33A
Discussion Worksheet 4

1. *Proof.* An approach to this problem, which you should always remember you can do, is to do some explicit computation. Indeed, if we let

$$S = \begin{bmatrix} b & a-1 \\ 1-a & b \end{bmatrix}$$

Then we have the equation

$$B = S^{-1}AS$$

Remember that S was the so called change-of-basis matrix. We can think of S as taking in a vector written in $\mathscr{B}$ coordinates and outputting that same vector in $\mathscr{S}$ coordinates. Here, $\mathscr{B}$ is the basis given in the problem and $\mathscr{S}$ is the standard basis $\vec{e_1}, \vec{e_2}$. That weird swirly thing is apparently a script S.

We know S, so we can compute S^{-1}. We know A too, so from this we can determine B by multiplying everything out.

I'll include all that nonsense below, but first let me explain a more conceptual way to get at this problem. Why is $B = S^{-1}AS$ true? Well let's first look at AS. Its columns are A applied to the basis vectors in $\mathscr{B}$. I'll call those

$$\vec{v_1} = \begin{bmatrix} b \\ 1-a \end{bmatrix}$$
$$\vec{v_2} = \begin{bmatrix} a-1 \\ b \end{bmatrix}$$

Then the first column of AS is $A\vec{v_1}$ and the second is $A\vec{v_2}$. In other words, the columns are $T(\vec{v_1})$ and $T(\vec{v_2})$. The columns of $S^{-1}(AS)$ are therefore S^{-1} applied to $T(\vec{v_i})$, which rewrite $T(\vec{v_i})$ from the standard basis to $\mathscr{B}$. The point being that the columns of B are meant to be $T(\vec{v_i})$ written in the basis $\mathscr{B}$. The above method gives a concrete way that will always work, but sometimes things end up working out nicely enough that there's an easier path. For instance, let's try computing $T(\vec{v_i})$.

$$\begin{aligned} T(\vec{v_1}) &= \begin{bmatrix} a & b \\ b & -a \end{bmatrix} \begin{bmatrix} b \\ 1-a \end{bmatrix} \\ &= \begin{bmatrix} ba + b(1-a) \\ b^2 - a(1-a) \end{bmatrix} \\ &= \begin{bmatrix} b \\ a^2 + b^2 - a \end{bmatrix} \end{aligned}$$

Remember that we are assuming $a^2 + b^2 = 1$. So this becomes

$$\begin{bmatrix} b \\ 1-a \end{bmatrix}$$

which is just $\vec{v_1}$! In other words, $T(\vec{v_1}) = \vec{v_1}$. We can similarly compute that $T(\vec{v_2}) = -\vec{v_2}$. Just multiply this out in the same way and replace any instance of $a^2 + b^2$ with 1.

So $T(\vec{v_1}) = \vec{v_1}$ and $T(\vec{v_2}) = -\vec{v_2}$. What then is $\vec{v_1}$ in the basis $\mathscr{B}$? Well $\vec{v_1}$ is the first element of this basis, so it's just $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ And $\vec{v_2}$ is the second element of this basis, so the representation of $-\vec{v_2}$ in $\mathscr{B}$ is given by $\begin{bmatrix} 0 \\ -1 \end{bmatrix}$ So we conclude that

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

This is the matrix we get in the standard basis when we reflect about the x axis! And indeed, the linear transformation T is reflection about the line spanned by $\vec{v_1}$. Notice that $\vec{v_1}$ and $\vec{v_2}$ are perpindicular and unit length, just like $\vec{e_1}$ and $\vec{e_2}$.

We were lucky in this case that we were able to reason out what the representation of $T(\vec{v_i})$ was in the basis $\mathscr{B}$. In general, when we can't just see what the representation will be, we have to multiply by S^{-1} as above. So let me now include the computation of $S^{-1}AS$. Notice that we already computed $A\vec{v_1}$ and $A\vec{v_2}$ above as $A\vec{v_1} = \vec{v_1}$ and $A\vec{v_2} = -\vec{v_2}$. So we have

$$AS = \begin{bmatrix} b & 1-a \\ 1-a & -b \end{bmatrix}$$

Furthermore, we can invert 2×2 matrices pretty easily.

$$S^{-1} = \frac{1}{b^2 - (a-1)(1-a)} \begin{bmatrix} b & 1-a \\ a-1 & B \end{bmatrix} = \frac{1}{2-2a} \begin{bmatrix} b & 1-a \\ a-1 & B \end{bmatrix}$$

Note that we needed $a \neq 1$ to divide by $2 - 2a$ here. This is a good sign! You should expect that all the hypotheses given will be used somewhere. Anyways, now we just have to multiply some matrices

$$\begin{aligned} S^{-1}AS &= \frac{1}{2-2a} \begin{bmatrix} b & 1-a \\ a-1 & b \end{bmatrix} \begin{bmatrix} b & 1-a \\ 1-a & -b \end{bmatrix} \\ &= \frac{1}{2-2a} \begin{bmatrix} b^2 + (1-a)^2 & b(1-a) - b(1-a) \\ (a-1) + b(1-a) & (a-1)(1-a) - b^2 \end{bmatrix} \\ &= \frac{1}{2-2a} \begin{bmatrix} b^2 + 1 - 2a + a^2 & 0 \\ 0 & -1 + 2a - a^2 - b^2 \end{bmatrix} \end{aligned}$$

Using again the assumption $a^2 + b^2 = 1$, we rewrite this as

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

which was the same answer as our other method. $\square$

2. *Proof.* Similarly to problem 1, there are two approaches here. There's a brute force approach, and there's an approach using some more conceptual reasoning. The latter is quicker, but it's always worth remembering that you can brute force in cases like this. That way, you can proceed even without catching the underlying concepts at play.

   Let's start with the conceptual approach. When we see something like $S^{-1}AS$, we should think about change of basis. For a linear transformation T, let me denote $[T]_{\mathscr{S}}$ as the representation of T in the standard basis. Let's name the columns of S as $\vec{v_1}$, $\vec{v_2}$. Since S is invertible, these columns form a basis, which we call $\mathscr{B}$. Then S is the change of basis matrix taking us from $\mathscr{B}$ to $\mathscr{S}$. Point being that we then have the relationship

   $$S^{-1}[T]_{\mathscr{S}}S = [T]_{\mathscr{B}}$$

   That looks a lot like what we're after! This leads us to try recontextualizing the problem as searching for a basis $\mathscr{B}$. Indeed, we let T be the linear transformation which is represented in the standard basis by $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ We are then looking for a basis $\mathscr{B}$ so that

   $$[T]_{\mathscr{B}} = \begin{bmatrix} 0 & b \\ 1 & d \end{bmatrix}$$

   where b and d can be anything.

   This is an equivalent problem to the one given, as the columns of the matrix S are precisely the basis we're looking for, but this rephrasing can help us out. For instance, let's now consider what it means that the first column of $[T]_{\mathscr{B}}$ is $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ Indeed, the first column of $[T]_{\mathscr{B}}$ is meant to be given by $[T(\vec{v_1})]_{\mathscr{B}}$, meaning the representation of $T(\vec{v_1})$ in the basis $\mathscr{B}$. Well what vector does $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ represent in the basis $\mathscr{B}$? By definition, it represents $0\vec{v_1} + 1\vec{v_2}$, which is just $\vec{v_2}$. Thus, saying that the first column of $[T]_{\mathscr{B}}$ is $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ is the same as saying that $T(\vec{v_1}) = \vec{v_2}$.

   So we've now reduced this problem to finding some vector $\vec{v_1}$ so that $\vec{v_1}$ and $T(\vec{v_1})$ form a basis. As we're in a two dimensional space, it suffices to show that they're linearly independent. Now let's just try some vectors. The easiest one I can think of is $\vec{v_1} = \vec{e_1}$. Indeed, $T(\vec{e_1}) = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$ and that's not a scalar multiple of $\vec{e_1}$. So these are linearly independent, and we let our basis be $\mathscr{B} = (\vec{e_1}, T(\vec{e_1}))$. As discussed, it follows then that the first column of $[T]_{\mathscr{B}}$ is $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ So we can take

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

   Now, how about the brute force approach? Well let's start by naming the elements of

S.

$$S = \begin{bmatrix} x & y \\ z & w \end{bmatrix}$$

We can compute S^{-1} as

$$S^{-1} = \frac{1}{xw - yz} \begin{bmatrix} w & -y \\ -z & x \end{bmatrix}$$

provided $xw - yz \neq 0$. So we want to solve the matrix equation

$$\frac{1}{xw - yz} \begin{bmatrix} w & -y \\ -z & x \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} x & y \\ z & w \end{bmatrix} = \begin{bmatrix} 0 & b \\ 1 & d \end{bmatrix}$$

where b, d can be anything. Because the second column of the right hand side is arbitrary, we can simplify our computation a bit and focus on the first columns. Formally, let's multiply the first standard basis vector in both sides

$$\frac{1}{xw - yz} \begin{bmatrix} w & -y \\ -z & x \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} x & y \\ z & w \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

The point is that we can focus as best we can on only the first column, Let's start multiplying it out.

$$\begin{aligned} \frac{1}{xw - yz} \begin{bmatrix} w & -y \\ -z & x \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} x & y \\ z & w \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} &= \frac{1}{xw - yz} \begin{bmatrix} w & -y \\ -z & x \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} x \\ z \end{bmatrix} \\ &= \frac{1}{xw - yz} \begin{bmatrix} w & -y \\ -z & x \end{bmatrix} \begin{bmatrix} x + 2z \\ 3x + 4z \end{bmatrix} \\ &= \frac{1}{xw - yz} \begin{bmatrix} w(x + 2z) - y(3x + 4z) \\ -z(x + 2z) + x(3x + 4z) \end{bmatrix} \end{aligned}$$

We want to solve for x, y, z, w so that $xw - yz \neq 0$ and so that the above vector equals $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$

That's rough, but we can do it. Ultimately, we need to find some actual numbers x, y, z, w, so let's just pick the most simplifying choices we can and hope it works. x appears in almost every single term, so let's take $x = 0$. Then we can rewrite the above as $yz \neq 0$ and

$$-\frac{1}{yz} \begin{bmatrix} 2zw - 4yz \\ -2z^2 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Every term here has a z, so it'll be fruitful to get rid of it. We can't set it equal to 0, lest we make $yz = 0$, so let's try the next best thing: $z = 1$. This yields

$$-\frac{1}{y} \begin{bmatrix} 2w - 4y \\ -2 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

So the bottom component tells us that $\frac{2}{y} = 1$. We then take $y = 2$. The first component then yields the equation

$$-w + 4 = 0$$

so we take $w = 4$. Putting all of this together, we get

$$S = \begin{bmatrix} 0 & 2 \\ 1 & 4 \end{bmatrix}$$

Note that we could have gotten this in the more conceptual way too, if we took $\vec{v_1} = \vec{e_2}$ instead of $\vec{v_1} = \vec{e_1}$. $\square$

3. *Proof.* If we stare at the hypotheses on A we're given long enough, we might notice something interesting.

$$A \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 3 \\ 6 \end{bmatrix} = 3 \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$
$$A \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} -2 \\ -1 \end{bmatrix} = - \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

So, on the specific vectors we're given in the problem, A just acts by scaling. Let me denote

$$\vec{v_1} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$
$$\vec{v_2} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

Then rewriting what we said above, we get

$$A\vec{v_1} = 3\vec{v_1}$$
$$A\vec{v_2} = -\vec{v_2}$$

Let's take a step back and imagine we had a different pair of conditions, say

$$B\vec{e_1} = 3\vec{e_1}$$
$$B\vec{e_2} = -\vec{e_2}$$

In this case, we can directly determine that the first column of B is $3\vec{e_1}$ and the second is $-\vec{e_2}$. That is,

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

This is diagonal!

The conditions on A are nearly identical as those on B, with the only difference between that they are with respect to the $\vec{v_i}$ rather than the $\vec{e_i}$. This suggests to us that if we replace the standard basis $\mathscr{S} = (\vec{e_1}, \vec{e_2})$ with the basis $\mathscr{B} = (\vec{v_1}, \vec{v_2})$, we should expect to get this diagonal matrix. But one thing to note, is $\mathscr{B}$ a basis? Check that $\vec{v_1}$ and $\vec{v_2}$ are linearly independent!

Now, why am I talking about bases at all? Well we want to show that A is similar to a diagonal matrix D. Being similar means that there is some invertible S so that

$$S^{-1}AS = D$$

which looks a lot like changing basis! Indeed, let me now change language to linear transformations. Let's say T is the linear map $\mathbb{R}^2 \longrightarrow \mathbb{R}^2$ given by A in the standard basis. That is, $[T]_{\mathscr{S}} = A$. Then $D = [T]_{\mathscr{B}}$ where $\mathscr{B}$ consists of the columns of S.

So let's indeed take

$$S = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$$

so that the columns of S are $\vec{v_1}$, $\vec{v_2}$. Now, let's consider $[T]_{\mathscr{B}}$. Its columns are given by $T(\vec{v_i})$ written in $\mathscr{B}$. So how do we write, say, $T(\vec{v_1})$ in the basis $\mathscr{B}$? Well $T(\vec{v_1}) = A\vec{v_1} = 3\vec{v_1}$. And $3\vec{v_1}$ is represented in $\mathscr{B}$ as $\begin{bmatrix} 3 \\ 0 \end{bmatrix}$ Similarly, $T(\vec{v_2}) = -\vec{v_2}$ is represented in $\mathscr{B}$ by $\begin{bmatrix} 0 \\ -1 \end{bmatrix}$ Thus, we determine

$$D = [T]_{\mathscr{B}} = \begin{bmatrix} 3 & 0 \\ 0 & -1 \end{bmatrix}$$

and

$$D = S^{-1}AS$$

for

$$S = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$$

□

4. *Proof.* Let's let

$$S = \begin{bmatrix} 1 & a \\ 0 & c \end{bmatrix}$$

so that its columns are the basis vectors we're given. We are given the matrix for T in the standard basis $\mathscr{S}$. Let's call let $\mathscr{B}$ be the basis given by the columns of S. Then we have

$$S^{-1}[T]_{\mathscr{S}}S = [T]_{\mathscr{B}}$$

and we know that

$$[T]_{\mathscr{S}} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

So we have

$$\begin{bmatrix} 1 & a \\ 0 & c \end{bmatrix}^{-1} \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 1 & a \\ 0 & c \end{bmatrix} = [T]_{\mathscr{B}}$$

Our goal is the matrix on the right hand side, we just need need to bash out the left hand side.

$$\begin{aligned}
\begin{bmatrix} 1 & a \\ 0 & c \end{bmatrix}^{-1} \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 1 & a \\ 0 & c \end{bmatrix} &= \frac{1}{c} \begin{bmatrix} c & -a \\ 0 & 1 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 1 & a \\ 0 & c \end{bmatrix} \\
&= \frac{1}{c} \begin{bmatrix} c & -a \\ 0 & 1 \end{bmatrix} \begin{bmatrix} a & a^2 + bc \\ c & ac + cd \end{bmatrix} \\
&= \frac{1}{c} \begin{bmatrix} 0 & c(a^2 + bc) - a(ac + cd) \\ c & ac + cd \end{bmatrix} \\
&= \begin{bmatrix} 0 & a^2 + bc - a^2 - ad \\ 1 & a + d \end{bmatrix} \\
&= \begin{bmatrix} 0 & bc - ad \\ 1 & a + d \end{bmatrix}
\end{aligned}$$

Notice that the first column is $\vec{e_2}$, much like in problem 2. This is for the same reason too, we apply T to the first basis vector $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and get the second basis vector $\begin{bmatrix} a \\ c \end{bmatrix}$ □

5. *Proof.* First off, a matrix is upper triangular if it looks like

$$\begin{bmatrix} * & * & \dots & \dots & * \\ 0 & * & \dots & \dots & * \\ 0 & 0 & * & \dots & \vdots \\ \vdots & & \ddots & & \vdots \\ 0 & \dots & \dots & 0 & * \end{bmatrix}$$

In algebraic terms, a matrix A is upper triangular if $A_{ij} = 0$ for any $i > j$. Visually, anything below the main diagonal is 0. For a 2×2 matrix, this means that it's of the form

$$\begin{bmatrix} * & * \\ 0 & * \end{bmatrix}$$

So let's suppose T had a basis $\mathscr{B} = (\vec{v_1}, \vec{v_2})$ for which its matrix representation $[T]_{\mathscr{B}}$ was upper triangular. That is,

$$[T]_{\mathscr{B}} = \begin{bmatrix} a & b \\ 0 & c \end{bmatrix}$$

This would imply that $T(\vec{v_1})$ is represented in $\mathscr{B}$ as $\begin{bmatrix} a \\ 0 \end{bmatrix}$ Now, $\begin{bmatrix} a \\ 0 \end{bmatrix}$ represents $a\vec{v_1} + 0\vec{v_2} = a\vec{v_1}$. Hence, this condition would imply that $T(\vec{v_1}) = avecv_1$.

Ok, let's unpack what that means. This means that there is some vector $\vec{v_1}$ in the plane so that T acts on $\vec{v_1}$ via scaling it by a factor of a. But hang on, T takes a vector $\vec{x}$ to its 90° counterclockwise rotation. Certainly it can't scale any vector right? If it did, then $\vec{v_1}$ and $T(\vec{v_1})$ would be on the same line through the origin. But that clearly can't happen for rotation right? Aha, what if $\vec{v_1} = \vec{0}$! Then this condition works, and in fact any a would work. But we need $\vec{v_1}, \vec{v_2}$ to be a basis, so certainly $\vec{v_1}$ cannot be $\vec{0}$. This tells us that it's impossible for T to have a basis which is upper triangular.

Non-essential but interesting food for though: What if we were doing linear algebra with complex numbers instead? Meaning the matrices and vectors had complex entries, and the scalars were complex numbers. Then we can use the same matrix

$$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

to define a "complex – linear" transformation $\mathbb{C}^2 \longrightarrow \mathbb{C}^2$. Does this have a basis where it's upper triangular? Diagonal? $\square$