Matrix Inversion & LU Factorization

Geometric Algorithms
Lecture 11

CAS CS 132

Objectives

» Demonstrate how to **invert** a matrix

» Motivate **matrix factorization** in general, and the LU factorization in specific

» Recall elementary row operations and connect them to matrices

» Look at the **LU factorization**, how to find it, and how to use it

Keywords

Matrix Inverse

Invertible Transformation

1–1 Correspondence

numpy.linalg.inv

Determinant

Invertible Matrix Theorem

elementary matrices

LU factorization

Matrix Inverses

Basic Algebra

$$2x = 10$$

Basic Algebra

$$2x = 10$$

How do we solve this equation?

Basic Algebra

$$2x = 10$$

How do we solve this equation?

Divide on both sides by 2 to get $x = 5$.

Basic Algebra

$$2x = 10$$

How do we solve this equation?

Divide on both sides by 2 to get $x = 5$.

Multiply each side by $\frac{1}{2}$ *a.k.a.* 2^{-1}.

Basic Algebra

$$2x = 10$$

How do we solve this equation?

Divide on both sides by 2 to get $x = 5$.

Multiply each side by $\frac{1}{2}$ *a.k.a.* 2^{-1}*.*

$\frac{1}{2}$ is the **reciprocal** or **multiplicative inverse** of 2.

Basic Algebra

$$2^{-1}(2x) = 2^{-1}(10)$$

How do we solve this equation?

Divide on both sides by 2 to get $x = 5$.

Multiply each side by $\frac{1}{2}$ *a.k.a.* 2^{-1}.

$\frac{1}{2}$ is the **reciprocal** or **multiplicative inverse** of 2.

Basic Algebra

$$1x = 5$$

How do we solve this equation?

Divide on both sides by 2 to get $x = 5$.

Multiply each side by $\frac{1}{2}$ *a.k.a.* 2^{-1}.

$\frac{1}{2}$ is the **reciprocal** or **multiplicative inverse** of 2.

Basic Algebra

$$x = 5$$

How do we solve this equation?

Divide on both sides by 2 to get $x = 5$.

Multiply each side by $\frac{1}{2}$ *a.k.a.* 2^{-1}.

$\frac{1}{2}$ is the **reciprocal** or **multiplicative inverse** of 2.

Wouldn't it be nice...

$$A\mathbf{x} = \mathbf{b}$$

Wouldn't it be nice...

$$A\mathbf{x} = \mathbf{b}$$

```
How do we solve this equation?
```

Wouldn't it be nice...

$$A\mathbf{x} = \mathbf{b}$$

How do we solve this equation?

Multiply each side by A^{-1} to get $\mathbf{x} = A^{-1}\mathbf{b}$.

Wouldn't it be nice...

$$A\mathbf{x} = \mathbf{b}$$

How do we solve this equation?

Multiply each side by A^{-1} to get $\mathbf{x} = A^{-1}\mathbf{b}$.

A^{-1} is the **multiplicative inverse** of A

Wouldn't it be nice...

$$A^{-1}A\mathbf{x} = A^{-1}\mathbf{b}$$

How do we solve this equation?

Multiply each side by A^{-1} to get $\mathbf{x} = A^{-1}\mathbf{b}$.

A^{-1} is the **multiplicative inverse** of A

Wouldn't it be nice...

$$I\mathbf{x} = A^{-1}\mathbf{b}$$

How do we solve this equation?

Multiply each side by A^{-1} to get $\mathbf{x} = A^{-1}\mathbf{b}$.

A^{-1} is the **multiplicative inverse** of A

Wouldn't it be nice...

$$\mathbf{x} = A^{-1}\mathbf{b}$$

How do we solve this equation?

Multiply each side by A^{-1} to get $\mathbf{x} = A^{-1}\mathbf{b}$.

A^{-1} is the **multiplicative inverse** of A

Matrix Inverses

Matrix Inverses

Definition. For a $n \times n$ matrix A, an **inverse** of A is a $n \times n$ matrix B such that

$$AB = I_n \text{ and } BA = I_n$$

Matrix Inverses

Definition. For a $n \times n$ matrix A, an **inverse** of A is a $n \times n$ matrix B such that

$$AB = I_n \text{ and } BA = I_n$$

A is **invertible** if it has an inverse. Otherwise it is **singular.**

Matrix Inverses

Definition. For a $n \times n$ matrix A, an **inverse** of A is a $n \times n$ matrix B such that

$$AB = I_n \text{ and } BA = I_n$$

A is **invertible** if it has an inverse. Otherwise it is **singular.**

Example. $\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}$

Example: Geometric

Reflection across the x_1–axis in $\mathbb{R}^2$ is it's own inverse.

Verify:

Example: No inverse

Verify:

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

Inverses are Unique

Theorem. If B and C are inverses of A, then $B = C$.

Verify:

Inverses are Unique

Theorem. If B and C are inverses of A, then $B = C$.

Verify:

If A is invertible, then we write A^{-1} for *the* inverse of A.

Unique Solutions

If $A\mathbf{x} = \mathbf{b}$ has a <u>unique</u> solution for any choice of $\mathbf{b}$, then it has

» <u>exactly one</u> solution for any choice of $\mathbf{b}$

Unique Solutions

If $A\mathbf{x} = \mathbf{b}$ has a unique solution for any choice of $\mathbf{b}$, then it has

» at least one solution for any choice of $\mathbf{b}$

» at most one solution for any choice of $\mathbf{b}$

Unique Solutions

If $A\mathbf{x} = \mathbf{b}$ has a <u>unique</u> solution for any choice of $\mathbf{b}$, then it has

» T is <u>onto</u>

» T is <u>one-to-one</u>

where T is implemented by A

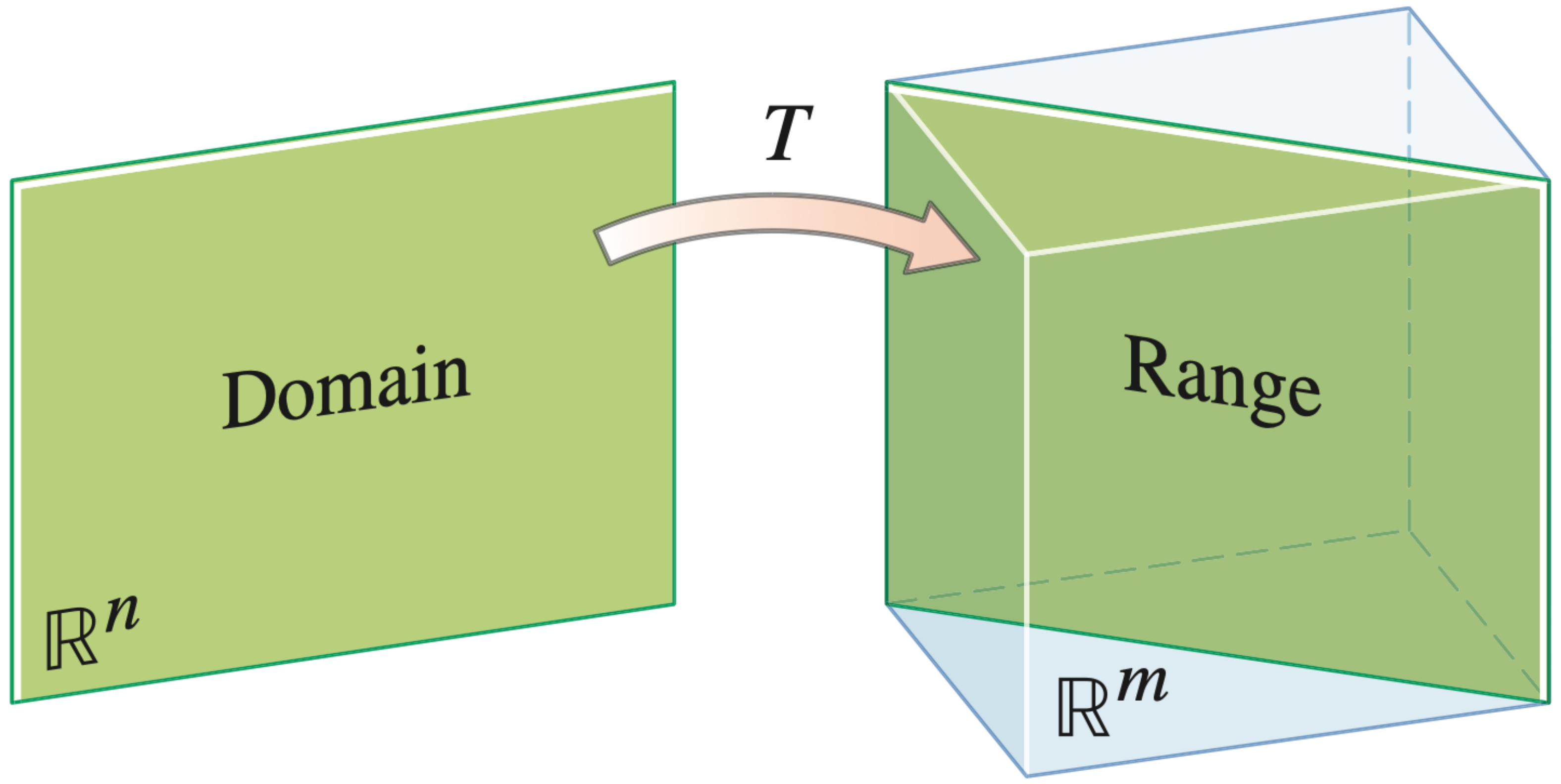
Recall: Onto Transformations

Recall: Onto Transformations

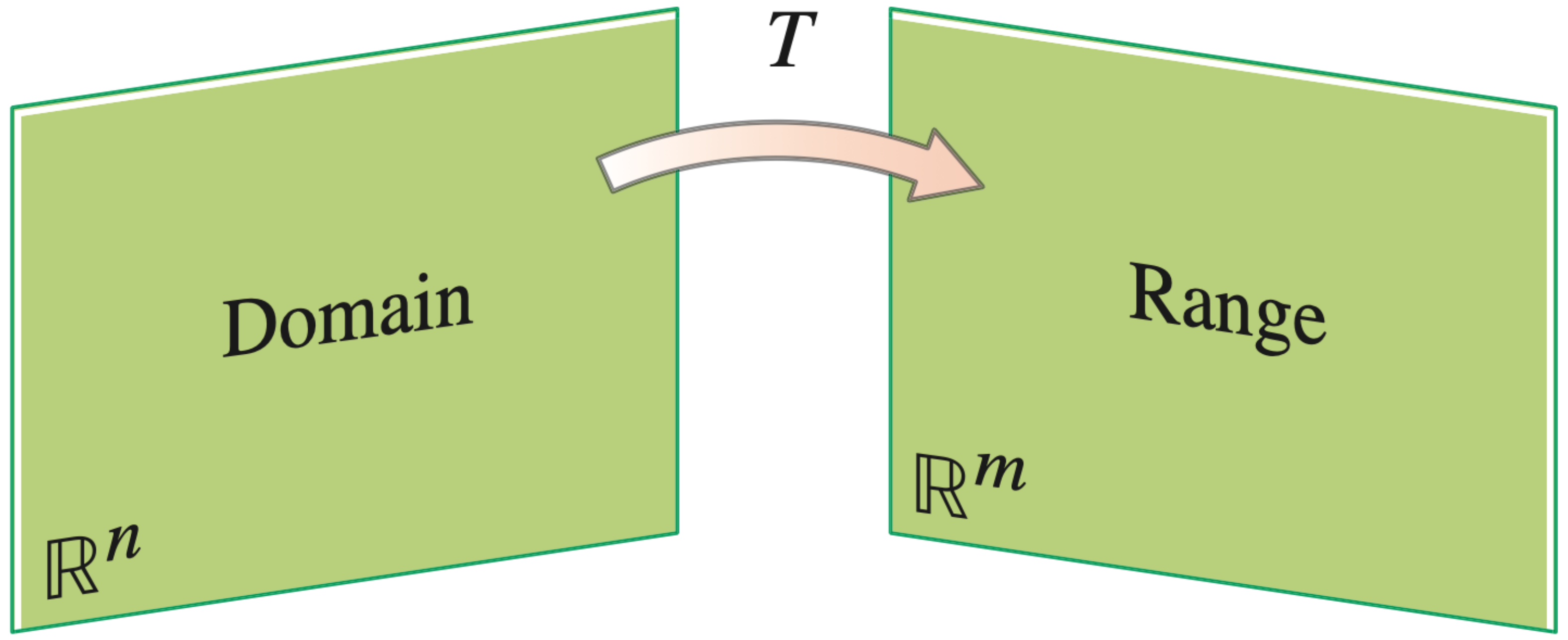
Definition. A transformation $T:\mathbb{R}^n \to \mathbb{R}^m$ is ***onto*** if any vector $\mathbf{b}$ in $\mathbb{R}^m$ is the image of at least one vector $\mathbf{v}$ in $\mathbb{R}^n$ (where $T(\mathbf{v}) = \mathbf{b}$).

Recall: Onto Transformations

Definition. A transformation $T: \mathbb{R}^n \to \mathbb{R}^m$ is ***onto*** if any vector $\mathbf{b}$ in $\mathbb{R}^m$ is the image of at least one vector $\mathbf{v}$ in $\mathbb{R}^n$ (where $T(\mathbf{v}) = \mathbf{b}$).



T is *not* onto $\mathbb{R}^m$

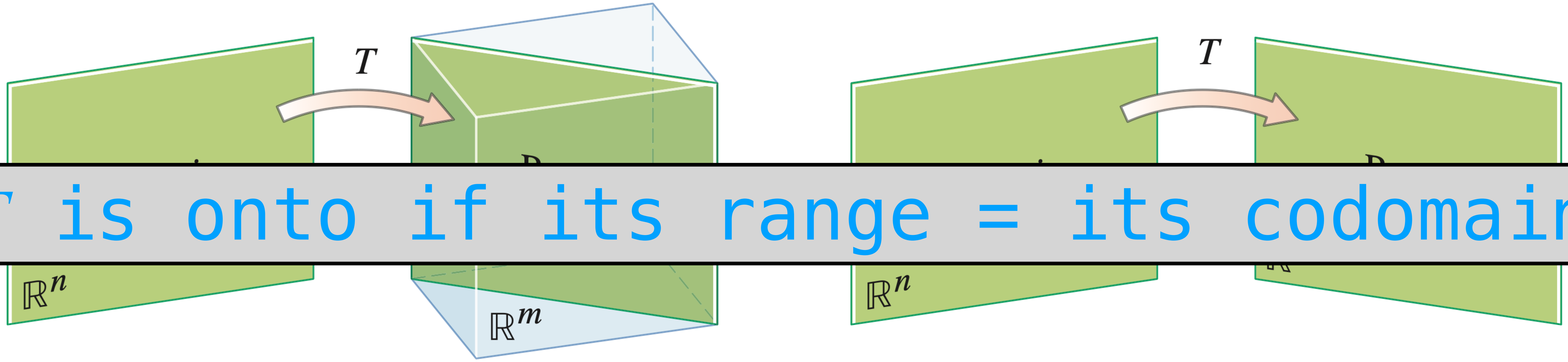


T is onto $\mathbb{R}^m$

image source: Linear Algebra and its Applications. Lay, Lay, and McDonald

Recall: Onto Transformations

Definition. A transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ is ***onto*** if any vector $\mathbf{b}$ in $\mathbb{R}^m$ is the image of at least one vector $\mathbf{v}$ in $\mathbb{R}^n$ (where $T(\mathbf{v}) = \mathbf{b}$).



T is onto if its range = its codomain

T is *not* onto $\mathbb{R}^m$

T is onto $\mathbb{R}^m$

image source: Linear Algebra and its Applications. Lay, Lay, and McDonald

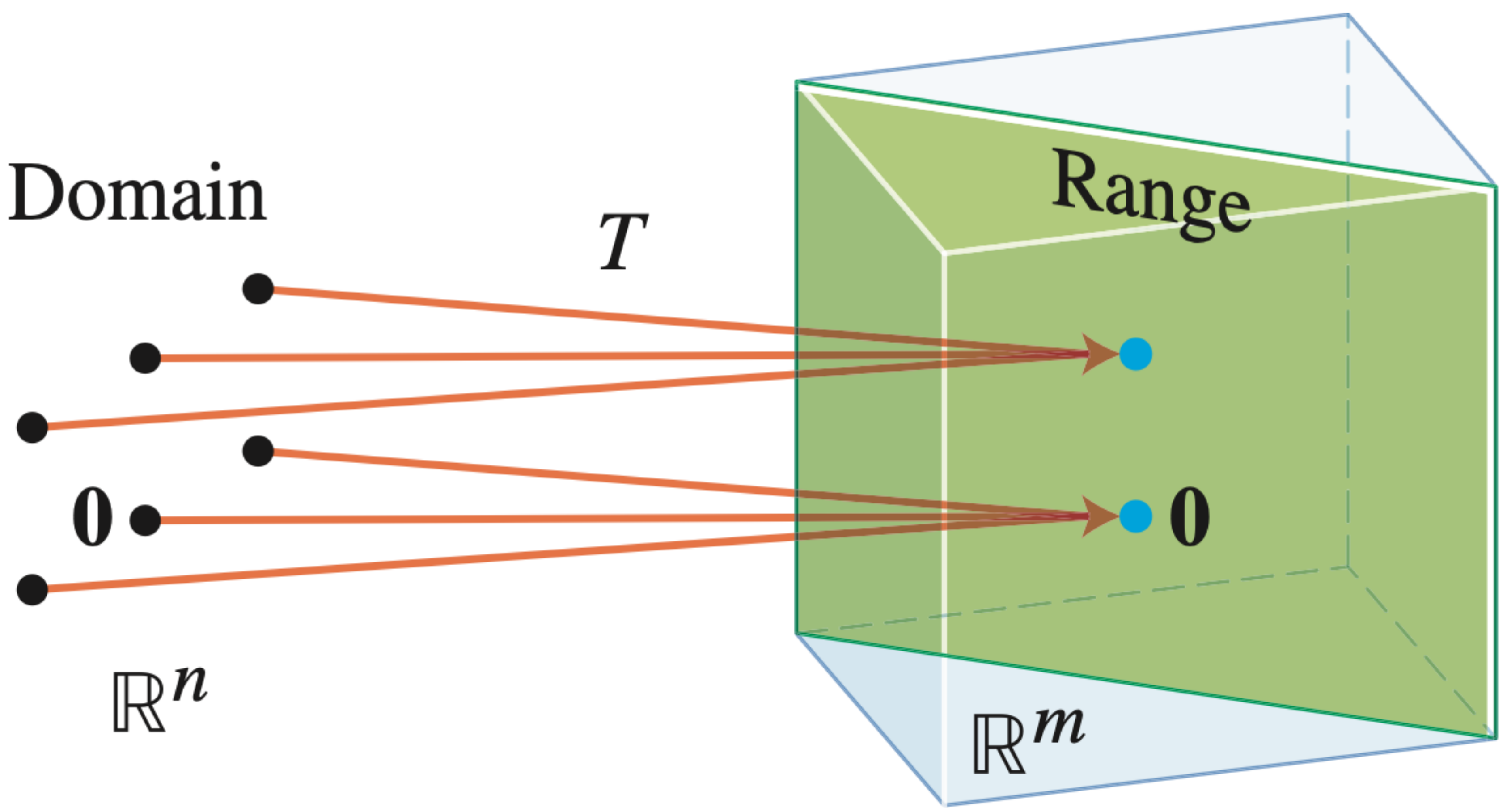
Recall: One-to-one Transformations

Recall: One-to-one Transformations

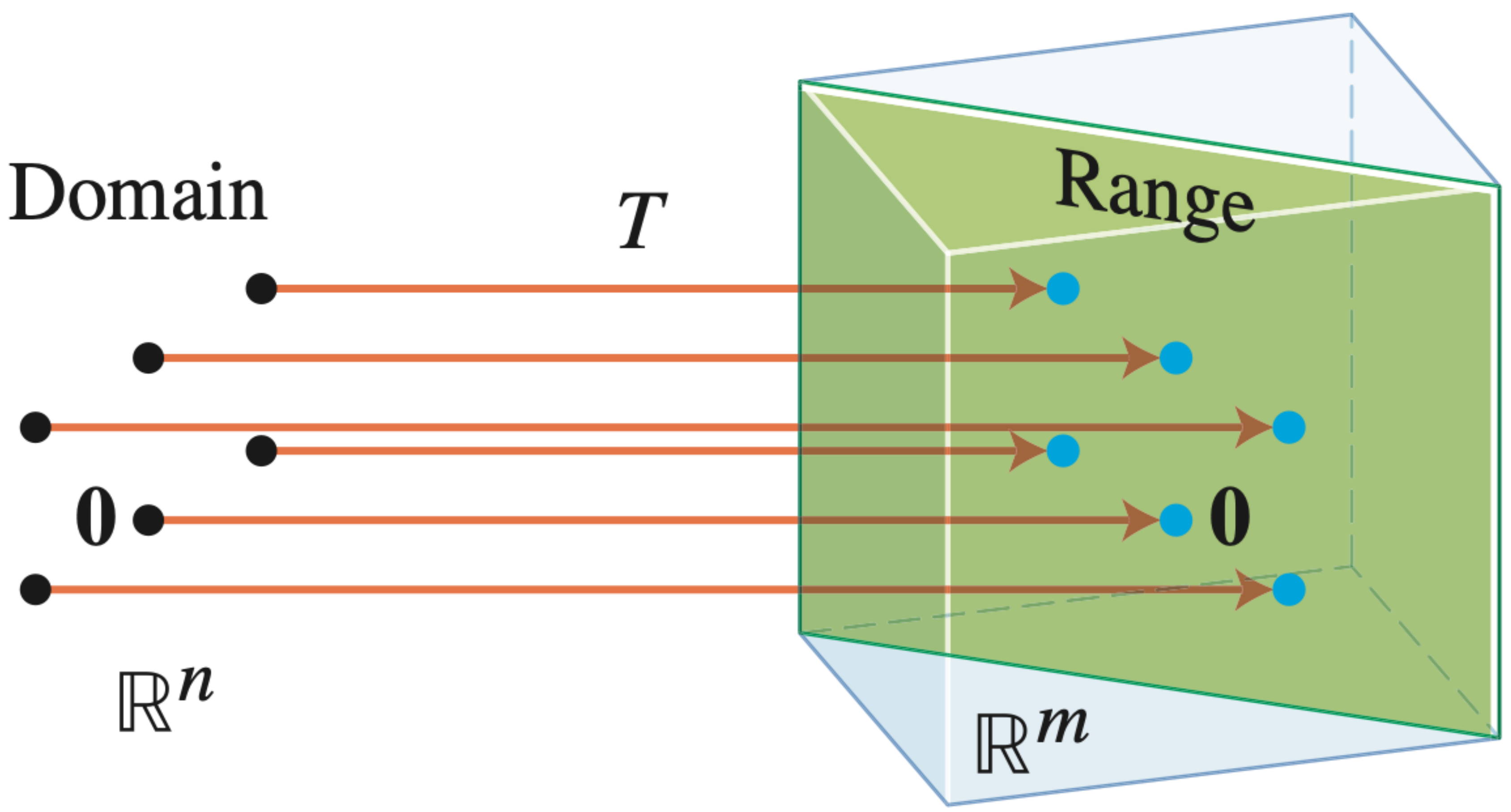
Definition. A transformation $T: \mathbb{R}^n \to \mathbb{R}^m$ is ***one-to-one*** if any vector $\mathbf{b}$ in $\mathbb{R}^m$ is the image of at most one vector $\mathbf{v}$ in $\mathbb{R}^n$ (where $T(\mathbf{v}) = \mathbf{b}$).

Recall: One-to-one Transformations

Definition. A transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ is ***one-to-one*** if any vector $\mathbf{b}$ in $\mathbb{R}^m$ is the image of at most one vector $\mathbf{v}$ in $\mathbb{R}^n$ (where $T(\mathbf{v}) = \mathbf{b}$).



T is *not* one-to-one



T is one-to-one

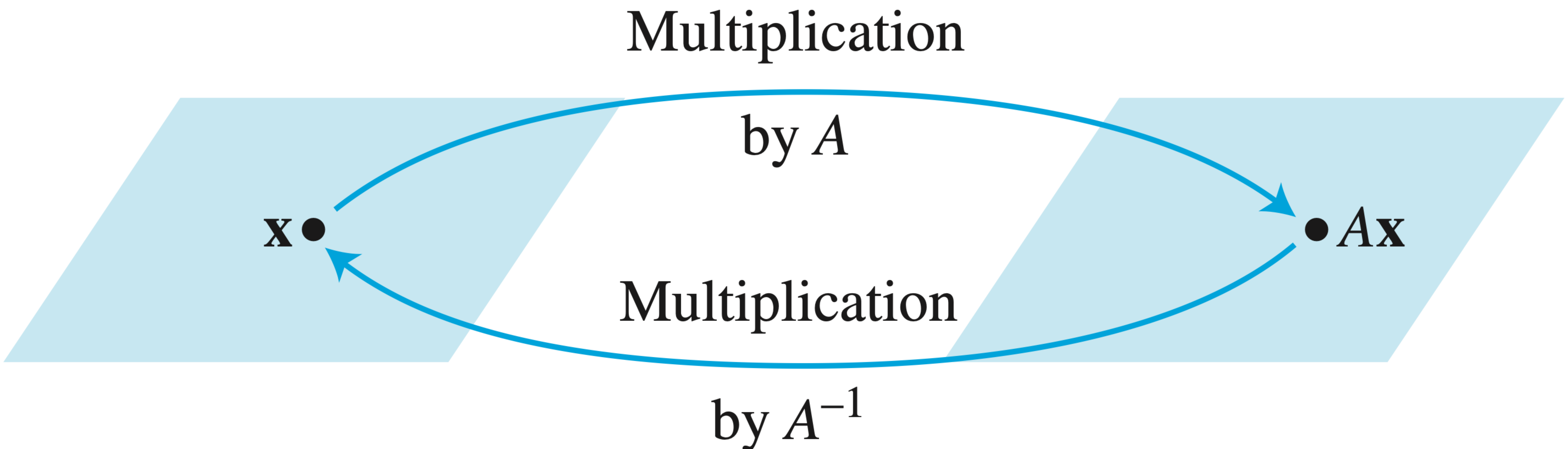
image source: Linear Algebra and its Applications. Lay, Lay, and McDonald

Connection to Transformations

Definition. A linear transformation $T: \mathbb{R}^n \to \mathbb{R}^n$ is **invertible** if there is a linear transformation S such that

$$S(T(\mathbf{v})) = \mathbf{v} \text{ and } T(S(\mathbf{v})) = \mathbf{v}$$

for any $\mathbf{v}$ in $\mathbb{R}^n$



Connection to Transformations

Connection to Transformations

Theorem. A $n \times n$ matrix A is invertible if and only if the matrix transformation $\mathbf{x} \mapsto A\mathbf{x}$ is invertible

Connection to Transformations

Theorem. A $n \times n$ matrix A is invertible if and only if the matrix transformation $\mathbf{x} \mapsto A\mathbf{x}$ is invertible

A matrix is invertible if it's possible to "undo" its transformation without "losing information"

Connection to Transformations

Theorem. A $n \times n$ matrix A is invertible if and only if the matrix transformation $\mathbf{x} \mapsto A\mathbf{x}$ is invertible

A matrix is invertible if it's possible to "undo" its transformation without "losing information"

Non-Example. Projection onto the x_1-axis

Connection to Transformations

Connection to Transformations

Definition. A transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ is a **one-to-one correspondence** (bijection) if any vector $\mathbf{b}$ in $\mathbb{R}^n$ is the image of **exactly** one vector $\mathbf{v}$ in $\mathbb{R}^n$ (where $T(\mathbf{v}) = \mathbf{b}$)

Connection to Transformations

Definition. A transformation $T: \mathbb{R}^n \to \mathbb{R}^n$ is a **one-to-one correspondence** (bijection) if any vector $\mathbf{b}$ in $\mathbb{R}^n$ is the image of **exactly** one vector $\mathbf{v}$ in $\mathbb{R}^n$ (where $T(\mathbf{v}) = \mathbf{b}$)

A transformation is a 1-1 correspondence if it is 1-1 and onto

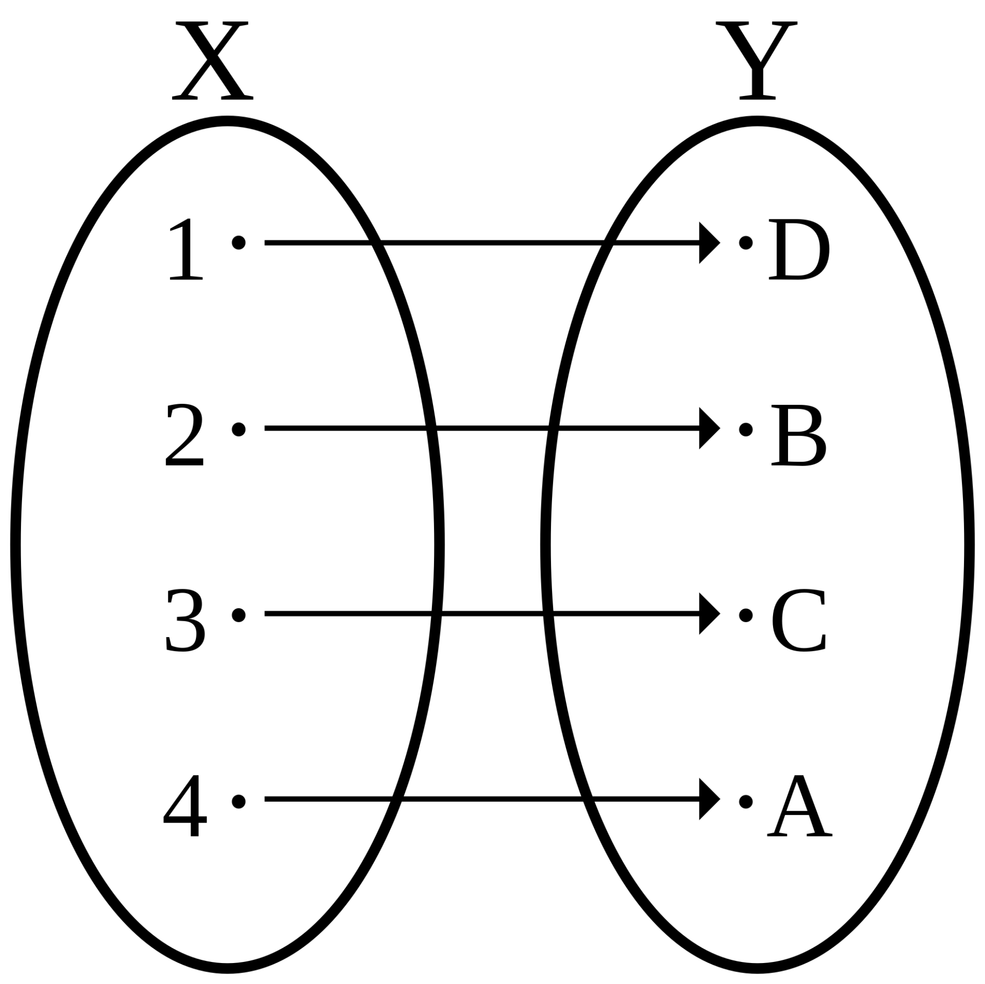
Connection to Transformations

Definition. A transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ is a **one-to-one correspondence** (bijection) if any vector $\mathbf{b}$ in $\mathbb{R}^n$ is the image of **exactly** one vector $\mathbf{v}$ in $\mathbb{R}^n$ (where $T(\mathbf{v}) = \mathbf{b}$)

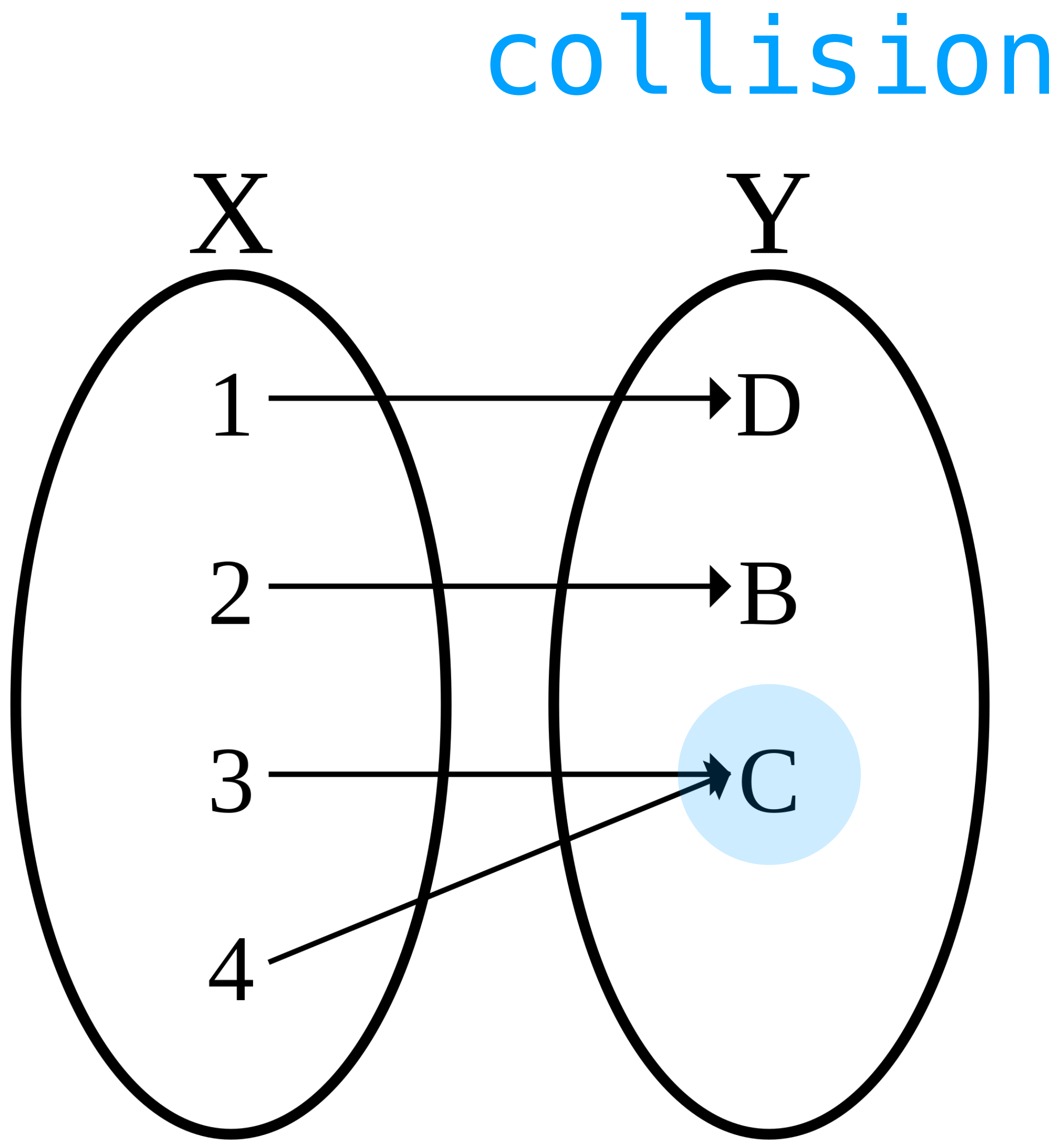
A transformation is a 1-1 correspondence if it is 1-1 and onto

Invertible transformations are 1-1 correspondences

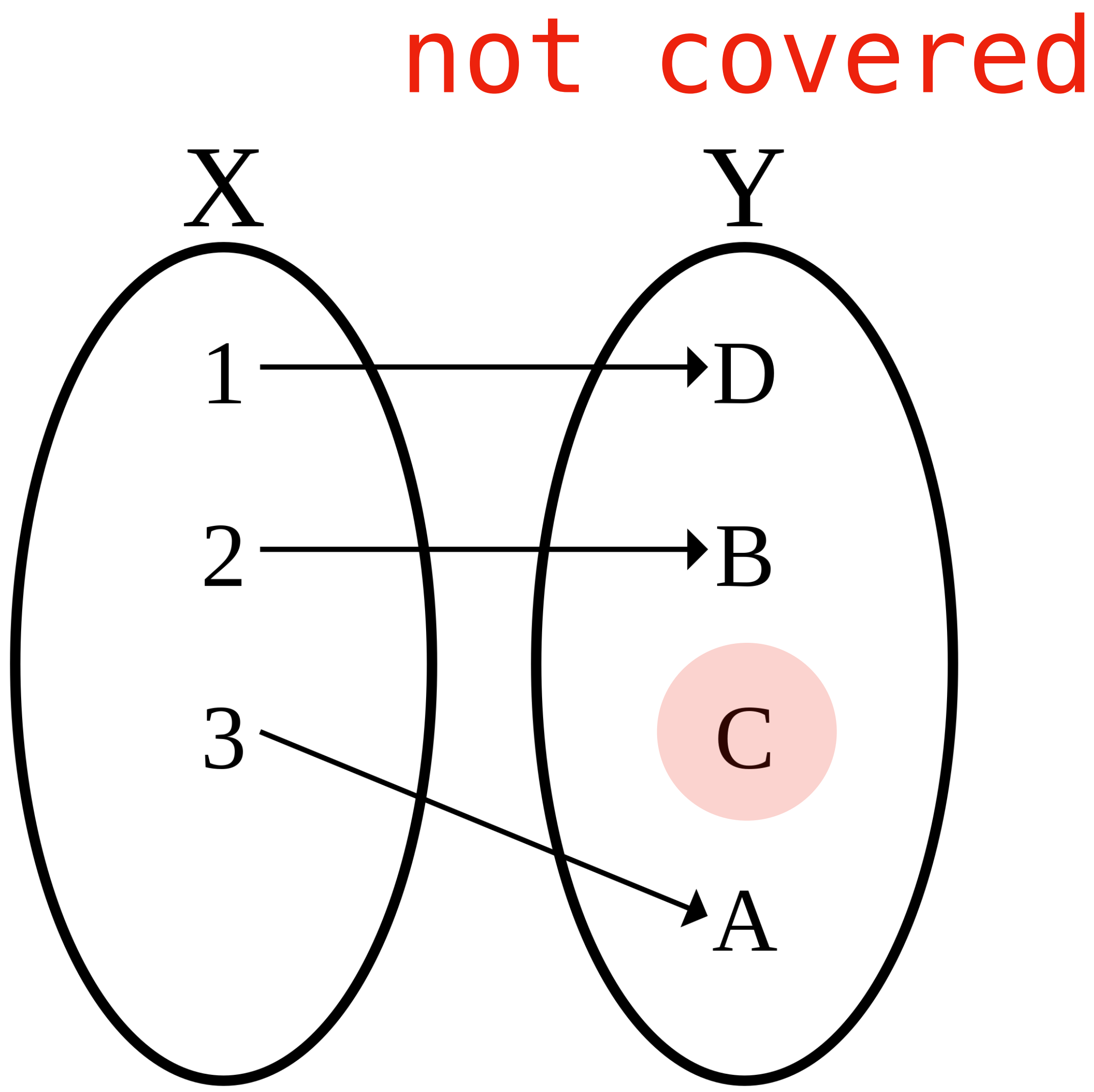
Kinds of Transformations (Pictorially)



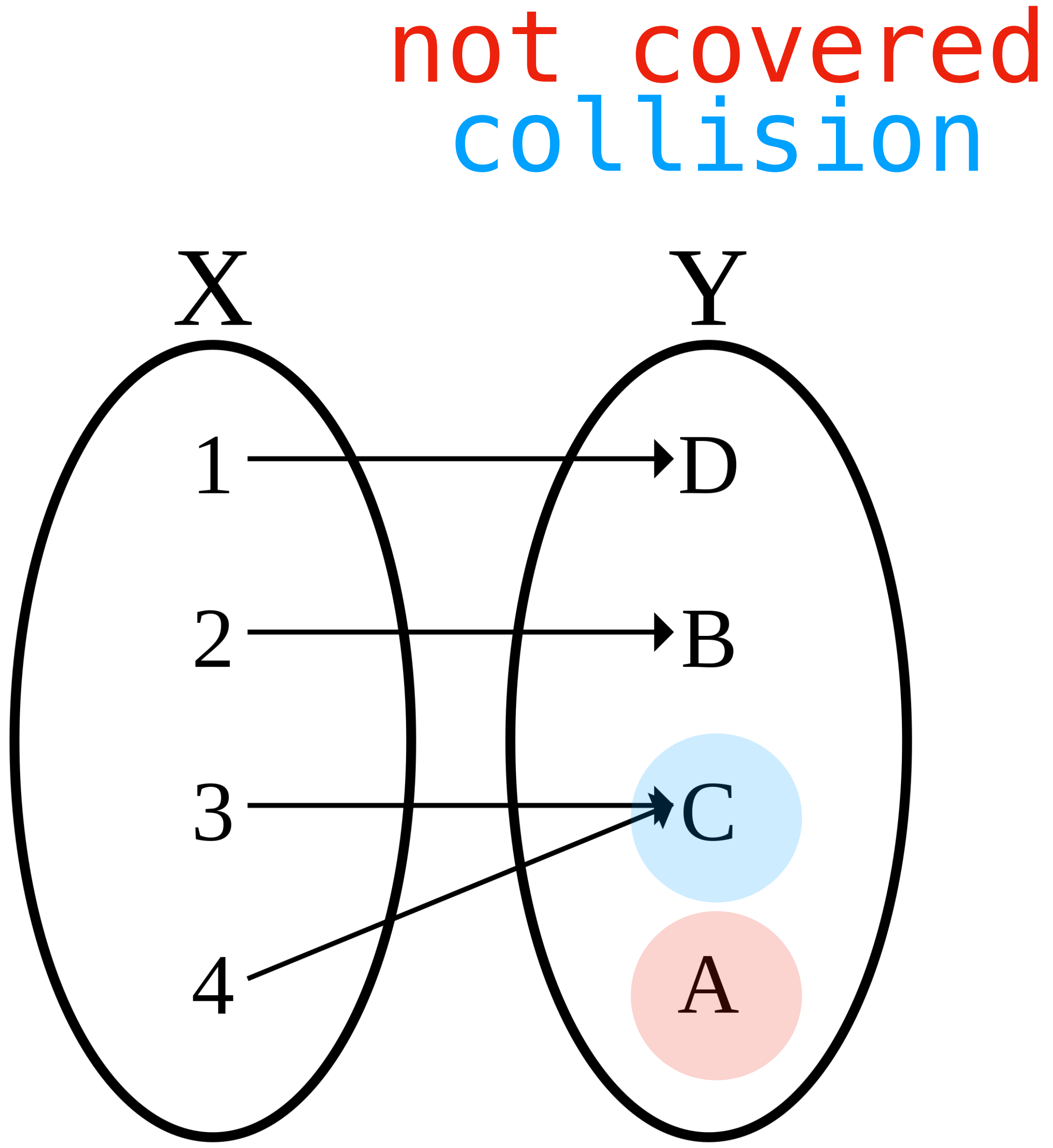
1–1 correspondence



onto, not 1–1



1–1 not onto



not 1–1, not onto

Computing Matrix Inverses

Fundamental Questions

How can we determine if a matrix has an inverse?

If a matrix has an inverse how do we compute it?

Fundamental Questions

Answer 1: Try to compute it

How can we determine if a matrix has an inverse?

If a matrix has an inverse how do we compute it?

Fundamental Questions

Answer 1: Try to compute it

How can we determine if a matrix has an inverse?

If a matrix has an inverse how do we compute it?

Answer 2: the Invertible Matrix Theorem (IMT)

In General

$$A \begin{bmatrix} \mathbf{b}_1 & \mathbf{b}_2 & \mathbf{b}_3 \end{bmatrix} = I$$

Can we solve for each $\mathbf{b}_i$?:

In General

$$A\mathbf{b}_1 = \mathbf{e}_1 \qquad A\mathbf{b}_2 = \mathbf{e}_2 \qquad A\mathbf{b}_3 = \mathbf{e}_3$$

If we want a matrix B such that $AB = I$, then the above equation must hold (in the case B has 3 columns)

In General

$$\mathbf{Ab}_1 = \mathbf{e}_1 \qquad \mathbf{Ab}_2 = \mathbf{e}_2 \qquad \mathbf{Ab}_3 = \mathbf{e}_3$$

If we want a matrix B such that $AB = I$, then the above equation must hold (in the case B has 3 columns)

We need to solve 3 matrix equations

How To: Matrix Inverses

Question. Find the inverse of an invertible $n \times n$ matrix A

Solution. Solve the equation $A\mathbf{x} = \mathbf{e}_i$ for every standard basis vector $\mathbf{e}_i$. Put those solutions $\mathbf{s}_1, \mathbf{s}_2, \ldots, \mathbf{s}_n$ into a single matrix

$$[\mathbf{s}_1 \quad \mathbf{s}_2 \quad \cdots \quad \mathbf{s}_n]$$

How To: Matrix Inverses

Question. Find the inverse of an invertible $n \times n$ matrix A

Solution. Row reduce the matrix $[A \quad I]$ to a matrix $[I \quad B]$. Then B is the inverse of A

This is really the same thing. It's a simultaneous reduction demo

Special Case: 2×2 Matrice Inverses

Special Case: 2×2 Matrice Inverses

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

Special Case: 2×2 Matrice Inverses

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

The **determinant** of a 2×2 matrix is the value $ad - bc$

Special Case: 2×2 Matrice Inverses

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

The **determinant** of a 2×2 matrix is the value $ad - bc$

The inverse is defined only if the determinant is nonzero

Special Case: 2×2 Matrice Inverses

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

The **determinant** of a 2×2 matrix is the value $ad - bc$

The inverse is defined only if the determinant is nonzero

(see the notes on linear transformations for more information about determinants)

Example

$$\begin{bmatrix} -6 & 14 \\ 3 & -7 \end{bmatrix}$$

Example

$$\begin{bmatrix} -6 & 14 \\ 3 & -7 \end{bmatrix}$$

Is the above matrix invertible?

Example

$$\begin{bmatrix} -6 & 14 \\ 3 & -7 \end{bmatrix}$$

Is the above matrix invertible?

No. The determinant is $(-6)(-7) - 14(3) = 42 - 42 = 0$

Algebra of Matrix Inverses

Algebraic Properties (Matrix Inverses)

Theorem. For a $n \times n$ invertible matrix A, the matrix A^{-1} is invertible and

$$(A^{-1})^{-1} = A$$

Verify:

Algebraic Properties (Matrix Inverses)

Theorem. For a $n \times n$ invertible matrix A, the matrix A^T is invertible and

$$(A^T)^{-1} = (A^{-1})^T$$

Verify:

Algebraic Properties (Matrix Inverses)

Theorem. For a $n \times n$ invertible matrices A and B, the matrix AB is invertible and

$$(AB)^{-1} = B^{-1}A^{-1}$$

Verify:

Question

Suppose that A *is a* $n \times n$ *invertible matrix such that* $A = A^T$ *and* B *is a* $m \times n$ *matrix*

Simplify the expression $A(BA^{-1})^T$ *using the algebraic properties we've seen*

Answer: B^T

$$A(BA^{-1})^T$$

$$A = A^T$$

Invertible Matrix Theorem

Motivation

Question. How do we know if a square matrix is invertible?

Answer. *Every* perspective we've taken so far can help us answer this question

Invertible Matrix Theorem

Theorem. Suppose A is a $n \times n$ invertible matrix. Then the following hold.

1. A^T is invertible

Invertible Matrix Theorem

Theorem. Suppose A is a $n \times n$ invertible matrix. Then the following hold.

2. $A\mathbf{x} = \mathbf{b}$ has at <u>least</u> one solution for every $\mathbf{b}$
3. $A\mathbf{x} = \mathbf{b}$ has at <u>most</u> one solution for every $\mathbf{b}$
4. $A\mathbf{x} = \mathbf{b}$ has at <u>exactly</u> one solution for every $\mathbf{b}$

Invertible Matrix Theorem

Theorem. Suppose A is a $n \times n$ invertible matrix. Then the following hold.

5. A has a pivot in every <u>column</u>
6. A has a pivot in every <u>row</u>
7. A is row equivalent to I_n

Invertible Matrix Theorem

Theorem. Suppose A is a $n \times n$ invertible matrix. Then the following hold.

8. $A\mathbf{x} = \mathbf{0}$ has only the trivial solution
9. The columns of A are linearly independent
10. The columns of A span $\mathbb{R}^n$

Invertible Matrix Theorem

Theorem. Suppose A is a $n \times n$ invertible matrix. Then the following hold.

11. The linear transformation $\mathbf{x} \mapsto A\mathbf{x}$ is onto
12. $\mathbf{x} \mapsto A\mathbf{x}$ is one-to-one
13. $\mathbf{x} \mapsto A\mathbf{x}$ is a one-to-one correspondence
14. $\mathbf{x} \mapsto A\mathbf{x}$ is invertible

Taking Stock: IMT

1. A is invertible
2. A^T is invertible
3. $A\mathbf{x} = \mathbf{b}$ has at least one solution for any $\mathbf{b}$
4. $A\mathbf{x} = \mathbf{b}$ has at most one solution for any $\mathbf{b}$
5. $A\mathbf{x} = \mathbf{b}$ has a unique solution for any $\mathbf{b}$
6. A has n pivots (per row and per column)
7. A is row equivalent to I
8. $A\mathbf{x} = \mathbf{0}$ has only the trivial solution
9. The columns of A are linearly independent
10. The columns of A span $\mathbb{R}^n$
11. The linear transformation $\mathbf{x} \mapsto A\mathbf{x}$ is onto
12. $\mathbf{x} \mapsto A\mathbf{x}$ is one-to-one
13. $\mathbf{x} \mapsto A\mathbf{x}$ is a one-to-one correspondence
14. $\mathbf{x} \mapsto A\mathbf{x}$ is invertible

These all express the **same thing**

(this is a stronger statement than we just verified)

Taking Stock: IMT

1. A is invertible
2. A^T is invertible
3. $A\mathbf{x} = \mathbf{b}$ has at least one solution for any $\mathbf{b}$
4. $A\mathbf{x} = \mathbf{b}$ has at most one solution for any $\mathbf{b}$
5. $A\mathbf{x} = \mathbf{b}$ has a unique solution for any $\mathbf{b}$
6. A has n pivots (per row and per column)
7. A is row equivalent to I
8. $A\mathbf{x} = \mathbf{0}$ has only the trivial solution
9. The columns of A are linearly independent
10. The columns of A span $\mathbb{R}^n$
11. The linear transformation $\mathbf{x} \mapsto A\mathbf{x}$ is onto
12. $\mathbf{x} \mapsto A\mathbf{x}$ is one-to-one
13. $\mathbf{x} \mapsto A\mathbf{x}$ is a one-to-one correspondence
14. $\mathbf{x} \mapsto A\mathbf{x}$ is invertible

These all express the **same thing**

(this is a stronger statement than we just verified)

!! only for square matrices !!

We get a lot of information for free

We get a lot of information for free

Theorem. If A is square, then

A **is 1–1** if and only if A **is onto**

We get a lot of information for free

Theorem. If A is square, then

A **is 1–1** if and only if A **is onto**

We only need to check one of these.

We get a lot of information for free

Theorem. If A is square, then

A **is 1–1** if and only if A **is onto**

We only need to check one of these.

Warning. Remember this only applies square matrices.

We get a lot of information for free

We get a lot of information for free

Theorem. If A is square, then

A **is invertible** $\equiv$ $A\mathbf{x} = \mathbf{0}$ **implies** $\mathbf{x} = \mathbf{0}$

We get a lot of information for free

Theorem. If A is square, then

A **is invertible** $\quad \equiv \quad$ $A\mathbf{x} = \mathbf{0}$ **implies** $\mathbf{x} = \mathbf{0}$

Invertibility is completely determined by how A behaves on $\mathbf{0}$.

Question (Conceptual)

True *or* ***False:*** *If A is invertible, and B is row equivalent to A (we can transform B into A by a sequence of row operations), the B is also invertible.*

Answer: True

Row reductions don't change the number of pivots.

Question

If $[\mathbf{a}_1 \quad \mathbf{a}_2 \quad \mathbf{a}_3]$ *is invertible, then is* $\left[(\mathbf{a}_1 + \mathbf{a}_2 - 2\mathbf{a}_3) \quad (\mathbf{a}_2 + 5\mathbf{a}_3) \quad \mathbf{a}_3\right]$ *also invertible? Justify your answer.*

Answer

Consider $[\mathbf{a}_1 \quad \mathbf{a}_2 \quad \mathbf{a}_3]^T$. We can get to $\left[(\mathbf{a}_1 + \mathbf{a}_2 - 2\mathbf{a}_3) \quad (\mathbf{a}_2 + 5\mathbf{a}_3) \quad \mathbf{a}_3\right]^T$ by row operations

LU Factorization

Matrix Factorization

Matrix Factorization

A **factorization** of a matrix A is an equation which expresses A as a product of one or more matrices, e.g.,

$$A = BC$$

Matrix Factorization

A **factorization** of a matrix A is an equation which expresses A as a product of one or more matrices, e.g.,

$$A = BC$$

So far, we've been given two factors and asked to find their product

Factorization is the harder direction

Reasons to Factorize

Reasons to Factorize

Writing A as the product of multiple matrices can

Reasons to Factorize

Writing A as the product of multiple matrices can

» make computing with A faster

Reasons to Factorize

Writing A as the product of multiple matrices can

» make computing with A faster

» make working with A easier

Reasons to Factorize

Writing A as the product of multiple matrices can

- make computing with A faster
- make working with A easier
- expose important information about A

Reasons to Factorize

Writing A as the product of multiple matrices can

» make computing with A faster LU Decomposition

» make working with A easier

» expose important information about A

The Problem

Question. For an matrix A, solve the equations

$$A\mathbf{x} = \mathbf{b}_1 \quad , \quad A\mathbf{x} = \mathbf{b}_2 \quad \ldots \quad A\mathbf{x} = \mathbf{b}_{k-1} \quad , \quad A\mathbf{x} = \mathbf{b}_k$$

In other words: we want to solve <u>a bunch</u> of matrix equations over the same matrix

The Problem

Question. For a matrix A, solve (for X) in the equation

$$AX = B$$

where X and B are matrices of appropriate dimension

This is (essentially) the same question

The Problem

Question. Solve $AX = B$

If A is *invertible*, then we have a solution:

Find A^{-1} and then $X = A^{-1}B$

The Problem

Question. Solve $AX = B$

If A is *invertible*, then we have a solution:

Find A^{-1} and then $X = A^{-1}B$

What if A^{-1} is not invertible?
Even if it is, can we do it faster?

LU Factorization at a High Level

Given a $m \times n$ matrix A, we are going to factorize A as

echelon form of A

$$A = \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ * & 1 & 0 & 0 \\ * & * & 1 & 0 \\ * & * & * & 1 \end{bmatrix}}_{L} \underbrace{\begin{bmatrix} \blacksquare & * & * & * & * \\ 0 & \blacksquare & * & * & * \\ 0 & 0 & 0 & \blacksquare & * \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}}_{U}$$

LU Factorization at a High Level

Given a $m \times n$ matrix A, we are going to factorize A as

echelon form of A

$$A = \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ * & 1 & 0 & 0 \\ * & * & 1 & 0 \\ * & * & * & 1 \end{bmatrix}}_{L} \underbrace{\begin{bmatrix} \blacksquare & * & * & * & * \\ 0 & \blacksquare & * & * & * \\ 0 & 0 & 0 & \blacksquare & * \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}}_{U}$$

Note. This applies to non-square matrices

What are "L" and "U"?

L stands for "lower" as in *lower triangular*

U stands for "upper" as in *upper triangular*

$$A = \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ * & 1 & 0 & 0 \\ * & * & 1 & 0 \\ * & * & * & 1 \end{bmatrix}}_{L} \underbrace{\begin{bmatrix} \blacksquare & * & * & * \\ 0 & \blacksquare & * & * \\ 0 & 0 & \blacksquare & * \\ 0 & 0 & 0 & \blacksquare \end{bmatrix}}_{U}$$

The Fundamental Question

$$A = L\boxed{U}$$ echelon form of A

The Fundamental Question

$$A = L\boxed{U}$$

echelon form of A

We know how to build U, that's just the forward phase of Gaussian elimination

The Fundamental Question

$$A = L\boxed{U}$$ echelon form of A

We know how to build U, that's just the forward phase of Gaussian elimination

How do we build L?

The Fundamental Question

$$A = L\boxed{U}$$ echelon form of A

We know how to build U, that's just the forward phase of Gaussian elimination

How do we build L?

The idea. L "implements" the row operations of the forward phase

Elementary Matrices

Recall: Elementary Row Operations

scaling	multiply a row by a number
interchange	switch two rows
replacement	add a scaled equation to another

The First Key Observation

The First Key Observation

Elementary row operations are **linear transformations (viewed as transformation on columns)**

The First Key Observation

Elementary row operations are **linear transformations (viewed as transformation on columns)**

Example: Scale row 2 by 5

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \xrightarrow{R_2 \leftarrow 5R_2} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 5a_{21} & 5a_{22} & 5a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

Example: Scaling

Restricted to one column, we see this is the above linear transformation

$$\begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} \mapsto \begin{bmatrix} v_1 \\ 5v_2 \\ v_3 \end{bmatrix}$$

Example: Scaling

$$\begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} \mapsto \begin{bmatrix} v_1 \\ 5v_2 \\ v_3 \end{bmatrix}$$

Let's build the matrix which implements it:

Another Example: Scaling + Replacement

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \longrightarrow \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ (a_{31} - 2a_{11}) & (a_{32} - 2a_{12}) & (a_{33} - 2a_{13}) \end{bmatrix}$$

$$R_3 \leftarrow (R_3 - 2R_1)$$

Another Example: Scaling + Replacement

$$R_3 \leftarrow (R_3 - 2R_1)$$

Elementary row operations are linear, so they are implemented by matrices

General Elementary Scaling Matrix

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & k & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

General Elementary Scaling Matrix

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & k & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

If we want to perform $R_3 \leftarrow kR_3$ then we need the identity matrix but with the entry $A_{33} = k$.

General Elementary Scaling Matrix

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & k & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

If we want to perform $R_3 \leftarrow kR_3$ then we need the identity matrix but with the entry $A_{33} = k$.

If we want to perform $R_i \leftarrow kR_i$ then we need the identity matrix but with then entry $A_{ii} = k$.

General Replacement Matrix

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ k & 0 & 0 & 1 \end{bmatrix}$$

General Replacement Matrix

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ k & 0 & 0 & 1 \end{bmatrix}$$

If we want to perform $R_4 \leftarrow R_4 + kR_1$, then we need the identity matrix but with the entry $A_{41} = k$.

General Replacement Matrix

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ k & 0 & 0 & 1 \end{bmatrix}$$

If we want to perform $R_4 \leftarrow R_4 + kR_1$, then we need the identity matrix but with the entry $A_{41} = k$.

If we want to perform $R_i \leftarrow R_i + kR_j$, then we need the identity matrix but with the entry $A_{ij} = k$.

General Swap Matrix

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

If we want to swap R_2 and R_3, then we need the identity matrix, but with R_2 and R_3 swapped.

Elementary Matrices

Definition. An **elementary matrix** is a matrix obtained by applying a single row operation to the identity matrix I.

Example.

How To: Finding Elementary Matrices

Question. Find the matrix implementing the elementary row operation `op`

Solution. Apply `op` to the identity matrix of the appropriate size

Products of Elementary Matrices

Products of Elementary Matrices

Taking stock:

Products of Elementary Matrices

Taking stock:

» Elementary matrices implement elementary row operations

Products of Elementary Matrices

<u>Taking stock:</u>

» Elementary matrices implement elementary row operations

» Remember that Matrix multiplication is transformation composition (i.e., do one then the other)

Products of Elementary Matrices

Taking stock:

» Elementary matrices implement elementary row operations

» Remember that Matrix multiplication is transformation composition (i.e., do one then the other)

So we can implement any sequence of row operations as a product of elementary matrices

How to: Matrices implementing Row Operations

Question. Find the matrix implementing a sequence of row operations op_1, op_2,...

Solution. Apply the row operations in sequence to the identity matrix of the appropriate size

Question

Find the matrix implementing the following sequence of elementary row operations on a $3 \times n$ *matrix.*

$$R_2 \leftarrow 3R_2$$

$$R_1 \leftarrow R_1 + R_2$$

$$R_2 \leftrightarrow R_3$$

Then multiply it with the all-ones 3×3 *matrix.*

Answer

$$\begin{bmatrix} 1 & 3 & 0 \\ 0 & 0 & 1 \\ 0 & 3 & 0 \end{bmatrix}$$

Second Key Observation

Second Key Observation

Elementary row operations are **invertible** linear transformations

Second Key Observation

Elementary row operations are **invertible** linear transformations

This also means the product of elementary matrices is invertible

$$(E_1E_2E_3E_4)^{-1} = E_4^{-1}E_3^{-1}E_2^{-1}E_1^{-1}$$

!! the order reverses !!

Question (Conceptual)

Describe the inverse transformation for each elementary row operation

Question (Conceptual)

Describe the inverse transformation for each elementary row operation

The inverse of scaling by k is scaling by $1/k$

Question (Conceptual)

Describe the inverse transformation for each elementary row operation

The inverse of scaling by k is scaling by $1/k$

The inverse of $R_i \leftarrow R_i + kR_j$ is $R_i \leftarrow R_i - kR_j$

Question (Conceptual)

Describe the inverse transformation for each elementary row operation

The inverse of scaling by k is scaling by $1/k$

The inverse of $R_i \leftarrow R_i + kR_j$ is $R_i \leftarrow R_i - kR_j$

The inverse of swapping is swapping again

Recall: Elementary Row Operations

scaling	multiply a row by a number
interchange	switch two rows
replacement	add a scaled equation to another

Recall: Elementary Row Operations

We only need these two for the forward phase

interchange switch two rows

replacement add a scaled equation to another

Recall: Elementary Row Operations

We'll assume we only need this

replacement add a scaled equation to another

Reminder: LU Factorization at a High Level

Given a $m \times n$ matrix A, we are going to factorize A as

Echelon form of A

$$A = \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ * & 1 & 0 & 0 \\ * & * & 1 & 0 \\ * & * & * & 1 \end{bmatrix}}_{L} \underbrace{\begin{bmatrix} \blacksquare & * & * & * & * \\ 0 & \blacksquare & * & * & * \\ 0 & 0 & 0 & \blacksquare & * \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}}_{U}$$

Gaussian Elimination and Elementary Matrices

$$A \sim A_1 \sim A_2 \sim \ldots \sim A_k$$

Consider a sequence of elementary row operations from A to an echelon form

Each step can be represent as a **product with an elementary matrix**

Gaussian Elimination and Elementary Matrices

$$A \sim E_1 A \sim E_2 E_1 A \sim \ldots \sim E_k E_{k-1} \ldots E_2 E_1 A$$

Gaussian Elimination and Elementary Matrices

$$A \sim E_1A \sim E_2E_1A \sim \ldots \sim E_kE_{k-1}\ldots E_2E_1A$$

This exactly tells us that if B is the final echelon form we get then

$B = (E_kE_{k-1}\ldots E_2E_1)A = EA$

where E implements a <u>sequence</u> of row operations. So:

Gaussian Elimination and Elementary Matrices

$$A \sim E_1 A \sim E_2 E_1 A \sim \ldots \sim E_k E_{k-1} \ldots E_2 E_1 A$$

This exactly tells us that if B is the final echelon form we get then

Invertible

$$B = (E_k E_{k-1} \ldots E_2 E_1) A = EA$$

where E implements a <u>sequence</u> of row operations. So:

Gaussian Elimination and Elementary Matrices

$$A \sim E_1A \sim E_2E_1A \sim \ldots \sim E_kE_{k-1}\ldots E_2E_1A$$

This exactly tells us that if B is the final echelon form we get then

$$B = \underbrace{(E_kE_{k-1}\ldots E_2E_1)}_{\text{Invertible}}A = EA$$

where E implements a <u>sequence</u> of row operations. So:

$$A = E^{-1}B = (E_1^{-1}E_2^{-1}\ldots E_{k-1}^{-1}E_k^{-1})B$$

LU Factorization Algorithm

LU Factorization Algorithm

```
FUNCTION LU_Factorization(A):
```

LU Factorization Algorithm

```
FUNCTION LU_Factorization(A):
    L ← identity matrix
```

LU Factorization Algorithm

```
FUNCTION LU_Factorization(A):
    L ← identity matrix
    U ← A
```

LU Factorization Algorithm

```
FUNCTION LU_Factorization(A):
    L ← identity matrix
    U ← A
    convert U to an echelon form by GE forward step # without swaps
```

LU Factorization Algorithm

```
FUNCTION LU_Factorization(A):
    L ← identity matrix
    U ← A
    convert U to an echelon form by GE forward step # without swaps
    FOR each row operation OP in the prev step:
```

LU Factorization Algorithm

```
FUNCTION LU_Factorization(A):
    L ← identity matrix
    U ← A
    convert U to an echelon form by GE forward step # without swaps
    FOR each row operation OP in the prev step:
        E ← the matrix implementing OP
```

LU Factorization Algorithm

```
FUNCTION LU_Factorization(A):
    L ← identity matrix
    U ← A
    convert U to an echelon form by GE forward step # without swaps
    FOR each row operation OP in the prev step:
        E ← the matrix implementing OP
        L ← L @ E^{-1}    # note the multiplication on the right
```

LU Factorization Algorithm

```
FUNCTION LU_Factorization(A):
    L ← identity matrix
    U ← A
    convert U to an echelon form by GE forward step # without swaps
    FOR each row operation OP in the prev step:
        E ← the matrix implementing OP
        L ← L @ E⁻¹    # note the multiplication on the right
    RETURN (L, U)
```

LU Factorization Algorithm

```
FUNCTION LU_Factorization(A):
    L ← identity matrix
    U ← A
    convert U to an echelon form by GE forward step # without swaps
    FOR each row operation OP in the prev step:
        E ← the matrix implementing OP
        L ← L @ E^{-1}    # note the multiplication on the right
    RETURN (L, U)
```

we'll see how to do this more efficiently

The forward part of Gaussian elimination <u>is</u> matrix factorization

The "L" Part

$$E = E_k E_{k-1} \cdots E_2 E_1$$

This a product of elementary matrices

So $L = E^{-1} = E_1^{-1} E_2^{-1} \ldots E_{k-1}^{-1} E_k^{-1}$ **!! the order reverses !!**

We won't prove this, but it's worth thinking about: **why is this lower triangular?**

And can we build this in a more efficient way?

demo

How To: LU Factorization by hand

Question. Find a LU Factorization for the matrix A (assuming no swaps)

Solution.

» Start with L as the identity matrix
» Find U by the forward part of GE
» For each operation $R_i \leftarrow R_i + kR_j$, set L_{ij} to $-k$

Analyzing Linear Algebra Algorithms

Analyzing Linear Algebra Algorithms

We will not use $O(\cdot)$ notation!

Analyzing Linear Algebra Algorithms

We will not use $O(\cdot)$ notation!

For numerics, we care about number of **FL**oating-oint **OP**erations (FLOPs):

>> addition
>> subtraction
>> multiplication
>> division
>> square root

Analyzing Linear Algebra Algorithms

We will not use $O(\cdot)$ notation!

For numerics, we care about number of **FL**oating-oint **OP**erations (FLOPs):

>> addition
>> subtraction
>> multiplication
>> division
>> square root

$2n$ vs. n is very different when $n \sim 10^{20}$

Analyzing LU Factorization

Dominant Terms

Dominant Terms

that said, we don't care about *exact* bounds

Dominant Terms

that said, we don't care about *exact* bounds

A function $f(n)$ is ***asymptotically equivalent*** to $g(n)$ if

$$\lim_{i \to \infty} \frac{f(i)}{g(i)} = 1$$

Dominant Terms

that said, we don't care about *exact* bounds

A function $f(n)$ is ***asymptotically equivalent*** to $g(n)$ if

$$\lim_{i \to \infty} \frac{f(i)}{g(i)} = 1$$

for polynomials, they are equivalent to their dominant term

Dominant Terms

the dominant term of a polynomial is the monomial with the highest degree

$$\lim_{i \to \infty} \frac{3x^3 + 100000x^2}{3x^3} = 1$$

$3x^3$ dominates the function even though the coefficient for x^2 is so large

How To: Solving systems with the LU

Question. Solve the equation $A\mathbf{x} = \mathbf{b}$ given that $A = LU$ is a LU factorization.

Solution. First solve $L\mathbf{x} = \mathbf{b}$ to get a solution $\mathbf{c}$, then solve $U\mathbf{x} = \mathbf{c}$ to get a solution $\mathbf{d}$.

Verify:

How To: Solving systems with the LU

Question. Solve the equation $A\mathbf{x} = \mathbf{b}$ given that $A = LU$ is a LU factorization.

Solution. First solve $L\mathbf{x} = \mathbf{b}$ to get a solution $\mathbf{c}$, then solve $U\mathbf{x} = \mathbf{c}$ to get a solution $\mathbf{d}$.

Why is this better than just solving $A\mathbf{x} = \mathbf{b}$?

FLOPs for Solving General Systems

The following FLOP estimates are based on $n \times n$ matrices

Gaussian Elimination: $\sim \frac{2n^3}{3}$ FLOPS

GE Forward: $\sim \frac{2n^3}{3}$ FLOPS

GE Backward: $\sim 2n^2$ FLOPS

Matrix Inversion: $\sim 2n^3$ FLOPS

Matrix–Vector Multiplication: $\sim 2n^2$ FLOPS

Solving by matrix inversion: $\sim 2n^3$ FLOPS

Solving by Gaussian elimination: $\sim \frac{2n^3}{3}$ FLOPS

FLOPS for solving LU systems

LU Factorization: $\sim \dfrac{2n^3}{3}$ FLOPS

Solving $L\mathbf{x} = \mathbf{b}$: $\sim 2n^2$ FLOPS (by "forward" elimination)

Solving $U\mathbf{x} = \mathbf{c}$: $\sim 2n^2$ FLOPS (already in echelon form)

Solving by LU Factorization: $\sim \dfrac{2n^3}{3}$ FLOPS

If you solve several matrix equations for the same matrix, LU factorization is faster than matrix inversion on the first equation, and the same (asymptotically) in later equations (and it works for rectangular matrices).

Other Considerations: Density

Other Considerations: Density

If A doesn't have to many entries (A is **sparse)**, then its likely that L and U won't either.

Other Considerations: Density

If A doesn't have to many entries (A is **sparse)**, then its likely that L and U won't either.

But A^{-1} may have ***many*** entries (A^{-1} is **dense**)

Other Considerations: Density

If A doesn't have to many entries (A is **sparse**), then its likely that L and U won't either.

But A^{-1} may have ***many*** entries (A^{-1} is **dense**)

Sparse matrices are faster to compute with and better with respect to storage.

Summary

Matrix inverses allow us to easily solve many matrixes equations over the same A

LU Factorizations allows us to do the same, but more generally more efficiently