1 Last time: one-to-one and onto linear transformations

Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a function.

The following mean the same thing:

- T is *linear* is the sense that $T(u + v) = T(u) + T(v)$ and $T(cv) = cT(v)$ for $u, v \in \mathbb{R}^n$, $c \in \mathbb{R}$.
- There is an $m \times n$ matrix A such that T has the formula $T(v) = Av$ for $v \in \mathbb{R}^n$.

If we are given a linear transformation T, then $T(v) = Av$ for the matrix

$$A = \left[\begin{array}{cccc} T(e_1) & T(e_2) & \dots & T(e_n) \end{array}\right]$$

where $e_i \in \mathbb{R}^n$ is the vector with a 1 in row i and 0 in all other rows.

Call A the *standard matrix* of T.

The following all mean the same thing for a function $f : X \to Y$.

- f is *one-to-one*.
- If $a, b \in X$ and $f(a) = f(b)$ then $a = b$.
- If $a, b \in X$ and $a \neq b$ then $f(a) \neq f(b)$.
- f does not send different inputs to the same output.

Similarly, the following all mean the same thing for a function $f : X \to Y$.

- f is *onto*.
- The range of f is equal to the codomain, i.e., $\mathrm{range}(f) = \{f(a) : a \in X\} = Y$.
- For each $y \in Y$ there is at least one $x \in X$ with $f(x) = y$.
- Every element of the codomain of f is an output for some input.

We can detect whether a linear transformation is one-to-one or onto by locating the pivot positions in its standard matrix (by row reducing).

Theorem. Suppose $T : \mathbb{R}^n \to \mathbb{R}^m$ is the linear transformation $T(v) = Av$ where A is an $m \times n$ matrix.

(1) T is one-to-one if and only if the columns of A are linearly independent, which happens precisely when A has a pivot position in every column.

(2) T is onto if and only if the span of the columns of A is $\mathbb{R}^m$, which happens precisely when A has a pivot position in every row.

2 Operators on linear transformations and matrices

Key point from last time and starting point of today: linear transformations $\mathbb{R}^n \to \mathbb{R}^m$ are uniquely represented by $m \times n$ matrices, and every $m \times n$ matrix corresponds to a linear transformation $\mathbb{R}^n \to \mathbb{R}^m$.

There are several simple, natural operations we can use to combine and alter linear transformations to get other linear transformations. The goal is to translate these function operations into matrix operations.

Sums and scalar multiples. Suppose $T : \mathbb{R}^n \to \mathbb{R}^m$ and $U : \mathbb{R}^n \to \mathbb{R}^m$ are two linear functions with the same domain and codomain. Their sum $T + U$ is the function $\mathbb{R}^n \to \mathbb{R}^m$ defined by

$$(T + U)(v) = T(v) + U(v) \qquad \text{for } v \in \mathbb{R}^n.$$

If $c \in \mathbb{R}$ is a scalar, then cT is the function $\mathbb{R}^n \to \mathbb{R}^m$ defined by

$$(cT)(v) = cT(v) \qquad \text{for } v \in \mathbb{R}^n.$$

Fact. Both $T + U$ and cT are linear transformations.

Proof. To see that $T + U$ is linear, we check that

$$(T + U)(u + v) = T(u + v) + U(u + v) = T(u) + T(v) + U(u) + U(v) = (T + U)(u) + (T + U)(v)$$

for $u, v \in \mathbb{R}^n$, and

$$(T + U)(av) = T(av) + U(av) = aT(v) + aU(v) = a(T + U)(v)$$

for $a \in \mathbb{R}$ and $v \in \mathbb{R}^n$. Since these properties hold, $T + U$ is linear.

The proof that cT is linear is similar. (Try this yourself!) □

Since sums and scalar multiples of linear functions are linear, it follows that differences $T - U$ and arbitrary linear combinations $aT + bU + cV + \dots$ of linear functions are linear.

Suppose T and U have standard matrices

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & & a_{2n} \\ \vdots & & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \qquad \text{and} \qquad B = \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & & b_{2n} \\ \vdots & & \ddots & \vdots \\ b_{m1} & b_{m2} & \dots & b_{mn} \end{bmatrix}$$

so that $T(v) = Av$ and $U(v) = Bv$.

Proposition. The standard matrix of $T + U$ is

$$A + B = \begin{bmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \dots & a_{1n} + b_{1n} \\ a_{21} + b_{21} & a_{22} + b_{22} & & a_{2n} + b_{2n} \\ \vdots & & \ddots & \vdots \\ a_{m1} + b_{m1} & a_{m2} + b_{m2} & \dots & a_{mn} + b_{mn} \end{bmatrix}.$$

The standard matrix of cT is

$$cA = \begin{bmatrix} ca_{11} & ca_{12} & \dots & ca_{1n} \\ ca_{21} & ca_{22} & & ca_{2n} \\ \vdots & & \ddots & \vdots \\ ca_{m1} & ca_{m2} & \dots & ca_{mn} \end{bmatrix}.$$

This is how we *define* sums and scalar multiples of matrices. These operations work in essentially the same way as for vectors: we can add matrices of the same size, by adding the entries in corresponding positions together, and we can multiply a matrix by a scalar c by multiplying all entries by c.

Example. We have

$$\begin{bmatrix} 4 & 0 & 5 \\ -1 & 3 & 2 \end{bmatrix} + \begin{bmatrix} 1 & 1 & 1 \\ 3 & 5 & 7 \end{bmatrix} = \begin{bmatrix} 5 & 1 & 6 \\ 2 & 8 & 9 \end{bmatrix}.$$

and

$$-\begin{bmatrix} 4 & 0 & 5 \\ -1 & 3 & 2 \end{bmatrix} + 2\begin{bmatrix} 1 & 1 & 1 \\ 3 & 5 & 7 \end{bmatrix} = \begin{bmatrix} -4 & 0 & -5 \\ 1 & -3 & -2 \end{bmatrix} + \begin{bmatrix} 2 & 2 & 2 \\ 6 & 10 & 14 \end{bmatrix} = \begin{bmatrix} -2 & 2 & -3 \\ 7 & 7 & 12 \end{bmatrix}.$$

Suppose T, U, V are linear transformations $\mathbb{R}^n \to \mathbb{R}^m$ with standard matrices A, B, C. Let $a, b \in \mathbb{R}$.

The following properties then hold:

Functions	Matrices
1. $T + U = U + V$	$A + B = B + A$.
2. $(T + U) + V = T + (U + V)$	$(A + B) + C = A + (B + C)$.
3. $T + 0 = T$ where $0 : \mathbb{R}^n \to \mathbb{R}^m$ is the map $0(v) = 0 \in \mathbb{R}^m$.	$A + 0 = A$.
4. $a(T + U) = aT + aU$	$a(A + B) = aA + aB$.
5. $(a + b)T = aT + bT$	$(a + b)A = aA + bA$.
6. $a(bT) = (ab)T$.	$a(bA) = (ab)A$.

Composition. Suppose $U : \mathbb{R}^n \to \mathbb{R}^m$ and $T : \mathbb{R}^m \to \mathbb{R}^k$ are linear.

Note that we assume the codomain of U is equal to the domain of T.

The *composition* $T \circ U$ is the function $\mathbb{R}^n \to \mathbb{R}^k$ given by

$$(T \circ U)(v) = T(U(v)) \qquad \text{for } v \in \mathbb{R}^n.$$

Fact. Since T and U are linear, $T \circ U$ is linear.

Proof. To see that $T \circ U$ is linear, we check that

$$(T \circ U)(u + v) = T(U(u + v)) = T(U(u) + U(v)) = T(U(u)) + T(U(v)) = (T \circ U)(u) + (T \circ U)(v)$$

for $u, v \in \mathbb{R}^n$, and

$$(T \circ U)(cv) = T(U(cv)) = T(cU(v)) = cT(U(v)) = c(T \circ U)(v)$$

for $c \in \mathbb{R}$ and $v \in \mathbb{R}^n$. □

Important note: $U \circ T$ is not defined unless $k = n$.

Even if $k = n$ so that both $T \circ U$ and $U \circ T$ are defined, there is no reason to expect that $T \circ U = U \circ T$.

Example. If $n = m = k = 1$ and $T(x) = 2x$ and $U(x) = x^2$, then

$$(T \circ U)(x) = T(x^2) = 2x^2 \qquad \text{but} \qquad (U \circ T)(x) = U(2x) = 4x^2.$$

Since $T \circ U$ is a linear transformation $\mathbb{R}^n \to \mathbb{R}^k$, there is a unique $k \times n$ matrix C such that

$$(T \circ U)(v) = Cv \qquad \text{for } v \in \mathbb{R}^n.$$

If A is the standard matricx of T and B is the standard matrix of U, then we define the *matrix product*

$$AB = C.$$

Note how this definition works: if A is $k \times m$ and B is $m \times n$ then we *define* AB to be the unique $k \times n$ matrix C such that $Cv = A(Bv)$ for all $v \in \mathbb{R}^n$.

How do we actually compute the rectangular array which is AB, from A and B?

Theorem. Suppose B has columns $b_1, b_2, \dots, b_n \in \mathbb{R}^m$ so that $B = \begin{bmatrix} b_1 & b_2 & \dots & b_n \end{bmatrix}$.

Then $AB = \begin{bmatrix} Ab_1 & Ab_2 & \dots & Ab_n \end{bmatrix}$. (This makes sense as A is $k \times m$.)

Proof. AB is the standard matrix of the linear function $T \circ U$, so

$$\begin{aligned} AB &= \left[\begin{array}{cccc} (T\circ U)(e_1) & (T\circ U)(e_2) & \cdots & (T\circ U)(e_n) \end{array}\right] = \left[\begin{array}{cccc} A(Be_1) & A(Be_2) & \cdots & A(Be_n) \end{array}\right] \\ &= \left[\begin{array}{cccc} Ab_1 & Ab_2 & \cdots & Ab_n \end{array}\right]. \end{aligned}$$

$\square$

Example. If $A = \begin{bmatrix} 2 & 3 \\ 1 & -5 \end{bmatrix}$ and $B = \begin{bmatrix} 4 & 3 & 6 \\ 1 & -2 & 3 \end{bmatrix}$, then $b_1 = \begin{bmatrix} 4 \\ 1 \end{bmatrix}$, $b_2 = \begin{bmatrix} 3 \\ -2 \end{bmatrix}$, $b_3 = \begin{bmatrix} 6 \\ 3 \end{bmatrix}$, so

$$AB = \left[\begin{array}{ccc} Ab_1 & Ab_2 & Ab_3 \end{array}\right] = \begin{bmatrix} 11 & 0 & 21 \\ -1 & 13 & -9 \end{bmatrix}.$$

The quick rule for computing AB: if the ith row of A and jth column of B are

$$\left[\begin{array}{cccc} a_1 & a_2 & \dots & a_m \end{array}\right] \qquad \text{and} \qquad \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

then the entry in the ith row and jth column of AB is

$$\left[\begin{array}{cccc} a_1 & a_2 & \dots & a_m \end{array}\right] \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} = a_1b_1 + a_2b_2 + \cdots + a_mb_m.$$

Example. Suppose $A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \\ 9 & 9 & 9 \end{bmatrix}$.

The entry in the 2nd row and 2nd column of AB is

$$5 \cdot 2 + 6 \cdot 5 + 7 \cdot 8 + 8 \cdot 9 = 10 + 30 + 56 + 72 = 168.$$

Write I_n for the $n \times n$ matrix

$$I_n = \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{bmatrix}$$

which has 1 in each diagonal position, and zeros in all other positions.

The matrix I_n is the standard matrix of the *identity map* $\mathbb{R}^n \to \mathbb{R}^n$.

This is the linear function T with $T(v) = v$ for all $v \in \mathbb{R}^n$.

Proposition. Let A, B, C be matrices.

Assume A is $m \times n$, B is $n \times l$, and C is $l \times k$.

Then $A(BC) = (AB)C$.

Proof. Use our first definition of matrix multiplication.

By this definition, AB and BC are the unique matrices such that $(AB)x = A(Bx)$ and $(BC)x = B(Cx)$.

In turn, $A(BC)$ is the unique matrix such that $(A(BC))x = A((BC)x) = A(B(Cx))$.

But $(AB)C$ is also the unique matrix such that $((AB)C)x = (AB)(Cx) = A(B(Cx))$.

Therefore $A(BC) = (AB)C$. □

Here are some easier properties. Suppose A, B, C are matrices and $r \in \mathbb{R}$.

- If A is $m \times n$ and B, C are $n \times l$ then $A(B + C) = AB + AC$.
- If A, B are $m \times n$ and C is $n \times l$ then $(A + B)C = AC + BC$.
- If A is $m \times n$ and B is $n \times l$ then $r(AB) = (rA)B = A(rB)$.
- If A is $m \times n$ then $I_m A = AI_n = A$.

3 Pathologies of matrix multiplication

Suppose A and B are matrices.

Four important observations:

1. The product AB is defined only if the number of columns of A is the number of rows of B.
2. Even if AB and BA are both defined, it typically happens that $AB \neq BA$.
3. $AB = AC$ does not imply $B = C$.
4. It can happen that $AB = 0 = \begin{bmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{bmatrix}$ even if both $A \neq 0$ and $B \neq 0$.

Example. We have

$$\begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix} = \left[\begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 0 \\ 1 \end{bmatrix} \right] = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

while

$$\begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} = \left[\begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 0 \\ 0 \end{bmatrix} \right] = \begin{bmatrix} 0 & 0 \\ 2 & 0 \end{bmatrix}.$$

If A and B are both *square* matrices of the same size (meaning they have the same number of rows and columns), and $AB = BA$, then we say that A and B commutes.

4 Matrix transpose

The *transpose* of an $m \times n$ matrix A is the $n \times m$ matrix A^T whose columns are the rows of A.

If a_{ij} is the entry in row i and column j of A, then this is the entry in row j and column i of A^T.

For example, if $A = \begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix}$ then $A^T = \begin{bmatrix} a & d \\ b & e \\ c & f \end{bmatrix}$.

The transpose of A is given by flipping A across the main diagonal, in order to interchange rows/columns.

Another example: if $C = \begin{bmatrix} 1 & 1 & 1 & 1 \\ -3 & 5 & -2 & 7 \\ 0 & 0 & 1 & 0 \end{bmatrix}$ then $C^T = \begin{bmatrix} 1 & -3 & 0 \\ 1 & 5 & 0 \\ 1 & -2 & 1 \\ 1 & 7 & 0 \end{bmatrix}$.

We finish this lecture by noting some basic properties of the transpose operation:

- $(A^T)^T = A$ since flipping twice does nothing.
- If A and B have the same size then $(A+B)^T = A^T + B^T$.
- If $c \in \mathbb{R}$ then $(cA)^T = c(A^T)$.
- If A is an $k \times m$ matrix and B is and $m \times n$ matrix then $(AB)^T = B^T A^T$.

To prove the last property, use our earlier results to compute the entries in ith row and jth column of the matrices on either side (in terms of the entries of A and B), and check that these are equal.

Question for later: how is the linear transformation with standard matrix A related to the linear transformation with standard matrix A^T? What is the transpose as an operation on linear transformations?

5 Vocabulary

Keywords from today's lecture:

1. **Sums, scalar multiples, and compositions** of linear functions.

   If $T : \mathbb{R}^n \to \mathbb{R}^m$ and $U : \mathbb{R}^n \to \mathbb{R}^m$ and $c \in \mathbb{R}$ then

   $$T + U : \mathbb{R}^n \to \mathbb{R}^m$$

   is the function with $(T + U)(v) = T(v) + U(v)$, and

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

   is the function with $(cT)(v) = c(T(v))$.

   If $T : \mathbb{R}^n \to \mathbb{R}^m$ and $U : \mathbb{R}^m \to \mathbb{R}^k$ then $U \circ T : \mathbb{R}^n \to \mathbb{R}^k$ is the function $(U \circ T)(v) = U(T(v))$.

2. **Sums, scalar multiples, and products** of matrices.

   If A and B are $m \times n$ matrices then $A + B$ is the $m \times n$ matrix whose entry in position (i, j) is $A_{ij} + B_{ij}$. If $c \in \mathbb{R}$ then cA is the matrix whose entry in position (i, j) is cA_{ij}.

   If A is $m \times n$ and B is $n \times k$ then AB is the $m \times k$ whose entry in position (i, j) is ith row of A (which is a $1 \times n$ matrix) times the jth column of B (which is a vector in $\mathbb{R}^n$).

   Example: $\begin{bmatrix} a & b \\ c & d \end{bmatrix} + \begin{bmatrix} w & x \\ y & z \end{bmatrix} = \begin{bmatrix} a+w & b+x \\ c+y & d+z \end{bmatrix}$.

   Example: $5 \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 5a & 5b \\ 5c & 5d \end{bmatrix}$.

   Example: $\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} w & x \\ y & z \end{bmatrix} = \begin{bmatrix} aw+by & ax+bz \\ cw+dy & cx+dz \end{bmatrix}$.

3. **Transpose** of a matrix.

   If A is an $m \times n$ matrix then its transpose A^T is the $n \times m$ with the same entries as A but with rows and columns interchanged.

   Example: $\begin{bmatrix} a & b & c \\ x & y & z \end{bmatrix}^T = \begin{bmatrix} a & x \\ b & y \\ c & z \end{bmatrix}$.

4. **Identity** matrix

   The $n \times n$ matrix $I = I_n$ with 1s on the diagonal and 0s off the diagonal.

   $AI = A$ and $IB = B$ for all matrices A with n columns and all matrices B with n rows.

   Example: $\begin{bmatrix} 1 \end{bmatrix}$, $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, and so on.