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 $h(\mathbf{x}) = (AB) * \mathbf{x}$

Definition If f and g are functional 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 $m \times n$ matrix, and define $f : \mathbb{F}^n \longrightarrow \mathbb{F}^m$ by $f(\mathbf{x}) = A\mathbf{x}$

Suppose A is an invertible matrix. Then 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 $A\mathbf{x} = \mathbf{b}$.

If A is invertible then, 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 $\mathbf{x}^T M = \mathbf{b}$ where

M =

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

Invertible matrices: why care?

Reason 1: Existence and uniqueness of solution to matrix-vector equations.
If A is invertible then, 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 $\mathbf{x}^T M = \mathbf{b}$ where

M =

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

Will this work for every vector $\mathbf{b}$?

- Is there a unique solution? If multiple solutions then we cannot be certain we have calculated true quantities.

Since M^T is an invertible matrix, the function $f(\mathbf{x}) = \mathbf{x} * M$ is an invertible function, so there is a unique solution for every vector $\mathbf{b}$.

Example 2: Sensor node with hardware components {radio,sensor,CPU,memory}.
Use three test periods

Invertible matrices: why care?

Reason 1: Existence and uniqueness of solution to matrix-vector equations.
If A is invertible then, for every right-hand side vector $\mathbf{b}$, the equation $A\mathbf{x} = \mathbf{b}$ has *exactly one solution*.

Example 2: Sensor node with hardware components $\{$radio,sensor,CPU,memory$\}$.
Use three test periods

- total power consumed in these test periods $\mathbf{b} = [140, 170, 60]$
- for each test period, vector says how long each hardware component worked:
  - $\mathbf{duration}_1$ = `Vec(D, 'radio':0.1, 'CPU':0.3)`
  - $\mathbf{duration}_2$ = `Vec(D, 'sensor':0.2, 'CPU':0.4)`
  - $\mathbf{duration}_3$ = `Vec(D, 'memory':0.3, 'CPU':0.1)`

To get $\mathbf{u}$, solve $A\mathbf{x} = \mathbf{b}$ where

$$A = \begin{bmatrix} \mathbf{duration}_1 \\ \hline \mathbf{duration}_2 \\ \hline \mathbf{duration}_3 \end{bmatrix}$$

Does this yield current draw for each hardware component?

- The matrix A is *not* invertible.
- In particular, the function $\mathbf{x} \mapsto A\mathbf{x}$ is not one-to-one.
- Therefore no guarantee that solution to $A\mathbf{x} = \mathbf{b}$ is unique
  $\Rightarrow$ we can't derive power per hardware component.
- Need to add more test periods....

Use **four** test periods

- total power consumed in these test periods $\mathbf{b} = [140, 170, 60, 170, 250]$

Invertible matrices: why care?

Reason 1: Existence and uniqueness of solution to matrix-vector equations.
If A is invertible then, for every right-hand side vector $\mathbf{b}$, the equation $A\mathbf{x} = \mathbf{b}$ has *exactly one solution*.
Example 2: Sensor node with hardware components $\{$radio,sensor,CPU,memory$\}$.
Use **four** test periods

- total power consumed in these test periods $\mathbf{b} = [140, 170, 60, 170, 250]$
- for each test period, vector says how long each hardware component worked:
  - $\mathbf{duration}_1 =$ `Vec(D, 'radio':0.1, 'CPU':0.3)`
  - $\mathbf{duration}_2 =$ `Vec(D, 'sensor':0.2, 'CPU':0.4)`
  - $\mathbf{duration}_3 =$ `Vec(D, 'memory':0.3, 'CPU':0.1)`
  - $\mathbf{duration}_4 =$ `Vec(D, 'memory':0.5, 'CPU':0.4)`

To get $\mathbf{u}$, solve $A\mathbf{x} = \mathbf{b}$ where

$$A = \begin{bmatrix} \mathbf{duration}_1 \\ \hline \mathbf{duration}_2 \\ \hline \mathbf{duration}_3 \\ \hline \mathbf{duration}_4 \end{bmatrix}$$

Does this yield current draw for each hardware component?

- This time the matrix A *is* invertible...
- so equation has exactly one solution.
- We can in principle find power consumption per component.

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: Suppose the matrix product AB is defined. Then AB is an invertible matrix if and only both A and B are invertible matrices.

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.

Example:

Product of invertible matrices is an invertible matrix

Proposition: Suppose the matrix product AB is defined. Then AB is an invertible matrix if and only both A and B are invertible matrices.

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.

Example:

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

Product of invertible matrices is an invertible matrix

Proposition: Suppose the matrix product AB is defined. Then AB is an invertible matrix if and only both A and B are invertible matrices.

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.
  Therefore $f \circ g$ is invertible
  so the matrix corresponding to $f \circ g$ (which is AB) is an invertible matrix

Product of invertible matrices is an invertible matrix

Proposition: Suppose the matrix product AB is defined. Then AB is an invertible matrix if and only both A and B are invertible matrices.

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.

- Conversely, suppose AB is an invertible matrix.
  Then the corresponding function $f \circ g$ is an invertible function.
  It follows that f and g must be invertible functions,
  so the corresponding matrices A and B must be invertible matrices.

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 function is onto?*

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