CHAPTER VII: LINEAR TRANSFORMATIONS

We have spoken somewhat mysteriously in Chapter I about matrix multiplication "doing something" to vectors. This chapter will make this idea explicit, with the definition of a function. A linear transformation is a particular kind of functoon, that will turn out to be matrix multiplicatoon.

Section B will discuss linear transformations of particular interest in geometry, what are called <u>rigid motions</u>. Section C returns to the difference equations of Chapter IC, along with a few other examples:

p 560

SECTION $\underline{\text{VII}}$ A: LINEAR TRANS- FORMATIONS and MATRICES

After we've identified a linear transformation as a particular type of function, the important result of this section is identifying a matrix for each linear transformation, that performs

p. 561

the transformation with matrix multiplication.

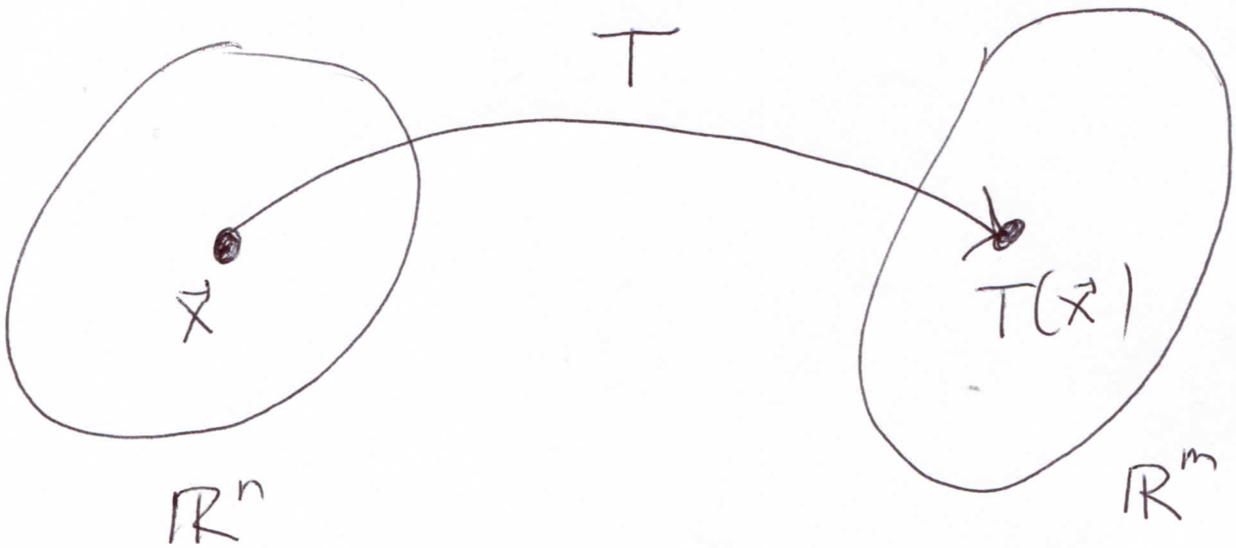
DEFINITION 7.1

A <u>function</u> is a set of instructions describing how to modify each of a set of points. More precisely, a **function f, from $\mathbb{R}^n$ to $\mathbb{R}^m$**, denoted

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

p. 562

is a map that associates to each $\vec{x}$ in $\mathbb{R}^n$ a unique $T(\vec{x})$ (reads "T of $\vec{x}$") in $\mathbb{R}^m$.



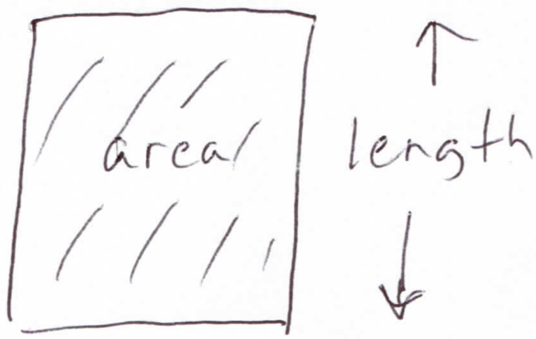
Example 7.2 If you're considering buying a square plot of land, you'd like to know its area.

Area is hard to measure, or even estimate, directly; you'd have to use a large collection of $(1\text{ ft}) \times (1\text{ ft})$ concrete squares and see how many are required to cover the plot of land.

But the length of a side of your square plot is easy to estimate: pace it off with your own two feet.

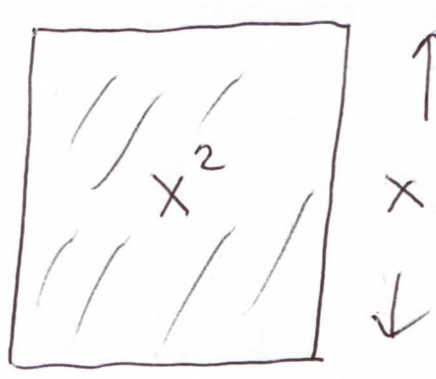
If you possess special knowledge such as

$$(\text{area of square}) = (\text{length of side})^2$$



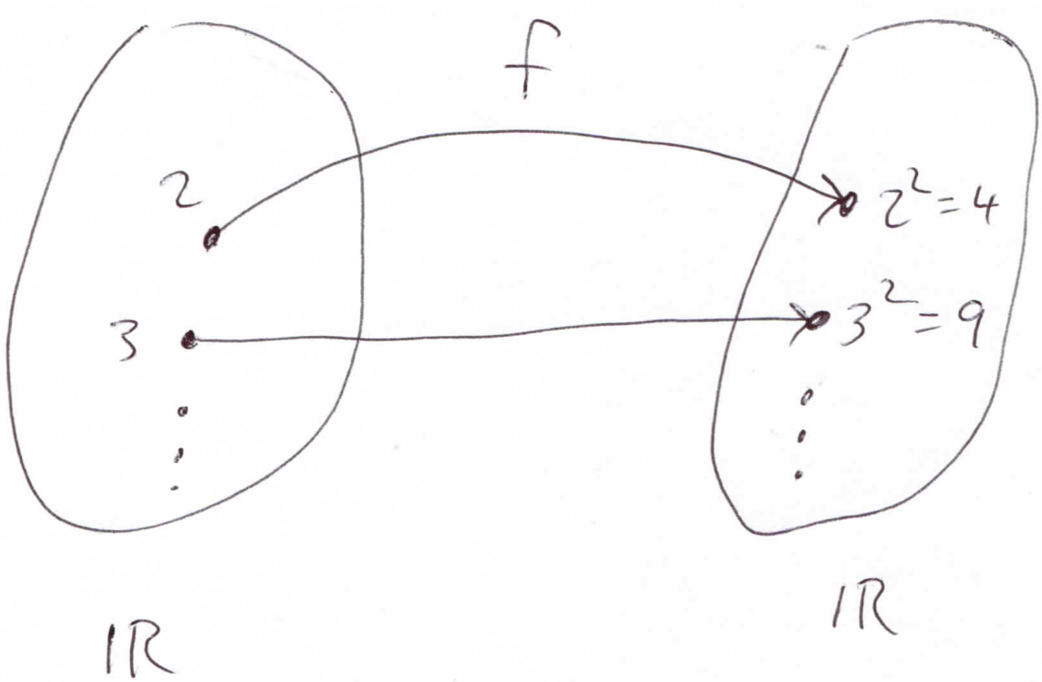
then the function

$$f(x) \equiv x^2$$



is of great interest;
f is a function that
associates something easy
to measure (length) with
something you want (area).

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



p.566

DEFINITION 7.3

A linear transformation from $\mathbb{R}^n \to \mathbb{R}^m$

is a function

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

such that

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

and

(2) $T(c\vec{u}) = cT(\vec{u})$,

for any $\vec{u}, \vec{v}$ in $\mathbb{R}^n$, real c.

Examples 7.4

(1) $T\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) \equiv \begin{bmatrix} x_1 + 2x_2 \\ 3x_1 + 4x_2 \\ 5x_1 + 6x_2 \end{bmatrix}$

is a function from $\mathbb{R}^2$ to $\mathbb{R}^3$ that can be shown (see Theorem 7.8) to be a linear transformation.

(2) $T(x) \equiv x^2$, from Example 7.2, is <u>not</u> linear, since

$$T(x+y) = (x+y)^2 = x^2 + 2xy + y^2$$
$$\neq x^2 + y^2 = T(x) + T(y),$$ unless

x or y is zero.

DISCUSSION 7.5

It can be shown (see Theorem 7.8) that the only function from $\mathbb{R}$ to $\mathbb{R}$ that's linear is

$$T(x) \equiv ax \quad (x \text{ in } \mathbb{R}),$$

for some fixed real number a.

This is related to the fact that the only subspaces of $\mathbb{R}^2$, besides $\{\vec{0}\}$ and $\mathbb{R}^2$, are of the form

$$\{(x, ax) \mid x \text{ is real}\};$$

that is, lines thru the origin.

Notice that, in Examples 7.4(1), that

$$T\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix};$$

This sort of representation — multiplication by a matrix — will turn out to be possible for any linear transformation. Let's give it a name.

p. 570

DEFINITION 7.6

If A is an $(m \times n)$ matrix,

define

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

by $T_A(\vec{x}) \equiv A \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$.

$A \equiv$ the **standard matrix** of T_A.

Example 7.7.

Let $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$. Then

$T_A : \mathbb{R}^3 \to \mathbb{R}^2$, with

$$T_A(x_1, x_2, x_3) = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

$$= \begin{bmatrix} x_1 + 2x_2 + 3x_3 \\ 4x_1 + 5x_2 + 6x_3 \end{bmatrix}.$$

Notice that A is (2×3);

in Definition 7.6, $n = 3$, $m = 2$.

THEOREM 7.8

A function $T: \mathbb{R}^n \to \mathbb{R}^m$ is linear if and only if

$$T = T_A ,$$

for some matrix A.

DISCUSSION 7.9

The interesting part of Theorem 7.8 is, given T, constructing the standard matrix A.

A clue may be found from the <u>standard basis</u>

$\vec{e}_1 \equiv (1, 0, 0, \ldots)$, $\vec{e}_2 \equiv (0, 1, 0, 0, \ldots)$, $\ldots$; see Definition 4.32.

Notice that

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \vec{e}_1 \equiv \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

$$= \begin{bmatrix} 1 \\ 4 \end{bmatrix}; \quad \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \vec{e}_2 =$$

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 5 \end{bmatrix}, \quad \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \vec{e}_3 = \begin{bmatrix} 3 \\ 6 \end{bmatrix}$$

In general, we may recover a matrix one column at a time by multiplying it by $\vec{e}_1, \vec{e}_2, \ldots$

LEMMA 7.10 If A is an $(m \times n)$ matrix, then, for $1 \le j \le n$

$$A\vec{e}_j = \begin{pmatrix} j^{th} \text{ column} \\ \text{of } A \end{pmatrix}$$

This will tell us how to construct the standard matrix A in Theorem 7.8

Proof of Theorem 7.8:

If A is a matrix, then 1.15 implies that T_A is linear.

Converrely, if $T : \mathbb{R}^n \to \mathbb{R}^m$ is linear, define

$$A \equiv \left[T(\vec{e}_1) \;\; T(\vec{e}_2) \;\; \cdots \;\; T(\vec{e}_n) \right];$$

that is, for $1 \leq j \leq n$, the j^{th} column of A is $T(\vec{e}_j)$.

For any $\vec{x} = (x_1, x_2, \ldots, x_n)$ in $\mathbb{R}^n$,

$$A \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = x_1 T(\vec{e}_1) + x_2 T(\vec{e}_2) + \cdots + x_n T(\vec{e}_n)$$

$=$ (by linearity) $T(x_1 \vec{e}_1 + x_2 \vec{e}_2 + \cdots)$

$= T(\vec{x})$. Thus $T = T_A$.

p. 576

Examples 7.11

For each of the following, find the standard matrix for T.

(1) $T(x_1, x_2, x_3, x_4) =$
$(x_1 - 2x_3 + x_4,\ x_2 + 3x_3 + 5x_4)$.

$T(\vec{e}_1) = T(1,0,0,0) = (1,0)$

$T(\vec{e}_2) = T(0,1,0,0) = (0,1)$

$T(\vec{e}_3) = (-2,3)$

$T(\vec{e}_4) = (1,5)$

Make $T(\vec{e}_1), T(\vec{e}_2), \ldots$ into

columns $\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} -2 \\ 3 \end{bmatrix}, \begin{bmatrix} 1 \\ 5 \end{bmatrix},$

then merge:

$$A = \begin{bmatrix} 1 & 0 & -2 & 1 \\ 0 & 1 & 3 & 5 \end{bmatrix}$$

It must be said that this standard matrix could've been constructed without Theorem 7.8, by writing all components as rows of a matrix:

$$T(\vec{x}) = \begin{bmatrix} x_1 - 2x_3 + x_4 \\ x_2 + 3x_3 + 5x_4 \end{bmatrix}$$

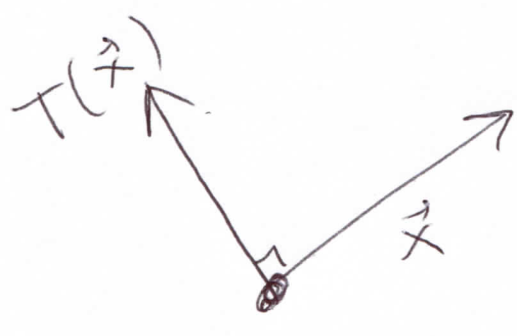
$$= \underbrace{\begin{bmatrix} 1 & 0 & -2 & 1 \\ 0 & 1 & 3 & 5 \end{bmatrix}}_{A} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix},$$

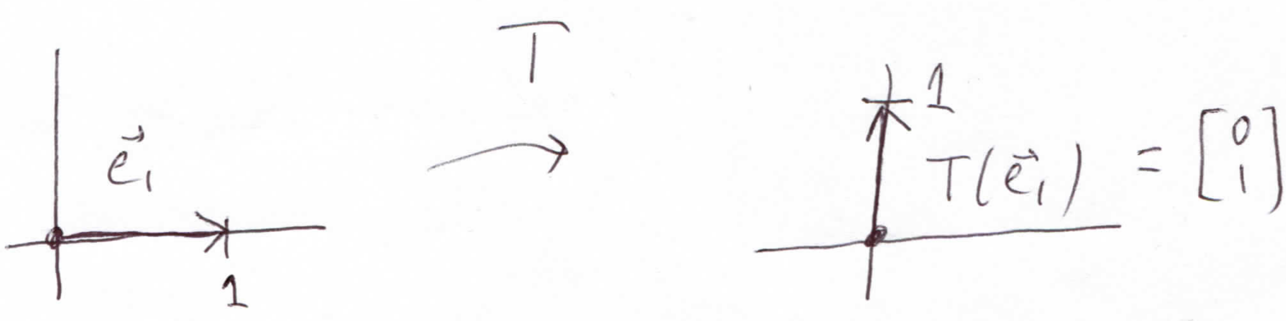
pulling off coefficients as in Gauss-Jordan elimination.

p. 578

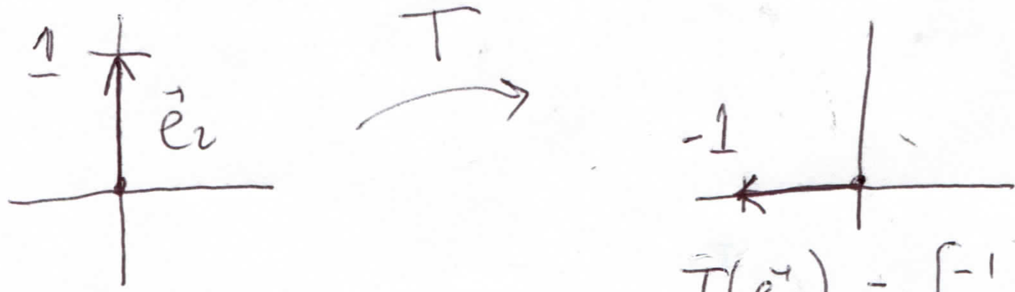
Here is an example where we really need Theorem 7.8.

(2) Define $T: \mathbb{R}^2 \to \mathbb{R}^2$ as counterclockwise rotation by 90 degrees; that is, $T(\vec{x})$ is $\vec{x}$ rotated 90 degrees counterclockwise.





p. 579



$$T(\vec{e}_2) = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$$

$$\to A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

Thus, IF T is linear, then

$$T(x_1, x_2) = A \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -x_2 \\ x_1 \end{bmatrix},$$

for any real x_1, x_2. See Exs. 1.25 (2).

Having obtained **A** by the wishful thinking of T being linear, let's show directly that $T(\vec{x}) = A\vec{x}$,

for any $\vec{x}$ in $\mathbb{R}^2$; that is,

we need to show that A

rotates $\vec{x}$ 90 degrees

counterclockwise when we multiply by A.

Notice first that

$$(x_1, x_2) \cdot \left(A \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) =$$

$$(x_1, x_2) \cdot (-x_2, x_1) = 0; \text{ thus}$$

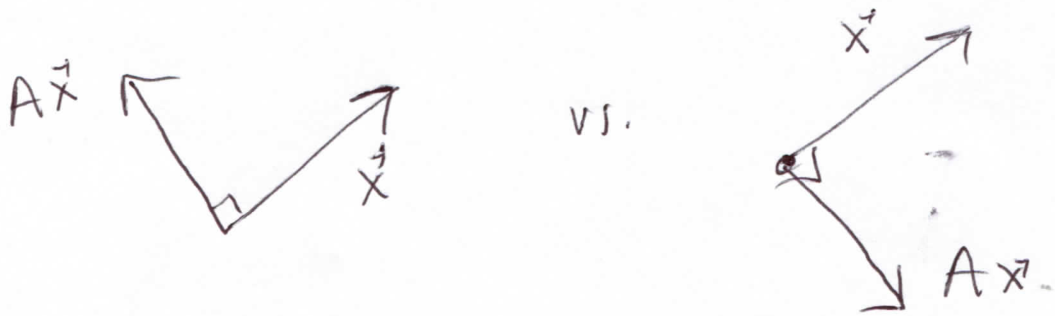
$(A\vec{x}) \perp \vec{x}$, for all $\vec{x}$ in $\mathbb{R}^2$.

Also notice that $\left\| A \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \right\| =$

$$\|(-x_2, x_1)\| = \sqrt{x_2^2 + x_1^2} = \|(x_1, x_2)\|$$

p. 581

We have shown that multiplication by A is a 90 degree rotation; all that remains is to verify the rotation is counterclockwise:



We saw a similar duality in the direction of the <u>cross product</u>; see the picture, after the "right-hand rule."

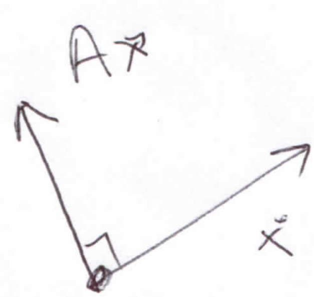
p582

Motivated by this, let's calculate, for $\vec{x} \equiv (x_1, x_2, 0)$,

$$\vec{x} \times (A\vec{x}) = \det \begin{bmatrix} \vec{i} & \vec{j} & \vec{k} \\ x_1 & x_2 & 0 \\ -x_2 & x_1 & 0 \end{bmatrix}$$

$$= (x_1^2 + x_2^2)\vec{k} = (0, 0, (x_1^2 + x_2^2));$$

since $(x_1^2 + x_2^2) > 0$, the "right-hand rule" or its subsequent discussion imply $(A\vec{x})$ is counterclockwise from $\vec{x}$, as desired.



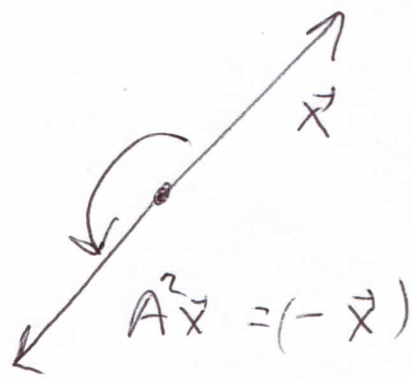
p. 583

Notice that

$$A^2 = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} = -I,$$

a rotation of $180 = 90 \times 2$

degrees;



$$A^3 = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} = -A,$$

rotating 270 degrees

counterclockwise, while

p. 584

$$A^4 = I,$$

a rotation of 360 degrees,

$$\vec{x} = A^4 \vec{x}$$

as in the mean joke "I used to be clueless, but now I'm turned around 360 degrees".