Eigenvalues and Eigenvectors
This chapter is based on notes written by Dr Helen Wilson

0.1 Eigenvalues and eigenvectors

0.1.1 Background: polynomials

An nth-order polynomial is a function of the form

$$f(x) = c_0 + c_1 x + c_2 x^2 + \cdots + c_n x^n$$

with $c_n \neq 0$.

If $f(\alpha) = 0$ then we say $x = \alpha$ is a root of the equation $f(x) = 0$, and $f(x)$ can be factored: $f(x) = (x - \alpha)q(x)$.

Every polynomial of order n (i.e. with highest power x^n) can be broken down into n factors:

$$c_0 + c_1 x + c_2 x^2 + \cdots + c_n x^n = c_n(x - \alpha_1)(x - \alpha_2)(x - \alpha_3) \cdots (x - \alpha_n)$$

although the numbers α_1, α_2 etc. don't have to be all different and may be complex.

If all the coefficients c_0, c_1, $\ldots c_n$ are real, then any complex roots appear in complex conjugate pairs: $a + ib$ and $a - ib$.

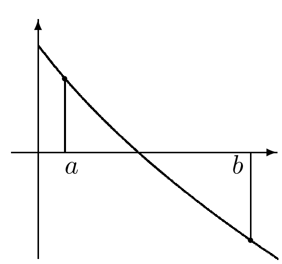
There are two things we need be able to do to find all the roots of a polynomial:

- Find a root α; and
- Given one root, find the new (smaller) polynomial $q(x)$.

Finding roots

To find roots of the equation $f(x) = 0$:

- Guess easy values: 0, 1, -1, 2, -2 and so on.
- If $f(a) = 0$ then a is a root
- If $f(a) > 0$ and $f(b) < 0$ then there is a root between a and b:



- Guess factors of c_0 (e.g. if $c_0 = 6$, try ± 1, ± 2, ± 3, ± 6).

Quadratics

You probably know the formula for quadratics, but there are nicer (and quicker) methods.

An easy one:

$$x^2 + 6x + 9 = (x + 3)^2$$

Now it follows that

$$x^2 + 6x + 13 = (x + 3)^2 + 4$$

so to solve $x^2 + 6x + 13 = 0$ we need

$$(x+3)^2+4 = 0 \quad (x+3)^2 = -4 \quad x+3 = \pm 2i \quad x = -3\pm 2i \text{ so } x^2+6x+13 = (x+3-2i)(x+3+2i).$$

Similarly,

$$x^2+6x+8 = (x+3)^2-1 \qquad x^2+6x+8 = 0 \implies (x+3)^2 = 1 \qquad x+3 = \pm 1 \qquad x = -2 \text{ or } -4.$$

This procedure:

$$x^2 + 2nx + c \quad \rightarrow \quad (x + n)^2 + c - n^2$$

is called **completing the square**.

If the numbers work out, there is an even quicker method. Look at the general factored quadratic (without a constant at the front):

$$(x - a)(x - b) = x^2 - (a + b) + ab$$

Now, looking at a specific case, e.g.

$$x^2 + 4x + 3 = 0$$

we want to match the two: so we need

$$\begin{aligned} &ab = 3 \\ &a + b = -4 \end{aligned}$$

The first one immediately suggests either 1 and 3 or −1 and −3; the second one tells us that −1 and −3 are the roots so

$$x^2 + 4x + 3 = (x + 1)(x + 3).$$

Another example:

$$x^2 - x - 12 = 0.$$

Now the factors may be 1 and 12, 2 and 6, or 3 and 4 (with a minus sign somewhere). The "sum" (which needs to be the difference because of the minus sign) should be 1 so we quickly reach

$$\text{Roots } x = 4 \text{ and } x = -3, \;\; x^2 - x - 12 = (x - 4)(x + 3).$$

Factorising polynomials

The **long division** method you learnt years ago for dividing large numbers works in a very similar way for polynomials.

Example: $f(x) = x^3 + 1$. We spot a root $x = -1$ which means there is a factor of $(x + 1)$, so $x^3 + 1 = (x + 1)q(x)$.

$$\begin{array}{rlrrrrrr}
 & & x^2 & - & x & + & 1 & \\ \hline
x+1 &) & x^3 & + & 0x^2 & + & 0x & + \ 1 \\
 & & x^3 & + & x^2 & & & \\ \hline
 & & & & -x^2 & + & 0x & \\
 & & & & -x^2 & - & x & \\ \hline
 & & & & & & x & + \ 1
\end{array}$$

At each stage we choose a multiple of our $(x+1)$ factor that matches the highest power in the remainder we are trying to get rid of; we write that multipying factor above the line; we write the multiple below; then we subtract to get the next remainder. In this case $q(x) = x^2 - x + 1$ which has complex roots.

Example: $f(x) = x^3 + 10x^2 + 31x + 30$. This time we don't immediately spot a factor, so try a few values. Note first that if $x > 0$ all the terms are positive so $f(x) > 0$; so all the roots must be negative.

$$\begin{aligned}
x = 0 &\quad : \quad f(0) = 30 \\
x = -1 &\quad : \quad f(-1) = -1 + 10 - 31 + 30 = 8 \\
x = -2 &\quad : \quad f(-2) = -8 + 40 - 62 + 30 = 0
\end{aligned}$$

so we can write $f(x) = (x + 2)q(x)$.

$$\begin{array}{rlrrrrrr}
 & & x^2 & + & 8x & + & 15 & \\ \hline
x+2 &) & x^3 & + & 10x^2 & + & 31x & + \ 30 \\
 & & x^3 & + & 2x^2 & & & \\ \hline
 & & & & 8x^2 & + & 31x & \\
 & & & & 8x^2 & + & 16x & \\ \hline
 & & & & & & 15x & + \ 30
\end{array}$$

In this case we have $q(x) = x^2 + 8x + 15 = (x + 3)(x + 5)$ so $f(x) = (x + 2)(x + 3)(x + 5)$.

There is another method; simply write down the most general form for $q(x)$ and then work out the coefficients of each power of x.

Example: $f(x) = x^3 + x^2 + x + 1$. Spot a root $x = -1$. Then write

$$f(x) = (x + 1)(ax^2 + bx + c) = \begin{array}{rrrrrrr} ax^3 & + & bx^2 & + & cx & & \\ & + & ax^2 & + & bx & + & c \end{array}$$

so we have these equations for the coefficients of the powers of x:

$$\begin{aligned}
x^3 &\quad : \quad a = 1 \\
x^2 &\quad : \quad a + b = 1 \implies b = 0 \\
x &\quad : \quad b + c = 1 \implies c = 1 \\
1 &\quad : \quad c = 1 \text{ which agrees.}
\end{aligned}$$

Thus $q(x) = x^2 + 1$ and $f(x) = (x + 1)(x + i)(x - i)$.

0.1.2 Motivation

If $\underline{\underline{A}}$ is a square matrix, there is often no obvious relationship between the vector $\underline{v}$ and its image $\underline{\underline{A}}\underline{v}$ under multiplication by $\underline{\underline{A}}$. However frequently there are some nonzero vectors that $\underline{\underline{A}}$ maps into scalar multiple of themselves. Such vectors (called eigenvectors) arise naturally in the study of vibrations, electrical systems, genetics, chemical reactions, quantum mechanics, mechanical stress, economics and geometry.

0.1.3 Definitions

For a square matrix $\underline{\underline{A}}$, if

$$\underline{\underline{A}}\,\underline{v} = \lambda \underline{v}$$

with $\underline{v} \neq \underline{0}$ then $\underline{v}$ is an **eigenvector** of $\underline{\underline{A}}$ with **eigenvalue** λ.

When is this possible?

$$\underline{\underline{A}}\,\underline{v} = \lambda \underline{\underline{I}}\,\underline{v} \qquad (\underline{\underline{A}} - \lambda \underline{\underline{I}})\,\underline{v} = \underline{0}$$

This is just an ordinary homogeneous linear system. Remember, if the determinant of the matrix on the left is non-zero then there is a unique solution $\underline{v} = \underline{0}$; for a **nontrivial** solution we need

$$\det(\underline{\underline{A}} - \lambda \underline{\underline{I}}) = 0.$$

Suppose this determinant is zero; then our equation will have an infinite number of solutions, $\underline{v}$ and any multiple of it.

The determinant $\det(\underline{\underline{A}} - \lambda \underline{\underline{I}})$ is a polynomial in λ of degree n, so there are at most n different eigenvalues.

A useful fact (not proved here) is $\det(\underline{\underline{A}}) = \lambda_1 \lambda_2 \cdots \lambda_n$.

Example

$$\underline{\underline{A}} = \begin{pmatrix} 5 & -2 \\ 9 & -6 \end{pmatrix}.$$

$$\begin{aligned} |\underline{\underline{A}} - \lambda \underline{\underline{I}}| &= \begin{vmatrix} 5-\lambda & -2 \\ 9 & -6-\lambda \end{vmatrix} = (5-\lambda)(-6-\lambda) - (-2)(9) \\ &= (-30 + \lambda + \lambda^2) + 18 = \lambda^2 + \lambda - 12 = (\lambda + 4)(\lambda - 3) \end{aligned}$$

so the matrix has eigenvalues $\lambda_1 = 3$, $\lambda_2 = -4$.

$\lambda_1 = 3$:

$$(\underline{\underline{A}} - \lambda_1 \underline{\underline{I}})\underline{v}_1 = \underline{0} \;\text{ and }\; \underline{v}_1 = \begin{pmatrix} a \\ b \end{pmatrix} \implies \begin{pmatrix} 2 & -2 \\ 9 & -9 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

$$a - b = 0 \qquad a = b \qquad \underline{v}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \text{ (or any multiple of this).}$$

Check:

$$\underline{\underline{A}}\,\underline{v}_1 = \begin{pmatrix} 5 & -2 \\ 9 & -6 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 3 \end{pmatrix} = 3 \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \lambda_1 \underline{v}_1.$$

$\lambda_2 = -4$:

$$(\underline{\underline{A}} - \lambda_2 \underline{\underline{I}})\underline{v}_2 = \underline{0} \text{ and } \underline{v}_2 = \begin{pmatrix} a \\ b \end{pmatrix} \implies \begin{pmatrix} 9 & -2 \\ 9 & -2 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

$$9a - 2b = 0 \qquad 9a = 2b \qquad \underline{v}_2 = \begin{pmatrix} 2 \\ 9 \end{pmatrix} \text{ (or any multiple of this).}$$

Check:

$$\underline{\underline{A}}\,\underline{v}_2 = \begin{pmatrix} 5 & -2 \\ 9 & -6 \end{pmatrix} \begin{pmatrix} 2 \\ 9 \end{pmatrix} = \begin{pmatrix} -8 \\ -36 \end{pmatrix} = -4 \begin{pmatrix} 2 \\ 9 \end{pmatrix} = \lambda_2 \underline{v}_2.$$

Another example

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

Again, we look for eigenvalues using the determinant:

$$|\underline{\underline{A}} - \lambda \underline{\underline{I}}| = \begin{vmatrix} 1-\lambda & -1 \\ 2 & -1-\lambda \end{vmatrix} = (1-\lambda)(-1-\lambda) + 2 = \lambda^2 + 1$$

so the two roots of this equation are $\lambda = \pm i$.

Eigenvector and eigenvalue properties

- Eigenvalue and eigenvector pair satisfy

$$\underline{\underline{A}}\,\underline{v} = \lambda \underline{v} \qquad \text{and} \qquad \underline{v} \neq \underline{0}.$$

- λ is allowed to be zero
- λ is allowed to be complex: but if $a + ib$ is an eigenvalue so is $a - ib$, and the eigenvectors corresponding to these eigenvalues will also be a complex conjugate pair (as long as $\underline{\underline{A}}$ is real)
- The same eigenvalue may appear more than once.
- Eigenvectors corresponding to different eigenvalues are linearly independent (and explain).

Example

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

First we need to solve the determinant equation:

$$\det(\underline{\underline{A}} - \lambda \underline{\underline{I}}) = 0$$

$$
\begin{aligned}
0 &= \begin{vmatrix} -\lambda & -1 & -3 \\ 2 & 3-\lambda & 3 \\ -2 & 1 & 1-\lambda \end{vmatrix} = (-\lambda)\begin{vmatrix} 3-\lambda & 3 \\ 1 & 1-\lambda \end{vmatrix} - (-1)\begin{vmatrix} 2 & 3 \\ -2 & 1-\lambda \end{vmatrix} + (-3)\begin{vmatrix} 2 & 3-\lambda \\ -2 & 1 \end{vmatrix} \\
&= (-\lambda)\{(3-\lambda)(1-\lambda)-3\} + \{2(1-\lambda)+6\} - 3\{2+2(3-\lambda)\} \\
&= (-\lambda)\{3-4\lambda+\lambda^2-3\} + \{2-2\lambda+6\} - \{6+18-6\lambda\} \\
&= (-1)\{-4\lambda^2+\lambda^3\} + \{8-2\lambda\} - \{6+18-6\lambda\} \\
&= -\lambda^3+4\lambda^2+4\lambda-16 = -(\lambda-2)(\lambda^2-2\lambda-8) = -(\lambda-2)(\lambda+2)(\lambda-4)
\end{aligned}
$$

This is zero if $\lambda = 2$, $\lambda = -2$ or $\lambda = 4$ so these are the three eigenvalues.

Now for each eigenvector in turn, we solve the equation $(\underline{\underline{A}} - \lambda\underline{\underline{I}})\underline{v} = \underline{0}$.

First eigenvector: $\lambda_1 = 2$.

$$
(\underline{\underline{A}} - 2\underline{\underline{I}})\underline{v}_1 = \underline{0} \qquad \begin{pmatrix} -2 & -1 & -3 \\ 2 & 1 & 3 \\ -2 & 1 & -1 \end{pmatrix}\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}
$$

We carry out a single Gaussian elimination step here ($r_2 \to r_2 + r_1$ and $r_3 \to r_3 - r_1$):

$$
\begin{pmatrix} -2 & -1 & -3 \\ 0 & 0 & 0 \\ 0 & 2 & 2 \end{pmatrix}\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}
$$

Now we can see that there is a zero row (which we expected because we made the determinant zero) and two equations to solve by back substitution. When we circle the leading term in each row there is no circle in the third column so we rename $c = \alpha$. Then r_3 gives

$$
\begin{aligned}
r_3 &\implies 2b + 2c = 0 \qquad b = -\alpha \\
r_1 &\implies -2a - b - 3c = 0 \qquad a = -\alpha
\end{aligned}
$$

and so the general eigenvector is

$$
\underline{v}_1 = \begin{pmatrix} -\alpha \\ -\alpha \\ \alpha \end{pmatrix} \qquad \text{and we can use any multiple of it:} \qquad \underline{v}_1 = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}.
$$

Second eigenvector: $\lambda_2 = -2$

$$
(\underline{\underline{A}} + 2\underline{\underline{I}})\underline{v}_2 = \underline{0} \qquad \begin{pmatrix} 2 & -1 & -3 \\ 2 & 5 & 3 \\ -2 & 1 & 3 \end{pmatrix}\begin{pmatrix} d \\ e \\ f \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}
$$

Again, we eliminate: $r_2 \to r_2 - r_1$ and $r_3 \to r_3 + r_1$:

$$
\begin{pmatrix} 2 & -1 & -3 \\ 0 & 6 & 6 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} d \\ e \\ f \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}
$$

Just like last time, when we circle the leading elements in the rows there is no circle in column 3 so we can set $f = \alpha$ and back substitute:

$$
\begin{aligned}
r_2 &\implies 6e + 6f = 0 \qquad e = -\alpha \\
r_1 &\implies 2d - e - 3f = 0 \qquad d = \alpha
\end{aligned}
$$

which gives us the solution

$$\begin{pmatrix} d \\ e \\ f \end{pmatrix} = \begin{pmatrix} \alpha \\ -\alpha \\ \alpha \end{pmatrix} \qquad \underline{v}_2 = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}.$$

Third eigenvector: $\lambda_3 = 4$

$$(\underline{\underline{A}} - 4\underline{\underline{I}})\underline{v}_3 = \underline{0} \qquad \begin{pmatrix} -4 & -1 & -3 \\ 2 & -1 & 3 \\ -2 & 1 & -3 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

This time it makes sense to use non-standard Gaussian elimination: $r_3 \to r_3 + r_2$ followed by $r_2 \to 2r_2 + r_1$.

$$\begin{pmatrix} -4 & -1 & -3 \\ 0 & -3 & 3 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

and, again, we can set $c = \alpha$

$$\begin{aligned} r_2 &\implies -3b + 3c = 0 \qquad b = \alpha \\ r_1 &\implies -4a - b - 3c = 0 \qquad a = -\alpha \end{aligned}$$

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} -\alpha \\ \alpha \\ \alpha \end{pmatrix} \qquad \underline{v}_3 = \begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix}.$$

Examples with repeated eigenvalues

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

Since this is a diagonal matrix, its eigenvalues are just the values on the diagonal (check with the full determinant if you're not convinced). In this case we have $\lambda = 1$, $\lambda = 2$ and $\lambda = 2$ again. The eigenvector for $\lambda = 1$ is perfectly normal:

$$(\underline{\underline{A}} - \underline{\underline{I}})\underline{v}_1 = \underline{0} \qquad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \qquad \underline{v}_1 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

When we come to solve with $\lambda = 2$ though, the eigenvector is not uniquely defined:

$$(\underline{\underline{A}} - 2\underline{\underline{I}})\underline{v}_2 = \underline{0} \qquad \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} d \\ e \\ f \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

which only tells us that $f = 0$ and doesn't fix d and e. In fact in this case we can choose any two *different* pairs d and e as our two eigenvectors for the repeated eigenvalue $\lambda = 2$,

$$\text{e.g. } \underline{v}_2 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \underline{v}_3 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \qquad \text{or} \qquad \underline{v}_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad \underline{v}_3 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$$

and in fact any *linear combination* $\alpha\underline{v}_2 + \beta\underline{v}_3$ will also be an eigenvector of $\underline{\underline{A}}$ with eigenvalue 2.

This doesn't always work, though. Here is another example:

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

Just like last time, the eigenvalues are $\lambda_1 = 1$, $\lambda_2 = \lambda_3 = 2$.

Look for the eigenvector of $\lambda_1 = 1$:

$$(\underline{\underline{A}} - \lambda_1 \underline{\underline{I}})\underline{v}_1 = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \qquad \underline{v}_1 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Now the eigenvector(s) of $\lambda_2 = 2$:

$$(\underline{\underline{A}} - \lambda_2 \underline{\underline{I}})\underline{v}_2 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Again, there is just one solution:

$$\underline{v}_2 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}.$$

In this case there are only two eigenvectors.

Review of eigenvalues and eigenvectors

- Eigenvalue and eigenvector pair satisfy

  $$\underline{\underline{A}}\,\underline{v} = \lambda \underline{v} \qquad \text{and} \qquad \underline{v} \neq \underline{0}.$$

  We find the eigenvalues by solving the polynomial $\det(\underline{\underline{A}} - \lambda\underline{\underline{I}}) = 0$ and then find each eigenvector by solving the linear system $(\underline{\underline{A}} - \lambda\underline{\underline{I}})\underline{v} = \underline{0}$.

- λ is allowed to be zero or complex.

- The same eigenvalue may appear more than once; if it does, we may have a choice of eigenvectors or a missing one.

- The product of the eigenvalues is the determinant

- The sum of the eigenvalues is the **trace** of the matrix; the sum of the elements on the leading diagonal.

0.1.4 Commuting matrices

A pair of matrices $\underline{\underline{A}}$ and $\underline{\underline{B}}$ are said to *commute* if $\underline{\underline{A}}\,\underline{\underline{B}} = \underline{\underline{B}}\,\underline{\underline{A}}$. If two $n \times n$ matrices commute, and they both have n distinct eigenvalues, then they have **the same eigenvectors.**

Proof

Look at an eigenvector of $\underline{\underline{A}}$. We know that $\underline{\underline{A}}\,\underline{v} = \lambda\underline{v}$. Now let $\underline{u} = \underline{\underline{B}}\,\underline{v}$. Then

$$\underline{\underline{A}}\,\underline{u} = \underline{\underline{A}}\,\underline{\underline{B}}\,\underline{v} = \underline{\underline{B}}\,\underline{\underline{A}}\,\underline{v} = \underline{\underline{B}}\lambda\underline{v} = \lambda\underline{\underline{B}}\,\underline{v} = \lambda\underline{u}$$

We've shown

$$\underline{\underline{A}}\,\underline{u} = \lambda\underline{u}$$

which means $\underline{u}$ is an eigenvector of $\underline{\underline{A}}$ with eigenvalue λ. Since the eigenvector of $\underline{\underline{A}}$ corresponding to λ was $\underline{v}$, it follows that $\underline{u}$ must be a multiple of $\underline{v}$:

$$\underline{u} = \mu\underline{v} \qquad \underline{\underline{B}}\,\underline{v} = \mu\underline{v}$$

so $\underline{v}$ is an eigenvector of $\underline{\underline{B}}$ with eigenvalue μ.

Example

$$\underline{\underline{A}} = \begin{pmatrix} 3 & 2 \\ -1 & 0 \end{pmatrix} \qquad \underline{\underline{B}} = \begin{pmatrix} 1 & -2 \\ 1 & 4 \end{pmatrix}$$

These commute:

$$\underline{\underline{A}}\,\underline{\underline{B}} = \begin{pmatrix} 3 & 2 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} 1 & -2 \\ 1 & 4 \end{pmatrix} = \begin{pmatrix} 5 & 2 \\ -1 & 2 \end{pmatrix} \qquad \underline{\underline{B}}\,\underline{\underline{A}} = \begin{pmatrix} 1 & -2 \\ 1 & 4 \end{pmatrix}\begin{pmatrix} 3 & 2 \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} 5 & 2 \\ -1 & 2 \end{pmatrix}$$

Let's look at $\underline{\underline{A}}$ first. The eigenvalues:

$$\det(\underline{\underline{A}} - \lambda\underline{\underline{I}}) = \begin{vmatrix} 3-\lambda & 2 \\ -1 & -\lambda \end{vmatrix} = (3-\lambda)(-\lambda)+2 = \lambda^2 - 3\lambda + 2 = (\lambda-1)(\lambda-2).$$

The eigenvalues are $\lambda_1 = 1$ and $\lambda_2 = 2$.

First eigenvector: $\lambda_1 = 1$.

$$(\underline{\underline{A}}-\underline{\underline{I}})\underline{v}_1 = \begin{pmatrix} 2 & 2 \\ -1 & -1 \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \implies a+b=0 \qquad \underline{v}_1 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

Second eigenvector: $\lambda_2 = 2$.

$$(\underline{\underline{A}}-2\underline{\underline{I}})\underline{v}_1 = \begin{pmatrix} 1 & 2 \\ -1 & -2 \end{pmatrix}\begin{pmatrix} c \\ d \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \implies c+2d=0 \qquad \underline{v}_2 = \begin{pmatrix} 2 \\ -1 \end{pmatrix}.$$

Now we see what effect $\underline{\underline{B}}$ has on these eigenvectors:

$$\underline{\underline{B}}\,\underline{v}_1 = \begin{pmatrix} 1 & -2 \\ 1 & 4 \end{pmatrix}\begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 3 \\ -3 \end{pmatrix} = 3\underline{v}_1 \qquad \underline{\underline{B}}\,\underline{v}_2 = \begin{pmatrix} 1 & -2 \\ 1 & 4 \end{pmatrix}\begin{pmatrix} 2 \\ -1 \end{pmatrix} = \begin{pmatrix} 4 \\ -2 \end{pmatrix} = 2\underline{v}_2.$$

Note

Because $\underline{\underline{A}}$ commutes with itself, it follows that $\underline{\underline{A}}$ and $\underline{\underline{A}}^2$ (and all other powers of $\underline{\underline{A}}$) have the same eigenvectors.

0.1.5 Symmetric matrices

If our real matrix $\underline{\underline{A}}$ is symmetric then all its eigenvalues are real.

The eigenvectors corresponding to any two different eigenvalues will be **orthogonal**, i.e. $\underline{v}_i \cdot \underline{v}_j = 0$ (or $\underline{v}_i^\top \underline{v}_j = 0$) if $\lambda_i \neq \lambda_j$.

For equal eigenvalues, there will still be a complete set of eigenvectors, and we can choose the constants in the eigenvectors for the repeated eigenvalue so that all the eigenvectors are orthogonal.

Example

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

First we look for the eigenvalues:

$$\begin{aligned} \det(\underline{\underline{A}} - \lambda \underline{\underline{I}}) &= \begin{vmatrix} -\lambda & 1 & 1 \\ 1 & -\lambda & 1 \\ 1 & 1 & -\lambda \end{vmatrix} = (-\lambda) \begin{vmatrix} -\lambda & 1 \\ 1 & -\lambda \end{vmatrix} - (1) \begin{vmatrix} 1 & 1 \\ 1 & -\lambda \end{vmatrix} + (1) \begin{vmatrix} 1 & -\lambda \\ 1 & 1 \end{vmatrix} \\ &= (-\lambda)\{\lambda^2 - 1\} - \{-\lambda - 1\} + \{1 + \lambda\} = -\lambda^3 + \lambda + \lambda + 1 + 1 + \lambda = -\lambda^3 + 3\lambda + 2 \end{aligned}$$

We can spot the root $\lambda = -1$, and factorise:

$$-\lambda^3 + 3\lambda + 2 = -(\lambda^3 - 3\lambda - 2) = -(\lambda + 1)(\lambda^2 - \lambda - 2) = -(\lambda + 1)(\lambda + 1)(\lambda - 2)$$

so we have the eigenvalue $\lambda = -1$ twice, and the eigenvalue $\lambda = 2$.

We'll find the straightforward one first:

$$(\underline{\underline{A}} - 2\underline{\underline{I}})\underline{v}_1 = \begin{pmatrix} -2 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

Gaussian elimination: $r_2 \to 2r_2 + r_1$; $r_3 \to 2r_3 + r_1$:

$$\begin{pmatrix} -2 & 1 & 1 \\ 0 & -3 & 3 \\ 0 & 3 & -3 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

Now the last row is clearly a multiple of r_2 so we can set $c = \alpha$ and back-substitute to get $b = \alpha$ and $a = \alpha$.

$$\underline{v}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

Now onto the repeated root:

$$(\underline{\underline{A}} + \underline{\underline{I}})\underline{v}_2 = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} d \\ e \\ f \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

All three rows are the same so Gaussian elimination would just leave us with the equation $d + e + f = 0$. Note that even at this stage we can see any vector we choose will be orthogonal to $\underline{v}_1$:

$$\underline{v}_1^\top \underline{v}_2 = \begin{pmatrix} 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} d \\ e \\ f \end{pmatrix} = d + e + f = 0.$$

Now we want to choose **two** vectors satisfying $d+e+f = 0$ which are orthogonal to one another.

Choosing the first one is easy: you have free choice. Keep it as simple as possible; use zeros where you can. The obvious choice is probably

$$\underline{v}_2 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}.$$

Now for the second, we need to choose a new vector that satisfies two things:

$$\underline{v}_3 = \begin{pmatrix} d' \\ e' \\ f' \end{pmatrix} \qquad d'+e'+f' = 0 \qquad \underline{v}_2^\top \underline{v}_3 = 0 \implies \begin{pmatrix} 1 & 0 & -1 \end{pmatrix} \begin{pmatrix} d' \\ e' \\ f' \end{pmatrix} = d'-f' = 0.$$

Now we have two linear equations in d', e' and f':

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & -1 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} d' \\ e' \\ f' \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

and we can solve this system by Gaussian elimination and back substitution in the usual way. $r_2 \to r_2 - r_1$:

$$\begin{pmatrix} 1 & 1 & 1 \\ 0 & -1 & -2 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} d' \\ e' \\ f' \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \qquad \begin{array}{ll} f' = \alpha & \\ -e' - 2f' = 0, & e' = -2\alpha \\ d' + e' + f' = 0, & d' = \alpha \end{array} \qquad \underline{v}_3 = \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}.$$

0.1.6 Power method

Iterative methods

Remember Newton's method for finding a root of the equation $f(x) = 0$:

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}.$$

This is an **iterative method**: we start with an initial guess x_0 and keep applying the rule until we get a root. Although it isn't guaranteed to work, if we start close enough to a root we will end up there.

Why do we need iterative methods?

We usually find eigenvalues by solving the **characteristic equation** of our $n \times n$ matrix $\underline{\underline{A}}$:

$$\det \left(\underline{\underline{A}} - \lambda \underline{\underline{I}} \right) = 0.$$

When n is large, this is a huge polynomial and it isn't easy to find the roots. If we really need all the roots, then using something like Newton's method may be the best way; but very often in engineering problems all we need is the largest eigenvalue (e.g. the dominant mode of vibration will produce the largest amplitudes).

Using eigenvectors as a basis

If we have a set of n linearly independent vectors in n-dimensional space then we can say that they form a **basis** of the space, and we can write any n-dimensional vector $\underline{z}$ can be written as a linear combination of them:

$$\underline{z} = \alpha_1 \underline{v}_1 + \alpha_2 \underline{v}_2 + \cdots + \alpha_n \underline{v}_n$$

for some constants α_1, α_2, $\ldots \alpha_n$.

In our case, if our matrix has all real eigenvalues and a complete set of eigenvectors (as, for example, if it is symmetric), then we can use the eigenvectors of our matrix as the basis, so that we know how the vectors change when we multiply by $\underline{\underline{A}}$:

$$\underline{\underline{A}}\,\underline{z} = \alpha_1 \lambda_1 \underline{v}_1 + \alpha_2 \lambda_2 \underline{v}_2 + \cdots + \alpha_n \lambda_n \underline{v}_n.$$

$$\underline{\underline{A}}^2 \underline{z} = \alpha_1 \lambda_1^2 \underline{v}_1 + \alpha_2 \lambda_2^2 \underline{v}_2 + \cdots + \alpha_n \lambda_n^2 \underline{v}_n.$$

We need just one more assumption to make our method work: the eigenvalue with the largest magnitude must be strictly larger than all the others: $|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq |\lambda_4| \geq \cdots$.

In this case, after many multiplications the first term will be the largest:

$$\underline{\underline{A}}^k \underline{z} \approx \alpha_1 \lambda_1^k \underline{v}_1.$$

Definition of the Power method

The basic method (given on the left) is: choose an initial guess for the eigenvector and call it $\underline{z}_0$. Then iteratively carry out this process:

```
for k = 1, 2, ...
    find z_k = A z_{k-1} = A^k z_0
    form c_k = z_0^T z_k
    form μ_k = c_k / c_{k-1}   (not for k = 1)
    μ_k → λ_1
next k
```

```
for k = 1, 2, ...
    find y_k = A z_{k-1}
    let μ_k = maximum element of y_k
    form z_k = (1/μ_k) y_k
    μ_k → λ_1 and z_k → v_1
next k
```

In practice it is best to normalise the vector $\underline{\underline{A}}^k \underline{z}_0$ so that it has 1 as its largest element (see the method on the right, above). This process of normalisation helps prevent errors from growing. We can also extract an estimate of the eigenvector at the end. This is the essence of the **PageRank** algorithm (except that the sort of matrix we used there always has largest eigenvalue 1 so (a) we don't need to renormalise and (b) we want the eigenvector, not the eigenvalue).

Example

Let's look at a nice, small 2×2 matrix and use as our starting guess a simple vector with only one non-zero element:

$$\underline{\underline{A}} = \begin{pmatrix} 4 & 1 \\ 2 & 3 \end{pmatrix} \qquad \underline{z}_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

We'll use the normalising method, and work to 2 decimal places.

$k = 1$

$$\underline{y}_1 = \begin{pmatrix} 4 & 1 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 4 \\ 2 \end{pmatrix} \qquad \mu_1 = 4 \qquad \underline{z}_1 = \begin{pmatrix} 1 \\ 0.5 \end{pmatrix}$$

$k = 2$

$$\underline{y}_2 = \begin{pmatrix} 4 & 1 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} 1 \\ 0.5 \end{pmatrix} = \begin{pmatrix} 4.5 \\ 3.5 \end{pmatrix} \qquad \mu_2 = 4.5 \qquad \underline{z}_2 = \begin{pmatrix} 1 \\ 0.78 \end{pmatrix}$$

$k = 3$

$$\underline{y}_3 = \begin{pmatrix} 4 & 1 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} 1 \\ 0.78 \end{pmatrix} = \begin{pmatrix} 4.78 \\ 4.33 \end{pmatrix} \qquad \mu_3 = 4.78 \qquad \underline{z}_3 = \begin{pmatrix} 1 \\ 0.91 \end{pmatrix}$$

$k = 4$

$$\underline{y}_4 = \begin{pmatrix} 4 & 1 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} 1 \\ 0.91 \end{pmatrix} = \begin{pmatrix} 4.91 \\ 4.72 \end{pmatrix} \qquad \mu_4 = 4.91 \qquad \underline{z}_4 = \begin{pmatrix} 1 \\ 0.96 \end{pmatrix}$$

$k = 5$

$$\underline{y}_5 = \begin{pmatrix} 4 & 1 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} 1 \\ 0.96 \end{pmatrix} = \begin{pmatrix} 4.96 \\ 4.88 \end{pmatrix} \qquad \mu_4 = 4.96 \qquad \underline{z}_4 = \begin{pmatrix} 1 \\ 0.98 \end{pmatrix}$$

$k = 6$

$$\underline{y}_6 = \begin{pmatrix} 4 & 1 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} 1 \\ 0.98 \end{pmatrix} = \begin{pmatrix} 4.98 \\ 4.95 \end{pmatrix} \qquad \mu_4 = 4.98 \qquad \underline{z}_4 = \begin{pmatrix} 1 \\ 0.99 \end{pmatrix}$$

and we can see that $\mu_k \to 5$ and $\underline{z}_k \to \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ as the process continues.

Further examples if needed

$$\underline{\underline{A}} = \begin{pmatrix} 3 & 2 \\ -1 & 0 \end{pmatrix} \qquad \underline{\underline{B}} = \begin{pmatrix} 10 & -18 \\ 6 & -11 \end{pmatrix} \qquad \underline{\underline{C}} = \begin{pmatrix} 1 & 2 \\ -2 & 5 \end{pmatrix}$$