Shifrin 1.4 Linear Transformations.

Definition. A function $T: \mathbb{R}^n \to \mathbb{R}^m$ is called a <u>linear transformation</u> or <u>linear map</u> if

1) $T(\vec{u}+\vec{v}) = T(\vec{u}) + T(\vec{v})$
2) $T(c\vec{v}) = cT(\vec{v})$.

These properties are so natural that one tends to assume they always hold.

Self-check. Which of the following are linear?

$x \longmapsto ax+b$

$x \longmapsto ax$

$x \longmapsto x^2$

$\vec{x} \mapsto \vec{v}\cdot\vec{x}$

$f(x) \mapsto f'(x)$

$f(x) \mapsto \int_0^1 f(x)\,dx$

$f(x) \mapsto \int f(x)\,dx$

Notice that if $\vec{x} = x_1\vec{e}_1 + \ldots + x_n\vec{e}_n$ ~~then~~ and T is linear, then

$$T(\vec{x}) = T(x_1\vec{e}_1 + \ldots + x_n\vec{e}_n)$$

$$= T(x_1\vec{e}_1) + \ldots + T(x_n\vec{e}_n)$$

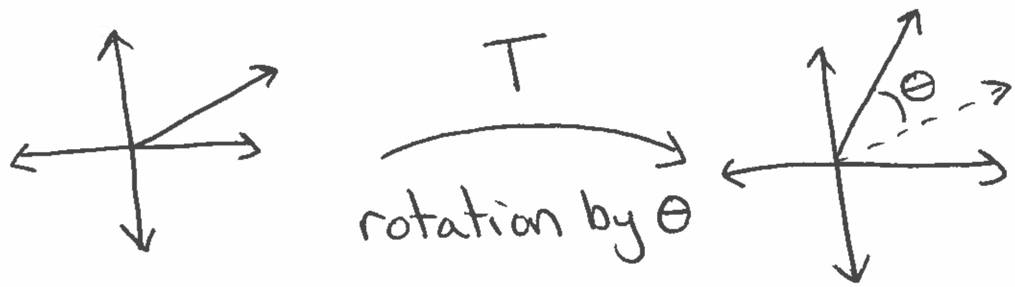
$$= x_1 T(\vec{e}_1) + \ldots + x_n T(\vec{e}_n).$$

This means that T is really determined by the vectors $T(\vec{e}_1), \ldots, T(\vec{e}_n)$. We encode $T: \mathbb{R}^n \to \mathbb{R}^m$ by the $m \times n$ array

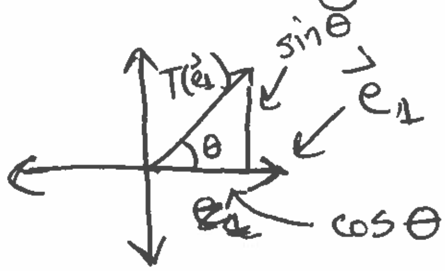
$$[T] = \underbrace{\begin{bmatrix} \uparrow & & \uparrow \\ T(\vec{e}_1) & \cdots & T(\vec{e}_n) \\ \downarrow & & \downarrow \end{bmatrix}}_{n \text{ columns}} \left.\vphantom{\begin{bmatrix} \uparrow \\ T \\ \downarrow \end{bmatrix}}\right\} m \text{ rows}$$

which is called the <u>standard</u> <u>matrix</u> for T.

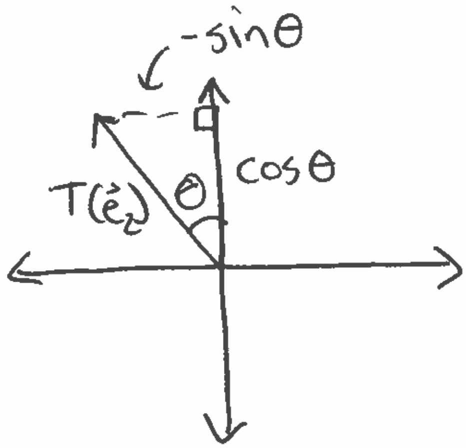
We now give some useful examples.



We can find $T(\vec{e}_1)$ and $T(\vec{e}_2)$ with some trigonometry



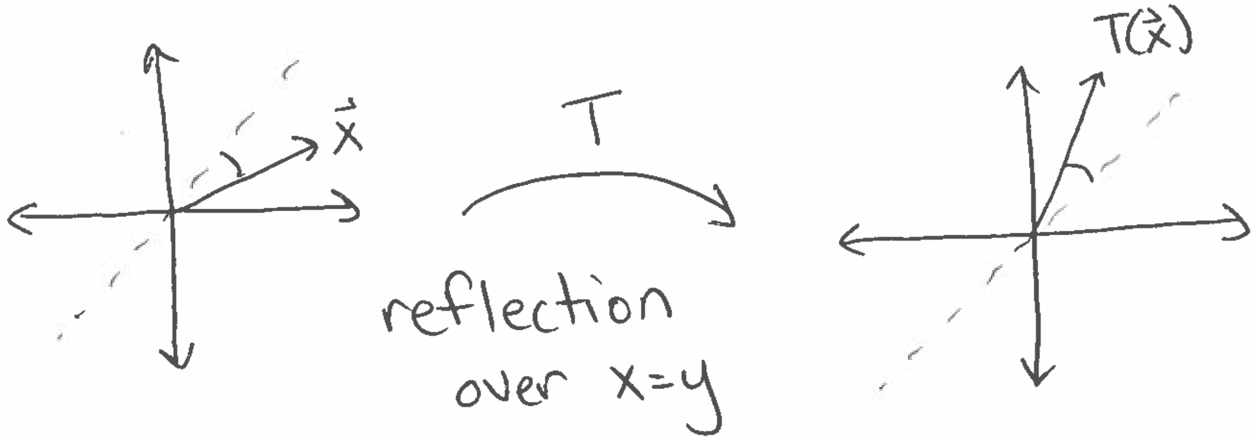
$$T(\vec{e}_1) = \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}$$



$$T(\vec{e}_2) = \begin{bmatrix} -\sin\theta \\ \cos\theta \end{bmatrix}$$

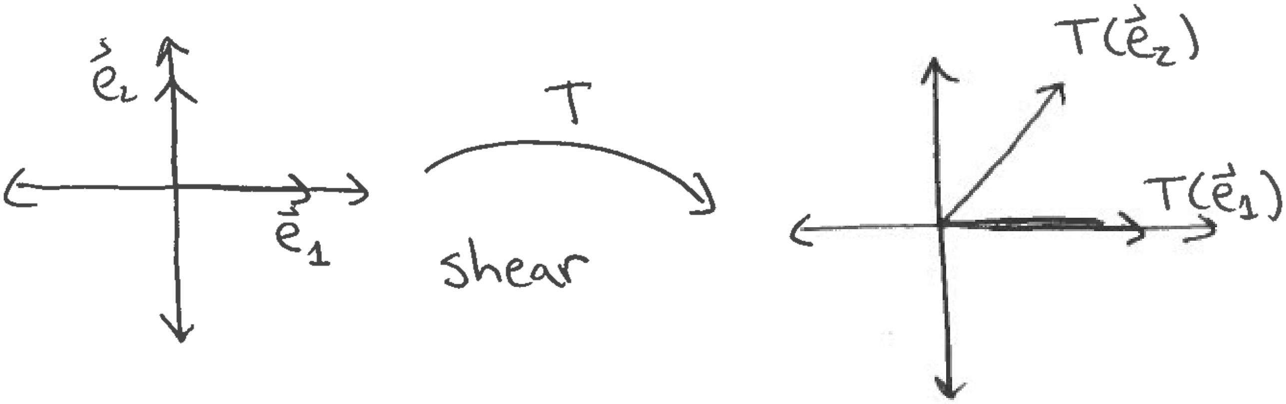
so

$$[T] = \begin{bmatrix} \uparrow & \uparrow \\ T(\vec{e}_1) & T(\vec{e}_2) \\ \downarrow & \downarrow \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$



Here $T\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} x_2 \\ x_1 \end{bmatrix}$. (swaps coords)

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



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

Definition. We call an $m \times n$ array of real numbers an $m \times n$ <u>matrix</u>, A.
We define (for a vector $\vec{x} \in \mathbb{R}^n$)

$$A\vec{x} = \begin{bmatrix} \leftarrow A_1 \rightarrow \\ \leftarrow A_m \rightarrow \end{bmatrix} \vec{x} = \begin{bmatrix} A_1 \cdot \vec{x} \\ \vdots \\ A_m \cdot \vec{x} \end{bmatrix} \in \mathbb{R}^m$$

or

$$(A\vec{x})_{j1} = \sum_{i=1}^{n} A_{ji} x_{i1}$$

Proposition. If $A = [T]$ is the standard matrix for a linear transformation T, then $A\vec{x} = T(\vec{x})$.

Proof. If $\vec{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = x_1 \vec{e}_1 + \dots + x_n \vec{e}_n$, then

$$T(\vec{x}) = x_1 T(\vec{e}_1) + \dots + x_n T(\vec{e}_n).$$

Now each $T(\vec{e}_j)$ can be written in terms of the standard basis for $\mathbb{R}^m$, which we call $\vec{d}_1, \dots, \vec{d}_m$ just this once. So

$$T(\vec{x}) = x_1\left(T(\vec{e}_1)_1 \vec{d}_1 + \dots + T(\vec{e}_1)_m \vec{d}_m\right)$$
$$+ \dots + x_n\left(T(\vec{e}_n)_1 d_1 + \dots + T(\vec{e}_n)_m \vec{d}_m\right)$$

and regrouping, we get

$$= \left(x_1 T(\vec{e}_1)_1 + \dots + x_n T(\vec{e}_n)_1\right) \vec{d}_1$$
$$+ \dots + \left(x_1 T(\vec{e}_1)_m + \dots + x_n T(\vec{e}_n)_m\right) \vec{d}_m$$

$$= \left(x_1 A_{11} + \dots x_n A_{1n}\right) \vec{d}_1$$
$$+ \dots + \left(x_1 A_{m1} + \dots x_n A_{mn}\right) \vec{d}_m$$

$$= \begin{bmatrix} A_1 \cdot \vec{x} \\ \vdots \\ A_m \cdot \vec{x} \end{bmatrix} \in \mathbb{R}^m, \text{ as claimed. } \square$$

Motto of the day: The linear transformation T and the standard matrix $[T]$ are not the same thing.

We will call the set of $m \times n$ matrices $\mathrm{Mat}_{m \times n}$ or $\mathcal{M}_{m \times n}$.

Look up: square matrix, diagonal entries, diagonal matrix, upper triangular lower triangular.

Definition. If $S, T : \mathbb{R}^n \to \mathbb{R}^m$ are linear maps, and $c \in \mathbb{R}$, we define

$$(cT)(\vec{x}) = c(T(\vec{x}))$$

$$(S+T)(\vec{x}) = S(\vec{x}) + T(\vec{x}).$$

There are corresponding operations on the standard matrices for these linear maps.

For $A \in \mathrm{Mat}_{m \times n}$, $c \in \mathbb{R}$, we have

$$cA = c\begin{bmatrix} a_{11} & & a_{1n} \\ & \ddots & \\ a_{m1} & & a_{mn} \end{bmatrix} = \begin{bmatrix} ca_{11} & & ca_{1n} \\ & \ddots & \\ ca_{m1} & & ca_{mn} \end{bmatrix}$$

And for $A, B \in \mathrm{Mat}_{m \times n}$,

$$A + B = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} + \begin{bmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & \ddots & \vdots \\ b_{m1} & \cdots & b_{mn} \end{bmatrix}$$

$$= \begin{bmatrix} a_{11}+b_{11} & \cdots & a_{1n}+b_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1}+b_{m1} & \cdots & a_{mn}+b_{mn} \end{bmatrix}$$

Proposition. $c[T] = [cT]$, $[S]+[T] = [S+T]$.

Proof. To be done in homework.

Example.

$$A = \begin{bmatrix} 1 & 2 \\ 4 & 6 \end{bmatrix} \quad B = \begin{bmatrix} -2 & 3 \\ 5 & -1 \end{bmatrix} \quad C = \begin{bmatrix} 1 & 2 & 1 \\ 3 & -4 & 7 \end{bmatrix}$$

Find $A+B$, $2C$. Can ~~you~~ you add $A+C$? (No!)

There is an algebra of matrices like the algebra of vectors.

Proposition. Let A, B, C be matrices in $Mat_{m\times n}$ and $c, d \in \mathbb{R}$, and 0 be the $m\times n$ matrix of 0s

1. $A + B = B + A$
2. $(A+B)+C = A+(B+C)$
3. $0 + A = A$
4. There is a matrix $-A$ so that $A + (-A) = 0$
5. $c(dA) = (cd)A$
6. $c(A+B) = cA + cB$
7. $(c+d)A = cA + dA$
8. $1A = A$

(so matrices are a real vector space).

Now let's go back to our basic idea.

$T : \mathbb{R}^n \to \mathbb{R}^m$ $\longleftrightarrow$ $[T] \in Mat_{m\times n}$

a linear transformation $\qquad$ an $m \times n$ matrix

If we have $S : \mathbb{R}^p \to \mathbb{R}^n$ and $T : \mathbb{R}^n \to \mathbb{R}^m$, we ~~must~~ can define the composition

$$T \circ S : \mathbb{R}^p \to \mathbb{R}^m \text{ by } (T \circ S)(\vec{x}) = T(S(\vec{x})).$$

Lemma. If $S : \mathbb{R}^p \to \mathbb{R}^n$ and $T : \mathbb{R}^n \to \mathbb{R}^m$ are linear transformations, then $T \circ S : \mathbb{R}^p \to \mathbb{R}^m$ is also a linear transformation.

Proof.
$$\begin{aligned}(T \circ S)(\vec{x} + \vec{y}) &= T(S(\vec{x} + \vec{y})) \\ &= T(S(\vec{x}) + S(\vec{y})) \\ &= T(S(\vec{x})) + T(S(\vec{y})) \\ &= (T \circ S)(\vec{x}) + (T \circ S)(\vec{y}).\end{aligned}$$

$$
\begin{aligned}
(T\circ S)(c\vec{x}) &= T(S(c\vec{x}))\\
&= T(cS(\vec{x}))\\
&= cT(S(\vec{x})). \qquad \square
\end{aligned}
$$

This raises a natural question: how are the standard matrices $[S] \in \mathrm{Mat}_{n\times p}$, $[T] \in \mathrm{Mat}_{m\times n}$ and $[T\circ S] \in \mathrm{Mat}_{m\times p}$ related?

Definition. Let A be an $m\times n$ matrix and B be an $n\times p$ matrix. The <u>matrix product</u> AB is defined by

$$
\begin{aligned}
(AB)_{ij} &= \sum_{k=1}^{n} a_{ik} b_{kj}\\
&= (i\text{th } \underline{\text{row}} \text{ of } A)\cdot(j\text{th } \underline{\text{column}} \text{ of } B)
\end{aligned}
$$

Example.

$$\begin{bmatrix} 1 & 3 \\ 2 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 4 & 1 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 4-3 & 1+3 \\ 8+1 & 2-1 \\ 4-1 & 1+1 \end{bmatrix} = \begin{bmatrix} 1 & 4 \\ 9 & 1 \\ 3 & 2 \end{bmatrix}$$

$$3\times\cancel{2} \qquad \cancel{2}\times 2 \qquad 3\times 2$$

Definition. If the matrix product AB exists, we say A, B are <u>conformable</u> for multiplication.

Lemma. A, B are conformable $\Leftrightarrow$ the # of columns of A = # of rows of B.

Proposition. Suppose $S : \mathbb{R}^p \to \mathbb{R}^n$ and $T : \mathbb{R}^n \to \mathbb{R}^m$ are linear transformations. Then

$$[T \circ S] = [T][S]$$

matrix product

Proof. Suppose that $A = [T]$ and $B = [S]$. For any $\vec{x} \in \mathbb{R}^p$, we know

$$
\begin{aligned}
[T \circ S]\vec{x} &= T(S(\vec{x})) \\
&= T(B\vec{x}) \\
&= A(B\vec{x})
\end{aligned}
$$

matrix vector mult

We must therefore show that

$$A(B\vec{x}) = (AB)\vec{x}$$

matrix vector mult | matrix matrix mult | matrix vector mult

Now

$$
\begin{aligned}
(A(B\vec{x}))_{j1} &= \sum_{i=1}^{n} A_{ji}(B\vec{x})_{i1} \\
&= \sum_{i=1}^{i=n} A_{ji}\left(\sum_{k=1}^{p} B_{ik}\vec{x}_{k1}\right)
\end{aligned}
$$

$$= \sum_{i=1}^{n} \sum_{k=1}^{p} A_{ji} B_{ik} \vec{x}_{k1}$$

$$= \sum_{k=1}^{p} \left(\sum_{i=1}^{n} A_{ji} B_{ik} \right) \vec{x}_{k1}$$

$$= \sum_{k=1}^{p} (AB)_{jk} \vec{x}_{k1}$$

$$= \left((AB) \vec{x} \right)_{j1} \qquad \square$$