Linear Transformations

Many abstract mathematical properties may be understood through the analogy of distortion of space. The simplest such distortions are the ones that alter the volumes of "solids" residing in that space by a constant factor.

For example, if $T: \mathbb{R}^3 \mapsto \mathbb{R}^3$ is given by $T(x,y,z) = (ax, by, cz)$ then, as we have seen earlier, T is a stretching distortion that transforms the unit sphere $x^2+y^2+z^2=1$ into the ellipsoid $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$. The distortion T modifies the volume of the sphere $\frac{4\pi}{3}$ to the volume of the ellipsoid $\frac{4\pi}{3}abc$. In fact, if S is the set of all points enclosed by a surface, then $\text{Volume}(T(S)) = abc\,\text{Volume}(S)$, where $T(S)$ is the image of S under T (i.e. the shape of the transformed solid)

Notice that a stretching distortion on a 1-D space $\mathbb{R}$ is necessarily of the form $T: \mathbb{R} \mapsto \mathbb{R}$, $T(x) = mx$ for some constant $m \in \mathbb{R}$. That is $T(x) = mx$ is just the equation of a line with slope m and y-intercept 0.

Observe that for any $\vec{v}, \vec{w} \in \mathbb{R}$ and any constant $a \in \mathbb{R}$

$$T(\vec{v}+\vec{w}) = m(\vec{v}+\vec{w}) = m\vec{v} + m\vec{w} = T(\vec{v}) + T(\vec{w}) \quad (1)$$

$$T(a\vec{v}) = m(a\vec{v}) = a(m\vec{v}) = a\,T(\vec{v}) \quad (2)$$

We are interested in describing all functions $T: \mathbb{R}^n \mapsto \mathbb{R}^m$ that satisfy properties (1) and (2) and make the following definition

Def: Let T be a function from $\mathbb{R}^n$ to $\mathbb{R}^m$, $T: \mathbb{R}^n \mapsto \mathbb{R}^m$, then T is a linear transformation if for any $\vec{v}, \vec{w} \in \mathbb{R}^n$ and any scalar $a \in \mathbb{R}$ we have

Additivity of vectors

$$T(\vec{v}+\vec{w}) = T(\vec{v}) + T(\vec{w}) \qquad (1)$$

Scalar multiplicativity

$$T(a\vec{v}) = a\,T(\vec{v}) \qquad (2)$$

Let's look at a few examples

Ex. Let $T: \mathbb{R}^2 \mapsto \mathbb{R}$ be given by $T(x,y) = 9x - 5y$.

Then T is linear:

Let $\vec{x} = (x_1, x_2)$, $\vec{y} = (y_1, y_2)$, then $T(\vec{x}+\vec{y}) = T((x_1,x_2)+(y_1,y_2)) =$

$= T((x_1+y_1, x_2+y_2)) = 9(x_1+y_1) - 5(x_2+y_2) = 9x_1 - 5x_2 + 9y_1 - 5y_2$

$= T(x_1, x_2) + T(y_1, y_2) = T(\vec{x}) + T(\vec{y})$. Thus, property (1) is satisfied.

To see that property (2) holds as well, take $a \in \mathbb{R}$. Then,

$T(a\vec{x}) = T(a(x_1,x_2)) = T(ax_1, ax_2) = 9(ax_1) - 5(ax_2) =$

$= a(9x_1 - 5x_2) = a\,T(x_1,x_2) = a\,T(\vec{x})$

Ex. Let $T: \mathbb{R}^2 \mapsto \mathbb{R}$ be given by $T(x,y) = \sqrt{x^2+y^2}$. Then T is not linear because it does not satisfy property (1)

Let $\vec{v} = (1,0)$ and $\vec{w} = (0,1)$

Then $T(\vec{v}+\vec{w}) = T(1,1) = \sqrt{1+1} = \sqrt{2}$, while $T(\vec{v}) + T(\vec{w}) =$

$= \sqrt{1^2+0^2} + \sqrt{0^2+1^2} = 1+1 = 2$

Ex. let $T(x,y) = (x+y, 3x, y)$. Then T is linear:

let $\vec{x} = (x_1, x_2)$ $\vec{y} = (y_1, y_2)$, then $T(\vec{x}+\vec{y}) = T((x_1+y_1, x_2+y_2))$

$= (x_1+y_1+x_2+y_2, 3(x_1+y_1), x_2+y_2) = ([x_1+x_2]+[y_1+y_2], 3x_1+3y_1, x_2+y_2)$

$= (x_1+x_2, 3x_1, x_2) + (y_1+y_2, 3y_1, y_2) = T(x_1, x_2) + T(y_1, y_2)$

$= T(\vec{x}) + T(\vec{y})$. This shows that property (1) holds.

let $a \in \mathbb{R}$, then $T(a\vec{x}) = T(a(x_1, x_2)) = T((ax_1, ax_2)) =$

$= (ax_1 + ax_2, 3ax_1, ax_2) = a(x_1+x_2, 3x_1, x_2) = aT(x_1, x_2)$

Hence (2) holds as well.

Comprehension check: Determine, which of the followings are linear transformations. How did you decide?

a) $T(x,y) = xy$

b) $T(x,y,z) = (3x-y, z+x, 4y+3z)$

c) $T(x,y,z) = (x-y, \sqrt{x}, 6z)$

Verifying which maps (transformations) are linear using the direct method above can be rather tedious. A simple theorem will allow us to classify linear maps $T: \mathbb{R}^n \mapsto \mathbb{R}^m$ at a glance.

Lemma 1: let $T: \mathbb{R}^n \mapsto \mathbb{R}^m$ be linear, then $T(x_1, \ldots, x_n) =$

$= (T_1(x_1, \ldots, x_n), \ldots, T_m(x_1, \ldots, x_n))$ where

each $T_i: \mathbb{R}^n \to \mathbb{R}$ is linear, Proof:

Since T is linear, $T(\vec{x}+\vec{y}) = T(\vec{x}) + T(\vec{y})$ for all $\vec{x},\vec{y} \in \mathbb{R}^n$, we have $T(\vec{x}+\vec{y}) = (T_1(\vec{x}+\vec{y}), \dots, T_m(\vec{x}+\vec{y})) =$
$= T(\vec{x}) + T(\vec{y}) = (T_1(\vec{x}), \dots, T_m(\vec{x})) + (T_1(\vec{y}), \dots, T_m(\vec{y})) =$
$= (T_1(\vec{x}) + T_1(\vec{y}), \dots, T_m(\vec{x}) + T_m(\vec{y}))$

In particular, $(T_1(\vec{x}+\vec{y}), \dots, T_m(\vec{x}+\vec{y})) = (T_1(\vec{x}) + T_1(\vec{y}), \dots, T_m(\vec{x}) + T_m(\vec{y}))$

Since two m-tuples are equal iff their components are equal, it follows that $T_i(\vec{x}+\vec{y}) = T_i(\vec{x}) + T_i(\vec{y})$ for $i \in \{1, \dots, m\}$

The proof that $T_i(a\vec{x}) = a T_i(\vec{x})$ holds for $a \in \mathbb{R}$ is similar. (you should supply the proof).

Lemma 2: Let $T: \mathbb{R}^n \to \mathbb{R}$ be linear. Then $T(x_1, \dots, x_n) = \sum_{i=1}^{n} a_i x_i$

Proof: Observe that the vector $\vec{x} = (x_1, \dots, x_n) = x_1\vec{e}_1 + \dots + x_n\vec{e}_n =$
$= \sum_{i=1}^{n} x_i \vec{e}_i$ where $\vec{e}_i$ is the i^{th} coordinate vector.

By properties (1) and (2) $T(x_1, \dots, x_n) = T(x_1\vec{e}_1 + \dots + x_n\vec{e}_n)$
$= T(x_1\vec{e}_1) + \dots + T(x_n\vec{e}_n) = x_1 T(\vec{e}_1) + \dots + x_n T(\vec{e}_n) = \sum_{i=1}^{n} x_i T(\vec{e}_i)$
where $T(\vec{e}_i) \in \mathbb{R}$ (why?). Letting $a_i = T(\vec{e}_i)$ we get that $T(x_1, \dots, x_n) = \sum_{i=1}^{n} x_i a_i = \sum_{i=1}^{n} a_i x_i$ as desired.

As illustration of lemma 1 suppose $T: \mathbb{R}^4 \mapsto \mathbb{R}^3$ is given by $T(x_1, x_2, x_3, x_4) = (x_1+x_2+x_3+x_4,\ 6x_3,\ 9x_2 - x_1)$

Then $T_1(x_1, x_2, x_3, x_4) = x_1+x_2+x_3+x_4$, $T_2(x_1, x_2, x_3, x_4) = 6x_3$

$T_3(x_1, x_2, x_3, x_4) = 9x_2 - x_1$ are linear maps with

$T_i(x_1, x_2, x_3, x_4) \in \mathbb{R}$ as you should verify.

(Lemma 2 makes this verification trivial. Why?)

We may now prove the theorem that characterizes all linear transformations on Euclidian spaces.

Theorem: let $T: \mathbb{R}^n \mapsto \mathbb{R}^m$ be any function. Then T is linear iff $T(x_1, \ldots, x_n) = \left(\sum_{i=1}^{n} a_{1i} x_i, \ldots, \sum_{i=1}^{n} a_{mi} x_i\right)$

proof: By lemma 1, T is linear iff $T(x_1, \ldots, x_n) =$
$= (T_1(x_1, \ldots, x_n), \ldots, T_m(x_1, \ldots, x_n))$ where $T_i: \mathbb{R}^n \mapsto \mathbb{R}$ is linear. By lemma 2 T_i is linear iff $T_i(x_1, \ldots, x_n) = \sum_{j=1}^{n} a_{ij} x_j$

Combining these observations establishes the desired result.

Comprehension check: which of the following functions are Linear maps?

a) $T(x_1, x_2, x_3) = (6x_2 + \pi x_3,\ 0,\ 7x_1 - \frac{1}{6}x_2,\ -2x_2)$

b) $T(x_1, x_2) = (6x_1 - \frac{\sqrt{3}}{2},\ 7x_2,\ 6x_2)$

c) $T(x_1, x_2, x_3) = (x_1 x_2 x_3)\sin(x_1 x_2 x_3)$

d) $T(x_1, x_2) = \big((\sin\theta + \tan\theta)x_1, x_2\big) \quad \theta \in \mathbb{R}$

e) $T(x_1, x_2) = (\sin\theta + x_1, x_2) \quad \theta \in \mathbb{R}$.

Composition of linear maps

Let $T: \mathbb{R}^p \mapsto \mathbb{R}^n$ and $S: \mathbb{R}^n \mapsto \mathbb{R}^m$ be linear maps, then $ST: \mathbb{R}^p \mapsto \mathbb{R}^m$ is a function defined by $S \circ T(x_1, \dots, x_p) = S(T(x_1, \dots, x_p))$.

Ex. Let $T: \mathbb{R}^2 \mapsto \mathbb{R}^3$ be given by $T(x_1, x_2) = (6x_1, x_2 - x_1, 6x_2)$ and $S: \mathbb{R}^3 \mapsto \mathbb{R}$ be given $S(x_1, x_2, x_3) = x_1 + 2x_2 + 3x_3$. Then T and S are linear maps. $ST: \mathbb{R}^2 \mapsto \mathbb{R}$ is given by $ST(x_1, x_2) = S(T(x_1, x_2)) = S(6x_1, x_2 - x_1, 6x_2) =$
$= 6x_1 + 2(x_2 - x_1) + 3(6x_2) = 4x_1 + 20x_2$

Ex. Let $T: \mathbb{R}^3 \mapsto \mathbb{R}^4$ be given by $T(x_1, x_2, x_3) =$
$= (x_1 - 2x_2, \pi x_3, 0, x_1 + x_2 + x_3)$ and $S: \mathbb{R}^4 \mapsto \mathbb{R}^2$ be given by $S(x_1, x_2, x_3, x_4) = (0, x_1 + 2x_2)$. Then $ST: \mathbb{R}^3 \mapsto \mathbb{R}^2$ is given by $ST(x_1, x_2, x_3) = S(T(x_1, x_2, x_3)) =$
$= S(x_1 - 2x_2, \pi x_3, 0, x_1 + x_2 + x_3) = (0, x_1 - 2x_2 + 2(\pi x_3)) =$
$= (0, x_1 - 2x_2 + 2\pi x_3)$

Notice that in both examples ST turned out to be a linear transformation. This is true in general.

Theorem: Let $T: \mathbb{R}^p \mapsto \mathbb{R}^n$ and $S: \mathbb{R}^n \mapsto \mathbb{R}^m$ be linear maps. Then $ST: \mathbb{R}^p \mapsto \mathbb{R}^m$ is linear.

Proof: $ST(\vec{x}+\vec{y}) = S(T(\vec{x}+\vec{y})) = S(T(\vec{x}) + T(\vec{y})) =$

$= S(T(\vec{x})) + S(T(\vec{y})) = ST(\vec{x}) + ST(\vec{y})$ for $\vec{x}, \vec{y} \in \mathbb{R}^p$

Hence additivity holds.

$ST(a\vec{x}) = S(T(a\vec{x})) = S(aT(\vec{x})) = aS(T(\vec{x})) = aST(\vec{x})$

for any $a \in \mathbb{R}$.

Hence scalar multiplicativity holds.

It follows that ST is linear as desired.

Composition of functions can be extended to three or more functions. Furthere more, composition of functions is associative:

Let $T_1: \mathbb{R}^{n_1} \mapsto \mathbb{R}^{n_2}$, $T_2: \mathbb{R}^{n_2} \mapsto \mathbb{R}^{n_3}$, and $T_3: \mathbb{R}^{n_3} \mapsto \mathbb{R}^{n_4}$, then $T_3(T_2T_1)(\vec{x}) = (T_3T_2)T_1(\vec{x})$ for all $\vec{x} \in \mathbb{R}^{n_1}$.

you can convince yourself of this fact by thinking of $T_3T_2T_1$ as a conveyer belt:

$$\vec{x} \mapsto \boxed{T_1} \mapsto T_1(\vec{x}) \mapsto \boxed{T_2} \mapsto T_2(T_1(\vec{x})) \mapsto \boxed{T_3} \downarrow T_3(T_2(T_1(\vec{x})))$$

Where each box represents a machine responsible for part of the assembly.

$(T_3T_2)T_1(\vec{x})$ can be represented thus:

$$\vec{x} \to \boxed{T_1} \mapsto T_1(\vec{x}) \mapsto \boxed{\boxed{T_2} \mapsto \boxed{T_3}} \mapsto (T_3T_2)\,T_1(\vec{x})$$

where the second box represents a single machine with subroutines T_2 and T_3.

$$\vec{x} \mapsto \boxed{\boxed{T_1} \mapsto \boxed{T_2}} \mapsto (T_2T_1)(\vec{x}) \mapsto \boxed{T_3} \mapsto T_3(T_2T_1)(\vec{x})$$

is a representation of $T_3(T_2T_1)(\vec{x})$. In this case, the functions T_1 and T_2 were replaced by a machine that has T_1 & T_2 as subroutines.

Both of the above conveyor belts produce the same end product. In other words, $(T_3T_2)T_1(\vec{x}) = T_3(T_2T_1)(\vec{x})$ for all $\vec{x} \in \mathbb{R}^n$. Since two functions are identical if they agree on every value are equivalent, it follows that $(T_3T_2)T_1 = T_3(T_2T_1)$.

<u>Addition of linear maps</u>

Let $T, S : \mathbb{R}^n \mapsto \mathbb{R}^m$ be linear. Define $T+S : \mathbb{R}^n \mapsto \mathbb{R}^m$ by $(T+S)(\vec{x}) = T(\vec{x}) + S(\vec{x})$ where $\vec{x} \in \mathbb{R}^n$ and $T(\vec{x}), S(\vec{x}) \in \mathbb{R}^m$

You should verify that $T+S$ is linear and that $T+S = S+T$

Ex. let $T : \mathbb{R}^2 \mapsto \mathbb{R}^3$ be given by $T(x_1, x_2) = (6x_1, 6x_1, x_2 - x_1)$ and $S : \mathbb{R}^2 \mapsto \mathbb{R}^3$ be given by $S(x_1, x_2) = (x_1 + 2x_2, -5x_1, 6x_1 + x_2)$

Then $(T+S)(\vec{x}) = (7x_1 + 2x_2,\ x_1,\ 5x_1 + 2x_2)$

Scalar multiplication of a linear map

Let $T: \mathbb{R}^n \mapsto \mathbb{R}^m$ be linear, then for $a \in \mathbb{R}$ define aT by $aT(\vec{x}) = a(T(\vec{x}))$. Clearly, $aT: \mathbb{R}^n \mapsto \mathbb{R}^m$ is linear.

Ex. let $T: \mathbb{R}^2 \mapsto \mathbb{R}^2$ be given by $T(x,y) = (y,-x)$, then $-6T: \mathbb{R}^2 \mapsto \mathbb{R}^2 = (-6y, 6x)$

Matrix representation of linear maps

Recall that any linear map $T: \mathbb{R}^n \mapsto \mathbb{R}^m$ is determined by its action on the coordinate unit vectors $\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n$. That is, $T(x_1, \dots, x_n) = T(x_1\vec{e}_1 + \dots + x_n\vec{e}_n) = x_1 T(\vec{e}_1) + \dots + x_n T(\vec{e}_n)$.
In the last equation, if we were to know the values $T(\vec{e}_1), \dots, T(\vec{e}_n)$, we would have been able to write out the explicit formula for T.

Ex. Suppose $T: \mathbb{R}^3 \mapsto \mathbb{R}^4$ is linear and $T(\vec{e}_1) = T(1,0,0) = (1,2,3,4)$, $T(\vec{e}_2) = T(0,1,0) = (-1,0,3,3)$, and $T(0,0,1) = (6,-1,0,2)$. Then $T(x_1,x_2,x_3) =$

$$= T\big(x_1(1,0,0) + x_2(0,1,0) + x_3(0,0,1)\big) =$$

$$= x_1\, T(1,0,0) + x_2\, T(0,1,0) + x_3\, T(0,0,1) =$$

$$= x_1(1,2,3,4) + x_2(-1,0,3,3) + x_3(6,-1,0,2) =$$

$$= (x_1 - x_2 + 6x_3,\ 2x_1 - x_3,\ 3x_1 + 3x_2,\ 4x_1 + 3x_2 + 2x_3)$$

Comprehenshion check: let $T: \mathbb{R}^2 \mapsto \mathbb{R}^2$ be linear. Suppose $T(\vec{e}_1) = (1,-1)$ and $T(\vec{e}_2) = (0,1)$. Find the explicit formula of T.

Notice that for $T: \mathbb{R}^n \mapsto \mathbb{R}^m$ linear, $T(\vec{e}_i) \in \mathbb{R}^m$ for $i \in \{1,\dots,n\}$. In particular, $T(\vec{e}_i)$ is a vector in $\mathbb{R}^m$. Hence $T(\vec{e}_i) = (a_{1i}, a_{2i}, \dots, a_{mi})$ for a suitable choice of $a_{1i}, \dots, a_{mi} \in \mathbb{R}$. Thus $T(\vec{e}_i) = a_{1i}\vec{f}_1 + a_{2i}\vec{f}_2 + \dots + a_{mi}\vec{f}_m$ where the $\vec{f}_j$ are coordinate vectors in $\mathbb{R}^m$ (just like the $\vec{e}_i$ are coordinate vectors in $\mathbb{R}^n$).

This allows us to 'encode' T in a matrix form:

Def: let $T: \mathbb{R}^n \mapsto \mathbb{R}^m$ be linear, then $M(T)$ is an $m \times n$ matrix

$$\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \quad \text{where} \quad T(e_i) = a_{1i}\vec{f}_1 + \dots + a_{mi}\vec{f}_m.$$

In other words

$$\begin{array}{c c} & \begin{array}{ccc} T(\vec{e}_1), & \dots, & T(\vec{e}_n) \end{array} \\ \begin{array}{c} \vec{f}_1 \\ \vdots \\ \vec{f}_m \end{array} & \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \end{array}$$

represents what happens to each vector $\vec{e}_i$ under T. For instance, $T(\vec{e}_1) = a_{11}\vec{f}_1 + \dots + a_{m1}\vec{f}_m$.

Ex. let $T: \mathbb{R}^2 \mapsto \mathbb{R}^3$ be given by $T(x_1, x_2) = (-x_2, 0, x_1)$.

Then $T(1,0) = (0,0,1)$ and $T(0,1) = (-1,0,0)$.

Thus $T(1,0) = \boxed{0}\cdot(1,0,0) + \boxed{0}\cdot(0,1,0) + \boxed{1}\cdot(0,0,1)$

and $T(0,1) = \boxed{-1}(1,0,0) + \boxed{0}(0,1,0) + \boxed{0}(0,0,1)$

Hence

$$\mathcal{M}(T) = \begin{array}{c|cc} & T(\vec{e}_1) & T(\vec{e}_2) \\ \hline (1,0,0) & 0 & -1 \\ (0,1,0) & 0 & 0 \\ (0,0,1) & 1 & 0 \end{array}$$

Notice that the first column is simply $T(1,0)^T = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$

and that the second column is $T(0,1)^T = \begin{pmatrix} -1 \\ 0 \\ 0 \end{pmatrix}$

Ex. let $T: \mathbb{R}^3 \mapsto \mathbb{R}^4$ be given by

$T(x_1, x_2, x_3) = (x_1 - x_2 + 6x_3,\ 2x_1 - x_3,\ 3x_1 + 3x_2,\ 4x_1 + 3x_2 + 2x_3)$

Then $T(1,0,0) = (1,2,3,4)$, $T(0,1,0) = (-1,0,3,3)$,

and $T(0,0,1) = (6,-1,0,2)$

$$\mathcal{M}(T) = \begin{array}{c|ccc} & T(1,0,0) & T(0,1,0) & T(0,0,1) \\ \hline (1,0,0,0) & 1 & -1 & 6 \\ (0,1,0,0) & 2 & 0 & -1 \\ (0,0,1,0) & 3 & 3 & 0 \\ (0,0,0,1) & 4 & 3 & 2 \end{array}$$

Notice again that the first column is just the transpose of the vector $T(1,0,0)$, the second column is the transpose of $T(0,1,0)$ and the third column is the transpose of $T(0,0,1)$.

In general, if $\{\vec{e}_1, \ldots, \vec{e}_n\}$ are unit (orthogonal) coordinate vectors of $\mathbb{R}^n$ and $\{\vec{f}_1, \ldots, \vec{f}_m\}$ are unit (orthogonal) coordinate vectors of $\mathbb{R}^m$, then the i^{th} column of $M(T)$ is the transpose of $T(\vec{e}_i)$.

To see that this is so, observe that $T(\vec{e}_i)$ is of the form $(a_{1i}, \ldots, a_{mi}) = a_{1i}\vec{f}_1 + \ldots + a_{mi}\vec{f}_m$. Thus, writing this sum in a column

$$\begin{matrix} a_{1i}\vec{f}_1 \\ + \\ \vdots \\ + \\ a_{mi}\vec{f}_m \end{matrix}$$

implies that

$$\begin{array}{c} \\ \vec{f}_1 \\ \vdots \\ \vec{f}_m \end{array} \begin{array}{c} \cdots\ T(\vec{e}_i)\ \cdots \\ \left(\begin{matrix} & a_{1i} & \\ \cdots & \vdots & \cdots \\ & a_{mi} & \end{matrix}\right) \end{array}.$$

When coordinate vectors are involved, finding $M(T)$ is easy:

Ex. Let $T: \mathbb{R}^5 \mapsto \mathbb{R}^4$ be given by

$$T(x_1, x_2, x_3, x_4, x_5) = (\sqrt{2}x_1 - 7x_2 + \pi x_3,\ x_5 - x_4,\ x_1 + 3x_3,\ x_5 + 2x_2)$$

To find the first column, simply locate x_1 together with its coefficient in each of the coordinates of $T(x_1, x_2, x_3, x_4, x_5)$:

$$T(x_1,x_2,x_3,x_4,x_5) = (\boxed{\sqrt{2}\,x_1} - 7x_2 + \pi x_3 + 0\cdot x_4 + 0\cdot x_5,$$
$$\boxed{0\cdot x_1} + 0\cdot x_2 + 0\cdot x_3 - x_4 + x_5,\ \boxed{x_1} + 0\cdot x_2 + 3x_3 + 0\cdot x_4 + 0\cdot x_5,$$
$$\boxed{0\cdot x_1} + 2x_2 + 0\cdot x_3 + 0\cdot x_4 + x_5)$$

Replacing x_1 with 1, we obtain the column vector $\begin{pmatrix}\sqrt{2}\\0\\1\\0\end{pmatrix}$.
This is the first column of $M(T)$.
Doing the same procedure with x_2 in place of x_1, we obtain the second column of $M(T)$ and so on.
In particular, the matrix $M(T)$ is given by

$$M(T) = \begin{pmatrix} \sqrt{2} & -7 & \pi & 0 & 0 \\ 0 & 0 & 0 & -1 & 1 \\ 1 & 0 & 3 & 0 & 0 \\ 0 & 2 & 0 & 0 & 1 \end{pmatrix}$$

<u>Remark</u>: This procedure works only when the coordinate vectors are orthogonal to one another and are of magnitude 1.

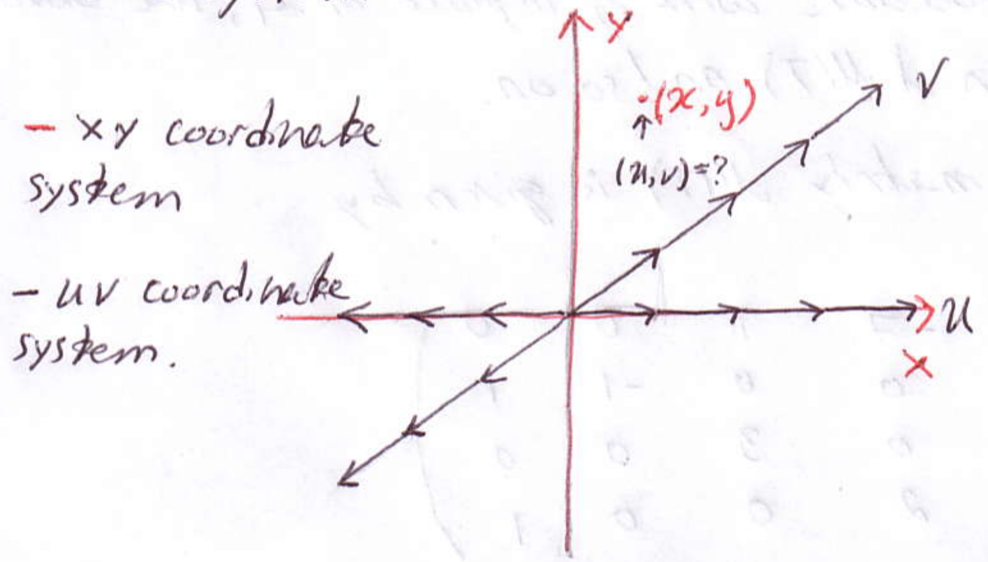
Ex. Let $T: \mathbb{R}^2 \to \mathbb{R}^2$ be given by $T(x,y) = (y,x)$.
What is the matrix of T with respect to the orthogonal coordinates $(1,0), (0,1)$? What is the matrix with respect to the "skewed" coordinate system generated by $(1,0), (1,1)$?

Solution:

with respect to the standard rectangular system generated by $(1,0), (0,1)$ the matrix is

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

To solve for the second matrix, note first the 'skewed' coordinate system



How might we locate the (u,v) coordinate in the uv system that determines the same physical point as the (x,y) coordinate?

$(x,y) = x\vec{e}_1 + y\vec{e}_2 = u\,\vec{e}_1 + v(\vec{e}_1+\vec{e}_2)$

where $\vec{e}_1+\vec{e}_2$ is the vector that generates the v-axis.

thus $(x,y) = x\vec{e}_1 + y\vec{e}_2 = (u+v)\vec{e}_1 + v\vec{e}_2 = (u+v, v)$.

This means that $x = u+v$ and $y = v \Rightarrow x = u+y \Rightarrow$

$\Rightarrow u = x-y$.

In other words, the physical location of a point with (x,y) coordinates is given by the coordinates $(x-y, y)$ in the 'skewed' uv system.

In particular $T(1,0) = (0,1) = -1(1,0) + 1(1,1)$ and

$$T(1,1) = (1,1) = 0\cdot(1,0) + 1(1,1)$$

The corresponding matrix is therefore

$$\begin{array}{cc} & \begin{array}{cc} T(1,0) & T(1,1) \end{array} \\ \begin{array}{c} (1,0) \\ (1,1) \end{array} & \begin{pmatrix} -1 & 0 \\ 1 & 1 \end{pmatrix} \end{array}$$

Notice that the columns are not transposes of $T(1,0)$ and $T(1,1)$

<u>Recovering the linear map from its matrix</u>

Let A be an $m \times n$ matrix. Can we recover the linear map $T: \mathbb{R}^n \mapsto \mathbb{R}^m$ which it represents? In general, we must know the coordinate system on which T acts to make this recovery. In Calc III however, we will always assume that the coordinate systems of $\mathbb{R}^n$ and $\mathbb{R}^m$ are $\{\vec{e}_1, \dots, \vec{e}_n\}$ and $\{\vec{f}_1, \dots, \vec{f}_m\}$ respectively. This will allow us to deduce the linear function $T: \mathbb{R}^n \mapsto \mathbb{R}^m$ from its $m \times n$ matrix A, Ex. let A be the 4×3 matrix

$$\begin{pmatrix} 1 & 6 & 9 \\ -2 & 12 & 18 \\ 3 & 1 & 0 \\ 0 & 1 & 7 \end{pmatrix}$$

What linear map does the matrix encode?

Solution:

From the matrix we know that $T(\vec{e}_1) = T(1,0,0) =$

$= 1(1,0,0,0) - 2(0,1,0,0) + 3(0,0,1,0) + 0(0,0,0,1)$

$= (1,-2,3,0)$.

Similarly, $T(\vec{e}_2) = (6,12,1,1)$, and $T(\vec{e}_3) = (9,18,0,7)$.

Hence $T(x_1,x_2,x_3) = T\left(\sum_{i=1}^{3} x_i \vec{e}_i\right) = \sum_{i=1}^{3} x_i T(\vec{e}_i) =$

$= x_1(1,-2,3,0) + x_2(6,12,1,1) + x_3(9,18,0,7) =$

$= (x_1 + 6x_2 + 9x_3, -2x_1 + 12x_2 + 18x_3, 3x_1 + x_2, x_2 + 7x_3)$

Remark: Although the above solution illustrates the underlying concept, $T(x_1,x_2,x_3)$ can be found very quickly as follows:

$$\begin{pmatrix} 1\,x_1 + 6\,x_2 + 9\,x_3 \\ -2\,x_1 + 12\,x_2 + 18\,x_3 \\ 3\,x_1 + 1\,x_2 + 0\,x_3 \\ 0\,x_1 + 1\,x_2 + 7\,x_3 \end{pmatrix}$$

Thus we can think of the matrix A as a column vector $\begin{pmatrix} x_1+6x_2+9x_3 \\ -2x_1+12x_2+18x_3 \\ 3x_1+x_2 \\ x_2+7x_3 \end{pmatrix}$ whose transpose is the desired linear map.

Comprehension check: let A be the 4×4 matrix

$$\begin{pmatrix} 1 & 0 & 9 & -\pi \\ 1 & 7 & 8 & \sqrt{2} \\ -6 & 0 & 3 & 8 \\ 2 & 10 & 3 & 1 \end{pmatrix}$$

Find the corresponding linear transformation.

Matrix addition

Our goal is to define matrix operations that will allow $\mathcal{M}(T)$ to simulate the linear map T for which it stands. In particular, given linear maps $S, T : \mathbb{R}^n \mapsto \mathbb{R}^m$, we would like to define matrix addition in such a way that

$$\mathcal{M}(T+S) = \mathcal{M}(T) + \mathcal{M}(S).$$

In other words, if T and S are translated (encoded by) into matrix form, there would not be any need to compute $T+S$ and then encode it in matrix form $\mathcal{M}(T+S)$. Instead, the matrix $\mathcal{M}(T+S)$ would be obtained by simply adding the matrices that we already have.

Before proving the general case, let's see an example.

Ex. let $T, S: \mathbb{R}^3 \mapsto \mathbb{R}^2$ be given by $T(x_1, x_2, x_3) = (x_1 + x_2, x_3)$ and $S(x_1, x_2, x_3) = (-x_3, x_1 + 3x_2)$. Then

$$\mathcal{M}(T) = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \text{ and } \mathcal{M}(S) = \begin{pmatrix} 0 & 0 & -1 \\ 1 & 3 & 0 \end{pmatrix}$$

Since $(T+S)(x_1, x_2, x_3) = (x_1 + x_2 - x_3, x_1 + 3x_2 + x_3)$,

$$\mathcal{M}(T+S) = \begin{pmatrix} 1 & 1 & -1 \\ 1 & 3 & 1 \end{pmatrix} =$$

$$= \begin{pmatrix} 1+0 & 1+0 & 0+(-1) \\ 0+1 & 0+3 & 1+0 \end{pmatrix}$$

If we define matrix addition coordinatewise, we get

$\mathcal{M}(T+S) = \mathcal{M}(T) + \mathcal{M}(S)$

This suggests the following proposition.

Proposition: Let $T, S: \mathbb{R}^n \mapsto \mathbb{R}^m$ be linear. Suppose

$$\mathcal{M}(T) = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \text{ and } \mathcal{M}(S) = \begin{pmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & & \vdots \\ b_{m1} & \cdots & b_{mn} \end{pmatrix}$$

then the matrix of $T+S$ is the $m \times n$ matrix

$$\mathcal{M}(T+S) = \begin{pmatrix} a_{11}+b_{11} & \cdots & a_{1n}+b_{1n} \\ \vdots & & \vdots \\ a_{m1}+b_{m1} & \cdots & a_{mn}+b_{mn} \end{pmatrix}$$

Proof: The i^{th} column of $\mathcal{M}(T+S)$ is determined from $(T+S)(\vec{e_i}) = T(\vec{e_i}) + S(\vec{e_i}) = \sum_{j=1}^{m} a_{ji}\vec{f_j} + \sum_{j=1}^{m} b_{ji}\vec{f_j}$ (1)

$= \sum_{j=1}^{m}(a_{ji}+b_{ji})\vec{f_j}$ where $\begin{pmatrix} a_{1i} \\ a_{2i} \\ \vdots \\ a_{mi} \end{pmatrix}$ is the i^{th} column of

$\mathcal{M}(T)$ and $\begin{pmatrix} b_{1i} \\ b_{2i} \\ \vdots \\ b_{mi} \end{pmatrix}$ is the i^{th} column of $\mathcal{M}(S)$

Equation (1) suggests that $\begin{pmatrix} a_{1i}+b_{1i} \\ a_{2i}+b_{2i} \\ \vdots \\ a_{mi}+b_{mi} \end{pmatrix}$ is the i^{th} column

of $\mathcal{M}(T+S)$. The result follows as desired.

Def: let A, B be $m\times n$ matrices. Define $A+B$ to be the $m\times n$ matrix, whose ijth entry, $(A+B)_{ij} = A_{ij} + B_{ij}$.
That is, if

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \text{ and } B = \begin{pmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & & \vdots \\ b_{m1} & \cdots & b_{mn} \end{pmatrix}$$

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

with this definition $\mathcal{M}(T+S) = \mathcal{M}(T) + \mathcal{M}(S)$. for any pair of linear transformations $T, S: \mathbb{R}^n \mapsto \mathbb{R}^m$.

<u>Matrix of scalar multiplication</u>

let $T: \mathbb{R}^n \mapsto \mathbb{R}^m$. Then for any constant $a \in \mathbb{R}$
aT is a linear map from $\mathbb{R}^n$ to $\mathbb{R}^m$. We want to define the product of a scalar and a matrix in such a way that
$\mathcal{M}(aT) = a\mathcal{M}(T)$
you should prove the following proposition.

Proposition: let $T: \mathbb{R}^n \mapsto \mathbb{R}^m$ be linear and suppose that

$$\mathcal{M}(T) = \begin{pmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & & \vdots \\ b_{m1} & \cdots & b_{mn} \end{pmatrix}. \text{ Then } \mathcal{M}(aT) = \begin{pmatrix} ab_{11} & \cdots & ab_{1n} \\ \vdots & & \vdots \\ ab_{m1} & \cdots & ab_{mm} \end{pmatrix}$$

for $a \in \mathbb{R}$.

Consequently we define a scalar-matrix multiplication as follows:

Def: Let B be an $m \times n$ matrix and $a \in \mathbb{R}$. Then aB is the matrix whose ij^{th} entry, $(aB)_{ij} = a B_{ij}$. That is, if

$$B = \begin{pmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & & \vdots \\ b_{m1} & \cdots & b_{mn} \end{pmatrix} \Rightarrow aB = \begin{pmatrix} ab_{11} & \cdots & ab_{1n} \\ \vdots & & \vdots \\ ab_{m1} & \cdots & ab_{mn} \end{pmatrix}$$

Multiplication of matrices

Let $S: \mathbb{R}^n \mapsto \mathbb{R}^m$ and $T: \mathbb{R}^p \mapsto \mathbb{R}^n$ be linear. We would like to define matrix multiplication in such a way that $\mathcal{M}(ST) = \mathcal{M}(S)\mathcal{M}(T)$, where, naturally, $\mathcal{M}(ST)$ is an $m \times p$ matrix and $\mathcal{M}(S)$ is an $m \times n$ matrix ($\mathcal{M}(T)$ is $n \times p$ matrix)
We will need the following proposition.

Proposition: let $S: \mathbb{R}^n \mapsto \mathbb{R}^m$ and $T: \mathbb{R}^p \mapsto \mathbb{R}^n$ be linear. Suppose

$$\mathcal{M}(S) = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}, \text{ and } \mathcal{M}(T) = \begin{pmatrix} b_{11} & \cdots & b_{1p} \\ \vdots & & \vdots \\ b_{n1} & \cdots & b_{np} \end{pmatrix}.$$

Then $\mathcal{M}(ST)$ is an $m \times p$ matrix $\begin{pmatrix} c_{11} & \cdots & c_{1p} \\ \vdots & & \vdots \\ c_{m1} & \cdots & c_{mp} \end{pmatrix}$

where $c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$

Proof: Let $\{\vec{e}_1, \dots, \vec{e}_p\}$, $\{\vec{f}_1, \dots, \vec{f}_n\}$, and $\{\vec{g}_1, \dots, \vec{g}_m\}$ be the coordinate vectors of $\mathbb{R}^p$, $\mathbb{R}^n$, and $\mathbb{R}^m$ respectively.

Then the j^{th} column of $\mathcal{M}(ST)$ is given by (deduced from)

$$ST(\vec{e}_j) = S(T(\vec{e}_j)) = S\left(\sum_{k=1}^{n} b_{kj} \vec{f}_k\right) = \sum_{k=1}^{n} b_{kj} S(\vec{f}_k) =$$

$$= \sum_{k=1}^{n} b_{kj} \left(\sum_{i=1}^{m} a_{ik} \vec{g}_i \right) = \sum_{i=1}^{m} \left(\sum_{k=1}^{n} a_{ik} b_{kj} \right) \vec{g}_i$$

In particular, the j^{th} column of $\mathcal{M}(ST)$ is

$$\begin{pmatrix} \sum_{k=1}^{n} a_{1k} b_{kj} \\ \sum_{k=1}^{n} a_{2k} b_{kj} \\ \vdots \\ \sum_{k=1}^{n} a_{mk} b_{kj} \end{pmatrix}$$

Thus, in order to make $\mathcal{M}(ST) = \mathcal{M}(S)\mathcal{M}(T)$ work, we need to define matrix multiplication like we do below

Def: Let A be an $m \times n$ matrix and let B be an $n \times p$ matrix. Then AB is the $m \times p$ matrix with

$$(AB)_{ij} = (A_{i1}, \ldots, A_{in}) \cdot (B_{1j}, \ldots, B_{nj})$$

<u>Another way to write the formula of a linear map</u>

Let $T: \mathbb{R}^n \mapsto \mathbb{R}^m$ be a linear map with $m \times n$ matrix

$$\mathcal{M}(T) = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}.$$ Recall that this representation

means that $T(\vec{e}_1) = (a_{11}, \ldots, a_{m1}), \ldots, T(\vec{e}_n) = (a_{1n}, \ldots, a_{mn})$. In other words, $\mathcal{M}(T) = \left([T(\vec{e}_1)]^T \ldots [T(\vec{e}_n)]^T \right)$ where $[T(\vec{e}_i)]^T$ is the transpose of $T(\vec{e}_i)$. In particular, the ith column of $\mathcal{M}(T)$ lists the components $a_{1i}, a_{2i}, \ldots, a_{mi}$ of the vector $T(\vec{e}_i) = a_{1i}\vec{f}_1 + a_{2i}\vec{f}_2 + \ldots + a_{mi}\vec{f}_m \in \mathbb{R}^m$

Let's define a matrix $\mathcal{M}(\vec{v})$ that lists the components of $\vec{v}$ in terms of the coordinate vectors.

Def: Let $\vec{v} \in \mathbb{R}^n$ $\vec{v} = v_1\vec{e}_1 + v_2\vec{e}_2 + \ldots + v_n\vec{e}_n$ then $\mathcal{M}(\vec{v}) = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}$

Ex. If $\vec{v} = (1, -3, 0, 7)$ $\mathcal{M}(\vec{v}) = \begin{pmatrix} 1 \\ -3 \\ 0 \\ 7 \end{pmatrix}$.

Notice that, whereas $\vec{v}$ is a vector, $\mathcal{M}(\vec{v})$ is a matrix.

Also observe that $\mathcal{M}(\vec{v})$ can be interpreted (with a slight violation of the distinction between vectors and matrices) as $\vec{v}^T$.

We will sometimes refer to $\mathcal{M}(\vec{v})$ as a column vector.

Let $(x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$. If $T: \mathbb{R}^n \mapsto \mathbb{R}^m$ is a linear map with matrix $\mathcal{M}(T)$ given above, then $\mathcal{M}(T)$ and $\mathcal{M}(x_1, x_2, \ldots, x_n) = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ can be multiplied by one another.

What might be the relationship between $\mathcal{M}(T(x_1, x_2, \ldots, x_n))$ and $\mathcal{M}(T)\,\mathcal{M}(x_1, \ldots, x_n)$?

(Notice that $\mathcal{M}(T(x_1, \ldots, x_n))$ is an $m \times 1$ matrix). The answer is supplied by the following proposition.

Proposition: Let $T: \mathbb{R}^n \mapsto \mathbb{R}^m$ be a linear map with $m \times n$ matrix $\mathcal{M}(T) = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$ Let $\vec{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$

Then $\mathcal{M}(T(\vec{x})) = \mathcal{M}(T)\,\mathcal{M}(\vec{x})$

Proof: $T(\vec{x}) = T(x_1, x_2, \ldots, x_n) = T\left(\sum_{i=1}^{n} x_i \vec{e_i}\right) =$

$= \sum_{i=1}^{n} x_i T(\vec{e_i}) = \sum_{i=1}^{n} x_i (a_{1i}, a_{2i}, \ldots, a_{mi}) =$

$= \sum_{i=1}^{n} (a_{1i}x_i, a_{2i}x_i, \ldots, a_{mi}x_i) =$

$= (a_{11}x_1, a_{21}x_1, \ldots, a_{m1}x_1) + (a_{12}x_2, a_{22}x_2, \ldots, a_{m2}x_2) + \ldots$

$+ (a_{1n}x_n, a_{2n}x_n, \ldots, a_{mn}x_n) = \left(\sum_{i=1}^{n} a_{1i}x_i, \sum_{i=1}^{n} a_{2i}x_i, \ldots, \sum_{i=1}^{n} a_{mi}x_i\right)$

$$\text{Thus } M(T(\vec{x})) = \begin{bmatrix} \sum_{i=1}^{n} a_{1i}x_i \\ \sum_{i=1}^{n} a_{2i}x_i \\ \vdots \\ \sum_{i=1}^{n} a_{mi}x_i \end{bmatrix}$$

clearly this is just the matrix $M(T)M(\vec{x})$.

<u>What does all of this mean?</u>

Ex. let $T: \mathbb{R}^3 \mapsto \mathbb{R}^2$ be given by $T(x_1, x_2, x_3) = (5x_1 + x_2, x_3)$

then $M(T(x_1, x_2, x_3)) = \begin{pmatrix} 5x_1 + x_2 \\ x_3 \end{pmatrix}$, $M(T) = \begin{pmatrix} 5 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

and $\mathcal{M}(x_1, x_2, x_3) = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$.

By the proposition above

$$\mathcal{M}(T(x_1, x_2, x_3)) = [T(x_1, x_2, x_3)]^T = \mathcal{M}(T)\,\mathcal{M}(x_1, x_2, x_3) =$$

$$= \mathcal{M}(T)\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 5 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.$$

In other words $[T(\vec{x})]^T = \begin{pmatrix} 5 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}[\vec{x}]^T$

This is the reason that many books simply say $T = \mathcal{M}(T)$.

That is if $A = \mathcal{M}(T)$ then a common statement is

$$T\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = A\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

<u>The distributive law of matrix multiplication</u>

Let A be an $m \times n$ matrix and let B, C be $n \times p$ matrices.

Observe that $A(B+C) = AB + AC$ because

$$[A(B+C)]_{ij} = (\text{row } i \text{ of } A) \cdot (\text{Column } j \text{ of } B+C) =$$

$$= (A_{i1}, A_{i2}, \ldots, A_{in}) \cdot ((B+C)_{1j}, (B+C)_{2j}, \ldots, (B+C)_{nj})$$

$$= (A_{i1}, A_{i2}, \ldots, A_{in}) \cdot [(B_{1j}, B_{2j}, \ldots, B_{nj}) + (C_{1j}, C_{2j}, \ldots, C_{nj})] =$$

$= (A_{i1}, A_{i2}, \dots, A_{in}) \cdot (B_{1j}, B_{2j}, \dots, B_{nj}) +$

$+ (A_{i1}, A_{i2}, \dots, A_{in}) \cdot (C_{1j}, C_{2j}, \dots, C_{nj}) =$

$= (\text{row } i \text{ of } A) \cdot (\text{column } j \text{ of } B) + (\text{row } i \text{ of } A) \cdot (\text{column } j \text{ of } C) =$

$= [AB + AC]_{ij}$

Thus, if $S: \mathbb{R}^n \mapsto \mathbb{R}^m$ and $T, P: \mathbb{R}^l \mapsto \mathbb{R}^n$ are linear then $\mathcal{M}(S(T+P)) = \mathcal{M}(ST + SP) = \mathcal{M}(ST) + \mathcal{M}(SP)$.

A similar result holds when A, B are $m \times n$ matrices and C is an $n \times p$ matrix (you should verify this).

<u>The range of a linear map</u>

Does the equation $x^2 = 4$ have a solution? Sure! $x = \pm 2$ are solutions. How about $x^2 = -4$? Clearly there are no real solutions. Generally, we say that if $f: \mathcal{U} \subset \mathbb{R} \mapsto \mathbb{R}$ is a function with domain $\mathcal{U}$, then the range of f is the set of all $y \in \mathbb{R}$ such that $f(x) = y$ for some $x \in \mathcal{U}$. We generalize this below

Def: Let $f: \mathcal{U} \subset \mathbb{R}^n \mapsto \mathbb{R}^m$ be a function, then the range of f, $R(f)$, is the set $\{\vec{y} \in \mathbb{R}^m ; f(\vec{x}) = \vec{y} \text{ for some } \vec{x} \in \mathcal{U}\}$

Can anything meaningful be said about ranges of linear maps?

Let $T: \mathbb{R}^n \mapsto \mathbb{R}^m$ be linear. Obviously $R(T) = \{T(\vec{x}) : \vec{x} \in \mathbb{R}^n\}$

That is every 'point' $\vec{y} \in \mathbb{R}^m$ is of the form $\vec{y} = T(\vec{x})$ for some $\vec{x} \in \mathbb{R}^n$.

Observe yet again that $T(\vec{x}) = T(x_1, \dots, x_n) =$

$= T(\sum_{i=1}^{n} x_i \vec{e_i}) = \sum_{i=1}^{n} x_i T(\vec{e_i})$ replacing each x_i with parameter t_i we see that the range of T, $R(T)$, is the set of all points with parameters $\sum_{i=1}^{n} t_i T(\vec{e_i})$.

Ex. let $T: \mathbb{R}^3 \mapsto \mathbb{R}^2$ be given by $T(x_1, x_2, x_3) =$

$= (x_1 + x_2, x_3)$. Then $T(1,0,0) = (1,0)$ and $T(0,1,0) = (1,0)$

and $T(0,0,1) = (0,1)$.

The range of T $R(T) = \{x_1 T(1,0,0) + x_2 T(0,1,0) + x_3 T(0,0,1)\}$

$= \{t_1(1,0) + t_2(1,0) + t_3(0,1) ; t_1, t_2, t_3 \in \mathbb{R}\} =$

$= \{(t_1+t_2)(1,0) + t_3(0,1) ; t_1, t_2, t_3 \in \mathbb{R}\} = \{s(1,0) + t(0,1) ; s, t \in \mathbb{R}\}$

Note that $R(T)$ is the set generated by the coordinate vectors $(1,0), (0,1)$. Hence $R(T) = \mathbb{R}^2$. In other words T is onto.

Ex. Let $T: \mathbb{R}^4 \mapsto \mathbb{R}^3$ be given by $T(x_1, x_2, x_3, x_4) =$

$= (x_1 + 2x_2,\ x_1 + 2x_2 + x_3,\ 0)$, then $R(T) = \{T(\vec{x}) : \vec{x} \in \mathbb{R}^4\}$

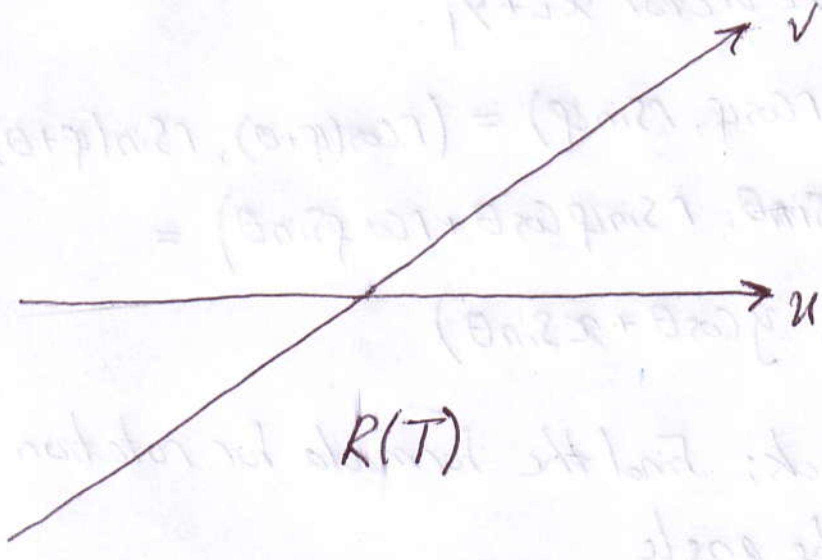
$= \{(x_1 + 2x_2,\ x_1 + 2x_2 + x_3,\ 0) : x_1, x_2, x_3 \in \mathbb{R}\} =$

$= \{(x_1, x_1, 0) + (2x_2, 2x_2, 0) + (0, x_3, 0) : x_1, x_2, x_3 \in \mathbb{R}\}$

$= \{r(1,1,0) + s(1,1,0) + t(0,1,0) : r, s, t \in \mathbb{R}\} =$

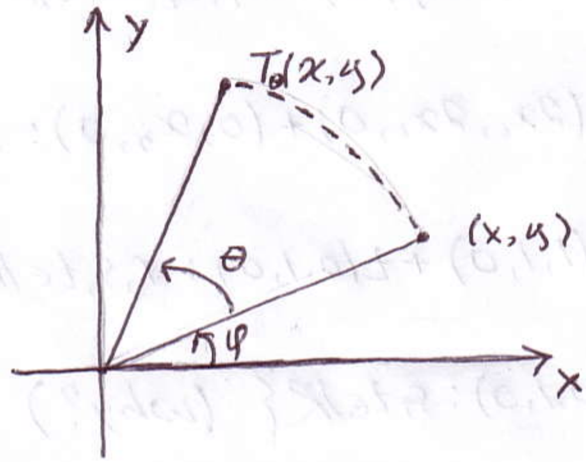
$= \{s(1,1,0) + t(0,1,0) : s, t \in \mathbb{R}\}$ (why?) Thus $R(T)$ is

the set of all points on the plane $s(1,1,0) + t(0,1,0)$.
This plane is spanned by vectors $(1,1,0)$ and $(0,1,0)$



Linear maps and rotation of space

The rotation of a point (x,y) in $\mathbb{R}^2$ counterclockwise about the origin by an angle θ is a linear transformation.



Let $(x,y) = (r\cos\varphi, r\sin\varphi)$ where $r = \sqrt{x^2+y^2}$ and φ is the angle from the x-axis to the vector $x\vec{i}+y\vec{j}$.

Then $T_\theta(x,y) = T_\theta(r\cos\varphi, r\sin\varphi) = \big(r\cos(\varphi+\theta), r\sin(\varphi+\theta)\big) =$

$= \big(r\cos\varphi\cos\theta - r\sin\varphi\sin\theta,\ r\sin\varphi\cos\theta + r\cos\varphi\sin\theta\big) =$

$= \big(x\cos\theta - y\sin\theta,\ y\cos\theta + x\sin\theta\big)$

Comprehension check: Find the formula for rotation about the origin by the angle

a) $\frac{\pi}{2}$

b) π.

Transformations of this form allow us to rotate graphs of functions.

In other words, if $f: \mathbb{R} \mapsto \mathbb{R}$ is a function, then $T_\theta(x, f(x))$ rotates $(x, f(x))$ by an angle θ, thus the image of $\{(x, f(x)) : x \in \mathbb{R}\}$ under T_θ, $T_\theta(\{(x, f(x)) : x \in \mathbb{R}\})$ is the graph of f rotated by θ.

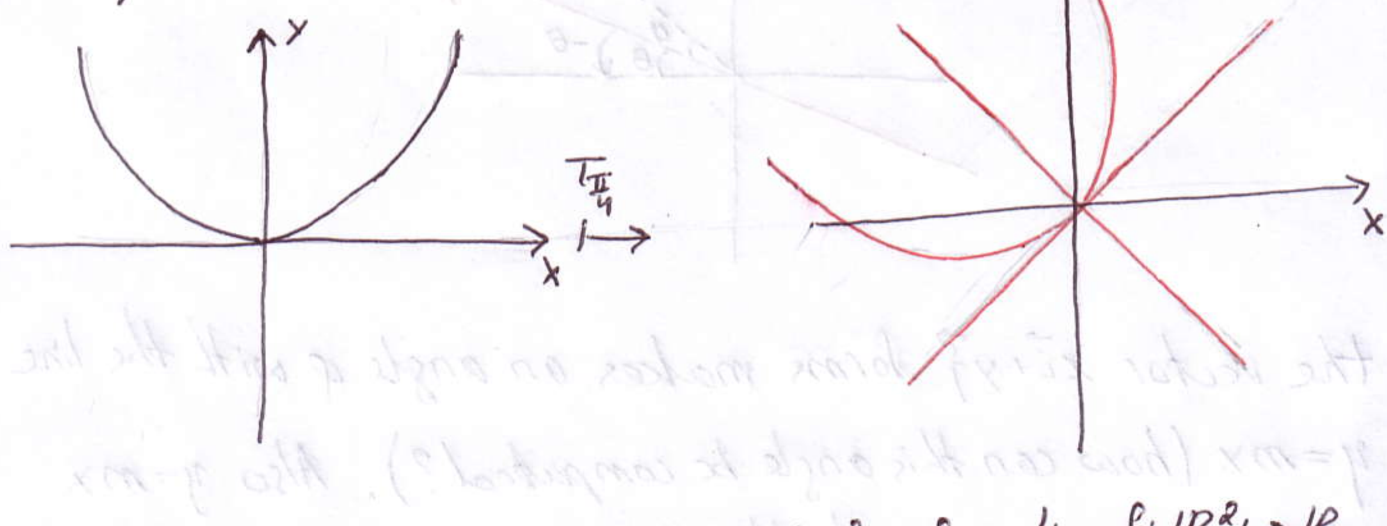
Ex. What is the image of the graph of $f(x) = x^2$ under the transformation $T_{\frac{\pi}{4}}$.

Solution: $T_{\frac{\pi}{4}}(x,y) = \left(x \cos\frac{\pi}{4} - y \sin\frac{\pi}{4},\ y \cos\frac{\pi}{4} + x \sin\frac{\pi}{4}\right) =$

$= \left(x\frac{1}{\sqrt{2}} - y\frac{1}{\sqrt{2}},\ y\frac{1}{\sqrt{2}} + x\frac{1}{\sqrt{2}}\right) = \frac{1}{\sqrt{2}}(x-y, x+y)$

Thus $T_{\frac{\pi}{4}}(\{(x, x^2) : x \in \mathbb{R}\}) = \left\{\left(\frac{x - x^2}{\sqrt{2}}, \frac{x + x^2}{\sqrt{2}}\right) : x \in \mathbb{R}\right\}$

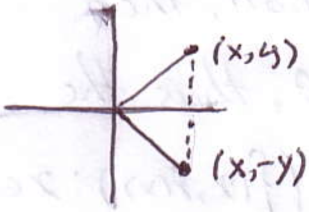
Graphically



Comprehension check: Given a graph of a function $f: \mathbb{R}^2 \mapsto \mathbb{R}$, what linear map $T_\theta: \mathbb{R}^3 \mapsto \mathbb{R}^3$ would rotate this graph by θ about the z-axis? Find an explicit formula.

Linear maps and reflections

Consider reflecting (x,y) in the x-axis

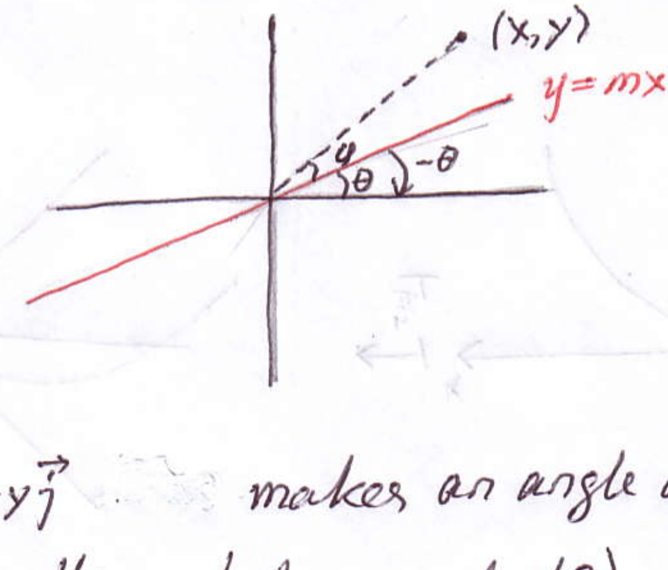


$$S(x,y) = (x,-y)$$

Clearly reflection through the x-axis is a linear map.

Comprehension check: If S is the linear map given by $S(x,y) = (x,-y)$, draw $S(\{(x,x^2) : x \in \mathbb{R}\})$. If $L : \mathbb{R}^2 \mapsto \mathbb{R}^2$ is a reflection through the y-axis find the explicit formula of L. Also draw $L(\{(x,x^2) : x \in \mathbb{R}\})$.

Generally, to reflect through the line $y = mx$ the following strategy may be implemented:



the vector $x\vec{i} + y\vec{j}$ makes an angle φ with the line $y = mx$ (how can this angle be computed?). Also $y = mx$ makes an angle θ with the x-axis. Thus, by preforming a rotation about the origin by $-\theta$, then reflecting by $-\varphi$ and then rotating back by θ, the desired reflection is achieved.

In other words, $T_\theta S T_{-\theta}$ is the transformation that performs the reflection through the line $y=mx$.

Invertible linear transformations

Recall that a function $f: \mathbb{R} \mapsto \mathbb{R}$ is invertible if there is some function $f^{-1}: \mathbb{R} \mapsto \mathbb{R}$ s.t. $f(f^{-1}(x)) = x$ and $f^{-1}(f(x)) = x$ for all $x \in \mathbb{R}$.

Generally, a function $f: \mathbb{R}^n \mapsto \mathbb{R}^m$ is invertible if there exists a function $f^{-1}: \mathbb{R}^m \mapsto \mathbb{R}^n$ s.t. $f^{-1}(f(\vec{x})) = \vec{x} \in \mathbb{R}^n$ and $f(f^{-1}(\vec{y})) = \vec{y} \in \mathbb{R}^m$.

Just like any other function, a linear map is invertible iff it is one-to-one and onto. Although it is beyond the scope of our lectures to prove this rigorously, no linear map $T: \mathbb{R}^m \mapsto \mathbb{R}^n$ is one-to-one if $m>n$ and no linear map $S: \mathbb{R}^m \mapsto \mathbb{R}^n$ is onto if $m<n$. It follows that, in order to be invertible, a linear map T must take $\mathbb{R}^n \mapsto \mathbb{R}^n$.

Def: Let $T: \mathbb{R}^n \mapsto \mathbb{R}^n$ be a linear map. Then T is invertible if there exists a linear function $T^{-1}: \mathbb{R}^n \mapsto \mathbb{R}^n$ such that $T(T^{-1}(\vec{x})) =$ $= T^{-1}(T(\vec{x})) = \vec{x}$ for any $\vec{x} \in \mathbb{R}^n$.

Ex. Let $I: \mathbb{R}^3 \mapsto \mathbb{R}^3$ be defined by $I(x_1, x_2, x_3) = (x_1, x_2, x_3)$. Then I is called the identity map (the function that does nothing). Since $I(I(\vec{x})) = \vec{x}$ for all $\vec{x}$. It follows that $I^{-1}=I$. We define $I_3=\mathcal{M}(I)$. Note that

$$\mathcal{M}(I)=\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix}=I_3.$$

In general, $I:\mathbb{R}^n\mapsto\mathbb{R}^n$ given by $I(\vec{x})=\vec{x}$ is invertible with $I^{-1}=I$. In that case, we define $I_n=\mathcal{M}(I)$. Note that I_n is the identity matrix with 1 on the diagonal and 0 everywhere else. In other words $(I_n)_{ij}=\begin{cases}1 \text{ if } i=j\\0 \text{ if } i\neq j\end{cases}$.

The identity map I allows us to restate the definition of invertibility as follows.

Def. $T:\mathbb{R}^n\mapsto\mathbb{R}^n$ is invertible if there exists $T^{-1}:\mathbb{R}^n\mapsto\mathbb{R}^n$ such that $TT^{-1}=T^{-1}T=I$ (why?)

Ex. Let T_θ be the linear map that rotates every point (x,y) by an angle θ about the origin. Then $T_\theta^{-1}=T_{-\theta}$ (why?).

Comprehenshion check: Let $S(x,y)=(x,-y)$. Find S^{-1}.

Ex. Let $L:\mathbb{R}^3\mapsto\mathbb{R}^3$ be given by $L(x,y,z)=(z,-x,y)$ Find L^{-1}.

Solution: $(x,y,z)=L^{-1}L(x,y,z)=L^{-1}(L(x,y,z))=L^{-1}(z,-x,y)$ letting $u=z$, $v=-x$, $w=y$, observe that $L^{-1}(z,-x,y)=(x,y,z)$ becomes $L^{-1}(u,v,w)=(-v,w,u)$. It follows that L^{-1} takes the first coordinate to the third, multiplies the second coordinate by -1 and moving in in front, and moves the third coordinate to second place. Thus $L^{-1}(x,y,z) = (-y, z, x)$.

Comprehenshion check: Let $L: \mathbb{R}^3 \mapsto \mathbb{R}^3$ be given by $L(x,y,z) = (2y, -z, \frac{1}{3}x)$. Find L^{-1}.

It is generally not so easy to find the inverse function of an invertible linear map $T: \mathbb{R}^n \mapsto \mathbb{R}^n$. It is useful to use the matrix representation of T, $\mathcal{M}(T)$:

Observe that if $T^{-1}: \mathbb{R}^n \mapsto \mathbb{R}^n$ is the inverse of $T: \mathbb{R}^n \mapsto \mathbb{R}^n$, then $\mathcal{M}(I) = \mathcal{M}(T^{-1}T) = \mathcal{M}(T^{-1})\mathcal{M}(T)$.

Similarly $\mathcal{M}(I) = \mathcal{M}(TT^{-1}) = \mathcal{M}(T)\mathcal{M}(T^{-1})$.

In other words

$$I_n = \mathcal{M}(T)\ \mathcal{M}(T^{-1}) = \mathcal{M}(T^{-1})\mathcal{M}(T).$$

This motivates the following definition:

Def: Let A be an $n \times n$ matrix. Then A is invertible if there is an $n \times n$ matrix B s.t. $AB = BA = I_n$. We call B "A^{-1}".

With this definition, we can say that for invertible $T: \mathbb{R}^n \mapsto \mathbb{R}^n$, $\mathcal{M}(T^{-1})$, the matrix of the inverse of T, is the inverse of the matrix of T. In symbols, $\mathcal{M}(T^{-1}) = [\mathcal{M}(T)]^{-1}$.

Because $T: \mathbb{R}^n \mapsto \mathbb{R}^n$ may be associated with $\mathcal{M}(T)\mathcal{M}(x)$. (In other words, $T(x) = \mathcal{M}(T)\mathcal{M}(x)$, if notation is abused a little), we can find T^{-1} by calculating $[\mathcal{M}(T)]^{-1}$!

As we will soon see, $\mathcal{M}(T)$ is invertible iff $\det(\mathcal{M}(T)) \neq 0$

Let $A = \mathcal{M}(T)$, then the inverse of A may be obtained using the theorem below:

Theorem: An $n \times n$ matrix A is invertible if and only if $\det(A) \neq 0$. When $\det(A) = |A| \neq 0$, the inverse is given by

$$A^{-1} = \frac{1}{|A|} \begin{pmatrix} (-1)^{1+1}|\overline{A}_{11}| & (-1)^{1+2}|\overline{A}_{12}| & \cdots & (-1)^{1+n}|\overline{A}_{1n}| \\ (-1)^{2+1}|\overline{A}_{21}| & (-1)^{2+2}|\overline{A}_{22}| & \cdots & (-1)^{2+n}|\overline{A}_{2n}| \\ \vdots & \vdots & & \vdots \\ (-1)^{n+1}|\overline{A}_{n1}| & (-1)^{n+2}|\overline{A}_{n2}| & \cdots & (-1)^{n+n}|\overline{A}_{nn}| \end{pmatrix}^T$$

where $\overline{A}_{ij}$ is the $(n-1)\times(n-1)$ matrix obtained by deleting row i and column j from A.

Ex. Find the inverse of $A = \begin{pmatrix} 6 & 1 \\ 3 & 1 \end{pmatrix}$

Solution: $\det(A) = |A| = 6 \cdot 1 - 1 \cdot 3 = 3 \neq 0$ hence A is invertible.

$$A^{-1} = \frac{1}{|A|} \begin{pmatrix} (-1)^{1+1}|\overline{A}_{11}| & (-1)^{1+2}|\overline{A}_{12}| \\ (-1)^{2+1}|\overline{A}_{21}| & (-1)^{2+2}|\overline{A}_{22}| \end{pmatrix}^T$$

Now $\overline{A}_{11} = \begin{pmatrix} \cancel{6} & \cancel{1} \\ \cancel{3} & 1 \end{pmatrix} = 1$, hence $|\overline{A}_{11}| = 1$, similarly

$\bar{A}_{12} = 3$ hence $|\bar{A}_{12}| = \det(\bar{A}_{12}) = 3$, $|\bar{A}_{21}| = 1$, and $|\bar{A}_{22}| = 6$

Thus $A^{-1} = \frac{1}{|A|}\begin{pmatrix} |\bar{A}_{11}| & -|\bar{A}_{12}| \\ -|\bar{A}_{21}| & |\bar{A}_{22}| \end{pmatrix}^T = \frac{1}{3}\begin{pmatrix} 1 & -3 \\ -1 & 6 \end{pmatrix}^T =$

$= \frac{1}{3}\begin{pmatrix} 1 & -1 \\ -3 & 6 \end{pmatrix}$.

Ex. Suppose $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is such that $|A| \neq 0$.

Find A^{-1}.

Solution: $|\bar{A}_{11}| = d$, $|\bar{A}_{12}| = c$, $|\bar{A}_{21}| = b$, $|\bar{A}_{22}| = a$

Hence $A^{-1} = \frac{1}{|A|}\begin{pmatrix} d & -c \\ -b & a \end{pmatrix}^T = \frac{1}{ad-bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$.

Ex. Calculate the inverse of $A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 2 & 3 \end{pmatrix}$.

Solution: $|A| = -\begin{vmatrix} 0 & 1 \\ 1 & 3 \end{vmatrix} = 1$,

$\bar{A}_{11} = \begin{pmatrix} 0 & 1 \\ 2 & 3 \end{pmatrix}$, $\bar{A}_{12} = \begin{pmatrix} 0 & 1 \\ 1 & 3 \end{pmatrix}$, ... $\bar{A}_{33} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$

so $|\bar{A}_{11}| = -2$, $|\bar{A}_{12}| = -1$, ..., $|\bar{A}_{33}| = 0$

Hence $A^{-1} = \frac{1}{1}\begin{pmatrix} (-2) & -(-1) & 0 \\ -3 & 0 & -(-1) \\ 1 & 0 & 0 \end{pmatrix}^T = \begin{pmatrix} -2 & -3 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$

Ex. Show that $(AB)^{-1} = B^{-1}A^{-1}$, where A and B are any two invertible matrices.

Solution:

$$(B^{-1}A^{-1})(AB) = ((B^{-1}A^{-1})A)B = (B^{-1}(A^{-1}A))B =$$

$$= (B^{-1}I_n)B = B^{-1}B = I_n.$$

where we know that $B^{-1}I_n = B^{-1}$ because, for any $n\times n$ matrix C, $I_nC = CI_n = C$. This can be verified by direct computation or as follows:

$C = \mathcal{M}(T)$ for some $T: \mathbb{R}^n \mapsto \mathbb{R}^n$ and $I_n = \mathcal{M}(I)$.

Hence $C = \mathcal{M}(T) = \mathcal{M}(TI) = \mathcal{M}(T)\,\mathcal{M}(I) = CI_n$.

Also $C = \mathcal{M}(T) = \mathcal{M}(IT) = \mathcal{M}(I)\mathcal{M}(T) = I_nC$.

Ex. Find the inverse of $H: \mathbb{R}^2 \mapsto \mathbb{R}^2$ where $H(x,y) = (x-y, 2x+6y)$

Solution: We may think of $H(x,y)$ as $\mathcal{M}(H)\mathcal{M}(x,y) =$

$= \begin{pmatrix} 1 & -1 \\ 2 & 6 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}$. Notice that $|\mathcal{M}(H)| = 1\cdot 6 - (-1)\cdot 2 = 8 \neq 0$

so $\mathcal{M}(H)$ is invertible with inverse $[\mathcal{M}(H)]^{-1} = \frac{1}{8}\begin{pmatrix} 6 & 1 \\ -2 & 1 \end{pmatrix}$.

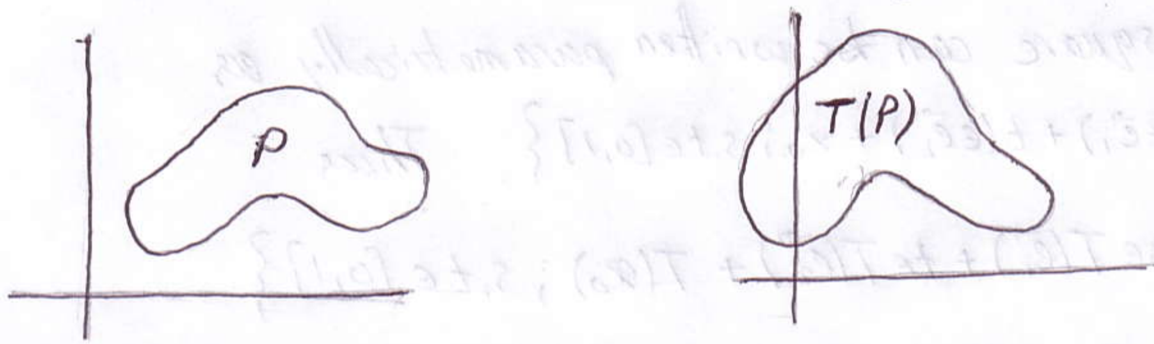
Hence $\mathcal{M}(H^{-1})\mathcal{M}(x,y) = [\mathcal{M}(H)]^{-1}\begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{8}\begin{pmatrix} 6 & 1 \\ -2 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}$

In particular, $H^{-1}(x,y) = \frac{1}{8}(6x+y, -2x+y)$.

Geometry of linear transformations

Area and Volume Properties

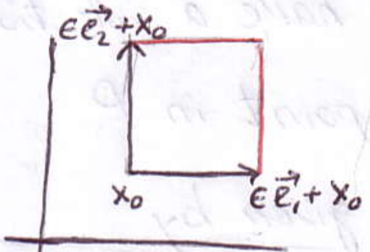
Given a linear transformation $T:\mathbb{R}^n \mapsto \mathbb{R}^n$ and $\vec{x} \in \mathbb{R}^n$ $T(\vec{x})$ is just another point in $\mathbb{R}^n$. If we have a collection P of points in $\mathbb{R}^n$ and apply T to each point in P, then we obtain a new collection of points given by $T(P) = \{T(\vec{x}) : \vec{x} \in P\}$ called the image of P under T.

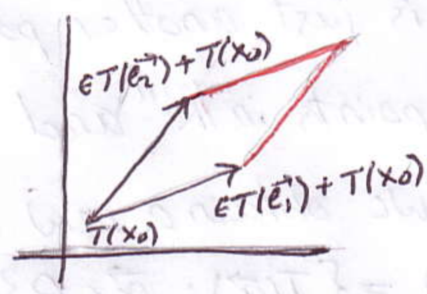


Ex. Let $T:\mathbb{R}^2 \mapsto \mathbb{R}^2$ be linear. Suppose $P = \{\vec{x} ; \vec{x} = s\vec{a} + t\vec{b} + \vec{x}_0,\ 0 \le s \le 1,\ 0 \le t \le 1\}$. Then $T(P) = \{T(\vec{x}) ; \vec{x} = s\vec{a} + t\vec{b} + x_0\}$ $= \{T(s\vec{a} + t\vec{b} + \vec{x}_0) ; t \in [0,1]\ s \in [0,1]\} = \{sT(\vec{a}) + tT(\vec{b}) + T(x_0) ; s,t \in [0,1]\}$. In other words, the image of a parallelogram under T is itself a parallelogram.

Comprehension check: let $T:\mathbb{R}^3 \mapsto \mathbb{R}^3$ be linear. What is $T(P)$ for $P = \{\vec{x} ; \vec{x} = s\vec{a} + t\vec{b} + u\vec{c} + \vec{x}_0,\ s,t,u \in [0,1]\}$?

To understand the relationship between the volume of an arbitrary 'solid' P and the volume of $T(P)$ consider $T: \mathbb{R}^2 \mapsto \mathbb{R}^2$. What does T do to a square of area ϵ^2?





Any such square can be written parametrically as
$S = \{ s(\epsilon\vec{e}_1) + t(\epsilon\vec{e}_2) + x_0 ; s,t \in [0,1] \}$. Thus
$T(S) = \{ s\epsilon T(\vec{e}_1) + t\epsilon T(\vec{e}_2) + T(x_0) ; s,t \in [0,1] \}$

The area of $T(S)$ is the area of a parallelogram with sides $\epsilon T(\vec{e}_1)$ and $\epsilon T(\vec{e}_2)$. In other words, area of $T(S)$ =

$$\left| \det \begin{pmatrix} \epsilon T(\vec{e}_1) \\ \epsilon T(\vec{e}_2) \end{pmatrix} \right|$$

Since the determinant of a matrix is the determinant of the transpose of the matrix we see that

$$\text{area of } T(S) = \epsilon^2 \det \begin{pmatrix} T(\vec{e}_1) \\ T(\vec{e}_2) \end{pmatrix} = \epsilon^2 \det \left([T(\vec{e}_1)]^T \ [T(\vec{e}_2)]^T \right)$$

$$= \left| \det(M(T)) \right| \text{Area}(S).$$

If P is an arbitrary region, P can be approximated by pixels of area ϵ^2, consequently, $T(P)$ can be approximated by pixels of the form $T(S)$ with area $\det(M(T))\epsilon^2$.

Remark: If $T(\vec{e}_1)$ and $T(\vec{e}_2)$ are linearly dependent, i.e. if $c_1 T(\vec{e}_1) + c_2 T(\vec{e}_2) = \vec{0}$ where not all the c_i are 0, $T(S)$ will be a line segment, with 0 area. In particular, $\det(M(T)) = 0$. This implies that T is not invertible, because several different points are collapsed into a single point. In other words, T is not invertible because T is not 1-1.

Generally, the same line of reasoning will show that if $T: \mathbb{R}^n \mapsto \mathbb{R}^n$ is linear and P is a solid in $\mathbb{R}^n$, then the volume of $T(P) = \det(M(T))\, Vol(P)$.

In particular, linear maps stretch space uniformly, modifying the volume by the constant $\det(M(T))$

Comprehension check: What happens if, for $T: \mathbb{R}^n \mapsto \mathbb{R}^n$, $T(\vec{e}_1), \dots, T(\vec{e}_n)$ are linearly dependent? How does it affect $\det(M(T))$?

The determinant of a product of two matrices (optional)

Let A, B be $n \times n$ matrices. What might be the relationship between $\det(AB)$ and $\det(A)$, $\det(B)$?

We may think of A^T as $M(T)$ for some linear map T.
(T would be defined by setting $T(\vec{e_i}) = (a_{1i}, a_{2i}, \ldots, a_{ni})$ for $i \in \{1, \ldots, n\}$, where $(a_{1i}, a_{2i}, \ldots, a_{ni})$ is the transpose of the i^{th} column of A^T).

The matrix B may be thought of as a representation of an n-D parallelepiped whose columns are the vectors that span it.

By the work done before, $\det(T(B))$ = volume of transformed parallelepiped $= \det(M(T)) \, Vol(B) = \det(M(T)) \det(B) =$
$= \det(A^T) \det(B) = \det(A) \det(B)$.

But $\det(AB) = \det(T(B)) = \det(A) \det(B)$

Thus the determinant of a product is the product of the determinants.

Proof of the inverse of matrix theorem (optional)

First, we will need to establish Cramer's rule, which is a procedure for solving a system of linear equations in n unknowns:

Def: A system of n equations in n unknowns is a collection of simultaneous equations of the form

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n &= b_n \end{aligned}$$

Observe that this system of equations can be written compactly as $A\vec{x}^T = \vec{b}$ where

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & & a_{nn} \end{pmatrix} \quad \vec{x}^T = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \quad \vec{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$$

Ex. the collection of simultaneous equations

$$\begin{aligned} 5x_1 - 3x_2 &= 1 \\ 6x_1 + x_2 &= -2 \end{aligned}$$

is of the form $A\vec{x}^T = \vec{b}$

where $A = \begin{pmatrix} 5 & -3 \\ 6 & 1 \end{pmatrix}$, $\vec{x}^{\,T} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ and $\vec{b}^{\,T} = \begin{pmatrix} 1 \\ -2 \end{pmatrix}$

Comprehension check: $\begin{cases} x_1 + 3x_2 = 1 \\ 6x_1 - 7x_2 + x_3 = 3 \\ x_2 + 3x_3 = 0 \end{cases}$ Find $A, \vec{x}^{\,T}, \vec{b}^{\,T}$.

Theorem (Cramer's rule): Let $A\vec{x}^{\,T} = \vec{b}^{\,T}$ be a system of n equations in n unknowns. Then, if $\det(A) \neq 0$, the system has a unique solution $\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ where $x_i = \dfrac{\det(B_i)}{\det(A)}$, B_i being the matrix of A with its i^{th} column replaced by $\vec{b}^{\,T}$.

For example, the system

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 &= b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 &= b_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 &= b_3 \end{aligned}$$

has the solution $x_1 = \dfrac{\begin{vmatrix} b_1 & a_{12} & a_{13} \\ b_2 & a_{22} & a_{23} \\ b_3 & a_{32} & a_{33} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}}$, $x_2 = \dfrac{\begin{vmatrix} a_{11} & b_1 & a_{13} \\ a_{21} & b_2 & a_{23} \\ a_{31} & b_3 & a_{33} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}}$

and $x_3 = \dfrac{\begin{vmatrix} a_{11} & a_{12} & b_1 \\ a_{21} & a_{22} & b_2 \\ a_{31} & a_{32} & b_3 \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}}$

Proof: In expanded form, the equations are

$$a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1 \quad (1)$$
$$a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2 \quad (2)$$
$$\vdots$$
$$a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n = b_n \quad (n)$$

multiplying (1) by $(-1)^{1+1}|\bar{A}_{11}|$, (2) by $(-1)^{2+1}|\bar{A}_{21}|$, ... (n) by $(-1)^{n+1}|\bar{A}_{n1}|$ and adding we get

$$\left(a_{11}(-1)^{1+1}|\bar{A}_{11}| + a_{21}(-1)^{2+1}|\bar{A}_{21}| + \ldots + a_{n1}(-1)^{n+1}|\bar{A}_{n1}|\right)x_1 +$$
$$+\left(a_{12}(-1)^{1+1}|\bar{A}_{11}| + a_{22}(-1)^{2+1}|\bar{A}_{21}| + \ldots + a_{n2}(-1)^{n+1}|\bar{A}_{n1}|\right)x_2 +$$
$$+ \ldots + \left(a_{n1}(-1)^{1+1}|\bar{A}_{11}| + a_{n2}(-1)^{2+1}|\bar{A}_{21}| + \ldots + a_{nn}(-1)^{n+1}|\bar{A}_{n1}|\right)x_n$$
$$= b_1(-1)^{1+1}|\bar{A}_{11}| + b_2(-1)^{2+1}|\bar{A}_{21}| + \ldots + b_n(-1)^{n+1}|\bar{A}_{n1}|$$

This long equation reduces to

$|A|x_1 = |B_1|$, or $x_1 = \dfrac{|B_1|}{|A|}$ (the coefficients of $x_2, x_3, \ldots, x_n$ sum to 0. the coefficient of x_2, for instance is $|C|$ where C is the $n \times n$ matrix obtained from A by replacing the first column of A by the second column of A).

A similar procedure shows that $x_i = \dfrac{|B_i|}{|A|}$.

To prove the inverse theorem, we'll need the following lemma.

Lemma: Let A be $m\times n$ matrix and B - an $n\times p$ matrix with columns $b_1^T, \dots, b_p^T$.

Then $AB = A(b_1^T b_2^T \dots b_p^T) = (Ab_1^T\ Ab_2^T \dots Ab_p^T)$.

In other words, the columns of AB are just the products Ab_i^T where b_i^T is the $n\times 1$ matrix whose entries are the column i entries of B.

Proof: A moment of thought will convince you of the validity of the lemma.

Inverse theorem: Let A be an $n\times n$ matrix with $\det A \neq 0$.

Then A is invertible with inverse $A^{-1} = (A^{-1})_{ij}$. The ij^{th} entry is $(-1)^{i+j} \dfrac{|\overline{A}_{ij}{}^T|}{|A|}$

Proof: A is invertible iff there exists a matrix $X = \begin{pmatrix} x_{11} & \dots & x_{1n} \\ \vdots & & \vdots \\ x_{n1} & \dots & x_{nn} \end{pmatrix}$ such that $AX = I_n$. In particular, $AX = (Ax_1^T\ Ax_2^T \dots Ax_n^T) = (e_1^T e_2^T \dots e_n^T)$. Thus $x_{ij} = \dfrac{|B_i(e_j^T)|}{|A|}$ where $B_i(e_j^T)$ is the matrix of A with its i^{th} column replaced by the transpose of e_j. By the work done on cofactor expansion, $|B_i(e_j^T)| = (-1)^{i+j}|\overline{A}_{ji}|$

where $\overline{A}_{ji}$ is the $(n-1)\times(n-1)$ matrix whose j^{th} row and i^{th} column were deleted. It follows that the transpose of the matrix

$$\frac{1}{|A|}\left((-1)^{i+j}|\overline{A}_{ij}|\right)$$ is the desired matrix X.