Summary

Quick summary of today's notes. Lecture starts on next page.

- A function $f : X \to Y$ is *invertible* if it is both one-to-one and onto.
- The *identity function* on a set X is the function $\mathrm{id}_X : X \to X$ with $\mathrm{id}_X(x) = x$ for all $x \in X$.

  If $f : X \to Y$ is a function then $f \circ \mathrm{id}_X = \mathrm{id}_Y \circ f = f$.
- A function $f : X \to Y$ is invertible if and only if there is a function $g : Y \to X$ such that $f \circ g = \mathrm{id}_Y$ and $g \circ f = \mathrm{id}_X$. If such a function g exists, then it is unique and we call it the *inverse* of f.

  If we know that $f : X \to Y$ is invertible, then we write $f^{-1} : Y \to X$ to denote its inverse.
- Let m and n be positive integers.

  The identity function on the set $\mathbb{R}^n$ is linear. Its standard matrix is the $n \times n$ *identity matrix*

$$I_n = \left[\begin{array}{cccc} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{array}\right].$$

- Suppose $T : \mathbb{R}^n \to \mathbb{R}^m$ is linear. Then T is invertible only if $m = n$.

  Suppose $T : \mathbb{R}^n \to \mathbb{R}^n$ is invertible. Then its inverse $T^{-1} : \mathbb{R}^n \to \mathbb{R}^n$ is also linear.
- A matrix is *invertible* if it is the standard matrix of an invertible linear transformation.

  A matrix can be invertible only if it is square, i.e., has the same number of rows and columns.
- Suppose A is invertible. Then A is the standard matrix of an invertible linear transformation T.

  We define the *inverse* A^{-1} of A to be the standard matrix of the linear transformation T^{-1}.

  If A is $n \times n$ then A^{-1} is also $n \times n$, and $AA^{-1} = A^{-1}A = I_n$.
- A square matrix A is invertible if and only if $\mathsf{RREF}(A) = I_n$ for some n.
- We can check if an $n \times n$ matrix A is invertible and try to compute A^{-1} at the same time.

  We do this as follows. First form the $n \times 2n$ matrix $\left[\begin{array}{cc} A & I_n \end{array}\right]$.

  Then compute the reduced echelon form of this matrix.

  If $\mathsf{RREF}\left(\left[\begin{array}{cc} A & I_n \end{array}\right]\right) = \left[\begin{array}{cc} I_n & B \end{array}\right]$ for some $n \times n$ matrix B, then $A^{-1} = B$.

  Otherwise, A is not invertible.
- A 2×2 matrix $A = \left[\begin{array}{cc} a & b \\ c & d \end{array}\right]$ is invertible if and only if $ad - bc \neq 0$, in which case

$$A^{-1} = \frac{1}{ad - bc}\left[\begin{array}{cc} d & -b \\ -c & a \end{array}\right].$$

1 Last time: adding and multiplying matrices

Suppose $T : \mathbb{R}^r \to \mathbb{R}^s$ and $U : \mathbb{R}^n \to \mathbb{R}^m$ are linear transformations with standard matrices A and B.

Let $c \in \mathbb{R}$ be a scalar.

1. The scalar multiple $cT : \mathbb{R}^r \to \mathbb{R}^s$ of T is the linear transformation with $(cT)(x) = cT(x)$.

   The standard matrix of cT is cA where cA is given by multiplying every entry in A by c, i.e.:

   Definition. cA is the matrix $\begin{bmatrix} ca_1 & ca_2 & \dots & ca_r \end{bmatrix}$ where $A = \begin{bmatrix} a_1 & a_2 & \dots & a_r \end{bmatrix}$.

2. If T and U have the same domain and codomain, meaning that $r = n$ and $s = m$ and that A and B have the same size, then we can form the sum $T + U : \mathbb{R}^n \to \mathbb{R}^m$ as the linear transformation with $(T + U)(x) = T(x) + U(x)$.

   The standard matrix of $T + U$ is $A + B$ where:

   Definition. $A + B$ is the matrix whose ith column is the sum of the ith columns of A and B.

3. If the domain of T is the codomain of U, meaning that $r = m$, then we can form the composition $T \circ U : \mathbb{R}^n \to \mathbb{R}^s$ as the linear transformation with $(T \circ U)(x) = T(U(x))$.

   The standard matrix of $T \circ U$ is the product AB where:

   Definition. The product AB of matrices A and B, where the number of columns of A is the number of rows of B, is the matrix $AB = \begin{bmatrix} Ab_1 & Ab_2 & \dots & Ab_n \end{bmatrix}$ where $B = \begin{bmatrix} b_1 & b_2 & \dots & b_n \end{bmatrix}$.

Some remarks.

$A + B$ is only defined if A and B are matrices of the same size.

We have $A + B = B + A$ when either side is defined.

AB is only defined if the number of columns of A is equal to the number of rows of B.

When defined, AB has the same number of rows as A, and the same number of columns as B.

Even if AB and BA are both defined, we may still have $AB \neq BA$.

Example. Let $A = \begin{bmatrix} a & b & c & d \\ e & f & g & h \\ i & j & k & l \end{bmatrix}$ be a 3×4 matrix.

Consider what happens when we multiply A on the left by various 3×3 matrices.

1. $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} A = \begin{bmatrix} a & b & c & d \\ i & j & k & l \\ e & f & g & h \end{bmatrix}$. Multiplication swaps rows 2 and 3 of A.
2. $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{bmatrix} A = \begin{bmatrix} a & b & c & d \\ 3e & 3f & 3g & 3h \\ i & j & k & l \end{bmatrix}$. Multiplication rescales row 2 by factor 3.
3. $\begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} A = \begin{bmatrix} a + 2i & b + 2j & c + 2k & d + 2l \\ e & f & g & h \\ i & j & k & l \end{bmatrix}$. This adds 2 times row 3 to row 1.

Moral: row operations on A correspond to multiplying A on the left by certain square matrices.

2 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.

If $A = \begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix}$ and $A^T = \begin{bmatrix} a & d \\ b & e \\ c & f \end{bmatrix}$. If $C = \begin{bmatrix} 1 & 1 & 1 & 1 \\ -3 & 5 & -2 & 7 \\ 0 & 0 & 1 & 0 \end{bmatrix}$ then $C^T = \begin{bmatrix} 1 & -3 & 0 \\ 1 & 5 & 0 \\ 1 & -2 & 1 \\ 1 & 7 & 0 \end{bmatrix}$.

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

Basic properties of the transpose operation:

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

Since matrices represent linear transformation, operations on matrices correspond to operations on linear transformations. For example, matrix multiplication corresponds to composition of linear functions.

It is reasonable to ask what operation the transpose corresponds to on linear transformations.

Given vectors $u = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}$ and $v = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$ in $\mathbb{R}^n$, define $(u, v) = u^T v = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n \in \mathbb{R}$.

This is often called the *dot product* of u and v. Note that $(u, v \pm w) = (u, v) \pm (u, w)$.

Important fact in the case when $u = v$: If $v \neq 0$ then $(v, v) > 0$. If $(v, v) = 0$ then $v = 0$.

Proof. If any entry $v_i \neq 0$ then $(v, v) = v_1^2 + v_2^2 + \cdots + v_n^2 \geq v_i^2 > 0$. □

Here is the answer to the question of what the transpose means in terms of linear transformations:

Proposition. Suppose $L : \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation. There exists a unique linear transformation $L^T : \mathbb{R}^m \to \mathbb{R}^n$, called the *transpose* of L, such that $(L(u), v) = (u, L^T(v))$ for all $u \in \mathbb{R}^n$ and $v \in \mathbb{R}^m$. The standard matrix of L^T is the transpose of the standard matrix of L.

Proof. Let A be the standard matrix of L. If define $L^T : \mathbb{R}^m \to \mathbb{R}^n$ by $L^T(x) = A^T x$, then

$$(L(u), v) = (Au, v) = (Au)^T v = u^T A^T v = (u, A^T v) = (u, L^T(v))$$

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

We must now show that L^T is the *unique* linear transformation with this property.

Suppose $T : \mathbb{R}^m \to \mathbb{R}^n$ is a linear transformation such that $(L(u), v) = (u, T(v))$ for $u \in \mathbb{R}^n$ and $v \in \mathbb{R}^m$.

To show the uniqueness of L^T, we must prove that $L^T = T$.

Let B be the standard matrix of T. Then for all $u \in \mathbb{R}^n$ and $v \in \mathbb{R}^m$ we have

$$0 = (L(u), v) - (L(u), v) = (u, A^T v) - (u, Bv) = (u, (A^T - B)v).$$

Note that we can choose u and v independently to be any vectors we want (in $\mathbb{R}^n$ and $\mathbb{R}^m$).

Therefore if we set $X = A^T - B$ and $u = Xv$, then it follows that $(Xv, Xv) = 0$ for all $v \in \mathbb{R}^m$.

This forces us to have $Xv = 0$ for all $v \in \mathbb{R}^m$. Therefore X must be the zero matrix. (Why?)

Therefore $A^T = B$ and $L^T = T$. □

3 Inverses

Let $f : X \to Y$ be a function with domain X and codomain Y.

Definition. The function f is *invertible* or *bijective* if f is both onto and one-to-one.

Here is a more direct definition of an invertible function:

Proposition. The function f is invertible iff for each $b \in Y$ there is exactly one $a \in X$ with $f(a) = b$.

Proof. f is onto iff for each $b \in Y$ there is at least one input $a \in X$ with $f(a) = b$.

f is one-to-one iff for each $b \in Y$ there is at most one input $a \in X$ with $f(a) = b$.

Therefore f is both onto and one-to-one iff the given condition holds. □

The *identity function* on a set X is the function $\mathrm{id}_X : X \to X$ with $\mathrm{id}_X(a) = a$ for all $a \in X$.

If $f : X \to Y$ and $g : Y \to X$ are any functions then $f \circ \mathrm{id}_X = f$ and $\mathrm{id}_X \circ g = g$.

Example. The identity function on $\mathbb{R}^n$ is the linear transformation $\mathrm{id}_{\mathbb{R}^n} : \mathbb{R}^n \to \mathbb{R}^n$ whose standard matrix is the $n \times n$ *identity matrix*

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

If AI_n is defined then $AI_n = A$. If I_nA is defined then $I_nA = A$.

Even more concretely, a function is invertible if and only if it has an *inverse* in the following sense.

Proposition. The function $f : X \to Y$ is invertible if and only if there is a function $f^{-1} : Y \to X$ such that $f \circ f^{-1} = \mathrm{id}_Y$ and $f^{-1} \circ f = \mathrm{id}_X$.

If there exists a function f^{-1} with these properties, then it is unique, and we call it the *inverse* of f.

Proof. We first prove that the inverse is unique, and then prove the first statement. This takes 3 steps.

1. Suppose $g : Y \to X$ and $h : Y \to X$ both have $f \circ g = f \circ h = \mathrm{id}_Y$ and $g \circ f = h \circ f = \mathrm{id}_X$.

   Then $g \circ (f \circ h) = g \circ \mathrm{id}_Y = g$ and $(g \circ f) \circ h = \mathrm{id}_X \circ h = h$.

   But $g \circ (f \circ h)$ and $(g \circ f) \circ h$ are the same: given any $b \in Y$, they both have output $g(f(h(b)))$.

   Therefore $g = h$, so if f has an inverse, then it is unique.

2. If f has an inverse $f^{-1} : Y \to X$, then for each $b \in Y$ we have $f(a) = b$ for $a = f^{-1}(b)$.

   This value of $a \in X$ is unique since if $f(a) = f(a') = b$ then

   $$a = \mathrm{id}_X(a) = (f^{-1} \circ f)(a) = f^{-1}(f(a)) = f^{-1}(f(a')) = (f^{-1} \circ f)(a') = \mathrm{id}_X(a') = a'.$$

   Therefore, by the previous proposition, if f has an inverse then f is invertible.

3. Suppose f is invertible. Define $f^{-1}(b)$ for $b \in Y$ to be the unique $a \in X$ such $f(a) = b$.

   This defines a function $f^{-1} : Y \to X$ with $f(f^{-1}(b)) = f(a) = b$ and $f^{-1}(f(a)) = a$.

   This is the same as saying $f \circ f^{-1} = \mathrm{id}_Y$ and $f^{-1} \circ f = \mathrm{id}_X$.

$\square$

Example. Suppose $T : \mathbb{R}^2 \to \mathbb{R}^2$ is the linear function $T(v) = Av$ for $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$.

We have $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 \\ 0 & -2 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \mathsf{RREF}(A)$.

This means A has a pivot position in every row and every column.

By results in previous lectures, we know that this implies that T is onto and one-to-one, i.e., bijective.

What is the inverse T^{-1} of T?

Note that $T^{-1}\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right)$ is the unique vector $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ such that $Ax = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$.

We can solve for x by row reducing the augmented matrix of this matrix equation:

$$\begin{bmatrix} 1 & 2 & 1 \\ 3 & 4 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 1 \\ 0 & -2 & -3 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & -2 \\ 0 & -2 & -3 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & 3/2 \end{bmatrix},$$

which means that the equation's unique solution is $x = \boxed{T^{-1}\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} -2 \\ 3/2 \end{bmatrix}}$.

Similarly, $T^{-1}\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right)$ is the unique vector $y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$ such that $Ay = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$.

We again solve by row reduction to reduced echelon form:

$$\begin{bmatrix} 1 & 2 & 0 \\ 3 & 4 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 0 \\ 0 & -2 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 1 \\ 0 & -2 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & -1/2 \end{bmatrix}$$

which means that $y = \boxed{T^{-1}\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ -1/2 \end{bmatrix}}$.

If we knew that T^{-1} were linear, then we could conclude that

$$\boxed{T^{-1}(v) = Bv \quad \text{for} \quad B = \begin{bmatrix} -2 & 1 \\ 3/2 & -1/2 \end{bmatrix}}.$$

This turns out to be the right formula for T^{-1}. To see why, just check that

$$AB = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}\begin{bmatrix} -2 & 1 \\ 3/2 & -1/2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \text{and} \quad BA = \begin{bmatrix} -2 & 1 \\ 3/2 & -1/2 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

so $T \circ T^{-1} = T^{-1} \circ T = \mathrm{id}_{\mathbb{R}^2}$.

It turns out that the inverse of an invertible linear transformation is always linear. Moreover, a linear transformation is invertible only if its standard matrix is square:

Proposition. If $T : \mathbb{R}^n \to \mathbb{R}^m$ is linear and invertible, then $n = m$ and T^{-1} is linear.

Proof. From results last time, we know that T is onto only if $n \geq m$ and one-to-one only if $n \leq m$. If T is both, then necessarily $n = m$.

Recall how T^{-1} is defined. For $u, v \in \mathbb{R}^n$ and $c \in \mathbb{R}$, we have

- $T^{-1}(u + v) =$ the unique vector x with $T(x) = u + v$.

  Since $T(T^{-1}(u) + T^{-1}(v)) = T(T^{-1}(u)) + T(T^{-1}(v)) = u + v$, it follows that $x = T^{-1}(u) + T^{-1}(v)$.
- $T^{-1}(cv) =$ the unique vector y with $T(y) = cv$.

  Since $T(cT^{-1}(v)) = cT(T^{-1}(v)) = cv$, it follows that $y = cT^{-1}(v)$.

These two items confirm that T^{-1} is linear. □

As usual, let's now translate the notion of invertibility for linear functions to matrices.

Definition. Let A be an $n \times n$ matrix and define $T : \mathbb{R}^n \to \mathbb{R}^n$ by $T(x) = Ax$. The matrix A is *invertible* if the function T is invertible. Its *inverse* is the unique matrix A^{-1} such that $T^{-1}(x) = A^{-1}x$.

This definition is natural enough, but a little abstract. Here is a more concrete formulation:

Proposition. Let A be an $n \times n$ matrix. The following mean the same thing:

(1) A is invertible.

(2) There is an $n \times n$ matrix A^{-1} such that $AA^{-1} = A^{-1}A = I_n$.

(3) For each $b \in \mathbb{R}^n$ the equation $Ax = b$ has a unique solution.

(4) $\mathsf{RREF}(A) = I_n$.

Proof. To show that (1)-(4) are equivalent, we show that (1) implies (2), (2) implies (3), (3) implies (4), and (4) implies (1). This chain of implications shows that any of the properties implies any of the others.

Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be the linear transformation $T(X) = Ax$.

If A is invertible, then T is invertible with linear inverse $T^{-1} : \mathbb{R}^n \to \mathbb{R}^n$.

In this case if A^{-1} is the standard matrix of T^{-1} then

$$(T \circ T^{-1})(x) = A(A^{-1}x) = (AA^{-1})x = x$$
$$(T^{-1} \circ T)(x) = A^{-1}(Ax) = (A^{-1}A)x = x$$

for all $x \in \mathbb{R}^n$, which means that $AA^{-1} = A^{-1}A = I_n$. (Why?) Thus (1) $\Rightarrow$ (2).

If there is an $n \times n$ matrix A^{-1} such that $A^{-1}A = AA^{-1} = I_n$ then for each $b \in \mathbb{R}^n$, the unique solution to $Ax = b$ is $x = A^{-1}Ax = A^{-1}b$. (Check this!) Thus (2) $\Rightarrow$ (3).

If $Ax = b$ has a solution for every $b \in \mathbb{R}^n$ then A has a pivot position in every row. The solution to $Ax = b$ is unique if and only if $Ax = 0$ has only trivial solutions, which happens if and only if A has a pivot position in every column.

Thus if $Ax = b$ has a unique solution for every b then we must have $\mathsf{RREF}(A) = I_n$, so (3) $\Rightarrow$ (4).

Finally, if $\mathsf{RREF}(A) = I_n$ then A has a pivot position in every row and every column, so the columns of A are linearly independent (meaning T is one-to-one) and also span $\mathbb{R}^n$ (meaning T is onto), so T is invertible, which means A is invertible. So (4) $\Rightarrow$ (1). □

A synonym for an invertible matrix is a *non-singular matrix*.

A matrix which is *not* invertible is sometimes called *singular*.

Example (Warning)**.**

Suppose $A = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$ and $B = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$. We have $AB = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = I_3$.

Neither A nor B is invertible, however.

The problem is there is no matrix A' such that $A'A = I_4$ and no matrix B' such that $BB' = I_3$.

A is the standard matrix of a linear transformation which is not one-to-one. (Why?)

B is the standard matrix of a linear transformation which is not onto. (Why?)

The following is a useful formula for small calculations.

Theorem. Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ be an arbitrary 2×2 matrix.

(1) A is invertible if and only if $ad - bc \neq 0$.

(2) If $ad - bc \neq 0$ then $A^{-1} = \frac{1}{ad-bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$.

Proof. If $ad - bc = 0$ and either $b = 0$ or $d = 0$, then A has a row or column of all zeros (why?), so $\mathsf{RREF}(A)$ is not the identity matrix and A is not invertible.

If $ad - bc = 0$ and $b \neq 0$ and $d \neq 0$, then A is row equivalent to

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \sim \begin{bmatrix} ad & bd \\ c & d \end{bmatrix} \sim \begin{bmatrix} ad & bd \\ -bc & -bd \end{bmatrix} \sim \begin{bmatrix} ad & bd \\ 0 & 0 \end{bmatrix}$$

so $\mathsf{RREF}(A)$ cannot be the identity matrix, so A is not invertible.

If $ad - bc \neq 0$, then you can just check that $A^{-1}A = AA^{-1} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$. Try this yourself:

$$\frac{1}{ad-bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} \quad & \quad \\ \quad & \quad \end{bmatrix}$$

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \frac{1}{ad-bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} = \frac{1}{ad-bc} \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} = \begin{bmatrix} \quad & \quad \\ \quad & \quad \end{bmatrix}.$$

$\square$

Example. Recall in our earlier example we showed that $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}^{-1} = \begin{bmatrix} -2 & 1 \\ 3/2 & -1/2 \end{bmatrix} = \frac{1}{-2} \begin{bmatrix} 4 & -2 \\ -3 & 1 \end{bmatrix}$.

Some properties of the inverse of a matrix:

Theorem. Let A and B be $n \times n$ matrices.

1. If A is invertible then $(A^{-1})^{-1} = A$.
2. If A and B are both invertible then AB is invertible with $(AB)^{-1} = B^{-1}A^{-1}$.
3. If A is invertible then A^T is invertible with $(A^T)^{-1} = (A^{-1})^T$.

Proof.

1. To get a matrix C with $CA^{-1} = A^{-1}C = I_n$, take $C = A$.
2. Remember from last time that matrix multiplication is *associative*: this means that no matter how we parenthesize the product of a bunch of matrices, we get the same thing.

   Use this as follows: if A and B are invertible then $(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AA^{-1} = I_n$.

   Likewise, $(B^{-1}A^{-1})(AB) = B^{-1}(A^{-1}A)B = B^{-1}B = I_n$. Therefore $(AB)^{-1} = B^{-1}A^{-1}$.
3. Observe that $A^T(A^{-1})^T = (A^{-1}A)^T = I_n^T = I_n$ and $(A^{-1})^T A^T = (AA^{-1})^T = I_n^T = I_n$.

$\square$

A corollary of this theorem is that the product of a list of $n \times n$ invertible matrices is itself invertible, with inverse the product of the inverses in reverse order. In symbols: $(ABC \cdots Z)^{-1} = Z^{-1} \cdots C^{-1}B^{-1}A^{-1}$.

Checking whether an $n \times n$ matrix is invertible is nearly the same process as computing its inverse.

Process to compute A^{-1}

Let A be an $n \times n$ matrix. Consider the $n \times 2n$ matrix $\left[\begin{array}{cc} A & I_n \end{array}\right]$.

If A is invertible then $\mathsf{RREF}\left(\left[\begin{array}{cc} A & I_n \end{array}\right]\right) = \left[\begin{array}{cc} I_n & A^{-1} \end{array}\right]$.

So to compute A^{-1}, row reduce $\left[\begin{array}{cc} A & I_n \end{array}\right]$ to reduced echelon form, and then take the last n columns.

Example. To find the inverse of $A = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 0 & 3 \\ 4 & -3 & 8 \end{bmatrix}$ we row reduce

$$\left[\begin{array}{ccc|ccc} 0 & 1 & 2 & 1 & 0 & 0 \\ 1 & 0 & 3 & 0 & 1 & 0 \\ 4 & -3 & 8 & 0 & 0 & 1 \end{array}\right] \sim \left[\begin{array}{ccc|ccc} 1 & 0 & 3 & 0 & 1 & 0 \\ 0 & 1 & 2 & 1 & 0 & 0 \\ 4 & -3 & 8 & 0 & 0 & 1 \end{array}\right]$$

$$\sim \left[\begin{array}{ccc|ccc} 1 & 0 & 3 & 0 & 1 & 0 \\ 0 & 1 & 2 & 1 & 0 & 0 \\ 0 & -3 & -4 & 0 & -4 & 1 \end{array}\right]$$

$$\sim \left[\begin{array}{ccc|ccc} 1 & 0 & 3 & 0 & 1 & 0 \\ 0 & 1 & 2 & 1 & 0 & 0 \\ 0 & 0 & 2 & 3 & -4 & 1 \end{array}\right]$$

$$\sim \left[\begin{array}{ccc|ccc} 1 & 0 & 3 & 0 & 1 & 0 \\ 0 & 1 & 0 & -2 & 4 & -1 \\ 0 & 0 & 2 & 3 & -4 & 1 \end{array}\right]$$

$$\sim \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -9/2 & 7 & -3/2 \\ 0 & 1 & 0 & -2 & 4 & -1 \\ 0 & 0 & 2 & 3 & -4 & 1 \end{array}\right] \sim \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -9/2 & 7 & -3/2 \\ 0 & 1 & 0 & -2 & 4 & -1 \\ 0 & 0 & 1 & 3/2 & -2 & 1/2 \end{array}\right].$$

Now check directly that $A^{-1} = \begin{bmatrix} -9/2 & 7 & -3/2 \\ -2 & 4 & -1 \\ 3/2 & -2 & 1/2 \end{bmatrix}$!

4 Vocabulary

Keywords from today's lecture:

1. **Identity function** $\mathrm{id}_X : X \to X$.

   The function with $\mathrm{id}_X(x) = x$ for all $x \in X$.

   If $X = \mathbb{R}^4$ then id_X is the linear transformation with standard matrix $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

2. **Invertible function**.

   A function $f : X \to Y$ such that for each $y \in Y$ there is exactly one element $x \in X$ with $f(x) = y$.

   This occurs precisely when there is an *inverse function* $g : Y \to X$ such that $g(f(x)) = x$ and $f(g(y)) = y$ for all $x \in X$ and $y \in Y$, that is, with $g \circ f = \mathrm{id}_X$ and $f \circ g = \mathrm{id}_Y$.

3. **Invertible matrix** and the *inverse* of a matrix.

   A matrix is invertible if it is square and is the standard matrix of an invertible linear transformation $\mathbb{R}^n \to \mathbb{R}^n$ for some n.

   An $n \times n$ matrix A is invertible precisely when $\mathsf{RREF}(A) = I_n$ is the $n \times n$ identity matrix.

   If A is invertible, then there exists a unique matrix A^{-1} such that $AA^{-1} = A^{-1}A = I_n$.

   The matrix A^{-1} is called the inverse of A.

   If A is $n \times n$ and invertible then $\mathsf{RREF}(\begin{bmatrix} A & I_n \end{bmatrix}) = \begin{bmatrix} I_n & A^{-1} \end{bmatrix}$.

   Row reducing gives the most efficient way of calculating A^{-1}.

   Example: If $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ then A is invertible if and only if $ad - bc \neq 0$.

   When this happens, $A^{-1} = \frac{1}{ad-bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$.