Matricies Give Linear Transformations

Consider the vector spaces $U = \mathbb{R}^n$ and $V = \mathbb{R}^m$, expressing elements of each as column vectors.
If A is an m by n matrix we can define an associated function

$$L : U \to V$$

according to the rule

$$L(\vec{u}) = A\vec{u}$$

Theorem
The function L, as defined above, is a linear transformation.

Proof: We have that

$$\begin{aligned} L(a\vec{u}_1 + b\vec{u}_2) &= A(a\vec{u}_1 + b\vec{u}_2) \\ &= a(A\vec{u}_1) + b(A\vec{u}_2) \\ &= aL(\vec{u}_1) + bL(\vec{u}_2) \end{aligned}$$

It is not at all uncommon to not distinguish the matrix A from the function L.
So people will talk about a matrix A being a linear transformation.

Examples

The zero transformation between the vector spaces $\mathbb{R}^2$ and $\mathbb{R}^3$ can be described by the matrix

$$A = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}$$

indeed, we can see that the associated linear transformation

$$L(\vec{v}) = A\vec{v}$$

will satisfy

$$L((x, y)) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

which always gives the zero vector.

The identity transformation between the vector spaces $\mathbb{R}^3$ and $\mathbb{R}^3$ can be described by the matrix:

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

indeed, we can see that the associated linear transformation

$$L(\vec{v}) = A\vec{v}$$

will satisfy

$$L((x, y, z)) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

which sends every vector to itself.

The *inclusion transformations* of

$$\mathbb{R}^2 \simeq U = \{(x, y, 0) \in \mathbb{R}^3\} \subset \mathbb{R}^3$$

between the vector spaces $\mathbb{R}^2$ and $\mathbb{R}^3$ can be described by the matrix:

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}$$

indeed, we can see that the associated linear transformation

$$L(\vec{v}) = A\vec{v}$$

will satisfy

$$L((x, y)) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ y \\ 0 \end{pmatrix}$$

is precisely the image of (x, y) in $\mathbb{R}^3$ under the described inclusion.

If we write

$$V = \mathbb{R}^3 = U \oplus W$$

with

$$U = \{(x, y, 0) \in \mathbb{R}^3\} \qquad W = \{(0, 0, z) \in \mathbb{R}^3\}$$

The projection transformation

$$P_{V,U} : \mathbb{R}^3 \to U \simeq \mathbb{R}^2$$

an be described by the matrix:

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

indeed, we can see that the associated linear transformation

$$L(\vec{v}) = A\vec{v}$$

will satisfy

$$L((x, y, z)) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix}$$

is precisely the image of (x, y) in $\mathbb{R}^3$ under the described projection.

If L is the linear transformation $\mathbb{R}^2 \to \mathbb{R}^2$ associated to the matrix A given by

$$A = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$$

then the inverse linear transformation

$$L^{-1} : \mathbb{R}^2 \to \mathbb{R}^2$$

is associated to the matrix

$$A^{-1} = \begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix}$$

indeed, we can see that the associated linear transformation

$$L(\vec{v}) = A\vec{v}$$

will satisfy

$$L^{-1}(L(\vec{v})) = A^{-1}(L(\vec{v})) = A^{-1}A\vec{v} = \mathrm{Id}\vec{v} = \vec{v}$$

for all $\vec{v} \in \mathbb{R}^2$, and likewise

$$L(L^{-1}(\vec{v})) = A(L^{-1}(\vec{v})) = AA^{-1}\vec{v} = \mathrm{Id}\vec{v} = \vec{v}$$

If L is the linear transformation $\mathbb{R}^2 \to \mathbb{R}^2$ associated to the matrix A given by

$$A = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$$

and if M is the linear transformation $\mathbb{R}^2 \to \mathbb{R}^2$ associated to the matrix B given by

$$B = \begin{pmatrix} 1 & 3 \\ 0 & 2 \end{pmatrix}$$

then the linear transformations

$$(3L + 2M) : \mathbb{R}^2 \to \mathbb{R}^2$$

is associated to the matrix

$$3A + 2B = \begin{pmatrix} 5 & 12 \\ 0 & 7 \end{pmatrix}$$

The linear transformation

$$L \circ M$$

is associated to the matrix

$$AB = \begin{pmatrix} 1 & 7 \\ 0 & 2 \end{pmatrix}$$

Both of the above can be checked in the same way

If a transformation comes from a matrix, everything we do transformations corresponds to doing something with matricies and vice versa.

certain linear transformation $L : \mathbb{R}^m \to \mathbb{R}^n$	$\leftrightarrow$	n by m matrix
zero transformation	$\leftrightarrow$	zero matrix
identity Transformation	$\leftrightarrow$	identity matrix
inclusion/projections	$\leftrightarrow$	specific matricies
linear combinations of transformations	$\leftrightarrow$	adding matricies
composition of transformations	$\leftrightarrow$	multiplying matricies
inverses of transformations	$\leftrightarrow$	inverses of matricies

We will eventually prove these things really do correspond in general.

Linear transformations for $\mathbb{R}^n$ don't seem any worse than matricies.
What if there are linear transformations not described by matricies?

We want to prove all linear transformations from $U = \mathbb{R}^n$ to $V = \mathbb{R}^m$ come from m by n matricies.

Two key facts you will prove on the assignment:

Theorem
Two linear transformations $L_1 : U \to V$ and $L_2 : U \to V$ are equal, if for any basis $\vec{e}_1, \ldots, \vec{e}_n$ we have

$$L_1(\vec{e}_1) = L_2(\vec{e}_1), \quad L_1(\vec{e}_2) = L_2(\vec{e}_2), \quad \cdots \quad L_1(\vec{e}_n) = L_2(\vec{e}_n)$$

Theorem
If $\vec{e}_1, \ldots, \vec{e}_n$ is any basis for U, and $\vec{w}_1, \ldots, \vec{w}_n$ any vectors in V, then there is a unique linear transformation $L : U \to V$ so that

$$L(\vec{e}_1) = \vec{w}_1, \quad L(\vec{e}_2) = \vec{w}_2, \quad \cdots \quad L(\vec{e}_n) = \vec{w}_n$$

Lemma

Let A be an n by m matrix, and denote by $\vec{a}_i$ the columns of A, so

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{pmatrix} = \begin{pmatrix} \vdots & \vdots & & \vdots \\ \vec{a}_1 & \vec{a}_2 & \cdots & \vec{a}_m \\ \vdots & \vdots & & \vdots \end{pmatrix} \qquad \vec{a}_i = \begin{pmatrix} a_{1i} \\ a_{2i} \\ \vdots \\ a_{ni} \end{pmatrix}$$

Then the linear transformation $L : \mathbb{R}^m \to \mathbb{R}^n$ is **the** transformation which takes the basis vectors

$$\vec{e}_1 = (1, 0, \ldots, 0), \quad \vec{e}_2 = (0, 1, \ldots, 0), \quad \ldots, \quad , \vec{e}_m = (0, 0, \ldots, 1)$$

to $\vec{a}_1, \ldots, \vec{a}_m$. That is $L(\vec{e}_i) = \vec{a}_i$

Proof (sketch):

from the assignment There is a unique linear transformation $M : \mathbb{R}^m \to \mathbb{R}^n$ for which $M(\vec{e}_i) = \vec{a}_i$.

So, we really just need to check that this one does that, and indeed,

$$L(\vec{e}_i) = A\vec{e}_i = \begin{pmatrix} \vdots & \vdots & & \vdots \\ \vec{a}_1 & \vec{a}_2 & \cdots & \vec{a}_m \\ \vdots & \vdots & & \vdots \end{pmatrix} \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \leftarrow \text{ith spot} = \begin{pmatrix} a_{1i} \\ a_{2i} \\ \vdots \\ a_{ni} \end{pmatrix} = \vec{a}_i$$

The key, is that when you multiply a matrix, by $\vec{e}_i$, you obtain the i-th column of A.

Matricies Give All Linear Transformations

Consider the vector spaces $U = \mathbb{R}^n$ and $V = \mathbb{R}^m$.

Theorem

If $L : U \to V$ is a linear transformation, then there exists a unique m by n matrix A so that:

$$L(\vec{u}) = A\vec{u}$$

Moreover, we can determine the i-th column $\vec{a}_i$ of A by computing

$$\vec{a}_i = L(\vec{e}_i)$$

for $\vec{e}_i = (0, \dots, 0, 1, 0, \dots, 0)$ the vector which has a 1 in precisely the ith coordinate.

Proof

We can definitely write down the matrix A, whose columns are $\vec{a}_i = L(\vec{e}_i)$.
By the lemma, there is a linear transformation M, associated to A given by:

$$M(\vec{v}) = A\vec{v}$$

But we have

$$M(\vec{e}_i) = \vec{a}_i = L(\vec{e}_i)$$

so these agree on the basis, so by our lemma they must be equal, $L = M$ so L comes from a matrix.

Example - Matrix of a linear transformation

We know from that result on the assignment, that we can specify a linear transformation by talking about what it does to any basis, so what happens if some jerk (hi guys!) doesn't use the basis $\vec{e}_i$?

Notice that $(1,2)$, $(3,4)$ is a basis of $\mathbb{R}^2$ (Why?)

Suppose $L : \mathbb{R}^2 \to \mathbb{R}^3$ is a linear transformation which satisfies

$$L((1,2)) = (1,1,1) \qquad L((3,4)) = (1,2,3)$$

what is the matrix for L?

First question, what is $L((1,0))$? then what is $L((0,1))$?

This sort of process is pretty much always works if the person gave you enough information to determine a unique transformation.

Linear Transformations $V \to W$ between finite dimensional vector spaces.

It is often convenient when doing concrete problems about a finite dimensional vector space to use the identification

$$a_1\vec{f}_1 + \cdots + a_n\vec{f}_n \leftrightarrow (a_1, \ldots, a_n)$$

If I pretend I am working with $\mathbb{R}^n$, can I pretend my linear transformations are matricies?

$$L : V \to W \qquad \Leftrightarrow \qquad A : \mathbb{R}^n \to \mathbb{R}^m$$

Linear Transformations $V \to W$ between finite dimensional vector spaces.

Suppose V is a finite dimensional vector space with a basis $\vec{f}_1, \ldots, \vec{f}_n$.
Suppose W is a finite dimensional vector space with a basis $\vec{g}_1, \ldots, \vec{g}_m$.

Then I can identify V with $\mathbb{R}^n$ and W with $\mathbb{R}^m$ by using:

$$L_V(a_1\vec{f}_1 + \cdots + a_n\vec{f}_n) \to (a_1, \ldots, a_n) \qquad L_W(b_1\vec{g}_1 + \cdots + b_m\vec{g}_m) \to (b_1, \ldots, b_m)$$

Now, every linear transformation $M : V \to W$ leads to a transformation

$$M' = L_W \circ M \circ L_V^{-1} : \mathbb{R}^n \to \mathbb{R}^m$$

So M' is a matrix!!

Every linear transformation (matrix) $M' : \mathbb{R}^n \to \mathbb{R}^m$ leads to a transformation

$$M'' = L_W^{-1} \circ M' \circ L_V : V \to W$$

Linear transformations give matricies and matricies give linear transformations!

Important: This association of a matrix depends on the choices of basis!!

This association gives a bijection

Key Idea The association between linear transformations and matricies is a well defined bijection once we pick a basis.

Theorem

The function $R : \mathrm{Hom}(V, W) \to \mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^m)$ given by $R : (M) = M'$ is a bijection with inverse given by $S : (M') = M''$.

Proof (sketch)

To check that its a bijection we only need to check that the compositions

$$R \circ S(M) = M \qquad S \circ R(M') = M'$$

indeed

$$R \circ S(M) = R(L_W \circ M \circ L_V^{-1}) = L_W^{-1} \circ (L_W \circ M \circ L_V^{-1}) \circ L_V = M$$

The other direction is similar.

If we start with M, we shall call the associated matrix A, **the matrix associated to** M with respect to the basis $\vec{f}_1, \ldots, \vec{f}_n$ for V and $\vec{g}_1, \ldots, \vec{g}_m$ for W.

Conversely, if we start with A, we shall call the associated linear transformation M, **the linear transformation associated to** A with respect to the basis $\vec{f}_1, \ldots, \vec{f}_n$ for V and $\vec{g}_1, \ldots, \vec{g}_m$ for W.

Concretely Associating a Matrix to a Linear Transformation

Suppose V and W are vector spaces with a basis $\vec{f}_1, \ldots, \vec{f}_n$ for V and $\vec{g}_1, \ldots, \vec{g}_m$ for W.

Lemma

If $M : V \to W$ is any linear transformation, then the matrix associated to M, with respect to the basis $\vec{f}_1, \ldots, \vec{f}_n$ for V and $\vec{g}_1, \ldots, \vec{g}_m$ for W is a matrix

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & & \\ \vdots & & \ddots & \vdots \\ a_{m1} & & \cdots & a_{mn} \end{pmatrix}$$

which satisfies

$$M(\vec{f}_j) = a_{1j}\vec{g}_1 + \cdots + a_{mj}\vec{g}_m = \sum_{i=1}^{m} a_{ij}\vec{g}_i$$

Conversely, given a matrix A, as above, the linear transformation $S(M'_A)$ associated to it with respect to the basis $\vec{f}_1, \ldots, \vec{f}_n$ for V and $\vec{g}_1, \ldots, \vec{g}_m$ for W is the unique one which satisfies:

$$M(\vec{f}_j) = a_{1j}\vec{g}_1 + \cdots + a_{mj}\vec{g}_m = \sum_{i=1}^{m} a_{ij}\vec{g}_i$$

Proof to come!

The above equation, $M(\vec{f}_j) = a_{1j}\vec{g}_1 + \cdots + a_{mj}\vec{g}_m = \sum_{i=1}^{m} a_{ij}\vec{g}_i$ therefore characterizes how we associate matricies to linear transformations **given a choice of basis**.

Connection to $\mathbb{R}^n$

It is worth noticing that the above lemma even applies directly to

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

indeed, it just says in this case that, when we use $\vec{e}_j$

$$\begin{aligned} L(\vec{e}_j) &= a_{1j}(1, 0, \ldots, 0) + a_{2j}(0, 1, \ldots, 0) + \cdots + a_{mj}(0, 0, \ldots, 1) \\ &= (a_{1j}, a_{2j}, \ldots, a_{mj}) \\ &= \vec{a}_j \end{aligned}$$

where $\vec{a}_j$ is as before the j-th column of A.

Proof of the lemma

Proof: We know that if $M : V \to W$ is a linear transformation, the matrix, A, we should be associating is the one from

$$M' = L_W \circ M \circ L_V^{-1} : \mathbb{R}^n \to \mathbb{R}^m$$

We recall (from what we just saw) that

$$L_W \circ M \circ L_V^{-1}(\vec{e}_j) = M'(\vec{e}_j) = (a_{1j}, a_{2j}, \dots, a_{mj})$$

where $\vec{a}_j = (a_{1j}, a_{2j}, \dots, a_{mj})$ is the j-th column of A. (This is just the definition of the associated matrix)

Now we compute

$$\begin{aligned}
M(\vec{f}_j) &= L_W^{-1} \circ L_W(M(\vec{f}_j)) && \text{composition with identity} \\
&= L_W^{-1} \circ L_W(M(L_V^{-1}(\vec{e}_j))) && L_V^{-1}(\vec{e}_j) = \vec{f}_j \\
&= L_W^{-1}((a_{1j}, a_{2j}, \dots, a_{mj})) && \text{from above} \\
&= a_{1j}\vec{g}_1 + \cdots + a_{mj}\vec{g}_m && \text{definition of } L_W^{-1}
\end{aligned}$$

This proves the one direction, the proof of the other direction can be done similarly, or by using the fact that S and R were inverses, we leave this as an exercise.

Summary picture of what is being described

This picture is how we associate a matrix to a transformation and a choice of bases

Suppose U has basis $\vec{e}_1, \dots, \vec{e}_n$ and V has basis $\vec{f}_1, \dots, \vec{f}_m$.

The matrix for a linear transformation $L : U \to V$ has the form:

$$
\begin{array}{ccccccccc}
L(\vec{e}_1) & L(\vec{e}_2) & L(\vec{e}_3) & \cdots & L(\vec{e}_j) & \cdots & L(\vec{e}_n) & \\
\downarrow & \downarrow & \downarrow & & \downarrow & & \downarrow & \\
\end{array}
$$

$$
\begin{array}{c}
\begin{pmatrix}
a_{11} & a_{12} & a_{13} & \cdots & a_{1j} & \cdots & a_{1n} \\
a_{21} & a_{22} & a_{23} & \cdots & a_{2j} & \cdots & a_{2n} \\
a_{31} & a_{32} & a_{33} & \cdots & a_{3j} & \cdots & a_{3n} \\
\vdots & \vdots & \vdots & & \vdots & & \vdots \\
a_{i1} & a_{i2} & a_{i3} & \cdots & a_{ij} & \cdots & a_{in} \\
\vdots & \vdots & \vdots & & \vdots & & \vdots \\
a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mj} & \cdots & a_{mn}
\end{pmatrix}
\end{array}
\begin{array}{l}
\leftarrow \vec{f}_1 \\
\leftarrow \vec{f}_2 \\
\leftarrow \vec{f}_3 \\
\vdots \\
\leftarrow \vec{f}_i \\
\vdots \\
\leftarrow \vec{f}_m
\end{array}
$$

According to the rule

$$L(\vec{e}_j) = \overbrace{a_{1j}\vec{f}_1 + a_{2j}\vec{f}_2 + a_{3j}\vec{f}_3 + \cdots + a_{ij}\vec{f}_i + \cdots + a_{mj}\vec{f}_m}^{\text{gives entries } j\text{th column}}$$

Because $\vec{f}_i$ are a basis, this expression is **unique**!

Because $\vec{e}_j$ are a basis, this expression **determines** L

Example

If V has a basis $\vec{e}_1, \vec{e}_2, \vec{e}_3$ and W has a basis $\vec{f}_1, \vec{f}_2, \vec{f}_3$ and the linear transformations

$$L : V \to W$$

satisfies

$$L(\vec{e}_1) = 2\vec{f}_1 + 3\vec{f}_3, \qquad L(\vec{e}_2) = 2\vec{f}_1 + \vec{f}_2 \qquad L(\vec{e}_3) = \vec{f}_2$$

then the matrix assoicated to L with respect to the basis $\vec{e}_1, \vec{e}_2, \vec{e}_3$ for V and $\vec{f}_1, \vec{f}_2, \vec{f}_3$ for W is

$$\begin{pmatrix} 2 & 2 & 0 \\ 0 & 1 & 1 \\ 3 & 0 & 0 \end{pmatrix}$$

It is important to realize if we use a different basis for either side, then we really should expect the matrix to change, how so, is something we will discuss in the future.

Example

If V has a basis $\vec{e}_1, \vec{e}_2, \vec{e}_3$ and the linear transformations

$$L : V \to V$$

satisfies

$$L(\vec{e}_1) = 2\vec{e}_1 + 3\vec{e}_3, \qquad L(\vec{e}_2) = 2\vec{e}_1 + \vec{e}_2 \qquad L(\vec{e}_3) = \vec{e}_2$$

then the matrix assoicated to L with respect to the basis $\vec{e}_1, \vec{e}_2, \vec{e}_3$ (for both the domain and codomain) is

$$\begin{pmatrix} 2 & 2 & 0 \\ 0 & 1 & 1 \\ 3 & 0 & 0 \end{pmatrix}$$

It is important to realize if the map goes from V to V you typically want the same basis for the domain and codomain.

Natural Questions About Linear Transformations/Matricies

- Given some description of a linear transformation $L : \mathbb{R}^n \to \mathbb{R}^m$, what is the matrix for L?
  The key is to figuring out where the standard basis goes, what this involves depends on what you were told, but usually involves solving a system of equations.
- Given some description of a linear transformation $L : V \to W$, and identifications of V, W with $\mathbb{R}^n$ and $\mathbb{R}^m$, what is the associated matrix for L?
  The key is to figuring out where the basis goes, what this involves depends on what you were told, but usually involves solving a system of equations.
- How does the identification of V and W with $\mathbb{R}^n$ and $\mathbb{R}^m$ effect the matrix for L?
  We will come back to this, it is also on the assignment.