Math 217: Change of Basis and All That...[1]

I. Coordinates.

Let V be a vector space with ordered basis $\mathcal{B} = (v_1, \ldots, v_n)$.

Every element v in V can be written *uniquely* as a linear combination of the basis elements:
$$v = a_1 v_1 + a_2 v_2 + \cdots + a_n v_n.$$
The scalars a_i's can be recorded in a column vector, called the $\mathcal{B}$-**coordinate column vector of** x:
$$\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}.$$

We use the notation $[v]_\mathcal{B}$ to denote this column vector. For short, we also call this column (or its entries $a_1, a_2, \ldots, a_n$) the "$\mathcal{B}$-coordinates" of v.

Every vector v corresponds to *exactly one* such column vector in $\mathbb{R}^n$, and vice versa. That is, for all intents and purposes, we have just **identified** the vector space V with the more familiar space $\mathbb{R}^n$.

EXAMPLE I: The vector space $\mathcal{P}_2$ of polynomials of degree at most two consists of all expressions of the form $a + bx + cx^2$. The basis $\{1, x, x^2\}$ can be used to identify $\mathcal{P}_2$ with $\mathbb{R}^3$:
$$\mathcal{P}_2 \longrightarrow \mathbb{R}^3 \qquad a + bx + cx^2 \mapsto \begin{bmatrix} a \\ b \\ c \end{bmatrix}.$$

The formal mathematical statement is that the mapping
$$V \xrightarrow{L_\mathcal{B}} \mathbb{R}^n \quad v \mapsto [v]_\mathcal{B}$$
is an *isomorphism of vector spaces.*[2] This is deeper than just matching up the elements of V and $\mathbb{R}^n$. More is true: the way we add and scalar multiply also match up:
$$[v + w]_\mathcal{B} = [v]_\mathcal{B} + [w]_\mathcal{B} \quad \text{and} \quad [cv]_\mathcal{B} = c[v]_\mathcal{B}.$$
We call this mapping the $\mathcal{B}$-**coordinate isomorphism.** Note that $L_\mathcal{B}(v)$ is another notation for $[v]_\mathcal{B}$.

Summary of I:
BY CHOOSING A BASIS, A VECTOR SPACE OF DIMENSION n CAN BE IDENTIFIED WITH $\mathbb{R}^n$.
THIS IS ONE IMPORTANT REASON BASES ARE USEFUL!

---

[1]With help from the students of Math 217, especially Christina Bolema, Suki Dasher, Beldon Lin, Stephen Lisius and Jordan Zhu (F17), Kai Xuan Shau (W19), Gabe Weber (F22), and Haytham Tang (W23).

[2]To practice defintions and easy proofs, you should recall the formal definition of *isomorphism* and verify this!

II. Comparing Coordinates in Different Bases.

What if we choose a different basis $\mathcal{A} = (u_1, \ldots, u_n)$ for V? Of course, this gives a different identification of V with $\mathbb{R}^n$, namely

$$V \xrightarrow{L_\mathcal{A}} \mathbb{R}^n \qquad v \mapsto [v]_\mathcal{A}.$$

Now there are two different ways of writing each vector x in V as a column vector, namely

$$[v]_\mathcal{A} \quad \text{and} \quad [v]_\mathcal{B}.$$

How are these columns related?

The secret is that there is an $n \times n$ matrix S such that

$$S[v]_\mathcal{B} = [v]_\mathcal{A}.$$

You should think of the matrix S as a machine that takes the $\mathcal{B}$-coordinate column of each vector v and converts it (by multiplication) into the $\mathcal{A}$-coordinate column of v. This special matrix S is called the **change of basis matrix**[3] **from $\mathcal{B}$ to $\mathcal{A}$.** We also denote it by $S_{\mathcal{B} \to \mathcal{A}}$ to emphasize that S operates on $\mathcal{B}$-coordinates to produce $\mathcal{A}$-coordinates.

But how do we find $S_{\mathcal{B} \to \mathcal{A}}$? **I recommend carefully memorizing the following definition:**

Definition II: *The change of basis matrix from $\mathcal{B}$ to $\mathcal{A}$ is the $n \times n$ matrix $S_{\mathcal{B} \to \mathcal{A}}$ whose columns are the elements of $\mathcal{B}$ expressed in $\mathcal{A}$. That is,*

$$S_{\mathcal{B} \to \mathcal{A}} = [[v_1]_\mathcal{A} \ [v_2]_\mathcal{A} \ \cdots \ [v_n]_\mathcal{A}].$$

EXAMPLE II. Another basis for $\mathcal{P}_2$ is $\mathcal{A} = (x + 1, x - 1, 2x^2)$. The change of basis matrix from $\mathcal{B} = (1, x, x^2)$ to $\mathcal{A}$ is

$$S_{\mathcal{B} \to \mathcal{A}} = \begin{bmatrix} 1/2 & 1/2 & 0 \\ -1/2 & 1/2 & 0 \\ 0 & 0 & 1/2 \end{bmatrix}.$$

Consider the polynomial $f = a + bx + cx^2 \in \mathcal{P}_2$. Then $[f]_\mathcal{B} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$ and

$$[f]_\mathcal{A} = S_{\mathcal{B} \to \mathcal{A}}[f]_\mathcal{B} = \begin{bmatrix} 1/2 & 1/2 & 0 \\ -1/2 & 1/2 & 0 \\ 0 & 0 & 1/2 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} (a+b)/2 \\ (b-a)/2 \\ c/2 \end{bmatrix}.$$

This represents the fact that f can also be written as $\frac{a+b}{2}(x+1) + \frac{(b-a)}{2}(x-1) + \frac{c}{2}(2x^2)$.

Caution: Do not confuse the roles of the bases $\mathcal{A}$ and $\mathcal{B}$ in Definition II. Of course, we could also convert from $\mathcal{A}$-coordinates to $\mathcal{B}$-coordinates in the same way, but the matrix that does this, denoted $S_{\mathcal{A} \to \mathcal{B}}$, is not

---

[3]"change of coordinates matrix" might be a better name, since it is the *coordinates* we actually change, but we stick to the book's name.

the same matrix[4] as $S_{\mathcal{B} \to \mathcal{A}}$. In general, $S_{\mathcal{B} \to \mathcal{A}}$ converts $\mathcal{B}$-coordinates to $\mathcal{A}$-coordinates and $S_{\mathcal{A} \to \mathcal{B}}$ does the opposite, so that

$$S_{\mathcal{A} \to \mathcal{B}} = S_{\mathcal{B} \to \mathcal{A}}^{-1}.$$

THE SECRET REVEALED: **Why** is the change of coordinates mapping

$$\mathbb{R}^n \longrightarrow \mathbb{R}^n \quad [v]_\mathcal{B} \mapsto [v]_\mathcal{A}$$

given by matrix multiplication? Put differently, how do we know that it is a *linear transformation*?

The point is that the change-of-coordinates map is the composition of the maps from Section I:

$$\mathbb{R}^n \xrightarrow{L_\mathcal{B}^{-1}} V \xrightarrow{L_\mathcal{A}} \mathbb{R}^n$$

$$[v]_\mathcal{B} \mapsto v \mapsto [v]_\mathcal{A}.$$

Both these identifications are isomorphisms, so their composition is also an isomorphism.[5] In particular, it is a linear transformation $\mathbb{R}^n \to \mathbb{R}^n$. As you know, every linear transformation $\mathbb{R}^n \to \mathbb{R}^n$ is given by matrix multiplication. The matrix here is the change of basis matrix $S_{\mathcal{B} \to \mathcal{A}}$.

CHECK YOUR UNDERSTANDING II: To better understand Definition II, we compute the columns of the change of basis matrix directly. Remember that $S_{\mathcal{B} \to \mathcal{A}}$ "eats" $\mathcal{B}$-coordinates and "spits out" $\mathcal{A}$-coordinates via the matrix multiplication

$$S_{\mathcal{B} \to \mathcal{A}} [v]_\mathcal{B} = [v]_\mathcal{A}.$$

So in particular, for each basis element v_j in $\mathcal{B}$, we have

$$S_{\mathcal{B} \to \mathcal{A}} [v_j]_\mathcal{B} = [v_j]_\mathcal{A}.$$

But what is $[v_j]_\mathcal{B}$ when $\mathcal{B}$ is the basis $(v_1, \ldots, v_n)$? It is, of course, the standard unit column vector $\vec{e}_j$. This says

$$S_{\mathcal{B} \to \mathcal{A}} \vec{e}_j = [v_j]_\mathcal{A}.$$

But for any matrix, including $S_{\mathcal{B} \to \mathcal{A}}$, multiplying by $\vec{e}_j$ will produce its j-th column (the "Unreasonably Useful Lemma"). This confirms that the j-th column of $S_{\mathcal{B} \to \mathcal{A}}$ is precisely v_j expressed in $\mathcal{A}$-coordinates.

Summary of II:
THE CHANGE OF BASIS MATRIX $S_{\mathcal{B} \to \mathcal{A}}$ IS THE MATRIX WHOSE j-TH COLUMN IS $[v_j]_\mathcal{A}$, WHERE v_j IS THE j-TH BASIS ELEMENT OF $\mathcal{B}$. FOR EVERY VECTOR v IN V, WE HAVE

$$S_{\mathcal{B} \to \mathcal{A}}[v]_\mathcal{B} = [v]_\mathcal{A}.$$

IN PARTICULAR, $S_{\mathcal{A} \to \mathcal{B}} = S_{\mathcal{B} \to \mathcal{A}}^{-1}$.

---

[4] When reading math, you should compute along with the text. Here: find $S_{\mathcal{A} \to \mathcal{B}}$ for the example in Example II. It is much easier to find than finding $S_{\mathcal{B} \to \mathcal{A}}$!

[5] Be sure you can prove this easy fact!

III. Using Bases to Represent Transformations.

Let V be an n-dimensional vector space, and $T : V \longrightarrow V$ a linear transformation.[6] The choice of basis $\mathcal{B}$ for V identifies both the source and target of T with $\mathbb{R}^n$. Thus T gets identified with a linear transformation $\mathbb{R}^n \to \mathbb{R}^n$, and hence with a matrix multiplication. This matrix is called the $\mathcal{B}$-**matrix of** T. You should think of it as a matrix that models the transformation T in the basis $\mathcal{B}$. A good way to understand is using the "commutative[7] diagram":

$$
\begin{array}{ccc}
V & \xrightarrow{\ T\ } & V \\
{\scriptstyle L_{\mathcal{B}}} \downarrow & & {\scriptstyle L_{\mathcal{B}}} \downarrow \\
\mathbb{R}^n & \xrightarrow{\ [T]_{\mathcal{B}}\ } & \mathbb{R}^n.
\end{array}
$$

Here, the horizontal map T is our original transformation and the map $[T]_{\mathcal{B}}$ is the "model" of our map T in the $\mathcal{B}$-coordinate world. The vertical arrows are our "modeling maps" converting vectors in V to $\mathcal{B}$-coordinates.

It is easy to write down the matrix $[T]_{\mathcal{B}}$ directly:

Definition III: *The matrix of T in the basis $\mathcal{B} = (v_1, \ldots, v_n)$ is the $n \times n$ matrix*

$$
\begin{bmatrix} T \end{bmatrix}_{\mathcal{B}} = \begin{bmatrix} [T(v_1)]_{\mathcal{B}} & [T(v_2)]_{\mathcal{B}} & \ldots & [T(v_n)]_{\mathcal{B}} \end{bmatrix},
$$

whose columns are the vectors $T(v_i)$ expressed in the basis $\mathcal{B}$.

This matrix is helpful for computation. Take any $v \in V$. To compute $T(v)$, we simply convert v to the column vector $[v]_{\mathcal{B}}$, then multiply by the $n \times n$ matrix $[T]_{\mathcal{B}}$. This gives a column vector which represents the element $T(v)$ in the basis $\mathcal{B}$. That is:

(1) $$[T(v)]_{\mathcal{B}} = [T]_{\mathcal{B}}[v]_{\mathcal{B}}.$$

EXAMPLE III: Consider the linear transformation $D : \mathcal{P}_2 \to \mathcal{P}_2$ that sends f to $\frac{df}{dx}$. The matrix of D in the basis $\mathcal{B} = (1, x, x^2)$ is

$$
[D]_{\mathcal{B}} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix}.
$$

So differentiation is identified with the matrix multiplication:

$$
[D(a + bx + cx^2)]_{\mathcal{B}} = [D]_{\mathcal{B}}[a + bx + cx^2]_{\mathcal{B}} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} b \\ 2c \\ 0 \end{bmatrix}.
$$

This represents the fact that $\frac{d}{dx}(a + bx + cx^2) = b + 2cx$.

CHECK YOUR UNDERSTANDING: For $\begin{bmatrix} T \end{bmatrix}_{\mathcal{B}}$ defined as in Definition III, you should verify the formula

$$
[T(v)]_{\mathcal{B}} = [T]_{\mathcal{B}}[v]_{\mathcal{B}}
$$

---

[6]Our textbook in Math 217 only uses matrices to represent linear transformations from a vector space *to itself.* You might want to think about how these ideas can be naturally extended to any linear transformation $V \to W$.

[7]This is a "commutative diagram" sense that the traversing around the square following the arrows from top left to bottom right has the *same effect* on every vector in V.

for all vectors v in V. A good first step is to check this formula for the elements v_i in the basis $\mathcal{B}$. Do you see why knowing the formula for the basis elements implies it for all v?

Summary of III: LET $V \xrightarrow{T} V$ BE A LINEAR TRANSFORMATION. THE CHOICE OF BASIS $\mathcal{B}$ IDEN-
TIFIES BOTH THE SOURCE AND TARGET WITH $\mathbb{R}^n$, AND THEREFORE THE MAPPING T WITH MATRIX
MULTIPLICATION BY $[T]_{\mathcal{B}}$. THE MATRIX $[T]_{\mathcal{B}}$ IS EASY TO REMEMBER: ITS j-TH COLUMN IS $[T(v_j)]_{\mathcal{B}}$.

IV. Comparing Matrices of a Transformation in Different Bases.

How are the matrices of the **same** transformation in two different bases related?

Fix a linear transformation $V \xrightarrow{T} V$. The choice of a basis $\mathcal{B}$ identifies this map with the matrix multiplication

$$\mathbb{R}^n \xrightarrow{[T]_{\mathcal{B}}} \mathbb{R}^n \qquad [v]_{\mathcal{B}} \mapsto [T]_{\mathcal{B}}[v]_{\mathcal{B}}.$$

A different basis $\mathcal{A}$ gives a different representation of T:

$$\mathbb{R}^n \xrightarrow{[T]_{\mathcal{A}}} \mathbb{R}^n \qquad [v]_{\mathcal{A}} \mapsto [T]_{\mathcal{A}}[v]_{\mathcal{A}}.$$

What is the relationship between the two $n \times n$ matrices

$$[T]_{\mathcal{A}} \quad \text{and} \quad [T]_{\mathcal{B}}?$$

The next theorem answers this question. **I recommend memorizing it carefully.**

Theorem IV: Let $S_{\mathcal{B} \to \mathcal{A}}$ be the change of basis matrix from $\mathcal{B}$ to $\mathcal{A}$, and let $S_{\mathcal{A} \to \mathcal{B}}$ be the change of basis matrix from $\mathcal{A}$ to $\mathcal{B}$. Then we have a matrix equality:

(2) $$[T]_{\mathcal{B}} = S_{\mathcal{A} \to \mathcal{B}}[T]_{\mathcal{A}} S_{\mathcal{B} \to \mathcal{A}}.$$

The book writes this in the equivalent form:

(3) $$S_{\mathcal{B} \to \mathcal{A}}[T]_{\mathcal{B}} = [T]_{\mathcal{A}} S_{\mathcal{B} \to \mathcal{A}}.$$

EXAMPLE IV: We continue the example from III. Using the basis $\mathcal{A} = (x + 1, x - 1, 2x^2)$ instead of $\mathcal{B}$, the matrix of the differentiation map D is

$$[D]_{\mathcal{A}} = \begin{bmatrix} 1/2 & 1/2 & 2 \\ -1/2 & -1/2 & 2 \\ 0 & 0 & 0 \end{bmatrix}.$$

The matrices $[D]_{\mathcal{A}}$ and $[D]_{\mathcal{B}}$ are related by

$$[D]_{\mathcal{B}} = S_{\mathcal{B} \to \mathcal{A}}^{-1}[D]_{\mathcal{A}} S_{\mathcal{B} \to \mathcal{A}}$$

where $S_{\mathcal{B} \to \mathcal{A}}$ is the change of basis matrix from before. Note that $S_{\mathcal{B} \to \mathcal{A}}^{-1}$ is the change of basis matrix from $\mathcal{A}$ to $\mathcal{B}$ so its columns are easy to find:

$$S_{\mathcal{B} \to \mathcal{A}}^{-1} = \begin{bmatrix} 1 & -1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}.$$

A good way to understand Theorem IV is using the commutative diagram

$$\begin{array}{ccc} \mathbb{R}^n & \xrightarrow{[T]_\mathcal{B}} & \mathbb{R}^n. \\ {\scriptstyle S_{\mathcal{B}\to\mathcal{A}}}\downarrow & & \downarrow{\scriptstyle S_{\mathcal{B}\to\mathcal{A}}} \\ \mathbb{R}^n & \xrightarrow{[T]_\mathcal{A}} & \mathbb{R}^n. \end{array}$$

The top row of the diagram represents the linear transformation T modeled using $\mathcal{B}$-coordinates. The bottom row of the diagram represents the linear transformation T modeled using $\mathcal{A}$-coordinates. The vertical arrows convert from that $\mathcal{B}$-coordinate world to the $\mathcal{A}$-coordinate world. Do you see how this diagram represents formula (3) in Theorem IV? How can you alter this diagram to represent formula (2)?

It is helpful to play around with diagrams to represent different models and statements about them. For example, can you add in arrows to incorporate[8] the original map $T : V \to V$?

PROOF OF THEOREM IV: We need to check that the two maps

$$[T]_\mathcal{B} \qquad \text{and} \qquad S_{\mathcal{A}\to\mathcal{B}}\, [T]_\mathcal{A}\, S_{\mathcal{B}\to\mathcal{A}}$$

are the same. What does it mean that two maps are the same? It means that the have the same source and target, and they take each element of the source to the same output element in the target.

We should think of the source of $[T]_\mathcal{B}$ as the $\mathcal{B}$-coordinate space $\mathbb{R}^n$. This is also the source of $S_{\mathcal{A}\to\mathcal{B}}[T]_\mathcal{A}S_{\mathcal{B}\to\mathcal{A}}$. Let us follow what this latter map does to an arbitrary vector in $\mathbb{R}^n$.

Take an arbitrary column vector in the source of $[T]_\mathcal{B}$, that is, in the $\mathcal{B}$-coordinate space $\mathbb{R}^n$. This arbitrary column vector represents some v in V—that is, we can write it as $[v]_\mathcal{B}$ for some unique $v \in V$. Let us follow the column vector $[v]_\mathcal{B}$ through the composition

$$\mathbb{R}^n \xrightarrow{S_{\mathcal{B}\to\mathcal{A}}} \mathbb{R}^n \xrightarrow{[T]_\mathcal{A}} \mathbb{R}^n \xrightarrow{S_{\mathcal{A}\to\mathcal{B}}} \mathbb{R}^n.$$

The first arrow converts $[v]_\mathcal{B}$ to $[v]_\mathcal{A}$. The next arrow is modeling T in $\mathcal{A}$-coordinates, so it takes $[v]_\mathcal{A}$ to $[T(v)]_\mathcal{A}$. Finally, the last map converts back to $\mathcal{B}$-coordinates, so we get $[T(v)]_\mathcal{B}$. Summarizing, the composite map takes $[v]_\mathcal{B}$ to $[T(v)]_\mathcal{B}$. This is exactly what the map $[T]_\mathcal{B}$ does! So they are the same.

ALTERNATE PROOF OF THEOREM IV: We want to prove

$$S_{\mathcal{B}\to\mathcal{A}}\, [T]_\mathcal{B} = [T]_\mathcal{A}S_{\mathcal{B}\to\mathcal{A}}.$$

These are two $n \times n$ matrices we want to show are equal. We do this column by column, by multiplying each matrix by the standard unit column vector $\vec{e}_j$ (on the right).

We first compute $S_{\mathcal{B}\to\mathcal{A}}\, [T]_\mathcal{B}\vec{e}_j$. Remember that $\vec{e}_j = [v_j]_\mathcal{B}$. By definition of the $\mathcal{B}$-matrix of T,

$$[T]_\mathcal{B}\vec{e}_j = [T]_\mathcal{B}[v_j]_\mathcal{B} = [T(v_j)]_\mathcal{B}.$$

But now multiplying by the change of basis matrix gives

$$S_{\mathcal{B}\to\mathcal{A}}[T(v_j)]_\mathcal{B} = [T(v_j)]_\mathcal{A}.$$

This is the j-th column of the matrix $S_{\mathcal{B}\to\mathcal{A}}\, [T]_\mathcal{B}$.

---

[8]Check out the solutions to Worksheet 15 for a great diagram made by F22 Math 217 student Matthew Anderson!

We now compare this column to the j-th column of $[T]_{\mathcal{A}}S$, or $[T]_{\mathcal{A}}S_{\mathcal{B}\to\mathcal{A}}\vec{e}_j$. Again, since $\vec{e}_j = [v_j]_{\mathcal{B}}$,

$$S\vec{e}_j = S_{\mathcal{B}\to\mathcal{A}}[v_j]_{\mathcal{B}} = [v_j]_{\mathcal{A}},$$

by definition of the change of basis matrix S. Multiplying by the $\mathcal{A}$-matrix of T, we will get $T(v_j)$ expressed in the basis $\mathcal{A}$. That is,

$$[T]_{\mathcal{A}}S_{\mathcal{B}\to\mathcal{A}}\vec{e}_j = [T]_{\mathcal{A}}[v_j]_{\mathcal{A}} = [T(v_j)]_{\mathcal{A}}.$$

This is the j-th column of the matrix $[T]_{\mathcal{A}}S_{\mathcal{B}\to\mathcal{A}}$. We conclude the matrices $S_{\mathcal{B}\to\mathcal{A}}[T]_{\mathcal{B}} = [T]_{\mathcal{A}}S_{\mathcal{B}\to\mathcal{A}}$ must be equal, since the corresponding columns are equal. DONE!

CHECK YOUR UNDERSTANDING: It is good exercise to multiply each of the products

$$S_{\mathcal{B}\to\mathcal{A}}[T]_{\mathcal{B}} \quad \text{and} \quad [T]_{\mathcal{A}}S_{\mathcal{B}\to\mathcal{A}}.$$

by *any* column vector on the right. You should think of the column as representing some vector in the basis $\mathcal{B}$, since both $S_{\mathcal{B}\to\mathcal{A}}$ and $[T]_{\mathcal{B}}$ "act on" $\mathcal{B}$-coordinate columns. What happens when each matrix acts? [9]

Summary of Sections I, II, III, and IV on Change of Basis and All That:

LET V BE AN n-DIMENSIONAL VECTOR SPACE.

I. CHOOSING A BASIS $\mathcal{B}$, WE IDENTIFY V WITH $\mathbb{R}^n$. EACH VECTOR v IS IDENTIFIED WITH ITS COLUMN VECTOR $[v]_{\mathcal{B}}$ OF $\mathcal{B}$-COORDINATES.

II. THE TRANSFORMATION $V \xrightarrow{T} V$ GETS IDENTIFIED WITH MATRIX MULTIPLICATION BY THE $\mathcal{B}$-MATRIX. THE $\mathcal{B}$-MATRIX IS DENOTED $[T]_{\mathcal{B}}$. THE $\mathcal{B}$-MATRIX IS EASY TO FIND: ITS j-TH COLUMN IS $T(v_j)$ EXPRESSED IN $\mathcal{B}$.

III. IF WE INSTEAD CHOOSE A DIFFERENT BASIS $\mathcal{A}$, WE GET A DIFFERENT IDENTIFICATION OF V WITH $\mathbb{R}^n$. THESE IDENTIFICATIONS ARE RELATED TO EACH OTHER AS FOLLOWS:

$$S_{\mathcal{B}\to\mathcal{A}}[v]_{\mathcal{B}} = [v]_{\mathcal{A}}$$

WHERE $S_{\mathcal{B}\to\mathcal{A}}$ IS THE CHANGE OF BASIS MATRIX FROM $\mathcal{B}$ TO $\mathcal{A}$. THE MATRIX $S_{\mathcal{B}\to\mathcal{A}}$ IS EASY TO FIND: ITS j-TH COLUMN IS THE jTH ELEMENT OF $\mathcal{B}$ EXPRESSED IN $\mathcal{A}$.

IV. LIKEWISE, THE $\mathcal{A}$-MATRIX AND $\mathcal{B}$-MATRIX OF A LINEAR TRANSFORMATION T ARE RELATED BY

$$[T]_{\mathcal{B}} = S_{\mathcal{B}\to\mathcal{A}}^{-1}[T]_{\mathcal{A}}S_{\mathcal{B}\to\mathcal{A}}.$$

---

[9]Here is the answer if you got stuck: We first see what $S_{\mathcal{B}\to\mathcal{A}}[T]_{\mathcal{B}}$ does to columns by first multiplying by $[T]_{\mathcal{B}}$, then by $S_{\mathcal{B}\to\mathcal{A}}$. Since $[T]_{\mathcal{B}}$ "eats" $\mathcal{B}$-columns, we think of the column as the $\mathcal{B}$-coordinates $[x]_{\mathcal{B}}$ of some vector x in V. The whole point of the $\mathcal{B}$-matrix of T is that the matrix product $[T]_{\mathcal{B}}[x]_{\mathcal{B}}$ represents $T(x)$ again, but as a column, expressed in the basis $\mathcal{B}$, namely $[T(x)]_{\mathcal{B}}$. Multiplying now by $S_{\mathcal{B}\to\mathcal{A}}$ changes the $\mathcal{B}$-coordinates to $\mathcal{A}$-coordinates, so that $S_{\mathcal{B}\to\mathcal{A}}[T(x)]_{\mathcal{B}} = [T(x)]_{\mathcal{A}}$. Putting these together, we have

$$S_{\mathcal{B}\to\mathcal{A}}[T]_{\mathcal{B}}[x]_{\mathcal{B}} = [T(x)]_{\mathcal{A}}.$$

You should similarly think through the meaning of the matrix product $[T]_{\mathcal{A}}S_{\mathcal{B}\to\mathcal{A}}[x]_{\mathcal{B}}$: what each matrix does to columns and what it means. We get the same result:

$$[T]_{\mathcal{A}}S_{\mathcal{B}\to\mathcal{A}}[x]_{\mathcal{B}} = [T]_{\mathcal{A}}[x]_{\mathcal{A}} = [T(x)]_{\mathcal{A}}.$$

V. Choosing Bases Wisely.

Why choose one basis over another?

Depending on your problem or data, one basis may be more convenient or illuminating.

For example, working with an *orthonormal* basis is often simpler.[10] If $\mathcal{A} = (\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_n)$ is orthonormal, then it is especially easy to write vectors in the basis $\mathcal{A}$. The coordinates of $\vec{x}$ in $(\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_n)$ will be

$$[\vec{x}]_{\mathcal{A}} = \begin{bmatrix} \vec{x} \cdot \vec{u}_1 \\ \vec{x} \cdot \vec{u}_2 \\ \vdots \\ \vec{x} \cdot \vec{u}_n \end{bmatrix}.$$

By contrast, if we try to write $\vec{x}$ in a non-orthnormal basis, we might have to work hard to find the coordinates, perhaps by solving a big linear system. An orthonormal basis is also helpful for computing projections.

The change of basis matrix $S_{\mathcal{B} \to \mathcal{A}}$ from any basis $\mathcal{B} = (\vec{v}_1, \ldots, \vec{v}_n)$ to an **orthonormal** basis is likewise easy to find: the ij-th entry will be $\vec{u}_i \cdot \vec{v}_j$. And more: if we used Gram Schmidt process to get the new basis, the change of basis matrix will be upper triangular.

EXAMPLE V: Let V be the subspace of $\mathbb{R}^3$ spanned by $\vec{v}_1 = \begin{bmatrix} 3 \\ 0 \\ 4 \end{bmatrix}$ and $\vec{v}_2 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$. Using the Gram

Schmidt orthogonalization process, we find a new basis $(\vec{u}_1, \vec{u}_2) = \left(\begin{bmatrix} 3/5 \\ 0 \\ 4/5 \end{bmatrix}, \frac{1}{\sqrt{26}} \begin{bmatrix} 4/5 \\ 5 \\ -3/5 \end{bmatrix} \right)$, which is

orthonormal. The change of basis matrix from $(\vec{v}_1, \vec{v}_2)$ to $(\vec{u}_1, \vec{u}_2)$ is easy to find:

$$S = \begin{bmatrix} \vec{v}_1 \cdot \vec{u}_1 & \vec{v}_2 \cdot \vec{u}_1 \\ \vec{v}_1 \cdot \vec{u}_2 & \vec{v}_2 \cdot \vec{u}_2 \end{bmatrix} = \begin{bmatrix} 5 & 7/5 \\ 0 & \sqrt{26}/5 \end{bmatrix}.$$

We can write the vector $\vec{v} = \begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix}$, for example, as a column using $(\vec{v}_1, \vec{v}_2)$ or $(\vec{u}_1, \vec{u}_2)$. In the basis

$(\vec{v}_1, \vec{v}_2)$, this is easy: $[\vec{v}]_{(\vec{v}_1, \vec{v}_2)} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$. In the basis $(\vec{u}_1, \vec{u}_2)$, we can use

$$[\vec{v}]_{(\vec{u}_1, \vec{u}_2)} = S[\vec{v}]_{(\vec{v}_1, \vec{v}_2)} = \begin{bmatrix} 5 & 7/5 \\ 0 & \sqrt{26}/5 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 18/5 \\ -\sqrt{26}/5 \end{bmatrix}.$$

This represents the fact $\vec{v}$ can be written as $\frac{18}{5}\vec{u}_1 + \frac{-\sqrt{26}}{5}\vec{u}_2$. Check it!

---

[10]Note: orthonormality only makes sense in $\mathbb{R}^n$, or some other inner product space.

Another nice basis is an *eigenbasis.* An eigenbasis for a linear transformation $V \xrightarrow{T} V$ is one in which the matrix is **diagonal.**[11] If the matrix of T is diagonal in a basis $\mathcal{B} = (v_1, v_2, \ldots v_n)$, say

$$[T]_{\mathcal{B}} = \begin{bmatrix} \lambda_1 & 0 & 0 & \ldots & 0 \\ 0 & \lambda_2 & 0 & \ldots & 0 \\ \vdots & & \ldots & & \vdots \\ ]0 & 0 & \ldots & 0 & \lambda_n \end{bmatrix},$$

then this says exactly that $T(v_i) = \lambda_i v_i$. That is, that the basis $\mathcal{B}$ consists of **eigenvectors** for T. Geometrically, we see T is stretching (or contracting) v_i by the scalar λ_i. The directions of stretching are the *eigenvectors* and the stretching factors λ_i are *eigenvalues.*

EXAMPLE: Let $\mathbb{R}^n \xrightarrow{T} \mathbb{R}^n$ be multiplication by the $n \times n$ matrix A. Suppose that $\mathcal{B} = (\vec{v}_1, \ldots, \vec{v}_n)$ is an eigenbasis for T_A, with eigenvalues $\lambda_1, \ldots, \lambda_n$ (possibly repeated). Then

$$[T]_{\mathcal{B}} = \begin{bmatrix} \lambda_1 & 0 & 0 & \ldots & 0 \\ 0 & \lambda_2 & 0 & \ldots & 0 \\ \vdots & & \ldots & & \vdots \\ 0 & 0 & \ldots & 0 & \lambda_n \end{bmatrix}$$

is a diagonal matrix. Comparing the matrices of this transformation in the standard basis $\mathcal{A} = (\vec{e}_1, \ldots, \vec{e}_n)$ and the eigenbasis $\mathcal{B} = (\vec{v}_1, \ldots, \vec{v}_n)$, we have

$$[T]_{\mathcal{B}} = S_{\mathcal{A} \to \mathcal{B}} [T]_{\mathcal{A}} S_{\mathcal{B} \to \mathcal{A}}.$$

On the other hand, it is easy to check (do it!) that $S_{\mathcal{B} \to \mathcal{A}} = \begin{bmatrix} \vec{v}_1 & \ldots & \vec{v}_n \end{bmatrix}$ and that $S_{\mathcal{A} \to \mathcal{B}}$ is its inverse. And of course, $[T]_{\mathcal{A}} = A$ [be sure you see why!]. So this statement can be written

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

where D is the diagonal matrix formed by the eigenvalues of A and S is the matrix formed by the corresponding (linearly independent) eigenvectors. Writing an expression like this is called *diagonalizing* the matrix A. Caution: Not every matrix can be diagonalized because not even transformation has an eigenbasis!

Summary of V. DEPENDING ON THE PROBLEM WE ARE TRYING TO SOLVE, THE CHOICE OF A CONVENIENT BASIS CAN MAKE THINGS MUCH CLEARER.

ORTHONORMAL BASES (WHICH ALWAYS EXIST) ARE NICE BECAUSE IT IS EASY TO FIND THE COORDINATES OF ANY VECTOR IN AN ORTHONORMAL BASIS (BY USING DOT PRODUCT).

AN EIGENBASIS IS NICE FOR ANALYZING A GIVEN TRANSFORMATION T, BECAUSE THE MATRIX OF T IN AN EIGENBASIS WILL BE DIAGONAL. HOWEVER, SOME TRANSFORMATIONS DO NOT HAVE EIGENBASES.

VI. Another Use of the change of basis Matrix.

The change of basis matrix has another use in the special case where V is a subspace of $\mathbb{R}^d$.

---

[11]Not every linear transformation has an eigenbasis, but as you can imagine, it is much easier to understand a linear transformation that does.

Suppose V is a subspace of $\mathbb{R}^d$ with basis $\mathcal{B} = (\vec{v}_1, \ldots, \vec{v}_n)$. Each $\vec{v}_j$ is a column vector, so we can line them up into a $d \times n$ matrix $[\vec{v}_1 \quad \ldots \quad \vec{v}_n]$.

Another basis for V, say $\mathcal{A} = (\vec{u}_1, \ldots, \vec{u}_n)$, will give another $d \times n$ matrix $[\vec{u}_1 \quad \ldots \quad \vec{u}_n]$. What is the relationship between these two $d \times n$ matrices?

The following theorem gives the answer. **I recommend you memorize it.**

Theorem VI: Let S be the change of basis matrix from $\mathcal{B}$ to $\mathcal{A}$. Then we have a matrix equality
$$[\vec{v}_1 \quad \ldots \quad \vec{v}_n] = [\vec{u}_1 \quad \ldots \quad \vec{u}_n]S_{\mathcal{B} \to \mathcal{A}}.$$

On the right side of the equation, we are taking the product of a $d \times n$ and an $n \times n$ matrix.

EXAMPLE VI: We continue with Example V. The space $V \subset \mathbb{R}^3$ has basis $(\vec{v}_1, \vec{v}_2) = \left(\begin{bmatrix} 3 \\ 0 \\ 4 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right)$ and

also $(\vec{u}_1, \vec{u}_2) = \left(\begin{bmatrix} 3/5 \\ 0 \\ 4/5 \end{bmatrix}, \frac{1}{\sqrt{26}} \begin{bmatrix} 4/5 \\ 5 \\ -3/5 \end{bmatrix} \right)$. So we have a matrix equation

$$\begin{bmatrix} 3 & 1 \\ 0 & 1 \\ 4 & 1 \end{bmatrix} = \begin{bmatrix} 3/5 & \frac{4}{5\sqrt{26}} \\ 0 & \frac{5}{\sqrt{26}} \\ 4/5 & \frac{-3}{5\sqrt{26}} \end{bmatrix} \begin{bmatrix} 5 & 7/5 \\ 0 & \sqrt{26}/5 \end{bmatrix}.$$

Because the basis $(\vec{u}_1, \vec{u}_2)$ is obtained from the $(\vec{v}_1, \vec{v}_2)$ by Gram Schmidt, this is the **QR factorization**. Note that the change of basis matrix R is upper triangular. (We called this S previously, but in this context it is traditional to call it R.)

PLEASE NOTE: This is a different use of the change of basis matrix, valid only for subspaces of $\mathbb{R}^d$.

PROOF OF THEOREM VI: We use "block multiplication." The "row" $[\vec{u}_1 \quad \ldots \quad \vec{u}_n]$ times the j-th column of S is the j-th column of the product. Say the j-th column of S is $\begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{nj} \end{bmatrix}$. Then the the j-th column of the product is $a_{1j}\vec{u}_1 + a_{2j}\vec{u}_2 + \cdots + a_{nj}\vec{u}_n$. But this is exactly $\vec{v}_j$, by the definition of S! Remember that by definition, the j-th column of S is the coordinate column of $\vec{v}_j$ expressed in $\mathcal{A}$. This completes the proof. DONE!