

Linear Transformations

Definition 1 A linear transformation is a function $L : V \rightarrow W$ from one vector space V to another W which satisfies:

1. $L(u + v) = L(u) + L(v)$
2. $L(au) = aL(u)$

for every $u, v \in V$ and scalar a .

Exercise 2 Let $L : V \rightarrow W$ be a function from a vector space V to a vector space W . Prove that L is a linear transformation iff

$$L(au + bv) = aL(u) + bL(v)$$

for any $u, v \in V$ and scalars a and b .

Theorem 3 Every linear transformation $L : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is determined by a matrix $A \in \text{MAT}_{m,n}$:

$$L(X) = AX$$

for every $X \in \mathbb{R}^n$

proof:

Given A since

$$A(X + Y) = AX + AY \text{ and } A(aX) = a(AX)$$

it is clear that $L(X) = AX$ is a linear transformation.

For the converse, assume $L : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is a linear transformation. The canonical basis for $\mathbb{R}^n$ is the sequence of columns of the $n \times n$ identity matrix, $I_{n \times n}$. So let $e_i = \text{col}_i(I_{n \times n})$. Note for any vector

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = x_1 e_1 + x_2 e_2 + \cdots + x_n e_n.$$

Let A be the matrix such that

$$L(e_i) = \text{col}_i(A)$$

for every $i = 1, \dots, n$. Then

$$\begin{aligned} L \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} &= L(x_1 e_1 + x_2 e_2 + \dots + x_n e_n) \\ &= x_1 L(e_1) + x_2 L(e_2) + \dots + x_n L(e_n) \\ &= x_1 \text{col}_1(A) + x_2 \text{col}_2(A) + \dots + x_n \text{col}_n(A) \\ &= A \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \end{aligned}$$

QED

Definition 4 For $L : V \rightarrow W$ a linear transformation, define the null space of L , $\text{null}(L)$, and the range space of L , $\text{range}(L)$ as follows:

1. $\text{null}(L) = \{v \in V : L(v) = z\}$
2. $\text{range}(L) = \{w \in W : \text{there exists } v \in V \text{ such that } L(v) = w\}$

Theorem 5 $\text{null}(L)$ and $\text{range}(L)$ are subspaces of V and W , respectively.

proof:

$\text{null}(L)$: Suppose $u, v \in \text{null}(L)$ and a, b scalars. Then $L(u) = z$ and $L(v) = z$ and

$$L(au + bv) = aL(u) + bL(v) = az + bz = z + z = z.$$

Hence $au + bv \in \text{null}(L)$

$\text{range}(L)$: Suppose $w_1, w_2 \in \text{range}(L)$ and a_1, a_2 are scalars. Then there exists $v_1, v_2 \in V$ such that $L(v_1) = w_1$ and $L(v_2) = w_2$. Then

$$L(a_1 v_1 + a_2 v_2) = a_1 L(v_1) + a_2 L(v_2) = a_1 w_1 + a_2 w_2$$

and so $a_1 w_1 + a_2 w_2 \in \text{range}(L)$.

QED

Definition 6 A function $L : V \rightarrow W$ is one-to-one iff $L(u) = L(v)$ implies $u = v$ for every $u, v \in V$.

Definition 7 A function $L : V \rightarrow W$ is onto iff for every $w \in W$ there exists $v \in V$ such that $L(v) = w$.

Definition 8 A function $L : V \rightarrow W$ is an isomorphism iff L is a one-to-one, onto, linear transformation.

Theorem 9 A linear transformation $L : V \rightarrow W$ is an isomorphism iff $\text{null}(L) = \{z\}$ and $\text{range}(L) = W$.

proof:

It is clear from the definitions that L is onto iff $\text{range}(L) = W$.

Claim L is one-to-one iff $\text{null}(L) = \{z\}$.

proof:

Assume L is one-to-one. Then $L(u) = z$ and $L(z) = z$ implies $u = z$ and so $\text{null}(L) = \{z\}$.

Conversely, assume $\text{null}(L) = \{z\}$. Then if $L(u) = L(v)$, then $L(u - v) = L(u) - L(v) = z$ and so $u - v \in \text{null}(L)$. Hence $u - v \in \{z\}$ and so $u - v = z$ and therefore $u = v$. So $L(u) = L(v)$ implies $u = v$ and therefore L is one-to-one. This proves the Claim.

QED

Theorem 10 Suppose A is an invertible $n \times n$ matrix. If $L : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is defined by $L(X) = AX$, then L is an isomorphism.

proof:

We already know from Theorem 3 that L is a linear transformation. To see that it is one-to-one, assume $L(X) = L(Y)$. Then $AX = AY$ implies $A^{-1}AX = A^{-1}AY$ implies $X = Y$. To see that it is onto, suppose $Y \in \mathbb{R}^n$ and let $X = A^{-1}Y$, then $L(X) = AX = AA^{-1}Y = Y$.

QED

Theorem 11 A linear transformation $L : V \rightarrow W$ is an isomorphism iff there exists a linear transformation $L^{-1} : W \rightarrow V$ such that $L \circ L^{-1} = \text{Id}_W$ and $L^{-1} \circ L = \text{Id}_V$. (Id_V is the identity map on V , i.e., $\text{Id}_V(u) = u$ all $u \in V$.)

proof:

First suppose that such an L^{-1} exists. To see that L is one-to-one, suppose $L(u) = L(v)$. Then $L^{-1}(L(u)) = L^{-1}(L(v))$, but

$$u = L^{-1}(L(u)) = L^{-1}(L(v)) = v$$

and so $u = v$. To see that L is onto, suppose $w \in W$. Let $v = L^{-1}(w)$, then $L(v) = L(L^{-1}(w)) = w$. Hence for every $w \in W$ there exists $v \in V$ such that $L(v) = w$ and so L is onto.

For the converse, suppose that L is an isomorphism. This means that L is both one-to-one and onto. It follows that for every $w \in W$ there exists a unique $v \in V$ such that $L(v) = w$. Define $L^{-1} : W \rightarrow V$ by the rule $L^{-1}(w) = v$ iff $L(v) = w$. It follows that $v = L^{-1}(w) = v = L^{-1}(L(v))$ for every $v \in V$ and so $L^{-1} \circ L = Id_V$; and also $w = L(v) = L(L^{-1}(w))$ for every $w \in W$ and so $L \circ L^{-1} = Id_W$. Finally we prove the L^{-1} is a linear transformation. Suppose $w_1, w_2 \in W$ and a_1, a_2 are scalars. Let $u = L^{-1}(a_1w_1 + a_2w_2)$ and let $v = a_1L^{-1}(w_1) + a_2L^{-1}(w_2)$, then

$$L(u) = L(L^{-1}(a_1w_1 + a_2w_2)) = a_1w_1 + a_2w_2$$

and

$$L(v) = L(a_1L^{-1}(w_1) + a_2L^{-1}(w_2)) = a_1L(L^{-1}(w_1)) + a_2L(L^{-1}(w_2)) = a_1w_1 + a_2w_2$$

hence $L(u) = L(v)$ and since L is one-to-one we know that $u = v$ and so

$$L^{-1}(a_1w_1 + a_2w_2) = u = v = a_1L^{-1}(w_1) + a_2L^{-1}(w_2)$$

QED

Exercise 12 Suppose $L : V \rightarrow W$ is a one-to-one linear transformation and $v_1, \dots, v_n$ are linearly independent vectors in V .

Prove that $L(v_1), \dots, L(v_n)$ are linearly independent vectors in W .

Exercise 13 Suppose $L : V \rightarrow W$ is a linear transformation and $v_1, \dots, v_n$ are vectors in V . Prove that

$$L(\text{span}(\{v_1, \dots, v_n\})) = \text{span}(\{L(v_1), \dots, L(v_n)\})$$

Definition 14 V and W are isomorphic iff there exists an isomorphism $L : V \rightarrow W$.

Theorem 15 V is isomorphic to W iff $\dim(V) = \dim(W)$.

proof:

This is true in general, but we only proof it in case the spaces have finite dimension.

Suppose that V is isomorphic to W . Let $v_1, \dots, v_n$ be a basis for V and define $w_1, \dots, w_n$ by $L(v_i) = w_i$ for $i = 1, \dots, n$. By exercise 12 $w_1, \dots, w_n$ are linearly independent (since L is one-to-one) and by exercise 13 they span W since L is onto.

For the converse suppose that $\dim(V) = \dim(W)$ and let $v_1, \dots, v_n$ be a basis for V and $w_1, \dots, w_n$ be a basis for W . Given any $v \in V$ there exists a unique sequence of scalars $a_1, \dots, a_n$ such that

$$v = a_1v_1 + \dots + a_nv_n$$

They exists because $v_1, \dots, v_n$ span V and they are unique because

$$v = a_1v_1 + \dots + a_nv_n \text{ and } v = b_1v_1 + \dots + b_nv_n$$

implies

$$z = (a_1 - b_1)v_1 + \dots + (a_n - b_n)v_n$$

and so by linear independence $a_1 = b_1, \dots, a_n = b_n$. Now define $L : V \rightarrow W$ by

$$L(a_1v_1 + \dots + a_nv_n) = a_1w_1 + \dots + a_nw_n$$

Now we check that L is linear. Suppose $v = a_1v_1 + \dots + a_nv_n$ and $w = b_1v_1 + \dots + b_nv_n$ and a, b are scalars. Then

$$\begin{aligned} L(av + bw) &= L((aa_1 + bb_1)v_1 + \dots + (aa_n + bb_n)v_n) \\ &= (aa_1 + bb_1)w_1 + \dots + (aa_n + bb_n)w_n \\ &= a(a_1w_1) + \dots + a_nw_n + b(b_1w_1 + \dots + b_nw_n) \\ &= aL(v) + bL(w) \end{aligned}$$

Next to see that L is one-to-one and onto, note that for any w there exists unique $a_1, a_2, \dots, a_n$ such that $w = a_1w_1 + \dots + a_nw_n$ and so $v = a_1v_1 + \dots + a_nv_n$ is the unique element of V such that $L(v) = w$.

QED

Theorem 16 Suppose $L : V \rightarrow W$ is a linear transformation. Then

$$\dim(V) = \dim(\text{null}(L)) + \dim(\text{range}(L)).$$

proof:

Let $u_1, \dots, u_n$ be a basis for $\text{null}(L)$ and $v_1, \dots, v_m$ be a basis for $\text{range}(L)$. Let $w_1, \dots, w_m$ be in V such that $L(w_i) = v_i$.

Claim. $u_1, \dots, u_n, w_1, \dots, w_m$ is a basis for V .

proof:

First we show they are linearly independent. Suppose

$$c_1u_1 + \dots + c_nu_n + d_1w_1 + \dots + d_mw_m = z$$

then applying L we get

$$z = L(z) = L(c_1u_1 + \dots + c_nu_n) + L(d_1w_1 + \dots + d_mw_m) = z + d_1v_1 + \dots + d_mv_m$$

because $c_1u_1 + \dots + c_nu_n \in \text{null}(L)$. Hence

$$d_1v_1 + \dots + d_mv_m = z$$

and since the v 's are linearly independent we get $d_1 = \dots = d_m = 0$. Hence $c_1u_1 + \dots + c_nu_n = z$ and since the u 's are linearly independent $c_1 = \dots = c_n = 0$. So, our set is linearly independent.

Now we must see that they span V . Suppose $v \in V$ is arbitrary and let $L(v) = d_1v_1 + \dots + d_mv_m$, then set $u = d_1w_1 + \dots + d_mw_m$ and notice that

$$L(u) = L(d_1w_1 + \dots + d_mw_m) = d_1L(w_1) + \dots + d_mL(w_m) = d_1v_1 + \dots + d_mv_m.$$

It follows that $L(u) = L(v)$ and so $L(u - v) = z$ since L is linear and therefore $v - u \in \text{null}(L)$. Therefore (since the u 's are basis for $\text{null}(L)$) there exists c 's such that

$$v - u = c_1u_1 + \dots + c_nu_n$$

and therefore

$$v = u + c_1u_1 + \dots + c_nu_n = d_1w_1 + \dots + d_mw_m + c_1u_1 + \dots + c_nu_n.$$

Thus $v \in \text{span}(\{u_1, \dots, u_n, w_1, \dots, w_m\})$. This proves the Claim.

QED

From the Claim it follows that the dimension of V is $n + m$ or in other words $\dim(\text{null}(L)) + \dim(\text{range}(L))$. This proves the Theorem.
QED

In the case that L comes from an $m \times n$ matrix A , i.e., $L : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is given by $L(X) = AX$ we have that

$$n = \dim(\mathbb{R}^n) = \dim(\text{null}(A)) + \dim(\text{range}(A))$$

Theorem 17 *Suppose $PAQ = D$ where*

1. *A and D are $m \times n$*
2. *P is $m \times m$, Q is $n \times n$, both invertible, and*
3. *all the entries of D are zero except the lefthand corner which is the $k \times k$ identity matrix, i.e.,*

$$D = \begin{bmatrix} I_{k \times k} & 0 \\ 0 & 0 \end{bmatrix}_{m \times n}$$

Then

1. *$\text{col}_{k+1}(Q), \text{col}_{k+2}(Q), \dots, \text{col}_n(Q)$ is a basis for $\text{null}(A)$,*
2. *$\text{col}_1(P^{-1}), \text{col}_2(P^{-1}), \dots, \text{col}_k(P^{-1})$ is a basis for $\text{range}(A)$,*
3. *$Y \in \text{range}(A)$ iff $\text{entry}_j(PY) = 0$ for all j with $k + 1 \leq j \leq m$, and*
4. *if $Y \in \text{range}(A)$ then $A(QX) = Y$ where X is any $n \times 1$ vector such that $\text{entry}_j(X) = \text{entry}_j(PY)$ for $j = 1, \dots, k$.*

proof:

Recall that for any j

$$\text{col}_j(BC) = B\text{col}_j(C)$$

and for the case C is the identity matrix

$$\text{col}_j(B) = \text{col}_j(BI) = B\text{col}_j(I)$$

Using $PAQ = D$ so $AQ = P^{-1}D$ we have that

$$A\text{col}_j(Q) = P^{-1}\text{col}_j(D).$$

Now when $1 \leq j \leq k$ we have that $\text{col}_j(D) = \text{col}_j(I_{m \times m})$ and so

$$1 \leq j \leq k \text{ implies } A\text{col}_j(Q) = P^{-1}\text{col}_j(D) = P^{-1}\text{col}_j(I_{m \times m}) = \text{col}_j(P^{-1})$$

and thus $\text{col}_j(P^{-1}) \in \text{range}(A)$.

But when $k+1 \leq j \leq n$ we have that $\text{col}_j(D) = Z_{m \times 1}$ and so

$$k+1 \leq j \leq n \text{ implies } A\text{col}_j(Q) = P^{-1}\text{col}_j(D) = P^{-1}Z_{m \times 1} = Z_{m \times 1}$$

and thus $\text{col}_j(Q) \in \text{null}(A)$.

Since the columns of an invertible matrix are linearly independent we get that

1. $\text{col}_{k+1}(Q), \text{col}_{k+2}(Q), \dots, \text{col}_n(Q)$ are $n-k$ linearly independent vectors in $\text{null}(A)$ and
2. $\text{col}_1(P^{-1}), \text{col}_2(P^{-1}), \dots, \text{col}_k(P^{-1})$ are k linearly independent vectors in $\text{range}(A)$.

Hence, $\dim(\text{null}(A)) \geq n-k$ and $\dim(\text{range}(A)) \geq k$. But Theorem 16 implies that

$$n = \dim(\mathbb{R}^n) = \dim(\text{null}(A)) + \dim(\text{range}(A))$$

and therefore $\dim(\text{null}(A)) = n-k$ and $\dim(\text{range}(A)) = k$. (If either were larger they could not add to n .) But any set of m linearly independent vectors in a space of dimension m are a basis, so this proves (1) and (2).

To see (3):

$$\begin{aligned} Y \in \text{range}(A) & \text{ iff } \text{there exists } X \text{ } AX = Y \\ & \text{ iff } \text{there exists } X \text{ } PAX = PY \\ & \text{ iff } \text{there exists } X \text{ } PAQQ^{-1}X = PY \\ & \text{ iff } \text{there exists } U \text{ } PAQU = PY \\ & \text{ iff } \text{there exists } U \text{ } DU = PY \\ & \text{ iff } PY \in \text{range}(D) \end{aligned}$$

But $\text{range}(D)$ is exactly the set of all those vectors whose coordinates $\geq k+1$ are zero, i.e.

$$D \begin{bmatrix} s_1 \\ s_2 \\ \vdots \\ s_n \end{bmatrix} = \begin{bmatrix} t_1 \\ t_2 \\ \vdots \\ t_m \end{bmatrix} \text{ iff } s_1 = t_1, \dots, s_k = t_k \text{ and } t_{k+1} = 0, \dots, t_m = 0$$

That is to say $Y \in \text{range}(A)$ iff $PY \in \text{range}(D)$ iff $\text{entry}_j(PY) = 0$ for all j with $k+1 \leq j \leq m$.

To see (4). Just note that $DX = PY$ implies $PAQX = PY$ implies $AQX = Y$.

QED