Linear Transformations

A linear transformation is a function whose domain and target space are vector spaces and the function respects the operations of vector addition and scalar multiplication. Although we shall restrict our attention to linear transformations of the form

$$T: \mathbb{R}^n \longrightarrow \mathbb{R}^m,$$

it should be noted that linear transformations can be rather important in more general situations -

for instance, the set of all infinitely differentiable functions,

$$f : \mathbb{R} \to \mathbb{R},$$

can be endowed with the structure of a vector space so that the derivative becomes a linear transformation, from this vector space to itself.

Suppose $T : \mathbb{R}^n \to \mathbb{R}^m$ is a function. We say that T is a linear transformation provided:

1) if $\vec{v}, \vec{w} \in \mathbb{R}^n$, then

$$T(\vec{v}+\vec{w}) = T(\vec{v}) + T(\vec{w}),$$

i.e. T respects vector addition, and

2) if $r \in \mathbb{R}$ and $\vec{v} \in \mathbb{R}^n$, then

$$T(r\vec{v}) = rT(\vec{v}),$$

i.e. T respects scalar multiplication.

An alternative definition of linear transformation, when

$$T: \mathbb{R}^n \to \mathbb{R}^m,$$

combines 1) and 2) above into a single property. This definition is:

$$T: \mathbb{R}^n \to \mathbb{R}^m$$

is a linear transformation provided:

If $\vec{v}, \vec{w} \in \mathbb{R}^n$ and $r, s \in \mathbb{R}$, then

$$T(r\vec{v} + s\vec{w}) = rT(\vec{v}) + sT(\vec{w}).$$

HW #29: Suppose $T: \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation and k is a positive integer. If $r_1, \ldots, r_k \in \mathbb{R}$ and $\vec{v}_1, \ldots, \vec{v}_k \in \mathbb{R}^n$ prove that

$$T\left(\sum_{i=1}^{k} r_i \vec{v}_i\right) = \sum_{i=1}^{k} r_i T(\vec{v}_i).$$

Hint: Use induction.

Ex 1: Define $T: \mathbb{R}^2 \to \mathbb{R}^2$ via $T([x,y]) = [x+y, x-y]$. We'll show that T is a linear transformation.

To this end, let $r_1, r_2 \in \mathbb{R}$ and $[x_1, y_1], [x_2, y_2] \in \mathbb{R}^2$. Then

$$T(r_1[x_1,y_1] + r_2[x_2,y_2])$$

$$= T([r_1x_1+r_2x_2, r_1y_1+r_2y_2])$$

$$= [r_1x_1+r_2x_2+r_1y_1+r_2y_2, r_1x_1+r_2x_2-r_1y_1-r_2y_2]$$

$$= [r_1(x_1+y_1)+r_2(x_2+y_2), r_1(x_1-y_1)+r_2(x_2-y_2)]$$

$$= [r_1(x_1+y_1), r_1(x_1-y_1)] + [r_2(x_2+y_2), r_2(x_2-y_2)]$$

$$= r_1[x_1+y_1, x_1-y_1] + r_2[x_2+y_2, x_2-y_2]$$

$$= r_1T([x_1,y_1]) + r_2T([x_2,y_2])$$

Ex 2: Actually, this is a non-example.

If $T: \mathbb{R}^2 \to \mathbb{R}^2$ is defined by

$T([x,y]) = [x+1, y]$,

then T is not a linear transformation.

The easy way to see this is to

note that $T([0,0]) = [1,0] \neq [0,0]$.

That this is sufficient to imply that

T is not a linear transformation follows

from

<u>HW #30</u>: Suppose $T: \mathbb{R}^n \to \mathbb{R}^m$ is a

linear transformation. Prove that

$T(\vec{0}_n) = \vec{0}_m$.

<u>Hint</u>: If $\vec{v} \in \mathbb{R}^n$, then $0\vec{v} = \vec{0}_n$.

The following Proposition provides a

very useful and general way of

construction a linear transformation

$T: \mathbb{R}^n \longrightarrow \mathbb{R}^m$.

<u>Proposition 24</u>: Suppose A is a $m \times n$ matrix. Define

$$T: \mathbb{R}^n \to \mathbb{R}^m$$

via

$$T([x_1, \cdots, x_n]) = \left(A \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \right)^t .$$

Then T is a linear transformation.

<u>Proof</u>: Suppose $r, s \in \mathbb{R}$ and $[x_1, \cdots, x_n], [y_1, \cdots, y_n] \in \mathbb{R}^n$. Then

$$T(r[x_1, \cdots, x_n] + s[y_1, \cdots, y_n])$$

$$= T([rx_1 + sy_1, \cdots, rx_n + ry_n])$$

$$= \left(A \begin{bmatrix} rx_1 + sy_1 \\ \vdots \\ rx_n + sy_n \end{bmatrix} \right)^t$$

$$= \left(A \left(r \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} + s \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \right) \right)^t$$

$$= \left(rA \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} + sA \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \right)^t$$

$$= r \left(A \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \right)^t + s \left(A \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \right)^t$$

$$= r\,T([x_1, \dots, x_n]) + s\,T([y_1, \dots, y_n]) .$$

We complete this introductory section by proving two Propositions which delineate the limitations on linear transformations.

Proposition 25: Suppose $T: \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation and $\{\vec{v}_1, \cdots, \vec{v}_n\}$ is a basis for $\mathbb{R}^n$.

Then T is completely determined by the vectors $T(\vec{v}_1), \cdots, T(\vec{v}_n)$ in $\mathbb{R}^m$.

Proof: Suppose $\vec{w} \in \mathbb{R}^n$. Since $\{\vec{v}_1, \cdots, \vec{v}_n\}$ is a basis for $\mathbb{R}^n$, there exist unique $r_1, \cdots, r_n \in \mathbb{R}$ such that $\vec{w} = \sum_{i=1}^{n} r_i \vec{v}_i$. Then, by HW #29,

$$T(\vec{w}) = T\left(\sum_{i=1}^{n} r_i \vec{v}_i\right) = \sum_{i=1}^{n} r_i T(\vec{v}_i).$$

Thus, the vectors $T(\vec{v}_1), \cdots, T(\vec{v}_n)$ completely determine the vector $T(\vec{w})$.

<u>Proposition 26</u>: Suppose $\{\vec{v}_1, \cdots, \vec{v}_n\}$ is a basis for $\mathbb{R}^n$ and $\vec{w}_1, \cdots, \vec{w}_n \in \mathbb{R}^m$. Then there is a unique linear transformation,

$$T: \mathbb{R}^n \to \mathbb{R}^m,$$

such that $T(\vec{v}_i) = \vec{w}_i$, for $i = 1, \cdots, n$.

<u>Proof</u>: The proof of this Proposition makes strong use of the fact that every element of $\mathbb{R}^n$ can be realized as a unique linear combination of the elements in $\{\vec{v}_1, \cdots, \vec{v}_n\}$. If $\{v_1, \cdots, v_n\}$ didn't span $\mathbb{R}^n$ - then the domain of the function we construct would not be $\mathbb{R}^n$. If $\sum_{i=1}^{n} r_i \vec{v}_i$ were to equal $\sum_{i=1}^{n} s_i \vec{v}_i$,

where $r_j \neq s_j$ for some $j \in \{1, \cdots, n\}$, then the "function" we construct could be multi-valued – so not a function at all. Fortunately, every element in $\mathbb{R}^n$ can be uniquely expressed as a linear combination of $\vec{v}_1, \cdots, \vec{v}_n$.

The paragraph above implies that

$$T\left(\sum_{i=1}^{n} r_i \vec{v}_i\right) = \sum_{i=1}^{n} r_i \vec{w}_i$$

is a well-defined function

$$T : \mathbb{R}^n \longrightarrow \mathbb{R}^m.$$

Note that if $j \in \{1, \cdots, n\}$ and we let

$$r_i = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j \end{cases},$$

then $\sum_{i=1}^{n} r_i \vec{v}_i = \vec{v}_j$ and $\sum_{i=1}^{n} r_i \vec{w}_i = \vec{w}_j$

— so

$$T(\vec{v}_j) = T\left(\sum_{i=1}^{n} r_i \vec{v}_i\right) = \sum_{i=1}^{n} r_i \vec{w}_i = \vec{w}_j .$$

Thus, $T(\vec{v}_i) = \vec{w}_i$, for $i = 1, \dots, n$.

We now show that T is a linear transformation. To this end, let $a, b \in \mathbb{R}$ and note that $\sum_{i=1}^{n} r_i \vec{v}_i$ and $\sum_{i=1}^{n} s_i \vec{v}_i$ represent two generic elements of $\mathbb{R}^n$. Then

$$T\left(a \sum_{i=1}^{n} r_i \vec{v}_i + b \sum_{i=1}^{n} s_i \vec{v}_i\right)$$

$$= T\left(\sum_{i=1}^{n} (ar_i + bs_i) \vec{v}_i\right)$$

$$= \sum_{i=1}^{n} (ar_i + bs_i) \vec{w}_i$$

$$= a\sum_{i=1}^{n} r_i \vec{w}_i + b\sum_{i=1}^{n} s_i \vec{w}_i$$

$$= aT\left(\sum_{i=1}^{n} r_i \vec{v}_i\right) + bT\left(\sum_{i=1}^{n} s_i \vec{v}_i\right).$$

This shows that T is a linear transformation.

HW #31: Let

$$\mathcal{S}_2 = \{\vec{e}_1 = [1,0], \vec{e}_2 = [0,1]\}$$

denote the standard basis for $\mathbb{R}^2$.

Construct a linear transformation

$$T : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$$

such that $T(\vec{e}_1) = [1,2]$

and $T(\vec{e}_2) = [3,4]$.

ker(T) and im(T)

Suppose $T: \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation. Then there are two important subspaces associated to T, ker(T) and im(T). These are defined via:

$\ker(T) = \{\vec{v} \in \mathbb{R}^n \mid T(\vec{v}) = \vec{0}_m\}$, and

$\mathrm{im}(T) = \{\vec{w} \in \mathbb{R}^m \mid T(\vec{v}) = \vec{w} \text{ for some } \vec{v} \in \mathbb{R}^n\}$.

ker(T) denotes the kernel of T while im(T) denotes the image of T. It should be noted that many texts use range(T) to denote the image of T. My reason for not using range(T) is due to the fact that the current meaning of range of a function, i.e. the range of a function is its image, is relatively recent. In the 1960's and 70's, the range of a function frequently meant the target space or codomain of a function – and there were a number of high quality linear algebra texts written in the 1960's and 70's.

<u>Proposition #27</u>: Suppose $T: \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation, then $\ker(T)$ is a subspace of $\mathbb{R}^n$.

<u>Proof</u>: HW #30 tells us that $T(\vec{0}_n) = \vec{0}_m$,

so $\vec{0}_n \in \ker(T)$.

If $\vec{v}_1, \vec{v}_2 \in \ker(T)$, then $T(\vec{v}_i) = \vec{0}_m$ for $i = 1, 2$. Now, since T is a linear transformation

$$T(\vec{v}_1 + \vec{v}_2) = T(\vec{v}_1) + T(\vec{v}_2) = \vec{0}_m + \vec{0}_m = \vec{0}_m$$

- so $\vec{v}_1 + \vec{v}_2 \in \ker(T)$.

Finally, if $r \in \mathbb{R}$ and $\vec{v} \in \ker(T)$, then $T(\vec{v}) = \vec{0}_m$. Since T is a linear transformation

$$T(r\vec{v}) = rT(\vec{v}) = r\vec{0}_m = \vec{0}_m$$

- so $r\vec{v} \in \ker(T)$.

<u>Proposition #28</u>: Suppose $T: \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation. Then

$\text{im}(T)$ is a subspace of $\mathbb{R}^m$.

<u>Proof</u>: HW #30 tells us that

$\vec{0}_m \in \text{im}(T)$.

If $\vec{w}_1, \vec{w}_2 \in \text{im}(T)$ there exist

$\vec{v}_1, \vec{v}_2 \in \mathbb{R}^n$ such that $T(\vec{v}_i) = \vec{w}_i$, for

$i = 1, 2$. Then, since T is a linear

transformation

$$T(\vec{v}_1 + \vec{v}_2) = T(\vec{v}_1) + T(\vec{v}_2) = \vec{w}_1 + \vec{w}_2$$

so $\vec{w}_1 + \vec{w}_2 \in \text{im}(T)$.

Finally, if $r \in \mathbb{R}$ and $\vec{w} \in \text{im}(T)$,

then there is some $\vec{v} \in \mathbb{R}^n$ such that

$T(\vec{v}) = \vec{w}$. Since T is a linear

transformation

$T(r\vec{v}) = rT(\vec{v}) = r\vec{w}$

so $r\vec{w} \in \text{im}(T)$.

Proposition #29: Suppose $T: \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation and $\{\vec{v}_1, \cdots, \vec{v}_n\}$ is a basis for $\mathbb{R}^n$. Then $\text{im}(T) = \text{span}\left(\{T(\vec{v}_1), \cdots, T(\vec{v}_n)\}\right)$.

HW #32: Prove Proposition #29.

Hint: If $\vec{w} \in \text{im}(T)$, then $T(\vec{v}) = \vec{w}$ for some $\vec{v} \in \mathbb{R}^n$. Write $\vec{v} = \sum_{i=1}^{n} r_i \vec{v}_i$ for some $r_1, \cdots, r_n \in \mathbb{R}$.

HW #33: Let $T: \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ denote the linear transformation defined by

$$T([x,y,z]) = \left(\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} \right)^t.$$

Determine ker(T) and im(T) for this T.

The Dimension Theorem

While Theorem 8 lays the foundation for Linear Algebra, The Dimension Theorem lays the foundation for important applications of Linear Algebra.

<u>Theorem #11</u>: (The Dimension Theorem)

Suppose $T: \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is a linear transformation. Then

$$\dim(\ker(T)) + \dim(\operatorname{im}(T)) = n.$$

Note that $n = \dim(\operatorname{domain}(T))$.

<u>Proof</u>: Suppose $\{\vec{v}_1, \cdots, \vec{v}_k\}$ is a basis for $\ker(T)$. (Note that this set is

empty if $\ker(T) = \{\vec{0}_n\}$.) Extend $\{\vec{v}_1, \cdots, \vec{v}_k\}$ to a basis for $\mathbb{R}^n$, $\{\vec{v}_1, \cdots, \vec{v}_k, \vec{v}_{k+1}, \cdots, \vec{v}_n\}$ say. This sets things up so that the proof of the Dimension Theorem boils down to showing $\{T(\vec{v}_{k+1}), \cdots, T(\vec{v}_n)\}$ is a basis for $\mathrm{im}(T)$. (Note that if $k=n$, then $\{T(\vec{v}_{k+1}), \cdots, T(\vec{v}_n)\}$ is empty. In this case, view the span of $\{T(\vec{v}_{k+1}), \cdots, T(\vec{v}_n)\}$ as being $\{\vec{0}_m\}$.)

Suppose now that $\vec{w} \in \mathrm{im}(T)$. Proposition #29 implies that

$$\vec{w} \in \mathrm{span}\left(\{T(\vec{v}_1), \cdots, T(\vec{v}_n)\}\right).$$

Since $\{\vec{v}_1, \cdots, \vec{v}_k\}$ is a basis for $\ker(T)$, $T(\vec{v}_i) = \vec{0}_m$ for $i = 1, \cdots, k$. It follows that

$$\vec{w} \in \operatorname{span}(\{T(\vec{v}_{k+1}), \cdots, T(\vec{v}_n)\}).$$

So $\{T(\vec{v}_{k+1}), \cdots, T(\vec{v}_n)\}$ spans $\operatorname{im}(T)$.

We are left to showing $\{T(\vec{v}_{k+1}), \cdots, T(\vec{v}_n)\}$ is a linearly independent set. To this end, suppose that

$$\sum_{i=k+1}^{n} c_i T(\vec{v}_i) = \vec{0}_m.$$

Then, since T is a linear transformation,

$$T\left(\sum_{i=k+1}^{n} c_i \vec{v}_i\right) = \vec{0}_m.$$

This implies that $\sum_{i=k+1}^{n} c_i \vec{v}_i \in \ker(T)$.

Since $\{\vec{v}_1, \cdots, \vec{v}_k\}$ is a basis for $\ker(T)$, there exist $c_1, \cdots, c_k \in \mathbb{R}$ such that

$$\sum_{i=k+1}^{n} c_i \vec{v}_i = \sum_{i=1}^{k} c_i \vec{v}_i .$$

This, in turn, implies that

$$\sum_{i=1}^{k} (-c_i) \vec{v}_i + \sum_{i=k+1}^{n} c_i \vec{v}_i = \vec{0}_n .$$

Now, since $\{\vec{v}_1, \cdots, \vec{v}_n\}$ is a basis for $\mathbb{R}^n$, we must have

$$-c_1 = \cdots = -c_k = c_{k+1} = \cdots = c_n = 0 .$$

Since $c_{k+1} = \cdots = c_n = 0$, we conclude that $\{T(\vec{v}_{k+1}), \cdots, T(\vec{v}_n)\}$ is a linearly independent set.

Thus, $\{T(\vec{v}_{k+1}), \cdots, T(\vec{v}_n)\}$ is a basis for $\mathrm{im}(T)$ — and we are done.

HW #34: In the special case, when

$$T : \mathbb{R}^n \to \mathbb{R}^n$$

is a linear transformation, it makes sense to consider when $\ker(T) = \operatorname{im}(T)$ - since both $\ker(T)$ and $\operatorname{im}(T)$ are subspaces of $\mathbb{R}^n$. Use the Dimension Theorem to prove: If $\ker(T) = \operatorname{im}(T)$, then n must be an even integer.

HW #35: Construct a linear transformation,

$$T : \mathbb{R}^2 \to \mathbb{R}^2,$$

such that $\ker(T) = \operatorname{im}(T)$.

Computing ker(T) and im(T)

This section is devoted to computing ker(T) and im(T) in the following case: A is a $m \times n$ matrix and T is defined via

$$T([x_1, \cdots, x_n]) = \left(A \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \right)^t .$$

As we shall see in the not too distant future - this seemingly special case will work for any linear transformation

$$T : \mathbb{R}^n \longrightarrow \mathbb{R}^m .$$

In this section, we'll discover that

the tools for constructing ker(T) and im(T), in this special case, have already been developed.

<u>Constructing ker(T)</u>: Recall that $[x_1, \cdots, x_n] \in \ker(T)$ if and only if $T([x_1, \cdots, x_n]) = \vec{0}_m$. In our situation, this means

$$\left(A \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \right)^t = \vec{0}_m .$$

Taking the transpose of both sides of the equation above yields

$$A \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} .$$

Letting $A = [a_{ij}]$, we have

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Thus, $\ker(T)$ is just the solution space of the system of m homogeneous linear equations in the variables $x_1, \cdots, x_n$ given below.

$$\mathcal{S}: \quad \begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= 0 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= 0 \end{aligned}$$

Observe that the matrix associated to $\mathcal{S}$, $A_{\mathcal{S}}$, is

$$A_{\mathcal{S}} = \left[A \,\middle|\, \begin{matrix} 0 \\ \vdots \\ 0 \end{matrix} \right].$$

Recall that to solve $\mathcal{S}$, thus compute $\ker(T)$, we transform $A_{\mathcal{S}}$ into a matrix in RREF. In our present situation, this corresponds to transforming A into B, where B is in RREF. Thus, the solution space of $\mathcal{S}$ is the solution space of $\tilde{\mathcal{S}}$, where

$$A_{\tilde{\mathcal{S}}} = \left[B \,\middle|\, \begin{matrix} 0 \\ \vdots \\ 0 \end{matrix} \right].$$

Recall now that the dimension of the solution space of $\mathcal{S}$, or $\dim(\ker(T))$, is just the number of independent variables in the generic solution for S. Recall further, that the number of independent variables equals the number of columns of B in which no pivot occurs. Put another way, the number of independent variables in the generic solution for S is $n -$(the number of columns of B having a pivot). But the number of columns of B having a pivot equals the number of rows of B having a pivot. This, in turn, equals the number of nonzero rows in B.

Now, recall that the nonzero rows in B forms a basis for the row space of A.

The upshot of this discussion is:

$\dim(\ker(T)) = n - \dim(\text{row space of } A)$.

For future use, the dimension of the row space of A is usually called the row rank of A - so

$\dim(\ker(T)) = n - (\text{row rank of } A)$.

<u>Constructing im(T)</u>: Let $\mathcal{S}_n = \{\vec{e}_1, \cdots, \vec{e}_n\}$ denote the standard basis for $\mathbb{R}^n$, i.e.

$$\vec{e}_i = [0, \cdots, 0, \underset{\substack{\uparrow \\ i\text{th component}}}{1}, 0, \cdots 0]$$

for $i = 1, \cdots, n$. According to Proposition #29,

$$im(T) = \text{span}(\{T(\vec{e}_1), \cdots, T(\vec{e}_n)\}).$$

Note that

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{array}{c} \begin{bmatrix} 0 \\ \vdots \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \end{array} \begin{array}{l} \\ \\ \leftarrow j \\ \\ \\ \end{array} = \begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix}.$$

This implies that

$$T(\vec{e}_j) = (j\text{-th column of } A)^t$$

$$= j\text{-th row of } (A^t)$$

for $j = 1, \cdots, n$. Since the row space of A^t is just the span of its rows, where these rows are viewed as vectors in $\mathbb{R}^m$, we conclude that

$$\mathrm{im}(T) = \text{row space of } A^t.$$

This implies that

$$\dim(\mathrm{im}(T)) = \text{row rank of } A^t .$$

<u>HW #36</u>: Repeat HW #33 - using the approaches provided in this section.

We'll complete this section by proving a result which may be rather surprising.

<u>Proposition #30</u>: Suppose A is a $m \times n$ matrix. Then $\text{row rank}(A) = \text{row rank}(A^t)$.

<u>Proof</u>: Define $T: \mathbb{R}^n \to \mathbb{R}^m$ via

$$T([x_1, \cdots, x_n]) = \left(A \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \right)^t .$$

Since T is a linear transformation,

$$n = \dim(\ker(T)) + \dim(\operatorname{im}(T)),$$

according to the Dimension Theorem.

Observations made in this section are:

$$\dim(\ker(T)) = n - (\text{row rank}(A))$$

and

$$\dim(\operatorname{im}(T)) = \text{row rank}(A^t).$$

Thus, we have

$$n = n - (\text{row rank}(A)) + (\text{row rank}(A^t)).$$

This last equation yields

$$\text{row rank}(A) = \text{row rank}(A^t)$$

upon simplification.

Coordinate Systems and Matrix Representations of Linear Transformations

Special Note: Unless stated to the contrary, from this point forward the term basis will mean ordered basis. This means that reindexing will not be allowed.

Coordinate Systems: Let $\mathcal{V} = \{\vec{v}_1, \cdots, \vec{v}_n\}$ be an ordered basis for $\mathbb{R}^n$. Then, if $\vec{w} \in \mathbb{R}^n$, there exist unique $c_1, \cdots, c_n \in \mathbb{R}$ such that

$$\vec{w} = \sum_{i=1}^{n} c_i \vec{v}_i .$$

As such, $c_1, \cdots, c_n$ may be viewed as the $\mathcal{V}$-coordinates of $\vec{w}$. This is denoted by

$$\vec{w} = [c_1, \cdots, c_n]_{\mathcal{V}}$$

-where $[c_1, \cdots, c_n]_{\mathcal{V}}$ represents $\sum_{i=1}^{n} c_i \vec{v}_i$.

Note that, if $\mathcal{S}_n = \{\vec{e}_1, \cdots, \vec{e}_n\}$ denotes the standard ordered basis for $\mathbb{R}^n$, then

$$[c_1, \cdots, c_n]_{\mathcal{S}_n} = [c_1, \cdots, c_n].$$

Thus, the $\mathcal{V}$-coordinates on $\mathbb{R}^n$ are a simple generalization of the standard coordinates on $\mathbb{R}^n$.

We'll provide an example, at the

end of this section, which suggests the value in employing generalized coordinate systems.

Matrix Representations of Linear Transformations:

Suppose now that $T: \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation, $\mathcal{V} = \{\vec{v}_1, \cdots, \vec{v}_n\}$ is an ordered basis for $\mathbb{R}^n$ and $\mathcal{W} = \{\vec{w}_1, \cdots, \vec{w}_m\}$ is an ordered basis for $\mathbb{R}^m$.

If $j \in \{1, \cdots, n\}$, then there exist unique $a_{1j}, a_{2j}, \cdots, a_{mj} \in \mathbb{R}$ such that

$$T(\vec{v}_j) = \sum_{i=1}^{m} a_{ij} \vec{w}_i .$$

Put another way, $T(\vec{v}_j) = [a_{1j}, \cdots, a_{mj}]_{\mathcal{W}}$.

Since this can be done for each $j \in \{1, \cdots, n\}$, we have

$$T(\vec{v}_1) = [a_{11}, \cdots, a_{m1}]_{\mathcal{W}},$$

$$T(\vec{v}_2) = [a_{12}, \cdots, a_{m2}]_{\mathcal{W}},$$

$$\vdots$$

$$T(\vec{v}_n) = [a_{1n}, \cdots, a_{mn}]_{\mathcal{W}}.$$

Let A denote the $m \times n$ matrix whose j-th column is the transpose of the $\mathcal{W}$-coordinates of $T(\vec{v}_j)$, for $j = 1, \cdots, n$.

I.e.

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

where a_{ij} is the i-th coordinate in the

$\mathcal{W}$-coordinates for $T(\vec{v}_j)$.

<u>Claim</u>:

$$T([x_1, \cdots, x_n]_{\mathcal{V}}) = \left(A \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}\right)^t_{\mathcal{W}}.$$

<u>Proof of the Claim</u>: For each $j \in \{1, \cdots, n\}$, let B_j denote the $n \times 1$ matrix whose only nonzero entry is a 1 in the j-th row.

I.e.

$$B_j = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \begin{array}{l} \\ \\ \\ \leftarrow j\text{-th row} \\ \\ \\ \\ \end{array}.$$

Note that AB_j is the $m \times 1$ matrix - which is just the j-th column of A. Recall that the j-th column of A is the transpose of the $\mathcal{W}$-coordinates of $T(\vec{v}_j)$. Thus, taking the transpose of the j-th column of A yields the $\mathcal{W}$-coordinates of $T(\vec{v}_j)$. It follows that

$$T(\vec{v}_j) = (AB_j)^t_{\mathcal{W}} .$$

Next, note that

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} x_1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ x_2 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \cdots + \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ x_n \end{bmatrix}$$

$$= \sum_{j=1}^{n} x_j B_j .$$

So

$$\left(A\begin{bmatrix}x_1\\ \vdots\\ x_n\end{bmatrix}\right)^t_{\mathcal{W}} = \left(A\left(\sum_{j=1}^{n} x_j B_j\right)\right)^t_{\mathcal{W}}$$

$$= \left(\sum_{j=1}^{n} x_j\,(AB_j)\right)^t_{\mathcal{W}}$$

$$= \sum_{j=1}^{n} x_j\,(AB_j)^t_{\mathcal{W}}$$

$$= \sum_{j=1}^{n} x_j\,T(\vec{v}_j).$$

Finally, since T is a linear transformation, we obtain

$$T([x_1,\dots,x_n]_{\mathcal{V}}) = T\left(\sum_{j=1}^{n} x_j \vec{v}_j\right)$$

$$= \sum_{j=1}^{n} x_j\,T(\vec{v}_j).$$

Thus, we arrive at

$$T([x_1, \cdots, x_n]_{\mathcal{V}}) = \sum_{j=1}^{n} x_j \, T(\vec{v}_j)$$

$$= \left(A \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \right)^t_{\mathcal{W}} .$$

This completes the proof of the Claim.

The matrix A, constructed above, is called the matrix which represents T with respect to the ordered bases $\mathcal{V}$ for $\mathbb{R}^n$ and $\mathcal{W}$ for $\mathbb{R}^m$.

Ex 1: Define $T: \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ via

$$T([x,y]) = [x+y, x-y].$$

We have already shown that T is a linear transformation - see Ex 1, page 268. Let's construct the matrix which represents T with respect to the standard basis for $\mathbb{R}^2$, i.e.

$$S_2 = \{\vec{e}_1 = [1,0], \vec{e}_2 = [0,1]\}.$$

Note that

$$T(\vec{e}_1) = T([1,0]) = [1,1] = [1,1]_{S_2}$$

and

$$T(\vec{e}_2) = T([0,1]) = [1,-1] = [1,-1]_{S_2}.$$

The desired matrix is

$$A = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}.$$

Note that the 1^{st} column of A is the transpose of the $\mathcal{S}_2$ coordinates of $T(\vec{e}_1)$ and the 2^{nd} column of A is the transpose of the $\mathcal{S}_2$ coordinates of $T(\vec{e}_2)$.

To check that this A works, note that

$$\left(A\begin{bmatrix} x \\ y \end{bmatrix}\right)^t = \left(\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}\right)^t$$

$$= \begin{bmatrix} x+y \\ x-y \end{bmatrix}^t = [x+y, x-y] = T([x,y]).$$

Ex 2: Once again we'll consider

$$T: \mathbb{R}^2 \to \mathbb{R}^2$$

defined by

$$T([x,y]) = [x+y, x-y].$$

Now, however, we'll construct the matrix representing T for a nonstandard ordered basis.

Let $\vec{v}_1 = [1+\sqrt{2}, 1]$ and

$$\vec{v}_2 = [1-\sqrt{2}, 1].$$

Observe that $\vec{v}_2$ is not a scalar multiple of $\vec{v}_1$ – so

$$\vec{v}_2 \notin \text{span}(\{\vec{v}_1\}).$$

This implies that $\{\vec{v}_1, \vec{v}_2\}$ is a linearly independent set of two vectors in $\mathbb{R}^2$ — so

$$\mathcal{V} = \{\vec{v}_1, \vec{v}_2\}$$

must be a basis for $\mathbb{R}^2$.

We'll construct the matrix which represents T with respect to the ordered basis $\mathcal{V}$. To this end, note that

$$T(\vec{v}_1) = T([1+\sqrt{2}, 1])$$

$$= [1+\sqrt{2}+1, 1+\sqrt{2}-1]$$

$$= [\sqrt{2}+2, \sqrt{2}]$$

$$= \sqrt{2}\,[1+\sqrt{2}, 1]$$

$$= \sqrt{2}\,\vec{v}_1 = [\sqrt{2}, 0]_{\mathcal{V}}$$

and

$$T(\vec{v}_2) = T([1-\sqrt{2}, 1])$$

$$= [1-\sqrt{2}+1, 1-\sqrt{2}-1]$$

$$= [2-\sqrt{2}, -\sqrt{2}]$$

$$= -\sqrt{2}\,[1-\sqrt{2}, 1]$$

$$= -\sqrt{2}\,\vec{v}_2 = [0, -\sqrt{2}]_{\mathcal{V}}.$$

Thus the matrix which represents T with respect to the ordered basis $\mathcal{V}$ is

$$B = \begin{bmatrix} \sqrt{2} & 0 \\ 0 & -\sqrt{2} \end{bmatrix}.$$

Put another way

$$T([s,t]_{\mathcal{V}}) = \left(\begin{bmatrix} \sqrt{2} & 0 \\ 0 & -\sqrt{2} \end{bmatrix} \begin{bmatrix} s \\ t \end{bmatrix} \right)^t_{\mathcal{V}} .$$

This can be verified as follows:

$$\begin{aligned} T([s,t]_{\mathcal{V}}) &= T(s\vec{v}_1 + t\vec{v}_2) \\ &= sT(\vec{v}_1) + tT(\vec{v}_2) \\ &= s\sqrt{2}\,\vec{v}_1 + t(-\sqrt{2})\,\vec{v}_2 \\ &= \sqrt{2}s\,\vec{v}_1 - \sqrt{2}t\,\vec{v}_2 \end{aligned}$$

while

$$\left(\begin{bmatrix} \sqrt{2} & 0 \\ 0 & -\sqrt{2} \end{bmatrix} \begin{bmatrix} s \\ t \end{bmatrix} \right)^t_{\mathcal{V}} = \begin{bmatrix} \sqrt{2}s \\ -\sqrt{2}t \end{bmatrix}^t_{\mathcal{V}}$$

$$= [\sqrt{2}s, -\sqrt{2}t]_{\mathcal{V}} = \sqrt{2}s\,\vec{v}_1 - \sqrt{2}t\,\vec{v}_2 .$$

Example 2 is meant to suggest how useful generalized coordinates can be. With respect to standard coordinates, as in Example 1, the matrix for T is not diagonal – making it more difficult to understand how T is acting on $\mathbb{R}^2$, from a geometric perspective. The fact that the matrix for T, with respect to $\mathcal{V}$, is diagonal makes it somewhat easier to understand how T acts on $\mathbb{R}^2$.

Indeed, let L_i denote the line through the origin in $\mathbb{R}^2$ containing

$\vec{v}_i$, for $i=1,2$. T acts on L_1 by stretching L_1 by a factor of $\sqrt{2}$. T acts on L_2 by first reflecting L_2 through the origin and then stretching by a factor of $\sqrt{2}$. Add to this that T is a linear transformation and, in this case, L_1 and L_2 are orthogonal (since $\vec{v}_1 \cdot \vec{v}_2 = 0$) – Example 2 provides us with a rather good understanding of the geometry of T.

In passing, I must acknowledge that I have given you no clue as where the basis $\mathcal{V}$, used in Example 2, came from. We'll discuss the subject of diagonalizing matrices after we have discussed determinants.

Finally, I should mention that, if

$$T: \mathbb{R}^n \to \mathbb{R}^m$$

is a linear transformation, there may be many distinct matrices which represent T – just consider Examples 1 and 2. However, these distinct matrices result from changing the ordered bases for $\mathbb{R}^n$ and $\mathbb{R}^m$.

In other words, once $\mathbb{R}^n$ and $\mathbb{R}^m$ are given fixed ordered bases, the matrix which represents T with respect to these given bases is unique. This follows from the coordinates of a vector, with respect to a given ordered basis, are unique.

<u>HW #37</u>: Define $T: \mathbb{R}^3 \to \mathbb{R}^2$ via

$T([x,y,z]) = [z, x+y]$.

Verify that T is a linear transformation and construct the matrix which represents T with respect to the standard bases for $\mathbb{R}^3$ and $\mathbb{R}^2$.

HW #38: Suppose $T: \mathbb{R}^n \to \mathbb{R}^m$ and $S: \mathbb{R}^m \to \mathbb{R}^k$ are two linear transformations. Prove that their composition

$$S \circ T: \mathbb{R}^n \longrightarrow \mathbb{R}^k$$

is also a linear transformation.

Working With Matrix Representations

According to HW #38, if $T: \mathbb{R}^n \to \mathbb{R}^m$ and $S: \mathbb{R}^m \to \mathbb{R}^k$ are two linear transformations, then their composition

$$S \circ T : \mathbb{R}^n \to \mathbb{R}^k$$

is also a linear transformation.

Many of the games played, using matrix representations, make use of the following Proposition.

Proposition #31: Suppose

$T: \mathbb{R}^n \to \mathbb{R}^m$ and $S: \mathbb{R}^m \to \mathbb{R}^k$

are two linear transformations,

$\mathcal{V}$ is an ordered basis for $\mathbb{R}^n$,

$\mathcal{W}$ is an ordered basis for $\mathbb{R}^m$,

$\mathcal{U}$ is an ordered basis for $\mathbb{R}^k$,

B is the matrix which represents T with respect to the bases $\mathcal{V}$ and $\mathcal{W}$, and A is the matrix which represents S with respect to the bases $\mathcal{W}$ and $\mathcal{U}$.

Then AB is the matrix which represents $S \circ T$ with respect to the bases $\mathcal{V}$ and $\mathcal{U}$.

<u>Proof:</u> We are told that

$$T([x_1, \cdots, x_n]_{\mathcal{V}}) = \left(B \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}\right)^t_{\mathcal{W}}$$

and

$$S([y_1, \cdots, y_m]_{\mathcal{W}}) = \left(A \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix}\right)^t_{\mathcal{U}} .$$

Note that if

$$[y_1, \cdots, y_m]_{\mathcal{W}} = \left(B \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}\right)^t_{\mathcal{W}}$$

then

$$\begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix} = B \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}.$$

It follows that

$$\begin{aligned} S \circ T([x_1, \cdots, x_n]_{\mathcal{V}}) &= S(T([x_1, \cdots, x_n]_{\mathcal{V}})) \\ &= S\left(\left(B \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \right)^t_{\mathcal{W}} \right) \\ &= \left(A \left(B \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \right) \right)^t_{\mathcal{U}} \\ &= \left(AB \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \right)^t_{\mathcal{U}}. \end{aligned}$$

This completes the proof.

HW #39: Verify Proposition #31 in the case

$$\mathbb{R}^3 \xrightarrow{T} \mathbb{R}^2 \xrightarrow{S} \mathbb{R}^2$$

where

$$T([x,y,z]) = \left(\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} \right)^t$$

and

$$S([s,t]) = \left(\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} s \\ t \end{bmatrix} \right)^t .$$

HW #40: The identity transformation $id: \mathbb{R}^n \to \mathbb{R}^n$ is defined to be

$$id([x_1, \cdots, x_n]) = [x_1, \cdots, x_n].$$

Prove that id is a linear transformation.

Change of Coordinate Matrices

According to HW #40, the identity transformation

$$id : \mathbb{R}^n \longrightarrow \mathbb{R}^n$$

is a linear transformation. Suppose that $\mathcal{V} = \{\vec{v}_1, \cdots, \vec{v}_n\}$ and $\mathcal{W} = \{\vec{w}_1, \cdots, \vec{w}_n\}$ are two ordered bases for $\mathbb{R}^n$ and that A is the matrix which represents id with respect to $\mathcal{V}$ on the domain $\mathbb{R}^n$ and $\mathcal{W}$ on the target $\mathbb{R}^n$, i.e.

$$id([x_1, \cdots, x_n]_{\mathcal{V}}) = \left(A \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \right)^t_{\mathcal{W}} .$$

Since

$$id([x_1,\cdots,x_n]_{\mathcal{V}})$$
$$= id\left(\sum_{i=1}^n x_i \vec{v}_i\right)$$
$$= \sum_{i=1}^n x_i \, id(\vec{v}_i)$$
$$= \sum_{i=1}^n x_i \vec{v}_i$$
$$= [x_1,\cdots,x_n]_{\mathcal{V}},$$

it follows that

$$[x_1,\cdots,x_n]_{\mathcal{V}} = \left(A\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}\right)^t_{\mathcal{W}}.$$

Since multiplication by A transforms $\mathcal{V}$-coordinates into $\mathcal{W}$-coordinates, A is frequently referred to as the matrix which changes $\mathcal{V}$-coordinates into $\mathcal{W}$-coordinates.

Frequently, the elements in a nonstandard basis for $\mathbb{R}^n$ are defined using standard coordinates. When this occurs, it is easy to construct the matrix which changes the nonstandard coordinates into standard coordinates. For example, consider the nonstandard basis for $\mathbb{R}^2$, used in Example 2 on page 308:

$$\mathcal{V} = \{\vec{v}_1 = [1+\sqrt{2}, 1], \vec{v}_2 = [1-\sqrt{2}, 1]\}.$$

Since $\vec{v}_1$ and $\vec{v}_2$ are defined using standard coordinates, we obtain

$$id(\vec{v}_1) = id\left([1+\sqrt{2}, 1]\right) = [1+\sqrt{2}, 1]_{\mathcal{S}_2}$$

$$id(\vec{v}_2) = id\left([1-\sqrt{2}, 1]\right) = [1-\sqrt{2}, 1]_{\mathcal{S}_2}.$$

Thus

$$id\left([x_1, x_2]_{\mathcal{V}}\right) = \left(\begin{bmatrix} 1+\sqrt{2} & 1-\sqrt{2} \\ 1 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right)^t_{\mathcal{S}_2}$$

or

$$[x_1, x_2]_{\mathcal{V}} = \left(\begin{bmatrix} 1+\sqrt{2} & 1-\sqrt{2} \\ 1 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right)^t_{\mathcal{S}_2}.$$

Thus

$$A = \begin{bmatrix} 1+\sqrt{2} & 1-\sqrt{2} \\ 1 & 1 \end{bmatrix}$$

changes $\mathcal{V}$-coordinates into $\mathcal{S}_2$-coordinates.

Changing $\mathcal{S}_2$-coordinates into $\mathcal{V}$-coordinates is facilitated using

<u>Proposition #32</u>: Suppose $\mathcal{V}$ and $\mathcal{W}$ are two ordered bases for $\mathbb{R}^n$ and A is the matrix which changes $\mathcal{V}$-coordinates into $\mathcal{W}$-coordinates. Then A^{-1} is the matrix which changes $\mathcal{W}$-coordinates into $\mathcal{V}$-coordinates.

<u>Proof</u>: We're told that

$$id\left([x_1,\cdots,x_n]_{\mathcal{V}}\right) = \left(A\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}\right)^t_{\mathcal{W}}.$$

Let B denote the matrix that changes $\mathcal{W}$-coordinates into $\mathcal{V}$-coordinates.

Then

$$id([y_1,\cdots,y_n]_{\mathcal{W}}) = \left(B\begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}\right)^t_{\mathcal{V}}.$$

Proposition #31 implies that

$$id \circ id([y_1,\cdots,y_n]_{\mathcal{W}}) = \left(AB\begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}\right)^t_{\mathcal{W}}.$$

Since $id \circ id([y_1,\cdots,y_n]_{\mathcal{W}}) = [y_1,\cdots,y_n]_{\mathcal{W}}$,

we obtain that

$$[y_1,\cdots,y_n]_{\mathcal{W}} = \left(AB\begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}\right)^t_{\mathcal{W}}.$$

However, it is clear that I_n is the matrix which changes $\mathcal{W}$-coordinates into $\mathcal{W}$-coordinates, so

$$AB = I_n$$

- due to uniqueness of matrix representations with respect to the same ordered bases.

Proposition 22, on page 262, now implies that $B = A^{-1}$ - which completes the proof.

We are now in a position to construct the $\mathcal{S}_2$-coordinate to $\mathcal{V}$-coordinate matrix where $\mathcal{V} = \{ \vec{v}_1 = [1+\sqrt{2}, 1], \vec{v}_2 = [1-\sqrt{2}, 1] \}$.

For this, recall that

$$A = \begin{bmatrix} 1+\sqrt{2} & 1-\sqrt{2} \\ 1 & 1 \end{bmatrix}$$

changes $\mathcal{V}$-coordinates into $\mathcal{S}_2$-coordinates.

By Proposition #32, we must construct A^{-1}.

$$\left[\begin{array}{cc|cc} 1+\sqrt{2} & 1-\sqrt{2} & 1 & 0 \\ 1 & 1 & 0 & 1 \end{array}\right]$$

$$\to \left[\begin{array}{cc|cc} 1 & 1 & 0 & 1 \\ 1+\sqrt{2} & 1-\sqrt{2} & 1 & 0 \end{array}\right]$$

$$\to \left[\begin{array}{cc|cc} 1 & 1 & 0 & 1 \\ 0 & -2\sqrt{2} & 1 & -1-\sqrt{2} \end{array}\right]$$

$$\to \left[\begin{array}{cc|cc} 1 & 1 & 0 & 1 \\ 0 & -1 & \frac{1}{2\sqrt{2}} & \frac{-1-\sqrt{2}}{2\sqrt{2}} \end{array}\right]$$

$$\longrightarrow \left[\begin{array}{cc:cc} 1 & 0 & \frac{1}{2\sqrt{2}} & 1-\frac{1+\sqrt{2}}{2\sqrt{2}} \\ 0 & -1 & \frac{1}{2\sqrt{2}} & -\frac{1+\sqrt{2}}{2\sqrt{2}} \end{array}\right]$$

$$\longrightarrow \left[\begin{array}{cc:cc} 1 & 0 & \frac{1}{2\sqrt{2}} & 1-\frac{1+\sqrt{2}}{2\sqrt{2}} \\ 0 & 1 & -\frac{1}{2\sqrt{2}} & \frac{1+\sqrt{2}}{2\sqrt{2}} \end{array}\right]$$

So $A^{-1} = \begin{bmatrix} \frac{1}{2\sqrt{2}} & 1-\frac{1+\sqrt{2}}{2\sqrt{2}} \\ -\frac{1}{2\sqrt{2}} & \frac{1+\sqrt{2}}{2\sqrt{2}} \end{bmatrix}$

$$= \begin{bmatrix} \frac{1}{2\sqrt{2}} & \frac{2\sqrt{2}-(1+\sqrt{2})}{2\sqrt{2}} \\ -\frac{1}{2\sqrt{2}} & \frac{1+\sqrt{2}}{2\sqrt{2}} \end{bmatrix}$$

$$= \begin{bmatrix} \frac{1}{2\sqrt{2}} & \frac{\sqrt{2}-1}{2\sqrt{2}} \\ -\frac{1}{2\sqrt{2}} & \frac{1+\sqrt{2}}{2\sqrt{2}} \end{bmatrix}$$

$$= \frac{1}{2\sqrt{2}} \begin{bmatrix} 1 & -(1-\sqrt{2}) \\ -1 & 1+\sqrt{2} \end{bmatrix}.$$

Thus

$$\frac{1}{2\sqrt{2}} \begin{bmatrix} 1 & -(1-\sqrt{2}) \\ -1 & 1+\sqrt{2} \end{bmatrix}$$

changes $\mathcal{S}_2$-coordinates into $\mathcal{V}$-coordinates.

Now recall that in Example 1 on page 306 we showed that the matrix representing $T: \mathbb{R}^2 \to \mathbb{R}^2$, where $T([x,y]) = [x+y, x-y]$, is

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

- with respect to the standard basis for $\mathbb{R}^2$.

We have now learned that

$$\begin{bmatrix} 1+\sqrt{2} & 1-\sqrt{2} \\ 1 & 1 \end{bmatrix}$$

changes $\mathcal{V}$-coordinates into $\mathcal{S}_2$-coordinates

and

$$\frac{1}{2\sqrt{2}}\begin{bmatrix} 1 & -(1-\sqrt{2}) \\ -1 & 1+\sqrt{2} \end{bmatrix}$$

changes $\mathcal{S}_2$-coordinates into $\mathcal{V}$-coordinates.

We would then expect that

$$\frac{1}{2\sqrt{2}}\begin{bmatrix} 1 & -(1-\sqrt{2}) \\ -1 & 1+\sqrt{2} \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} 1+\sqrt{2} & 1-\sqrt{2} \\ 1 & 1 \end{bmatrix}$$

represents T with respect to $\mathcal{V}$-coordinates.

As to why, note that the matrices

$$\frac{1}{2\sqrt{2}}\begin{bmatrix} 1 & -(1-\sqrt{2}) \\ -1 & 1+\sqrt{2} \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 1+\sqrt{2} & 1-\sqrt{2} \\ 1 & 1 \end{bmatrix},$$

being change of coordinates matrices, both represent

id — and $id \circ T \circ id = T$. Also,

$\begin{bmatrix} 1+\sqrt{2} & 1-\sqrt{2} \\ 1 & 1 \end{bmatrix}$ transforms $\mathcal{V}$-coordinates

into $\mathcal{S}_2$-coordinates,

$\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$ accepts $\mathcal{S}_2$-coordinates and

yields the $\mathcal{S}_2$-coordinates of the image

of T and, finally,

$\frac{1}{2\sqrt{2}} \begin{bmatrix} 1 & -(1-\sqrt{2}) \\ -1 & 1+\sqrt{2} \end{bmatrix}$ transforms $\mathcal{S}_2$-coordinates

into $\mathcal{V}$-coordinates.

To confirm these expectations,

$$\frac{1}{2\sqrt{2}} \begin{bmatrix} 1 & -(1-\sqrt{2}) \\ -1 & 1+\sqrt{2} \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 1+\sqrt{2} & 1-\sqrt{2} \\ 1 & 1 \end{bmatrix}$$

$$= \frac{1}{2\sqrt{2}} \begin{bmatrix} 1-(1-\sqrt{2}) & 1+(1-\sqrt{2}) \\ -1+(1+\sqrt{2}) & -1-(1+\sqrt{2}) \end{bmatrix} \begin{bmatrix} 1+\sqrt{2} & 1-\sqrt{2} \\ 1 & 1 \end{bmatrix}$$

$$= \frac{1}{2\sqrt{2}} \begin{bmatrix} \sqrt{2} & 2-\sqrt{2} \\ \sqrt{2} & -(2+\sqrt{2}) \end{bmatrix} \begin{bmatrix} 1+\sqrt{2} & 1-\sqrt{2} \\ 1 & 1 \end{bmatrix}$$

$$= \frac{1}{2\sqrt{2}} \begin{bmatrix} \sqrt{2}+2+2-\sqrt{2} & \sqrt{2}-2+2-\sqrt{2} \\ \sqrt{2}+2-(2+\sqrt{2}) & \sqrt{2}-2-(2+\sqrt{2}) \end{bmatrix}$$

$$= \frac{1}{2\sqrt{2}} \begin{bmatrix} 4 & 0 \\ 0 & -4 \end{bmatrix} = \begin{bmatrix} \sqrt{2} & 0 \\ 0 & -\sqrt{2} \end{bmatrix}$$

Recall that $\begin{bmatrix} \sqrt{2} & 0 \\ 0 & -\sqrt{2} \end{bmatrix}$ is the matrix representing T, with respect to $\mathcal{V}$, in Example 2 on page 308.

<u>HW #41</u>: Define $T: \mathbb{R}^2 \to \mathbb{R}^2$ via

$T([x,y]) = [y,x]$.

(i) Verify that T is a linear transformation.

(ii) Construct the matrix which represents T with respect to $\mathcal{S}_2$.

(iii) Let $\mathcal{V} = \{ \vec{v}_1 = [1,1], \vec{v}_2 = [1,-1] \}$.

Construct the matrix which represents T with respect to $\mathcal{V}$.

(iv) Construct the matrix which transforms $\mathcal{V}$-coordinates into $\mathcal{S}_2$-coordinates.

(v) Construct the matrix which transforms $\mathcal{S}_2$-coordinates into $\mathcal{V}$-coordinates.

(vi) Use the matrices constructed in parts (ii), (iv) and (v) to construct the matrix constructed in (iii).

(vii) Use the matrix constructed in (iii) to discuss the geometry of T.