Inverses for Matrices that do NOT have Inverses

Teaching Students How to Think like Mathematicians

Professor Jeff Stuart

Mathematics Department
Pacific Lutheran University
Tacoma, WA 98447 USA
jeffrey.stuart@plu.edu

Pacific Lutheran University

January 2010

Everyone knows:

If A is $m \times n$, and we start with the augmented matrix

$$[A \mid I_m]$$

and perform row operations to obtain

$$[rref(A) \mid M]$$

where $rref(A)$ denotes the reduced row echelon form of A, then A is invertible if and only if $rref(A) = I_m$, in which case, $m = n$ and $M = A^{-1}$.

Someone in every class asks:

What does M tell us when $rref(A) \neq I_m$?

- How is this question answered?

Someone in every class asks:

What does M tell us when $rref(A) \neq I_m$?

- How is this question answered?
- The common answer: Nothing obvious.

Someone in every class asks:

What does M tell us when $rref(A) \neq I_m$?

- How is this question answered?
- The common answer: Nothing obvious.
- The other common answer: Nothing useful.

Someone in every class asks:

What does M tell us when $rref(A) \neq I_m$?

- How is this question answered?
- The common answer: Nothing obvious.
- The other common answer: Nothing useful.
- The less common answer: It's too complicated.

Someone in every class asks:

What does M tell us when $rref(A) \neq I_m$?

- How is this question answered?
- The common answer: Nothing obvious.
- The other common answer: Nothing useful.
- The less common answer: It's too complicated.
- The best answer: Hmmm...Great Question. Let's explore that.

NOT the Goal of this Talk

My goal is NOT to answer the mathematical question about M.

- The answer is too complicated for a fifteen minute talk.

NOT the Goal of this Talk

My goal is NOT to answer the mathematical question about M.

- The answer is too complicated for a fifteen minute talk.
- The answer is mostly well-known, there is a huge literature devoted to generalized inverses.

NOT the Goal of this Talk

My goal is NOT to answer the mathematical question about M.

- The answer is too complicated for a fifteen minute talk.
- The answer is mostly well-known, there is a huge literature devoted to generalized inverses.
- The answer is developed in an expository paper I have written for a good student in a first linear algebra course. E-mail me.

The Goal of this Talk

My goal for this talk is to show how I use this question as a means **to introduce students to how to think like mathematicians.**

- This can be accomplished in the course of one lecture.

The Goal of this Talk

My goal for this talk is to show how I use this question as a means **to introduce students to how to think like mathematicians.**

- This can be accomplished in the course of one lecture.
- This illuminates the framework of

The Goal of this Talk

My goal for this talk is to show how I use this question as a means **to introduce students to how to think like mathematicians.**

- This can be accomplished in the course of one lecture.
- This illuminates the framework of
    - Start with a known property or definition.

The Goal of this Talk

My goal for this talk is to show how I use this question as a means **to introduce students to how to think like mathematicians.**

- This can be accomplished in the course of one lecture.
- This illuminates the framework of
  - Start with a known property or definition.
  - Generalize by relaxing a condition.

The Goal of this Talk

My goal for this talk is to show how I use this question as a means **to introduce students to how to think like mathematicians.**

- This can be accomplished in the course of one lecture.
- This illuminates the framework of
  - Start with a known property or definition.
  - Generalize by relaxing a condition.
  - Explore simple or semi-simple examples.

The Goal of this Talk

My goal for this talk is to show how I use this question as a means **to introduce students to how to think like mathematicians.**

- This can be accomplished in the course of one lecture.
- This illuminates the framework of
  - Start with a known property or definition.
  - Generalize by relaxing a condition.
  - Explore simple or semi-simple examples.
  - Find the limitations of the generalization.

The Goal of this Talk

My goal for this talk is to show how I use this question as a means **to introduce students to how to think like mathematicians.**

- This can be accomplished in the course of one lecture.
- This illuminates the framework of
  - Start with a known property or definition.
  - Generalize by relaxing a condition.
  - Explore simple or semi-simple examples.
  - Find the limitations of the generalization.
  - Decide what properties you really want from the original.

The Goal of this Talk

My goal for this talk is to show how I use this question as a means **to introduce students to how to think like mathematicians.**

- This can be accomplished in the course of one lecture.
- This illuminates the framework of
    - Start with a known property or definition.
    - Generalize by relaxing a condition.
    - Explore simple or semi-simple examples.
    - Find the limitations of the generalization.
    - Decide what properties you really want from the original.
    - Look for a better (less obvious?) relaxation.

The Goal of this Talk

My goal for this talk is to show how I use this question as a means **to introduce students to how to think like mathematicians.**

- This can be accomplished in the course of one lecture.
- This illuminates the framework of
  - Start with a known property or definition.
  - Generalize by relaxing a condition.
  - Explore simple or semi-simple examples.
  - Find the limitations of the generalization.
  - Decide what properties you really want from the original.
  - Look for a better (less obvious?) relaxation.
- Along the way, review and reinforce important topical material.

Along the Way – Review and Reinforce

The Marvelous Connections between:

- Gaussian Elimination

$$A \longrightarrow (\text{row op's}) \longrightarrow B$$

- Matrix Multiplication

$$B = PA$$

  where P is invertible
- Linear Combinations of Rows: The rows of B are linear combos of the rows of A with the coefficients lurking in P.

Along the Way – Review and Reinforce

The Marvelous Connections between:

- The row space, column space and null space of A
- The row space, column space and null space of A^T
- The row space and null space of $rref(A)$.
- Subspaces of the row space of P

Partitioned Matrices and Matrix Multiplication

$A \longrightarrow rref(A)$ means $rref(A) = PA$ for some invertible matrix P
so

$$P[A \mid I_m] = [PA \mid PI_m] = [rref(A) \mid P]$$

Comparing that to

$$[A \mid I_m] \longrightarrow [rref(A) \mid M]$$

suggests $M = P$.

A Semi-Simple Example

Suppose that $A = \begin{bmatrix} 2 & 2 \\ 1 & 1 \end{bmatrix}$

Then taking the operations $\frac{1}{2}R_1$ and $R_2 - \frac{1}{2}R_1$ give

$$[A \mid I_2] \rightarrow \left[\begin{array}{cc|cc} 1 & 1 & 1/2 & 0 \\ 0 & 0 & -1/2 & 1 \end{array}\right]$$

Alternatively, the operation $R_1 \leftrightarrow R_2$ followed by $R_2 - 2R_1$ gives

$$[A \mid I_2] \rightarrow \left[\begin{array}{cc|cc} 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & -2 \end{array}\right]$$

so M is not unique! Also, neither $MA = I_2$ nor $AM = I_2$ holds for these two choices of M.

The Obvious Relaxation

- Perhaps the condition $AB = BA = I_n$ in the definition of an inverse is too strict.

The Obvious Relaxation

- Perhaps the condition $AB = BA = I_n$ in the definition of an inverse is too strict.
- What if we only require $AB = I_m$ or $BA = I_n$?

The Obvious Relaxation

- Perhaps the condition $AB = BA = I_n$ in the definition of an inverse is too strict.
- What if we only require $AB = I_m$ or $BA = I_n$?
- If $A = \mathbf{0}_{m \times n}$, there is obviously no such B. (Good!)

The Obvious Relaxation

- Perhaps the condition $AB = BA = I_n$ in the definition of an inverse is too strict.
- What if we only require $AB = I_m$ or $BA = I_n$?
- If $A = \mathbf{0}_{m \times n}$, there is obviously no such B. (Good!)
- If A is $m \times n$, then $AB = I_m$ or $BA = I_n$ forces B to be $n \times m$.

The Obvious Relaxation

- Perhaps the condition $AB = BA = I_n$ in the definition of an inverse is too strict.
- What if we only require $AB = I_m$ or $BA = I_n$?
- If $A = \mathbf{0}_{m \times n}$, there is obviously no such B. (Good!)
- If A is $m \times n$, then $AB = I_m$ or $BA = I_n$ forces B to be $n \times m$.
- If $AB = I_m$, then $x = Bb$ always solves $Ax = b$ so $Ax = b$ is always consistent, and hence, $rank(A) = m$.

The Obvious Relaxation

- Perhaps the condition $AB = BA = I_n$ in the definition of an inverse is too strict.
- What if we only require $AB = I_m$ or $BA = I_n$?
- If $A = \mathbf{0}_{m \times n}$, there is obviously no such B. (Good!)
- If A is $m \times n$, then $AB = I_m$ or $BA = I_n$ forces B to be $n \times m$.
- If $AB = I_m$, then $x = Bb$ always solves $Ax = b$ so $Ax = b$ is always consistent, and hence, $rank(A) = m$.
- If $BA = I_n$, then $Ax = \mathbf{0}$ implies $x = I_n x = BAx = B\mathbf{0} = \mathbf{0}$, and hence, $rank(A) = n$.

The Obvious Relaxation

- Perhaps the condition $AB = BA = I_n$ in the definition of an inverse is too strict.
- What if we only require $AB = I_m$ or $BA = I_n$?
- If $A = \mathbf{0}_{m \times n}$, there is obviously no such B. (Good!)
- If A is $m \times n$, then $AB = I_m$ or $BA = I_n$ forces B to be $n \times m$.
- If $AB = I_m$, then $x = Bb$ always solves $Ax = b$ so $Ax = b$ is always consistent, and hence, $rank(A) = m$.
- If $BA = I_n$, then $Ax = \mathbf{0}$ implies $x = I_n x = BAx = B\mathbf{0} = \mathbf{0}$, and hence, $rank(A) = n$.
- B exists if and only if $rank(A) = \min\{m, n\}$.

What inverse properties do we seek?

- (matrix algebra) $AB = BA = I$.
- (matrix arithmetic) Bb always uniquely solves $Ax = b$.
- (algorithm) B is constructed by elementary row operations from A.

What inverse properties do we seek?

- (matrix algebra) $AB = BA = I$.
- (matrix arithmetic) Bb always uniquely solves $Ax = b$.
- (algorithm) B is constructed by elementary row operations from A.
- When $rank(A) = m$, we can find B so that $AB = I_n$ and Bb always solves $Ax = b$.

What inverse properties do we seek?

- (matrix algebra) $AB = BA = I$.
- (matrix arithmetic) Bb always uniquely solves $Ax = b$.
- (algorithm) B is constructed by elementary row operations from A.
- When $rank(A) = m$, we can find B so that $AB = I_n$ and Bb always solves $Ax = b$.
- When $rank(A) = n$, we can find B so that $BA = I_n$ and Bb uniquely solves *consistent* $Ax = b$.

What inverse properties do we seek?

- (matrix algebra) $AB = BA = I$.
- (matrix arithmetic) Bb always uniquely solves $Ax = b$.
- (algorithm) B is constructed by elementary row operations from A.
- When $rank(A) = m$, we can find B so that $AB = I_n$ and Bb always solves $Ax = b$.
- When $rank(A) = n$, we can find B so that $BA = I_n$ and Bb uniquely solves *consistent* $Ax = b$.
- When $rank(A) < \min\{m, n\}$, can we find an analog to $AB = BA = I$ that always has a solution?

What inverse properties do we seek?

- (matrix algebra) $AB = BA = I$.
- (matrix arithmetic) Bb always uniquely solves $Ax = b$.
- (algorithm) B is constructed by elementary row operations from A.
- When $rank(A) = m$, we can find B so that $AB = I_n$ and Bb always solves $Ax = b$.
- When $rank(A) = n$, we can find B so that $BA = I_n$ and Bb uniquely solves *consistent* $Ax = b$.
- When $rank(A) < \min\{m, n\}$, can we find an analog to $AB = BA = I$ that always has a solution?
- When $rank(A) < \min\{m, n\}$, can we find a B so Bb uniquely solves *consistent* $Ax = b$?

What inverse properties do we seek?

- (matrix algebra) $AB = BA = I$.
- (matrix arithmetic) Bb always uniquely solves $Ax = b$.
- (algorithm) B is constructed by elementary row operations from A.
- When $rank(A) = m$, we can find B so that $AB = I_n$ and Bb always solves $Ax = b$.
- When $rank(A) = n$, we can find B so that $BA = I_n$ and Bb uniquely solves *consistent* $Ax = b$.
- When $rank(A) < \min\{m, n\}$, can we find an analog to $AB = BA = I$ that always has a solution?
- When $rank(A) < \min\{m, n\}$, can we find a B so Bb uniquely solves *consistent* $Ax = b$?
- Can we use elementary row operations to build B from A?

A Better Relaxation

It looks so deceptively simple in retrospect:

- $ABA = A$

A Better Relaxation

It looks so deceptively simple in retrospect:

- $ABA = A$
- Contains $AB = I_m$ and $BA = I_n$ and $AB = BA = I$!

A Better Relaxation

It looks so deceptively simple in retrospect:

- $ABA = A$
- Contains $AB = I_m$ and $BA = I_n$ and $AB = BA = I$!
- $ABA = A$ means Bb always solves a consistent $Ax = b$.

A Better Relaxation

It looks so deceptively simple in retrospect:

- $ABA = A$

- Contains $AB = I_m$ and $BA = I_n$ and $AB = BA = I$!
- $ABA = A$ means Bb always solves a consistent $Ax = b$.
- B can always be constructed using elementary row **and/or** elementary column operations **and** a single matrix multiplication.

A Better Relaxation

It looks so deceptively simple in retrospect:

- $ABA = A$
- Contains $AB = I_m$ and $BA = I_n$ and $AB = BA = I$!
- $ABA = A$ means Bb always solves a consistent $Ax = b$.
- B can always be constructed using elementary row **and/or** elementary column operations **and** a single matrix multiplication.
- Symmetry of $AB = BA = I$ suggests that we want **both** $ABA = A$ and $BAB = B$.

A Better Relaxation

It looks so deceptively simple in retrospect:

- $ABA = A$

- Contains $AB = I_m$ and $BA = I_n$ and $AB = BA = I$!
- $ABA = A$ means Bb always solves a consistent $Ax = b$.
- B can always be constructed using elementary row **and/or** elementary column operations **and** a single matrix multiplication.

- Symmetry of $AB = BA = I$ suggests that we want **both** $ABA = A$ and $BAB = B$.
  - This forces $rank(A) = rank(B)$.

An Imperfect Relaxation

The last step of thinking like a mathematician: **why we might want more and why there is more to explore:**

- $Ax = b$ **not** always consistent.

An Imperfect Relaxation

The last step of thinking like a mathematician: **why we might want more and why there is more to explore:**

- $Ax = b$ **not** always consistent.
- Consistent $Ax = b$ usually does **not** have a unique solution.

An Imperfect Relaxation

The last step of thinking like a mathematician: **why we might want more and why there is more to explore:**

- $Ax = b$ **not** always consistent.
- Consistent $Ax = b$ usually does **not** have a unique solution.
- B is usually **not** unique.

An Imperfect Relaxation

The last step of thinking like a mathematician: **why we might want more and why there is more to explore:**

- $Ax = b$ **not** always consistent.
- Consistent $Ax = b$ usually does **not** have a unique solution.
- B is usually **not** unique.
- B is usually **not** M in $[rref(A) \mid M]$.

An Imperfect Relaxation

The last step of thinking like a mathematician: **why we might want more and why there is more to explore:**

- $Ax = b$ **not** always consistent.
- Consistent $Ax = b$ usually does **not** have a unique solution.
- B is usually **not** unique.
- B is usually **not** M in $[rref(A) \mid M]$.
- Bb usually does **not** have certain desirable properties:

An Imperfect Relaxation

The last step of thinking like a mathematician: **why we might want more and why there is more to explore:**

- $Ax = b$ **not** always consistent.
- Consistent $Ax = b$ usually does **not** have a unique solution.
- B is usually **not** unique.
- B is usually **not** M in $[rref(A) \mid M]$.
- Bb usually does **not** have certain desirable properties:
  - least squares solution $\cdots$ Bb minimizes $\|Ax - b\|_2$

An Imperfect Relaxation

The last step of thinking like a mathematician: **why we might want more and why there is more to explore:**

- $Ax = b$ **not** always consistent.
- Consistent $Ax = b$ usually does **not** have a unique solution.
- B is usually **not** unique.
- B is usually **not** M in $[rref(A) \mid M]$.
- Bb usually does **not** have certain desirable properties:
    - least squares solution $\cdots$ Bb minimizes $\|Ax - b\|_2$
    - minimum norm solution $\cdots$ $\|Bb\|_2$ is minimum among all solutions to consistent $Ax = b$.

An Imperfect Relaxation

The last step of thinking like a mathematician: **why we might want more and why there is more to explore:**

- $Ax = b$ **not** always consistent.
- Consistent $Ax = b$ usually does **not** have a unique solution.
- B is usually **not** unique.
- B is usually **not** M in $[rref(A) \mid M]$.
- Bb usually does **not** have certain desirable properties:
  - least squares solution $\cdots$ Bb minimizes $\|Ax - b\|_2$
  - minimum norm solution $\cdots$ $\|Bb\|_2$ is minimum among all solutions to consistent $Ax = b$.
- B such that Bb is least squares and/or minimum norm solution is not usually constructible without using much more sophisticated tools.

Think "**Moore-Penrose**".
(Give yourself an A^+ for making it this far?)

Thank you!

These slides will be available on the author's website:

http://www.plu.edu/~stuartjl/

Just Google: "Stuart" AND "Matrices without Inverses"