A linear transformation is a *mapping* (or a *function*) which *preserves the structure of a vector space*. It *respects linearity* (essentially, it takes "flat" objects to "flat" objects.)

Definition

Let V, W be vector spaces. A function $T : V \to W$ is said to be a *linear transformation* if

(i) $T(v + w) = T(v) + T(w), \quad \forall v, w \in V$

(ii) $T(cv) = cT(v) \quad \forall v \in V, c \in \mathbb{R}$

Coordinates with respect to a Basis

Definition

Suppose $\mathcal{B} = \{b_1, \ldots, b_n\}$ is a basis for V and $x \in V$. The *coordinates of x relative to $\mathcal{B}$*(or the *$\mathcal{B}$-coordinates of x*) are the weights $c_1, \ldots, c_n$ such that

$$x = c_1 b_1 + \ldots + c_n b_n.$$

The vector $(c_1, \ldots, c_n) \in \mathbb{R}^n$ is denoted by $[x]_{\mathcal{B}}$, and is called the *coordinate vector of x relative to $\mathcal{B}$* or the *$\mathcal{B}$-coordinate vector of x*. The mapping

$$x \to [x]_{\mathcal{B}}$$

is called the *coordinate mapping (determined by $\mathcal{B}$)*.

Theorem

Let $\mathcal{B} = \{b_1, \ldots, b_n\}$ be a basis for a vector space V. The coordinate mapping

$$x \to [x]_{\mathcal{B}}$$

is an invertible linear transformation from V to $\mathbb{R}^n$.

Any linear transformation can be completely determined by what is does to a fixed basis, in the sense that,

Proposition

If $\mathcal{B} = \{b_1, \ldots, b_n\}$ is a basis for a vector space V, and if $S, T : V \to W$ are linear transformations, then $S = T$ iff

$$S(b_i) = T(b_i), \quad \forall i = 1, \ldots, n.$$

Proposition

Every linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ is a matrix transformation. In other words, there exists a unique $m \times n$ matrix A such that

$$T(\mathbf{x}) = A\mathbf{x}, \quad \forall \mathbf{x} \in \mathbb{R}^n$$

A is called the *standard matrix for the linear transformation* T.

The Change-of-coordinates in $\mathbb{R}^n$

Definition

Let $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ be an ordered basis of $\mathbb{R}^n$. The matrix

$$P_{\mathcal{B}} = [\mathbf{b}_1 \quad \ldots \quad \mathbf{b}_n]$$

formed using the basis vectors $\mathbf{b}_1, \ldots, \mathbf{b}_n$ as columns, is called the *change-of-coordinates* matrix from $\mathcal{B}$ to the standard basis in $\mathbb{R}^n$.

The change-of-coordinates matrix takes a coordinate vector with respect to the $\mathcal{B}$ basis and tranforms it to standard coordinates. So if $\mathbf{x}$ is a vector in $\mathbb{R}^n$, then

$$\mathbf{x} = P_{\mathcal{B}}[\mathbf{x}]_{\mathcal{B}}.$$

Proposition

Let $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ be an ordered basis of $\mathbb{R}^n$. Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be the coordinate transformation which sends

$$\mathbf{x} \mapsto [\mathbf{x}]_{\mathcal{B}}.$$

The change-of-coordinates matrix $P_{\mathcal{B}}$ is the standard matrix of the inverse T^{-1} of the coordinate transformation.

The standard matrix of the coordinate transformation T is $P_{\mathcal{B}}^{-1}$.

Particular Case: The $\mathcal{B}$-matrix

Definition

Let $T : V \to V$ be a linear transformation from a vector space to itself. Let $\mathcal{B} = \{v_1, \ldots, v_n\}$ be an ordered basis for V. There is a unique matrix $[T]_{\mathcal{B}}$, which we call the $\mathcal{B}$-matrix of T such that

$$[T(v)]_{\mathcal{B}} = [T]_{\mathcal{B}}[v]_{\mathcal{B}}, \quad \forall v \in V.$$

Further, $[T]_{\mathcal{B}}$ is obtained using the formula

$$[T]_{\mathcal{B}} = [[T(v_1)]_{\mathcal{B}} \quad \ldots \quad [T(v_n)]_{\mathcal{B}}]$$

Proposition

Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be a linear transformation. Let A be the standard matrix of T. Let $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ be any ordered basis of $\mathbb{R}^n$. Let

$$P_{\mathcal{B}} = [\mathbf{b}_1 \quad \cdots \quad \mathbf{b}_n]$$

be the change-of-coordinates matrix from $\mathcal{B}$ to the standard basis in $\mathbb{R}^n$. Then the $\mathcal{B}$-matrix of T is $P_{\mathcal{B}}^{-1} A P_{\mathcal{B}}$.

standard matrix of T

Proof:

For every $\mathbf{x} \in \mathbb{R}^n$,

$$[T]_{\mathcal{B}} = P_{\mathcal{B}}^{-1} A P_{\mathcal{B}}$$

$$\begin{aligned} [T(\mathbf{x})]_{\mathcal{B}} &= [A\mathbf{x}]_{\mathcal{B}} \\ &= P_{\mathcal{B}}^{-1} A\mathbf{x} \\ &= P_{\mathcal{B}}^{-1} A P_{\mathcal{B}} [\mathbf{x}]_{\mathcal{B}}. \end{aligned}$$

Application: Reflection across the plane $x + y + z = 0$

$[T]_\beta = P_\beta^{-1} A P_\beta$

$(-2, -3, 5)$

$\boxed{A = P_\beta [T]_\beta P_\beta^{-1}}$

We choose v_1 perpendicular to the plane, v_2, v_3 in the plane (computations are easier if we choose v_2 and v_3 perpendicular to each other).

$v_1 = (1, 1, 1)$

$v_2 = (1, 0, -1)$

$v_3 = (1, -2, 1)$

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

$T(v_1) = (-1, -1, -1)$

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

$$\beta = \{v_1, v_2, v_3\} = \{\underline{(1,1,1)}, \underline{(1,0,-1)}, \underline{(1,-2,1)}\}$$

$$[T(v_1)]_\beta = \{\underline{(-1,-1,-1)}\}_\beta = (-1,0,0)$$

$$T(v_2) = v_2 \Rightarrow [T(v_2)]_\beta = (0,1,0)$$

$$v_2 = 0 \times (1,1,1) + 1(1,0,-1) + 0(1,-2,1)$$

$$A = P_\beta [T]_\beta P_\beta^{-1}$$

$[T]_\beta$ Similar to A.

$$= \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & -2 \\ 1 & -1 & 1 \end{bmatrix} \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1/3 & 1/3 & 1/3 \\ 1/2 & 0 & -1/2 \\ 1/6 & -1/3 & 1/6 \end{bmatrix}$$

$$= \begin{bmatrix} -1 & 1 & 1 \\ -1 & 0 & -2 \\ -1 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1/3 & 1/3 & 1/3 \\ 1/2 & 0 & -1/2 \\ 1/6 & -1/3 & 1/6 \end{bmatrix} = \begin{bmatrix} 1/3 & -2/3 & -2/3 \\ -2/3 & 1/3 & -2/3 \\ -2/3 & -2/3 & 1/3 \end{bmatrix}$$

What are coordinates of the point $(-2,-3,5)$ after reflection?

$$\begin{bmatrix} 1/3 & -2/3 & -2/3 \\ -2/3 & 1/3 & -2/3 \\ -2/3 & -2/3 & 1/3 \end{bmatrix} \begin{bmatrix} -2 \\ -3 \\ -5 \end{bmatrix} = \begin{bmatrix} \\ \\ \\ \end{bmatrix}$$

— Finish yourself.

Definition

Two $n \times n$ matrix A and B are said to be *similar* if there exists an invertible matrix P such that $B = P^{-1}AP$.

Matrix of a Linear Transformation

Proposition (M)

Let V, W be vector spaces. Let $\mathcal{B} = \{v_1, \ldots, v_n\}$ be an ordered basis for V and $\mathcal{C} = \{w_1, \ldots, w_m\}$ be an ordered basis for W. Let $T : V \to W$ be a linear transformation. There exists a unique $m \times n$ matrix A such that

$$[T(v)]_{\mathcal{C}} = A[v]_{\mathcal{B}}, \quad \text{for every } v \in V.$$

Further, we have

$$A = [[T(v_1)]_{\mathcal{C}} \quad \ldots \quad [T(v_n)]_{\mathcal{C}}]$$

$$\begin{tikzcd} V \arrow[r, "f"] \arrow[rr, bend right, "g \circ f"'] & W \arrow[r, "g"] & Z \end{tikzcd} \qquad \begin{tikzcd} U \arrow[r, "T"] \arrow[rr, bend right, "S \circ T"'] & V \arrow[r, "S"] & W \end{tikzcd}$$

Proposition

Let U, V and W be vector spaces. Let $T : U \to V$ and $S : V \to W$ be linear transformations. Then the composite $S \circ T : U \to W$ is also a linear transformation.

Proof: Let $u_1, u_2 \in U$.

$$S \circ T(u_1 + u_2) = S(T(u_1 + u_2))$$

$$= S(T(u_1) + T(u_2)) \quad (\text{because } T \text{ is L.T.})$$

$$= \underline{S(T(u_1))} + S(T(u_2))$$

(because S is a L.T.)

$$= S \circ T(u_1) + S \circ T(u_2)$$

Since u_1, u_2 were arbitrary, $S \circ T$ satisfies condition 1 of the definition of a L.T.

Next let $u \in U$, $c \in \mathbb{R}$.

$$S \circ T(cu)$$

$$= S(\underline{T(cu)})$$

$$= S(c\,T(u))$$

$$= c\,S(T(u)) = c\,S \circ T(u).$$

As the choice of u and c were arbitrary, $S \circ T$ satisfies condition ② of defn of LT.

$\therefore \quad S \circ T : U \longrightarrow W$ is a linear transformation.

Immediate Corollary:

If $T_1 : V_1 \to V_2$,

$T_2 : V_2 \to V_3$, ...

$T_n : V_n \to V_{n+1}$

Then the composite $T_n \circ T_{n-1} \circ \cdots \circ T_1 : V_1 \to V_{n+1}$ is a linear transformation.

Exercise: Proof by induction.

$$
\begin{tikzcd}
V_1 \arrow[r, "T_1"'] \arrow[rrrr, bend right, "T_n \circ T_{n-1} \circ \cdots \circ T_1"'] & V_2 \arrow[r, "T_2"'] & V_3 \cdots & V_n \arrow[r, "T_n"'] & V_{n+1}
\end{tikzcd}
$$

Proof of Proposition (M)

This proof is slightly tricky, and I expect that only a few of you can come up with it on your own at this point. As you develop mathematical maturity though, you will realize that the idea behind it is almost trivial.

The idea: Give both coordinate mappings names and look at what is happening at the level of Euclidean spaces.

$$
\begin{tikzcd}
S_1^{-1}(x) \in V \arrow[r, "T"] \arrow[d, shift left, "S_1 \text{ coordinate map}"'] & W \ni T(S^{-1}(x)) \arrow[d, shift left, "S_2 \text{ coordinate}"] \\
\mathbb{R}^n \ni x \arrow[u, shift left] \arrow[r] & \mathbb{R}^m \ni S_2(T(S^{-1}(x))) \arrow[u, shift left]
\end{tikzcd}
$$

$$v \mapsto [v]_\beta \qquad w \mapsto [w]_\varphi$$

$$x \in \mathbb{R}^n$$

$$x \longmapsto \underbrace{S_2(T(S_1^{-1}(x)))}_{S_2 \circ T \circ S_1^{-1}} \in \mathbb{R}^m$$

Let A be the standard matrix of $\underline{S_2 \circ T \circ S_1^{-1}} : \mathbb{R}^n \to \mathbb{R}^m$

$$S_2 \circ T \circ S_1^{-1}(x) = A x, \quad \forall x \in \mathbb{R}^n$$

$$A = \left[S_2 \circ T \circ S_1^{-1}(e_1) \quad S_2 \circ T \circ S_1^{-1}(e_2) \quad \cdots \quad S_2 \circ T \circ S_1^{-1}(e_n) \right]$$

$$[T(v_1)]_{e_\varphi} = S_2(T(v_1)) \qquad S_2 : \omega \longmapsto [\omega]_{e_\varphi}$$

$$= \left[[T(S_1^{-1}(e_1))]_{e_\varphi} \cdots [T(S_1^{-1}(e_n))]_{e_\varphi} \right]$$

$$= \left[[T(v_1)]_{e_\varphi} \cdots [T(v_n)]_{e_\varphi} \right]$$

$$S_1(v_1) = (1, 0, \ldots, 0) = e_1$$

$$\Rightarrow S_1^{-1}(e_1) = v_1$$

$$v_1 = 1 \cdot v_1 + 0 \cdot v_2 + \cdots + 0 \cdot v_n$$

$$S_1(v_j) = (0, \ldots, \underset{\substack{\uparrow \\ j\text{-th}}}{1}, \ldots, 0) = e_j$$

Claim: A is the matrix of T with respect to β and $\mathcal{C}$.

Wts: $[T(v)]_{\mathcal{C}} = A[v]_\beta$, $\forall v \in V$.

Let $v \in V$ arbitrary.

$$A[v]_\beta = \left[[T(v_1)]_\varphi \cdots [T(v_n)]_\varphi \right] \times [v]_\beta$$

Let $[v_\beta] = (c_1, \ldots, c_n)$

$$A[v]_\beta = \left[[T(v_1)]_\varphi \cdots [T(v_n)]_\varphi \right] \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}$$

$$= c_1 [T(v_1)]_{e_\varphi} + \cdots + c_n [T(v_n)]_{e_\varphi} \quad -(1)$$

$$T(v) = T(c_1 v_1 + \cdots + c_n v_n)$$

$$= c_1 T(v_1) + \cdots + c_n T(v_n)$$

$$\Rightarrow [T(v)]_{e_\varphi} = c_1 [T(v_1)]_{e_\varphi} + \cdots + c_n [T(v_n)]_{e_\varphi} \quad -(2)$$

From 1 & 2,

we obtain

$$[T(v)]_\gamma = A[v]_\beta .$$

Check

Let S_1 be the coordinate mapping from V to $\mathbb{R}^n$ with respect to the basis $\mathcal{B}$, and S_2 be the coordinate mapping from W to $\mathbb{R}^m$ with respect to the basis $\mathcal{C}$.

Then $S_2 \circ T \circ S_1^{-1}$ is a linear transformation from $\mathbb{R}^n \to \mathbb{R}^m$. Let A be the standard matrix of $S_2 \circ T \circ S_1^{-1}$.

Then for every $v \in V$,

$$S_2 \circ T \circ S_1^{-1}([v]_{\mathcal{B}}) = A[v]_{\mathcal{B}}$$

Hence for every $v \in V$,

$$A[v]_{\mathcal{B}} = S_2(T(v)) = [T(v)]_{\mathcal{C}}.$$

Now,

$$\begin{aligned} A &= [A\mathbf{e}_1 \quad \ldots A\mathbf{e}_n] \\ &= [A[v_1]_{\mathcal{B}} \quad \ldots A[v_n]_{\mathcal{B}}] \\ &= [[T(v_1)]_{\mathcal{C}} \quad \ldots \quad [T(v_n)]_{\mathcal{C}}] \end{aligned}$$

This unique matrix A is called is called the *matrix of T* with respect to the bases $\mathcal{B}$ and $\mathcal{C}$, and is denoted by $[T]_{\mathcal{B},\mathcal{C}}$.

Let us look at some examples of how to find this matrix, when working with matrices and/or polynomials.