvectors of the following matrix:

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

Solution.

♦

4.3 TWO POINTS

Firstly, while there is no room for argument about the eigen*values* of a given matrix, eigen*vectors* are only unique in terms of their direction. This is maybe clear from the definitions, but also note that if $(\boldsymbol{x}, \lambda)$ is an eigen pair then so is $(k\boldsymbol{x}, \lambda)$ for any non-zero constant k as

$$A(k\boldsymbol{x}) = \underline{\hspace{10cm}}\ .$$

The second point is that eigenvalues need not be real, and they can be complex even when A is real. If they are complex then they occur in complex conjugate pairs. Consider the following example.

Example 4.3.1

Find the eigenvalues and eigenvectors of the following matrix:

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

Solution.

♦

4.4 SYSTEMS OF 1^{st} ORDER ODEs

Suppose that the concentration y_1, y_2 of two reacting chemicals is given by the pair of equations

$$\begin{aligned} \dot{y}_1 &= \frac{dy_1}{dt} = y_1 - 2y_2 \\ \dot{y}_2 &= \frac{dy_2}{dt} = -3y_1 \ . \end{aligned}$$

A more compact form of writing these equations would be as a matrix product, by defining the vector $\boldsymbol{y} = (y_1, y_2)$, and the *matrix of coefficients* $A = \begin{pmatrix} 1 & -2 \\ -3 & 0 \end{pmatrix}$. In this way, we can replace the above pair of equations with the single matrix equation

$$\dot{\boldsymbol{y}} = \frac{d\boldsymbol{y}}{dt} = A\boldsymbol{y} \ .$$

Generalising, suppose that we have N variables $y_1, y_2, \ldots, y_N$ whose rate of change with time depends on some linear combination of these variables. Let $\boldsymbol{y} = (y_1, y_2, \ldots, y_N)$ and write the system of N first order linear ODEs as a single matrix DE, namely

$$\frac{d\boldsymbol{y}}{dt} = A\boldsymbol{y} \ . \tag{2}$$

Looking for a solution, we extend our successful idea of using an ansatz, and in fact let it take the same form, except now involving vectors:

$$\boldsymbol{y} = \boldsymbol{x}e^{\lambda t} \implies \dot{\boldsymbol{y}} = \lambda \boldsymbol{x}e^{\lambda t} \ .$$

In fact, what we are doing with this vector equation is taking all of the individual ansatz, $y_i = x_i e^{\lambda t}$ for $i = 1$ to N, and putting them into a vector. Now $\dot{\boldsymbol{y}} = A\boldsymbol{y}$ becomes $\lambda \boldsymbol{x}e^{\lambda t} = A\boldsymbol{x}e^{\lambda t}$. As $e^{\lambda t} \neq 0$ for any value of t or λ, we must have $\lambda \boldsymbol{x} = A\boldsymbol{x}$.

In other words, the solution of our system of ODEs requires the eigenvalues λ and the associated eigenvectors $\boldsymbol{x}$ of the matrix of coefficients A. Because of the linearity of (2) its general solution is

$$\boldsymbol{y} = \sum_{i=1}^{N} c_i x_i e^{\lambda_i t} \ ,$$

where the c_i are constants.

Example 4.4.1

Find the general solution of

$$\frac{d}{dt}\boldsymbol{y} = \begin{pmatrix} 4 & 0 & -1 \\ 3 & 1 & -1 \\ 2 & 0 & 1 \end{pmatrix} \boldsymbol{y} \quad \text{with initial conditions} \quad y(0) = \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix} \ .$$

Solution.

♦

You should give some thought to what this means. How does each of the variables behave? What have we actually done?

4.5 EXTENSION TO 2^{nd} ORDER PROBLEMS

The extension is quite straightforward, but there are a few subtleties in the method which we develop through the following example. You should put a box around the key points, as there is no summary to follow.

Example 4.5.1

Consider a system of three identical Hookean springs with spring constant k, and two identical masses of mass m, with $m/k = 1$, connected as follows. The first spring is connected to the ceiling, the first mass is connected to the bottom of the first spring, the second spring to the bottom of the mass, the second mass to the bottom of the second spring, and the third spring to the bottom of the second mass and to the floor. The springs hang vertically; the vertical displacements from equilibrium of the two masses are y_1 and y_2, and are positive downwards. If the masses oscillate freely, find the displacements at any positive time t.

Solution.

♦

Let us interpret our result. The system has two modes of oscillation, which may occur together or separately (depending on the values of the arbitrary constants), as follows.

1. When $\sigma^2 = \lambda_1 = -1$, the eigenvector $(1,1)^T$ means that both masses have the same displacement, ie they ______________________________, with frequency ________.

2. When $\sigma^2 = \lambda_2 = -3$, the eigenvector $(1,-1)^T$ means that the two masses have equal but opposite displacements, in other words their movements are 180° out of phase. This mode has a higher frequency of __________.

Since the initial conditions give the values of a_1, a_2, b_1, b_2, we should be able to start the experiment in such a way that we see either pure in-phase motion ($a_2 = b_2 = 0$), pure anti-phase motion ($a_1 = b_1 = 0$), or some combination thereof.

4.6 REMARKS ON REPEATED EIGENVALUES

An $n \times n$ matrix always has n eigenvalues, but some of them may be repeated. In that case, the matrix may have $< n$ linearly independent[2] eigenvectors. However, the matrix may still have the full complement of eigenvectors[3] as does the matrix

$$A = \begin{pmatrix} 3 & 0 \\ 0 & 3 \end{pmatrix} .$$

For consider

$$0 = \det (A - \lambda I) = \begin{vmatrix} 3-\lambda & 0 \\ 0 & 3-\lambda \end{vmatrix} = (3-\lambda)^2 \implies \lambda = 3, 3$$

(ie one eigenvalue, 3, which is repeated). But the matrix A does have a pair of linearly independent eigenvectors, for example

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 0 \\ 1 \end{pmatrix} .$$

On the other hand, if we are solving a system of n ODEs and the $n \times n$ matrix of coefficients has repeated eigenvectors *and* has less than n linearly independent eigenvectors, then our method of solution will not work in its current form.

Example 4.6.1

Find the eigenvalues and eigenvectors of

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

Solution.

[2] We have already observed that if $\boldsymbol{x}$ is an eigenvector then $k\boldsymbol{x}$ is also, for any $k \neq 0$, but the key thing is generating linearly independent eigenvectors.

[3] There is a test to determine how may eigenvectors a matrix has, but this is for the future.

♦

As an **aside**, we can take a look at why this happens. Consider the ODE

$$\frac{d^2y}{dt^2} - 2\frac{dy}{dt} + y = 0 \ .$$

The auxiliary equation is $m^2 - 2m + 1 = 0$ with repeated root $m = 1$. We learned how to deal with this, finding that the general solution is

$$y = c_1 e^t + c_2 t e^t \ . \tag{3}$$

We now write this one second order ODE as two first order ODEs[4]. First, we set

$$y = y_1 \quad \text{and} \quad y_2 = \frac{dy_1}{dt} \ ,$$

so that equation (3) implies that

$$\frac{d^2y_1}{dt^2} = \frac{dy_2}{dt} = 2\frac{dy_1}{dt} - y_1 = 2y_2 - y_1 \ .$$

So in matrix form our one second order ODE becomes the system of two ODEs

$$\frac{d}{dt}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 2 \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \ .$$

We solved such systems by starting with

$$\boldsymbol{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \boldsymbol{x} e^{\lambda t} \ ,$$

which excludes the second part of (3) $\implies$ the matrix must be an eigenvector short.

[4]This is how you would solve this second order ODE using the built-in ODE solvers of MATLAB®, for example.

4.7 PRINCIPAL AXES OF AN ELLIPSE

A circle is inscribed on the face of a block of material. The block is then subject to compression and shear forces which deform the circle into an ellipse.

By choosing the *principal axes* of the ellipse as new coordinate axes we can replace these forces with compressions only. In other words, we wish to have a new coordinate system in which the ellipse is not set at an angle to the axes, but in which the major and minor axes of the ellipse are aligned with the coordinate axes. A picture might look like the following.

Let us see what the theory would need. Suppose that our ellipse is given by

$$4y^2 + 2yz + 4z^2 = 30\ .$$

The presence of the '$2yz$' term immediately tells us that this ellipse is at angle to the (x, y)-axes. It is this "cross term" that we wish to eliminate by choosing new coordinates (u, w) with which to describe the ellipse. To do so, we first write the ellipse as a *quadratic form*[5], ie put

$$Q = 4y^2 + 2yz + 4z^2 = \begin{pmatrix} y \\ z \end{pmatrix}^T \begin{pmatrix} a & b \\ b & c \end{pmatrix} \begin{pmatrix} y \\ z \end{pmatrix}$$

for some a,b and c (note the symmetry of the two bs). Expanding the RHS we get

[5]If you don't know what this means, then the following few lines should be enough to tell you.

So

$$a = \_\_\_\_\_, \quad 2b = \_\_\_\_\_ \implies b = \_\_\_\_\_, \quad c = \_\_\_\_\_ \ .$$

Annotate the following for a more easily-remembered shortcut:

$$Q = 4y^2 + 2yz + 4z^2 = \begin{pmatrix} y & z \end{pmatrix} \begin{pmatrix} 4 & 1 \\ 1 & 4 \end{pmatrix} \begin{pmatrix} y \\ z \end{pmatrix} \ .$$

As we said above, we need to change to new variables u and w so that

$$Q = \begin{pmatrix} u & w \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \begin{pmatrix} u \\ w \end{pmatrix} \tag{4}$$

$$= \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_$$

$$= \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_$$

because $\lambda_1 u^2 + \lambda_2 w^2 = 30$ is an ellipse with the u and w coordinate axes as its principle axes (note the absence of any cross term). The next section outlines how to define (u, w) and $\lambda_{1,2}$ in terms of the eigenvalues and eigenvectors of the matrix of the quadratic form, and to relate (u, w) to (y, z). This *diagonalization* of matrices — so called because everything off the diagonal becomes zero — is important not only for describing strains as mentioned above, but also for a range of numerical techniques.

4.8 DIAGONALIZING A MATRIX

An $n \times n$ matrix A is called *diagonalizable* if and only if there is a non-singular matrix P and a diagonal matrix D such that,

$$P^{-1}AP = D \ .$$

The following are all equivalent definitions of the diagonalizability of A:

$$P^{-1}AP = \underline{\hspace{20em}} \, .$$

Aside: note that such a P exists $\iff$ A has n linearly independent eigenvectors.

For example, for a 2×2 matrix A, $AP = PD$ gives,

$$A \begin{pmatrix} \boldsymbol{x}_1 & \boldsymbol{x}_2 \end{pmatrix} = \begin{pmatrix} \boldsymbol{x}_1 & \boldsymbol{x}_2 \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} ,$$

where the column vector $\boldsymbol{x}_i$ is the i^{th} column of P and λ_i is the i^{th} diagonal element of D. Multiplying out

$$\begin{pmatrix} A\boldsymbol{x}_1 & A\boldsymbol{x}_2 \end{pmatrix} = \begin{pmatrix} \lambda_1 \boldsymbol{x}_1 + 0 & 0 + \lambda_2 \boldsymbol{x}_2 \end{pmatrix} ,$$

and since these two matrices are equal, their columns are equal, which upon examination reveals that the $\lambda_i, \boldsymbol{x}_i$ are the eigenvalue/vector pairs of A.

To diagonalize the $n \times n$ matrix A:

1. Find the eigenvalue/eigenvector pairs $\lambda_i, \boldsymbol{x}_i$.
2. Form a matrix P whose columns are the eigenvectors.
3. Form a diagonal matrix D whose diagonal elements are the eigenvalues *in the same order in which the eigenvectors were placed in P.*
4. Then $P^{-1}AP = D$.

Although it seems as if we don't really use P, the matrix of eigenvectors, we will shortly see that P is essential for finding the principal axes of an ellipse (amongst other uses).

Example 4.8.1

Diagonalise the following matrix:

$$A = \begin{pmatrix} 4 & 1 \\ 1 & 4 \end{pmatrix} .$$

Solution.

♦

Example 4.8.2

Return to the motivating example from the beginning of this lecture, and find the principle axis of the ellipse in question.

***Solution*.**

♦

The new coordinate axes (u, w) can be sketched on the old (y, z) axes by considering

$$\begin{pmatrix} u \\ w \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 0 \\ 1 \end{pmatrix} .$$

To find the principal axes of an ellipse:

1. Write the equation for the ellipse in matrix form as on page 23.
2. Work through the box on page 25 to diagonalise the matrix.
3. Now normalise the eigenvectors to get a new matrix P (as in above example, and formalised in next lecture).
4. Use the diagonal matrix in the quadratic form to find the equation of the ellipse in the new (principal axes) coordinate system (as above).
5. Use P to find how the new coordinate system relates to the old (as above).
6. Sketch the new axes on your diagram, if required.

See also the thoroughly-worked example in the next lecture.

4.9 ORTHOGONAL MATRICES

In the previous lecture it was essential that the matrix P had the property that its transpose was equal to its inverse. Matrices with the property that $P^T = P^{-1}$ are called *orthogonal*. In this short section, we examine a little bit of the theory behind orthogonal matrices. From the definition, a simple test as to whether a given matrix is orthogonal is to check if you get the identity matrix when you multiply the matrix by its transpose.

Example 4.9.1

Is the following matrix P orthogonal?

$$P = \frac{1}{5}\begin{pmatrix} 4 & 3 \\ -3 & 4 \end{pmatrix} .$$

Solution.

♦

So it looks like this matrix is a good candidate for being orthogonal. However, we would normally have to show that both $AB = I$ and $BA = I$ to establish that $B = A^{-1}$. For *matrices* doing one is actually enough. So in fact $\boxed{\text{if } P^TP = I \text{ then } P \text{ is orthogonal}}$.

Any orthogonal matrix P can be written in terms of its columns, each column being a column vector: $P = (\boldsymbol{x}_1\ \boldsymbol{x}_2\ \boldsymbol{x}_3)$, so that $P^T P = I$ gives

$$\begin{pmatrix} \boldsymbol{x}_1 \\ \boldsymbol{x}_2 \\ \boldsymbol{x}_3 \end{pmatrix} \begin{pmatrix} \boldsymbol{x}_1 & \boldsymbol{x}_2 & \boldsymbol{x}_3 \end{pmatrix} = \begin{pmatrix} \boldsymbol{x}_1^T\boldsymbol{x}_1 & \boldsymbol{x}_1^T\boldsymbol{x}_2 & \boldsymbol{x}_1^T\boldsymbol{x}_3 \\ \boldsymbol{x}_2^T\boldsymbol{x}_1 & \boldsymbol{x}_2^T\boldsymbol{x}_2 & \boldsymbol{x}_2^T\boldsymbol{x}_3 \\ \boldsymbol{x}_3^T\boldsymbol{x}_1 & \boldsymbol{x}_3^T\boldsymbol{x}_2 & \boldsymbol{x}_3^T\boldsymbol{x}_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

where $\boldsymbol{x}_i^T \boldsymbol{x}_j \equiv \boldsymbol{x}_i \cdot \boldsymbol{x}_j$, (ie the dot product).

By inspection,

$$\boldsymbol{x}_1 \cdot \boldsymbol{x}_1 = \boldsymbol{x}_2 \cdot \boldsymbol{x}_2 = \boldsymbol{x}_3 \cdot \boldsymbol{x}_3 = 1\ .$$

In other words, all the columns of P have ____________________. Also,

$$\boldsymbol{x}_i \cdot \boldsymbol{x}_j = 0\ , \quad i \neq j\ ,$$

ie all the columns of P are _________________________.

So when diagonalising a matrix we use

$$P = \begin{pmatrix} \hat{\boldsymbol{x}}_1 & \hat{\boldsymbol{x}}_2 & \ldots & \hat{\boldsymbol{x}}_n \end{pmatrix}$$

where $\hat{\boldsymbol{x}}_i$ is the normalised eigenvector $\boldsymbol{x}_i / \|\boldsymbol{x}_i\|$.

Example 4.9.2

Show that for eigenvalues of a symmetric matrix A, different eigenvectors are always mutually perpendicular provided they have different eigenvalues.

Solution.

♦

Note that if $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ share the same eigenvalue then the eigenvectors can be *chosen* to be perpendicular, but this is beyond the scope of this course.

Example 4.9.3

Apply everything we have learned to the ellipse

$$Q = 6y^2 + 4yz + 9z^2 = 20 \ .$$

Solution.