Math 115A - Week 4
Textbook sections: 2.3-2.4
Topics covered:

- A quick review of matrices
- Co-ordinate matrices and composition
- Matrices as linear transformations
- Invertible linear transformations (isomorphisms)
- Isomorphic vector spaces

* * * * *

A quick review of matrices

- An $m \times n$ *matrix* is a collection of $mn$ scalars, organized into $m$ rows and $n$ columns:

$$A = \begin{pmatrix} A_{11} & A_{12} & \dots & A_{1n} \\ A_{21} & A_{22} & \dots & A_{2n} \\ & & \vdots & \\ A_{m1} & A_{m2} & \dots & A_{mn} \end{pmatrix}.$$

  If $A$ is a matrix, then $A_{jk}$ refers to the scalar entry in the $j^{th}$ row and $k^{th}$ column. Thus if

$$A := \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$$

  then $A_{11} = 1$, $A_{12} = 2$, $A_{21} = 3$, and $A_{22} = 4$.

- (The word "matrix" is late Latin for "womb"; it is the same root as maternal or matrimony. The idea being that a matrix is a receptacle for holding numbers. Thus the title of the recent Hollywood movie "the Matrix" is a play on words).

- A special example of a matrix is the $n \times n$ *identity matrix* $I_n$, defined by

$$I_n := \begin{pmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ & & \vdots & \\ 0 & 0 & \ldots & 1 \end{pmatrix}$$

or equivalently that $(I_n)_{jk} := 1$ when $j = k$ and $(I_n)_{jk} := 0$ when $j \neq k$.

- If $A$ and $B$ are two $m \times n$ matrices, the sum $A + B$ is another $m \times n$ matrix, defined by adding each component separately, for instance

$$(A + B)_{11} := A_{11} + B_{11}$$

and more generally

$$(A + B)_{jk} := A_{jk} + B_{jk}.$$

If $A$ and $B$ have different shapes, then $A + B$ is left undefined.

- The scalar product $cA$ of a scalar $c$ and a matrix $A$ is defined by multiplying each component of the matrix by $c$:

$$(cA)_{jk} := cA_{jk}.$$

- If $A$ is an $m \times n$ matrix, and $B$ is an $l \times m$ matrix, then the matrix product $BA$ is an $l \times n$ matrix, whose co-ordinates are given by the formula

$$(BA)_{jk} = B_{j1}A_{1k} + B_{j2}A_{2k} + \ldots + B_{jm}A_{mk} = \sum_{i=1}^{m} B_{ji}A_{ik}.$$

Thus for instance if

$$A := \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$$

and

$$B := \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}$$

then

$$(BA)_{11} = B_{11}A_{11} + B_{12}A_{21}; \quad (BA)_{12} = B_{11}A_{12} + B_{12}A_{22}$$

$$(BA)_{21} = B_{21}A_{11} + B_{22}A_{21}; \quad (BA)_{22} = B_{21}A_{12} + B_{22}A_{22}$$

and so

$$BA = \begin{pmatrix} B_{11}A_{11} + B_{12}A_{21} & B_{11}A_{12} + B_{12}A_{22} \\ B_{21}A_{11} + B_{22}A_{21} & B_{21}A_{12} + B_{22}A_{22} \end{pmatrix}$$

or in other words

$$\begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix} \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} = \begin{pmatrix} B_{11}A_{11} + B_{12}A_{21} & B_{11}A_{12} + B_{12}A_{22} \\ B_{21}A_{11} + B_{22}A_{21} & B_{21}A_{12} + B_{22}A_{22} \end{pmatrix}.$$

If the number of columns of $B$ does not equal the number of rows of $A$, then $BA$ is left undefined. Thus for instance it is possible for $BA$ to be defined while $AB$ remains undefined.

- This matrix multiplication rule may seem strange, but we will explain why it is natural below.

- It is an easy exercise to show that if $A$ is an $m \times n$ matrix, then $I_m A = A$ and $AI_n = A$. Thus the matrices $I_m$ and $I_n$ are multiplicative identities, assuming that the shapes of all the matrices are such that matrix multiplication is defined.

* * * * *

Co-ordinate matrices and composition

- Last week, we introduced the notion of a linear transformation $T : X \to Y$. Given two linear transformations $T : X \to Y$ and $S : Y \to Z$, where the target space of $T$ matches up with the initial space of $S$, their *composition* $ST : X \to Z$, defined by

$$ST(v) = S(Tv)$$

is also a linear transformation; this is easy to check and I'll leave it as an exercise. Also, if $I_X : X \to X$ is the identity on $X$ and $I_Y : Y \to Y$ is the identity on $Y$, it is easy to check that $TI_X = T$ and $I_Y T = T$.

- **Example** Suppose we are considering combinations of two molecules: methane $CH_4$ and water $H_2O$. Let $X$ be the space of all linear combinations of such molecules, thus $X$ is a two-dimensional space with $\alpha := (methane, water)$ as an ordered basis. (A typical element of $X$ might be $3 \times methane + 2 \times water$). Let $Y$ be the space of all linear combinations of Hydrogen, Carbon, and Oxygen atoms; this is a three-dimensional space with $\beta := (hydrogen, carbon, oxygen)$ as an ordered basis. Let $Z$ be the space of all linear combinations of electrons, protons, and neutrons, thus it is a three-dimensional space with $\gamma := (electron, proton, neutron)$ as a basis. There is an obvious linear transformation $T : X \to Y$, defined by starting with a collection of molecules and breaking them up into component atoms. Thus

$$T(methane) = 4 \times hydrogen + 1 \times carbon$$

$$T(water) = 2 \times hydrogen + 1 \times oxygen$$

and so $T$ has the matrix

$$[T]_\alpha^\beta = [T]_{(methane,water)}^{(hydrogen,carbon,oxygen)} = \begin{pmatrix} 4 & 2 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Similarly, there is an obvious linear transformation $S : Y \to Z$, defined by starting with a collection of atoms and breaking them up into component particles. Thus

$$S(hydrogen) = 1 \times electron + 1 \times proton$$

$$S(carbon) = 6 \times electron + 6 \times proton + 6 \times neutron$$

$$S(oxygen) = 8 \times electron + 8 \times proton + 8 \times neutron.$$

Thus

$$[S]_\beta^\gamma = [S]_{(hydrogen,carbon,oxygen)}^{(electron,proton,neutron)} = \begin{pmatrix} 1 & 6 & 8 \\ 1 & 6 & 8 \\ 0 & 6 & 8 \end{pmatrix}.$$

The composition $ST : X \to Z$ of $S$ and $T$ is thus the transformation which sends molecules to their component particles. (Note that even though $S$ is to the left of $T$, the operation $T$ is applied *first*. This rather unfortunate fact occurs because the conventions of mathematics place the operator $T$ before the operand $x$, thus we have $T(x)$ instead of $(x)T$. Since all the conventions are pretty much entrenched, there's not much we can do about it). A brief calculation shows that

$$ST(methane) = 10 \times electron + 10 \times proton + 6 \times neutron$$

$$ST(water) = 10 \times electron + 10 \times proton + 8 \times neutron$$

and hence

$$[ST]_\alpha^\gamma = [ST]_{(methane,water)}^{(electron,proton,neutron)} = \begin{pmatrix} 10 & 10 \\ 10 & 10 \\ 6 & 8 \end{pmatrix}.$$

Now we ask the following question: how are these matrices $[T]_\alpha^\beta$, $[S]_\beta^\gamma$, and $[ST]_\alpha^\gamma$ related?

- Let's consider the 10 entry on the top left of $[ST]_\alpha^\gamma$. This number measures how many electrons there are in a methane molecule. From the matrix of $[T]_\alpha^\beta$ we see that each methane molecule has 4 hydrogen, 1 carbon, and 0 oxygen atoms. Since hydrogen has 1 electron, carbon has 6, and oxygen has 8, we see that the number of electrons in methane is

$$4 \times 1 + 1 \times 6 + 0 \times 8 = 10.$$

Arguing similarly for the other entries of $[ST]_\alpha^\gamma$, we see that

$$[ST]_\alpha^\gamma = \begin{pmatrix} 4 \times 1 + 1 \times 6 + 0 \times 8 & 2 \times 1 + 0 \times 6 + 1 \times 8 \\ 4 \times 1 + 1 \times 6 + 0 \times 8 & 2 \times 1 + 0 \times 6 + 1 \times 8 \\ 4 \times 0 + 1 \times 6 + 0 \times 8 & 2 \times 0 + 0 \times 6 + 1 \times 8 \end{pmatrix}.$$

But this is just the matrix product of $[S]_\beta^\gamma$ and $[T]_\alpha^\beta$:

$$[ST]_\alpha^\gamma = \begin{pmatrix} 1 & 6 & 8 \\ 1 & 6 & 8 \\ 0 & 6 & 8 \end{pmatrix} \begin{pmatrix} 4 & 2 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} = [S]_\beta^\gamma [T]_\alpha^\beta.$$

- More generally, we have

- **Theorem 1.** Suppose that $X$ is $l$-dimensional and has an ordered basis $\alpha = (u_1, \ldots, u_l)$, $Y$ is $m$-dimensional and has an ordered basis $\beta = (v_1, \ldots, v_m)$, and $Z$ is $n$-dimensional and has a basis $\gamma$ of $n$ elements. Let $T : X \to Y$ and $S : Y \to Z$ be linear transformations. Then

$$[ST]^\gamma_\alpha = [S]^\gamma_\beta [T]^\beta_\alpha.$$

- **Proof.** The transformation $T$ has a co-ordinate matrix $[T]^\beta_\alpha$, which is an $m \times l$ matrix. If we write

$$[T]^\beta_\alpha =:= \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1l} \\ a_{21} & a_{22} & \ldots & a_{2l} \\ & & \vdots & \\ a_{m1} & a_{m2} & \ldots & a_{ml} \end{pmatrix}$$

then we have

$$\begin{array}{ll} Tu_1 & = a_{11}v_1 + a_{21}v_2 + \ldots + a_{m1}v_m \\ Tu_2 & = a_{12}v_1 + a_{22}v_2 + \ldots + a_{m2}v_m \\ & \vdots \\ Tu_l & = a_{1l}v_1 + a_{2l}v_2 + \ldots + a_{ml}v_m \end{array}$$

We write this more compactly as

$$Tu_i = \sum_{j=1}^m a_{ji} v_j \text{ for } i = 1, \ldots, l.$$

- Similarly, $S$ has a co-ordinate matrix $[S]^\gamma_\beta$, which is an $n \times m$ matrix. If

$$[S]^\gamma_\beta := \begin{pmatrix} b_{11} & b_{12} & \ldots & b_{1m} \\ b_{21} & b_{22} & \ldots & b_{2m} \\ & & \vdots & \\ b_{n1} & b_{m2} & \ldots & b_{nm} \end{pmatrix}.$$

then

$$Sv_j = \sum_{k=1}^n b_{kj} w_k \text{ for } j = 1, \ldots, m.$$

Now we try to understand how $ST$ acts on the basis $u_1, \ldots, u_l$. Applying $S$ to both sides of the $T$ equations, and using the fact that $S$ is linear, we obtain

$$STu_i = \sum_{j=1}^{m} a_{ji} Sv_j.$$

Applying our formula for $Sv_j$, we obtain

$$STu_i = \sum_{j=1}^{m} a_{ji} \sum_{k=1}^{n} b_{kj} w_k$$

which we can rearrange as

$$STu_i = \sum_{k=1}^{n} (\sum_{j=1}^{m} b_{kj} a_{ji}) w_k.$$

Thus if we define

$$c_{ki} := \sum_{j=1}^{m} b_{kj} a_{ji} = b_{k1} a_{1i} + b_{k2} a_{2i} + \ldots + b_{km} a_{mi}$$

then we have

$$STu_i = \sum_{k=1}^{n} c_{ki} w_k$$

and hence

$$[ST]_\alpha^\gamma = \begin{pmatrix} c_{11} & c_{12} & \ldots & c_{1l} \\ c_{21} & c_{22} & \ldots & c_{2l} \\ & & \vdots & \\ c_{n1} & c_{m2} & \ldots & c_{nl} \end{pmatrix}.$$

However, if we perform the matrix multiplication

$$\begin{pmatrix} b_{11} & b_{12} & \ldots & b_{1m} \\ b_{21} & b_{22} & \ldots & b_{2m} \\ & & \vdots & \\ b_{n1} & b_{m2} & \ldots & b_{nm} \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1l} \\ a_{21} & a_{22} & \ldots & a_{2l} \\ & & \vdots & \\ a_{m1} & a_{m2} & \ldots & a_{ml} \end{pmatrix}$$

we get exactly the same matrix (this is because of our formula for $c_{ki}$ in terms of the $b$ and $a$ co-efficients). This proves the theorem. $\square$

- This theorem illustrates *why* matrix multiplication is defined in that strange way - multiplying rows against columns, etc. It also explains why we need the number of columns of the left matrix to equal the number of rows of the right matrix; this is like how to compose two transformations $T : X \to Y$ and $S : Y \to Z$ to form a transformation $ST : X \to Z$, we need the target space of $T$ to equal to the initial space of $S$.

* * * * *

Comparison between linear transformations and matrices

- To summarize what we have done so far:

- Given a vector space $X$ and an ordered basis $\alpha$ for $X$, one can write vectors $v$ in $V$ as column vectors $[v]^\alpha$. Given two vector spaces $X$, $Y$, and ordered bases $\alpha$, $\beta$ for $X$ and $Y$ respectively, we can write linear transformations $T : X \to Y$ as matrices $[T]^\beta_\alpha$. The action of $T$ then corresponds to matrix multiplication by $[T]^\gamma_\beta$:

$$[Tv]^\beta = [T]^\beta_\alpha [v]^\alpha;$$

  i.e. we can "cancel" the basis $\alpha$. Similarly, composition of two linear transformations corresponds to matrix multiplication: if $S : Y \to Z$ and $\gamma$ is an ordered basis for $Z$, then

$$[ST]^\gamma_\alpha = [S]^\gamma_\beta [T]^\beta_\alpha$$

  i.e. we can "cancel" the basis $\beta$.

- Thus, by using bases, one can understand the behavior of linear transformations in terms of matrix multiplication. This is not quite saying that linear transformations are the same as matrices, for two reasons: firstly, this correspondence only works for finite dimensional spaces $X$, $Y$, $Z$; and secondly, the matrix you get depends on the basis you choose - a single linear transformation can correspond to many different matrices, depending on what bases one picks.

- To clarify the relationship between linear transformations and matrices let us once again turn to the scalar case, and now consider currency conversions. Let $X$ be the space of US currency - this is the one-dimensional space which has $(dollar)$ as an (ordered) basis; $(cent)$ is also a basis. Let $Y$ be the space of British currency (with $(pound)$ or $(penny)$ as a basis; $pound = 100 \times penny$), and let $Z$ be the space of Japanese currency (with $(yen)$ as a basis). Let $T : X \to Y$ be the operation of converting US currency to British, and $S : Y \to Z$ the operation of converting British currency to Japanese, thus $ST : X \to Z$ is the operation of converting US currency to Japanese (via British).

- Suppose that one dollar converted to half a pound, then we would have

$$[T]^{(pound)}_{(dollar)} = (0.5),$$

or in different bases

$$[T]^{(pound)}_{(cent)} = (0.005); \quad [T]^{(penny)}_{(cent)} = (0.5); \quad [T]^{(penny)}_{(dollar)} = (50).$$

Thus the same linear transformation $T$ corresponds to many different $1 \times 1$ matrices, depending on the choice of bases both for the domain $X$ and the range $Y$. However, conversion works properly no matter what basis you pick (as long as you are consistent), e.g.

$$[v]^{(dollar)} = (6) \Rightarrow [Tv]^{(pound)} = [T]^{(pound)}_{(dollar)} [v]^{(dollar)} = (0.5)(6) = (3).$$

Furthermore, if each pound converted to 200 yen, so that

$$[S]^{(yen)}_{(pound)} = (200)$$

then we can work out the various matrices for $ST$ by matrix multiplication (which in the $1 \times 1$ case is just scalar multiplication):

$$[ST]^{(yen)}_{(dollar)} = [S]^{(yen)}_{(pound)} [T]^{(pound)}_{(dollar)} = (200)(0.5) = (100).$$

One can of course do this computation in different bases, but still get the same result, since the intermediate basis just cancels itself out at the end:

$$[ST]^{(yen)}_{(dollar)} = [S]^{(yen)}_{(penny)} [T]^{(penny)}_{(dollar)} = (2)(50) = (100)$$

etc.

- You might amuse yourself concocting a vector example of currency conversion - for instance, suppose that in some country there was more than one type of currency, and they were not freely interconvertible. A US dollar might then convert to $x$ amounts of one currency plus $y$ amounts of another, and so forth. Then you could repeat the above computations except that the scalars would have to be replaced by various vectors and matrices.

- One basic example of a linear transformation is the *identity transformation* $I_V : V \to V$ on a vector space $V$, defined by $I_V v = v$. If we pick any basis $\beta = (v_1, \ldots, v_n)$ of $V$, then of course we have

$$I_V v_1 = 1 \times v_1 + 0 \times v_2 + \ldots + 0 \times v_n$$

$$I_V v_2 = 0 \times v_1 + 1 \times v_2 + \ldots + 0 \times v_n$$

$$\ldots$$

$$I_V v_n = 0 \times v_1 + 0 \times v_2 + \ldots + 1 \times v_n$$

and thus

$$[I_V]_\beta^\beta = \begin{pmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ & & \vdots & \\ 0 & 0 & \ldots & 1 \end{pmatrix} = I_n.$$

Thus the identity transformation is connected to the identity matrix.

* * * * *

Matrices as linear transformations.

- We have now seen how linear transformations can be viewed as matrices (after selecting bases, etc.). Conversely, every matrix can be viewed as a linear transformation.

- **Definition** Let $A$ be an $m \times n$ matrix. Then we define the linear transformation $L_A : \mathbf{R}^n \to \mathbf{R}^m$ by the rule

$$L_A x := Ax \text{ for all } x \in \mathbf{R}^n,$$

where we think of the vectors in $\mathbf{R}^n$ and $\mathbf{R}^m$ as column vectors.

- **Example** Let $A$ be the matrix

$$A := \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix}.$$

Then $L_A : \mathbf{R}^2 \to \mathbf{R}^3$ is the linear transformation

$$L_A \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 + 2x_2 \\ 3x_1 + 4x_2 \\ 5x_1 + 6x_2 \end{pmatrix}.$$

- It is easily checked that $L_A$ is indeed linear. Thus for every $m \times n$ matrix $A$ we can associate a linear transformation $L_A : \mathbf{R}^n \to \mathbf{R}^m$. Conversely, if we let $\alpha$ be the standard basis for $\mathbf{R}^n$ and $\beta$ be the standard basis for $\mathbf{R}^m$, then for every linear transformation $T : \mathbf{R}^n \to \mathbf{R}^m$ we can associate an $m \times n$ matrix $[T]_\alpha^\beta$. The following simple lemma shows that these two operations invert each other:

- **Lemma 2.** Let the notation be as above. If $A$ is an $m \times n$ matrix, then $[L_A]_\alpha^\beta = A$. If $T : \mathbf{R}^n \to \mathbf{R}^m$ is a linear transformation, then $L_{[T]_\alpha^\beta} = T$.

- **Proof** Let $\alpha = (e_1, e_2, \ldots, e_n)$ be the standard basis of $\mathbf{R}^n$. For any column vector

$$x = \begin{pmatrix} x_1 \\ \ldots \\ x_n \end{pmatrix}$$

in $\mathbf{R}^n$, we have

$$x = x_1 e_1 + \ldots x_n e_n$$

and thus

$$[x]^\alpha = \begin{pmatrix} x_1 \\ \ldots \\ x_n \end{pmatrix} = x.$$

Thus $[x]^\alpha = x$ for all $x \in \mathbf{R}^n$. Similarly we have $[y]^\beta = y$ for all $y \in \mathbf{R}^m$.

- Now let $A$ be an $m \times n$ matrix, and let $x \in \mathbf{R}^n$. By definition

$$L_A x = Ax$$

On the other hand, we have

$$[L_A x]^\beta = [L_A]^\beta_\alpha [x]^\alpha$$

and hence (by the previous discussion)

$$L_A x = [L_A]^\beta_\alpha x.$$

Thus

$$[L_A]^\beta_\alpha x = Ax \text{ for all } x \in \mathbf{R}^n.$$

If we apply this with $x$ equal to the first basis vector $\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$, we see that the first column of the matrices $[L_A]^\beta_\alpha$ and $A$ are equal. Similarly we see that all the other columns of $[L_A]^\beta_\alpha$ and $A$ match, so that $[L_A]^\beta_\alpha = A$ as desired.

- Now let $T : \mathbf{R}^n \to \mathbf{R}^m$ be a linear transformation. Then for any $x \in \mathbf{R}^n$

$$[Tx]^\beta = [T]^\beta_\alpha [x]^\alpha$$

which by previous discussion implies that

$$Tx = [T]^\beta_\alpha x = L_{[T]^\beta_\alpha} x.$$

Thus $T$ and $L_{[T]^\beta_\alpha}$ are the same linear transformation, and the lemma is proved. □

- Because of the above lemma, any result we can say about linear transformations, one can also say about matrices. For instance, the following result is trivial for linear transformations:

- **Lemma 3. (Composition is associative)** Let $T : X \to Y$, $S : Y \to Z$, and $R : Z \to W$ be linear transformations. Then we have $R(ST) = (RS)T$.

- **Proof.** We have to show that $R(ST)(x) = (RS)T(x)$ for all $x \in X$. But by definition

$$R(ST)(x) = R((ST)(x)) = R(S(T(x)) = (RS)(T(x)) = (RS)T(x)$$

as desired. □

- **Corollary 4. (Matrix multiplication is associative)** Let $A$ be an $m \times n$ matrix, $B$ be a $l \times m$ matrix, and $C$ be a $k \times l$ matrix. Then $C(BA) = (CB)A$.

- **Proof** Since $L_A : \mathbf{R}^n \to \mathbf{R}^m$, $L_B : \mathbf{R}^m \to \mathbf{R}^l$, and $L_C : \mathbf{R}^l \to \mathbf{R}^k$ are linear transformations, we have from the previous Lemma that

$$L_C(L_B L_A) = (L_C L_B) L_A.$$

Let $\alpha, \beta, \gamma, \delta$ be the standard bases of $\mathbf{R}^n$, $\mathbf{R}^m$, $\mathbf{R}^l$, and $\mathbf{R}^k$ respectively. Then we have

$$[L_C(L_B L_A)]_\alpha^\delta = [L_C]_\gamma^\delta [L_B L_A]_\alpha^\gamma = [L_C]_\gamma^\delta ([L_B]_\beta^\gamma [L_A]_\alpha^\beta) = C(BA)$$

while

$$[(L_C L_B) L_A]_\alpha^\delta = [L_C L_B]_\beta^\delta [L_A]_\alpha^\beta = ([L_C]_\gamma^\delta [L_B]_\beta^\gamma)[L_A]_\alpha^\beta = (CB)A$$

using Lemma 2. Combining these three identities we see that $C(BA) = (CB)A$. $\square$

- The above proof may seem rather weird, but it managed to prove the matrix identity $C(BA) = (CB)A$ without having to do lots and lots of matrix multiplication. Exercise: try proving $C(BA) = (CB)A$ directly by writing out $C$, $B$, $A$ in co-ordinates and expanding both sides!

- We have just shown that matrix multiplication is associative. In fact, all the familiar rules of algebra apply to matrices (e.g. $A(B + C) = AB + AC$, and $A$ times the identity is equal to $A$) provided that all the matrix operations make sense, of course. (The shapes of the matrices have to be compatible before one can even begin to add or multiply them together). The one important caveat is that matrix multiplication is not commutative: $AB$ is usually not the same as $BA$! Indeed there is no guarantee that these two matrices are the same shape (or even that they are both defined at all).

- Some other properties of $A$ and $L_A$ are stated below. As you can see, the proofs are similar to the ones above.

- If $A$ is an $m \times n$ matrix and $B$ is an $l \times m$ matrix, then $L_{BA} = L_B L_A$. Proof: Let $\alpha, \beta, \gamma$ be the standard bases of $\mathbf{R}^n$, $\mathbf{R}^m$, $\mathbf{R}^l$ respectively. Then $L_B L_A$ is a linear transformation from $\mathbf{R}^n$ to $\mathbf{R}^l$, and so

$$[L_B L_A]_\alpha^\gamma = [L_B]_\beta^\gamma [L_A]_\alpha^\beta = BA,$$

  and so by taking $L$ of both sides and using Lemma 2, we obtain $L_B L_A = L_{BA}$ as desired.

- If $A$ is an $m \times n$ matrix, and $B$ is another $m \times n$ matrix, then $L_{A+B} = L_A + L_B$. Proof: $L_A + L_B$ is a linear transformation from $\mathbf{R}^n$ to $\mathbf{R}^m$. Let $\alpha$, $\beta$ be the standard bases of $\mathbf{R}^n$ and $\mathbf{R}^m$ respectiely. Then

$$[L_A + L_B]_\alpha^\beta = [L_A]_\alpha^\beta + [L_B]_\alpha^\beta = A + B$$

  and so by taking $L$ of both sides and using Lemma 2, we obtain $L_{A+B} = L_A + L_B$ as desired.

* * * * *

Invertible linear transformations

- We have already dealt with the concepts of a linear transformation being one-to-one, and of being onto. We now combine these two concepts to that of a transformation being *invertible*.

- **Definition.** Let $T : V \to W$ be a linear transformation. We say that a linear transformation $S : W \to V$ is the *inverse* of $T$ if $TS = I_W$ and $ST = I_V$. We say that $T$ is *invertible* if it has an inverse, and call the inverse $T^{-1}$; thus $TT^{-1} = I_W$ and $T^{-1}T = I_V$.

- **Example** Let $T : \mathbf{R}^3 \to \mathbf{R}^3$ be the doubling transformation $Tv := 2v$. Let $S : \mathbf{R}^3 \to R^3$ be the halving transformation $Sv := v/2$. Then $S$ is the inverse of $T$: $ST(v) = S(2v) = (2v)/2 = v$, while $TS(v) = T(v/2) = 2(v/2) = v$, thus both $ST$ and $TS$ are the identity on $\mathbf{R}^3$.

- Note that this definition is symmetric: if $S$ is the inverse of $T$, then $T$ is the inverse of $S$.

- Why do we call $S$ *the* inverse of $T$ instead of just *an* inverse? This is because every transformation can have at most one inverse:

- **Lemma 6**. Let $T : V \to W$ be a linear transformation, and let $S : W \to V$ and $S' : W \to V$ both be inverses of $T$. Then $S = S'$.

- **Proof**
$$S = SI_W = S(TS') = (ST)S' = I_V S' = S'.$$
$\square$

- Not every linear transformation has an inverse:

- **Lemma 7**. If $T : V \to W$ has an inverse $S : W \to V$, then $T$ must be one-to-one and onto.

- **Proof** Let's show that $T$ is one-to-one. Suppose that $Tv = Tv'$; we have to show that $v = v'$. But by applying $S$ to both sides we get $STv = STv'$, thus $I_V v = I_V v'$, thus $v = v'$ as desired. Now let's show that $T$ is onto. Let $w \in W$; we have to find $v$ such that $Tv = w$. But $w = I_W w = TSw = T(Sw)$, so if we let $v := Sw$ then we have $Tv = w$ as desired. $\square$.

- Thus, for instance, the zero transformation $T : \mathbf{R}^3 \to \mathbf{R}^3$ defined by $Tv = 0$ is not invertible.

- The converse of Lemma 7 is also true:

- **Lemma 8**. If $T : V \to W$ is a one-to-one and onto linear transformation , then it has an inverse $S : W \to V$, which is also a linear transformation.

- **Proof** Let $T : V \to W$ be one-to-one and onto. Let $w$ be any element of $W$. Since $T$ is onto, we have $w = Tv$ for some $v$ in $V$; since $T$ is one-to-one; this $v$ is unique (we can't have two different elements $v$, $v'$ of $V$ such that $Tv$ and $Tv'$ are both equal to $w$). Let us define $Sw$ as equal to this $v$, thus $S$ is a transformation from $W$ to $V$. For any $w \in W$, we have $w = Tv$ and $Sw = v$ for some $v \in V$, and hence $TSw = w$; thus $TS$ is the identity $I_W$.

- Now we show that $ST = I_V$, i.e. that for every $v \in V$, we have $STv = v$. Since we already know that $TS = I_W$, we have that $TSw = w$ for all $w \in W$. In particular we have $TSTv = Tv$, since $Tv \in W$. But since $T$ is invertible, this implies that $STv = v$ as desired.

- Finally, we show that $S$ is linear, i.e. that it preserves addition and scalar multiplication. We'll just show that it preserves addition, and leave scalar multiplication as an exercise. Let $w, w' \in W$; we need to show that $S(w + w') = Sw + Sw'$. But we have

$$T(S(w+w')) = (TS)(w+w') = I_W(w+w') = I_W w + I_W w' = TSw + TSw' = T(Sw+Sw');$$

since $T$ is one-to-one, this implies that $S(w+w') = Sw+Sw'$ as desired. The preservation of scalar multiplication is proven similarly. $\square$

- Thus a linear transformation is invertible if and only if it is one-to-one and onto. Invertible linear transformations are also known as *isomorphisms*.

- **Definition** Two vector spaces $V$ and $W$ are said to be *isomorphic* if there is an invertible linear transformation $T : V \to W$ from one space to another.

- **Example** The map $T : \mathbf{R}^3 \to P_2(\mathbf{R})$ defined by

$$T(a, b, c) := ax^2 + bx + c$$

is easily seen to be linear, one-to-one, and onto, and hence an isomorphism. Thus $\mathbf{R}^3$ and $P_2(\mathbf{R})$ are isomorphic.

- Isomorphic spaces tend to have almost identical properties. Here is an example:

- **Lemma 9**. Two finite-dimensional spaces $V$ and $W$ are isomorphic if and only if $\dim(V) = \dim(W)$.

- **Proof** If $V$ and $W$ are isomorphic, then there is an invertible linear transformation $T : V \to W$ from $V$ to $W$, which by Lemma 7 is one-to-one and onto. Since $T$ is one-to-one, $\text{nullity}(T) = 0$. Since $T$ is onto, $\text{rank}(T) = \dim(W)$. By the dimension theorem we thus have $\dim(V) = \dim(W)$.

- Now suppose that $\dim(V)$ and $\dim(W)$ are equal; let's say that $\dim(V) = \dim(W) = n$. Then $V$ has a basis $\{v_1, \ldots, v_n\}$, and $W$ has a basis

$\{w_1, \ldots, w_n\}$. By Theorem 6 of last week's notes, we can find a linear transformation $T : V \to W$ such that $Tv_1 = w_1, \ldots, Tv_n = w_n$. By Theorem 3 of last week's notes, $w_1, \ldots, w_n$ must then span $R(T)$. But since $w_1, \ldots, w_n$ span $W$, we have $R(T) = W$, i.e. $T$ is onto. By Lemma 2 of last week's notes, $T$ is therefore one-to-one, and hence is an isomorphism. Thus $V$ and $W$ are isomorphic. $\square$

- Every basis leads to an isomorphism. If $V$ has a finite basis $\beta = (v_1, \ldots, v_n)$, then the co-ordinate map $\phi_\beta : V \to \mathbf{R}^n$ defined by

$$\phi_\beta(x) := [x]^\beta$$

is a linear transformation (see last week's homework), and is invertible (this was discussed in last week's notes, where we noted that we can reconstruct $x$ from $[x]^\beta$ and vice versa). Thus $\phi_\beta$ is an isomorphism between $V$ to $\mathbf{R}^n$. In the textbook $\phi_\beta$ is called the *standard representation of $V$ with respect to $\beta$*.

- Because of all this theory, we are able to essentially equate finite-dimensional vector spaces $V$ with the standard vector spaces $\mathbf{R}^n$, to equate vectors $v \in V$ with their co-ordinate vectors $[v]^\alpha \in \mathbf{R}^n$ (provided we choose a basis $\alpha$ for $V$) and linear transformations $T : V \to W$ from one finite-dimensional space to another, with $n \times m$ matrices $[T]_\alpha^\beta$. This means that, for finite-dimensional linear algebra at least, we can reduce everything to the study of column vectors and matrices. This is what we will be doing for the rest of this course.

* * * * *

Invertible linear transformations and invertible matrices

- An $m \times n$ matrix $A$ has an *inverse* $B$, if $B$ is an $n \times m$ matrix such that $BA = I_n$ and $AB = I_m$. In this case we call $A$ an *invertible matrix*, and denote $B$ by $A^{-1}$.

- **Example.** If

$$A = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 4 \end{pmatrix}$$

then

$$A^{-1} = \begin{pmatrix} 1/2 & 0 & 0 \\ 0 & 1/3 & 0 \\ 0 & 0 & 1/4 \end{pmatrix}$$

is the inverse of $A$, as can be easily checked.

- The relationship between invertible linear transformations and invertible matrices is the following:

- **Theorem 10.** Let $V$ be a vector space with finite ordered basis $\alpha$, and let $W$ be a vector space with finite ordered basis $\beta$. Then a linear transformation $T : V \to W$ is invertible if and only if the matrix $[T]_\alpha^\beta$ is invertible. Furthermore, $([T]_\alpha^\beta)^{-1} = [T^{-1}]_\beta^\alpha$

- **Proof.** Suppose that $V$ is $n$-dimensional and $W$ is $m$-dimensional; this makes $[T]_\alpha^\beta$ an $m \times n$ matrix.

- First suppose that $T : V \to W$ has an inverse $T^{-1} : W \to V$. Then

$$[T]_\alpha^\beta [T^{-1}]_\beta^\alpha = [TT^{-1}]_\beta^\beta = [I_W]_\beta^\beta = I_m$$

while

$$[T^{-1}]_\beta^\alpha [T]_\alpha^\beta = [T^{-1}T]_\alpha^\alpha = [I_V]_\alpha^\alpha = I_n,$$

thus $[T^{-1}]_\beta^\alpha$ is the inverse of $[T]_\alpha^\beta$ and so $[T]_\alpha^\beta$ is invertible.

- Now suppose that $[T]_\alpha^\beta$ is invertible, with inverse $B$. We'll prove shortly that there exists a linear transformation $S : W \to V$ with $[S]_\beta^\alpha = B$. Assuming this for the moment, we have

$$[ST]_\alpha^\alpha = [S]_\beta^\alpha [T]_\alpha^\beta = B[T]_\alpha^\beta = I_n = [I_V]_\alpha^\alpha$$

and hence $ST = I_V$. A similar argument gives $TS = I_W$, and so $S$ is the inverse of $T$ and so $T$ is invertible.

- It remains to show that we can in fact find a transformation $S : W \to V$ with $[S]_\beta^\alpha = B$. Write $\alpha = (v_1, \ldots, v_n)$ and $\beta = (w_1, \ldots, w_m)$. Then we want a linear transformation $S : W \to V$ such that

$$Sw_1 = B_{11}v_1 + \ldots + B_{1n}v_n$$

$$Sw_2 = B_{21}v_1 + \ldots + B_{2n}v_n$$

$$\vdots$$

$$Sw_m = B_{m1}v_1 + \ldots + B_{mn}v_n.$$

But we can do this thanks to Theorem 6 of last week's notes. $\square$

- **Corollary 11**. An $m \times n$ matrix $A$ is invertible if and only if the linear transformation $L_A : \mathbf{R}^n \to \mathbf{R}^m$ is invertible. Furthermore, the inverse of $L_A$ is $L_{A^{-1}}$.

- **Proof.** If $\alpha$ is the standard basis for $\mathbf{R}^n$ and $\beta$ is the standard basis for $\mathbf{R}^m$, then
$$[L_A]_\alpha^\beta = A.$$
Thus by Theorem 10, $A$ is invertible if and only if $L_A$ is. Also, from Theorem 10 we have
$$[L_A^{-1}]_\beta^\alpha = ([L_A]_\beta^\alpha)^{-1} = A^{-1} = [L_{A^{-1}}]_\beta^\alpha$$
and hence
$$L_A^{-1} = L_{A^{-1}}$$
as desired. $\square$

- **Corollary 12.** In order for a matrix $A$ to be invertible, it must be square (i.e. $m = n$).

- **Proof**. This follows immediately from Corollary 11 and Lemma 9. $\square$

- On the other hand, not all square matrices are invertible; for instance the zero matrix clearly does not have an inverse. More on this in a later week.