1 Eigenvalues and eigenvectors - wtf?

In general, a linear transformation T maps any line $L \subset \mathbb{R}^n$ to some other line in $\mathbb{R}^n$. Of course, there are a bazillion different lines in $\mathbb{R}^n$, so there is no systematic way to understand T by looking at random lines. The key is to try to see if T maps any lines L to itself. Eigenvectors are vectors that are the one-element bases for such lines (and as such, must be nonzero). If T does in fact map L to itself then it must do so by scalar multiplication; it maps L to itself by the formula $T\vec{x} = \lambda\vec{x}$. This scalar λ is what an eigenvalue is. Notice that different kinds of eigenvalues have different qualitative behavior. If the eigenvalue is 1, then $T\vec{x} = 1\vec{x} = \vec{x}$, so T acts as the identity along L. If $\lambda = 0$ then $T\vec{x} = 0\vec{x} = 0$, so T is collapsing L to the origin. When $\lambda = -1$, then $T\vec{x} = -1\vec{x} = -\vec{x}$, so T flips, or reflects, the line. When $|\lambda| < 1$ then T is some kind of contraction and when $|\lambda| > 1$, T is dialating L.

Exercise 1.1. *Suppose T is a linear transformation, and λ is an eigenvalue for T. Show the set of vectors such that $T\vec{x} = \lambda\vec{x}$ is a subspace.*

Such a space is called the eigenspace of λ.

An *eigenbasis* for T is a basis $\{\vec{x}_1, ..., \vec{x}_n\}$ of $\mathbb{R}^n$ such that each $\vec{x}_i$ is an eigenvector for T. Here are the important facts:

· Many linear transformations don't have an eigenbasis!
· Linear transformations that have an eigenbasis are easy to understand. They are *diagonalizable*

2 Change of basis and eigenbasis

First we need to see why linear transformations that have an eigenbasis are easy to understand. Let $\{\vec{x}_1, ..., \vec{x}_n\}$ be a basis β and let $T : \mathbb{R}^n \to \mathbb{R}^n$ be linear transformation. Then the matrix that represents T in β coordinates is

$$[T]_\beta = \begin{bmatrix} | & & | \\ [T\vec{x}_1]_\beta & \cdots & [T\vec{x}_n]_\beta \\ | & & | \end{bmatrix}$$

Now if β is an eigenbasis, then each of the $\vec{x}_i$'s is an eigenvector with eigenvalue say λ_i. So $[T\vec{x}_i]_\beta = [\lambda_i\vec{x}_i]_\beta = \begin{bmatrix} 0 \\ \vdots \\ \lambda_i \\ \vdots \\ 0 \end{bmatrix}$ which is the vector consisting of all zeros except for the ith component which is λ_i. Plugging this in to our formula for

$[T]_\beta$ yields

$$\begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}$$

aka the diagonal matrix whose ith diagonal component is λ_i.

Exercise 2.1. *Suppose $T : \mathbb{R}^4 \to \mathbb{R}^4$ is a linear transformation and $\beta = \{\vec{v}_1, \vec{v}_2, \vec{v}_3, \vec{v}_4\}$ is an eigenbasis for T with corresponding eigenvalues $2, 4, 0$, and -1. What is $[T]_\beta$?*

solution:

$$[T]_\beta = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}.$$

3 Using an eigenbasis to find matrix representations

Suppose there are two islands, β and γ. On γ, the inhabitants keep lush groves of cheese trees and keep griddles well larded and hot for making grilled cheese sandwiches. On β however the foolish inhabitants long ago forgot how to care for their cheese trees. The orchards are overrun with thick weeds and the gnarled trees only occasionally exude sour a whey from their leaves. Their griddles lie unused, cold and rusty. Even the oldest, wrinkliest elders cannot recall how to operate them.

In short, γ has grilled cheese sandwiches, β does not. So, how can we describe the process for producing a grilled cheese sandwich on β? It would be something like
1) Sail from β to γ
2) Make a grilled cheese sandwich on γ
3) Sail back from γ to β

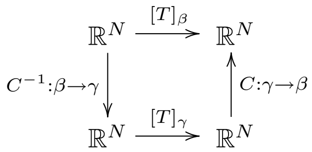
This is an analogy for our situation. Instead of being an island with grilled cheese sandwiches, γ is an eigenbasis for T. We know what $[T]_\gamma$ is. It is the diagonal matrix where the ith entry on the diagonal is the eigenvalue of the ith member of γ. So to produce $[T]_\beta \vec{x}$ we need to do something like

1) Convert the coordinates of $\vec{x}$ from β coordinates to γ coordinates
2) Apply $[T]_\gamma$
3) Convert the coordinates of this vector back from γ to β coordinates.

The change of basis matrix $C = \begin{bmatrix} | & & | \\ [\vec{v}_1]_\beta & \cdots & [\vec{v}_n]_\beta \\ | & & | \end{bmatrix}$ is the change of basis matrix converting γ coordinates to β coordinates. As matrix-functions this corresponds to:
1) Multiply by change of basis matrix, C^{-1}.
2) Multiply by $[T]_\gamma$
3) Multiply by the opposite change of basis matrix, C.

Remember that composition of functions is from right to left. So we find that $[T]_\beta = C[T]_\gamma C^{-1}$.

You may find it helpful to recall the following diagram.



Here C is the change of basis matrix from γ to β coordinates.

Exercise 3.1. *Suppose $\beta = \{\vec{v}_1, \vec{2}, \vec{v}_3\}$ is a basis for $\mathbb{R}^3$. $T : \mathbb{R}^3 \to \mathbb{R}^3$ is a linear transformation such that:*
$T(\vec{v}_1 + \vec{v}_3) = 2\vec{v}_1 + 2\vec{v}_3$
$T(\vec{v}_1 - 2\vec{v}_2 + \vec{v}_3) = -\vec{v}_1 + 2\vec{v}_2 - \vec{v}_3$
$T(-3\vec{v}_2 + \vec{v}_3) = -9\vec{v}_2 + 3\vec{v}_3$.
Find the matrix representation of T with respect to the basis β.

solution: Notice that the vectors $\vec{v}_1 + \vec{v}_3, \vec{v}_1 - 2\vec{v}_2 + \vec{v}_3, -3\vec{v}_2 + \vec{v}_3$ are eigenvectors of T with eigenvalues $2, -1$, and 3 respectively. Also they are linearly independent so they form a basis for $\mathbb{R}^3$. So lets call this basis $\beta' = \{\vec{w}_1, \vec{w}_2, \vec{w}_3\} = \{\vec{v}_1 + \vec{v}_3, \vec{v}_1 - 2\vec{v}_2 + \vec{v}_3, -3\vec{v}_2 + \vec{v}_3\}$. So β' is what we've been calling an eigenbasis.

The matrix representing T with respect to the basis β' is

$$[T]'_\beta = \begin{bmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 3 \end{bmatrix}.$$

Now for the change of basis matrix. The key is that we know the β coordinates of the β' basis elements. For instance: $[\vec{w}_1]_\beta = [\vec{v}_1 + \vec{v}_3]_\beta = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$. Then we can find the change of basis matrix that converts β' coordinates to β coordinates.

$$C = \begin{bmatrix} | & | & | \\ [\vec{w}_1]_\beta & [\vec{w}_2]_\beta & [\vec{w}_3]_\beta \\ | & | & | \end{bmatrix}$$

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

We also need the change of basis matrix that converts β coordinates to β' coordinates which is C^{-1}:

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

$$\text{Then } [T]_\beta = C[T]'_\beta C^{-1}$$

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

$$= \frac{1}{2} \begin{bmatrix} 1 & 3 & 9 \\ 12 & -2 & -24 \\ -5 & 3 & 15 \end{bmatrix}.$$

Lets do another problem, a little more geometric.

Exercise 3.2. *Let P be the plane spanned by the vectors* $\begin{bmatrix} 2 \\ 1 \\ 0 \\ 1 \end{bmatrix}$ *and* $\begin{bmatrix} 1 \\ 2 \\ 0 \\ 1 \end{bmatrix}$. *Let $T : \mathbb{R}^4 \to \mathbb{R}^4$ be projection onto P. Find the matrix that represents T.*

solution: First some relevent facts about projection matrices:

· Projections **always** have an eigenbasis. The recipe to find one is to combine a basis for the subspace V you are projecting onto and a basis for $V^\perp$, the orthogonal complement.

· The elements of this basis that lie in V are eigenvectors whose eigenvalues are 1 (fixed points), and the basis elements that came from the basis for $V^\perp$ have eigenvalue equal to 0 (they are in the nullspace).

· with respect to this basis the matrix looks like:

$$\left[\begin{array}{c|c} I & 0 \\ \hline 0 & 0 \end{array}\right],$$

where I is k×k identity matrix where k=dim(V).

OK, now to our problem. We're already given a basis for P, it is $\{ \begin{bmatrix} 2 \\ 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 0 \\ 1 \end{bmatrix} \}$, so we must just find a basis for $P^\perp$. We get a basis by solving for the nullspace of the matrix

$$\begin{bmatrix} 1 & 2 & 0 & 1 \\ 2 & 1 & 0 & 1 \end{bmatrix}.$$

The augmented matrix reduces to

$$\left[\begin{array}{cccc|c} 1 & 0 & 0 & \frac{1}{3} & 0 \\ 0 & 1 & 0 & \frac{1}{3} & 0 \end{array}\right].$$

The pivot equations yield $x_1 = -\frac{1}{3}x_4$ and $x_2 = -\frac{1}{3}x_4$. So

$$P^\perp = \{ \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} -\frac{1}{3}x_4 \\ -\frac{1}{3}x_4 \\ x_3 \\ x_4 \end{bmatrix} = x_3 \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} -\frac{1}{3} \\ -\frac{1}{3} \\ 0 \\ 1 \end{bmatrix} \}$$

And we find that a basis for $P^\perp$ is $\{ \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -\frac{1}{3} \\ -\frac{1}{3} \\ 0 \\ 1 \end{bmatrix} \}$.

Thus, an eigenbasis is

$$\beta = \{ \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -\frac{1}{3} \\ -\frac{1}{3} \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 0 \\ 1 \end{bmatrix} \}$$

with associated eigenvalues $0, 0, 1$ and 1.

The change of basis matrix from β coordinates to standard coordinates is

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

The change of basis matrix from standard coordinates to β coordinates is C^{-1} which I'm too lazy to compute. As stated before

$$[T]_\beta = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

and the answer is

$$[T]_s t = C[T]_\beta C^{-1}$$

Why Diagonalizability is equivalent to an eigenbasis

Suppose A is diagonalizable, so $A = CDC^{-1}$ for some diagonal matrix D and invertible matrix C. Then the columns of C form an eigenbasis for A.

· First off they form a basis for $\mathbb{R}^n$. Why? well they must be linearly independent since C is invertible, and there are n of them.

· They are eigenvectors. Why? The ith column of C is equal to $C \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix}$ of course.

So:

$$A(\text{ ith column of } C) = CDC^{-1}(\text{ ith column of } C)$$

$$= CDC^{-1}C \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix}$$

$$= CD \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix}$$

and since D is diagonal, the ith column of D is just $\lambda_i \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots 0 \end{bmatrix}$,

$$= C\lambda_i \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix}$$

$$= \lambda_i(\text{ ith column of } C)$$

so the ith column of C is an eigenvector of A with eigenvalue the ith diagonal component of D.

OK, now lets do the reverse. Suppose we have an eigenbasis $\{\vec{v}_1, ..., \vec{v}_n\}$ with associated eigenvalues $\lambda_1, ..., \lambda_n$. Let D be the diagonal matrix whose ith vertical component is equal to λ_i and let $C = \begin{bmatrix} | & & | \\ \vec{v}_1 & \cdots & \vec{v}_n \\ | & & | \end{bmatrix}$.

Then I claim that $C^{-1}AC = D$. To do this it is enough to check that both matrices evaluate the same way on the standard vectors $\begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix}$.

$$D \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} = \text{ ith column of D}$$

$$= \lambda_i \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix}.$$

But also

$$C^{-1}AC \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix}$$

$$= C^{-1}A \text{ ith column of C}$$

$$= C^{-1}A\vec{v}_i$$

and remember that $\vec{v}_i$ was an eigenvalue of A with eigenvalue λ_i.

$$= C^{-1}\lambda_i\vec{v}_i = \lambda_i C^{-1}\vec{v}_i$$

And since $C \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} = \vec{v}_i$, $\begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} = C^{-1}\vec{v}_i$

$$= \lambda_i \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix}$$

So we conclude that $C^{-1}AC = D$. Flipping the C's over to the other side we get that $A = CDC^{-1}$, so A is diagonalizable.

Combining bases for eigenspaces into one mammoth basis

Here is an extra topic we didn't have time to cover in section. Suppose λ_1 and λ_2 are two *distinct* eigenvalues for a linear transformation T. Then in class I said that if we take a basis for E_{λ_1} and E_{λ_2} and combine them into a super collection of vectors, then this collection is linearly independent. Lets prove this.

Suppose then that $\beta_{\lambda_1} = \{\vec{v}_1, ..., \vec{v}_n\}$ is a basis for the eigenspace E_{λ_1} and $\beta_{\lambda_2} = \{\vec{w}_1, ..., \vec{w}_m\}$ is a basis for the eigenspace E_{λ_2}. We want to show that the collection

$$\{\vec{v}_1, ..., \vec{v}_n, \vec{w}_1, ..., \vec{w}_m\}$$

is linearly independent. If

$$c_1\vec{v}_1 + \cdots + c_n\vec{v}_n + d_1\vec{w}_1 + \cdots + d_m\vec{w}_m = \vec{0}$$

for some scalars $c_1, ..., c_n, d_1, ..., d_m = 0$. We must show they are all zero. Here is what we are tempted to do. If

$$c_1\vec{v}_1 + \cdots + c_n\vec{v}_n + d_1\vec{w}_1 + \cdots + d_m\vec{w}_m = \vec{0}$$

then

$$c_1\vec{v}_1 + \cdots + c_n\vec{v}_n = -d_1\vec{w}_1 - \cdots - d_m\vec{w}_m$$

apply T to both sides and distribute T using linearity:

$$c_1T\vec{v}_1 + \cdots + c_nT\vec{v}_n = -d_1T\vec{w}_1 - \cdots - d_mT\vec{w}_m = \vec{0}$$

and now use the fact that all the $\vec{v}_i$ are eigenvectors with eigenvalue λ_1 and all the $\vec{w}_i$ are eigenvectors with eigenvalue λ_2 to conclude that

$$c_1(\lambda_1\vec{v}_1) + \cdots + c_n(\lambda_1\vec{v}_n) = -d_1(\lambda_2\vec{w}_1) - \cdots - d_m(\lambda_2\vec{w}_m)$$

Now factor out the λ_1 and the λ_2:

$$\lambda_1 c_1\vec{v}_1 + \cdots + c_n\vec{v}_n = \lambda_2(-d_1\vec{w}_1 - \cdots - d_m\vec{w}_m).$$

Now, *provided that* $c_1\vec{v}_1 + \cdots + c_n\vec{v}_n = -d_1\vec{w}_1 - \cdots - d_m\vec{w}_m$ is nonzero, we conclude that $\lambda_1 = \lambda_2$, which is a contradiction. The only other option then is for $c_1\vec{v}_1 + \cdots + c_n\vec{v}_n = -d_1\vec{w}_1 - \cdots - d_m\vec{w}_m$ to be zero. But in that case since $\{\vec{w}_1, ..., \vec{w}_m\}$ and $\{\vec{v}_1, ..., \vec{v}_m\}$ are bases, all the coefficients must be zero. That's what we needed to show, so we're done.