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.

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

NBZ.

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 $j \in \{1 \ldots n\}$.

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

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

Example 7.

Let T be a linear transformation from

the m-tuple space F^m into

the n-tuple space F^n.

Theorem 1 tells us that T is uniquely determined by

the sequence of vectors $\beta_1 \dots \beta_m$ where

$\beta_i = T\epsilon_i$ for $i \in \{1 \dots m\}$.

(i.e. T is uniquely determined by the images

of the standard basis vectors.)

So, how does this vector act?

Let $\alpha = (x_1 \dots x_m)$.

Then $T\alpha = x_1\beta_1 + \cdots + x_m\beta_m$.

If B is the $m \times n$ matrix

that has row vectors $\beta_1, \dots \beta_m$

the transnformation can be written as

$T\alpha = \alpha B$.

i.e. if $\beta_i = (B_{i1}, \dots B_{in})$ then

$$T(x_1, \dots, x_n) = [x_1 \cdots x_m] \begin{bmatrix} B_{11} & \cdots & B_{1n} \\ \vdots & & \vdots \\ B_{m1} & \cdots & B_{mn} \end{bmatrix}.$$

This is a very explicit description

of the linear transforamtion.

In Section 3.4

we look at the relationship b/w

linear transformations and matrices in detail;

Here, we skip it because the matrix B
is on the right of the vector
and that can cause confusion;

The point of the example is that
one can find a reasonably simple description
of all linear tranformatinos
from F^m to F^n.

Story:

- If T is a linear transformation from V into W
  then the range of T is
  not only a subset of W
  it is a *subspace* of W.

- Let R_T be the range of T
  i.e. the set of all vectors β in W
  such that $\beta = T\alpha$ for some α in V.
  $\{\beta : \beta = T\alpha, \text{for } \alpha \in V\}$

Claim. R_T is a subspace.
Proof.

- Let β_1 and β_2 be in R_T and let c be a scalar.
  (we will show $c\beta_1 + \beta_2$ is also in R_T)

  - Then, there are vectors
    α_1 and α_2 in V such that

    $T\alpha_1 = \beta_1$ and
    $T\alpha_2 = \beta_2$.

  - Now, since T is linear, it holds that

    $$T(c\alpha_1 + \alpha_2) = cT\alpha_1 + T\alpha_2 = c\beta_1 + \beta_2$$

    which shows that $c\beta_1 + \beta_2$ is also in R_T.

■

Story:

- Another interesting subspace associtaed with
  the linear transformation T
  is the set N consisting of the vectors
  α in V such that $T\alpha = 0$,
  $N \coloneqq \{\alpha \in V : T\alpha = 0\}$
  (me: the null space?)

Claim. N is a subspace of V.

Proof.
(a) $T(0) = 0$ so that N is non-empty
(b) if $T\alpha_1 = T\alpha_2 = 0$
then
$$T(c\alpha_1 + \alpha_2) = cT\alpha_1 + T\alpha_2 = c0 + 0 = 0$$
so that $c\alpha_1 + \alpha_2 \in N$.

■

Definition.
Let
V and W be vector spaces oven the field F and
T be a linear transformation from V to W.

The **null space** of T is the set of all vectors α in V

such that $T\alpha = 0$

i.e. $T \coloneqq \{\alpha \in V : T\alpha = 0\}$.

If V is finite dimensional

the **rank** of T is the dimension of the range of T

i.e. $\text{rank}(T) \coloneqq \dim(\{T\alpha : \alpha \in V\})$

Story:

The following is one of the most important results in linear algebra.

Theorem 2.

Let

V and W be vector spaces over the field F and

T be a linear transformation

from V to W.

Suppose V is finite-dimensional.

Then

$\text{rank}(T) + \text{nullity}(T) = \dim V$.

Proof.

Let $\{\alpha_1, \ldots \alpha_k\}$ be a basis for N (the null space of T).

There are vectors

$\alpha_{k+1}, \ldots \alpha_n$ in V such that

$\{\alpha_1 \ldots \alpha_n\}$ is a basis for V.

Assertion. We assert that $\{T\alpha_{k+1}, \ldots T\alpha_n\}$ is a basis for

the range of T.

(we prove this shortly)

Strategy/final argument:

How does this help?

If r is the rank of T

the assertion above says that

$r = n - k$

(because the range of k

has a basis with $n - k$ elements)

Since k is the nullity, and n the dimension of V

this suffices to establish the theorem.

Therefore, it suffices to prove the assertion

to establish the result.

Claim. $\{T\alpha_{k+1}, \ldots T\alpha_n\}$ span the range of T.

The vectors $T\alpha_1, \ldots T\alpha_n$ certainly span the range of T

(because any vector in β in V can be

written as a linear combination of α_is

and $T\beta$ is then a linear combination of $T\alpha_i$s)

Clearly, $T\alpha_{k+1}, \ldots T\alpha_n$ also span the range

(because $T\alpha_j = 0$ for all $j \leq k$).

It remains to establish that these are also

linearly independent.

Claim. $T\alpha_{k+1} \ldots T\alpha_n$ are linearly independent.

Consider scalars c_i such that

$\sum_{i\in\{k+1,\ldots n\}} c_i(T\alpha_i) = 0$.

This says

$$T\left(\quad \sum \quad c_i\alpha_i \right) = 0$$

$\Big(\sum_{i \in \{k+1,\dots n\}} \Big)$

and so the vector $\sum_{i \in \{k+1,\dots n\}} c_i \alpha_i$ is also
in the null space of T.

Since $\alpha_1, \dots \alpha_k$ form a basis for N (the null space)
there must be scalars $b_1, \dots b_k$ such that

$\alpha = \sum_{i \in \{1 \dots k\}} b_i \alpha_i$.

Thus,

$$\sum_{i \in \{1 \dots k\}} b_i \alpha_i - \sum_{j \in \{k+1,\dots n\}} c_j \alpha_j = 0$$

and now since $\alpha_1, \dots \alpha_n$ are linearly independent
it must be the case that
$b_1 = \dots = b_k = c_{k+1} = \dots = c_n = 0$.

■

Theorem 3.
If A is an $m \times n$ matrix with entries in the field F
then
$\text{row rank}(A) = \text{column rank}(A)$.

Proof.
Strategy: We obtain an expression relating
the dimension of the solution space of $AX = 0$ (nullity) and m (dimension of the initial vector space)
with the column and row rank of A

The first part uses Theorem 2 above and
the second part uses a fact about row rank from Chapter 2.

Let T be the linear transformation from
$F^{n \times 1}$ to $F^{m \times 1}$
defined by $T(X) = AX$.

The null space of T
is the solution space for the system
$AX = 0$
(i.e. the set of all column matrices
X such that $AX = 0$).

The range of T is the set of all $m \times 1$ column matrices
Y such that $AX = Y$ has a solution for X.
(i.e. Ys such that there is an X such that
$AX = Y$)

Let $A_1, \dots A_n$ be the columns of A.

Then
$AX = x_1 A_1 + \dots + x_n A_n$

so that the range of T is the
subpsace spanned by the columns of A.

(i.e. the range of T is the column space of A.)

Thus
$\text{rank}(T) = \text{column rank}(A)$.

From Theorem 2 (above),

EXAMPLE 15. Let A be an $m \times n$ matrix and let S be the solution space for the homogeneous system $AX = 0$ (Example 7). Let R be a row-reduced echelon matrix which is row-equivalent to A. Then S is also the solution space for the system $RX = 0$. If R has r non-zero rows, then the system of equations $RX = 0$ simply expresses r of the unknowns $x_1, \dots, x_n$ in terms of the remaining $(n - r)$ unknowns x_j. Suppose that the leading

we have that if S is the solution space for the system $AX = 0$
then it holds that
$\dim S + \text{column rank}(A) = n$.

Recall Example 15 of Chapter 2:
We showed that if r is the dimension of the row space of A
then the solution space S has a basis
consisting of $n - r$ vectors:

$$\dim S = n - \text{row rank}(A).$$

Combining these, we have
$\text{row rank}(A) = \text{column rank}(A)$.

■

Story:
The proof of Theorem 3
depends on calculations concerning
systems of linear equations.

There is a more conceptual proof
that does not rely on such calculations.

We look at it in Section 3.7

§ 3.2 The Algebra of Linear Transformations

Story:
In the study of linear transformations
from V to W
it is of fundamental importance
that the set of these transformations
inherit a natural vector space structure.

The set of linear transformations
from a space V into itself
has even more algebraic structure
because
ordinary composition of functions
provides a 'multiplication' of such transformations.

We shall explore these ideas in this section.

Theorem 4.
Let
- V and W be vector spaces over the field F, and
- T and U be linear transformations
from V into W.

Then, the function $(T + U)$ defined by
$(T + U)(\alpha) = T\alpha + U\alpha$
is a linear transformation
from V to W.

Further
if c is any element of F
the function cT defined by
$(cT)(\alpha) = c(T\alpha)$
is a linear transformation
from V to W.

Finally, consider the set of all linear transformations
from V to W

non-zero entries of the non-zero rows occur in columns $k_1, \ldots, k_r$. Let J be the set consisting of the $n - r$ indices different from $k_1, \ldots, k_r$:

$$J = \{1, \ldots, n\} - \{k_1, \ldots, k_r\}.$$

The system $RX = 0$ has the form

$$\begin{array}{ccc} x_{k_1} + \sum_J c_{1j}x_j & = & 0 \\ \vdots \quad\quad \vdots & & \vdots \\ x_{k_r} + \sum_J c_{rj}x_j & = & 0 \end{array}$$

where the c_{ij} are certain scalars. All solutions are obtained by assigning (arbitrary) values to those x_j's with j in J and computing the corresponding values of $x_{k_1}, \ldots, x_{k_r}$. For each j in J, let E_j be the solution obtained by setting $x_j = 1$ and $x_i = 0$ for all other i in J. We assert that the $(n - r)$ vectors E_j, j in J, form a basis for the solution space.

Since the column matrix E_j has a 1 in row j and zeros in the rows indexed by other elements of J, the reasoning of Example 13 shows us that the set of these vectors is linearly independent. That set spans the solution space, for this reason. If the column matrix T, with entries $t_1, \ldots, t_n$, is in the solution space, the matrix

$$N = \sum_J t_j E_j$$

is also in the solution space *and* is a solution such that $x_j = t_j$ for each j in J. The solution with that property is unique; hence, $N = T$ and T is in the span of the vectors E_j.

from V to W
together with the addition and scalar multiplication
defined above.
This forms a vector space over the field F.

Proof.

Let T and U be as stated in the premise.

Claim. $(T + U)$ is a linear transformation.
Then, we have the following:

$$\begin{aligned}(T+U)(c\alpha+\beta) &= T(c\alpha+\beta)+U(c\alpha+\beta)\\ &= c(T\alpha)+T\beta+c(U\alpha)+U\beta\\ &= c(T\alpha+U\alpha)+(T\beta+U\beta)\\ &= c(T+U)(\alpha)+(T+U)(\beta)\end{aligned}$$

Claim. Similarly, cT is a linear transformation.
Observe,

$$\begin{aligned}(cT)(d\alpha+\beta) &= c[T(d\alpha+\beta)]\\ &= c[d(T\alpha)+T\beta]\\ &= cd(T\alpha)+c(T\beta)\\ &= d[c(T\alpha)]+c(T\beta)\\ &= d[(cT)\alpha]+(cT)\beta\end{aligned}$$

It remains to establish that
the set of linear transformations of V to W
(together with these operations)
is a vector space.

This is a matter of verifying the conditions
on vector addition and
scalar multiplication.

The book leaves it to the reader—we will leave add it in the problem set.
HW

The text gives the following hint/makes the following remark:

- The zero vector in this space
  will be the zero transformation
  (i.e. the transformation that sends
  every vector of V into the zero vector in W)

- each of the properties of the two operations
  follows from the corresponding property of
  the operations in W.

■

Remark: The text mentions another way of looking at this theorem

- If one defines sum and scalar multiplation as we did
  then the set of *all* functions from V to W
  becomes a vector space over the field F.

  - This has nothing to do with the fact that
    V is a vector space over F
    (only that V is a non-empty set).

- When V is a vector space
  one can define a linear transformation
  from V to W.

  Theorem 4 says that
  the linear transformatinos are a *subspace* of
  the *space* of all functions from V to W.

Notation: We use $L(V, W)$ to denote
the space of linear transformations from
V to W.

(The text emphasises that
L is only defined when
V and W are vector spaces over the *same* field)

Theorem 5.

Let
V be an n-dimensional vector space over F, and
W be an m-dimensional vector space over F.

Then
the space $L(V, W)$ is finite-dimensional
and has dimension mn.

Proof.

Strategy: We construct a very simple basis for L.

Let
$\mathcal{B} = \{\alpha_1 \dots \alpha_n\}$ and
$\mathcal{B}' = \{\beta_1 \dots \beta_m\}$
be ordered bases for V and W, respectively.

For each pair of integers
(p, q) with $1 \leq p \leq m$ and $1 \leq q \leq n$
one can define a linear transformation
$E^{p,q}$ from V to W by

$$\begin{aligned} E^{p,q}(\alpha_i) &= \begin{cases} 0, & \text{if } i \neq q \\ \beta_p, & \text{if } i = q \end{cases} \\ &= \delta_{iq}\beta_p. \end{aligned}$$

(i.e. map the vector α_q to β_p and map the rest to the 0 vector)

According to Theorem 1
there is a unique transformation from V to W
satisfying these conditions.

Claim: The mn transformations $E^{p,q}$ form a basis for $L(V, W)$.

Let T be a linear transformation from V to W.
For each $j \in \{1 \dots n\}$
let $A_{ij}, \dots A_{mj}$ be the coordinates of the vector
$T\alpha_j$ in the ordered basis $\mathcal{B}'$
i.e.
$$T\alpha_j = \sum_{p \in \{1 \dots m\}} A_{pj}\beta_p. \qquad [3\text{-}1]$$

We want to show that

$$T = \sum_{p \in \{1 \dots m\}, q \in \{1 \dots m\}} A_{pq}E^{p,q}. \qquad [3\text{-}2]$$

Proof:
Let U be the linear transformation
on the RHS above (in [3-2]).

Then for each j, it holds that

$$\begin{aligned} U\alpha_j &= \sum_p \sum_q A_{pq}E^{p,q}(\alpha_j) \\ &= \sum_p \sum_q A_{pq}\delta_{jq}\beta_p \\ &= \sum_{p=1}^{m} A_{pj}\beta_p \\ &= T\alpha_j \end{aligned}$$

and therefore $U = T$.

□

From [3-2] it follows that $E^{p,q}$ span $L(V, W)$
(one can write every linear transformation
as a linear combination of $E^{p,q}$).

It remains to show that $E^{p,q}$ are linearly independent.

Proof.
We want to show that
no linear combination of E^{pq}s
can give the zero vector
(unless all coefficients are zero).

This is also follows from [3-2].

Suppose $U = \sum_{pq} A_{pq} E^{p,q}$
is the zero transformation;

Then
$U\alpha_j = 0$, i.e.
$\sum_p A_{pj}\beta_j = 0$.

Now, independence of β_js implies that
$A_{pj} = 0$ for every p and j.

□

This shows $E^{p,q}$ form a basis, as claimed.

□

The size of the basis is mn, and this means
the space $L(V, W)$ has dimension mn as asserted.

■

Theorem 6.
Let
- V, W and Z be vector spaces over the field F.
- T be a linear transformation from V to W and
- U be a linear transformation from W to Z

Then the composed function UT
defined by $UT(\alpha) = U(T(\alpha))$
is a linear transformation
from V to Z.

Proof.

$$\begin{aligned}(UT)(c\alpha + \beta) &= U[T(c\alpha + \beta)] \\ &= U(cT\alpha + T\beta) \\ &= c[U(T\alpha)] + U(T\beta) \\ &= c(UT)(\alpha) + UT(\beta)\end{aligned}$$

■

Story:
Henceforth
we will primarily focus on
linear transformations

T from V to V

And so we use the following simpler notation:

"T is a linear operator on V"

Definition.

If V is a vector space over the field F

a **linear operator on** V is

a linear transformation from V to V.

NB: In Theorem 6

when $V = W = Z$,

U and T are linear operators on V,

it turns out that

their composition UT is again

a linear operator on V.

NB2: Thus the space $L(V, V)$

has a "multiplication" defined on it

by composition.

In this case

the operator TU is also defined

and

one should note that in general

$UT \neq TU$, i.e. $UT - TU \neq 0$.

NB3: If T is a linear operator on V then

one can compose T with T.

Notation. We use $T^2 = TT$, and in general, $T^n = T \cdots T$ (n times)

for $n = 1,2,3 \ldots$

We define $T^0 = I$ if $T \neq 0$.

Lemma.

Let

V be a vector space over the field F

U, T_1 and T_2 be linear operators on V

c be an element of F.

(a) $IU = UI = U$

(b)

$U(T_1 + T_2) = UT_1 + UT_2$

$(T_1 + T_2)U = T_1U + T_2U$

(c) $c(UT_1) = (cU)T_1 = U(cT_1)$

Proof.

(a) Immediate; stated for emphasis

(b) $[U(T_1 + T_2)](\alpha) = U[(T_1 + T_2)(\alpha)]$

$= U(T_1\alpha + T_2\alpha)$

$= U(T_1\alpha) + U(T_2\alpha)$

$= (UT_1)(\alpha) + (UT_2)(\alpha)$

so that $U(T_1 + T_2) = UT_1 + UT_2$.

Similarly,

$[(T_1 + T_2)U](\alpha) = (T_1 + T_2)(U\alpha)$

$= T_1(U\alpha) + T_2(U\alpha)$

$= (T_1U)(\alpha) + (T_2U)(\alpha)$

so that $(T_1 + T_2)U = T_1U + T_2U$.

(NB: Note that the proofs of these two distributive laws

do not use the fact that T_1 and T_2 are linear

and

the proof of the second one does not use the fact

that U is linear either.)

<--- Why not?

We do use the fact that

$(T_1 + T_2)(\alpha) = T_1(\alpha) + T_2(\alpha)$

which I thought was coming from

linearity—

no, this is simply how we defined

(c) HW

$T_1 + T_2$

Story:

The text says that
the contents of this lemma &
a portion of Theorem 5
tell us that the vector space $L(V, V)$
together with the composition operation
is what is known as a linear algebra with identity.
This will be discussed in Chapter 4.

Example 8.

If A is an $m \times n$ matrix
with entries in F
we have the linear transformation
T defined by $T(X) = AX$
from $F^{n\times 1}$ to $F^{m\times 1}$.

If B is a $p \times m$ matrix
(with entries in F)
we have the linear transformation
U defined by $U(Y) = BY$
from $F^{m\times 1}$ to $F^{p\times 1}$.

The composition UT is easily described:

$$\begin{aligned}(UT)(X) &= U(T(X)) \\ &= U(AX) \\ &= B(AX) \\ &= (BA)X\end{aligned}$$

Thus, UT is "left multiplication by the product matrix BA".

Example 9.

Let
F be a field,
V be the vector space of all polynomial functions
from F to F,
D be the differentiation operator
defined in Example 2 and
T be the linear operator "multiplication by x":
$(Tf)(x) = xf(x)$.

Then $DT \neq TD$.

In fact $DT - TD = I$.
HW

Example 10.

Let $\mathcal{B} = \{\alpha_1, \dots \alpha_n\}$ be an ordered basis for a vector space V.
Consider the linear operators $E^{p,q}$ (that also arose in the proof of Theorem 5):
$E^{p,q}(\alpha_i) = \delta_{iq}\alpha_p$.
(it takes the basis vector q and replaces it with p
and sends all other basis vectors to the zero vector)

These n^2 linear operators
form a basis for the space of linear operators on V.

What is $E^{p,q}E^{r,s}$?

We have

We have

$(E^{p,q}E^{r,s})(\alpha_i) = E^{p,q}(\delta_{is}\alpha_r)$
$= \delta_{is}E^{p,q}(\alpha_r)$
$= \delta_{is}\delta_{rq}\alpha_p$

Therefore

we have

$$E^{p,q}E^{r,s} = \begin{cases} \mathbf{0}, & \text{if } r \neq q \\ E^{p,s}, & \text{if } q = r. \end{cases}$$

Let T be a linear operator on V.

We showed in the proof of Theorem 5 that if

$A_j = [T\alpha_j]_{\mathcal{B}}$
$A = [A_1, \ldots, A_n]$

then

$T = \sum_{pq} A_{pq}E^{p,q}$.

If

$$U = \sum_{rs} B_{rs}E^{r,s}$$

is another linear operator on V then

the last lemma tells us that

$$TU = \left(\sum_{pq} A_{pq}E^{p,q}\right)\left(\sum_{rs} B_{rs}E^{r,s}\right) = \sum_{pqrs} A_{pq}B_{rs}E^{p,q}E^{r,s}$$

and as we already noted, the only terms that contribute to the sum

are the ones where $q = r$

and

since $E^{p,r}E^{r,s} = E^{p,s}$

(easily verified by using the definition)

it holds that

$$TU = \sum_{ps}\left(\sum_{r} A_{pr}B_{rs}\right)E^{p,s}.$$

Thus, the effect of composing T and U

is to multiply the matrices A and B.

Story:

- In our discussion of algebraic operations

  with linear transformations

  we have not yet said anything about invertibility.

  One specific question of interest is this:

  For which linear operators T on the space V

  does there exist a linear operator T^{-1}

  such that $TT^{-1} = T^{-1}T = I$?

Notation.

The function T from V to W

is called **invertible** if

there exists

a function U from W to V

such that

UT is the identity function on V and

TU is the identity function on W.

Proposition.

If T is invertible
the function U is unique and
is denoted by T^{-1}.

Furthermore
T is invertible
iff
(1) T is 1: 1, (i.e. $T\alpha = T\beta$ implies $\alpha = \beta$)
(2) T is onto (i.e. the range of T is (all of) W).

Proof? The text defers the proof to the appendix; so we assume it for now.

Theorem 7.

Let
V and W be vector spaces
over the field F and
T be a linear transformation from
V to W.

If T is invertible, then
the inverse function T^{-1}
is a linear transformation from W to V.

Proof.

(NB: We assume the proposition above, and prove T^{-1} is a linear transformation).

Let
β_1, β_2 be vectors in W, and
c be a scalar.

We will show that
$T^{-1}(c\beta_1 + \beta_2) = cT^{-1}\beta_1 + T^{-1}\beta_2$.

Let $\alpha_i = T^{-1}\beta_i$, $i = 1,2$, i.e.
let α_i be the unique vector in V such that
$T\alpha_i = \beta_i$.
(otherwise T^{-1} will not be unique; contradicting the proposition above)

Since T is linear, we have
$T(c\alpha_1 + \alpha_2) = cT\alpha_1 + T\alpha_2 = c\beta_1 + \beta_2$.

Thus, $c\alpha_1 + \alpha_2$ is the unique vector in V
that is sent by T to $c\beta_1 + \beta_2$.
(again, otherwise T^{-1} would not be unique)

We therefore have
$T^{-1}(c\beta_1 + \beta_2) = c\alpha_1 + \alpha_2 = c(T^{-1}\beta_1) + T^{-1}\beta_2$
as we asserted.

This showes T^{-1} is linear.

■

Story:

- Suppose we have
  an invertible linear transformation T from V to W
  and
  an invertible linear transformation U from W to Z.
- What can one say about the inverse of UT?

Claim. UT is invertible and $(UT)^{-1} = T^{-1}U^{-1}$.

Proof. $(UT)^{-1}(UT) = T^{-1}U^{-1}UT = I$ and $(UT)(UT)^{-1} = UTT^{-1}U^{-1} = I$
so $(UT)^{-1}$ is both a left and right inverse of UT

thus UT is invertible.

(NB: The proof did not rely on whether U^{-1}, T^{-1} are 1:1 or onto nor does it rely on linearity)

■

Story:

- If T is linear
  then $T(\alpha - \beta) = T\alpha - T\beta$
- Thus,
  $T\alpha = T\beta$ iff $T(\alpha - \beta) = 0$.
- This enormously simplifies the verification that T is 1:1.

Notation.

A linear transformation T is called **non-singular**
if $T\gamma = 0$ implies $\gamma = 0$
(i.e. the null space of T is $\{0\}$).

Claim. Evidently, T is 1:1
iff
T is non-singular

Proof idea:
Clearly, if T is not 1:1, there would be two distinct elements
being mapped to the same element
and that would mean T is singular
Conversely, if T is singular,
there are at least two distinct elements
that are mapped to the same element
thus T is not 1:1

□

Story:

- The extension of this claim is that non-singular linear transformations
  are those that preserve linear independence.

Theorem 8.

Let T be a linear transformation from V to W.
Then
T is non-singular iff
T carries
each linearly independent subset of V
into
a linearly independent subset of W.

Proof.

$\Rightarrow$
(i.e. T is non-singular implies independent set to independent set)

Let S be a linearly independent subset of V.

Claim. If $\alpha_1, \ldots, \alpha_k$ are vectors in S then
the vectors $T\alpha_1, \ldots T\alpha_k$ are linearly independent
Proof.
Suppose there are some scalars $c_1 \ldots c_k$ such that
$c_1(T\alpha_1) + \cdots + c_k(T\alpha_k) = 0$.
That would mean
$T(c_1\alpha_1 + \cdots + c_k\alpha_k) = 0$.
Finally, since T is non-singular,
$c_1\alpha_1 + \cdots + c_k\alpha_k = 0$.
It follows that each $c_i = 0$.
Thus the vectors $T\alpha_1, \cdots T\alpha_k$ are linearly independent.

□

The claim shows that the image of S under T is independent, as asserted.

$\Leftarrow$
(i.e. Independent set to independent set implies T is non-singular)

Let α be a non-zero vector in V.
Define $S \coloneqq \{\alpha\}$ which is independent (only one vector in the set).
The image of S is the set $\{T\alpha\}$ is
guaranteed to be independent (by the premise)

Therefore, $T\alpha \neq 0$
(because, recall the set consisting of the zero vector alone
is dependent).

Thus, the null space of T is the zero subspace, i.e.
T is non-singular, as asserted.

■

Story [Me]:

- We said a linear transformation T is 1:1 iff T is non-singular
  and we also said that
  T is invertible iff T is 1:1 and onto.

- We now explore whether
  a singular T can have any kind of inverse.

Example 11.
Let F be a subfield of the complex numbers
(or a field of characteristic zero).

Let V be the space of polynomial functions over F.

Consider
the differentiation operator D
and
the "multiplication by x" operator T
from Example 9.

Let's look at operator D first.
Since D sends all constants into 0
D is singular.

However, V is not finite dimensional and
the range of D is all of V.

It turns out that it is possible to define a
right inverse for D.

For instance
if E is the indefinite integral operator:

$$E(c_0 + c_1x + \cdots + c_nx^n) = c_0x + \frac{1}{2}c_1x^2 + \cdots + \frac{1}{n+1}c_nx^{n+1}$$

then
E is a linear operator on V and
$DE = I$

So, D has a right inverse (even though it is singular).

On the other hand
$ED \neq I$ because ED sends the constants into 0.

(Me: So, it is unclear if D has a left inverse.)

Now, going back to operator T,
note that T is in a "reverse situation" somehow.

If $xf(x) = 0$ for all x,
then $f = 0$.
(HW)

Thus,
T is non-singular.

And, it is possible to find a left inverse for T.
For instance,
if U is the operation "remove th constant term and divide by x":
$U(c_0 + c_1x + \cdots + c_nx^n) = c_1 + c_2x + \cdots + c_nx^{n-1}$
then U is a linear operator on V
and $UT = I$. [easy to verify]

But $TU \neq I$ <--- TODO: understand this step
(since every function in the range of TU
is in the range of T
which is the space of polynomial functions f
such that $f(0) = 0$.)