Quiz

Give the definition of matrix-matrix multiplication (A times B) in which one of the matrices (which?) is broken into rows and columns (which?).
In your definition, specify how the output matrix is obtained from the input matrices A and B.
Illustrate your definition with an explicit numeric example.
Your definition should not explicitly mention dot-products or linear combinations.

What is the relation between matrix-matrix multiplication and function composition?
Illustrate with an explicit numeric example.

Review of concept of linear transformation

A function $f : \mathcal{V} \longrightarrow \mathcal{W}$ where $\mathcal{V}$ and $\mathcal{W}$ are vector spaces is a linear transformation if

Property L1: For every vector $\mathbf{v}$ in $\mathcal{V}$ and every scalar α in $\mathbb{F}$, $f(\alpha\,\mathbf{v}) = \alpha\, f(\mathbf{v})$

Property L2: For every two vectors $\mathbf{u}$ and $\mathbf{v}$ in $\mathcal{V}$, $f(\mathbf{u} + \mathbf{v}) = f(\mathbf{u}) + f(\mathbf{v})$

Things to note:

Most common example is: Given $R \times C$ matrix M,

function is $f : \mathbb{F}^C \longrightarrow \mathbb{F}^R$ with rule $f(\mathbf{x}) = M\mathbf{x}$

We proved that such a function is a linear transformation.

But this is not only kind of linear transformation:

- $\mathcal{V}$ and $\mathcal{W}$ can be arbitrary vector spaces (not just $\mathbb{F}^C$ or $\mathbb{F}^R$)
- The function f can be described in another way, e.g. "rotation by angle $\pi/3$ within the plane Span $\{[1, 2, 3], [9, 1, 7]\}$"

Nevertheless, it's true that if domain of f is a subspace of $\mathbb{F}^C$ and co-domain is a subspace of $\mathbb{F}^R$ and f is a linear transformation

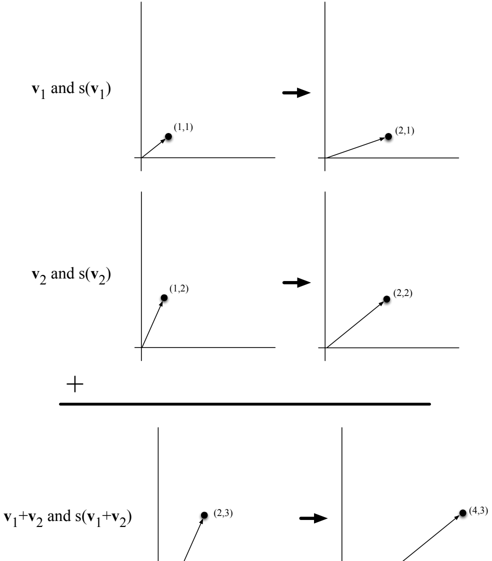
then f can be represented as matrix-vector multiplication.

Which functions are linear transformations?

Define $s([x, y])$ = stretching by two in horizontal direction

$$\text{Property L1:} \quad s(\mathbf{v}_1 + \mathbf{v}_2) = s(\mathbf{v}_1) + s(\mathbf{v}_2)$$

$$\text{Property L2:} \quad s(\alpha\, \mathbf{v}) = \alpha\, s(\mathbf{v})$$



Since the function $s(\cdot)$ satisfies Properties L1 and L2, it is a linear transformation.

Similarly can show rotation by θ degrees is a linear transformation.

What about translation?

$$t([x, y]) = [x, y] + [1, 2]$$

This function violates Property L1. For example:

$$t([4, 5] + [2, -1]) = t([6, 4]) = [7, 6]$$

but

$$t([4, 5]) + t([2, -1]) = [5, 7] + [3, 1] = [8, 8]$$

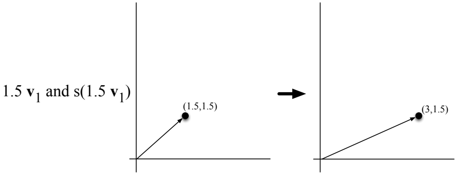
Which functions are linear transformations?

Define $s([x,y])$ = stretching by two in horizontal direction

$$\text{Property L1: } s(\mathbf{v}_1 + \mathbf{v}_2) = s(\mathbf{v}_1) + s(\mathbf{v}_2)$$

$$\text{Property L2: } s(\alpha\, \mathbf{v}) = \alpha\, s(\mathbf{v})$$

$\times\, 1.5$



Since the function $s(\cdot)$ satisfies Properties L1 and L2, it is a linear transformation.

Similarly can show rotation by θ degrees is a linear transformation.

What about translation?

$$t([x,y]) = [x,y] + [1,2]$$

This function violates Property L1. For example:

$$t([4,5] + [2,-1]) = t([6,4]) = [7,6]$$

but

$$t([4,5]) + t([2,-1]) = [5,7] + [3,1] = [8,8]$$

Since $t(\cdot)$ violates Property L1 for at least one input, it is not a linear transformation.

Can similarly show that $t(\cdot)$ does not satisfy Property L2.

Linear transformations: Pushing linear combinations through the function

Defining properties of linear transformations:

Property L1: $f(\alpha\,\mathbf{v}) = \alpha\, f(\mathbf{v})$

Property L2: $f(\mathbf{u} + \mathbf{v}) = f(\mathbf{u}) + f(\mathbf{v})$

Proposition: For a linear transformation f,
for any vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$ in the domain of f and any scalars $\alpha_1, \ldots, \alpha_n$,

$$f(\alpha_1\,\mathbf{v}_1 + \cdots + \alpha_n\,\mathbf{v}_n) = \alpha_1\, f(\mathbf{v}_1) + \cdots + \alpha_n\, f(\mathbf{v}_n)$$

"Proof": Consider the case of $n = 2$.

$$\begin{aligned} f(\alpha_1\,\mathbf{v}_1 + \alpha_2\,\mathbf{v}_2) &= f(\alpha_1\,\mathbf{v}_1) + f(\alpha_2\,\mathbf{v}_2) && \text{by Property L2} \\ &= \alpha_1\, f(\mathbf{v}_1) + \alpha_2\, f(\mathbf{v}_2) && \text{by Property L1} \end{aligned}$$

Proof for general n is similar. QED

linear transformations: Pushing linear combinations through the function

Proposition: For a linear transformation f,
$f(\alpha_1\, \mathbf{v}_1 + \cdots + \alpha_n\, \mathbf{v}_n) = \alpha_1\, f(\mathbf{v}_1) + \cdots + \alpha_n\, f(\mathbf{v}_n)$

Example: $f(\mathbf{x}) = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} * \mathbf{x}$

Verify that $f(10\,[1,-1] + 20\,[1,0]) = 10\, f([1,-1]) + 20\, f([1,0])$

$$\begin{aligned}
&\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \left(10\,[1,-1] + 20\,[1,0]\right) \\
&= \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \left([10,-10] + [20,0]\right) \\
&= \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} [30,-10] \\
&= 30\,[1,3] - 10[2,4] \\
&= [30,90] - [20,40]
\end{aligned}$$

$$\begin{aligned}
&10 \left(\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} * [1,-1]\right) + 20 \left(\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} * [1,0]\right) \\
&= 10\,([1,3] - [2,4]) + 20\,(1[1,3]) \\
&= 10\,[-1,-1] + 20\,[1,3] \\
&= [-10,-10] + [20,60] \\
&= [10,50]
\end{aligned}$$

From function to matrix, revisited

We saw a method to derive a matrix from a function:

Given a function $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$, we want a matrix M such that $f(\mathbf{x}) = M * \mathbf{x}$....

- Plug in the standard generators $\mathbf{e}_1 = [1, 0, \ldots, 0, 0], \ldots, \mathbf{e}_n = [0, \ldots, 0, 1]$
- Column i of M is $f(\mathbf{e}_i)$.

This works correctly whenever such a matrix M really exists:

Proof: If there is such a matrix then f is linear:

- (Property L1) $f(\alpha\,\mathbf{v}) = \alpha\, f(\mathbf{v})$ and
- (Property L2) $f(\mathbf{u} + \mathbf{v}) = f(\mathbf{u}) + f(\mathbf{v})$

Let $\mathbf{v} = [\alpha_1, \ldots, \alpha_n]$ be any vector in $\mathbb{R}^n$.

We can write $\mathbf{v}$ in terms of the standard generators.

$$\begin{aligned}
\mathbf{v} &= \alpha_1\,\mathbf{e}_1 + \cdots + \alpha_n\,\mathbf{e}_n \\
\text{so} & \\
f(\mathbf{v}) &= f(\alpha_1\,\mathbf{e}_1 + \cdots + \alpha_n\,\mathbf{e}_n) \\
&= \alpha_1\, f(\mathbf{e}_1) + \cdots + \alpha_n\, f(\mathbf{e}_n) \\
&= \alpha_1\,(\text{column } 1 \text{ of } M) + \cdots + \alpha_n\,(\text{column} n \text{ of } M) = M * \mathbf{v} \qquad \textit{QED}
\end{aligned}$$

A linear transformation maps zero vector to zero vector

Lemma: If $f : \mathcal{U} \longrightarrow \mathcal{V}$ is a linear transformation then f maps the zero vector of $\mathcal{U}$ to the zero vector of $\mathcal{V}$.

Proof: Let $\mathbf{0}$ denote the zero vector of $\mathcal{U}$, and let $\mathbf{0}_{\mathcal{V}}$ denote the zero vector of $\mathcal{V}$.

$$f(\mathbf{0}) = f(\mathbf{0} + \mathbf{0}) = f(\mathbf{0}) + f(\mathbf{0})$$

Subtracting $f(\mathbf{0})$ from both sides, we obtain

$$\mathbf{0}_{\mathcal{V}} = f(\mathbf{0})$$

QED

linear transformations and zero vectors: Kernel

Definition: *Kernel* of a linear transformation f is $\{\mathbf{v} \ : \ f(\mathbf{v}) = \mathbf{0}\}$

Written Ker f

For a function $f(\mathbf{x}) = M * \mathbf{x}$,

$$\text{Ker } f = \text{Null } M$$

Kernel and one-to-one

One-to-One Lemma: A linear transformation is one-to-one if and only if its kernel is a trivial vector space.

Proof: Let $f : \mathcal{U} \longrightarrow \mathcal{V}$ be a linear transformation. We prove two directions.

- Suppose Ker f contains some nonzero vector $\mathbf{u}$, so $f(\mathbf{u}) = \mathbf{0}_\mathcal{V}$.
  Because a linear transformation maps zero to zero, $f(\mathbf{0}) = \mathbf{0}_\mathcal{V}$ as well,
  so f is not one-to-one.

- Suppose Ker $f = \{\mathbf{0}\}$.
  Let $\mathbf{v}_1, \mathbf{v}_2$ be any vectors such that $f(\mathbf{v}_1) = f(\mathbf{v}_2)$.
  Then $f(\mathbf{v}_1) - f(\mathbf{v}_2) = \mathbf{0}_\mathcal{V}$
  so, by linearity, $f(\mathbf{v}_1 - \mathbf{v}_2) = \mathbf{0}_\mathcal{V}$,
  so $\mathbf{v}_1 - \mathbf{v}_2 \in$ Ker f.
  Since Ker f consists solely of $\mathbf{0}$,
  it follows that $\mathbf{v}_1 - \mathbf{v}_2 = \mathbf{0}$, so $\mathbf{v}_1 = \mathbf{v}_2$.

QED

Kernel and one-to-one

One-to-One Lemma A linear transformation is one-to-one if and only if its kernel is a trivial vector space.

Define the function $f(\mathbf{x}) = A * \mathbf{x}$.

If Ker f is trivial (i.e. if Null A is trivial)

then a vector $\mathbf{b}$ is the image under f of at most one vector.

That is, at most one vector $\mathbf{u}$ such that $A * \mathbf{u} = \mathbf{b}$
That is, the solution set of $A * \mathbf{x} = \mathbf{b}$ has at most one vector.

linear transformations that are onto?

Question: How can we tell if a linear transformation is onto?

Recall: for a function $f : \mathcal{V} \longrightarrow \mathcal{W}$, the *image* of f is the set of all images of elements of the domain:

$$\{f(\mathbf{v}) \ : \mathbf{v} \in \mathcal{V}\}$$

(You might know it as the "range" but we avoid that word here.)

The image of function f is written Im f

"Is function f is onto?" same as "is Im f = co-domain of f?"

Example: *Lights Out*

Define $f([\alpha_1, \alpha_2, \alpha_3, \alpha_4]) = \left[\begin{array}{c|c|c|c} \boxed{\begin{matrix}\bullet & \bullet \\ \bullet & \ \end{matrix}} & \boxed{\begin{matrix}\bullet & \bullet \\ \ & \bullet \end{matrix}} & \boxed{\begin{matrix}\bullet & \ \\ \bullet & \bullet \end{matrix}} & \boxed{\begin{matrix}\ & \bullet \\ \bullet & \bullet \end{matrix}} \end{array}\right] * [\alpha_1, \alpha_2, \alpha_3, \alpha_4]$

Im f is set of configurations for which 2×2 *Lights Out* can be solved,
so "f is onto" means "2×2 *Lights Out* can be solved for every configuration"

Can 2×2 *Lights Out* be solved for every configuration? What about 5×5?

Linear transformations that are onto?

“Is function f is onto?” same as “is $\operatorname{Im} f =$ co-domain of f?”

First step in understanding how to tell if a linear transformation f is onto:

- study the image of f

Proposition: The image of a linear transformation $f : \mathcal{V} \longrightarrow \mathcal{W}$ is a vector space

The image of a linear transformation is a vector space

Proposition: The image of a linear transformation $f : \mathcal{V} \longrightarrow \mathcal{W}$ is a vector space

Recall: a set $\mathcal{U}$ of vectors is a vector space if

V1: $\mathcal{U}$ contains a zero vector,

V2: for every vector $\mathbf{w}$ in $\mathcal{U}$ and every scalar α, the vector $\alpha\,\mathbf{w}$ is in $\mathcal{U}$

V3: for every pair of vectors $\mathbf{w}_1$ and $\mathbf{w}_2$ in $\mathcal{U}$, the vector $\mathbf{w}_1 + \mathbf{w}_2$ is in $\mathcal{U}$

Proof:

V1: Since the domain $\mathcal{V}$ contains a zero vector $\mathbf{0}_{\mathcal{V}}$ and $f(\mathbf{0}_{\mathcal{V}}) = \mathbf{0}_{\mathcal{W}}$, the image of f includes $\mathbf{0}_{\mathcal{W}}$. This proves Property V1.

V2: Suppose some vector $\mathbf{w}$ is in the image of f.
That means there is some vector $\mathbf{v}$ in the domain $\mathcal{V}$ that maps to $\mathbf{w}$: $f(\mathbf{v}) = \mathbf{w}$.
By Property L1, for any scalar α, $f(\alpha\,\mathbf{v}) = \alpha\, f(\mathbf{v}) = \alpha\,\mathbf{w}$
so $\alpha\,\mathbf{w}$ is in the image. This proves Property V2.

V3: Suppose vectors $\mathbf{w}_1$ and $\mathbf{w}_2$ are in the image of f.
That is, there are vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ in the domain such that $f(\mathbf{v}_1) = \mathbf{w}_1$ and $f(\mathbf{v}_2) = \mathbf{w}_2$.
By Property L2, $f(\mathbf{v}_1 + \mathbf{v}_2) = f(\mathbf{v}_1) + f(\mathbf{v}_2) = \mathbf{w}_1 + \mathbf{w}_2$
so $\mathbf{w}_1 + \mathbf{w}_2$ is in the image. This proves Property V3.

QED

linear transformations that are onto?

We've shown

Proposition: The image of a linear transformation $f : \mathcal{V} \longrightarrow \mathcal{W}$ is a vector space

This proposition shows that, for a linear transformation f, Im f is always a subspace of the co-domain $\mathcal{W}$.

The function is onto if Im f includes all of $\mathcal{W}$.

In a couple of weeks we will have a way to tell.

From function inverse to matrix inverse

Matrices A and $B \Rightarrow$ functions $f(\mathbf{y}) = A * \mathbf{y}$ and $g(\mathbf{x}) = B * \mathbf{x}$ and $h(\mathbf{x}) = (AB) * \mathbf{x}$

Def. If f and g are inverses of each other, we say A and B are matrix inverses of each other.

Example: An elementary row-addition matrix

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \Rightarrow \text{function } f([x_1, x_2, x_3]) = [x_1, x_2 + 2x_1, x_3])$$

Function adds twice the first entry to the second entry.

Functional inverse: subtracts the twice the first entry from the second entry:

$$f^{-1}([x_1, x_2, x_3]) = [x_1, x_2 - 2x_1, x_3]$$

Thus the inverse of A is

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

This matrix is also an elementary row-addition matrix.

Matrix inverse

If A and B are matrix inverses of each other, we say A and B are *invertible* matrices.

Can show that a matrix has at most one inverse.

We denote the inverse of matrix A by A^{-1}.

(A matrix that is not invertible is sometimes called a *singular* matrix, and an invertible matrix is called a *nonsingular* matrix.)

Invertible matrices: why care?

Reason 1: Existence and uniqueness of solution to matrix-vector equations.

Let A be an invertible $m \times n$ matrix, and define $f : \mathbb{F}^n \longrightarrow \mathbb{F}^m$ by $f(\mathbf{x}) = A\mathbf{x}$

Since A is an invertible matrix, f is an invertible function. Then f is one-to-one and onto:

- Since f is onto, for any m-vector $\mathbf{b}$ there is *some* vector $\mathbf{u}$ such that $f(\mathbf{u}) = \mathbf{b}$. That is, there is at least one solution to the matrix-vector equation $A\mathbf{x} = \mathbf{b}$.
- Since f is one-to-one, for any m-vector $\mathbf{b}$ there is *at most one* vector $\mathbf{u}$ such that $f(\mathbf{u}) = \mathbf{b}$. That is, there is at most one solution to $A\mathbf{x} = \mathbf{b}$.

For every right-hand side vector $\mathbf{b}$, the equation $A\mathbf{x} = \mathbf{b}$ has *exactly one solution*.

Example 1: Industrial espionage. Given the vector $\mathbf{b}$ specifying the amount of each resource consumed, figure out quantity of each product JunkCo has made.

Solve vector-matrix equation $M\mathbf{x} = \mathbf{b}$ where

M =

	garden gnome	hula hoop	slinky	silly putty	salad shooter
metal	0	0	.25	0	.15
concrete	1.3	0	0	0	0
plastic	.2	1.5	0	.3	.5
water	.8	.4	.2	.7	.4
electricity	.4	.3	.7	.5	.8

Invertible matrices: why care?

Reason 2: Algorithms for solving matrix-vector equation $A\mathbf{x} = \mathbf{b}$ are simpler if we can assume A is invertible.

Later we learn two such algorithms.

We also learn how to cope if A is not invertible.

Reason 3:
Invertible matrices play a key role in *change of basis*.

Change of basis is important part of linear algebra

- used e.g. in image compression;
- we will see it used in adding/removing perspective from an image.

Product of invertible matrices is an invertible matrix

Proposition: If A and B are invertible and AB is defined then AB is invertible.

Example:

$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$ correspond to functions

$f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ and $g : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$

$$f\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 + x_2 \\ x_2 \end{bmatrix}$$

$$g\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_1 + x_2 \end{bmatrix}$$

f is an invertible function.

g is an invertible function.

The functions f and g are invertible so the function $f \circ g$ is invertible.

By the Matrix-Multiplication Lemma, the function $f \circ g$ corresponds to the matrix product

$AB = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$ so that matrix is invertible.

Product of invertible matrices is an invertible matrix

Proposition: If A and B are invertible and AB is defined then AB is invertible.

Example:

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 4 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 5 & 0 & 1 \end{bmatrix}$$

Multiplication by matrix A adds 4 times first element to second element:

$f([x_1, x_2, x_3]) = [x_1, x_2 + 4x_1, x_3])$

This function is invertible.

Multiplication by matrix B adds 5 times first element to third element:

$g([x_1, x_2, x_3]) = [x_1, x_2, x_3 + 5x_1]$

This function is invertible

By Matrix Multiplication Lemma, multiplication by matrix AB corresponds to composition of functions $f \circ g$: $(f \circ g)([x_1, x_2, x_3]) = [x_1, x_2 + 4x_1, x_3 + 5x_1]$

The function $f \circ g$ is also an invertible function.... so AB is an invertible matrix.

Proof: Define the functions f and g by $f(\mathbf{x}) = A\mathbf{x}$ and $g(\mathbf{x}) = B\mathbf{x}$.

- Suppose A and B are invertible matrices.
  Then the corresponding functions f and g are invertible.

Product of invertible matrices is an invertible matrix

Proposition: If A and B are invertible and AB is defined then AB is invertible.

Example:

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \qquad B = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 1 & 0 \end{bmatrix}$$

The product is $AB = \begin{bmatrix} 4 & 5 & 1 \\ 10 & 11 & 4 \\ 16 & 17 & 7 \end{bmatrix}$

which is *not* invertible

so at least one of A and B is not invertible

and in fact $\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$ is *not* invertible.

Proof: Define the functions f and g by $f(\mathbf{x}) = A\mathbf{x}$ and $g(\mathbf{x}) = B\mathbf{x}$.

- Suppose A and B are invertible matrices.
  Then the corresponding functions f and g are invertible.

Product of invertible matrices is an invertible matrix

Proposition: If A and B are invertible and AB is defined then AB is invertible.

Proof: Define the functions f and g by $f(\mathbf{x}) = A\mathbf{x}$ and $g(\mathbf{x}) = B\mathbf{x}$.

- Suppose A and B are invertible matrices.
  Then the corresponding functions f and g are invertible.
  Therefore $f \circ g$ is invertible
  so the matrix corresponding to $f \circ g$ (which is AB) is an invertible matrix. QED

Matrix inverse

Lemma: If the $R \times C$ matrix A has an inverse A^{-1} then AA^{-1} is the $R \times R$ identity matrix.

Proof: Let $B = A^{-1}$. Define $f(\mathbf{x}) = A\mathbf{x}$ and $g(\mathbf{y}) = B\mathbf{y}$.

- By the Matrix-Multiplication Lemma, $f \circ g$ satisfies $(f \circ g)(\mathbf{x}) = AB\mathbf{x}$.
- On the other hand, $f \circ g$ is the identity function,
- so AB is the $R \times R$ identity matrix.

QED

Matrix inverse

Lemma: If the $R \times C$ matrix A has an inverse A^{-1} then AA^{-1} is identity matrix.

What about the converse?

Conjecture: If AB is an indentity matrix then A and B are inverses...?

Counterexample:

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$$

$$AB = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$A * [0, 0, 1]$ and $A * [0, 0, 0]$ both equal $[0, 0]$, so null space of A is not trivial,
so the function $f(\mathbf{x}) = A\mathbf{x}$ is not one-to-one,
so f is not an invertible function.

Shows: $AB = I$ is *not* sufficient to ensure that A and B are inverses.

Matrix inverse

Lemma: If the $R \times C$ matrix A has an inverse A^{-1} then AA^{-1} is identity matrix.

What about the converse?

FALSE Conjecture: If AB is an indentity matrix then A and B are inverses...?

Corollary: Matrices A and B are inverses of each other if and only if both AB and BA are identity matrices.

Matrix inverse

Lemma: If the $R \times C$ matrix A has an inverse A^{-1} then AA^{-1} is identity matrix.

Corollary: A and B are inverses of each other iff both AB and BA are identity matrices.

Proof:

- Suppose A and B are inverses of each other. By lemma, AB and BA are identity matrices.
- Suppose AB and BA are both identity matrices.
  Define $f(\mathbf{y}) = A * \mathbf{y}$ and $g(\mathbf{x}) = B * \mathbf{x}$
  - Because AB is identity matrix, by Matrix-Multiplication Lemma, $f \circ g$ is the identity function.
  - Because BA is identity matrix, by Matrix-Multiplication Lemma, $g \circ f$ is the identity function.
  - This proves that f and g are functional inverses of each other, so A and B are matrix inverses of each other.

QED

Matrix inverse

Question: How can we tell if a matrix M is invertible?

Partial Answer: By definition, M is an invertible matrix if the function $f(\mathbf{x}) = M\mathbf{x}$ is an invertible function, i.e. if the function is one-to-one and onto.

- *One-to-one:* Since the function is linear, we know by the One-to-One Lemma that the function is one-to-one if its kernel is trivial, i.e. if the null space of M is trivial.
- *Onto:* We haven't yet answered the question *how we can tell if a linear transformation is onto?*

If we knew how to tell if a linear transformation is onto, therefore, we would know how to tell if a matrix is invertible.