ES.1803 Topic 16 Notes

Jeremy Orloff

16 Eigenvalues, diagonalization, decoupling

This note covers topics that will take us several classes to get through. While we we will look at $n \times n$ matrices, most of our computational examples will use 2×2 matrices. These have almost all the features of bigger square matrices and they are computationally much easier.

16.1 Etymology:

This is from a Wikipedia discussion page: The word **eigen** in German or Dutch translates as 'inherent', 'characteristic', 'private'. So an eigenvector of a matrix is characteristic or inherent to the matrix. The word eigen is also translated as 'own' with the same sense as the meanings above. That is the eigenvector of a matrix is the matrix's 'own vector'.

In English you sometimes see eigenvalues called special or characteristic values.

16.2 Definition

For a square matrix M, an eigenvalue is a number (scalar) that satisfies the equation

$$M\,\mathbf{v} = \lambda\,\mathbf{v} \text{ for some non-zero vector } \mathbf{v}. \tag{1}$$

The vector $\mathbf{v}$ is called a non-zero eigenvector corresponding to λ. We will call Equation 16.1 the eigenvector equation.

Comments:
1. Using the symbol λ for the eigenvalue is a fairly common practice when looking at generic matrices. If the eigenvalue has a physical interpretation, we'll often use a corresponding letter. For example, in population matrices the eigenvalues are growth rates, so we'll often denote them using r or k.

2. Eigenvectors are not unique. That is, if $\mathbf{v}$ is an eigenvector with eigenvalue λ then so is any multiple of $\mathbf{v}$. Indeed, the set of all eigenvectors with eigenvalue λ is clearly a vector space. (You should convince yourself of this!)

16.3 Why eigenvectors are special

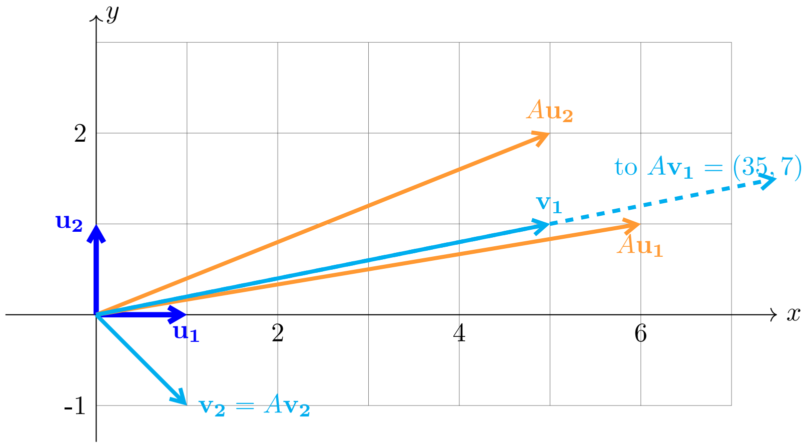
Example 16.1. Let $A = \begin{bmatrix} 6 & 5 \\ 1 & 2 \end{bmatrix}$. We will explore how A transforms vectors and what makes an eigenvector special. We will see that A scales and rotates most vectors, but only scales eigenvectors. That is, eigenvectors lie on lines that are unmoved by A.

Take $\mathbf{u_1} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \Rightarrow A\mathbf{u_1} = \begin{bmatrix} 6 \\ 1 \end{bmatrix}$; Take $\mathbf{u_2} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \Rightarrow A\mathbf{u_2} = \begin{bmatrix} 5 \\ 2 \end{bmatrix}$.

We see that A scales and turns most vectors.

Now take $\mathbf{v_1} = \begin{bmatrix} 5 \\ 1 \end{bmatrix} \Rightarrow A\mathbf{v_1} = \begin{bmatrix} 35 \\ 7 \end{bmatrix} = 7\mathbf{v_1}$. By the definition in Equation 1, this shows that $\mathbf{v_1}$ is an eigenvector with eigenvalue 7. The eigenvector is special since A scales it by 7, but does not rotate it.

Likewise, $\mathbf{v_2} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ then $A\mathbf{v_2} = \mathbf{v_2}$. So $\mathbf{v_2}$ is an eigenvector with eigenvalue 1. The eigenvector $\mathbf{v_2}$ is really special, it is unmoved by A.



Example 16.1: Action of the matrix A on vectors

The following example shows how knowing eigenvalues and eigenvectors simplifies calculations with a matrix. In fact, you don't even need the matrix once you know all of its eigenvalues and eigenvectors.

Example 16.2. Suppose A is a 2×2 matrix that has eigenvectors $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ 3 \end{bmatrix}$ with eigenvalues 2 and 4 respectively.

(a) Compute $A \begin{bmatrix} 1 \\ 2 \end{bmatrix}$.

Solution: Since $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ is an eigenvector, this follows directly from the definition of eigenvectors: $A \begin{bmatrix} 1 \\ 2 \end{bmatrix} = 2 \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$.

(b) Compute $A \left(\begin{bmatrix} 1 \\ 2 \end{bmatrix} + \begin{bmatrix} 1 \\ 3 \end{bmatrix} \right)$.

Solution: This uses the definition of eigenvector plus linearity:

$$A \left(\begin{bmatrix} 1 \\ 2 \end{bmatrix} + \begin{bmatrix} 1 \\ 3 \end{bmatrix} \right) = A \begin{bmatrix} 1 \\ 2 \end{bmatrix} + A \begin{bmatrix} 1 \\ 3 \end{bmatrix} = 2 \begin{bmatrix} 1 \\ 2 \end{bmatrix} + 4 \begin{bmatrix} 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 6 \\ 16 \end{bmatrix}.$$

(c) Compute $A \left(3 \begin{bmatrix} 1 \\ 2 \end{bmatrix} + 5 \begin{bmatrix} 1 \\ 3 \end{bmatrix} \right)$.

Solution: Again this uses the definition of eigenvector plus linearity:

$$A\left(3\begin{bmatrix}1\\2\end{bmatrix}+5\begin{bmatrix}1\\3\end{bmatrix}\right)=3A\begin{bmatrix}1\\2\end{bmatrix}+5A\begin{bmatrix}1\\3\end{bmatrix}=6\begin{bmatrix}1\\2\end{bmatrix}+20\begin{bmatrix}1\\3\end{bmatrix}=\begin{bmatrix}26\\72\end{bmatrix}.$$

(d) Compute $A\begin{bmatrix}0\\1\end{bmatrix}$.

Solution: We first decompose $\begin{bmatrix}0\\1\end{bmatrix}$ into eigenvectors:

$$\begin{bmatrix}0\\1\end{bmatrix}=\begin{bmatrix}1\\3\end{bmatrix}-\begin{bmatrix}1\\2\end{bmatrix}.$$

Now we can once again use the definition of eigenvector plus linearity:

$$A\begin{bmatrix}0\\1\end{bmatrix}=A\left(\begin{bmatrix}1\\3\end{bmatrix}-\begin{bmatrix}1\\2\end{bmatrix}\right)=A\begin{bmatrix}1\\3\end{bmatrix}-A\begin{bmatrix}1\\2\end{bmatrix}=4\begin{bmatrix}1\\3\end{bmatrix}+2\begin{bmatrix}1\\2\end{bmatrix}=\begin{bmatrix}2\\8\end{bmatrix}.$$

Example 16.3. Any rotation in three dimensions is around some axis. The vector along this axis is fixed by the rotation, i.e., it is an eigenvector with eigenvalue 1.

16.4 Computational algorithm

We start by summarizing the method. We will justify it and give examples below.

Computational method:
1. The eigenvalues of A are the roots of the characteristic equation

$$\det(A-\lambda I)=0 \tag{2}$$

2. The corresponding eigenspace of A is $\text{Null}(A-\lambda I)$.

Notes. 1. Again, we call Equation 2 the characteristic equation. (Eigenvalues are sometimes called characteristic values.) It allows us to find the eigenvalues and eigenvectors separately in a two step process.

2. The eigenspace is so-called, because it is the vector subspace which consists of all eigenvectors corresponding to λ.

3. **Notation:** For simplicity we will sometimes use the notation $|A|=\det(A)$. So the characteristic equation can be written $|A-\lambda I|=0$.

16.4.1 Justification of the computational algorithm

First we recall the following basic fact about square matrices from Topic 15.

Fact: The null space of A is nontrivial exactly when $\det(A)=0$.

Next, we manipulate the eigenvalue equation (Equation 1) so that finding eigenvectors becomes finding null vectors. Suppose, λ is an eigenvalue and $\mathbf{v}$ is a corresponding nonzero eigenvector. Then, starting with the eigenequation we have:

$$A\,\mathbf{v}=\lambda\,\mathbf{v}\ \Leftrightarrow\ A\,\mathbf{v}=\lambda\,I\mathbf{v}\ \Leftrightarrow\ A\,\mathbf{v}-\lambda\,I\mathbf{v}=0\ \Leftrightarrow\ (A-\lambda\,I)\mathbf{v}=0.$$

Since $\mathbf{v} \neq 0$, the last equation just above says $A - \lambda I$ has a nontrivial null space. So our fact about determinants and null spaces tells us that λ is an eigenvalue if and only if $\det(A - \lambda I) = 0$, i.e., if and only if λ is a root of the characteristic equation. This justifies Step 1 in the algorithm.

Likewise, the equation $(A - \lambda I)\mathbf{v} = 0$ says that $\mathbf{v}$ is an eigenvector corresponding to λ if and only if $\mathbf{v}$ is in $\text{Null}(A - \lambda I)$. This justifies Step 2 in the algorithm.

16.4.2 Examples

Example 16.4. Find the eigenvalues of the matrix $A = \begin{bmatrix} 6 & 5 \\ 1 & 2 \end{bmatrix}$. For each eigenvalue find a basis of the corresponding eigenspace.

Solution: Step 1. Find the eigenvalues λ: $|A - \lambda I| = 0$ (characteristic equation)

$$A - \lambda I = \begin{bmatrix} 6 & 5 \\ 1 & 2 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} = \begin{bmatrix} 6-\lambda & 5 \\ 1 & 2-\lambda \end{bmatrix}.$$

Taking the determinant and setting it to 0 gives

$$\det(A - \lambda I) = (6-\lambda)(2-\lambda) - 5 = \lambda^2 - 8\lambda + 7 = 0.$$

The roots of this are $\boxed{\lambda = 7,\ 1.}$

Step 2. For each eigenvalue, find basis vectors for the eigenspace, i.e., find a basis of $\text{Null}(A - \lambda I)$.

$\lambda_1 = 7$: $A - \lambda I = \begin{bmatrix} -1 & 5 \\ 1 & -5 \end{bmatrix}$. This has RREF $R = \begin{bmatrix} 1 & -5 \\ 0 & 0 \end{bmatrix}$. The null space is 1 dimensional, a basis is $\mathbf{v_1} = \begin{bmatrix} 5 \\ 1 \end{bmatrix}$.

$\lambda_1 = 1$: $A - \lambda I = \begin{bmatrix} 5 & 5 \\ 1 & 1 \end{bmatrix}$. This has RREF $R = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$. The null space is 1 dimensional, a basis is $\mathbf{v_2} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$.

Remember, any scalar multiple of these eigenvectors is also an eigenvector with the same eigenvalue.

Let's reemphasize a key point:

Example 16.5. Eigenspaces are null spaces. Consider the matrix

$$A = \begin{bmatrix} 4 & 8 & -2 & 2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix}$$

Find the eigenvalues and eigenspaces of A.

Solution: This is a 4×4 matrix, but the characteristic equation is not hard to find.

$$|A - \lambda I| = \begin{vmatrix} 4-\lambda & 8 & -2 & 2 \\ 0 & -\lambda & 0 & 0 \\ 0 & 0 & 1-\lambda & 1 \\ 0 & 0 & 1 & 1-\lambda \end{vmatrix} = (4-\lambda)(-\lambda)(\lambda^2 - 2\lambda) = -\lambda^2(4-\lambda)(\lambda-2).$$

So the eigenvalues are $\lambda = 0, 0, 4, 2$.

Eigenspace for $\lambda = 0$:

We must find Null(A): The RREF of A is $R = \begin{bmatrix} 1 & 2 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$.

Using this we see that Null(A) (eigenspace for $\lambda = 0$) is 2 dimensional and has basis

$$\left\{ \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ -1 \\ 1 \end{bmatrix} \right\}$$

Let's highlight that Null(A) is nontrivial means $\lambda = 0$ is an eigenvalue.

For the other two eigenvalues we must find Null($A - 4I$) and Null($A - 2I$). This is not hard and you should do it as an exercise. We get:

The eigenspace for $\lambda = 4$ has basis $\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \right\}$. The eigenspace for $\lambda = 2$ has basis $\left\{ \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} \right\}$.

Notes.

Trick. In the 2×2 case we don't have to write out the RREF to find the eigenvector. Notice that the entries in our eigenvectors come from the entries in one row of the matrix. The eigenvector is the column vector with entries: right entry of the row, minus the left entry. For example, if $A - \lambda I = \begin{bmatrix} -1 & 5 \\ 1 & -5 \end{bmatrix}$, then, using the top row, we see that $\mathbf{v} = \begin{bmatrix} 5 \\ 1 \end{bmatrix}$ is a basis vector for Null($A - \lambda I$). If you think about this a moment, you'll see why it must be the case.

Matlab: In Matlab the function `eig(A)` returns the eigenvectors and eigenvalues of a matrix.

16.5 Complex eigenvalues

If the eigenvalues are complex, then the eigenvectors are complex. Otherwise there is no difference in the algebra.

Example 16.6. Find the eigenvalues and basic eigenvectors of the matrix $A = \begin{bmatrix} 3 & 4 \\ -4 & 3 \end{bmatrix}$.

Solution: Step 1. Find the eigenvalues λ: $|A - \lambda I| = 0$ **(characteristic equation)**

$$\det\left(\begin{bmatrix} 3 & 4 \\ -4 & 3 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix}\right) = \begin{vmatrix} 3-\lambda & 4 \\ -4 & 3-\lambda \end{vmatrix} = (3-\lambda)^2 + 16 = 0.$$

So the eigenvalues are $\boxed{\lambda = 3 \pm 4i.}$

Step 2. Find corresponding basic eigenvectors. That is, find a basis of $\text{Null}(A - \lambda I)$. (In the 2×2 case, we can do that without doing row reduction.)

$\lambda_1 = 3 + 4i$: $(A - \lambda I) = \begin{bmatrix} -4i & 4 \\ -4 & -4i \end{bmatrix}$. Take $\boxed{\mathbf{v_1} = \begin{bmatrix} 1 \\ i \end{bmatrix}}$.

$\lambda_2 = 3 - 4i$: $(A - \lambda I) = \begin{bmatrix} 4i & 4 \\ -4 & 4i \end{bmatrix}$. Take $\boxed{\mathbf{v_2} = \begin{bmatrix} 1 \\ -i \end{bmatrix}}$.

Time saver: Notice that the eigenvalues and eigenvectors come in complex conjugate pairs. Knowing this, there is no need to do a computation to find the second member of each pair.

Example 16.7. Find the eigenvalues and basic eigenvectors of the matrix $A = \begin{bmatrix} 1 & -4 \\ 5 & 5 \end{bmatrix}$.

Solution: Step 1. Find λ (eigenvalues): $|A - \lambda I| = 0$ **(characteristic equation)**

$$\begin{vmatrix} 1-\lambda & -4 \\ 5 & 5-\lambda \end{vmatrix} = \lambda^2 - 6\lambda + 25 = 0 \Rightarrow \boxed{\lambda = 3 \pm 4i.}$$

Step 2. Find corresponding basic eigenvectors (basis of $\text{Null}(A - \lambda I)$):

$\lambda_1 = 3 + 4i$: $(A - \lambda I) = \begin{bmatrix} -2-4i & -4 \\ 5 & 2-4i \end{bmatrix}$. Take $\boxed{\mathbf{v_1} = \begin{bmatrix} 4 \\ -2-4i \end{bmatrix}}$.

$\lambda_2 = 3 - 4i$: Take $\boxed{\mathbf{v_2} = \overline{\mathbf{v}}_\mathbf{1} = \begin{bmatrix} 4 \\ -2+4i \end{bmatrix}}$.

16.6 Repeated eigenvalues

When a matrix has repeated eigenvalues the eigenvectors are not as well behaved as when the eigenvalues are distinct. There are two main examples

Example 16.8. (Defective case) Find the eigenvalues and basic eigenvectors of the matrix $A = \begin{bmatrix} 3 & 1 \\ 0 & 3 \end{bmatrix}$.

Solution: Step 1. Find the eigenvalues λ: $|A - \lambda I| = 0$ **(characteristic equation)**

$$\begin{vmatrix} 3-\lambda & 1 \\ 0 & 3-\lambda \end{vmatrix} = (\lambda - 3)^2 = 0 \Rightarrow \boxed{\lambda = 3,\ 3.}$$

Step 2. Find the basic eigenvectors (basis of $\text{Null}(A - \lambda I)$):

$\lambda_1 = 3$: $\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \mathbf{v_1}$. This is already in RREF. It has one free variable, so the null space is 1 dimensional. We can take a basis vector: $\boxed{\mathbf{v_1} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}}$.

We have two eigenvalues but only one independent eigenvector, so we call this case defective or incomplete. In linear algebra, there is a lot to explore with defective matrices. In 18.03, we will not go into a lot of detail about them.

Example 16.9. (Complete case) Find the eigenvalues and basic eigenvectors of the matrix $A = \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix}$.

Solution: Step 1. Find the eigenvalues λ: $|A - \lambda I| = 0$ **(characteristic equation)**

$$\begin{vmatrix} 3-\lambda & 0 \\ 0 & 3-\lambda \end{vmatrix} = (\lambda - 3)^2 = 0 \Rightarrow \boxed{\lambda = 3,\ 3.}$$

Step 2. Find corresponding basic eigenvectors (basis of Null$(A - \lambda I)$):

$\lambda_1 = 3$: $\quad A - \lambda I = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$.

This equation shows that every vector in $\mathbf{R}^2$ is an eigenvector. That is, the eigenvalue $\lambda = 3$ has a two dimensional eigenspace. We can pick any two independent vectors as a basis, e.g., $\mathbf{v_1} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\mathbf{v_2} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$. (These are the simplest choices, but any two independent vectors would work!)

Because we have as many independent eigenvectors as eigenvalues, we call this case complete.

16.7 Diagonal matrices

In this section we will see how easy it is to work with diagonal matrices. In later sections we will see how working with eigenvalues and eigenvectors of a matrix is like turning it into a diagonal matrix.

Example 16.10. Consider the diagonal matrix $B = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$

Convince yourself that $B \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} 2u \\ 3v \end{bmatrix}$. That is B scales the first coordinate by 2 and the second coordinate by 3.

We can write this as

$$B \begin{bmatrix} 1 \\ 0 \end{bmatrix} = 2 \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \text{and} \quad B \begin{bmatrix} 0 \\ 1 \end{bmatrix} = 3 \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

This is exactly the definition of eigenvectors. That is, $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ are eigenvectors with eigenvalues 2 and 3 respectively. We state this an an important fact.

Important fact. For a diagonal matrix, the diagonal entries are the eigenvalues and the eigenvectors are the standard basis vectors.

Example 16.11. The matrix $A = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 4 \end{bmatrix}$ has eigenvalues and corresponding basic eigenvectors

$$\begin{aligned} \lambda &= 2, \qquad 3, \qquad 4 \\ \mathbf{v} &= \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \end{aligned}$$

You can check this by multiplying A times each eigenvector.

Example 16.12. For the matrix A in the previous example, compute $\det A$, A^2, A^5.

Solution: $\det(A)$ = product of diagonal entries = 24.

$A^2 = \begin{bmatrix} 2^2 & 0 & 0 \\ 0 & 3^2 & 0 \\ 0 & 0 & 4^2 \end{bmatrix}$. Likewise, $A^5 = \begin{bmatrix} 2^5 & 0 & 0 \\ 0 & 3^5 & 0 \\ 0 & 0 & 4^5 \end{bmatrix}$

16.8 Diagonalization

Diagonalization is a way to make a matrix almost as easy to work with as a diagonal matrix.

Theorem. Diagonalization theorem. Suppose the $n \times n$ matrix A has n independent eigenvectors. Then, we can write

$$A = S\Lambda S^{-1},$$

where S is a matrix whose columns are the n independent eigenvectors and Λ is the diagonal matrix whose diagonal entries are the corresponding eigenvalues.

The proof is below. We illustrate this first with our standard example.

Example 16.13. We know the matrix $A = \begin{bmatrix} 6 & 5 \\ 1 & 2 \end{bmatrix}$ has eigenvalues 7 and 1 with corresponding basic eigenvectors $\mathbf{v_1} = \begin{bmatrix} 5 & 1 \end{bmatrix}^T$ and $\mathbf{v_2} = \begin{bmatrix} -1 & 1 \end{bmatrix}^T$

We put the eigenvectors as the columns of a matrix S and the eigenvalues as the entries of a diagonal matrix Λ.

$$S = \begin{bmatrix} \mathbf{v_1} & \mathbf{v_2} \end{bmatrix} = \begin{bmatrix} 5 & -1 \\ 1 & 1 \end{bmatrix}, \qquad \Lambda = \begin{bmatrix} 7 & 0 \\ 0 & 1 \end{bmatrix}$$

The diagonalization theorem says that

$$A = S\Lambda S^{-1} = \begin{bmatrix} 5 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 7 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1/6 & 1/6 \\ -1/6 & 5/6 \end{bmatrix}.$$

This is called the diagonalization of A. Note the form: a diagonal matrix Λ surrounded by S and S^{-1}.

Proof of the diagonalization theorem. We will do this for the matrix in the example above. It should be clear that this proof carries over to any $n \times n$ matrix with n independent eigenvectors.

The equation $A = S\Lambda S^{-1}$ can be rewritten as $AS = S\Lambda$. We will show this is true by showing that both sides have the same effect when multiplying any vector. That is,

$$AS\mathbf{v} = S\Lambda\mathbf{v}$$

for any vector $\mathbf{v}$

First, let $\mathbf{e_1} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $\mathbf{e_2} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ be the standard basis vectors of $\mathbf{R}^2$. Since every vector is a linear combination of the basis vectors, it is enough to show

$$AS\mathbf{e_1} = S\Lambda\mathbf{e_1} \quad \text{and} \quad AS\mathbf{e_2} = S\Lambda\mathbf{e_2}.$$

Recall that multiplying a matrix times a column vector results in a linear combination of the columns. In our case,

$$S\mathbf{e_1} = \begin{bmatrix} \mathbf{v_1}\, \mathbf{v_2} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \mathbf{v_1}, \quad \text{and} \quad S\mathbf{e_2} = \begin{bmatrix} \mathbf{v_1}\, \mathbf{v_2} \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \mathbf{v_2}.$$

Now we can check that $AS\mathbf{e_1} = S\Lambda\mathbf{e_1}$:

$$AS\mathbf{e_1} = A\mathbf{v_1} = 7\mathbf{v_1} \quad \text{and} \quad S\Lambda\mathbf{e_1} = S \begin{bmatrix} 7 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = S \begin{bmatrix} 7 \\ 0 \end{bmatrix} = 7\mathbf{v_1}.$$

The equation $A\mathbf{v_1} = 7\mathbf{v_1}$ follows because $\mathbf{v_1}$ is an eigenvector of A with eigenvalue 7. Thus we have shown that $AS\mathbf{e_1} = S\Lambda\mathbf{e_1}$. In exactly the same way, we can show that $AS\mathbf{e_2} = S\Lambda\mathbf{e_2}$.

Thus we can conclude that $AS = S\Lambda$. So, $A = S\Lambda S^{-1}$.

In general, the steps for diagonalizing an $n \times n$ matrix A are:

1. Find the eigenvalues $\lambda_1, \dots, \lambda_n$ and corresponding basic eigenvectors $\mathbf{v_1}, \dots, \mathbf{v_n}$.
2. Make the matrix of eigenvectors $S = \begin{bmatrix} \mathbf{v_1} & \mathbf{v_2} & \cdots & \mathbf{v_n} \end{bmatrix}$
3. Make the diagonal matrix of eigenvalues $\Lambda = \begin{bmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & \cdots & \lambda_n \end{bmatrix}$

The diagonalization is: $A = S\Lambda S^{-1}$.

Note: Diagonalization requires that A have a full complement of eigenvectors. If A is defective, it can't be diagonalized.

We have the following important formula

$$\det(\mathbf{A}) = \textbf{ product of its eigenvalues.}$$

This follows easily from the diagonalization formula

$$\det(A) = \det(S\Lambda S^{-1}) = \det(S)\det(\Lambda)\det(S^{-1}) = \det(\Lambda) = \text{ product of diagonal entries.}$$

Example 16.14. Consider the matrix $A = \begin{bmatrix} 5 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 7 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 5 & -1 \\ 1 & 1 \end{bmatrix}^{-1}$.

(a) What are the eigenvalues and eigenvectors of A.

(b) Compute $\det A$, A^2, A^5.

Solution: For ease of writing, let $S = \begin{bmatrix} 5 & -1 \\ 1 & 1 \end{bmatrix}$ and $\Lambda = \begin{bmatrix} 7 & 0 \\ 0 & 1 \end{bmatrix}$. So, $A = S\Lambda S^{-1}$.

(a) The columns of S are eigenvectors and the diagonal entries of Λ are the corresponding eigenvalues. We have eigenpairs

$$\lambda = 7,\ \mathbf{v} = \begin{bmatrix} 5 \\ 1 \end{bmatrix} \qquad \text{and} \qquad \lambda = 1,\ \mathbf{v} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}.$$

(b) We have $\det A = \det \Lambda = 7$. We also have

$$A^2 = S\Lambda S^{-1} \cdot S\Lambda S^{-1} = S\Lambda^2 S^{-1} = S \begin{bmatrix} 7^2 & 0 \\ 0 & 1^2 \end{bmatrix} S^{-1}.$$

Likewise $A^5 = S\Lambda^5 S^{-1} = S \begin{bmatrix} 7^5 & 0 \\ 0 & 1^5 \end{bmatrix} S^{-1}$.

16.9 Diagonal matrices and uncoupled algebraic systems

Example 16.15. (An uncoupled algebraic system) Consider the system

$$\begin{aligned} 7u \qquad &= 1 \\ v &= 3 \end{aligned}$$

The variables u and v are uncoupled. That is, they never occur in the same equation. We can solve the system by finding each variable separately: $u = 1/7$, $v = 3$.

Example 16.16. Now consider the system

$$\begin{aligned} 6x &+ 5y = 2 \\ x &+ 2y = 4. \end{aligned}$$

In matrix form this is

$$\begin{bmatrix} 6 & 5 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \end{bmatrix} \tag{3}$$

The matrix $A = \begin{bmatrix} 6 & 5 \\ 1 & 2 \end{bmatrix}$ is the same matrix as in Examples 16.1 and 16.4 above. In this system the variables x and y are coupled. We will explain the logic of decoupling later. For this example, we will decouple the equations using some magical choices involving eigenvectors.

The examples above showed that the eigenvalues of A are 7 and 1 with eigenvectors $\begin{bmatrix} 5 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$. We write all vectors in terms of the eigenvectors by making the change of variables

$$\begin{bmatrix} x \\ y \end{bmatrix} = u \begin{bmatrix} 5 \\ 1 \end{bmatrix} + v \begin{bmatrix} -1 \\ 1 \end{bmatrix} \Leftrightarrow x = 5u - v; \quad y = u + v.$$

For the future, we note: $\begin{bmatrix} 2 \\ 4 \end{bmatrix} = \begin{bmatrix} 5 \\ 1 \end{bmatrix} + 3 \begin{bmatrix} -1 \\ 1 \end{bmatrix}$.

Converting our equation from x and y to u and v we get

$$\begin{bmatrix} 6 & 5 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 6 & 5 \\ 1 & 2 \end{bmatrix} \left(u \begin{bmatrix} 5 \\ 1 \end{bmatrix} + v \begin{bmatrix} -1 \\ 1 \end{bmatrix} \right) = 7u \begin{bmatrix} 5 \\ 1 \end{bmatrix} + v \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

Thus,

$$\begin{bmatrix} 6 & 5 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \end{bmatrix} \qquad \Leftrightarrow \qquad 7u \begin{bmatrix} 5 \\ 1 \end{bmatrix} + v \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 5 \\ 1 \end{bmatrix} + 3 \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

It is easy to see that the last system is the same as the equations

$$7u = 1 \qquad v = 3.$$

In u, v coordinates the system is diagonal and easy to solve.

16.10 Introduction to matrix methods for solving systems of DEs

In this section we will solve linear, homogeneous, constant coefficient systems of differential equations using the matrix methods we have developed. For now we will just consider matrices with real, distinct eigenvalues. In the next topic we will look at complex and repeated eigenvalues.

As with constant coefficient DEs, we will use the method of optimism to discover a systematic technique for solving systems of DEs. We start by giving the general 2×2 linear, homogeneous, constant coefficient system of DEs. It has the form

$$\begin{array}{rcccc} x' & = & ax & + & by \\ y' & = & cx & + & dy. \end{array} \qquad (4)$$

Here a, b, c, d are constants and $x(t)$, $y(t)$ are the unknown functions we need to solve for.

There are a number of important things to note.

1. We can write Equation 4 in matrix form

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \Leftrightarrow \mathbf{x}' = A\mathbf{x} \qquad (5)$$

where $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ and $\mathbf{x} = \begin{bmatrix} x \\ y \end{bmatrix}$.

2. The system is homogeneous. You can see this by taking Equation 4 and putting all the x and y on the left side so that the right side becomes all zeros.

3. The system is linear. You should be able to check directly that a linear combination of solutions to Equation 5 is also a solution.

We illustrate the method of optimism for solving Equation 5 with an example.

Example 16.17. Solve the linear, homogeneous, constant coefficient system

$$\mathbf{x}' = A\mathbf{x}, \qquad \text{where} \qquad \mathbf{x} = \begin{bmatrix} x \\ y \end{bmatrix} \qquad \text{and} \qquad A = \begin{bmatrix} 6 & 5 \\ 1 & 2 \end{bmatrix}.$$

Solution: Using the method of optimism we try a solution

$$\mathbf{x} = e^{\lambda t}\mathbf{v},$$

where λ is a constant and $\mathbf{v}$ is a constant vector. Substituting the trial solution into both sides of the DE we get

$$\lambda e^{\lambda t}\mathbf{v} = e^{\lambda t}A\mathbf{v} \;\Leftrightarrow\; A\mathbf{v} = \lambda\mathbf{v}.$$

This is none other than the eigenvalue/eigenvector equation. So solving the system amounts to finding eigenvalues and eigenvectors. From our previous examples we know the eigenvalues and eigenvectors of A. We get two solutions.

$$\mathbf{x_1} = e^t \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \qquad \mathbf{x_2} = e^{7t} \begin{bmatrix} 5 \\ 1 \end{bmatrix}.$$

The general solution is the span of these solutions:

$$\mathbf{x} = c_1\mathbf{x_1} + c_2\mathbf{x_2} = c_1 e^t \begin{bmatrix} -1 \\ 1 \end{bmatrix} + c_2 e^{7t} \begin{bmatrix} 5 \\ 1 \end{bmatrix}$$

The solutions $\mathbf{x_1}$ and $\mathbf{x_2}$ are called modal or basic solutions.

Now that we know where the method of optimism leads, we can do a second example starting directly with finding eigenvalues and eigenvectors

Example 16.18. Find the general solution to the system

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} 3 & 4 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

Solution: First find eigenvalues and basic eigenvectors.

Characteristic equation: $\begin{vmatrix} 3-\lambda & 4 \\ 1 & 3-\lambda \end{vmatrix} = \lambda^2 - 6\lambda + 5 = 0 \;\Rightarrow\; \lambda = 1,\, 5.$

Basic eigenvectors: (basis of $\text{Null}(A - \lambda I)$):

$\lambda = 1$: $(A - \lambda I) = \begin{bmatrix} 2 & 4 \\ 1 & 2 \end{bmatrix}$. Take $\mathbf{v_1} = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$.

$\lambda = 5$: $(A - \lambda I) = \begin{bmatrix} -2 & 4 \\ 1 & - \end{bmatrix}$. Take $\mathbf{v_2} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$.

We have two modal solutions: $\mathbf{x_1} = e^t\mathbf{v_1}$ and $\mathbf{x_2} = e^{5t}\mathbf{v_2}$.

The general solution is $\mathbf{x} = c_1\mathbf{x_1} + c_2\mathbf{x_2} = c_1 e^t \begin{bmatrix} -2 \\ 1 \end{bmatrix} + c_2 e^{5t} \begin{bmatrix} 2 \\ 1 \end{bmatrix}$.

16.11 Decoupling systems of DEs

Example 16.19. (**An uncoupled system**) Consider the system

$$\begin{aligned} u'(t) &= 7u(t) \\ v'(t) &= v(t) \end{aligned}$$

Since u and v don't have any effect on each other, we say that u and v are uncoupled. It's easy to see the solution to this system is

$$\begin{aligned} u(t) &= c_1 e^{7t} \\ v(t) &= c_2 e^{t} \end{aligned}$$

In matrix form we have

$$\begin{bmatrix} u' \\ v' \end{bmatrix} = \begin{bmatrix} 7 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix}.$$

The coefficient matrix has eigenvalues 7 and 1, with basic eigenvectors $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$. The general solution to the system of DEs is

$$\begin{bmatrix} u \\ v \end{bmatrix} = c_1 e^{7t} \begin{bmatrix} 1 \\ 0 \end{bmatrix} + c_2 e^{t} \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

We see an uncoupled system has a diagonal coefficient matrix and the basic eigenvectors are the standard basis vectors. All in all, it's simple and easy to work with.

The following example shows how to decouple a coupled system. After seeing this example, we will redo it, in a cleaner, more memorable way.

Example 16.20. Consider once again the system from Example 16.17

$$\begin{aligned} x' &= 6x + 5y \\ y' &= x + 2y. \end{aligned} \Leftrightarrow \mathbf{x}' = A\mathbf{x}, \quad \text{where} \quad \mathbf{x} = \begin{bmatrix} x \\ y \end{bmatrix}, \quad A = \begin{bmatrix} 6 & 5 \\ 1 & 2 \end{bmatrix} \tag{6}$$

In this system the variables x and y are coupled. Make a change of variable that converts this to a decoupled system.

Solution: From Example 16.17 we know the eigenvalues are 7 and 1, basic eigenvectors are $\begin{bmatrix} 5 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$, and the general solution is $\begin{bmatrix} x \\ y \end{bmatrix} = c_1 e^{7t} \begin{bmatrix} 5 \\ 1 \end{bmatrix} + c_2 e^{t} \begin{bmatrix} -1 \\ 1 \end{bmatrix}$.

Notice that $c_1 e^{7t}$ and $c_2 e^{t}$ in the above solution are just u and v from the previous example. So we can write

$$\begin{bmatrix} x(t) \\ y(t) \end{bmatrix} = u(t) \begin{bmatrix} 5 \\ 1 \end{bmatrix} + v(t) \begin{bmatrix} -1 \\ 1 \end{bmatrix}. \tag{7}$$

This is a change of variables.

Let's rewrite the system in Equation 6 in terms of u, v. Using Equation 7, we get

$$\mathbf{x}' = \begin{bmatrix} x' \\ y' \end{bmatrix} = u' \begin{bmatrix} 5 \\ 1 \end{bmatrix} + v' \begin{bmatrix} -1 \\ 1 \end{bmatrix} \quad \text{and} \quad A \begin{bmatrix} x \\ y \end{bmatrix} = A \left(u \begin{bmatrix} 5 \\ 1 \end{bmatrix} + v \begin{bmatrix} -1 \\ 1 \end{bmatrix} \right) = 7u \begin{bmatrix} 5 \\ 1 \end{bmatrix} + v \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

The last equality follows because $\begin{bmatrix} 5 & 1 \end{bmatrix}^T$ and $\begin{bmatrix} -1 & 1 \end{bmatrix}^T$ are eigenvectors of A.

Equating the two sides we get

$$u' \begin{bmatrix} 5 \\ 1 \end{bmatrix} + v' \begin{bmatrix} -1 \\ 1 \end{bmatrix} = 7u \begin{bmatrix} 5 \\ 1 \end{bmatrix} + v \begin{bmatrix} -1 \\ 1 \end{bmatrix}.$$

Comparing the coefficients of the eigenvectors we get

$$\begin{aligned} u' &= 7u \\ v' &= v \end{aligned} \Leftrightarrow \begin{bmatrix} u' \\ v' \end{bmatrix} = \begin{bmatrix} 7 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix}$$

That is, in terms of u and v the system is uncoupled. Note that the eigenvalues of A are precisely the diagonal entries of the uncoupled system.

16.11.1 Decoupling in general

Though it's somewhat disguised, the key to the previous example was diagonalization. Bringing this to the forefront makes the example cleaner and less complicated.

Suppose A is written in diagonalized form: $A = S\Lambda S^{-1}$, where, as usual, S is a matrix with the eigenvectors of A as columns and Λ is the diagonal matrix with the corresponding eigenvalues as entries.

Decoupling: Suppose we have the system $\mathbf{x}' = A\mathbf{x}$, then the change of variables

$$\mathbf{u} = S^{-1}\mathbf{x}$$

converts the coupled system into an uncoupled system $\mathbf{u}' = \Lambda\mathbf{u}$.

Proof. The key is diagonalization: the system $\mathbf{x}' = A\mathbf{x}$ can be written

$$\mathbf{x}' = S\Lambda S^{-1}\mathbf{x} \quad \Leftrightarrow \quad S^{-1}\mathbf{x}' = \Lambda S^{-1}\mathbf{x}.$$

Now, letting $\mathbf{u} = S^{-1}\mathbf{x}$ converts this to the uncoupled system

$$\mathbf{u}' = \Lambda\mathbf{u}.$$

Since this is an uncoupled equation, making the change of variables $\mathbf{u} = S^{-1}\mathbf{x}$ is called decoupling the system.

To end this section, we'll walk through the previous example, being more explicit about the use of diagonalization.

Example 16.21. Decouple the system in Example 16.20 using the diagonalized form of A.

Solution: The system in Example 16.20 is $\mathbf{x}' = A\mathbf{x}$.

Let $S = \begin{bmatrix} 5 & -1 \\ 1 & 1 \end{bmatrix}$ = the matrix with eigenvectors of A as columns.

Let $\Lambda = \begin{bmatrix} 7 & 0 \\ 0 & 1 \end{bmatrix}$ = the diagonal matrix with the eigenvalues of A as diagonal entries.

Diagonalization says that $A = S\Lambda S^{-1}$.

The decoupling change of variables is $\mathbf{u} = S^{-1}\mathbf{x}$. We can write this as

$$\mathbf{x} = S\mathbf{u} \quad \text{or} \quad \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 5 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = u \begin{bmatrix} 5 \\ 1 \end{bmatrix} + v \begin{bmatrix} -1 \\ 1 \end{bmatrix}.$$

This is exactly the change of variables used in Example 16.20.

The decoupled system is

$$\mathbf{u}' = \Lambda\mathbf{u} \quad \text{or} \quad \begin{bmatrix} u' \\ v' \end{bmatrix} = \begin{bmatrix} 7 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix},$$

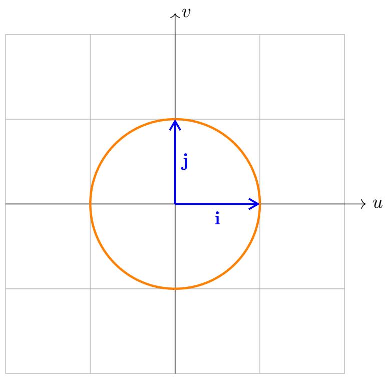
which is exactly the decoupled system found in Example 16.20.

16.12 Appendix: symmetric matrices

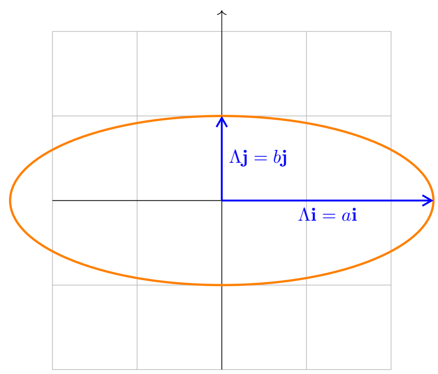
This section is optional. We won't ask about it on psets or tests. The first example in this section is a nice exercise in thinking about matrix multiplication as a way to transform vectors.

Example 16.22. Geometry of symmetric matrices. This is a fairly complex example showing how we can use the diagonal matrix $\Lambda = \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}$ and the rotation matrix $R = \begin{bmatrix} \cos\theta & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}$ to convert a circle to an ellipse as shown in the figures below.

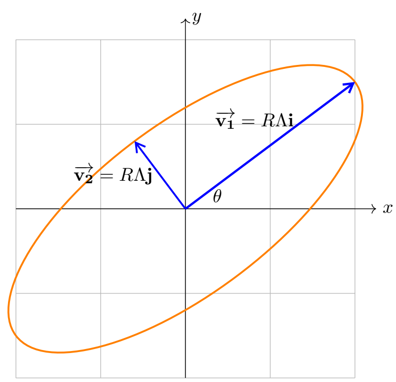
To do this, we think of matrix multiplication as a linear transformation. The diagonal matrix Λ transforms the circle by scaling the x and y directions by a and b respectively. This creates the ellipse in Figure (b), which is oriented with the axes. The rotation matrix R then rotates this ellipse to the general ellipse in Figure (c).



(a) Unit circle



(b) Ellipse made by scaling the axes by a and b respectively



(c) Ellipse in (b) rotated by θ

In coordinates $R\Lambda$ maps the unit circle $u^2 + v^2 = 1$ to the ellipse shown in (c). That is,

$$R\Lambda \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} u \\ v \end{bmatrix} = \Lambda^{-1} R^{-1} \begin{bmatrix} x \\ y \end{bmatrix}$$

Example 16.23. Spectral theorem. The previous example transforms the unit circle in uv-coordinates into an ellipse in xy-coordinates. In terms of inner products and transposes this becomes

$$\begin{aligned} 1 &= \left\langle \begin{bmatrix} u \\ v \end{bmatrix}, \begin{bmatrix} u \\ v \end{bmatrix} \right\rangle \\ &= \left\langle \Lambda^{-1} R^{-1} \begin{bmatrix} x \\ y \end{bmatrix}, \Lambda^{-1} R^{-1} \begin{bmatrix} x \\ y \end{bmatrix} \right\rangle \\ &= \begin{bmatrix} x \\ y \end{bmatrix}^T (\Lambda^{-1} R^{-1})^T \Lambda^{-1} R^{-1} \begin{bmatrix} x \\ y \end{bmatrix} \\ &= \begin{bmatrix} x \\ y \end{bmatrix}^T R \Lambda^{-2} R^{-1} \begin{bmatrix} x \\ y \end{bmatrix} \end{aligned}$$

The last equality uses the facts that for a rotation matrix $R^T = R^{-1}$ and for a diagonal matrix $\Lambda^T = \Lambda$.

Call the matrix occurring in the last two lines above A. That is,

$$A = R\Lambda^{-2}R^{-1} = (\Lambda^{-1}R^{-1})^T \Lambda^{-1} R^{-1}.$$

We then have the equation of the ellipse is

$$1 = \begin{bmatrix} x \\ y \end{bmatrix}^T A \begin{bmatrix} x \\ y \end{bmatrix}.$$

The matrix A has the following properties

1. It is symmetric

2. Its eigenvalues are a^{-2} and b^{-2}

3. Its eigenvectors are the the vectors $\overrightarrow{\mathbf{v_1}}$ and $\overrightarrow{\mathbf{v_2}}$ along the axes of the ellipse (see figure (c) above).

4. Its eigenvectors are orthogonal.

Proof.

1. This is clear from the formula $A = B^T B$ where $B = \Lambda^{-1} R^{-1}$.

2. This is clear from the diagonalization $A = R\Lambda^{-2}R^{-1}$. (Remember the eigenvalues are in the diagonal matrix Λ^{-2}.

3. We need to show that A transforms $\overrightarrow{\mathbf{v_1}}$ to a multiple of itself. This also follows by considering the action of each term in the diagonalization in turn (see the figures): R^{-1} moves $\overrightarrow{\mathbf{v_1}}$ to the x-axis; then Λ^{-2} scales the x-axis by a^{-2}; and finally R rotates the x-axis back the line along $\overrightarrow{\mathbf{v_1}}$. Using symbols

$$A\overrightarrow{\mathbf{v_1}} = R\Lambda^{-2}R^{-1}\overrightarrow{\mathbf{v_1}} = R\Lambda^{-2}a\mathbf{i} = R(a^{-2}a\mathbf{i}) = a^{-2}\overrightarrow{\mathbf{v_1}}$$

The properties of A are general properties of symmetric matrices.

Spectral theorem. A symmetric matrix A has the following properties.

1. It has real eigenvalues.

2. Its eigenvectors are mutually orthogonal.

Because of the connection to the axes of ellipses this is also called the principal axis theorem.