Linear Algebra – Lessons 5-6

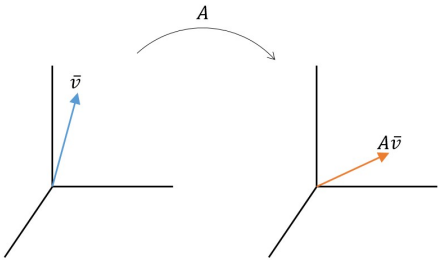
Matrix diagonalization

After learning about matrices as linear transformations and as tools for solving linear systems of equations, today we will discuss two key aspects in linear algebra – eigenvectors and eigenvalues. These new concepts will prove invaluable when considering dynamical systems, as you will see in the courses "Dynamical systems and the neuron" and "Theoretical and computational neuroscience" A and B.

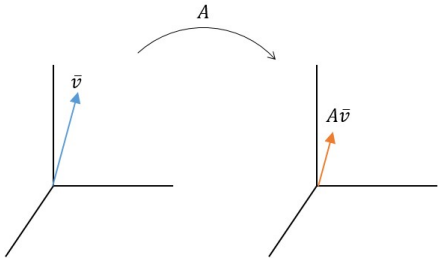
- We will start by discussing eigenvectors and eigenvalues.
- Then we will talk about changing bases.
- Finally, we will combine the two concepts and talk about the basis of the eigenvectors.

1. Eigenvectors and eigenvalues

In general, we think of a matrix A as representing some linear transformation – it takes a vector $\bar{v}$ as input, and gives some new vector $A\bar{v}$ as output:



We are interested in those special cases where the output vector is in the same direction as the input:



Definition: A non-zero vector $\bar{v}$ is called "an eigenvector of A" if $\bar{v}$ is parallel to $A\bar{v}$:

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

λ is called the eigenvalue of the eigenvector $\bar{v}$.
λ is a scalar, and it determines by how much A scales $\bar{v}$.

Eigen in German means "own". Indeed, eigenvectors and eigenvalues are "characteristic" of a matrix.

1.1 Intuitive examples

1.1.1 Projection onto a plane (e.g., the xy plan)

What are the eigenvalues and eigenvectors of the projection onto the plane?

Let's take $A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$. Any vector in the XY plane satisfies:

$$A\bar{v} = \bar{v} \quad \rightarrow \quad \lambda = 1$$

So we see that it's possible for many eigenvectors to have the same eigenvalue.

But there is another eigenvalue. Any vector that is perpendicular to the plane (i,e., a vector along $\hat{k}$) gives:

$$A\bar{v} = 0 \quad \rightarrow \lambda = 0$$

Notice that the last example also satisfies the definition of eigenvectors, since:

$$A\bar{v} = 0\bar{v}$$

So we see that 0 is a perfectly legitimate eigenvalue.

1.1.2 Permutation matrix

Let $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. This matrix switches between the two axes. What are its eigenvectors and eigenvalues? If we think about it geometrically, it seems reasonable that vectors on the diagonal will be eigenvectors. These are the only vectors that "don't mind" if you switch the x and y coordinates. Indeed:

$$\bar{v} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad A\bar{v} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \rightarrow \lambda = 1$$

$$\bar{v} = \begin{pmatrix} -1 \\ 1 \end{pmatrix} \quad A\bar{v} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \end{pmatrix} \rightarrow \lambda = -1$$

1.1.3 Shear

Now let A represent a 2D shear, for example:

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

What are its eigenvectors?

The first column is $\begin{pmatrix} 2 \\ 0 \end{pmatrix}$, which is just a scaled version of $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$, so $\bar{v} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ is an eigenvector with $\lambda = 2$.

Which vector can we guess might be another eigenvector? Well, the shear stretches the component along $\hat{\imath}$, and also adds a component along $\hat{\imath}$ to the vectors along $\hat{\jmath}$. It therefore makes sense that if we look for a vector that is already between the two principal axes ($\hat{\imath}, \hat{\jmath}$) it will be an eigenvector. Indeed, if we look at $\bar{v} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$:

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

Its eigenvalue is of course 1.

Soon we will see how to find these systematically.

1.2 The sum of the eigenvalues equals the trace, their product – the determinant

Two important theorems relate the matrix eigenvalues to known characteristics of the matrix, its trace and determinant.

The sum of the eigenvalues equals the trace: $\sum_i \lambda_i = Tr(A)$

The product of the eigenvalues equals the determinant: $\Pi_i \lambda_i = \det(A)$

***If A is diagonalizable**, then the number of non-zero eigenvalues equals the rank[1]:* $\sum_{i=1}^{n}(\lambda_i \neq 0) = rank(A)$

[1] This statement is true only if the matrix is diagonalizable (that is, it has n linearly independent eigenvectors). In the general case (you can read about it here), Let A be an $n \times n$ matrix. The number of linearly independent eigenvectors corresponding to the eigenvalue 0 is equal to $dim(\ker(A))$. By the rank-nullity theorem you know

Notice that this is true for the example above. For example, for the permutation matrix:

$$tr(A) = 0 \qquad (= 1 + (-1))$$
$$\det(A) = -1 \quad (= 1 \cdot -1)$$

But how can we find these eigenvectors and eigenvalues systematically?

1.3 A systematic procedure for finding eigenvectors and eigenvalues

Step 1: The characteristic polynomial

We start with the eigenvector definition:

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

You might think we should treat it like a system of linear equations, but it doesn't really resemble any case we have seen so far, since both $\bar{v}$ and λ are unknown.

First, note that we have two kinds of multiplication we here (matrix-vector and scalar-vector).

Remember we can think of $\lambda\bar{v}$ as $\lambda I\bar{v} = \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \dots & 0 \\ 0 & 0 & \lambda \end{pmatrix}\bar{v}$, so:

$$(A - \lambda I)\bar{v} = 0$$

Now keep good track of the next argument.

We still don't know what $\bar{v}$ is, but we do know something about it: it is a non-zero vector that is sent to 0. In other words, **$\bar{v}$ is in the kernel of $(A - \lambda I)$**. But this means that the kernel is not empty, so **this new matrix is singular**.

This means that:

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

In this new equation, $\bar{v}$ does not appear, so we can now just solve for λ.

We call the resulting equation the characteristic polynomial of A. For an $n \times n$ matrix, $\det(A - \lambda I)$ is a polynomial of degree n in λ.

To summarize this reasoning:

$$\bar{v} \neq 0, \;\; \bar{v} \in ker(A - \lambda I) \quad \rightarrow \quad ker(A - \lambda I) \neq \{0\} \quad \rightarrow \quad \det(A - \lambda I) = 0$$

Step 2: Solve for each λ

Solving $\det(A - \lambda I) = 0$ means finding all the possible values of λ that satisfy this condition.

Once we find some specific λ_0, we will plug it back in the equation, and solve for $\bar{v}$, since we now have a regular homogeneous equation:

$$(A - \lambda_0 I)\bar{v} = 0$$

So for any eigenvalue λ_0 we can try to find a relevant eigenvector (or several independent eigenvectors). The eigenvectors we find for λ_0 are called a basis for the eigenspace of λ_0.

---

that $dim(ker(A)) + rank = n$. If A is diagonalizable, this reduces to the theorem above. Here is a case of a matrix with only 0 as its eigenvalue (algebraic multiplicity of 2), but with $rank = 1$: $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. By the end of the course, you can read about a relation between the number of non-zero eigenvalues of A and the matrix A^TA here.

Example
Let's look at the same shear example from above:

$$A = \begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix}$$
$$\det(A - \lambda I) = det\begin{pmatrix} 2-\lambda & 1 \\ 0 & 1-\lambda \end{pmatrix} = (2-\lambda)(1-\lambda) = 0$$
$$\lambda_1 = 2, \quad \lambda_2 = 1$$

Now, to find an eigenvector for λ_1:

$$\begin{pmatrix} 2-\lambda_1 & 1 \\ 0 & 1-\lambda_1 \end{pmatrix}\bar{v} = \begin{pmatrix} 0 & 1 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

This is equivalent to a single equation:

$$v_2 = 0$$

This means there is no constraint of v_1, and we can choose whichever v_1 we want. For example $v_1 = 1$.
So we can have this eigenvector-eigenvalue pair:

$$\bar{v}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \lambda_1 = 2$$

This makes sense, because just by looking at the matrix we can see that $\hat{\imath}$ is simply scaled by 2.

To find an eigenvector for λ_2:

$$\begin{pmatrix} 2-\lambda_2 & 1 \\ 0 & 1-\lambda_2 \end{pmatrix}\bar{v} = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

These are actually two equations:

$$v_1 + v_2 = 0 \text{ and } 0 = 0$$
$$\rightarrow v_1 = -v_2$$
$$\rightarrow \bar{v}_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \lambda_2 = 1$$

Note that it's not accurate to say "find the eigenvector" of some eigenvalue. We always have freedom in scaling the eigenvector (if $\bar{v}$ is an eigenvector with some eigenvalue, so is $k\bar{v}$).
In some cases, we have even greater freedom. For example, what would happen if you look for eigenvectors of the identity matrix I?

Three important theorems

Theorem 1: A square matrix A is invertible iff 0 is not an eigenvalue of A.

Theorem 2: Eigenvectors that correspond to **distinct** eigenvalues are linearly independent.
Note that the opposite is not true: you might have similar eigenvalues that still have linearly independent eigenvectors.

Theorem 3: For a triangular matrix, the eigenvalues are the entries on its main diagonal.

1.4 The multiplicity of an eigenvalue

Definition: The algebraic multiplicity of an eigenvalue λ_0 is the number of times the factor $(\lambda - \lambda_0)$ appears in the characteristic polynomial.
For example, what is the multiplicity of the eigenvalue $\lambda = 1$ in the 3D identity matrix? The answer is 3.

Definition: The geometric multiplicity of an eigenvalue λ_0 is the number of independent eigenvectors associated with it.

Note that the algebraic multiplicity and geometric multiplicity of a given eigenvalue can differ. However, the geometric multiplicity can never exceed the algebraic multiplicity. Moreover, by definition, every eigenvalue must have at least one independent eigenvector associated with it (therefore, the geometrical multiplicity is never 0).

1.5 Eigenvalues can be complex numbers

Let's look at a rotation 90 degrees counterclockwise.

Question: What would be the matrix that represents this transformation?

Answer: Just by looking where the basis vectors $\hat{\imath}$ and $\hat{\jmath}$ land, we can infer that $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$

Now, if we use the theorems from before:

$$tr(A) = 0 + 0 = \lambda_1 + \lambda_2$$
$$\det(A) = 1 = \lambda_1 \lambda_2$$

We can already see what the eigenvalues will be. If we do it systematically:

$$\det(A - \lambda I) = \det\begin{pmatrix} -\lambda & -1 \\ 1 & -\lambda \end{pmatrix} = \lambda^2 + 1 = 0$$
$$\rightarrow \lambda_1 = i, \lambda_2 = -i$$

As you can see, even though all matrix entries were real, the eigenvalues are complex (and in this case, since this is an anti-symmetric matrix, purely imaginary).

Question: What are the eigenvectors corresponding to λ_1, λ_2?

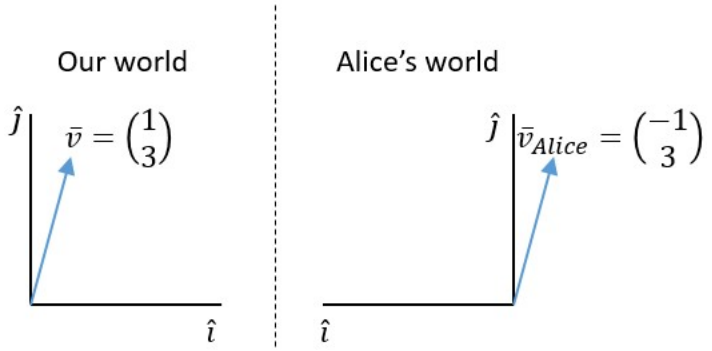
Answer: As long as we're talking about real matrices and vectors, there are no eigenvectors for this matrix. To talk about the eigenvectors of the rotation matrix, we must move on to talking about vectors over the complex field, $\bar{v} \in \mathbb{C}$, where the vector's entries can be complex. We won't do this now.

2. Change of basis

2.1 Different bases can describe the same vector

To make good use of eigenvalues and eigenvectors, we introduce the concept of change of basis. Throughout the course, when we write vectors like $\begin{pmatrix} 1 \\ 3 \end{pmatrix}$ or matrices like $\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$, we know how to interpret them geometrically. You know how to relate the vector $\begin{pmatrix} 1 \\ 3 \end{pmatrix}$ to the correct point in the 2D plane. The reason for this is that we are implicitly using the same coordinate system – we all agree on the orientation of the basis vectors and their length.

Other coordinate systems are also possible. For example, think of Alice Through the Looking Glass. After Alice walks through the mirror, the left-right axis (or $\hat{\imath}$) is flipped:



Alice uses a different coordinate system to describe the same vector. For her, getting to the point $\bar{v}$ entails going 1 step **opposite** to $\hat{\imath}$ and 3 step in the direction of $\hat{\jmath}$.
We can think of coordinate systems as different languages. In the case of Alice's coordinate system, the translation between languages is easy:

$$\bar{u} = \begin{pmatrix}u_1\\u_2\end{pmatrix} \rightarrow \quad \bar{u}_{Alice} = \begin{pmatrix}-u_1\\u_2\end{pmatrix}$$

Spoiler alert: Choosing some arbitrary basis seems to be a silly thing to do. Soon we'll see that moving to a new basis can have many advantages. The basis that is defined by the eigenvectors of a matrix is especially useful.

Definition: A basis for a space is a sequence of vectors $\bar{e}_1, \bar{e}_2, \ldots, \bar{e}_d$ with two properties:

- They are independent (there aren't too many vectors in the basis)
- They span the space (there aren't too few vectors in the basis)

In other words, a basis is a linearly independent spanning set of vectors.

Note that given a basis for some space, the number of vectors in the basis is the dimension of the space.

Definition (reminder): The space spanned by the vectors $\bar{e}_1, \bar{e}_2, \ldots, \bar{e}_d$ consists of all linear combinations of these vectors.

2.2 Changing bases using matrices

2.2.1 The change of basis matrix

Let's take some arbitrary basis, say: $\hat{e}_1 = \begin{pmatrix}3\\1\end{pmatrix}, \hat{e}_2 = \begin{pmatrix}-1\\2\end{pmatrix}$. Notice that these two new basis vectors are written **in our standard basis**, in the language we are used to.

Now, let $\bar{v} = \begin{pmatrix}1\\2\end{pmatrix}_{new}$ be a vector in this new coordinate system. Notice that we write "new" in the subscript to remind ourselves that these coordinates are relative to the new coordinate system.

How is this vector represented in our standard coordinate system?

The entries of $\bar{v}$ tells us that we have to take 1 step in the direction of $\hat{e}_1$ and 2 steps in the direction of $\hat{e}_2$:

$$\bar{v} = 1\begin{pmatrix}3\\1\end{pmatrix} + 2\begin{pmatrix}-1\\2\end{pmatrix} = \begin{pmatrix}3-2\\1+4\end{pmatrix} = \begin{pmatrix}1\\5\end{pmatrix}$$

We can make this translation process even easier. When we first introduced matrices, we said that we can think of $A\bar{x}$ as a linear combination of the columns of A. For example:

$$A = \begin{pmatrix}a & b\\c & d\end{pmatrix}, \bar{x} = \begin{pmatrix}x_1\\x_2\end{pmatrix} \rightarrow \quad A\bar{x} = x_1\begin{pmatrix}a\\c\end{pmatrix} + x_2\begin{pmatrix}b\\d\end{pmatrix}$$

So in fact, we can construct a "change of basis" matrix that translates from the new basis to the standard basis:

$$P = \begin{pmatrix}3 & -1\\1 & 2\end{pmatrix}$$

The change of basis matrix is the matrix whose columns are the new basis vectors.

Given any vector in the new basis, we can write it down in our standard basis like this:

$$\bar{v} = P\bar{v}_{new}$$

Notice that P is written using **our own coordinates**[2].

This matrix P is called the change of basis matrix. It tells us that the vector $\hat{\imath}_{new}$ is mapped to $\begin{pmatrix}3\\1\end{pmatrix}$ and the vector $\hat{\jmath}_{new}$ is mapped to $\begin{pmatrix}-1\\2\end{pmatrix}$.

Question: Does this feel unintuitive?
If it does, it might be helpful to use the following metaphor. P is like a dictionary that translates words from a foreign language (i.e., vectors written in a different basis) to our own language. In order for us to understand it, the dictionary itself has to be in our own language.

2.2.2 The inverse of the change of basis matrix
Now we know how to change basis from the new basis to our own basis.
What matrix can we use in order to change bases in the other direction, from our basis to the new basis?
The answer is simple: it's P^{-1}, the inverse of P. Can you see why?

2.3 Change of basis of a matrix
We know how to transform a vector from one basis to the other.
How can we transform an entire matrix from one basis to the other?

We will start with what we already know – transforming vectors.
Let's say we have a vector $\bar{v}_{new}$ in the new basis, and a matrix A in the old, standard basis. Instead of transforming the matrix A to a new basis, we can transform $\bar{v}_{new}$ to the standard basis:

$$P\bar{v}_{new}$$

[2] If it were written in the new basis, it would be $P_{new} = \begin{pmatrix}1 & 0\\0 & 1\end{pmatrix}$, because each column is just the corresponding (new) basis vector.

Now we can apply the matrix A as usual:

$$AP\bar{v}_{new}$$

But since we started with a vector in the new basis, we want the output vector to also be in the new basis. For this, we simply have to multiply the result by P^{-1}:

$$\underbrace{P^{-1}AP}_{\substack{A\ in\ the \\ new\ basis}}\ \bar{v}_{new}$$

But notice that this is simply the composite matrix $P^{-1}AP$ multiplying the vector $\bar{v}_{new}$. After seeing this, we can infer that the representation of a matrix A in a new basis defined by the columns of P is:

$$A_{new} = P^{-1}AP$$

A and A_{new} are called "similar matrices". **Similar matrices represent the same linear transformation in different bases.**

Concrete example

For a concrete example, let's use the new basis from 2.2.1. The basis vectors were: $\hat{e}_1 = \begin{pmatrix}3\\1\end{pmatrix}, \hat{e}_2 = \begin{pmatrix}-1\\2\end{pmatrix}$, the change of basis matrix was $P = \begin{pmatrix}3 & -1\\1 & 2\end{pmatrix}$, and the vector we looked at was $\bar{v}_{new} = \begin{pmatrix}1\\2\end{pmatrix}_{new}$. Let $A = \begin{pmatrix}-1 & 0\\0 & 2\end{pmatrix}$ be the matrix that we wish to apply to $\bar{v}_{new}$. We cannot take $A\bar{v}_{new}$ because A and $\bar{v}_{new}$ are not expressed in the same coordinate systems. So following the logic from above, we take:

$$\bar{v} = P\bar{v}_{new} = \begin{pmatrix}3 & -1\\1 & 2\end{pmatrix}\begin{pmatrix}1\\2\end{pmatrix} = \begin{pmatrix}1\\5\end{pmatrix}$$

$$A\bar{v} = AP\bar{v}_{new} = \begin{pmatrix}-1 & 0\\0 & 2\end{pmatrix}\begin{pmatrix}1\\5\end{pmatrix} = \begin{pmatrix}-1\\10\end{pmatrix}$$

To go back to the new basis, we need P^{-1}, which is[3]:

$$P^{-1} = \frac{1}{7}\begin{pmatrix}2 & 1\\-1 & 3\end{pmatrix}$$

So:

$$P^{-1}AP\bar{v}_{new} = \frac{1}{7}\begin{pmatrix}2 & 1\\-1 & 3\end{pmatrix}\begin{pmatrix}-1\\10\end{pmatrix} = \frac{1}{7}\begin{pmatrix}8\\31\end{pmatrix}$$

So A in the new basis is actually:

$$A_{new} = P^{-1}AP = \frac{1}{7}\begin{pmatrix}2 & 1\\-1 & 3\end{pmatrix}\begin{pmatrix}-1 & 0\\0 & 2\end{pmatrix}\begin{pmatrix}3 & -1\\1 & 2\end{pmatrix}$$

3. Matrix diagonalization[4]

3.1 The matrix in its eigenbasis

Given a matrix A, it seems like we have no reason to transform it to a different basis. But in fact, one basis will prove very useful –the eigenbasis. If we choose the eigenvectors as the basis vectors, various computation becomes much easier.

[3] Remember that for a 2x2 matrix

$$P^{-1} = \frac{1}{\det(P)}\begin{pmatrix}d & -b\\-c & a\end{pmatrix}$$

[4] This section is based on lecture 22 from Strang's lectures.

Let A be some matrix with n independent eigenvectors $\{\bar{v}_i\}_{i=1}^n$ and their associated eigenvalues $\{\lambda_i\}_{i=1}^n$. Let P be a matrix of the eigenvectors (this is the change of basis matrix from the eigenbasis to the standard basis):

$$P = \begin{pmatrix} | & | & | \\ \bar{v}_1 & \dots & \bar{v}_n \\ | & | & | \end{pmatrix}$$

We will show that:

$$P^{-1}AP = \Lambda$$

where $\Lambda = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \dots & 0 \\ 0 & 0 & \lambda_n \end{pmatrix}$ is a matrix with the corresponding eigenvalues on the diagonal.

Each column of P is an eigenvector of A. Now:

$$AP = A\begin{pmatrix} | & | & | \\ \bar{v}_1 & \dots & \bar{v}_n \\ | & | & | \end{pmatrix}$$

Since $\bar{v}_i$ are eigenvectors, we have $A\bar{v}_i = \lambda_i\bar{v}_i$

$$\begin{pmatrix} | & | & | \\ \lambda_1\bar{v}_1 & \dots & \lambda_n\bar{v}_n \\ | & | & | \end{pmatrix}$$

We can rewrite this as matrix multiplication[5]:

$$= \begin{pmatrix} | & | & | \\ \bar{v}_1 & \dots & \bar{v}_n \\ | & | & | \end{pmatrix}\begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \dots & 0 \\ 0 & 0 & \lambda_n \end{pmatrix} = P\Lambda$$

So we have:

$$AP = P\Lambda$$

Now multiply by P^{-1} from the left to get:

$$P^{-1}AP = \Lambda$$

What does this mean? **When we represent the matrix A in its eigenbasis, it is diagonal**. Any entry off the diagonal is 0. This makes perfect sense – the first column of a matrix tells us what happens to the first basis vector. In this case, the first basis vector is $\bar{v}_1$, and what A does is simply to multiply it by λ_1.

Representing a matrix in its eigenbasis is called diagonalizing the matrix.

Notice that we could equally get:

$$A = P\Lambda P^{-1}$$

3.2 Condition for diagonalizability

Notice a very important point: by using P^{-1} we assumed that P is invertible.

Question: What are the conditions for which a matrix P is invertible?
Answer: One sufficient condition was that the columns of P must be linearly independent.

[5] Remember this from Exercise 02? If you want to multiply each column of A by a different number λ_i, you have to multiply it from the right: $A\begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \dots & 0 \\ 0 & 0 & \lambda_n \end{pmatrix}$

A matrix is diagonalizable iff it has n independent eigenvectors.

Based on one of the theorems above - *A sufficient condition for being diagonalizable:*
If $A_{n\times n}$ has n different eigenvalues, it is sure to have n independent eigenvectors, and is therefore diagonalizable.

3.3 What is diagonalization good for?

What's special about a diagonal matrix? Think about some non-diagonal matrix, like $\begin{pmatrix}1 & 2\\ 0 & 1\end{pmatrix}$. This matrix represent a shear. While it keeps $\hat{\imath}$ in place, it takes any vector with a $\hat{\jmath}$ components, and adds to it some component along the $\hat{\imath}$ direction. In other words, it "mixes" the coordinates of the vector: having some x-component before the transformation results in having some y-component after the transformation.

This is not the case for a diagonal matrix. In a diagonal matrix coordinates never get "mixed". The directions are independent of each other.

Now, think about a system of linear equations. Wouldn't it be great if the coefficient matrix was diagonal? Solving the system then would be very easy, as every unknown is already isolated:

$$\begin{pmatrix}\lambda_1 & 0 & 0 & 0\\ 0 & \lambda_2 & 0 & 0\\ 0 & 0 & \dots & 0\\ 0 & 0 & 0 & \lambda_n\end{pmatrix}\begin{pmatrix}x_1\\ x_2\\ \dots\\ x_n\end{pmatrix}=\begin{pmatrix}b_1\\ b_2\\ \dots\\ b_n\end{pmatrix}$$

3.3.1 Powers of A

Changing to the eigenbasis is very useful when we want to calculate the power of a matrix. For example, let $A=\begin{pmatrix}1 & 3\\ 2 & 5\end{pmatrix}$. What is A^{100}?
In the standard basis, to find A^{100} we have to multiply A by itself a hundred times:

$$A^{100}=A\cdot A\cdot A\cdot \dots \cdot A$$

Instead, we can move to the eigenbasis, calculate the power there, and return to our old basis.

For example, finding A^2 is simple when we diagonalize the matrix:

$$A^2=AA=(P\Lambda P^{-1})(P\Lambda P^{-1})=P\Lambda\Lambda P^{-1}=P\Lambda^2P^{-1}$$

And in general, raising A to any power k:

$$A^k=P\Lambda^kP^{-1}$$

This is easy to calculate because:

$$\Lambda^k=\begin{pmatrix}\lambda_1 & 0 & 0 & 0\\ 0 & \lambda_2 & 0 & 0\\ 0 & 0 & \dots & 0\\ 0 & 0 & 0 & \lambda_n\end{pmatrix}^k=\begin{pmatrix}\lambda_1^k & 0 & 0 & 0\\ 0 & \lambda_2^k & 0 & 0\\ 0 & 0 & \dots & 0\\ 0 & 0 & 0 & \lambda_n^k\end{pmatrix}$$

3.3.2 Representing vectors in the eigenbasis

So far we've seen that moving to the new eigenbasis can be useful. In fact, we can enjoy the benefits of the eigenbasis even without changing the basis. Instead, we can use the eigenbasis differently, as we show next.

When we have n independent eigenvectors $\{\bar{v}_i\}_{i=1}^n$ for some matrix A, they form a **basis** for the vector space. This means we can represent any vector as a unique linear combination of $\bar{v}_1, \bar{v}_2, \dots, \bar{v}_n$. For any $\bar{u} \in \mathbb{R}^n$:

$$\bar{u} = a_1\bar{v}_1 + a_2\bar{v}_2 + \cdots + a_n\bar{v}_n = \sum_{i=1}^{n} a_i\bar{v}_i$$

Calculating $A\bar{u}$ should be very easy now:

$$A\bar{u} = A\left(\sum_{i=1}^{n} a_i\bar{v}_i\right) = \sum_{i=1}^{n} a_iA\bar{v}_i = \sum_{i=1}^{n} a_i\lambda_i\bar{v}_i$$

Again we see this "no-mixing coordinates" effect – What A does is simply scale each coefficient in the linear combination of the eigenvectors by the relevant eigenvalue.

3.4 Matrix functions (specifically, matrix exponential)

This is a good time to talk about matrix functions. A matrix function takes as input a square $n \times n$ matrix and gives as output a square $n \times n$ matrix.

For example, $f(A) = 2A^{17} + \frac{1}{2}A^2 + I$ is a matrix polynomial. Of course, you see that if A is diagonalizable, then it might be easier to calculate such a matrix polynomial.

Some of you already learned that some functions can be represented as an infinite sum of power terms (in fact, those that are indefinitely differentiable, meaning the derivative $\frac{df(x)}{dx}$, $\frac{d^2(x)}{dx^2}$ etc. are well-defined). For example: $\sin(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots$. This infinite sum is called the Taylor expansion of the function.

For a general function, we use the next theorem:

Theorem: If a function $f(x)$ has a Taylor expansion ($f(x) = f(0) + f'(0)\cdot x + f''(0)\cdot\frac{x^2}{2!} + \cdots$), then the matrix function $f(A)$ is defines by substituting x by A.
Note that this definition of $f(A)$ is true even if the matrix is not diagonalizable.

Example: $\sin(A) = \left(A - \frac{A^3}{3!} + \frac{A^5}{5!} - \frac{A^7}{7!} + \cdots\right)$

3.4.1 Matrix function used on an eigenvector

While thinking about actually calculating so many terms sounds like a terrible idea, things become easier when we think about eigenvectors.
Let $\bar{v}$ be an eigenvector of A with eigenvalue λ. Then for a matrix function $f(A)$:

$$\underbrace{f(A)}_{matrix}\ \bar{v} = \underbrace{f(\lambda)}_{number}\ \bar{v}$$

Question: Can you see why this is true? (Hint: think about the Taylor expansion)

This fact is easier to see for the case of diagonal matrices, as we show below.

3.4.2 Matrix function of diagonal matrices

Let A be a diagonalizable matrix with $A = P\Lambda P^{-1}$. In this case, applying the power series definition we find that:

$$f(A) = P\begin{pmatrix} f(\lambda_1) & \dots & 0 \\ \dots & \dots & \dots \\ 0 & \dots & f(\lambda_n) \end{pmatrix}P^{-1}$$

From this we can see that if A had eigenvalues $\lambda_1, \lambda_2, \dots \lambda_n$, then the eigenvalues of $f(A)$ are $f(\lambda_i)$.

3.4.3 The matrix exponential

One specific example that tends to appears in real life applications is the matrix exponential, such as e^{At} (where t is a parameter that denotes time). It often appears when we solve a system of linear differential equations.

The Taylor expansion of the exponential is $e^x = \sum_{k=0}^{\infty} \frac{1}{k!}x^k$.

In the calculus course you've seen that for a time-dependent variable $y(t)$, the 1D equation:

$$\dot{y} = ay$$

Has the following solution[6]:

$$y(t) = e^{at}y(0)$$

(you can check this: put this $y(t)$ back in the equation and make sure that the equation holds).

Now imagine that y is no longer a variable, but a vector of variables, all related to each other by a set of coupled differential equations[7]:

$$\dot{\bar{y}} = A\bar{y}$$

For example:

$$\begin{pmatrix} \frac{dy_1}{dt} \\ \frac{dy_2}{dt} \\ \frac{dy_3}{dt} \end{pmatrix} = \begin{pmatrix} -1 & 2 & 2 \\ 3 & -2 & 4 \\ 0 & 1 & -1 \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$$

Solving this system means finding three functions, $y_1(t), y_2(t), y_3(t)$ that will satisfy these three equations. In this example you can see that as each variable increases, it tends to lower its own rate of change in time. In contrast, it tends to increase the rate of change of the other variables.

---

[6] Note that the unknown here is actually a function. We are looking for a function that satisfies this relation. The number $y(0)$ is called "the initial condition", and is usually given.

[7] "Coupled" simply means that each variable equation depends on the rest of the variables.

Back to the general case. Let's take some inspiration from the 1D case, and "guess" that the solution is given by:

$$\bar{y}(t) = e^{At}\bar{y}(0)$$

Or written with the Taylor expantion:

$$\bar{y}(t) = \left(I + At + \frac{1}{2}(At)^2 + \frac{1}{3!}(At)^3 + \cdots\right)\bar{y}(0)$$

First: To verify that it is indeed the solution, we have to plug it back in the original equation (the left-hand side), and see if we get what we expected (the right-hand side). We now compute $\frac{d}{dt}\bar{y}(t)$ and hope that we get $A\bar{y}$:

$$\frac{d}{dt}\bar{y}(t) = \frac{d}{dt}\left(e^{At}\bar{y}(0)\right) = \left(A + A^2t + \frac{1}{2!}A^3t^2 + \cdots\right)\bar{y}(0)$$

$$= A\left(I + At + \frac{1}{2}(At)^2 + \frac{1}{3!}(At)^3 + \cdots\right)\bar{y}(0) = A\underbrace{e^{At}\bar{y}(0)}_{\bar{y}(t)} = A\bar{y}(t)$$

And indeed, we got the right-hand side of the equation. This tells us that our guess was correct, and that this really is a solution to the system.

As we know, actually calculating e^{At} may be hard. It might be easier to use the eigenbasis:

$$e^{At} = Pe^{\Lambda t}P^{-1} = P\begin{pmatrix} e^{\lambda_1 t} & \dots & 0 \\ \dots & \dots & \dots \\ 0 & \dots & e^{\lambda_n t} \end{pmatrix}P^{-1}\text{t}$$

In this new basis the equations are no longer coupled.

Finally, if we are given the initial vector $\bar{y}(0)$, we can use it to find a simple form for the solution at any time point t. First, we have to describe the initial vector $\bar{y}(0)$ as a linear combination of the eigenvectors $\bar{u}_1, \bar{u}_2, \dots, \bar{u}_n$:

$$\bar{y}(0) = c_1\bar{u}_1 + c_2\bar{u}_2 + \cdots + c_n\bar{u}_n$$

And because $A\bar{u}_i = \lambda_i\bar{u}_i$, we know that $e^{At}\bar{u}_i = e^{\lambda_i t}\bar{u}_i$. Therefore, the solution is:
will be:

$$e^{At}\bar{y}(0) = c_1e^{\lambda_1 t}\bar{u}_1 + c_2e^{\lambda_2 t}\bar{u}_2 + \cdots + c_ne^{\lambda_n t}\bar{u}_n$$

3.5 Using the eigenbasis to solve a dynamic system

In your homework, as well as in the course "Dynamical systems and the neuron" you will see a concrete example for how the eigenbasis, and the powers of a matrix in particular, are useful for studying dynamical systems.

An interesting side note (for the interested only, we haven't talked about this in class)

In one of your homework assignments you learned about nilpotent matrices (a nilpotent matrix is a matrix for which there is a scalar k such that $A^k = 0$). For a nilpotent matrix, the infinite sum of powers in the Taylor expansion will actually become finite. For example, look at $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. You can show that $A^2 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$. Therefore, $e^{At} = I + At + 0 + 0 + 0 \dots = I + At = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}$.

Note that in this example the matrix doesn't have n independent eigenvectors, so we could not use the eigenbasis to describe any arbitrary vector.

4. Example problem – Fibonacci numbers[8]

The Fibonacci numbers are the numbers in the following series:

$$0,1,1,2,3,5,8,13\ldots$$

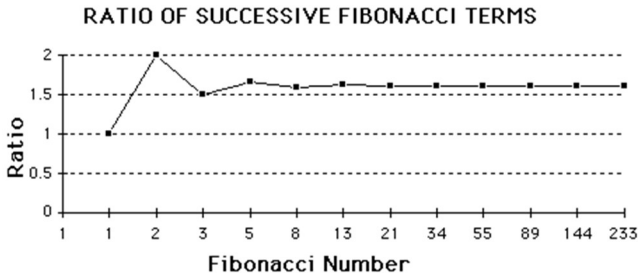
Every number is the sum of the preceding two numbers:

$$F_{k+2} = F_{k+1} + F_k$$

This is in fact a dynamical system. You can think about the index k as denoting (discrete) points in time. All the Fibonacci numbers are determined by the first two numbers in the series $F_0 = 0$ and $F_1 = 1$.

- What is F_{100}?
- Can we find a general formula for F_N?
- Can we say how fast the Fibonacci numbers are growing?

If we simply calculate the ratio between every two successive numbers we can see that the ratio is somewhat noisy in the beginning, but slowly stabilizes around 1.618…:



- Why is that?

If we want to use linear algebra to answer these questions, we would like to formulate the problem as some matrix-vector multiplication. We start by constructing the following vector:

$$\bar{u}_k = \begin{pmatrix} F_{k+1} \\ F_k \end{pmatrix}$$

Now:

$$\bar{u}_{k+1} = A\bar{u}_k$$

$$\underbrace{\begin{pmatrix} F_{k+2} \\ F_{k+1} \end{pmatrix}}_{\bar{u}_{k+1}} = A \begin{pmatrix} F_{k+1} \\ F_k \end{pmatrix}$$

What should be this matrix A?

$$\begin{pmatrix} F_{k+2} \\ F_{k+1} \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} F_{k+1} \\ F_k \end{pmatrix}$$

So we found:

[8] This example is taken from Strang's lecture 22 (at 34:30).

$$\bar{u}_{k+1} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \bar{u}_k$$

Of course:

$$\bar{u}_{100} = A^{100}\bar{u}_0$$

But calculating A^{100} seems like a horrible idea. To make it easier, we'll change to the eigenbasis of A. Let's call the eigenvectors $\bar{v}_1$ and $\bar{v}_2$. If we decompose $\bar{u}_1$ to its components in the eigenbasis, we find:

$$\bar{u}_{100} = A^{100}(c_1\bar{v}_1 + c_2\bar{v}_2)$$
$$\bar{u}_{100} = c_1\lambda_1^{100}\bar{v}_1 + c_2\lambda_2^{100}\bar{v}_2$$

So what are the eigenvalues and eigenvectors of this matrix?

$$\det(A - \lambda I) = \det\begin{pmatrix} 1-\lambda & 1 \\ 1 & -\lambda \end{pmatrix} = \lambda^2 - \lambda - 1 = 0$$

And we find that:

$$\lambda_{1,2} = \frac{1 \pm \sqrt{5}}{2}$$

You can see that $\lambda_1 \approx 1.618, \lambda_2 \approx -0.618$.

If we plug this back to the equation, we get:

$$\bar{u}_{100} \approx c_1(1.618)^{100}\bar{v}_1 + c_2 \underbrace{(-0.618)^{100}}_{\approx 0}\bar{v}_2$$

You can see that the system evolves by consecutive changes of the two different eigenvectors, $\bar{v}_1$ and $\bar{v}_2$. These are also called the modes of the system, or the eigenmodes of the system.

Back to the system. From the last equation you can see that because $|\lambda_2| < 1$, the second term vanishes. In other words, for large n, the system is governed by its first mode.

Now you can see that the Fibonacci numbers grow approximately with a factor of 1.618.

Finally, to fully solve the problem, we can find the eigenvectors. You can check and find out that these are:

$$\bar{v}_1 = \begin{pmatrix} \lambda_1 \\ 1 \end{pmatrix}, \quad \bar{v}_2 = \begin{pmatrix} \lambda_2 \\ 1 \end{pmatrix}$$

We know that $\bar{u}_0 = \begin{pmatrix} F_1 \\ F_0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$. To complete the solution, we have to represent it as a linear combination of the two eigenvectors:

$$\bar{u}_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} = c_1\begin{pmatrix} \lambda_1 \\ 1 \end{pmatrix} + c_2\begin{pmatrix} \lambda_2 \\ 1 \end{pmatrix}$$

$$\begin{pmatrix} \lambda_1 & \lambda_2 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$
$$\lambda_1 c_1 + \lambda_2 c_2 = 1$$
$$c_1 + c_2 = 0 \quad \rightarrow \quad c_2 = -c_1$$

$$\lambda_1 c_1 + \lambda_2(-c_1) = 1$$
$$c_1(\lambda_1 - \lambda_2) = 1$$

$$c_1 = \frac{1}{\lambda_1 - \lambda_2}$$

And therefore:

$$c_2 = \frac{1}{\lambda_2 - \lambda_1}$$

6. All real symmetric matrices are diagonalizable

An important class of matrices are the real symmetric matrices (i.e., $A = A^T$).

Theorem: Any real symmetric matrix is diagonalizable.

We will discuss such matrices in more detail when we talk about orthogonality.

7. Left and right eigenvectors

We haven't talked about this yet, but I will mention this at the beginning of next class.

The eigenvectors we talked about so far are also called the right eigenvectors of A, because they appear on the right of the matrix when we write $A\bar{v} = \lambda\bar{v}$.

The left eigenvectors also exist. These are the **row** vectors $\bar{w}^T$ that satisfy:

$$\bar{w}^T A = \lambda\bar{w}^T$$

(Make sure you understand why multiplying a row vector by a matrix gives another row vector. You can think about a row vector simply as a $1 \times m$ matrix).

Interestingly, just as the columns of P were the right eigenvectors of A, the rows of P^{-1} are the left eigenvectors of A.

Note: The left eigenvectors of A are actually the right eigenvectors of A^T. This can be easily shown by:

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

Taking the transpose on both sides, we get:

$$(A\bar{v})^T = \lambda\bar{v}^T$$
$$\bar{v}^T A^T = \lambda\bar{v}^T$$

So we see that the left and right eigenvectors of A actually share the same eigenvalue.

Additional resources

- Useful handouts on eigenvalues and eigenvectors from Harvard, which you can also find on Moodle: http://www.math.harvard.edu/archive/20_spring_05/handouts/ch05_notes.pdf
- A beautiful and useful visualization of the Fibonacci numbers and linear algebra: http://setosa.io/ev/eigenvectors-and-eigenvalues/
  Note also that this is a different way of thinking about the story behind Fibonacci numbers, which may be more intuitive in thinking about $\hat{\imath}, \hat{\jmath}$.

Another resource I found but haven't understood yet:

- A short and interesting paper about imaginary eigenvalues: http://vixra.org/pdf/1306.0113v1.pdf