Chapter 3. Eigenvalues, diagonalisation and some applications

This material is a reduced treatment of what is in Anton & Rorres chapter 6, chapter 5 (mostly for the case of 3 dimensions) and sections 5.4, 6.5 and 10.5.

3.1 Orthogonal diagonalisation

We move on now to consider linear transformations $f \colon \mathbb{R}^3 \to \mathbb{R}^3$ that are given by diagonal matrices in some choice of frame or orthonormal basis. (We could do this in \mathbb{R}^2 and it would be slightly easier, or we could do it in \mathbb{R}^n for any n but that seems too abstract.)

If we start with a 3×3 diagonal matrix

$$A = \begin{bmatrix} \lambda_1 & 0 & 0\\ 0 & \lambda_2 & 0\\ 0 & 0 & \lambda_3 \end{bmatrix}$$

we get a linear transformation $f \colon \mathbb{R}^3 \to \mathbb{R}^3$, $f(\mathbf{x}) = A\mathbf{x}$ that has

$$f(\mathbf{i}) = \begin{bmatrix} \lambda_1 & 0 & 0\\ 0 & \lambda_2 & 0\\ 0 & 0 & \lambda_3 \end{bmatrix} \begin{bmatrix} 1\\ 0\\ 0 \end{bmatrix} = \begin{bmatrix} \lambda_1\\ 0\\ 0 \end{bmatrix} = \lambda_1 \mathbf{i}$$

and we can see similarly that

$$f(\mathbf{j}) = \lambda_2 \mathbf{j}$$
 and $f(\mathbf{k}) = \lambda_3 \mathbf{k}$.

We want to look at linear transformations f (or 3×3 matrices A) that behave like this, not on the standard basis $\mathbf{i}, \mathbf{j}, \mathbf{k}$ but instead on some other orthonormal basis $\mathbf{u}, \mathbf{v}, \mathbf{w}$.

3.1.1 Proposition. Suppose we have a 3×3 matrix A, and a corresponding linear transformation $f \colon \mathbb{R}^3 \to \mathbb{R}^3$ given by the rule $f(\mathbf{x}) = A\mathbf{x}$.

Then there is an orthonormal basis $\mathbf{u}, \mathbf{v}, \mathbf{w}$ and scalars $\lambda_1, \lambda_2, \lambda_3$ so that

$$f(\mathbf{u}) = \lambda_1 \mathbf{u}, f(\mathbf{v}) = \lambda_2 \mathbf{v} \text{ and } f(\mathbf{w}) = \lambda_3 \mathbf{w}$$

if only if there is an orthogonal matrix P so that

$$A = P \begin{bmatrix} \lambda_1 & 0 & 0\\ 0 & \lambda_2 & 0\\ 0 & 0 & \lambda_3 \end{bmatrix} P^t$$

Idea of the proof. We'll prove some of this, though none of it is really hard.

Suppose we have the orthonormal basis $\mathbf{u}, \mathbf{v}, \mathbf{w}$ and the scalars $\lambda_1, \lambda_2, \lambda_3$. Where do we find *P*?

Well it is the same change of basis matrix we used before in connection with rotations.

$$P = \begin{bmatrix} u_1 & v_1 & w_1 \\ u_3 & v_2 & w_3 \\ u_3 & v_3 & w_3 \end{bmatrix}$$

(where the columns come from $\mathbf{u}, \mathbf{v}, \mathbf{w}$). Recall that

$$P\mathbf{i} = \mathbf{u}, P\mathbf{j} = \mathbf{v}, P\mathbf{k} = \mathbf{w},$$

and

$$P^t \mathbf{u} = \mathbf{i}, P^t \mathbf{v} = \mathbf{j}, P^t \mathbf{w} = \mathbf{k}.$$

Look at the matrix

$$B = P \begin{bmatrix} \lambda_1 & 0 & 0\\ 0 & \lambda_2 & 0\\ 0 & 0 & \lambda_3 \end{bmatrix} P^t$$

(which we hope to show coincides with A).

If we calculate $B\mathbf{u}$ we get

$$B\mathbf{u} = P \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix} P^t \mathbf{u}$$
$$= P \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix} \mathbf{i}$$
$$= P(\lambda_1 \mathbf{i})$$
$$= \lambda_1 P^t \mathbf{i}$$
$$= \lambda_1 \mathbf{u}$$
$$= f(\mathbf{u}) = A\mathbf{u}$$

Similarly

$$B\mathbf{v} = A\mathbf{v}$$
 and $B\mathbf{v} = A\mathbf{v}$

Since every $\mathbf{x} \in \mathbb{R}^3$ can be written as a combination

$$\mathbf{x} = (\mathbf{x} \cdot \mathbf{u})\mathbf{u} + (\mathbf{x} \cdot \mathbf{v})\mathbf{v} + (\mathbf{x} \cdot \mathbf{w})\mathbf{w}$$

of $\mathbf{u}, \mathbf{v}, \mathbf{w}$ we can show now quite easily that $B\mathbf{x} = A\mathbf{x}$ always.

If we use this for $\mathbf{x} = \mathbf{i}$ we find

first column of
$$B =$$
 first column of A

and we can show the other columns must coincide by taking $\mathbf{x} = \mathbf{j}$ and $\mathbf{x} = \mathbf{k}$. So B = A.

Eigenvalues, etc

3.1.2 Observation. If

$$A = P \begin{bmatrix} \lambda_1 & 0 & 0\\ 0 & \lambda_2 & 0\\ 0 & 0 & \lambda_3 \end{bmatrix} P$$

then $A^t = A$. (We call a matrix with this property of being equal to its own transpose a *symmetric matrix*.)

Proof. Recall that the transpose of a product is the product of the transposes taken in the reverse order. So if A is as above then

$$A^{t} = (P^{t})^{t} \begin{bmatrix} \lambda_{1} & 0 & 0 \\ 0 & \lambda_{2} & 0 \\ 0 & 0 & \lambda_{3} \end{bmatrix}^{t} P^{t}$$
$$= P \begin{bmatrix} \lambda_{1} & 0 & 0 \\ 0 & \lambda_{2} & 0 \\ 0 & 0 & \lambda_{3} \end{bmatrix} P^{t}$$
$$= A \qquad \Box$$

3.1.3 Theorem. The $n \times n$ symmetric matrices A are exactly those that can be written

$$A = PDP^t$$

for an $n \times n$ orthogonal matrix P and an $n \times n$ diagonal matrix D.

3.1.4 Remark. The proof of this theorem is going to be beyond us. One bit is very easy and we've already given it in the 'Observation'. That is we showed (at least in the 3×3 case) that if $A = PDP^t$, then $A^t = A$.

The other part is the hard part and what we'll do is explain some of the ideas that you need to know if you want to use the result. That is, we'll explain how to find P and D starting with a symmetric A. We won't give any explanation of why it is always possible (in theory) to find P and D. We'll stick mainly to the 3×3 case, and the explanation here also skips over a complication that can arise (when there are only 2 different eigenvalues, rather than the usual 3).

To some extent Proposition 3.1.1 already has the bones of what we need, but it helps to have some more terminology to explain what to do.

3.1.5 Definition. A vector $\mathbf{v} \in \mathbb{R}^n$ is called an *eigenvector* for a square $n \times n$ matrix A if

- $\mathbf{v} \neq \mathbf{0}$, and
- Av is a multiple of v

So there is a scalar λ so that

$$A\mathbf{v} = \lambda \mathbf{v}$$

The number λ is called the *eigenvalue* (some people may use the name 'characteristic value') for the eigenvector v of A.

3.1.6 Theorem. The eigenvalues of an $n \times n$ matrix A are exactly the solutions of the characteristic equation for A, which is the equation

$$\det(A - \lambda I_n) = 0$$

Proof. If λ is an eigenvalue for A, that means there is an eigenvector v for the eigenvalue λ . So, $A\mathbf{v} = \lambda \mathbf{v}$, and then we can say that

$$A\mathbf{v} = \lambda I_n \mathbf{v},$$

and that can be rearranged to say

$$(A - \lambda I_n)\mathbf{v} = \mathbf{0}$$

Since v is an eigenvector, then v is not the obvious solution v = 0 of that equation. This is a way of recognising that $A - \lambda I_n$ is not invertible. (If $A - \lambda I_n$ had an inverse matrix, then the equation $(A - \lambda I_n)v = 0$ would have only the zero solution.)

But another way to say that $A - \lambda I_n$ is not invertible is to say

$$\det(A - \lambda I_n) = 0$$

(see Theorem 1.4.5).

This shows that if λ is an eigenvalue, then λ solves the characteristic equation.

To go the other way, that is to show that if λ solves the characteristic equation then λ must be an eigenvalue, is not a very different line of argument. So we'll skip that, as we've said enough to show the connection between the concept of an eigenvalue and the solutions of the characteristic equation.

3.1.7 Proposition (A slightly incomplete statement). If A is a symmetric matrix then we can write $A = PDP^t$ (with P orthogonal and D diagonal) if we take

- D to be the diagonal matrix with the eigenvalues of A along the diagonal
- *P* to be a matrix where the columns are orthonormal eigenvectors of *A* for the eigenvalues (in the same order as we take the eigenvalues along the diagonal of *D*).

We will not prove this. The idea is to take unit vector eigenvectors. So we find the eigenvalues from the characteristic equation, and then we get eigenvectors for each eigenvalue. Next we divide the eigenvectors by their length to make them unit vectors as well as eigenvectors.

The eigenvectors are usually automatically perpendicular to one another and so orthonormal. The more complicated case is where there is more than one eigenvector for the same eigenvalue. To explain that a bit more, any nonzero multiple of an eigenvector is always another eigenvector for the same eigenvalue. (If $A\mathbf{v} = \lambda \mathbf{v}$, then $A(2\mathbf{v}) = \lambda(2\mathbf{v})$). So $2\mathbf{v}$ is again an eigenvector for the eigenvalue λ , and we can change the factor 2 to any nonzero factor.) By 'more than one eigenvector' I mean one where there are eigenvectors that are not just multiples of each other, and still belong to the same eigenvalue λ .

Those cases are a bit more tricky to work out. But maybe we will manage not to pick any of those more complicated examples!

5

Eigenvalues, etc

3.1.8 Example. For

$$A = \begin{bmatrix} 1 & 2\\ 2 & 4 \end{bmatrix}$$

find an orthogonal matrix P and a diagonal matrix D so that $A = PDP^t$. Solution: We want the eigenvalues and so we should work out the characteristic equation

$$\det(A - \lambda I_2) = 0$$

We have

$$A - \lambda I_2 = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 - \lambda & 2 \\ 2 & 4 - \lambda \end{bmatrix}$$

So we have $det(A - \lambda I_2) = (1 - \lambda)(4 - \lambda) - 4 = (\lambda - 1)(\lambda - 4) - 4 = \lambda^2 - 5\lambda + 4 - 4$ and the characteristic equation works out as

$$\lambda^2 - 5\lambda = 0$$

This is a quadratic equation (as it will be for every 2×2 matrix) and so we can solve it. We can do it in this case by factoring the equation as

$$\lambda(\lambda - 5) = 0$$

So the two solutions (the two eigenvalues) are $\lambda = 0$ and $\lambda = 5$.

Aside: Be careful not to divide out λ without taking account of the fact that $\lambda = 0$ is a solution.

This is only a 2×2 example, rather than 3×3 as we used for illustration before. So what we have now is the two eigenvalues $\lambda_1 = 0$ and $\lambda_2 = 5$. We should take our diagonal matrix D to be

$$D = \begin{bmatrix} \lambda_1 & 0\\ 0 & \lambda_2 \end{bmatrix} = \begin{bmatrix} 0 & 0\\ 0 & 5 \end{bmatrix}$$

but to find P we need eigenvectors.

For $\lambda = \lambda_1 = 0$ (the first eigenvalue) we want a nonzero vector $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \in \mathbb{R}^2$ so that $A\mathbf{v} = \lambda_1 \mathbf{v}$. In this case that is $A\mathbf{v} = 0\mathbf{v}$, so $A\mathbf{v} = \mathbf{0}$. We can find that by row-reducing the augmented matrix

$$\left[\begin{array}{rrr}1&2&:0\\2&4&:0\end{array}\right]$$

The first step is to subtract 2 times row 1 from row 2, to get the new row 2:

$$\left[\begin{array}{rrr}1&2&:0\\0&0&:0\end{array}\right]$$

This is already completely row reduced (reduced row echelon form) and it means that we really only have one equation

$$v_1 + 2v_2 = 0$$

(as the second row gives 0 = 0 — so tells us nothing). We can say that $v_1 = -2v_2$ and v_2 is a free variable. If we take $v_2 = 1$ then we get $v_1 = -2$ and the nonzero solution

$$\mathbf{v} = \begin{bmatrix} -2\\1 \end{bmatrix}$$

What would have happened if we took a different value for the free variable v_2 ? If we took $v_2 = 20$ we would just get 20 times the above vector. So just a multiple of the same eigenvector, not really different.

What we do want to do is to normalise that eigenvector to get one of length 1. In vector notation we have

$$\mathbf{v} = (-2)\mathbf{i} + \mathbf{j}$$

and we want to take

$$\frac{\mathbf{v}}{\|\mathbf{v}\|} = \frac{-2\mathbf{i} + \mathbf{j}}{\sqrt{(-2)^2 + 1^2}} = \frac{1}{\sqrt{5}}(-2\mathbf{i} + \mathbf{j}) = -\frac{2}{\sqrt{5}}\mathbf{i} + \frac{1}{\sqrt{5}}\mathbf{j}$$

as our normalised eigenvector.

Now, for the other eigenvalue $\lambda = \lambda_2 = 5$ we want a nonzero vector \mathbf{v} (not the same \mathbf{v} as we have a moment ago) so that $A\mathbf{v} = 5\mathbf{v}$. We write that as $A\mathbf{v} = 5I_2\mathbf{v}$ or $(A - 5I_2)\mathbf{v} = \mathbf{0}$. We can see that

$$A - 5I_2 = \begin{bmatrix} 1 & 2\\ 2 & 4 \end{bmatrix} - 5 \begin{bmatrix} 1 & 0\\ 0 & 1 \end{bmatrix} = \begin{bmatrix} -4 & 2\\ 2 & -1 \end{bmatrix}$$

and this time we want to row reduce

$$\left[\begin{array}{rrr} -4 & 2 & :0\\ 2 & -1 & :0 \end{array}\right]$$

We should divide row 1 by -4 (to get 1 in the top left corner) and we have

$$\left[\begin{array}{rrr} 1 & -1/2 & : 0 \\ 2 & -1 & : 0 \end{array}\right]$$

and then replace row 2 by OldRow $2 - 2 \times \text{OldRow1}$ to get

$$\left[\begin{array}{rrr}1 & -1/2 & : 0\\0 & 0 & : 0\end{array}\right]$$

Again we have just one equation $v_1 - (1/2)v_2 = 0$. Again we should take v_2 free. This time $v_1 = (1/2)v_2$ and if we pick $v_2 = 1$ we get the eigenvector

$$\frac{1}{2}\mathbf{i} + \mathbf{j}$$

(Actually it might save us some bother to take $v_2 = 2$ instead.) We should normalise this to

$$\frac{(1/2)\mathbf{i} + \mathbf{j}}{\|(1/2)\mathbf{i} + \mathbf{j}\|} = \frac{(1/2)\mathbf{i} + \mathbf{j}}{\sqrt{1/4} + 1}$$
$$= \frac{(1/2)\mathbf{i} + \mathbf{j}}{\sqrt{5/4}}$$
$$= \frac{(1/2)\mathbf{i} + \mathbf{j}}{\sqrt{5/2}}$$
$$= \frac{2}{\sqrt{5}}((1/2)\mathbf{i} + \mathbf{j})$$
$$= \frac{1}{\sqrt{5}}\mathbf{i} + \frac{2}{\sqrt{5}}\mathbf{j}$$

We now have the two eigenvectors to make into the columns of P. The result is

$$P = \begin{bmatrix} -\frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{bmatrix}$$

Notice that the columns of P are perpendicular to one another. As we have normalised the columns to be unit vectors, we have that P is an orthogonal matrix. We do then get $A = PDP^t$ or

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} = \begin{bmatrix} -\frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & 5 \end{bmatrix} \begin{bmatrix} -\frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{bmatrix}^t = \begin{bmatrix} -\frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & 5 \end{bmatrix} \begin{bmatrix} -\frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{bmatrix}$$

If you are doubtful that the theory really holds out (or just want to check that we made no slips in the calculations) we can verify

$$\begin{bmatrix} -\frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & 5 \end{bmatrix} \begin{bmatrix} -\frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{bmatrix} = \begin{bmatrix} 0 & \sqrt{5} \\ 0 & 2\sqrt{5} \end{bmatrix} \begin{bmatrix} -\frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$$

The theory says that this should always work for a symmetric A. One thing that could possibly go wrong for the 2×2 case is that the characteristic equation (the quadratic equation to solve for the eigenvalues) might have complex roots. Well, that never will happen if A is symmetric.

Let's try a 3×3 example to see how it goes. You can probably see that the calculations are fairly long even for the 2×2 case, though we could shorten what is written above by leaving out some of the chatty bits.

3.1.9 Example. Find the eigenvalues of the matrix

$$A = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 3 \end{bmatrix}$$

Solution: We need to work out the characteristic equation $det(A - \lambda I_3) = 0$ and find its solutions. First

$$A - \lambda I_3 = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 3 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 - \lambda & 1 & 1 \\ 1 & 2 - \lambda & 1 \\ 1 & 1 & 3 - \lambda \end{bmatrix}$$

We can expand the determinant along the first row

$$det(A - \lambda I_3) = (2 - \lambda) det \begin{bmatrix} 2 - \lambda & 1 \\ 1 & 3 - \lambda \end{bmatrix} - 1 det \begin{bmatrix} 1 & 1 \\ 1 & 3 - \lambda \end{bmatrix} + 1 det \begin{bmatrix} 1 & 2 - \lambda \\ 1 & 1 \end{bmatrix}$$

= $(2 - \lambda)((2 - \lambda)(3 - \lambda) - 1) - (3 - \lambda - 1)$
 $+1 - (2 - \lambda)$
= $-(\lambda - 2)((\lambda - 2)(\lambda - 3) - 1) - (2 - \lambda) - 1 + \lambda$
= $-(\lambda - 2)(\lambda^2 - 5\lambda + 6 - 1) - 3 + 2\lambda$
= $-(\lambda - 2)(\lambda^2 - 5\lambda + 5) - 3 + 2\lambda$
= $-(\lambda^3 - 5\lambda^2 + 5\lambda - 2\lambda^2 + 10\lambda - 10) - 3 + 2\lambda$
= $-(\lambda^3 - 7\lambda^2 + 15\lambda - 10) - 3 + 2\lambda$
= $-(\lambda^3 - 7\lambda^2 + 13\lambda - 7)$

Now Theorem 3.1.3 says that there have to be 3 real solutions to this (3 eigenvalues for the symmetric matrix A). However, it does not help us immediately with finding the solutions. In fact every cubic equation has at least one real root, but the formula for the roots of a cubic is too messy to be useful.

In practice the best hope for finding the roots is to use the **remainder theorem**. That says that if $\lambda = \lambda_1$ is a solution of the (polynomial) equation

$$\lambda^3 - 7\lambda^2 + 13\lambda - 7 = 0$$

then $\lambda - \lambda_1$ must divide $\lambda^3 - 7\lambda^2 + 13\lambda - 7$. The only nice way to find a solution is to hope we can spot one, and the only way to do that is to try the divisors of 7 (the constant term). Those divisors are 1, -1, 7 and -7.

For 1 we get

$$(\lambda^3 - 7\lambda^2 + 13\lambda - 7)|_{\lambda=1} = 1 - 7 + 13 - 7 = 0$$

and so the remainder theorem says $\lambda - 1$ divides $\lambda^3 - 7\lambda^2 + 13\lambda - 7$. We can divide it in using long division of polynomials

$$\lambda - 1 \frac{\lambda^2 - 6\lambda + 7}{\lambda^3 - 7\lambda^2 + 13\lambda - 7}$$

$$\underline{\lambda^3 - \lambda^2} - 6\lambda^2 + 13\lambda$$

$$\underline{-6\lambda^2 + 6\lambda} - 7$$

$$\underline{7\lambda - 7} - 7$$

$$0$$

Eigenvalues, etc

So we have

$$\lambda^3 - 7\lambda^2 + 13\lambda - 7 = (\lambda - 1)(\lambda^2 - 6\lambda + 7)$$

and the roots of $\lambda^3 - 7\lambda^2 + 13\lambda - 7 = 0$ (our characteristic equation) are $\lambda = 1$ and the roots of the quadratic $\lambda^2 - 6\lambda + 7$. Those roots of the quadratic are

$$\lambda = \frac{6 \pm \sqrt{6^2 - 4(7)}}{2} = \frac{6 \pm \sqrt{8}}{2} = 3 \pm \sqrt{\frac{8}{4}} = 3 \pm \sqrt{2}$$

So we now have all the eigenvalues

$$\lambda = 1, \quad \lambda = 3 + \sqrt{2} \text{ and } \lambda = 3 - \sqrt{2}$$

as required.

So we are finished what was asked.

Notice that to find P so that $A = P^t DP$ we would need to find unit eigenvectors for each of these three eigenvalues — so it would take quite a while. With the help of Mathematica, I can tell you the answers. For $\lambda = 1$, the eigenvector is

$$-\frac{1}{\sqrt{2}}\mathbf{i}+\frac{1}{\sqrt{2}}\mathbf{j},$$

for $\lambda = 3 + \sqrt{2}$ it is

$$\frac{1}{2}\mathbf{i} + \frac{1}{2}\mathbf{j} + \frac{1}{\sqrt{2}}\mathbf{k}$$

and for $\lambda = 3 - \sqrt{2}$ it is

$$-\frac{1}{2}\mathbf{i} - \frac{1}{2}\mathbf{j} + \frac{1}{\sqrt{2}}\mathbf{k}$$

So if

$$P = \begin{bmatrix} -\frac{1}{\sqrt{2}} & \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{\sqrt{2}} & \frac{1}{2} & -\frac{1}{2} \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \text{ and } D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 + \sqrt{2} & 0 \\ 0 & 0 & 3 - \sqrt{2} \end{bmatrix}$$

then $A = PDP^t$.

3.1.10 Remark. Theorem 3.1.3 is quite useful because it is so simple to apply. Symmetric matrices are easy to recognise. The fact that they can be expressed as diagonal matrices in a new orthonormal basis is useful.

In the case n = 3, this theorem is often called the 'Principal Axis Theorem' because of an interpretation it has in mechanics. In mechanics there is a symmetric matrix called the 'inertia matrix' associated with a solid object and it has to do with rotating the object around axes through the centre of mass. The axes where there will be no wobble or vibration are the axes in the direction of the eigenvectors. The 'Principal Axis Theorem' says there are always 3 such axes (though there could be more if the object is very symmetric).

The next topic is to do something similar for matrices that are not symmetric. In this case things become more complicated because there is no longer a 'Principal Axis Theorem'.

3.2 Diagonalisable matrices

3.2.1 Definition. An $n \times n$ matrix A is called *diagonalisable* if there is an invertible matrix S and diagonal matrix D (also $n \times n$ matrices) so that

$$A = SDS^{-1}$$

The matrix A is called *orthogonally diagonalisable* if we can take S to be an orthogonal matrix (so that S corresponds to what we had as P^t before).

- **3.2.2 Remarks.** (i) We've seen that the orthogonally diagonalisable matrices are exactly the symmetric matrices A (those with $A = A^t$). That is what Theorem 3.1.3 tells us.
- (ii) In the case where A is symmetric we take S = P where the columns of P are normalised eigenvectors of A. That means the columns of S are an orthonormal basis made up of eigenvectors.
- (iii) In general if $A = SDS^{-1}$, the columns of S must be eigenvectors for A.

If, say,

$$S = \begin{bmatrix} u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \\ u_3 & v_3 & w_3 \end{bmatrix}$$

we don't necessarily have $\mathbf{u} = u_1\mathbf{i} + u_2\mathbf{j} + u_3\mathbf{k}$, $\mathbf{v} = v_1\mathbf{i} + v_2\mathbf{j} + v_3\mathbf{k}$, $\mathbf{w} = w_1\mathbf{i} + w_2\mathbf{j} + w_3\mathbf{k}$ perpendicular to each other (and there is no longer any great advantage in them being normalised to have length 1).

We could explain what is going on with $A = SDS^{-1}$ in terms of changing axes from the usual ones to axes which are parallel to the vectors \mathbf{u} , \mathbf{v} and \mathbf{w} but this brings in ideas that are rather harder to follow. The new axes are no longer perpendicular axes and all we can say is that they are in directions that use up all 3 dimensions — so no one is in the plane of the other 2.

(iv) We can go about trying to write any square matrix as $A = SDS^{-1}$ in much the same way as we did before for the case of symmetric matrices.

Step 1 is to find the eigenvalues of A by solving the characteristic equation $det(A - \lambda I_n) = 0$. Things can go wrong at step 1. Even for 2×2 matrices, the quadratic equation we get could have only 1 root (both roots the same) or the roots could be complex numbers.

Complex numbers should not be such a problem, and things certainly work more often if we deal with comples matrices — matrices where the entries are allowed to be complex numbers as well as real numbers. However, that would be another level of complication on top of what we have done and I don't want to tackle that. It is actually not that much more difficult, but still it seems better not to go into it for this course.

If step 1 works out *and* it happens that we get n different real eigenvalues, then we can succeed in finding D and S so that $A = SDS^{-1}$. Take D to be the diagonal matrix with the

eigenvalues along the diagonal and S to have the eigenvectors as its columns. (You need to take the eigenvectors in the same order at the eigenvalues.)

(v) One case where the strategy works out quite nicely is for triangular matrices (upper or lower triangular) with all the diagonal entries different.

Take for example

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{bmatrix}$$

The characteristic equation $det(A - \lambda I_3) = 0$ is very easy to solve in this case because

$$A - \lambda I_3 = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 - \lambda & 2 & 3 \\ 0 & 4 - \lambda & 5 \\ 0 & 0 & 6 - \lambda \end{bmatrix}$$

and so

$$\det(A - \lambda I_3) = (1 - \lambda)(4 - \lambda)(6 - \lambda)$$

Thus the eigenvalues of the (triangular) matrix A are

 $\lambda_1 = 1, \quad \lambda_2 = 4 \text{ and } \lambda_3 = 6$

There was nothing special about the numbers 1, 2 and 6. For triangular matrices the eigenvalues are always going to work out to be the entries along the diagonal.

To find S we need eigenvectors. So, for $\lambda = 1$ we need to row reduce $[A - \lambda I_3 : \mathbf{0}] = [A - I_3 : \mathbf{0}]$, which is

	$\left[\begin{array}{c} 0\\ 0\\ 0\\ 0\end{array}\right]$	2 3 0	${3 \atop {5} \atop {5}}$	$\left[\begin{array}{c} : \ 0 \\ : \ 0 \\ : \ 0 \end{array}\right]$
Divide first row by 2:				
	$\left[\begin{array}{c} 0\\ 0\\ 0\\ 0\end{array}\right]$	$\begin{array}{c} 1 \\ 3 \\ 0 \end{array}$	$3/2 \\ 5 \\ 5$	$\left. \begin{array}{c} : \ 0 \\ : \ 0 \\ : \ 0 \end{array} \right]$
Row $2 - 3 \times$ Row 1:	_			_
	0	1	3/2	:0
	0	0	1/2	:0
	0	0	5	:0
Row 2 times 2:				
	[0	1	3/2	:0]
	0	0	1	:0
	0	0	5	:0

Row $3 - 5 \times$ Row 2:

In equations we have

$$\begin{cases} 0 & 1 & 3/2 & : 0 \\ 0 & 0 & 1 & : 0 \\ 0 & 0 & 0 & : 0 \end{cases}$$
$$\begin{cases} v_2 & + & \frac{3}{2}v_3 & = & 0 \\ & v_3 & = & 0 \end{cases}$$

and that means $v_3 = 0$ and $v_2 = 0$ but v_1 free. So $(v_1, v_2, v_3) = (1, 0, 0)$ makes up the first column of S.

We won't go through the similar calculations in detail but the results should be (2, 3, 0) for $\lambda = 4$ and (16, 25, 10) for $\lambda = 6$. (Any multiples of these vectors would work just as well.) So if

	1	0	0			1	2	16
D =	0	4	0	,	S =	0	3	25
	0	0	6			0	0	10

(columns of S are the eigenvectors) we have $A = SDS^{-1}$.

(vi) There are very simple matrices that are not diagonalisable (and it does not always help to allow complex numbers). One such example is

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

We can work out its eigenvalues by looking at the characteristic equation, but according to the logic of the example we have just done we know how it will turn out. The eigenvalues are the diagonal entries (since A is upper triangular) and so we just get $\lambda = 0$. If you like you can say that 0 is an eigenvalue twice.

So if we could write $A = SDS^{-1}$ the diagonal matrix D has to have the eigenvalues along the diagonal. In this case that means D has to be the zero matrix. But then SDS^{-1} works out as zero, and that is not the same as A.

So this A is **not** diagonalisable.

(vii) Now for some comments to explain some reasons that diagonalisable matrices are handy. At least it is handy if we already know S and D so that $A = SDS^{-1}$ (with D diagonal as usual).

It is very easy to calculate powers of A. Look first at A^2 . We have

$$A^{2} = AA = SDS^{-1}SDS^{-1} = SD(S^{-1}S)DS^{-1} = SDI_{n}DS^{-1} = SDDS^{-1} = SD^{2}S^{-1}$$
$$A^{3} = A^{2}A = SD^{2}S^{-1}SDS^{-1} = SD^{2}(S^{-1}S)DS^{-1} = SD^{2}I_{n}DS^{-1} = SD^{2}DS^{-1} = SD^{3}S^{-1}$$

and it is not hard to see that for every power A^n we have

$$A^n = SD^n S^{-1}$$

For small powers it is a minor convenience that powers of diagonal matrices are so easy to calculate, but for big powers this is a major saving. If (say for the 3×3 case)

$$D = \begin{bmatrix} \lambda_1 & 0 & 0\\ 0 & \lambda_2 & 0\\ 0 & 0 & \lambda_3 \end{bmatrix}$$

then

$$D^{n} = \begin{bmatrix} \lambda_{1}^{n} & 0 & 0\\ 0 & \lambda_{2}^{n} & 0\\ 0 & 0 & \lambda_{3}^{n} \end{bmatrix}$$

and your calculator can find powers of numbers.

(viii) Here we discuss the **exponential of a matrix**, a square matrix.

First we recall briefly that the ordinary exponential e^x of a numerical variable x can be expressed by an infinite series

$$e^x = 1 + x + \frac{1}{2!}x^2 + \frac{1}{3!}x^3 + \cdots$$

and sometimes it is convenient to use the notation exp(x) to mean exactly the same as e^x .

This kind of infinite sum needs to be defined using limits. So that series expression for e^x says

$$\exp(x) = e^x = \lim_{N \to \infty} 1 + x + \frac{1}{2!}x^2 + \frac{1}{3!}x^3 + \dots + \frac{1}{N!}x^N$$

A limit like this does not have to exist for every x. It is easy to see that the limit does exist for x = 0 and $e^0 = 1$, but for the exponential series it is known that the series converges (quite quickly in fact) for every x.

By analogy with this series for e^x , if A is an $n \times n$ matrix then we define

$$e^{A} = \lim_{N \to \infty} I_{n} + A + \frac{1}{2!}A^{2} + \frac{1}{3!}A^{3} + \dots + \frac{1}{N!}A^{N}$$

and we also write this as

$$\exp(A) = e^A = I_n + A + \frac{1}{2!}A^2 + \frac{1}{3!}A^3 + \cdots$$

(an infinite series again, but this time the terms to be added up are matrices).

It would be desirable to prove that this limit always exists, but we will not do it as it becomes quite easy only with concepts that we won't develop. What we can do is give a fairly convincing way to see that e^A makes sense if A is diagonalisable, and this also contains a good way to calculate e^A .

Suppose $A = SDS^{-1}$ is diagonalisable. We've already seen that $A^k = SD^kS^{-1}$ for all exponents k = 1, 2, ... So we can write

$$I_n + A + \frac{1}{2!}A^2 + \frac{1}{3!}A^3 + \dots + \frac{1}{N!}A^N$$

= $SS^{-1} + SDS^{-1} + \frac{1}{2!}SD^2S^{-1} + \frac{1}{3!}SD^3S^{-1} + \dots + \frac{1}{N!}SD^NS^{-1}$
= $S(I_n + D + \frac{1}{2!}D^2 + \frac{1}{3!}D^3 + \dots + \frac{1}{N!}D^N)S^{-1}$

Now

$$\lim_{N \to \infty} I_n + D + \frac{1}{2!}D^2 + \frac{1}{3!}D^3 + \dots + \frac{1}{N!}D^N$$

is easy to calculate. Say we take n = 3 as an illustration and $D = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix}$. Then

$$\begin{split} I_n + D &+ \frac{1}{2!}D^2 + \frac{1}{3!}D^3 + \dots + \frac{1}{N!}D^N \\ &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} + \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix} + \frac{1}{2!}\begin{bmatrix} \lambda_1^2 & 0 & 0 \\ 0 & \lambda_2^2 & 0 \\ 0 & 0 & \lambda_3^2 \end{bmatrix} + \dots + \frac{1}{N!}\begin{bmatrix} \lambda_1^N & 0 & 0 \\ 0 & \lambda_2^N & 0 \\ 0 & 0 & \lambda_3^N \end{bmatrix} \\ &= \begin{bmatrix} 1 + \lambda_1 + \frac{1}{2!}\lambda_1^2 + \dots + \frac{1}{N!}\lambda_1^N & 0 & 0 \\ 0 & 1 + \lambda_2 + \frac{1}{2!}\lambda_2^2 + \dots + \frac{1}{N!}\lambda_2^N & 0 \\ 0 & 0 & 1 + \lambda_3 + \dots + \frac{1}{N!}\lambda_3^N \end{bmatrix}$$

So now we can see what the limit of this is as $N \to \infty$ and we get

$$e^{D} = \begin{bmatrix} e^{\lambda_{1}} & 0 & 0\\ 0 & e^{\lambda_{2}} & 0\\ 0 & 0 & e^{\lambda_{3}} \end{bmatrix}$$

From this it is a short step (which we will not justify totally) that we can multiply the limit by S on the left and S^{-1} on the right to get

$$e^{A} = Se^{D}S^{-1} = S\begin{bmatrix} e^{\lambda_{1}} & 0 & 0\\ 0 & e^{\lambda_{2}} & 0\\ 0 & 0 & e^{\lambda_{3}} \end{bmatrix} S^{-1}$$

(where $A = SDS^{-1}$ and D is as above). This kind of calculation is valid for diagonalisable matrices of any size (not just 3×3 as we used for illustration).

In the next topic we will see that matrix exponentials can be used for solving differential equations.

3.3 Linear Differential Equations

We will take a rather restricted look at differential equations, aiming for an aspect where matrices can be used effectively.

A first order linear differential equation with constant coefficients is an equation of the form

$$\frac{dy}{dx} - ay = b$$

where a and b are constants. The unknown in a differential equation is a function y = y(x) and the reason it is called a 'differential equation' is that the equation involves not just the values of y (and possibly x) but also derivatives of the unknown function. We call this one first order because the highest derivative that occurs is the first derivative dy/dx of the unknown function. Later we will talk about second order equations which involve the second derivative d^2y/dx^2 . We won't actually deal with higher order equations, but third order ones would have the third derivative in them, and so on.

Sometimes these differential equations are called 'ordinary differential equations' (abbreviated ODE sometimes) to distinguish them from 'partial differential equations' (abbreviated PDE). Both are important for many different applications, but we will not deal with PDEs at all and what w say about ODEs is quite limited. Just to explain what a PDE is, it is an equation where the unknown is a function of more than one variable and the equation involves 'partial' derivatives of the unknown function. In case you have not heard anything about partial derivatives yet, here is a very very brief explanation.

We did discuss functions of more than one variable back at the start of this chapter, though we moved fairly quickly to vector-valued (linear) functions of a vector variable. A vector variable is the same as several scalar variables at once. As an example, we can have functions

$$y = f(x_1, x_2) = x_1^4 x_2 + 3x_1^2 x_2^2 + x_1^5 - x_2^6$$

of two variables. The partial derivative of this with respect to the x_1 -variable is what you get by differentiating the right hand side with respect to x_1 while treating all the other variables as constant (in this case only the variable x_2 as there are only 2 variables). Another way to explain it is that we fix some value (constant value) for x_2 while we differentiate with respect to x_1 . The notation for the partial derivative is $\frac{\partial y}{\partial x_1}$ and in the example it turns out to be

$$\frac{\partial y}{\partial x_1} = 4x_1^3 x_2 + 6x_1 x_2^2 + 5x_1^4 - 0$$

On the other hand the partial derivative with respect to x_2 is

$$\frac{\partial y}{\partial x_2} = x_1^4 + 6x_1^2x_2 + 0 - 6x_2^5$$

This brief explanation of how to calculate them is not enough to allow you to understand the ideas around partial derivatives. But anyhow, a PDE is an equation that involves partial derivatives of an unknown function (as well as the function itself and the variables x_1 , x_2 and more if there are more than 2 variables involved).

We've said that the example equation at the start of this discussion should have constant coefficients (a restriction that is not really necessary for this case) but we also said it was called a *linear* equation. Let's explain now what we mean by the word linear and why it is used at all.

The reason

$$\frac{dy}{dx} - ay = b$$

is called linear is that the left hand side

$$\frac{dy}{dx} - ay$$

depends on the unknown function y in a linear way. What we mean my this is that if we define an operation (or transformation) on functions y = y(x) by

$$Ty = \frac{dy}{dx} - ay$$

then the transformation T has the same key properties we had in Theorem 2.5.1. That is

- T(y+z) = T(y) + T(z) (if y = y(x) and z = z(x) are functions); and
- T(ky) = kT(y) (for y = y(x) a function and k a constant).

Recall that we first introduced linear transformation $f : \mathbb{R}^n \to \mathbb{R}^m$ as arising from an $m \times n$ matrix A by

$$f(\mathbf{x}) = A\mathbf{x}.$$

Later we showed in Theorem 2.5.1 that the two 'linearity' properties were a way to distinguish linear transformations from more general functions $f \colon \mathbb{R}^n \to \mathbb{R}^m$. What we had in Theorem 2.5.1 was a more abstract way of distinguishing linear transformations, one that is not tied to the use of the standard coordinates, or the standard basis for \mathbb{R}^n and \mathbb{R}^m .

If you look in the book, you can see that there is a more abstract approach where one can replace \mathbb{R}^n and \mathbb{R}^m by 'vector spaces' where you can add 'vectors' and multiply them by scalars. In this more abstract approach, our

$$Ty = \frac{dy}{dx} - ay$$

would fit right in as just one example, and we would have a more complete explanation of why we call it 'linear'.

3.3.1 Method of integrating factors. We now explain how to find all the solutions of first order linear equations

$$\frac{dy}{dx} - ay = b$$

by a method called *integrating factors*. In fact the method can be applied even if a and b are not constants but are allowed to be functions of x, but it is particularly easy to use for the case

where a and b are constants. For the situation where we are going to bring matrices into our considerations, the restriction to constant coefficients will be more crucial We now explain how to find all the solutions of first order linear equations

$$\frac{dy}{dx} - ay = b$$

by a method called *integrating factors*. In fact the method can be applied even if a and b are not constants but are allowed to be functions of x, but it is particularly easy to use for the case where a and b are constants. For the situation where we are going to bring matrices into our considerations, the restriction to constant coefficients will be more crucial.

The method is to multiply the equation by

$$e^{\int -a\,dx} = e^{-ax}.$$

When we do that we get

$$e^{-ax}\frac{dy}{dx} - ae^{-ax}y = be^{-ax}$$

and the whole point of this trick is that the left hand side is now the derivative of a product. From the product rule we have

$$\frac{d}{dx}\left(e^{-ax}y\right) = e^{-ax}\frac{dy}{dx} - ae^{-ax}y$$

and so the equation we now have can be rewritten

$$\frac{d}{dx}\left(e^{-ax}y\right) = be^{-ax}$$

So we get

$$e^{-ax}y = \int be^{-ax} \, dx = -\frac{b}{a}e^{-ax} + C$$

with C some constant. Multiply now by e^{ax} on both sides to get

$$y = -\frac{b}{a} + Ce^{ax}$$

3.3.2 Example. Find all solutions of

$$\frac{dy}{dx} - 5y = 3$$

Solution: Multiply by the integrating factor

$$e^{\int -5\,dx} = e^{-5x}$$

to get

$$e^{-5x}\frac{dy}{dx} - 5e^{-5x}y = 3e^{-5x}$$
$$\frac{d}{dx}(e^{-5x}y) = 3e^{-5x}$$
$$e^{-5x}y = \int 3e^{-5x} dx$$
$$= -\frac{3}{5}e^{-5x} + C$$
$$y = -\frac{3}{5} + Ce^{5x}$$

Note that the solution involves a constant C which can be anything.

3.3.3 Remark. We say that we have found the general solution y for the differential equation $\frac{dy}{dx} - 5y = 3$. In an application, where we would want to know y precisely, we need some more information to pin down y.

A fairly typical case is a case where we know one value of y in addition to the fact that y satisfies the differential equation. This kind of problem is called an 'initial value problem'. An example would be to find y given that $\frac{dy}{dx} - 5y = 3$ and y(0) = 0. We found above that y has to have the general form $y = -\frac{3}{5} + Ce^{5x}$ for some constant C, but then we can plug in x = 0 to see that

$$0 = y(0) = -\frac{3}{5} + Ce^0 = -\frac{3}{5} + C$$

and that tell us that C = 3/5. The solution to the initial value problem is then

$$y = -\frac{3}{5} + \frac{3}{5}e^{5x}.$$

3.3.4 Remark. A consequence of linearity of equations like

$$\frac{dy}{dx} - ay = b \tag{3.3.1}$$

is that there is a relationship between solutions of this equation and solutions of the associated 'homogeneous equation' where the right hand side is replaced by 0. That is with

$$\frac{dy}{dx} - ay = 0 \tag{3.3.2}$$

The relationship can be expressed in a few ways. One way is this. If we take one solution $y = y_0$ for (3.3.1) (the inhomogeneous equation), we know

$$\frac{dy_0}{dx} - ay_0 = b \tag{3.3.3}$$

Now subtract (3.3.3) from (3.3.1) and rearrange the result using linearity to get

$$\frac{d}{dx}(y - y_0) - a(y - y_0) = 0$$

What this says is that two solutions of the inhomogeneous equation (3.3.1) have a difference $y - y_0$ that is a solution of the associated homogeneous equation (3.3.2). Another way to say that is that if we somehow know one 'particular solution' y_0 for (3.3.1), then the general solution y for (3.3.1) is

 $y = y_0 + (\text{general solution of homogeneous equation (3.3.2)})$

This gives a strategy for solving linear equations. It is not really so useful for these first order linear equations because the method of integrating factors just works out all the solutions, but it helps a lot with second order linear problems (which we will come to soon).

The strategy is this:

- somehow find one 'particular' solution for the inhomogeneous equation (by guesswork or systematic guesswork this can often be done);
- look for the general solution of the associated homogeneous equation (with 0 on the right):

(This can be easier to deal with than having to cope with the right hand side at the same time.)

• the general solution of the inhomogeneous equation is then

(particular solution) + (general solution of homogeneous equation)

3.3.5 Second Order Linear. We now move on to second order linear differential equations. We will only deal with the case of constant coefficients (which makes things much easier) and we will also discuss only the homogeneous case. The strategy we have just outlined about 'particular' solutions + general solutions of the associated homogeneous equation is a very common way to approach these second order problems. So we will be dealing with most of the significant issues by looking at the homogeneous case.

That means we will discuss equations of the type

$$\frac{d^2y}{dx^2} + b\frac{dy}{dx} + cy = 0$$

where b and c are constants. We could allow for a constant coefficient in front of the second derivative term. But, we can divide across by that coefficient to get an equation like the one above. (Well we can do that if we are not dividing by 0. If the coefficient of d^2y/dx^2 was 0, then we would have a first order problem, not a second order one.)

Now there is a trick to reduce to a first order problem, but at the expense of getting two equations (a system of equations).

The trick is to introduce a new name y_1 for y and a temporary

$$y_2 = \frac{dy}{dx} = \frac{dy_1}{dx}$$

Then $\frac{d^2y}{dx^2} = \frac{d}{dx}\left(\frac{dy}{dx}\right) = \frac{dy_2}{dx}$. We can then rewrite the second order equation as

$$\frac{dy_2}{dx} + by_2 + cy_1 = 0$$

but we need to keep in mind the connection between y_2 and y_1 also. We get a system

$$\begin{cases} \frac{dy_1}{dx} = y_2\\ \frac{dy_2}{dx} + by_2 + cy_1 = 0 \end{cases}$$

or

$$\begin{cases} \frac{dy_1}{dx} = y_2\\ \frac{dy_2}{dx} = -cy_1 - by_2 \end{cases}$$

Using matrices we can write this system as a single equation between two column matrices

$$\begin{bmatrix} \frac{dy_1}{dx} \\ \frac{dy_2}{dx} \end{bmatrix} = \begin{bmatrix} y_2 \\ -cy_1 - by_2 \end{bmatrix}$$

and using matrix multiplication we can write that as

$$\begin{bmatrix} \frac{dy_1}{dx} \\ \frac{dy_2}{dx} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -c & -b \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

Next we treat the unknown in this equation as a vector valued function

$$\mathbf{y} = \mathbf{y}(x) = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

and we agree to say that differentiating such a function means differentiating each component. That is we define

$$\frac{d}{dx} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} \frac{dy_1}{dx} \\ \frac{dy_2}{dx} \end{bmatrix}$$

Now we can write our matrix equation (which came about from the system of two equations) as a differential equation for the vector-valued unknown $\mathbf{y}(x) = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$. We get

$$\frac{d\mathbf{y}}{dx} = \begin{bmatrix} 0 & 1\\ -c & -b \end{bmatrix} \mathbf{y}$$

20

If we use A to stand for the matrix

$$A = \begin{bmatrix} 0 & 1\\ -c & -b \end{bmatrix}$$

then our equation is

$$\frac{d\mathbf{y}}{dx} = A\mathbf{y}$$

To make it look even more like what we had before we write this

$$\frac{d\mathbf{y}}{dx} - A\mathbf{y} = \mathbf{0} \tag{3.3.4}$$

This looks so temptingly similar to the first order linear (homogeneous) equation

$$\frac{dy}{dx} - ay = 0$$

that we might like to try and use an integrating factor method to solve the new version. When a is a scalar we multiplied by the integrating factor $e^{\int -a \, dx} = e^{-ax}$ and so maybe we should multiply across by e^{-Ax} .

Since we conveniently discussed exponentials of matrices already, we can makes sense of e^{-Ax} . The order of matrix multiplication matters a lot and we multiply $\frac{dy}{dx} - Ay = 0$ by e^{-Ax} on the left. As e^{-Ax} is a 2 × 2 matrix and the equation is an equality of 2 × 1 column matrices, we have to multiply on the left. We get

$$e^{-Ax}\frac{d\mathbf{y}}{dx} - e^{-Ax}A\mathbf{y} = e^{-Ax}\mathbf{0}$$

or

$$e^{-Ax}\frac{d\mathbf{y}}{dx} - e^{-Ax}A\mathbf{y} = \mathbf{0}$$

It turns out to be the case that

$$\frac{d}{dx}e^{-Ax} = -e^{-Ax}A$$

if we interpret differentiation of a matrix-valued function of x to mean differentiation of each entry separately.

Why is that true? Well we are not really in a position to show it in general, but it is easy enough to see that it works out when A is replaced by a diagonal matrix D. If $D = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$ then

$$e^{-Dx} = \exp\left(\begin{bmatrix}-\lambda_1 x & 0\\ 0 & -\lambda_2 x\end{bmatrix}\right) = \begin{bmatrix}e^{-\lambda_1 x} & 0\\ 0 & e^{-\lambda_2 x}\end{bmatrix}$$

So

$$\frac{d}{dx}e^{-Dx} = \begin{bmatrix} \frac{d}{dx}e^{-\lambda_1 x} & 0\\ 0 & \frac{d}{dx}e^{-\lambda_2 x} \end{bmatrix}$$
$$= \begin{bmatrix} -\lambda_1 e^{-\lambda_1 x} & 0\\ 0 & -\lambda_2 e^{-\lambda_2 x} \end{bmatrix}$$
$$= -\begin{bmatrix} e^{-\lambda_1 x} & 0\\ 0 & e^{-\lambda_2 x} \end{bmatrix} \begin{bmatrix} \lambda_1 & 0\\ 0 & \lambda_2 \end{bmatrix} = -e^{-Dx}D$$

It is possible to come up with a modification of this argument to show $\frac{d}{dx}e^{-Ax} = -e^{-Ax}A$ when $A = SDS^{-1}$ is diagonalisable, but actually it is true for every square matrix A. Using $\frac{d}{dx}e^{-Ax} = -e^{-Ax}A$ we can rewrite our equation (3.3.4) as

$$e^{-Ax}\frac{d\mathbf{y}}{dx} + \frac{d}{dx}\left(e^{-Ax}\right)\mathbf{y} = \mathbf{0}$$
(3.3.5)

What we need now is a product rule for differentiating matrix products. We will not check it works, but it is true that if U = U(x) and V = V(x) are matrix functions so that the matrix product UV makes sense, then

$$\frac{d}{dx}(UV) = \left(\frac{dU}{dx}\right)V + U\frac{dV}{dx}$$

It is important to keep the order here so that U always stays to the left of V.

Using this product rule, we can rewrite (3.3.5) as

$$\frac{d}{dx}\left(e^{-Ax}\mathbf{y}\right) = \mathbf{0}$$

and it should be clear that the only vector functions that have derivative $\mathbf{0} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ are constants.

So we get

$$e^{-Ax}\mathbf{y} = \text{ constant vector } = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$$

To find y, we multiply by the inverse matrix of e^{-Ax} , which turns out to be e^{Ax} . We get

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = e^{Ax} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$$
(3.3.6)

as the solution of our system of differential equations (3.3.4).

Of course we should work this out further so it does not involve the matrix exponential. Let us assume (to make our life easier) that we are always in the case where A is diagonalisable. So we can write $A = SDS^{-1}$ with $D = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$ diagonal. Then $Ax = SDS^{-1}x = S(Dx)S^{-1}$

Eigenvalues, etc

(because x is a scalar) and $Dx = \begin{bmatrix} \lambda_1 x & 0 \\ 0 & \lambda_2 x \end{bmatrix}$ diagonal. We have then

$$e^{Ax} = Se^{Dx}S^{-1} = S\begin{bmatrix} e^{\lambda_1 x} & 0\\ 0 & e^{\lambda_2 x} \end{bmatrix} S^{-1}$$

When we go to use this in (3.3.6) the first thing will be to multiply out

$$S^{-1} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$$

It will make life easier for us if we just write the result as

$$\begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} = S^{-1} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$$

for two new constants α_1 and α_2 . So now out solution (3.3.6) comes to

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = e^{Ax} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = S \begin{bmatrix} e^{\lambda_1 x} & 0 \\ 0 & e^{\lambda_2 x} \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} = S \begin{bmatrix} \alpha_1 e^{\lambda_1 x} & 0 \\ 0 & \alpha_2 e^{\lambda_2 x} \end{bmatrix}$$

Recall now, finally, that the columns of S are eigenvectors for A belonging to the eigenvalues λ_1 and λ_2 . So S has the form

$$S = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 \end{bmatrix}$$

where v_1 and v_2 are the eigenvectors written as columns. That means that the solution works out as

$$\mathbf{y} = \alpha_1 e^{\lambda_1 x} \mathbf{v}_1 + \alpha_2 e^{\lambda_2 x} \mathbf{v}_2$$

We now summarise what all these calculations show:

3.3.6 Theorem. Assume that $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ is a diagonalisable matrix, with eigenvalues λ_1 and λ_2 and corresponding eigenvectors \mathbf{v}_1 and \mathbf{v}_2 . Then the solutions to the system of linear differential equations

$$\begin{cases} \frac{dy_1}{dx} = a_{11}y_1 + a_{12}y_2\\ \frac{dy_2}{dx} = a_{21}y_1 + a_{22}y_2 \end{cases}$$

are given by

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \alpha_1 e^{\lambda_1 x} \mathbf{v}_1 + \alpha_2 e^{\lambda_2 x} \mathbf{v}_2$$

where α_1 , α_2 are arbitrary constants.

3.3.7 Example. We did start with a second order equation

$$\frac{d^2y}{dx^2} + b\frac{dy}{dx} + cy = 0$$

and we rewrote it as a system

$$\frac{d\mathbf{y}}{dx} = \begin{bmatrix} 0 & 1\\ -c & -b \end{bmatrix} \mathbf{y}$$

where

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} y \\ dy/dx_2 \end{bmatrix}$$

According to the above, we need the eigenvalues and eigenvectors of the matrix

$$A = \begin{bmatrix} 0 & 1\\ -c & -b \end{bmatrix}$$

So we should look at the characteristic equation $det(A - \lambda I_2) = 0$. We have

$$A - \lambda I_2 = \begin{bmatrix} 0 & 1 \\ -c & -b \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} -\lambda & 1 \\ -c & -b - \lambda \end{bmatrix}$$

and so

$$\det(A - \lambda I_2) = -\lambda(-b - \lambda) + c = \lambda^2 + b\lambda + c$$

Notice then the close similarity between the differential equation and the characteristic equation. Replace the second derivative term by λ^2 , the derivative term by λ and the y term by 1.

To be specific we take the example

$$\frac{d^2y}{dx^2} - 5\frac{dy}{dx} + 4y = 0$$

so the characteristic equations is

$$\lambda^2 - 5\lambda + 4 = 0.$$

This factors as $(\lambda - 1)(\lambda - 4) = 0$ and so the eigenvalues are $\lambda = 1$ and $\lambda = 4$.

According to the recipe we also need the eigenvectors. We might say $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$ for $\lambda = 1$ and $\mathbf{w} = \begin{bmatrix} w_1 \\ v_2 \end{bmatrix}$ for $\lambda = 4$. Then we get

and $\mathbf{w} = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}$ for $\lambda = 4$. Then we get

$$\begin{bmatrix} y \\ \frac{dy}{dx} \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \alpha_1 e^x \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} + \alpha_2 e^{4x} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}$$

We don't really need to work out the eigenvectors in this case because we can see that

$$y = \alpha_1 v_1 e^x + \alpha_2 w_1 e^{4x}$$

Eigenvalues, etc

and we can just regard $\alpha_1 v_1$ and $\alpha_2 w_1$ as some constants. We get

$$y = C_1 e^x + C_2 e^{4x}$$

for constants C_1 and C_2 , where $\lambda_1 = 1$ and $\lambda_2 = 4$ are the roots of the equation $\lambda^2 - 5\lambda + 4 = 0$ that comes from the equation.

It is worth pointing out though that we have ignored the possibility that the quadratic might have complex roots (or just one root).

3.3.8 Example. Here is a somewhat applied example.

Two competing species live on the same small island and each one affects the growth rate of the other (by competing for the same food, say). If thir populations at time t are $x_1(t)$ and $x_2(t)$, a model for their growth rates says

$$\begin{cases} x'_1(t) = -3x_1(t) + 6x_2(t) \\ x'_2(t) = x_1(t) - 2x_2(t) \end{cases}$$

At time t = 0, $x_1(0) = 500$ and $x_2(0) = 200$.

Find $x_1(t)$ and $x_2(t)$.

Solution: We can write the model in matrix form as

$$\begin{bmatrix} x_1'(t) \\ x_2'(t) \end{bmatrix} = \frac{d}{dt} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} = \begin{bmatrix} -3 & 6 \\ 1 & -2 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}$$

We need the eigenvalues and eigenvectors for $A = \begin{bmatrix} -3 & 6 \\ 1 & -2 \end{bmatrix}$ in order to write down the general solution of this system of first order linear differential equations (and later we need to use the information about $x_1(0)$ and $x_2(0)$ to find the constants).

We have

$$A - \lambda I_2 = \begin{bmatrix} -3 & 6\\ 1 & -2 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0\\ 0 & 1 \end{bmatrix} = \begin{bmatrix} -3 - \lambda & 6\\ 1 & -2 - \lambda \end{bmatrix}$$

and so

$$\det(A - \lambda I_2) = (-3 - \lambda)(-2 - \lambda) - 6 = (\lambda + 3)(\lambda + 2) - 6 = \lambda^2 + 5\lambda + 6 - 6 = \lambda(\lambda + 5)$$

The eigenvalues are $\lambda_1 = 0$ and $\lambda_2 = -5$.

To find the eigenvalue for $\lambda_1 = 0$ we should row reduce $[A : \mathbf{0}]$. The answer should be the vector $\begin{bmatrix} 2\\1 \end{bmatrix}$.

For
$$\lambda_2 = -5$$
, we have $A - \lambda_2 I_2 = A + 5I_2 = \begin{bmatrix} 2 & 6 \\ 1 & 3 \end{bmatrix}$ and we should row reduce

$$\begin{bmatrix} 2 & 6 & : 0 \\ 1 & 3 & : 0 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 3 & : 0 \\ 1 & 3 & : 0 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 3 & : 0 \\ 0 & 0 & : 0 \end{bmatrix}$$

So that eigenvector is $\begin{bmatrix} -3\\1 \end{bmatrix}$. The general solution is then

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} = \alpha_1 e^0 \begin{bmatrix} 2 \\ 1 \end{bmatrix} + \alpha_2 e^{-5t} \begin{bmatrix} -3 \\ 1 \end{bmatrix} = \begin{bmatrix} 2\alpha_1 - 3\alpha_2 e^{-5t} \\ \alpha_1 + \alpha_2 e^{-5t} \end{bmatrix}$$

If we put t = 0 we are supposed to get

$$\begin{bmatrix} 500\\200 \end{bmatrix} = \begin{bmatrix} x_1(0)\\x_2(0) \end{bmatrix} = \begin{bmatrix} 2\alpha_1 - 3\alpha_2\\\alpha_1 + \alpha_2 \end{bmatrix}$$

We then have 2 simultaneous equations to solve for α_1 and α_2 . We could set that up as a matrix to be row reduced, but anyhow the solution is $\alpha_1 = 220$, $\alpha_2 = -20$.

So the answer to the example is

$$\begin{aligned} x_1(t) &= 440 + 60e^{-5t} \\ x_2(t) &= 220 - 20e^{-5t} \end{aligned}$$

(Aside: I'm not sure how realistic the model was but maybe it is. As time passes, that is as t get large and positive, $\lim_{t\to\infty} e^{-5t} = 0$ and the population will almost become (440, 220). For negative t there is a time when $x_2(t) < 0$ and that can't be right.)

3.4 Least square fit

We now discuss a topic that is rather important for applications.

Suppose we do an experiment and find some data points that we know (from the theory of the experiment) are supposed to lie on a line. Say we found n data points (x_1, y_1) , (x_2, y_2) , $\dots (x_n, y_n)$ and they are supposed to lie on a line y = mx + c. How do we find the right line?

If we had two points we would just find the line through the two points, but we will often have more than two data observations and (unless we fiddled the data) it is unlikely they will be on any one line. What then is the 'best' estimate of the line?

Well, that does depend on what you mean by 'best' but the least squares approach is often considered to be an appropriate interpretation of best. What it means is to choose the line that makes the sum of the square of the vertical distance from the data points to the line as small as possible. Here is a picture that is supposed to show the idea. The blobs represent 4 data points and the line should be so that the sum of the squares of the vertical segments joining the blobs to the line is smaller than for any other line.



This is not the only interpretation of 'best' that could be considered. It is reasonably appropriate if the data is such that the x-values x_1, x_2, \ldots, x_n are correct while the experimental errors are in the measurements of the y-values y_1, y_2, \ldots, y_n . The idea is that the 'correct' y-values are $y_1^*, y_2^*, \ldots, y_n^*$ and that these are so that the 'corrected' data points

$$(x_1, y_1^*), (x_2, y_2^*), \dots, (x_n, y_n^*)$$

lie on the 'correct' line. Moreover the most likely line is the one where we should make the smallest overall correction to the data — that is the one that make

$$(y_1 - y_1^*)^2 + (y_2 - y_2^*)^2 + \dots + (y_n - y_n^*)^2$$

as small as possible.

We take the line as y = mx + c (so c is the constant term and m is the coefficient of x or the slope). If all the data points were on a line, it would be the line that solved all the equations

$$mx_1 + c = y_1$$
$$mx_2 + c = y_2$$
$$\vdots$$
$$mx_n + c = y_n$$

In matrix form we can write this as

$$\begin{bmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots \\ x_n & 1 \end{bmatrix} \begin{bmatrix} m \\ c \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

or

$$X\mathbf{L} = \mathbf{y}$$

where we use the notation

$$X = \begin{bmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots \\ x_n & 1 \end{bmatrix}, \quad \mathbf{L} = \begin{bmatrix} m \\ c \end{bmatrix} \text{ and } \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

As we said already, once n > 2 we have too many equations for just 2 unknowns m and c. There is very little chance that there is any solution for L. The approach is to modify y to get y^* so that we do get a solution to

$$X\mathbf{L} = \mathbf{y}$$

and

$$\|\mathbf{y} - \mathbf{y}^*\|^2 = (y_1 - y_1^*)^2 + (y_2 - y_2^*)^2 + \dots + (y_n - y_n^*)^2$$

is as small as possible.

We'll take n = 3 for the purpose of explaining how that is done. We need some understanding of where we are allowed to take y^* so as to make $XL = y^*$ solvable. We can write

$$X\mathbf{L} = \begin{bmatrix} x_1 & 1\\ x_2 & 1\\ x_n & 1 \end{bmatrix} \begin{bmatrix} m\\ c \end{bmatrix} = m \begin{bmatrix} x_1\\ x_2\\ x_n \end{bmatrix} + c \begin{bmatrix} 1\\ 1\\ 1 \end{bmatrix}$$

What we can say then is that as m and c change we always find that XL lies in the plane in space \mathbb{R}^3 through the origin and containing the two vectors

$$\begin{bmatrix} x_1 \\ x_2 \\ x_n \end{bmatrix} = x_1 \mathbf{i} + x_2 \mathbf{j} + x_3 \mathbf{k} \text{ and } \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \mathbf{i} + \mathbf{j} + \mathbf{k}$$

So the best choice of y^* will be the point in that plane you get by dropping a perpendicular from y onto the plane. (That will give the point y^* on the plane closest to y.)

We can go about calculating that point, but there is a trick to avoid that. The difference $\mathbf{y} - \mathbf{y}^*$ should be a vector perpendicular to the plane. So $\mathbf{y} - \mathbf{y}^*$ is perpendicular to both $x_1\mathbf{i} + x_2\mathbf{j} + x_3\mathbf{k}$ and to $\mathbf{i} + \mathbf{j} + \mathbf{k}$. We can write that in matrix notation as

$$\begin{bmatrix} x_1 & x_2 & x_3 \\ 1 & 1 & 1 \end{bmatrix} (\mathbf{y} - \mathbf{y}^*) = \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \mathbf{0}$$

and that is just $X^t(\mathbf{y} - \mathbf{y}^*) = \mathbf{0}$. So we find that

$$X^t \mathbf{y} = X^t \mathbf{y}^*$$

So instead of finding \mathbf{y}^* and then solving

$$X\mathbf{L} = \mathbf{y}^*$$

we multiply this equation by X^t on the left. That gives

$$X^t X \mathbf{L} = X^t \mathbf{y}^*$$

which is the same as

 $X^t X \mathbf{L} = X^t \mathbf{y}$

Now $X^t X$ is just a 2 × 2 matrix and $X^t y$ is a 2 × 1 matrix. We are down to 2 equations in 2 unknowns m and c. They are known as the *normal equations*.

Summary: To find the line y = mx + c that is the best least squares fit to the data points

$$(x_1, y_1^*), (x_2, y_2^*), \dots, (x_n, y_n^*)$$

solve the normal equations

$$X^t X \mathbf{L} = X^t \mathbf{y}$$

where

$$X = \begin{bmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots \\ x_n & 1 \end{bmatrix}, \quad \mathbf{L} = \begin{bmatrix} m \\ c \end{bmatrix} \text{ and } \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

3.4.1 Example. Find the equation of the line that is the best least squares fit to the data points

(2,1), (5,2), (7,3), (8,3)

Solution: We take

$$X = \begin{bmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots \\ x_n & 1 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 5 & 1 \\ 7 & 1 \\ 8 & 1 \end{bmatrix}, \quad \mathbf{L} = \begin{bmatrix} m \\ c \end{bmatrix} \text{ and } \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 3 \end{bmatrix}$$

and solve the normal equations $X^t X \mathbf{L} = X^t \mathbf{y}$.

We need to calculate

$$X^{t}X = \begin{bmatrix} 2 & 5 & 7 & 8 \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 5 & 1 \\ 7 & 1 \\ 8 & 1 \end{bmatrix} = \begin{bmatrix} 142 & 22 \\ 22 & 4 \end{bmatrix}$$

and

$$X^{t}\mathbf{y} = \begin{bmatrix} 2 & 5 & 7 & 8 \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \\ 3 \end{bmatrix} = \begin{bmatrix} 57 \\ 9 \end{bmatrix}$$

And now solve

$$\begin{bmatrix} 142 & 22\\ 22 & 4 \end{bmatrix} \begin{bmatrix} m\\ c \end{bmatrix} = \begin{bmatrix} 57\\ 9 \end{bmatrix}$$

We can do that by row reducing

$$\begin{bmatrix} 142 & 22 & :57 \\ 22 & 4 & :9 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 11/71 & :57/142 \\ 22 & 4 & :9 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 11/71 & :57/142 \\ 0 & 42/71 & :12/71 \end{bmatrix}$$
$$\rightarrow \begin{bmatrix} 1 & 11/71 & :57/142 \\ 0 & 1 & :2/7 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & :5/14 \\ 0 & 1 & :2/7 \end{bmatrix}$$

So m = 5/14 and c = 2/7. The line is

$$y = \frac{5}{14}x + \frac{2}{7}$$

and here is a picture of the line with the points



3.5 Markov Matrices

One thing we could have mentioned earlier is that a matrix always has the same eigenvalues as its transpose. If A is an $n \times n$ (square) matrix we do know (from 1.4 (iv)) that $\det(A^t) = \det(A)$. It follows quite easily that A and A^t have the same characteristic equation. The reason is that

$$A^t - \lambda I_n = A^t - \lambda I_n^t = (A - \lambda I_n)^t$$

and so $\det(A^t - \lambda I_n) = \det(A - \lambda I_n)$. So the characteristic equation $\det(A^t - \lambda I_n) = 0$ is the same equation as $\det(A - \lambda I_n) = 0$.

One place where this little fact is useful is in studying Markov matrices. Markov matrices are square matrices which have

• all entries ≥ 0 , and

Eigenvalues, etc

• in each column, the sum of the entries in the column is 1

This is a 3×3 Markov matrix

$$A = \begin{bmatrix} 0.2 & 0 & 0.4 \\ 0.2 & 0.5 & 0.1 \\ 0.6 & 0.5 & 0.5 \end{bmatrix}$$

If we take the transpose of this Markov matrix and multiply it by the column of all 1's, we find

$$A^{t} \begin{bmatrix} 1\\1\\1 \end{bmatrix} = \begin{bmatrix} 0.2 & 0.2 & 0.6\\0 & 0.5 & 0.1\\0.4 & 0.1 & 0.5 \end{bmatrix} \begin{bmatrix} 1\\1\\1 \end{bmatrix} = \begin{bmatrix} 1\\1\\1 \end{bmatrix}$$

(because the row sums in the transpose are all equal 1).

What this shows is that if A is a Markov matrix, then the column of all 1's is an eigenvector of A^t with the eigenvalue 1. So there must be an eigenvector for the matrix A with the eigenvalue 1, that is a vector v with

 $A\mathbf{v} = \mathbf{v}$

That vector \mathbf{v} fixed by A is rather important in considerations of Markov matrices. There are many applications of Markov matrices (for example in finding what happens with genetic types over many generations) but we will stop our study of linear algebra in this module here, without looking into these interesting topics!