Chapter 3—Linear Transformations (§3.1)

Friday, February 6, 2026 10:31 am

§ 3.1 Linear Transformations

Story.

We now introduce linear transformations
the objects that we shall study in most of the remainder
of this book (course?).

The reader may find it helpful to read (or reread)
the discussion of functions in the Appendix
(since this chapter uses that terminology
freely)

Me: I haven't yet, but we will stop and revisit it
if something becomes unclear as we read.

Let us start with the definition of a linear transformation

Definition.

Let V and W be vector spaces over the field F.

A linear transformation from V into W
is a function T from V to W such that

$$T(c\alpha + \beta) = c(T\alpha) + T\beta$$

for all α and β in V and all scalars c in F.

Story.

Why this definition will hopefully become clear.

Let us start by looking at some examples that satisfy this definition.

Example 1.

If V is any vector space
the idendity transformation, I
defined by $I\alpha = \alpha$
is a linear transformation from V to V.

The **zero transformation** 0
defined by $0\alpha = 0$
is a linear transformation from
V into V.

Example 2.

Let F be a field and let V be the space of polynomial functions
f from F to F, given by

$$f(x) = c_0 + c_1 x + \cdots + c_k x^k$$

Let

$$(Df)(x) \coloneqq c_1 + 2c_2 x + \cdots + k c_k x^{k-1}.$$

Then D is a linear transformation from V to V
(the differentiation transformation).

Example 3.

Let A be a fixed
$m \times n$ matrix with entries in some field F.

The function T defined by
$T(X) = AX$
is a linear transformation from $F^{n\times 1}$ to $F^{m\times 1}$.

The function U defined by $U(\alpha) = \alpha A$
is a linear transformation from F^m to F^n.

Example 4.

Let P be a fixed $m \times m$ matrix
with entries in the field F.

Let Q be a fixed $n \times n$ matrix over F.

Let Q be a fixed $n \times n$ matrix over F.

Define a function T from $F^{m \times n}$ to itself by
$T(A) = PAQ.$

Then T is a linear transformation because

$$\begin{aligned} T(cA + B) &= P(cA + B)Q \\ &= (cPA + PB)Q \\ &= cPAQ + PBQ \\ &= cT(A) + T(B). \end{aligned}$$

Example 5.

Let R be the field of real numbers
Let V be the space of all functions from R to R
(that are continuous).

Define by T the transformation

$$(Tf)(x) = \int_0^x f(t)dt.$$

Then,
T is a linear transformation from V to V.

NB: The function Tf is not only continuous
but also a continuous first derivative.

Remark: Linearity of integration
is one of its fundamental properties.

Story:

- The book says: the reader should have no difficulty
  verifying that the transformatinos in
  Examples 1, 2, 3 and 5 are linear.
  (HW)
- We will expand our list of examples considerably
  as we learn more about linear transformations.

NB. Note that if T is a linear transformation from V to W

then $T(0) = 0$

(this follows from the defintion:

$$T(0) = T(0+0) = T(0) + T(0)).$$

Remark:

- What we are calling linear may be slightly different
  from what is considered in say physics etct.

  For instance, suppose V is just the real line.
  Then one might call a particular transformation T (from V to V)
  to be linear if its graph is a straight line.
  But for us, the straight line must pass through the origin.

Story:

- In addition to the $T(0) = 0$ property,
  we also observe another important property
  of linear transformations.

NB.

A linear transformation 'preserves' linear combinations,

i.e.

if $\alpha_1, \dots \alpha_n$ are vectors in V

and

$c_1 \dots c_n$ are scalars

then

$$T(c_1\alpha_1 + \cdots + c_n\alpha_n) = c_1(T\alpha_1) + \cdots + c_n(T\alpha_n).$$

This follows readily from the definition.

E.g.

$$\begin{aligned} T(c_1\alpha_1 + c_2\alpha_2) &= c_1(T\alpha_1) + T(c_2\alpha_2) \\ &= c_1(T\alpha_1) + c_2(T\alpha_2). \end{aligned}$$

Theorem 1.

Let

- V be a finite-dimensional vector space
  over th enfield F
- $\{\alpha_1 \dots \alpha_n\}$ be an ordered basis for V.
- W be a vector space over the same field F and
- $\beta_1, \dots, \beta_n$ be any vectors in W.

Then

there is precisely one linear transformation

T from V to W such that

$T\alpha_j = \beta_j$ for $j \in \{1 \dots n\}$.

Proof.

Strategy:

To prove there is some linear tranformation

T with $T\alpha_j = \beta_j$ we proceed as follows

(i) we first define a map T

(ii) show that this map does the required mapping

(iii) show that T is linear

(iv) show that T is unique.

NB: Given α in V

there is a unique n-tuple $(x_1 \dots x_n)$

such that

$$\alpha = x_1\alpha_1 + \cdots + x_n\alpha_n.$$

Defn.

For this vector α, define

$$T\alpha = x_1\beta_1 + \cdots + x_n\beta_n.$$

NB.

T is a well-defined rule for associating with each vector α in V

a vector $T\alpha$ in W.

NB2.

From the definition of T

it is clear that $T\alpha_j = \beta_j$ for each j.

Story: So this means the mapping is as we would like it to be

It remains to show that the transformatino T is linear.

Claim: T is a linear transformation.

Proof.

Let $\beta = y_1\alpha_1 + \cdots + y_n\alpha_n$

be in V and let c be any scalar.

Now, we have that

$$c\alpha + \beta = (cx_1 + y_1)\alpha_1 + \cdots + (cx_n + y_n)\alpha_n$$

and so by definition

$$T(c\alpha + \beta) = (cx_1 + y_1)\beta_1 + \cdots + (cx_n + y_n)\beta_n$$

On the other hand

$$c(T\alpha) + T\beta = c\sum_i x_i\beta_i + \sum_i y_i\beta_i = \sum_i (cx_i + y_i)\beta_i$$

and therefore

$$T(c\alpha + \beta) = c(T\alpha) + T\beta.$$

□

Story: The final step is to show uniqueness of T.

Claim: T is unique.

Proof

Suppose U is a linear transformation

from V to W with

$U\alpha_j = \beta_j$ for $j \in \{1 \dots n\}$.

Then, for the vector

$\alpha = \sum_{i=1}^{n} x_i\alpha_i$

it holds that

$$U\alpha = U\left(\sum_i x_i\alpha_i\right) = \sum_i x_i(U\,\alpha_i) = \sum_i x_i\beta_i$$

But this is exactly the rule/map T that we defined above.

This shows that the linear transformation T

with $T\alpha_j = \beta_j$ is unique.

□

␣

■

Story:

- Theorem 1
  as must be evident
  is very elementary
  - Yet, it is so basic that the book has stated it formally.
- The concept of a function is very general.
  - If V and W are (non-zero) vector spaces
    then there are very many functions
    from V to W.

  - Theorem 1 helps establish that
    the ones that are linear
    are extremely special.

Example 6.

The vectors

$\alpha_1 = (1,2)$ and $\alpha_2 = (3,4)$

are lineraly independent and

therefore form a basis for R^2.

According to Theorem 1,

there is a unique linear transformation

from R^2 to R^3 such that

$$T\alpha_1 = (3, 2, 1)$$
$$T\alpha_2 = (6, 5, 4).$$

We should therefore be abel to find $T(\epsilon_1)$, i.e. $T(1,0)$.

To this end,

we find scalars c_1, c_2 such that

$\epsilon_1 = c_1\alpha_1 + c_2\alpha_2$ and then we know that $T\epsilon_1 = c_1 T\alpha_1 + c_2 T\alpha_2$.

Using $(1,0) = c_1(1,2) + c_2(3,4)$
we get $c_1 = -2$ and $c_2 = 1$.

Thus
$$T(1,0) = -2(3,2,1) + (6,5,4) = (0,1,2).$$