

Linear Algebra

Review

Ngoc Hoang Luong

Faculty of Computer Science
University of Information Technology (UIT)
Vietnam National University - Ho Chi Minh City (VNU-HCM)

Maths for Computer Science, Fall 2021

Machine Learning and Linear Algebra



¹<https://xkcd.com/1838/>

The contents of this document are taken mainly from the following sources:

- ▶ Gilbert Strang. Linear Algebra and Learning from Data.
<https://math.mit.edu/~gs/learningfromdata/>
- ▶ Gilbert Strang. Introduction to Linear Algebra.
<http://math.mit.edu/~gs/linearalgebra/>
- ▶ Gilbert Strang. Linear Algebra for Everyone.
<http://math.mit.edu/~gs/everyone/>

Table of Contents

- 1 Matrix-Vector Multiplication Ax
- 2 Matrix-Matrix Multiplication AB
- 3 The Four Fundamental Subspaces of A : $\mathbf{C}(A)$, $\mathbf{C}(A^T)$, $\mathbf{N}(A)$, $\mathbf{N}(A^T)$
- 4 Elimination and $A = LU$
- 5 Orthogonal Matrices, Subspaces, and Projections

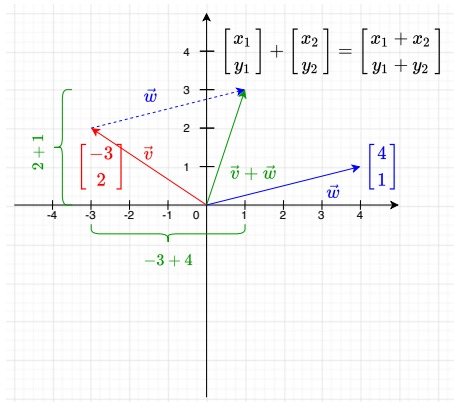
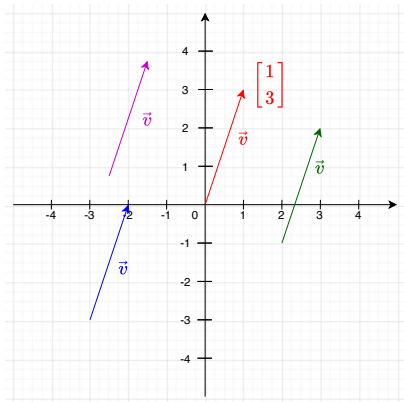
Table of Contents

- 1 Matrix-Vector Multiplication Ax
- 2 Matrix-Matrix Multiplication AB
- 3 The Four Fundamental Subspaces of A : $\mathbf{C}(A)$, $\mathbf{C}(A^\top)$, $\mathbf{N}(A)$, $\mathbf{N}(A^\top)$
- 4 Elimination and $A = LU$
- 5 Orthogonal Matrices, Subspaces, and Projections

- ▶ Vectors are arrays of numerical values.
- ▶ Each numerical value is referred to as *coordinate*, *component*, *entry*, or *dimension*.
- ▶ The number of components is the vector *dimensionality*.
- ▶ e.g., a vector representation of a person: 25 years old (Age), making 30 dollars an hour (Salary), having 5 years of experience (Experience): $[25, 30, 6]$.
- ▶ Vectors are special objects that can be added together and multiplied by scalars to produce another object of the same kind.

- ▶ **Geometric vectors** are often visualized as a quantity that has a **magnitude** as well as a **direction**.
- ▶ e.g., the velocity of a person moving at 1 meter/second in the eastern direction and 3 meters/second in the northern direction can be described as a directed line from the origin to $(1, 3)$.
- ▶ The **tail** of the vector is at the origin. The **head** is at $(1, 3)$.
- ▶ Geometric vectors can have arbitrary tails.
- ▶ Two geometric vectors can be added, such that $\mathbf{x} + \mathbf{y} = \mathbf{z}$ is another geometric vector.
- ▶ Multiplication by a scalar $\lambda \mathbf{x}, \lambda \in \mathbb{R}$, is also a geometric vector.

Vectors



- ▶ Polynomials are vectors. Adding two polynomials results in another polynomial. Multiplied by a scalar, the result is also a polynomial.

Vectors

- ▶ Polynomials are vectors. Adding two polynomials results in another polynomial. Multiplied by a scalar, the result is also a polynomial.
- ▶ Audio signals are also vectors. Addition of two audio signals and scalar multiplication result in new audio signals.

- ▶ Polynomials are vectors. Adding two polynomials results in another polynomial. Multiplied by a scalar, the result is also a polynomial.
- ▶ Audio signals are also vectors. Addition of two audio signals and scalar multiplication result in new audio signals.
- ▶ Elements of \mathbb{R}^n (tuples of n real numbers) are vectors. For example,

$$\mathbf{a} = \begin{bmatrix} 6 \\ 14 \\ -3 \end{bmatrix} \in \mathbb{R}^3$$

is a triplet of numbers. Adding two vectors $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$ component-wise results in another vectors $\mathbf{a} + \mathbf{b} = \mathbf{c} \in \mathbb{R}^n$.
Multiplying $\mathbf{a} \in \mathbb{R}^n$ by $\lambda \in \mathbb{R}$ results in a scaled vector $\lambda \mathbf{a} \in \mathbb{R}^n$.

Basic Operations with Vectors

- ▶ Vector of the same dimensionality can be added or subtracted.
- ▶ Consider two d -dimensional vectors:

$$\mathbf{x} + \mathbf{y} = \begin{bmatrix} x_1 \\ \dots \\ x_d \end{bmatrix} + \begin{bmatrix} y_1 \\ \dots \\ y_d \end{bmatrix} = \begin{bmatrix} x_1 + y_1 \\ \dots \\ x_d + y_d \end{bmatrix} \quad \mathbf{x} - \mathbf{y} = \begin{bmatrix} x_1 \\ \dots \\ x_d \end{bmatrix} - \begin{bmatrix} y_1 \\ \dots \\ y_d \end{bmatrix} = \begin{bmatrix} x_1 - y_1 \\ \dots \\ x_d - y_d \end{bmatrix}$$

- ▶ Vector addition is commutative: $\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$.

Basic Operations with Vectors

- ▶ A vector $\mathbf{x} \in \mathbb{R}^d$ can be scaled by a factor $a \in \mathbb{R}$ as follows

$$\mathbf{v} = a\mathbf{x} = a \begin{bmatrix} x_1 \\ \dots \\ x_d \end{bmatrix} = \begin{bmatrix} ax_1 \\ \dots \\ ax_d \end{bmatrix}$$

- ▶ Scalar multiplication operation scales the “length” of the vector, but does not change the “direction” (i.e., relative values of different components)

Basic Operations with Vectors

- ▶ The **dot product** between two vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^d$ is the sum of the element-wise multiplication of their individual components.

$$\mathbf{x} \cdot \mathbf{y} = \sum_{i=1}^d x_i y_i$$

- ▶ The dot product is commutative:

$$\mathbf{x} \cdot \mathbf{y} = \sum_{i=1}^d x_i y_i = \sum_{i=1}^d y_i x_i = \mathbf{y} \cdot \mathbf{x}$$

- ▶ The dot product is distributive:

$$\mathbf{x} \cdot (\mathbf{y} + \mathbf{z}) = \mathbf{x} \cdot \mathbf{y} + \mathbf{x} \cdot \mathbf{z}$$

Basic Operations with Vectors

- ▶ The dot product of a vector with itself produces the squared Euclidean norm. The norm defines the vector length and is denoted by $\| \cdot \|$:

$$\|x\|^2 = \mathbf{x} \cdot \mathbf{x} = \sum_{i=1}^d x_i^2$$

- ▶ The Euclidean norm of $x \in \mathbb{R}^d$ is defined as

$$\|x\|_2 = \sqrt{\sum_{i=1}^d x_i^2} = \sqrt{\mathbf{x} \cdot \mathbf{x}}$$

and computes the Euclidean distance of \mathbf{x} from the origin.

- ▶ The Euclidean norm is also known as the L_2 -norm.

Basic Operations with Vectors

- ▶ A generalization of the Euclidean norm is the L_p -norm, denoted by $\|\cdot\|_p$:

$$\|x\|_p = \left(\sum_{i=1}^d |x_i|^p \right)^{(1/p)}$$

where p is a positive value.

- ▶ When $p = 1$, we have the Manhattan norm, or the L_1 -norm.

Basic Operations with Vectors

- ▶ Vectors can be **normalized** to unit length by dividing them with their norm:

$$\mathbf{x}' = \frac{\mathbf{x}}{\|\mathbf{x}\|} = \frac{\mathbf{x}}{\sqrt{\mathbf{x} \cdot \mathbf{x}}}$$

- ▶ The resulting vector is a **unit vector**.
- ▶ The squared Euclidean distance $\mathbf{x}, \mathbf{y} \in \mathbb{R}^d$ can be shown to be the dot product of $\mathbf{x} - \mathbf{y}$ with itself:

$$\|\mathbf{x} - \mathbf{y}\|^2 = (\mathbf{x} - \mathbf{y}) \cdot (\mathbf{x} - \mathbf{y}) = \sum_{i=1}^d (x_i - y_i)^2$$

- ▶ **Cauchy-Schwarz Inequality:** the dot product between a pair of vectors is bounded above by the product of their lengths.

$$\left| \sum_{i=1}^d x_i y_i \right| = |\mathbf{x} \cdot \mathbf{y}| \leq \|\mathbf{x}\| \|\mathbf{y}\|$$

- ▶ **Triangle Inequality:** Consider the triangle formed by the origin, \mathbf{x} , and \mathbf{y} , the side length $\|\mathbf{x} - \mathbf{y}\|$ is no greater than the sum $\|\mathbf{x}\| + \|\mathbf{y}\|$ of the other two sides.

Basic Operations with Vectors

- ▶ Consider the triangle created by the origin, x , and y . Find the angle θ between x and y .

Basic Operations with Vectors

- ▶ Consider the triangle created by the origin, \mathbf{x} , and \mathbf{y} . Find the angle θ between \mathbf{x} and \mathbf{y} .
- ▶ The side lengths of this triangle are: $a = \|\mathbf{x}\|$, $b = \|\mathbf{y}\|$, and $c = \|\mathbf{x} - \mathbf{y}\|$. Using the cosine law, we have:

$$\begin{aligned}\cos(\theta) &= \frac{a^2 + b^2 - c^2}{2ab} = \frac{\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 - \|\mathbf{x} - \mathbf{y}\|^2}{2\|\mathbf{x}\|\|\mathbf{y}\|} \\ &= \frac{\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 - (\mathbf{x} - \mathbf{y}) \cdot (\mathbf{x} - \mathbf{y})}{2\|\mathbf{x}\|\|\mathbf{y}\|} \\ &= \frac{\mathbf{x} \cdot \mathbf{y}}{\|\mathbf{x}\|\|\mathbf{y}\|}\end{aligned}$$

Basic Operations with Vectors

- ▶ Consider the triangle created by the origin, \mathbf{x} , and \mathbf{y} . Find the angle θ between \mathbf{x} and \mathbf{y} .
- ▶ The side lengths of this triangle are: $a = \|\mathbf{x}\|$, $b = \|\mathbf{y}\|$, and $c = \|\mathbf{x} - \mathbf{y}\|$. Using the cosine law, we have:

$$\begin{aligned}\cos(\theta) &= \frac{a^2 + b^2 - c^2}{2ab} = \frac{\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 - \|\mathbf{x} - \mathbf{y}\|^2}{2\|\mathbf{x}\|\|\mathbf{y}\|} \\ &= \frac{\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 - (\mathbf{x} - \mathbf{y}) \cdot (\mathbf{x} - \mathbf{y})}{2\|\mathbf{x}\|\|\mathbf{y}\|} \\ &= \frac{\mathbf{x} \cdot \mathbf{y}}{\|\mathbf{x}\|\|\mathbf{y}\|}\end{aligned}$$

- ▶ Two vectors are **orthogonal** if their dot product is 0.

Basic Operations with Vectors

- ▶ Consider the triangle created by the origin, \mathbf{x} , and \mathbf{y} . Find the angle θ between \mathbf{x} and \mathbf{y} .
- ▶ The side lengths of this triangle are: $a = \|\mathbf{x}\|$, $b = \|\mathbf{y}\|$, and $c = \|\mathbf{x} - \mathbf{y}\|$. Using the cosine law, we have:

$$\begin{aligned}\cos(\theta) &= \frac{a^2 + b^2 - c^2}{2ab} = \frac{\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 - \|\mathbf{x} - \mathbf{y}\|^2}{2\|\mathbf{x}\|\|\mathbf{y}\|} \\ &= \frac{\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 - (\mathbf{x} - \mathbf{y}) \cdot (\mathbf{x} - \mathbf{y})}{2\|\mathbf{x}\|\|\mathbf{y}\|} \\ &= \frac{\mathbf{x} \cdot \mathbf{y}}{\|\mathbf{x}\|\|\mathbf{y}\|}\end{aligned}$$

- ▶ Two vectors are **orthogonal** if their dot product is 0.
- ▶ The vector $\mathbf{0}$ is considered orthogonal to every vector.

Definition

With $m, n \in \mathbb{N}$, a real-valued (m, n) matrix \mathbf{A} is an $m \cdot n$ -tuple of elements $a_{ij}, i = 1, \dots, m, j = 1, \dots, n$, which is ordered according to a rectangular scheme consisting of m rows and n columns:

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}, \quad a_{ij} \in \mathbb{R}$$

$\mathbb{R}^{m \times n}$ is the set of all real-valued (m, n) -matrices.

$\mathbf{A} \in \mathbb{R}^{m \times n}$ can also be represented as $\mathbf{a} \in \mathbb{R}^{mn}$ by stacking all n columns of the matrix into a long vector.

- ▶ A matrix has the same number of rows as columns is a **square** matrix. Otherwise, it is a **rectangular** matrix.
- ▶ A matrix having more rows than columns is referred to as *tall*, while a matrix having more columns than rows is referred to as *wide* or *fat*.
- ▶ A scalar can be considered as a 1×1 “matrix”.
- ▶ A d -dimensional vector can be considered a $1 \times d$ matrix when it is treated as a **row vector**.
- ▶ A d -dimensional vector can be considered a $d \times 1$ matrix when it is treated as a **column vector**.
- ▶ By defaults, vectors are assumed to be column vectors.

Matrix-Vector Multiplication

$$\begin{bmatrix} \cos 90^\circ & \sin 90^\circ \\ -\sin 90^\circ & \cos 90^\circ \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$
2

²<https://xkcd.com/184/>

Matrix-Vector Multiplication Ax

- ▶ Multiply A times x using rows of A .

$$\begin{bmatrix} 2 & 3 \\ 2 & 4 \\ 3 & 7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2x_1 + 3x_2 \\ 2x_1 + 4x_2 \\ 3x_1 + 7x_2 \end{bmatrix} = \begin{bmatrix} a_1^* x \\ a_2^* x \\ a_3^* x \end{bmatrix}$$

Ax = dot products of rows of A with x .

Matrix-Vector Multiplication Ax

- ▶ Multiply A times x using rows of A .

$$\begin{bmatrix} 2 & 3 \\ 2 & 4 \\ 3 & 7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2x_1 + 3x_2 \\ 2x_1 + 4x_2 \\ 3x_1 + 7x_2 \end{bmatrix} = \begin{bmatrix} \mathbf{a}_1^* x \\ \mathbf{a}_2^* x \\ \mathbf{a}_3^* x \end{bmatrix}$$

Ax = dot products of rows of A with x .

- ▶ Multiply A times x using columns of A .

$$\begin{bmatrix} 2 & 3 \\ 2 & 4 \\ 3 & 7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1 \begin{bmatrix} 2 \\ 2 \\ 3 \end{bmatrix} + x_2 \begin{bmatrix} 3 \\ 4 \\ 7 \end{bmatrix} = x_1 \mathbf{a}_1 + x_2 \mathbf{a}_2$$

Ax = combination of columns of \mathbf{a}_1 , \mathbf{a}_2 (of A) scaled by scalars x_1 , x_2 respectively.

Linear Combinations of Columns

Ax

Ax is a linear combination of the columns of A .

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = x_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} + \dots + x_n \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix}$$
$$Ax = x_1 \mathbf{a}_1 + x_2 \mathbf{a}_2 + \dots + x_n \mathbf{a}_n$$

Linear Combinations of Columns

Ax

Ax is a linear combination of the columns of A .

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = x_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} + \dots + x_n \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix}$$
$$Ax = x_1 \mathbf{a}_1 + x_2 \mathbf{a}_2 + \dots + x_n \mathbf{a}_n$$

Column space of $A = \mathbf{C}(A)$ = all vectors Ax

= all linear combinations of the columns

Column Space of A

$$A\mathbf{x} = \begin{bmatrix} 2 & 3 \\ 2 & 4 \\ 3 & 7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1 \begin{bmatrix} 2 \\ 2 \\ 3 \end{bmatrix} + x_2 \begin{bmatrix} 3 \\ 4 \\ 7 \end{bmatrix}$$

- ▶ Each $A\mathbf{x}$ is a vector in the \mathbb{R}^3 space.
- ▶ All combinations $A\mathbf{x} = x_1\mathbf{a}_1 + x_2\mathbf{a}_2$ produce what part of \mathbb{R}^3 ?

Column Space of A

$$A\mathbf{x} = \begin{bmatrix} 2 & 3 \\ 2 & 4 \\ 3 & 7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1 \begin{bmatrix} 2 \\ 2 \\ 3 \end{bmatrix} + x_2 \begin{bmatrix} 3 \\ 4 \\ 7 \end{bmatrix}$$

- ▶ Each $A\mathbf{x}$ is a vector in the \mathbb{R}^3 space.
- ▶ All combinations $A\mathbf{x} = x_1\mathbf{a}_1 + x_2\mathbf{a}_2$ produce what part of \mathbb{R}^3 ?
- ▶ Answer: a **plane**, containing:
 - the line of all vectors $x_1\mathbf{a}_1$,
 - the line of all vectors $x_2\mathbf{a}_2$,
 - the sum of any vector on one line + any vector on the other line, filling out an **infinite plane** containing the two lines, but not the whole \mathbb{R}^3 .

Column Space of A

$$A\mathbf{x} = \begin{bmatrix} 2 & 3 \\ 2 & 4 \\ 3 & 7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1 \begin{bmatrix} 2 \\ 2 \\ 3 \end{bmatrix} + x_2 \begin{bmatrix} 3 \\ 4 \\ 7 \end{bmatrix}$$

- ▶ Each $A\mathbf{x}$ is a vector in the \mathbb{R}^3 space.
- ▶ All combinations $A\mathbf{x} = x_1\mathbf{a}_1 + x_2\mathbf{a}_2$ produce what part of \mathbb{R}^3 ?
- ▶ Answer: a **plane**, containing:
 - the line of all vectors $x_1\mathbf{a}_1$,
 - the line of all vectors $x_2\mathbf{a}_2$,
 - the sum of any vector on one line + any vector on the other line, filling out an **infinite plane** containing the two lines, but not the whole \mathbb{R}^3 .

Definition

The combinations of the columns fill out the column space of A .

Column Space of A

$$A\mathbf{x} = \begin{bmatrix} 2 & 3 \\ 2 & 4 \\ 3 & 7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1 \begin{bmatrix} 2 \\ 2 \\ 3 \end{bmatrix} + x_2 \begin{bmatrix} 3 \\ 4 \\ 7 \end{bmatrix}$$

► $\mathbf{C}(A)$ is plane.

Column Space of A

$$A\mathbf{x} = \begin{bmatrix} 2 & 3 \\ 2 & 4 \\ 3 & 7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1 \begin{bmatrix} 2 \\ 2 \\ 3 \end{bmatrix} + x_2 \begin{bmatrix} 3 \\ 4 \\ 7 \end{bmatrix}$$

- ▶ $\mathbf{C}(A)$ is plane.
- ▶ The plane includes $(0,0)$, produced when $x_1 = x_2 = 0$.

Column Space of A

$$A\mathbf{x} = \begin{bmatrix} 2 & 3 \\ 2 & 4 \\ 3 & 7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1 \begin{bmatrix} 2 \\ 2 \\ 3 \end{bmatrix} + x_2 \begin{bmatrix} 3 \\ 4 \\ 7 \end{bmatrix}$$

- ▶ $\mathbf{C}(A)$ is plane.
- ▶ The plane includes $(0,0)$, produced when $x_1 = x_2 = 0$.
- ▶ The plane includes $(5,6,10) = \mathbf{a}_1 + \mathbf{a}_2$ and $(-1,-2,-4) = \mathbf{a}_1 - \mathbf{a}_2$. Every combination $x_1\mathbf{a}_1 + x_2\mathbf{a}_2$ is in $\mathbf{C}(A)$.

Column Space of A

$$A\mathbf{x} = \begin{bmatrix} 2 & 3 \\ 2 & 4 \\ 3 & 7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1 \begin{bmatrix} 2 \\ 2 \\ 3 \end{bmatrix} + x_2 \begin{bmatrix} 3 \\ 4 \\ 7 \end{bmatrix}$$

- ▶ $\mathbf{C}(A)$ is plane.
- ▶ The plane includes $(0,0)$, produced when $x_1 = x_2 = 0$.
- ▶ The plane includes $(5,6,10) = \mathbf{a}_1 + \mathbf{a}_2$ and $(-1,-2,-4) = \mathbf{a}_1 - \mathbf{a}_2$. Every combination $x_1\mathbf{a}_1 + x_2\mathbf{a}_2$ is in $\mathbf{C}(A)$.
- ▶ The probability the plane does not include a random point **rand**(3,1)? Which points are in the plane?

Column Space of A

$$A\mathbf{x} = \begin{bmatrix} 2 & 3 \\ 2 & 4 \\ 3 & 7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1 \begin{bmatrix} 2 \\ 2 \\ 3 \end{bmatrix} + x_2 \begin{bmatrix} 3 \\ 4 \\ 7 \end{bmatrix}$$

- ▶ $\mathbf{C}(A)$ is plane.
- ▶ The plane includes $(0,0)$, produced when $x_1 = x_2 = 0$.
- ▶ The plane includes $(5,6,10) = \mathbf{a}_1 + \mathbf{a}_2$ and $(-1,-2,-4) = \mathbf{a}_1 - \mathbf{a}_2$. Every combination $x_1\mathbf{a}_1 + x_2\mathbf{a}_2$ is in $\mathbf{C}(A)$.
- ▶ The probability the plane does not include a random point **rand**(3,1)? Which points are in the plane?

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

\mathbf{b} is in $\mathbf{C}(A)$ exactly when $A\mathbf{x} = \mathbf{b}$ has a solution \mathbf{x} .

\mathbf{x} shows how to express \mathbf{b} as a combination of the columns of A .

Column Space of A

- $\mathbf{b} = (1, 1, 1)$ is not in $\mathbf{C}(A)$ because

$$A\mathbf{x} = \begin{bmatrix} 2 & 3 \\ 2 & 4 \\ 3 & 7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \quad \text{is unsolvable.}$$

Column Space of A

- ▶ $\mathbf{b} = (1, 1, 1)$ is not in $\mathbf{C}(A)$ because

$$A\mathbf{x} = \begin{bmatrix} 2 & 3 \\ 2 & 4 \\ 3 & 7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \quad \text{is unsolvable.}$$

- ▶ What is the column space of A_2 ?

$$\begin{bmatrix} 2 & 3 & 5 \\ 2 & 4 & 6 \\ 3 & 7 & 10 \end{bmatrix}$$

Column Space of A

- $\mathbf{b} = (1, 1, 1)$ is not in $\mathbf{C}(A)$ because

$$A\mathbf{x} = \begin{bmatrix} 2 & 3 \\ 2 & 4 \\ 3 & 7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \quad \text{is unsolvable.}$$

- What is the column space of A_2 ?

$$\begin{bmatrix} 2 & 3 & 5 \\ 2 & 4 & 6 \\ 3 & 7 & 10 \end{bmatrix}$$

- $\mathbf{a}_3 = \mathbf{a}_1 + \mathbf{a}_2$, is already in $\mathbf{C}(A)$, the plane of \mathbf{a}_1 and \mathbf{a}_2 .
- Including this **dependent** column does not go beyond $\mathbf{C}(A)$.
- $\mathbf{C}(A_2) = \mathbf{C}(A)$.

Column Space of A

- $\mathbf{b} = (1, 1, 1)$ is not in $\mathbf{C}(A)$ because

$$A\mathbf{x} = \begin{bmatrix} 2 & 3 \\ 2 & 4 \\ 3 & 7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \quad \text{is unsolvable.}$$

- What is the column space of A_2 ?

$$\begin{bmatrix} 2 & 3 & 5 \\ 2 & 4 & 6 \\ 3 & 7 & 10 \end{bmatrix} \quad \begin{array}{l} \bullet \mathbf{a}_3 = \mathbf{a}_1 + \mathbf{a}_2, \text{ is already in } \mathbf{C}(A), \text{ the plane of } \mathbf{a}_1 \text{ and } \mathbf{a}_2. \\ \bullet \text{ Including this } \mathbf{dependent} \text{ column does not go beyond } \mathbf{C}(A). \\ \bullet \mathbf{C}(A_2) = \mathbf{C}(A). \end{array}$$

- What is the column space of A_3 ?

$$\begin{bmatrix} 2 & 3 & 1 \\ 2 & 4 & 1 \\ 3 & 7 & 1 \end{bmatrix}$$

Column Space of A

- $\mathbf{b} = (1, 1, 1)$ is not in $\mathbf{C}(A)$ because

$$A\mathbf{x} = \begin{bmatrix} 2 & 3 \\ 2 & 4 \\ 3 & 7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \quad \text{is unsolvable.}$$

- What is the column space of A_2 ?

$$\begin{bmatrix} 2 & 3 & 5 \\ 2 & 4 & 6 \\ 3 & 7 & 10 \end{bmatrix} \quad \begin{aligned} &\bullet \mathbf{a}_3 = \mathbf{a}_1 + \mathbf{a}_2, \text{ is already in } \mathbf{C}(A), \text{ the plane of } \mathbf{a}_1 \text{ and } \mathbf{a}_2. \\ &\bullet \text{ Including this **dependent** column does not go beyond } \mathbf{C}(A). \\ &\bullet \mathbf{C}(A_2) = \mathbf{C}(A). \end{aligned}$$

- What is the column space of A_3 ?

$$\begin{bmatrix} 2 & 3 & 1 \\ 2 & 4 & 1 \\ 3 & 7 & 1 \end{bmatrix} \quad \begin{aligned} &\bullet \mathbf{a}_3 = (1, 1, 1) \text{ is not in the plane } \mathbf{C}(A). \\ &\bullet \text{ Visualize the } xy\text{-plane and a third vector } (x_3, y_3, z_3) \text{ out of the plane (meaning that } z_3 \neq 0). \\ &\bullet \mathbf{C}(A_3) = \mathbb{R}^3. \end{aligned}$$

► Subspaces of \mathbb{R}^3 :

- The zero vector $(0, 0, 0)$.
- A line of all vectors $x_1 \mathbf{a}_1$.
- A plane of all vectors $x_1 \mathbf{a}_1 + x_2 \mathbf{a}_2$.
- The whole \mathbb{R}^3 with all vectors $x_1 \mathbf{a}_1 + x_2 \mathbf{a}_2 + x_3 \mathbf{a}_3$.

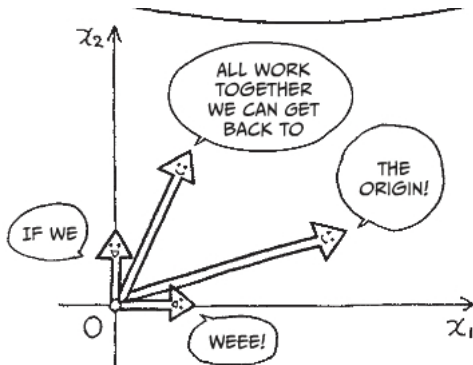
Column Spaces of \mathbb{R}^3

- ▶ Subspaces of \mathbb{R}^3 :
 - The zero vector $(0, 0, 0)$.
 - A line of all vectors $x_1 \mathbf{a}_1$.
 - A plane of all vectors $x_1 \mathbf{a}_1 + x_2 \mathbf{a}_2$.
 - The whole \mathbb{R}^3 with all vectors $x_1 \mathbf{a}_1 + x_2 \mathbf{a}_2 + x_3 \mathbf{a}_3$.
- ▶ Vectors $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ need to be **independent**. The only combination that gives the zero vector is $0\mathbf{a}_1 + 0\mathbf{a}_2 + 0\mathbf{a}_3$.

Column Spaces of \mathbb{R}^3

- ▶ Subspaces of \mathbb{R}^3 :
 - The zero vector $(0, 0, 0)$.
 - A line of all vectors $x_1 \mathbf{a}_1$.
 - A plane of all vectors $x_1 \mathbf{a}_1 + x_2 \mathbf{a}_2$.
 - The whole \mathbb{R}^3 with all vectors $x_1 \mathbf{a}_1 + x_2 \mathbf{a}_2 + x_3 \mathbf{a}_3$.
- ▶ Vectors $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ need to be **independent**. The only combination that gives the zero vector is $0\mathbf{a}_1 + 0\mathbf{a}_2 + 0\mathbf{a}_3$.
- ▶ The zero vector is in every subspace.

Linear Dependence



LINEAR DEPENDENCE

3

³<https://mathsci2.appstate.edu/sjg/class/2240/hf14.html>

Independent Columns, Basis, and Ranks of A

Definition

A **basis** for a subspace is a full set of independent vectors: All vectors in the space are combinations of the basis vector.

Independent Columns, Basis, and Ranks of A

Definition

A **basis** for a subspace is a full set of independent vectors: All vectors in the space are combinations of the basis vector.

Create a matrix C whose columns come directly from A :

- ▶ If column 1 of A is not all zero, put it into C .
- ▶ If column 2 of A is not a multiple of column 1, put it into C .
- ▶ If column 3 of A is not a combination of columns 1 and 2, put it into C . *Continue.*

Independent Columns, Basis, and Ranks of A

Definition

A **basis** for a subspace is a full set of independent vectors: All vectors in the space are combinations of the basis vector.

Create a matrix C whose columns come directly from A :

- ▶ If column 1 of A is not all zero, put it into C .
- ▶ If column 2 of A is not a multiple of column 1, put it into C .
- ▶ If column 3 of A is not a combination of columns 1 and 2, put it into C . *Continue.*
- ▶ At the end, C will have r columns ($r \leq n$). They are independent columns, and they are a “basis” for the column space $\mathbf{C}(A)$.

Independent Columns, Basis, and Ranks of A

$$\text{If } A = \begin{bmatrix} 1 & 3 & 8 \\ 1 & 2 & 6 \\ 0 & 1 & 2 \end{bmatrix} \text{ then } C =$$

Independent Columns, Basis, and Ranks of A

$$\text{If } A = \begin{bmatrix} 1 & 3 & 8 \\ 1 & 2 & 6 \\ 0 & 1 & 2 \end{bmatrix} \text{ then } C = \begin{bmatrix} 1 & 3 \\ 1 & 2 \\ 0 & 1 \end{bmatrix} \quad \begin{array}{l} n = 3 \text{ columns in } A \\ r = 2 \text{ columns in } C \end{array}$$

$$\text{If } A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{bmatrix} \text{ then } C =$$

Independent Columns, Basis, and Ranks of A

$$\text{If } A = \begin{bmatrix} 1 & 3 & 8 \\ 1 & 2 & 6 \\ 0 & 1 & 2 \end{bmatrix} \text{ then } C = \begin{bmatrix} 1 & 3 \\ 1 & 2 \\ 0 & 1 \end{bmatrix} \quad \begin{array}{l} n = 3 \text{ columns in } A \\ r = 2 \text{ columns in } C \end{array}$$

$$\text{If } A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{bmatrix} \text{ then } C = A \quad \begin{array}{l} n = 3 \text{ columns in } A \\ r = 3 \text{ columns in } C \end{array}$$

$$\text{If } A = \begin{bmatrix} 1 & 2 & 5 \\ 1 & 2 & 5 \\ 1 & 2 & 5 \end{bmatrix} \text{ then } C =$$

Independent Columns, Basis, and Ranks of A

$$\text{If } A = \begin{bmatrix} 1 & 3 & 8 \\ 1 & 2 & 6 \\ 0 & 1 & 2 \end{bmatrix} \text{ then } C = \begin{bmatrix} 1 & 3 \\ 1 & 2 \\ 0 & 1 \end{bmatrix} \quad \begin{array}{l} n = 3 \text{ columns in } A \\ r = 2 \text{ columns in } C \end{array}$$

$$\text{If } A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{bmatrix} \text{ then } C = A \quad \begin{array}{l} n = 3 \text{ columns in } A \\ r = 3 \text{ columns in } C \end{array}$$

$$\text{If } A = \begin{bmatrix} 1 & 2 & 5 \\ 1 & 2 & 5 \\ 1 & 2 & 5 \end{bmatrix} \text{ then } C = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \quad \begin{array}{l} n = 3 \text{ columns in } A \\ r = 1 \text{ columns in } C \end{array}$$

Independent Columns, Basis, and Ranks of A

$$\text{If } A = \begin{bmatrix} 1 & 3 & 8 \\ 1 & 2 & 6 \\ 0 & 1 & 2 \end{bmatrix} \text{ then } C = \begin{bmatrix} 1 & 3 \\ 1 & 2 \\ 0 & 1 \end{bmatrix} \quad \begin{array}{l} n = 3 \text{ columns in } A \\ r = 2 \text{ columns in } C \end{array}$$

$$\text{If } A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{bmatrix} \text{ then } C = A \quad \begin{array}{l} n = 3 \text{ columns in } A \\ r = 3 \text{ columns in } C \end{array}$$

$$\text{If } A = \begin{bmatrix} 1 & 2 & 5 \\ 1 & 2 & 5 \\ 1 & 2 & 5 \end{bmatrix} \text{ then } C = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \quad \begin{array}{l} n = 3 \text{ columns in } A \\ r = 1 \text{ columns in } C \end{array}$$

- ▶ The number r counts independent columns.
- ▶ It is the “dimension” of the column space of A and C (same space).

Definition

The **rank** of a matrix is the **dimension** of its column space.

Rank Factorization $A = CR$

- ▶ The matrix C connects to A by a third matrix R : $A = CR$.
- ▶ $A \in \mathbb{R}^{m \times n}$, $C \in \mathbb{R}^{m \times r}$, $R \in \mathbb{R}^{r \times n}$

Rank Factorization $A = CR$

- ▶ The matrix C connects to A by a third matrix R : $A = CR$.
- ▶ $A \in \mathbb{R}^{m \times n}$, $C \in \mathbb{R}^{m \times r}$, $R \in \mathbb{R}^{r \times n}$

$$A = \begin{bmatrix} 1 & 3 & 8 \\ 1 & 2 & 6 \\ 0 & 1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 3 \\ 1 & 2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 2 \end{bmatrix} = CR$$

Rank Factorization $A = CR$

- ▶ The matrix C connects to A by a third matrix R : $A = CR$.
- ▶ $A \in \mathbb{R}^{m \times n}$, $C \in \mathbb{R}^{m \times r}$, $R \in \mathbb{R}^{r \times n}$

$$A = \begin{bmatrix} 1 & 3 & 8 \\ 1 & 2 & 6 \\ 0 & 1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 3 \\ 1 & 2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 2 \end{bmatrix} = CR$$

- ▶ C multiplies the first column of R produces column 1 of A .
- ▶ C multiplies the second column of R produces column 2 of A .
- ▶ C multiplies the third column of R produces column 3 of A .
- ▶ Combinations of the columns of C produce the columns of A
→ Put the right numbers in R .

Rank Factorization $A = CR$

- ▶ The matrix C connects to A by a third matrix R : $A = CR$.
- ▶ $A \in \mathbb{R}^{m \times n}$, $C \in \mathbb{R}^{m \times r}$, $R \in \mathbb{R}^{r \times n}$

$$A = \begin{bmatrix} 1 & 3 & 8 \\ 1 & 2 & 6 \\ 0 & 1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 3 \\ 1 & 2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 2 \end{bmatrix} = CR$$

- ▶ C multiplies the first column of R produces column 1 of A .
- ▶ C multiplies the second column of R produces column 2 of A .
- ▶ C multiplies the third column of R produces column 3 of A .
- ▶ Combinations of the columns of C produce the columns of A
→ Put the right numbers in R .

Definition

$R = \mathbf{rref}(A)$ = row-reduced echelon form of A .

Rank Factorization $A = CR$

$$A = \begin{bmatrix} 1 & 3 & 8 \\ 1 & 2 & 6 \\ 0 & 1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 3 \\ 1 & 2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 2 \end{bmatrix} = CR$$

- ▶ The matrix R has $r = 2$ rows \mathbf{r}_1^* , \mathbf{r}_2^* .

Rank Factorization $A = CR$

$$A = \begin{bmatrix} 1 & 3 & 8 \\ 1 & 2 & 6 \\ 0 & 1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 3 \\ 1 & 2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 2 \end{bmatrix} = CR$$

- ▶ The matrix R has $r = 2$ rows \mathbf{r}_1^* , \mathbf{r}_2^* .
- ▶ Multiply row 1 of C with R , we get $\mathbf{r}_1^* + 3\mathbf{r}_2^* \rightarrow$ row 1 of A .
- ▶ Multiply row 2 of C with R , we get $\mathbf{r}_1^* + 2\mathbf{r}_2^* \rightarrow$ row 2 of A .
- ▶ Multiply row 3 of C with R , we get $0\mathbf{r}_1^* + 1\mathbf{r}_2^* \rightarrow$ row 3 of A .

Rank Factorization $A = CR$

$$A = \begin{bmatrix} 1 & 3 & 8 \\ 1 & 2 & 6 \\ 0 & 1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 3 \\ 1 & 2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 2 \end{bmatrix} = CR$$

- ▶ The matrix R has $r = 2$ rows \mathbf{r}_1^* , \mathbf{r}_2^* .
- ▶ Multiply row 1 of C with R , we get $\mathbf{r}_1^* + 3\mathbf{r}_2^* \rightarrow$ row 1 of A .
- ▶ Multiply row 2 of C with R , we get $\mathbf{r}_1^* + 2\mathbf{r}_2^* \rightarrow$ row 2 of A .
- ▶ Multiply row 3 of C with R , we get $0\mathbf{r}_1^* + 1\mathbf{r}_2^* \rightarrow$ row 3 of A .
- ▶ R has independent rows: No row is a combination of the other rows.
Hint: Look at the zeros and ones in R - the identity matrix I in R .

Rank Factorization $A = CR$

$$A = \begin{bmatrix} 1 & 3 & 8 \\ 1 & 2 & 6 \\ 0 & 1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 3 \\ 1 & 2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 2 \end{bmatrix} = CR$$

- ▶ The matrix R has $r = 2$ rows \mathbf{r}_1^* , \mathbf{r}_2^* .
- ▶ Multiply row 1 of C with R , we get $\mathbf{r}_1^* + 3\mathbf{r}_2^* \rightarrow$ row 1 of A .
- ▶ Multiply row 2 of C with R , we get $\mathbf{r}_1^* + 2\mathbf{r}_2^* \rightarrow$ row 2 of A .
- ▶ Multiply row 3 of C with R , we get $0\mathbf{r}_1^* + 1\mathbf{r}_2^* \rightarrow$ row 3 of A .
- ▶ R has independent rows: No row is a combination of the other rows.
Hint: Look at the zeros and ones in R - the identity matrix I in R .
- ▶ The rows of R are a **basis for the row space** of A .
- ▶ Notation: The row space of matrix $A = \mathbf{C}(A^\top)$.

Rank Factorization $A = CR$

- 1 The r columns of C are independent (by their construction).
- 2 Every column of A is a combination of those r columns of C (because $A = CR$).
- 3 The r rows of R are independent (they contain the matrix I_r).
- 4 Every row of A is a combination of those r rows of R (because $A = CR$).

Rank Factorization $A = CR$

- 1 The r columns of C are independent (by their construction).
- 2 Every column of A is a combination of those r columns of C (because $A = CR$).
- 3 The r rows of R are independent (they contain the matrix I_r).
- 4 Every row of A is a combination of those r rows of R (because $A = CR$).

Key facts:

- ▶ The r columns of C is a **basis** for $\mathbf{C}(A)$: dimension r .
- ▶ The r rows of R is a **basis** for $\mathbf{C}(A^\top)$: dimension r .

Notice

The number of independent columns = The number of independent rows.
The column space and row space of A both have dimension r .
The column rank of A = The row rank of A .

Question: If an $n \times n$ matrix A has n independent columns, then $C = ?$, $R = ?$

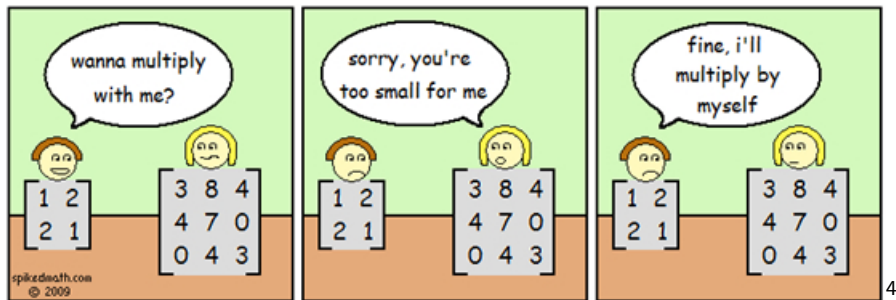
Question: If an $n \times n$ matrix A has n independent columns, then $C = ?$, $R = ?$

Answer: $C = A$, $R = I$.

Table of Contents

- 1 Matrix-Vector Multiplication Ax
- 2 Matrix-Matrix Multiplication AB
- 3 The Four Fundamental Subspaces of A : $\mathbf{C}(A)$, $\mathbf{C}(A^\top)$, $\mathbf{N}(A)$, $\mathbf{N}(A^\top)$
- 4 Elimination and $A = LU$
- 5 Orthogonal Matrices, Subspaces, and Projections

Matrix-Matrix Multiplication AB



4

⁴<https://mathsci2.appstate.edu/sjg/class/2240/hf14.html>

Compute AB by Inner Products

- ▶ **Inner products** (rows times columns) produce each of the numbers in $AB = C$:

$$\begin{bmatrix} \cdot & \cdot & \cdot \\ a_{21} & a_{22} & a_{23} \\ \cdot & \cdot & \cdot \end{bmatrix} \begin{bmatrix} \cdot & \cdot & b_{13} \\ \cdot & \cdot & b_{23} \\ \cdot & \cdot & b_{33} \end{bmatrix} = \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & c_{23} \\ \cdot & \cdot & \cdot \end{bmatrix}$$

- ▶ $c_{ij} = (\text{row } i \text{ of } A) \cdot (\text{column } j \text{ of } B)$

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj} = \sum_{k=1}^n a_{ik}b_{kj} = \mathbf{a}_i^* \mathbf{b}_j$$

- ▶ **Outer products** (columns times rows) produce **rank one matrices**.

$$\mathbf{uv}^\top = \begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix} \begin{bmatrix} 3 & 4 & 6 \end{bmatrix} = \begin{bmatrix} 6 & 6 & 8 \\ 6 & 6 & 8 \\ 3 & 4 & 6 \end{bmatrix}$$

- ▶ An $m \times 1$ matrix (a column \mathbf{u}) times a $1 \times p$ matrix (a row \mathbf{v}^\top) gives an $m \times p$ matrix.

- ▶ **Outer products** (columns times rows) produce **rank one matrices**.

$$\mathbf{uv}^\top = \begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix} \begin{bmatrix} 3 & 4 & 6 \end{bmatrix} = \begin{bmatrix} 6 & 6 & 8 \\ 6 & 6 & 8 \\ 3 & 4 & 6 \end{bmatrix}$$

- ▶ An $m \times 1$ matrix (a column \mathbf{u}) times a $1 \times p$ matrix (a row \mathbf{v}^\top) gives an $m \times p$ matrix.
- ▶ All columns of \mathbf{uv}^\top are multiples of \mathbf{u} .
- ▶ All rows of \mathbf{uv}^\top are multiples of \mathbf{v}^\top .

- ▶ **Outer products** (columns times rows) produce **rank one matrices**.

$$\mathbf{uv}^\top = \begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix} \begin{bmatrix} 3 & 4 & 6 \end{bmatrix} = \begin{bmatrix} 6 & 6 & 8 \\ 6 & 6 & 8 \\ 3 & 4 & 6 \end{bmatrix}$$

- ▶ An $m \times 1$ matrix (a column \mathbf{u}) times a $1 \times p$ matrix (a row \mathbf{v}^\top) gives an $m \times p$ matrix.
- ▶ All columns of \mathbf{uv}^\top are multiples of \mathbf{u} .
- ▶ All rows of \mathbf{uv}^\top are multiples of \mathbf{v}^\top .
- ▶ The column space of \mathbf{uv}^\top is the line through \mathbf{u} .
- ▶ The row space of \mathbf{uv}^\top is the line through \mathbf{v} .

- ▶ **Outer products** (columns times rows) produce **rank one matrices**.

$$\mathbf{uv}^\top = \begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix} \begin{bmatrix} 3 & 4 & 6 \end{bmatrix} = \begin{bmatrix} 6 & 6 & 8 \\ 6 & 6 & 8 \\ 3 & 4 & 6 \end{bmatrix}$$

- ▶ An $m \times 1$ matrix (a column \mathbf{u}) times a $1 \times p$ matrix (a row \mathbf{v}^\top) gives an $m \times p$ matrix.
- ▶ All columns of \mathbf{uv}^\top are multiples of \mathbf{u} .
- ▶ All rows of \mathbf{uv}^\top are multiples of \mathbf{v}^\top .
- ▶ The column space of \mathbf{uv}^\top is the line through \mathbf{u} .
- ▶ The row space of \mathbf{uv}^\top is the line through \mathbf{v} .
- ▶ All non-zero matrices \mathbf{uv}^\top have rank one.

$AB = \text{Sum of Rank-1 Matrices}$

- ▶ The product AB is the sum of columns \mathbf{a}_k times rows \mathbf{b}_k^* .

$$AB = \begin{bmatrix} | & & | \\ \mathbf{a}_1 & \dots & \mathbf{a}_n \\ | & & | \end{bmatrix} \begin{bmatrix} - & \mathbf{b}_1^* & - \\ & \vdots & \\ - & \mathbf{b}_n^* & - \end{bmatrix} = \mathbf{a}_1 \mathbf{b}_1^* + \mathbf{a}_2 \mathbf{b}_2^* + \dots + \mathbf{a}_n \mathbf{b}_n^*$$

$AB = \text{Sum of Rank-1 Matrices}$

- ▶ The product AB is the sum of columns \mathbf{a}_k times rows \mathbf{b}_k^* .

$$AB = \begin{bmatrix} | & & | \\ \mathbf{a}_1 & \dots & \mathbf{a}_n \\ | & & | \end{bmatrix} \begin{bmatrix} - & \mathbf{b}_1^* & - \\ & \vdots & \\ - & \mathbf{b}_n^* & - \end{bmatrix} = \mathbf{a}_1 \mathbf{b}_1^* + \mathbf{a}_2 \mathbf{b}_2^* + \dots + \mathbf{a}_n \mathbf{b}_n^*$$

- ▶ Example:

$$\begin{bmatrix} 1 & 0 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} 2 & 4 \\ 0 & 5 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \end{bmatrix} \begin{bmatrix} 2 & 4 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} \begin{bmatrix} 0 & 5 \end{bmatrix} = \begin{bmatrix} 2 & 4 \\ 6 & 12 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 5 \end{bmatrix} = \begin{bmatrix} 2 & 4 \\ 6 & 17 \end{bmatrix}$$

Insight from Column times Row

- ▶ Looking for the important part of a matrix A .
- ▶ Factor A into CR and look at the pieces $c_k r_k^*$ of $A = CR$.
- ▶ Factoring A into CR is the reverse of multiplying $CR = A$.
- ▶ The inside information about A is not visible until A is factored.

- ▶ Looking for the important part of a matrix A .
- ▶ Factor A into CR and look at the pieces $c_k r_k^*$ of $A = CR$.
- ▶ Factoring A into CR is the reverse of multiplying $CR = A$.
- ▶ The inside information about A is not visible until A is factored.

Important Factorizations

- 1 $A = LU$: elimination
- 2 $A = QR$: orthogonalization
- 3 $S = Q\Lambda Q^\top$: eigenvalues and orthonormal eigenvectors
- 4 $A = X\Lambda X^{-1}$: diagonalization
- 5 $A = U\Sigma V^\top$: Singular Value Decomposition (SVD)

Inverse Matrices

- ▶ The square matrix A is invertible if there exists a matrix A^{-1} that

$$A^{-1}A = I \text{ and } AA^{-1} = I$$

Inverse Matrices

- ▶ The square matrix A is invertible if there exists a matrix A^{-1} that

$$A^{-1}A = I \text{ and } AA^{-1} = I$$

- ▶ The matrix A cannot have two different inverses. Suppose $BA = I$ and also $AC = I$. Then $B = C$.

$$B(AC) = (BA)C \text{ gives } BI = IC \text{ or } B = C.$$

Inverse Matrices

- ▶ The square matrix A is invertible if there exists a matrix A^{-1} that

$$A^{-1}A = I \text{ and } AA^{-1} = I$$

- ▶ The matrix A cannot have two different inverses. Suppose $BA = I$ and also $AC = I$. Then $B = C$.

$$B(AC) = (BA)C \text{ gives } BI = IC \text{ or } B = C.$$

- ▶ If A is invertible, the one and only solution to $Ax = b$ is $x = A^{-1}b$.

Inverse Matrices

- ▶ The square matrix A is invertible if there exists a matrix A^{-1} that

$$A^{-1}A = I \text{ and } AA^{-1} = I$$

- ▶ The matrix A cannot have two different inverses. Suppose $BA = I$ and also $AC = I$. Then $B = C$.

$$B(AC) = (BA)C \text{ gives } BI = IC \text{ or } B = C.$$

- ▶ If A is invertible, the one and only solution to $Ax = b$ is $x = A^{-1}b$.
- ▶ If $Ax = 0$ for a nonzero vector x , then A has no inverse.

Inverse Matrices

- ▶ The square matrix A is invertible if there exists a matrix A^{-1} that

$$A^{-1}A = I \text{ and } AA^{-1} = I$$

- ▶ The matrix A cannot have two different inverses. Suppose $BA = I$ and also $AC = I$. Then $B = C$.

$$B(AC) = (BA)C \text{ gives } BI = IC \text{ or } B = C.$$

- ▶ If A is invertible, the one and only solution to $Ax = b$ is $x = A^{-1}b$.
- ▶ If $Ax = 0$ for a nonzero vector x , then A has no inverse.
- ▶ If A and B are invertible then so is AB . The inverse of AB is

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

Table of Contents

- 1 Matrix-Vector Multiplication Ax
- 2 Matrix-Matrix Multiplication AB
- 3 The Four Fundamental Subspaces of A : $\mathbf{C}(A)$, $\mathbf{C}(A^T)$, $\mathbf{N}(A)$, $\mathbf{N}(A^T)$
- 4 Elimination and $A = LU$
- 5 Orthogonal Matrices, Subspaces, and Projections

Example 1

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix} = \mathbf{u}\mathbf{v}^\top$$

- ▶ Column space $\mathbf{C}(A)$ is the line through $\mathbf{u} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$.

Example 1

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix} = \mathbf{u}\mathbf{v}^\top$$

- ▶ Column space $\mathbf{C}(A)$ is the line through $\mathbf{u} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$.
- ▶ Row space $\mathbf{C}(A^\top)$ is the line through $\mathbf{v} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$.

Example 1

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix} = \mathbf{u}\mathbf{v}^\top$$

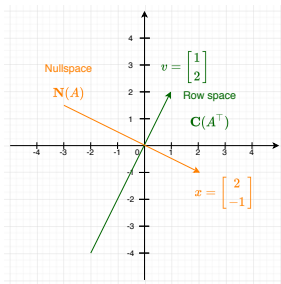
- ▶ Column space $\mathbf{C}(A)$ is the line through $\mathbf{u} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$.
- ▶ Row space $\mathbf{C}(A^\top)$ is the line through $\mathbf{v} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$.
- ▶ Nullspace $\mathbf{N}(A)$ is the line through $\mathbf{x} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$. $A\mathbf{x} = \mathbf{0}$.

Example 1

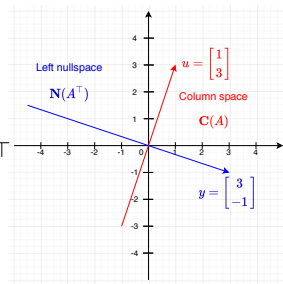
$$A = \begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix} = \mathbf{u}\mathbf{v}^\top$$

- ▶ Column space $\mathbf{C}(A)$ is the line through $\mathbf{u} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$.
- ▶ Row space $\mathbf{C}(A^\top)$ is the line through $\mathbf{v} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$.
- ▶ Nullspace $\mathbf{N}(A)$ is the line through $\mathbf{x} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$. $A\mathbf{x} = \mathbf{0}$.
- ▶ Left nullspace $\mathbf{N}(A^\top)$ is the line through $\mathbf{y} = \begin{bmatrix} 3 \\ -1 \end{bmatrix}$. $A^\top\mathbf{y} = \mathbf{0}$.

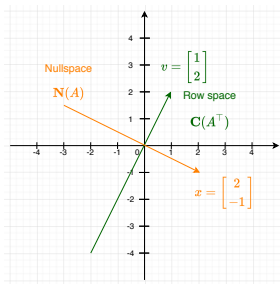
Example 1



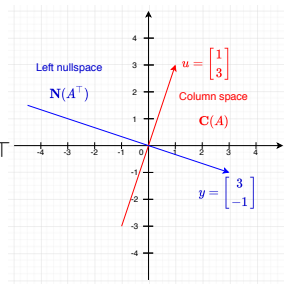
$$A = \begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix} = uv^T$$



Example 1



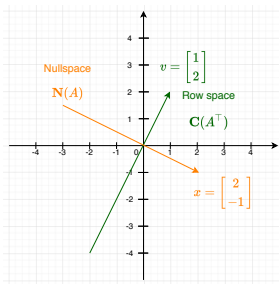
$$A = \begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix} = uv^T$$



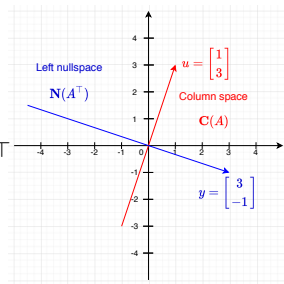
Definition

The column space $C(A)$ contains all combinations of the columns of A .

Example 1



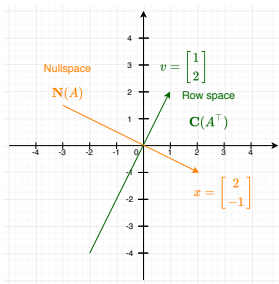
$$A = \begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix} = uv^T$$



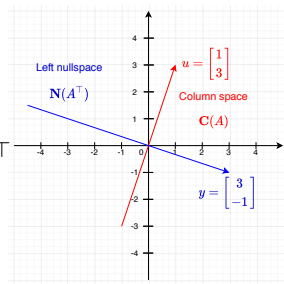
Definition

The column space $\mathbf{C}(A)$ contains all combinations of the columns of A .
The row space $\mathbf{C}(A^T)$ contains all combinations of the columns of A^T .

Example 1



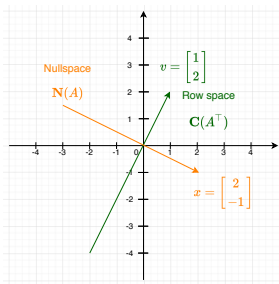
$$A = \begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix} = uv^T$$



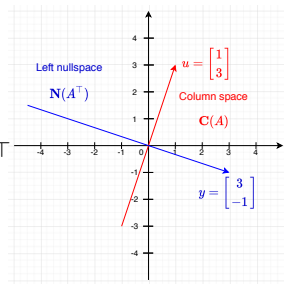
Definition

The column space $\mathbf{C}(A)$ contains all combinations of the columns of A .
The row space $\mathbf{C}(A^T)$ contains all combinations of the columns of A^T .
The nullspace $\mathbf{N}(A)$ contains all solutions x to $Ax = 0$.

Example 1



$$A = \begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix} = uv^T$$



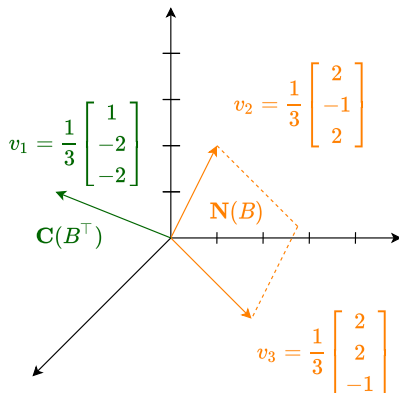
Definition

The column space $\mathbf{C}(A)$ contains all combinations of the columns of A .
The row space $\mathbf{C}(A^T)$ contains all combinations of the columns of A^T .
The nullspace $\mathbf{N}(A)$ contains all solutions x to $Ax = 0$.
The left nullspace $\mathbf{N}(A)$ contains all solutions y to $A^T y = 0$.

Example 2

$$B = \begin{bmatrix} 1 & -2 & -2 \\ 3 & -6 & -6 \end{bmatrix}$$

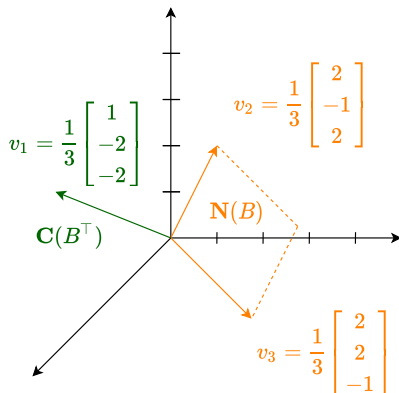
- $Bx = 0$ has solutions $x_1 = (2, 1, 0)$
and $x_2 = (2, 0, 1)$.



Example 2

$$B = \begin{bmatrix} 1 & -2 & -2 \\ 3 & -6 & -6 \end{bmatrix}$$

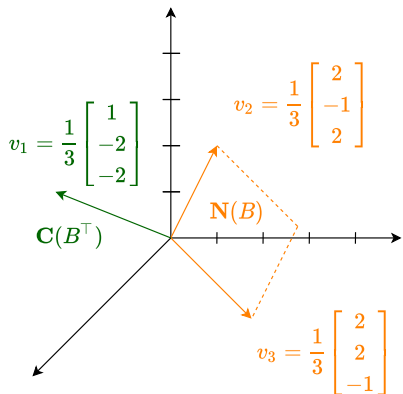
- ▶ $Bx = 0$ has solutions $x_1 = (2, 1, 0)$ and $x_2 = (2, 0, 1)$.
- ▶ The row space $\mathbf{C}(B^\top)$ is the infinite line through $v_1 = \frac{1}{3}(1, -2, -2)$.



Example 2

$$B = \begin{bmatrix} 1 & -2 & -2 \\ 3 & -6 & -6 \end{bmatrix}$$

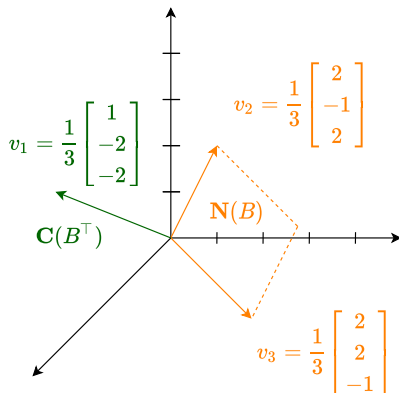
- ▶ $Bx = 0$ has solutions $x_1 = (2, 1, 0)$ and $x_2 = (2, 0, 1)$.
- ▶ The row space $\mathbf{C}(B^\top)$ is the infinite line through $v_1 = \frac{1}{3}(1, -2, -2)$.
- ▶ x_1 and x_2 are in the same plane with $v_2 = \frac{1}{3}(2, -1, 2)$ and $v_3 = \frac{1}{3}(2, 2, -1)$.



Example 2

$$B = \begin{bmatrix} 1 & -2 & -2 \\ 3 & -6 & -6 \end{bmatrix}$$

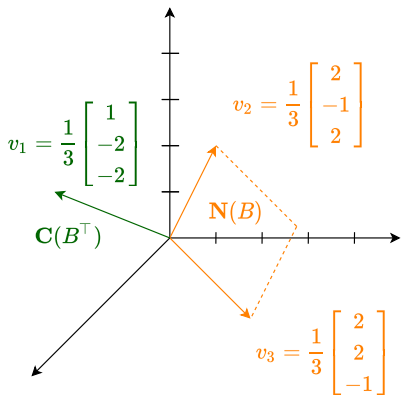
- ▶ $Bx = 0$ has solutions $x_1 = (2, 1, 0)$ and $x_2 = (2, 0, 1)$.
- ▶ The row space $\mathbf{C}(B^T)$ is the infinite line through $v_1 = \frac{1}{3}(1, -2, -2)$.
- ▶ x_1 and x_2 are in the same plane with $v_2 = \frac{1}{3}(2, -1, 2)$ and $v_3 = \frac{1}{3}(2, 2, -1)$.
- ▶ The nullspace $\mathbf{N}(B)$ has an **orthonormal basis** v_2 and v_3 , is the infinite plane of v_2 and v_3 .



Example 2

$$B = \begin{bmatrix} 1 & -2 & -2 \\ 3 & -6 & -6 \end{bmatrix}$$

- ▶ $Bx = 0$ has solutions $x_1 = (2, 1, 0)$ and $x_2 = (2, 0, 1)$.
- ▶ The row space $\mathbf{C}(B^T)$ is the infinite line through $v_1 = \frac{1}{3}(1, -2, -2)$.
- ▶ x_1 and x_2 are in the same plane with $v_2 = \frac{1}{3}(2, -1, 2)$ and $v_3 = \frac{1}{3}(2, 2, -1)$.
- ▶ The nullspace $\mathbf{N}(B)$ has an **orthonormal basis** v_2 and v_3 , is the infinite plane of v_2 and v_3 .
- ▶ v_1, v_2, v_3 : an orthonormal basis for \mathbb{R}^3 .



Subspaces of A

$$\text{If } Ax = \mathbf{0} \text{ then } \begin{bmatrix} \text{row } 1 \\ \vdots \\ \text{row } m \end{bmatrix} \begin{bmatrix} x \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$$

- ▶ x is orthogonal to every row of A .

Subspaces of A

$$\text{If } Ax = \mathbf{0} \text{ then } \begin{bmatrix} \text{row } 1 \\ \vdots \\ \text{row } m \end{bmatrix} \begin{bmatrix} x \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$$

- ▶ x is orthogonal to every row of A .
- ▶ Every x in the nullspace of A is orthogonal to the row space of A .

Subspaces of A

$$\text{If } Ax = \mathbf{0} \text{ then } \begin{bmatrix} \text{row } 1 \\ \vdots \\ \text{row } m \end{bmatrix} \begin{bmatrix} x \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$$

- ▶ x is orthogonal to every row of A .
- ▶ Every x in the nullspace of A is orthogonal to the row space of A .
- ▶ Every y in the nullspace of A^T is orthogonal to the column space of A .

Subspaces of A

$$\text{If } Ax = \mathbf{0} \text{ then } \begin{bmatrix} \text{row } 1 \\ \vdots \\ \text{row } m \end{bmatrix} \begin{bmatrix} x \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$$

- ▶ x is orthogonal to every row of A .
- ▶ Every x in the nullspace of A is orthogonal to the row space of A .
- ▶ Every y in the nullspace of A^T is orthogonal to the column space of A .

$$\begin{array}{ccccc} \mathbf{N}(A) & \perp & \mathbf{C}(A^T) & \mathbf{N}(A^T) & \perp & \mathbf{C}(A) \\ \text{Dimensions} & & n - r & r & & m - r & r \end{array}$$

- ▶ Two orthogonal subspaces. The dimensions add to n and to m .

Table of Contents

- 1 Matrix-Vector Multiplication Ax
- 2 Matrix-Matrix Multiplication AB
- 3 The Four Fundamental Subspaces of A : $\mathbf{C}(A)$, $\mathbf{C}(A^\top)$, $\mathbf{N}(A)$, $\mathbf{N}(A^\top)$
- 4 Elimination and $A = LU$
- 5 Orthogonal Matrices, Subspaces, and Projections

$Ax = b$ by Elimination

The usual order:

- ▶ Column 1.

$$[A \mid \mathbf{b}] = \left[\begin{array}{cccc|c} 2 & 1 & -1 & 2 & 5 \\ 4 & 5 & -3 & 6 & 9 \\ -2 & 5 & -2 & 6 & 4 \\ 4 & 11 & -4 & 8 & 2 \end{array} \right] \rightarrow \left[\begin{array}{cccc|c} 2 & 1 & -1 & 2 & 5 \\ 0 & 3 & -1 & 2 & -1 \\ 0 & 6 & -3 & 8 & 9 \\ 0 & 9 & -2 & 4 & -8 \end{array} \right]$$

$Ax = b$ by Elimination

The usual order:

- ▶ Column 1.
 - Row 1 is the first pivot row.

$$[A \mid \mathbf{b}] = \left[\begin{array}{cccc|c} 2 & 1 & -1 & 2 & 5 \\ 4 & 5 & -3 & 6 & 9 \\ -2 & 5 & -2 & 6 & 4 \\ 4 & 11 & -4 & 8 & 2 \end{array} \right] \rightarrow \left[\begin{array}{cccc|c} 2 & 1 & -1 & 2 & 5 \\ 0 & 3 & -1 & 2 & -1 \\ 0 & 6 & -3 & 8 & 9 \\ 0 & 9 & -2 & 4 & -8 \end{array} \right]$$

$Ax = b$ by Elimination

The usual order:

► Column 1.

- Row 1 is the first pivot row.
- Multiply row 1 by numbers $l_{21}, l_{31}, \dots, l_{n1}$ and subtract from rows $2, 3, \dots, n$ of A respectively.

$$\text{Multipliers } l_{21} = \frac{a_{21}}{a_{11}} \quad l_{31} = \frac{a_{31}}{a_{11}} \quad \dots \quad l_{n1} = \frac{a_{n1}}{a_{11}}$$

$$[A \mid \mathbf{b}] = \left[\begin{array}{cccc|c} 2 & 1 & -1 & 2 & 5 \\ 4 & 5 & -3 & 6 & 9 \\ -2 & 5 & -2 & 6 & 4 \\ 4 & 11 & -4 & 8 & 2 \end{array} \right] \rightarrow \left[\begin{array}{cccc|c} 2 & 1 & -1 & 2 & 5 \\ 0 & 3 & -1 & 2 & -1 \\ 0 & 6 & -3 & 8 & 9 \\ 0 & 9 & -2 & 4 & -8 \end{array} \right]$$

$Ax = b$ by Elimination

The usual order:

- ▶ Column 2.

$$\left[\begin{array}{cccc|c} 2 & 1 & -1 & 2 & 5 \\ 0 & 3 & -1 & 2 & -1 \\ 0 & 6 & -3 & 8 & 9 \\ 0 & 9 & -2 & 4 & -8 \end{array} \right] \rightarrow \left[\begin{array}{cccc|c} 2 & 1 & -1 & 2 & 5 \\ 0 & 3 & -1 & 2 & -1 \\ 0 & 0 & -1 & 4 & 11 \\ 0 & 0 & 1 & -2 & -5 \end{array} \right]$$

$Ax = b$ by Elimination

The usual order:

- ▶ Column 2.
 - The **new** row 2 is the second pivot row.

$$\left[\begin{array}{cccc|c} 2 & 1 & -1 & 2 & 5 \\ 0 & 3 & -1 & 2 & -1 \\ 0 & 6 & -3 & 8 & 9 \\ 0 & 9 & -2 & 4 & -8 \end{array} \right] \rightarrow \left[\begin{array}{cccc|c} 2 & 1 & -1 & 2 & 5 \\ 0 & 3 & -1 & 2 & -1 \\ 0 & 0 & -1 & 4 & 11 \\ 0 & 0 & 1 & -2 & -5 \end{array} \right]$$

$Ax = b$ by Elimination

The usual order:

► Column 2.

- The **new** row 2 is the second pivot row.
- Multiply row 2 by numbers $l_{32}, l_{42}, \dots, l_{n2}$ and subtract from rows 3, 4, \dots, n of A respectively.

$$\text{Multipliers } l_{32} = \frac{a_{32}}{a_{22}} \quad l_{42} = \frac{a_{42}}{a_{22}} \quad \dots \quad l_{n2} = \frac{a_{n2}}{a_{22}}$$

$$\left[\begin{array}{cccc|c} 2 & 1 & -1 & 2 & 5 \\ 0 & 3 & -1 & 2 & -1 \\ 0 & 6 & -3 & 8 & 9 \\ 0 & 9 & -2 & 4 & -8 \end{array} \right] \rightarrow \left[\begin{array}{cccc|c} 2 & 1 & -1 & 2 & 5 \\ 0 & 3 & -1 & 2 & -1 \\ 0 & 0 & -1 & 4 & 11 \\ 0 & 0 & 1 & -2 & -5 \end{array} \right]$$

$Ax = b$ by Elimination

The usual order:

- ▶ Column 3.

$$\left[\begin{array}{cccc|c} 2 & 1 & -1 & 2 & 5 \\ 0 & 3 & -1 & 2 & -1 \\ 0 & 0 & -1 & 4 & 11 \\ 0 & 0 & 1 & -2 & -5 \end{array} \right] \rightarrow \left[\begin{array}{cccc|c} \color{red}{2} & 1 & -1 & 2 & 5 \\ 0 & \color{red}{3} & -1 & 2 & -1 \\ 0 & 0 & \color{red}{-1} & 4 & 11 \\ 0 & 0 & 0 & \color{red}{2} & 6 \end{array} \right] = [U \mid \mathbf{c}]$$

$Ax = b$ by Elimination

The usual order:

- ▶ Column 3.
 - The **new** row 3 is the third pivot row.

$$\left[\begin{array}{cccc|c} 2 & 1 & -1 & 2 & 5 \\ 0 & 3 & -1 & 2 & -1 \\ 0 & 0 & -1 & 4 & 11 \\ 0 & 0 & 1 & -2 & -5 \end{array} \right] \rightarrow \left[\begin{array}{cccc|c} \mathbf{2} & 1 & -1 & 2 & 5 \\ 0 & \mathbf{3} & -1 & 2 & -1 \\ 0 & 0 & \mathbf{-1} & 4 & 11 \\ 0 & 0 & 0 & \mathbf{2} & 6 \end{array} \right] = [U \mid \mathbf{c}]$$

$Ax = b$ by Elimination

The usual order:

► Column 3.

- The **new** row 3 is the third pivot row.
- Multiply row 3 by numbers $l_{43}, l_{53}, \dots, l_{n3}$ and subtract from rows 4, 5, \dots, n of A respectively.

$$\text{Multipliers } l_{43} = \frac{a_{43}}{a_{33}} \quad l_{53} = \frac{a_{53}}{a_{33}} \quad \dots \quad l_{n3} = \frac{a_{n3}}{a_{33}}$$

$$\left[\begin{array}{cccc|c} 2 & 1 & -1 & 2 & 5 \\ 0 & 3 & -1 & 2 & -1 \\ 0 & 0 & -1 & 4 & 11 \\ 0 & 0 & 1 & -2 & -5 \end{array} \right] \rightarrow \left[\begin{array}{cccc|c} \color{red}{2} & 1 & -1 & 2 & 5 \\ 0 & \color{red}{3} & -1 & 2 & -1 \\ 0 & 0 & \color{red}{-1} & 4 & 11 \\ 0 & 0 & 0 & \color{red}{2} & 6 \end{array} \right] = [U \mid \mathbf{c}]$$

$Ax = b$ by Elimination

The usual order:

► Column 3.

- The **new** row 3 is the third pivot row.
- Multiply row 3 by numbers $l_{43}, l_{53}, \dots, l_{n3}$ and subtract from rows 4, 5, \dots, n of A respectively.

$$\text{Multipliers } l_{43} = \frac{a_{43}}{a_{33}} \quad l_{53} = \frac{a_{53}}{a_{33}} \quad \dots \quad l_{n3} = \frac{a_{n3}}{a_{33}}$$

$$\left[\begin{array}{cccc|c} 2 & 1 & -1 & 2 & 5 \\ 0 & 3 & -1 & 2 & -1 \\ 0 & 0 & -1 & 4 & 11 \\ 0 & 0 & 1 & -2 & -5 \end{array} \right] \rightarrow \left[\begin{array}{cccc|c} \color{red}{2} & 1 & -1 & 2 & 5 \\ 0 & \color{red}{3} & -1 & 2 & -1 \\ 0 & 0 & \color{red}{-1} & 4 & 11 \\ 0 & 0 & 0 & \color{red}{2} & 6 \end{array} \right] = [U \mid \mathbf{c}]$$

- Columns 3 to n : Eliminating on A until obtaining the **upper triangular** U : n pivots on its **diagonal**.

$Ax = b$ by Elimination

$$2x_1 + x_2 - x_3 + 2x_4 = 5$$

$$3x_2 - x_3 + 2x_4 = -1$$

$$-x_3 + 4x_4 = 11$$

$$2x_4 = 6$$

By back substitution, we get

$$x_4 = 3, \quad x_3 = 1, \quad x_2 = -2, \quad x_1 = 1$$

Lower Triangular L and Upper Triangular U

- Elimination on $Ax = b$ produces the upper triangular matrix

$$U = \begin{bmatrix} 2 & 1 & -1 & 2 \\ 0 & 3 & -1 & 2 \\ 0 & 0 & -1 & 4 \\ 0 & 0 & 0 & 2 \end{bmatrix}$$

Lower Triangular L and Upper Triangular U

- Elimination on $Ax = b$ produces the upper triangular matrix

$$U = \begin{bmatrix} 2 & 1 & -1 & 2 \\ 0 & 3 & -1 & 2 \\ 0 & 0 & -1 & 4 \\ 0 & 0 & 0 & 2 \end{bmatrix}$$

- and the lower triangular matrix

$$L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ l_{21} & 1 & 0 & 0 \\ l_{31} & l_{32} & 1 & 0 \\ l_{41} & l_{42} & l_{43} & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ -1 & 2 & 1 & 0 \\ 2 & 3 & -1 & 1 \end{bmatrix}$$

Lower Triangular L and Upper Triangular U

- Elimination on $Ax = b$ produces the upper triangular matrix

$$U = \begin{bmatrix} 2 & 1 & -1 & 2 \\ 0 & 3 & -1 & 2 \\ 0 & 0 & -1 & 4 \\ 0 & 0 & 0 & 2 \end{bmatrix}$$

- and the lower triangular matrix

$$L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ l_{21} & 1 & 0 & 0 \\ l_{31} & l_{32} & 1 & 0 \\ l_{41} & l_{42} & l_{43} & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ -1 & 2 & 1 & 0 \\ 2 & 3 & -1 & 1 \end{bmatrix}$$

- Elimination factors A into a lower triangular L times an upper triangular U .

$$A = LU$$

The Factorization $A = LU$

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ l_{21} & 1 & 0 & 0 \\ l_{31} & l_{32} & 1 & 0 \\ l_{41} & l_{42} & l_{43} & 1 \end{bmatrix} \begin{bmatrix} \text{pivot row 1} \\ \text{pivot row 2} \\ \text{pivot row 3} \\ \text{pivot row 4} \end{bmatrix} = \begin{bmatrix} 2 & 1 & -1 & 2 \\ 4 & 5 & -3 & 6 \\ -2 & 5 & -2 & 6 \\ 4 & 11 & -4 & 8 \end{bmatrix}$$

$$= \begin{bmatrix} 1 \\ l_{21} \\ l_{31} \\ l_{41} \end{bmatrix} [\text{pivot row 1}] + \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & x & x & x \\ 0 & x & x & x \\ 0 & x & x & x \end{bmatrix}$$

$$l_{ij} = \frac{a_{ij}}{a_{jj}}$$

$$= \begin{bmatrix} 1 \\ 2 \\ -1 \\ 2 \end{bmatrix} \begin{bmatrix} 2 & 1 & -1 & 2 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & x & x & x \\ 0 & x & x & x \\ 0 & x & x & x \end{bmatrix}$$

$$= \begin{bmatrix} 2 & 1 & -1 & 2 \\ 4 & 2 & -2 & 4 \\ -2 & -1 & 1 & -2 \\ 4 & 2 & -2 & 4 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 3 & -1 & 2 \\ 0 & 6 & -3 & 8 \\ 0 & 9 & -2 & 4 \end{bmatrix}$$

The Factorization $A = LU$

$$\begin{aligned}
 A &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ l_{21} & 1 & 0 & 0 \\ l_{31} & l_{32} & 1 & 0 \\ l_{41} & l_{42} & l_{43} & 1 \end{bmatrix} \begin{bmatrix} \text{pivot row 1} \\ \text{pivot row 2} \\ \text{pivot row 3} \\ \text{pivot row 4} \end{bmatrix} = \begin{bmatrix} 2 & 1 & -1 & 2 \\ 4 & 5 & -3 & 6 \\ -2 & 5 & -2 & 6 \\ 4 & 11 & -4 & 8 \end{bmatrix} \\
 &= \begin{bmatrix} 1 \\ l_{21} \\ l_{31} \\ l_{41} \end{bmatrix} [\text{pivot row 1}] + \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & x & x & x \\ 0 & x & x & x \\ 0 & x & x & x \end{bmatrix} \quad \boxed{l_{ij} = \frac{a_{ij}}{a_{jj}}} \\
 &= \begin{bmatrix} 1 \\ 2 \\ -1 \\ 2 \end{bmatrix} \begin{bmatrix} 2 & 1 & -1 & 2 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & x & x & x \\ 0 & x & x & x \\ 0 & x & x & x \end{bmatrix} \\
 &= \begin{bmatrix} 2 & 1 & -1 & 2 \\ 4 & 2 & -2 & 4 \\ -2 & -1 & 1 & -2 \\ 4 & 2 & -2 & 4 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 3 & -1 & 2 \\ 0 & 6 & -3 & 8 \\ 0 & 9 & -2 & 4 \end{bmatrix}
 \end{aligned}$$

The first step reduces the 4×4 problem to a 3×3 problem by removing $l_1 u_1^*$.

The Factorization $A = LU$

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ l_{21} & 1 & 0 & 0 \\ l_{31} & l_{32} & 1 & 0 \\ l_{41} & l_{42} & l_{43} & 1 \end{bmatrix} \begin{bmatrix} \text{pivot row 1} \\ \text{pivot row 2} \\ \text{pivot row 3} \\ \text{pivot row 4} \end{bmatrix} = l_1 u_1^* + \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 3 & -1 & 2 \\ 0 & 6 & -3 & 8 \\ 0 & 9 & -2 & 4 \end{bmatrix}$$

$$= l_1 u_1^* + \begin{bmatrix} 0 \\ 1 \\ l_{32} \\ l_{42} \end{bmatrix} [\text{pivot row 2}] + \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & x & x \\ 0 & 0 & x & x \end{bmatrix}$$

$$l_{ij} = \frac{a_{ij}}{a_{jj}}$$

$$= l_1 u_1^* + \begin{bmatrix} 0 \\ 1 \\ 2 \\ 3 \end{bmatrix} [0 \quad 3 \quad -1 \quad 2] + \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & x & x \\ 0 & 0 & x & x \end{bmatrix}$$

$$= l_1 u_1^* + \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 3 & -1 & 2 \\ 0 & 6 & -2 & 4 \\ 0 & 9 & -3 & 6 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 4 \\ 0 & 0 & 1 & -2 \end{bmatrix}$$

The Factorization $A = LU$

$$\begin{aligned}
 A &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ l_{21} & 1 & 0 & 0 \\ l_{31} & l_{32} & 1 & 0 \\ l_{41} & l_{42} & l_{43} & 1 \end{bmatrix} \begin{bmatrix} \text{pivot row 1} \\ \text{pivot row 2} \\ \text{pivot row 3} \\ \text{pivot row 4} \end{bmatrix} = l_1 u_1^* + \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 3 & -1 & 2 \\ 0 & 6 & -3 & 8 \\ 0 & 9 & -2 & 4 \end{bmatrix} \\
 &= l_1 u_1^* + \begin{bmatrix} 0 \\ 1 \\ l_{32} \\ l_{42} \end{bmatrix} [\text{pivot row 2}] + \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & x & x \\ 0 & 0 & x & x \end{bmatrix} \quad \boxed{l_{ij} = \frac{a_{ij}}{a_{jj}}} \\
 &= l_1 u_1^* + \begin{bmatrix} 0 \\ 1 \\ 2 \\ 3 \end{bmatrix} [0 \quad 3 \quad -1 \quad 2] + \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & x & x \\ 0 & 0 & x & x \end{bmatrix} \\
 &= l_1 u_1^* + \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 3 & -1 & 2 \\ 0 & 6 & -2 & 4 \\ 0 & 9 & -3 & 6 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 4 \\ 0 & 0 & 1 & -2 \end{bmatrix}
 \end{aligned}$$

The second step reduces the 3×3 problem to a 2×2 problem by removing $l_2 u_2^*$.



UIT
UNIVERSITY OF INFORMATION TECHNOLOGY

The Factorization $A = LU$

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ l_{21} & 1 & 0 & 0 \\ l_{31} & l_{32} & 1 & 0 \\ l_{41} & l_{42} & l_{43} & 1 \end{bmatrix} \begin{bmatrix} \text{pivot row 1} \\ \text{pivot row 2} \\ \text{pivot row 3} \\ \text{pivot row 4} \end{bmatrix} = l_1 u_1^* + l_2 u_2^* + \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 4 \\ 0 & 0 & 1 & -2 \end{bmatrix}$$

$$= l_1 u_1^* + l_2 u_2^* + \begin{bmatrix} 0 \\ 0 \\ 1 \\ l_{43} \end{bmatrix} [\text{pivot row 3}] + \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & x \end{bmatrix} \quad \boxed{l_{ij} = \frac{a_{ij}}{a_{jj}}}$$

$$= l_1 u_1^* + l_2 u_2^* + \begin{bmatrix} 0 \\ 0 \\ 1 \\ -1 \end{bmatrix} [0 \quad 0 \quad -1 \quad 4] + \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & x \end{bmatrix}$$

$$= l_1 u_1^* + l_2 u_2^* + \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 4 \\ 0 & 0 & 1 & -4 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}$$

$$= l_1 u_1^* + l_2 u_2^* + l_3 u_3^* + l_4 u_4^*$$

The Factorization $A = LU$

$$\begin{aligned}
 A &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ l_{21} & 1 & 0 & 0 \\ l_{31} & l_{32} & 1 & 0 \\ l_{41} & l_{42} & l_{43} & 1 \end{bmatrix} \begin{bmatrix} \text{pivot row 1} \\ \text{pivot row 2} \\ \text{pivot row 3} \\ \text{pivot row 4} \end{bmatrix} = l_1 u_1^* + l_2 u_2^* + \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 4 \\ 0 & 0 & 1 & -2 \end{bmatrix} \\
 &= l_1 u_1^* + l_2 u_2^* + \begin{bmatrix} 0 \\ 0 \\ 1 \\ l_{43} \end{bmatrix} [\text{pivot row 3}] + \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & x \end{bmatrix} \quad \boxed{l_{ij} = \frac{a_{ij}}{a_{jj}}} \\
 &= l_1 u_1^* + l_2 u_2^* + \begin{bmatrix} 0 \\ 0 \\ 1 \\ -1 \end{bmatrix} [0 \quad 0 \quad -1 \quad 4] + \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & x \end{bmatrix} \\
 &= l_1 u_1^* + l_2 u_2^* + \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 4 \\ 0 & 0 & 1 & -4 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix} \\
 &= l_1 u_1^* + l_2 u_2^* + l_3 u_3^* + l_4 u_4^*
 \end{aligned}$$

The third step reduces the 2×2 problem to a single number by removing $l_3 u_3^*$.

Elimination and $A = LU$

- ▶ Start from $[A \ \mathbf{b}] = [LU \ \mathbf{b}]$.

Elimination and $A = LU$

- ▶ Start from $[A \ \mathbf{b}] = [LU \ \mathbf{b}]$.
- ▶ Elimination produces $[U \ L^{-1}\mathbf{b}] = [U \ \mathbf{c}]$.

Elimination and $A = LU$

- ▶ Start from $[A \ \mathbf{b}] = [LU \ \mathbf{b}]$.
- ▶ Elimination produces $[U \ L^{-1}\mathbf{b}] = [U \ \mathbf{c}]$.
- ▶ Elimination on $A\mathbf{x} = \mathbf{b}$ produces the equation $U\mathbf{x} = \mathbf{c}$ that are ready for back substitution.

Elimination and $A = LU$

- ▶ Start from $[A \ \mathbf{b}] = [LU \ \mathbf{b}]$.
- ▶ Elimination produces $[U \ L^{-1}\mathbf{b}] = [U \ \mathbf{c}]$.
- ▶ Elimination on $A\mathbf{x} = \mathbf{b}$ produces the equation $U\mathbf{x} = \mathbf{c}$ that are ready for back substitution.
- ▶ $A = LU = \sum l_i \mathbf{u}_i^* = \text{sum of rank one matrices.}$

Table of Contents

- 1 Matrix-Vector Multiplication Ax
- 2 Matrix-Matrix Multiplication AB
- 3 The Four Fundamental Subspaces of A : $\mathbf{C}(A)$, $\mathbf{C}(A^\top)$, $\mathbf{N}(A)$, $\mathbf{N}(A^\top)$
- 4 Elimination and $A = LU$
- 5 Orthogonal Matrices, Subspaces, and Projections

Orthogonality

- ▶ Orthogonal \sim perpendicular.

Orthogonality

- ▶ Orthogonal \sim perpendicular.
- ▶ Orthogonal vectors \mathbf{x} and \mathbf{y} :

$$\mathbf{x}^\top \mathbf{y} = x_1y_1 + x_2y_2 + \cdots + x_ny_n = 0$$

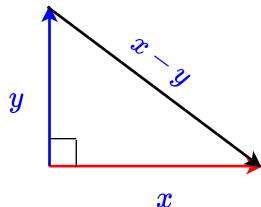
Orthogonality

- ▶ Orthogonal \sim perpendicular.
- ▶ Orthogonal vectors \mathbf{x} and \mathbf{y} :

$$\mathbf{x}^\top \mathbf{y} = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n = 0$$

Law of Cosines: θ is the angle between \mathbf{x} and \mathbf{y} :

$$\|\mathbf{x} - \mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 - 2\|\mathbf{x}\|\|\mathbf{y}\|\cos\theta$$



Orthogonality

- ▶ Orthogonal \sim perpendicular.
- ▶ Orthogonal vectors x and y :

$$x^\top y = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n = 0$$

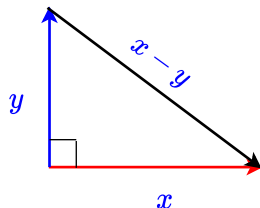
Law of Cosines: θ is the angle between x and y :

$$\|x - y\|^2 = \|x\|^2 + \|y\|^2 - 2\|x\|\|y\|\cos\theta$$

Orthogonal vectors have $\cos\theta = 0$.

Pythagoras Law:

$$\begin{aligned}\|x - y\|^2 &= \|x\|^2 + \|y\|^2 \\ (x - y)^\top (x - y) &= x^\top x + y^\top y \\ x^\top x + y^\top y - x^\top y - y^\top x &= x^\top x + y^\top y \\ x^\top y &= 0\end{aligned}$$



Orthogonal Basis

- ▶ Orthogonal basis for a subspace: Every pair of basis vectors has $\mathbf{v}_i^\top \mathbf{v}_j = 0$

Orthogonal Basis

- ▶ Orthogonal basis for a subspace: Every pair of basis vectors has $\mathbf{v}_i^\top \mathbf{v}_j = 0$
- ▶ Orthonormal basis: Orthogonal basis of unit vectors: Every $\mathbf{v}_i^\top \mathbf{v}_i = 1$ (length 1).

Orthogonal Basis

- ▶ Orthogonal basis for a subspace: Every pair of basis vectors has $\mathbf{v}_i^\top \mathbf{v}_j = 0$
- ▶ Orthonormal basis: Orthogonal basis of unit vectors: Every $\mathbf{v}_i^\top \mathbf{v}_i = 1$ (length 1).
- ▶ From orthogonal to orthonormal, divide every basis vector \mathbf{v}_i by its length $\|\mathbf{v}_i\|$.

Orthogonal Basis

- ▶ Orthogonal basis for a subspace: Every pair of basis vectors has $\mathbf{v}_i^\top \mathbf{v}_j = 0$
- ▶ Orthonormal basis: Orthogonal basis of unit vectors: Every $\mathbf{v}_i^\top \mathbf{v}_i = 1$ (length 1).
- ▶ From orthogonal to orthonormal, divide every basis vector \mathbf{v}_i by its length $\|\mathbf{v}_i\|$.
- ▶ The standard basis is orthogonal (and orthonormal) in \mathbb{R}^n :

$$\text{Standard basis } \mathbf{i}, \mathbf{j}, \mathbf{k} \text{ in } \mathbb{R}^3 \quad \mathbf{i} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \quad \mathbf{j} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \quad \mathbf{k} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Orthogonal Basis

- ▶ Orthogonal basis for a subspace: Every pair of basis vectors has $\mathbf{v}_i^\top \mathbf{v}_j = 0$
- ▶ Orthonormal basis: Orthogonal basis of unit vectors: Every $\mathbf{v}_i^\top \mathbf{v}_i = 1$ (length 1).
- ▶ From orthogonal to orthonormal, divide every basis vector \mathbf{v}_i by its length $\|\mathbf{v}_i\|$.
- ▶ The standard basis is orthogonal (and orthonormal) in \mathbb{R}^n :

$$\text{Standard basis } \mathbf{i}, \mathbf{j}, \mathbf{k} \text{ in } \mathbb{R}^3 \quad \mathbf{i} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \quad \mathbf{j} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \quad \mathbf{k} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

- ▶ Every subspace of \mathbb{R}^n has an orthogonal basis.

Orthogonal Subspaces

- ▶ Subspace **S** is orthogonal to subspace **T**: Every vector in **S** is orthogonal to every vector in **T**.

Orthogonal Subspaces

- ▶ The row space $\mathbf{C}(A^\top)$ is orthogonal to the nullspace $\mathbf{N}(A)$.

$$A\mathbf{x} = \begin{bmatrix} \text{row } 1 \\ \vdots \\ \text{row } m \end{bmatrix} \begin{bmatrix} \mathbf{x} \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$$

Orthogonal Subspaces

- ▶ The row space $\mathbf{C}(A^\top)$ is orthogonal to the nullspace $\mathbf{N}(A)$.

$$A\mathbf{x} = \begin{bmatrix} \text{row } 1 \\ \vdots \\ \text{row } m \end{bmatrix} \begin{bmatrix} \mathbf{x} \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$$

- ▶ The column space $\mathbf{C}(A)$ is orthogonal to the left nullspace $\mathbf{N}(A^\top)$.

$$A^\top \mathbf{y} = \begin{bmatrix} (\text{column } 1)^\top \\ \vdots \\ (\text{column } m)^\top \end{bmatrix} \begin{bmatrix} \mathbf{y} \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$$

Orthogonal Subspaces

- ▶ Every vector \mathbf{v} in \mathbb{R}^n has a row space component \mathbf{v}_{row} and a nullspace component \mathbf{v}_{null} : $\mathbf{v} = \mathbf{v}_{row} + \mathbf{v}_{null}$

$$A\mathbf{x} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Orthogonal Subspaces

- ▶ Every vector \mathbf{v} in \mathbb{R}^n has a row space component \mathbf{v}_{row} and a nullspace component \mathbf{v}_{null} : $\mathbf{v} = \mathbf{v}_{row} + \mathbf{v}_{null}$

$$A\mathbf{x} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

- ▶ The row space $\mathbf{C}(A^\top)$ is the plane of all vectors $\beta_1 \mathbf{a}_1^* + \beta_2 \mathbf{a}_2^*$.

Orthogonal Subspaces

- ▶ Every vector \mathbf{v} in \mathbb{R}^n has a row space component \mathbf{v}_{row} and a nullspace component \mathbf{v}_{null} : $\mathbf{v} = \mathbf{v}_{row} + \mathbf{v}_{null}$

$$A\mathbf{x} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

- ▶ The row space $\mathbf{C}(A^\top)$ is the plane of all vectors $\beta_1 \mathbf{a}_1^* + \beta_2 \mathbf{a}_2^*$.
- ▶ The nullspace $\mathbf{N}(A)$ is the line through $\mathbf{u} = (0, 0, 1)$: all vectors $\beta_3 \mathbf{u}$

$$\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} \in \mathbb{R}^3 \quad \mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \underbrace{\beta_1 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + \beta_2 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}}_{\mathbf{v}_{row}} + \underbrace{\beta_3 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}}_{\mathbf{v}_{null}}$$

Orthogonal Subspaces

- ▶ Every vector \mathbf{v} in \mathbb{R}^n has a row space component \mathbf{v}_{row} and a nullspace component \mathbf{v}_{null} : $\mathbf{v} = \mathbf{v}_{row} + \mathbf{v}_{null}$

$$A\mathbf{x} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

- ▶ The row space $\mathbf{C}(A^\top)$ is the plane of all vectors $\beta_1 \mathbf{a}_1^* + \beta_2 \mathbf{a}_2^*$.
- ▶ The nullspace $\mathbf{N}(A)$ is the line through $\mathbf{u} = (0, 0, 1)$: all vectors $\beta_3 \mathbf{u}$

$$\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} \in \mathbb{R}^3 \quad \mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \underbrace{\beta_1 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + \beta_2 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}}_{\mathbf{v}_{row}} + \underbrace{\beta_3 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}}_{\mathbf{v}_{null}}$$

- ▶ Dimensions: $\dim \mathbf{C}(A^\top) + \dim \mathbf{N}(A) = r + (n - r) = n$.

Orthogonal Subspaces

- ▶ Every vector \mathbf{v} in \mathbb{R}^n has a row space component \mathbf{v}_{row} and a nullspace component \mathbf{v}_{null} : $\mathbf{v} = \mathbf{v}_{row} + \mathbf{v}_{null}$

$$A\mathbf{x} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

- ▶ The row space $\mathbf{C}(A^\top)$ is the plane of all vectors $\beta_1 \mathbf{a}_1^* + \beta_2 \mathbf{a}_2^*$.
- ▶ The nullspace $\mathbf{N}(A)$ is the line through $\mathbf{u} = (0, 0, 1)$: all vectors $\beta_3 \mathbf{u}$

$$\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} \in \mathbb{R}^3 \quad \mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \underbrace{\beta_1 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + \beta_2 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}}_{\mathbf{v}_{row}} + \underbrace{\beta_3 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}}_{\mathbf{v}_{null}}$$

- ▶ Dimensions: $\dim \mathbf{C}(A^\top) + \dim \mathbf{N}(A) = r + (n - r) = n$.
- ▶ A row space basis (r vectors) and a nullspace basis ($n - r$ vectors) produces a basis for the whole \mathbb{R}^n (n vectors).

The Big Picture

Fundamental Theorem in Linear Algebra

The row space and nullspace of A are orthogonal complements in \mathbb{R}^n .

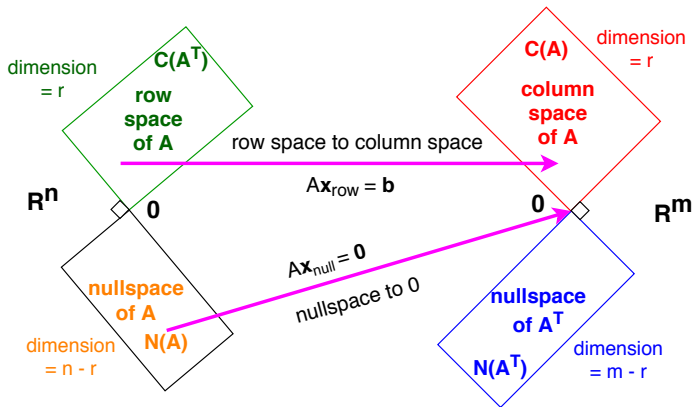
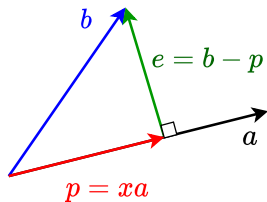


Figure: Two pairs of orthogonal subspaces.

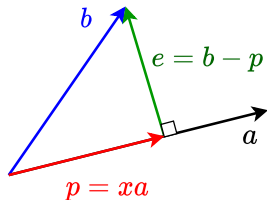
Projection onto a Line



► $e = b - p$

► $p = xa$

Projection onto a Line



► $e = b - p$

► $p = xa$

► Because e is orthogonal to a :

$$a^\top e = 0$$

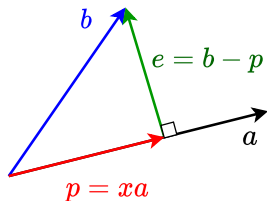
$$a^\top (b - p) = 0$$

$$a^\top (b - xa) = 0$$

$$xa^\top a = a^\top b$$

$$x = \frac{a^\top b}{a^\top a}$$

Projection onto a Line



► $e = b - p$

► $p = xa$

► Because e is orthogonal to a :

$$a^\top e = 0$$

$$a^\top (b - p) = 0$$

$$a^\top (b - xa) = 0$$

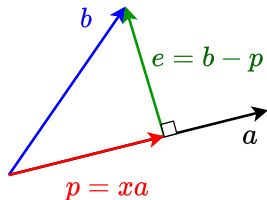
$$xa^\top a = a^\top b$$

$$x = \frac{a^\top b}{a^\top a}$$

► Therefore, $p = ax = a \frac{a^\top b}{a^\top a}$

► There is a **projection matrix** P that $p = Pb$.

Projection onto a Line



- ▶ $e = b - p$
- ▶ $p = xa$
- ▶ Because e is orthogonal to a :

$$a^\top e = 0$$

$$a^\top (b - p) = 0$$

$$a^\top (b - xa) = 0$$

$$xa^\top a = a^\top b$$

$$x = \frac{a^\top b}{a^\top a}$$

- ▶ Therefore, $p = ax = a \frac{a^\top b}{a^\top a}$
- ▶ There is a **projection matrix** P that $p = Pb$.

$$P = \frac{aa^\top}{a^\top a}$$

Projection onto a Line

$$P = \frac{\mathbf{a}\mathbf{a}^\top}{\mathbf{a}^\top\mathbf{a}}$$

- ▶ Column space of A : matrix-vector multiplication $A\mathbf{x} \in \mathbf{C}(A)$.
- ▶ $\mathbf{p} = P\mathbf{b}$. What is the column space $\mathbf{C}(P)$?

Projection onto a Line

$$P = \frac{\mathbf{a}\mathbf{a}^\top}{\mathbf{a}^\top\mathbf{a}}$$

- ▶ Column space of A : matrix-vector multiplication $A\mathbf{x} \in \mathbf{C}(A)$.
- ▶ $\mathbf{p} = P\mathbf{b}$. What is the column space $\mathbf{C}(P)$?
- ▶ $\mathbf{C}(P)$ is the line through \mathbf{a} .
- ▶ Is P symmetric?

Projection onto a Line

$$P = \frac{\mathbf{a}\mathbf{a}^\top}{\mathbf{a}^\top\mathbf{a}}$$

- ▶ Column space of A : matrix-vector multiplication $A\mathbf{x} \in \mathbf{C}(A)$.
- ▶ $\mathbf{p} = P\mathbf{b}$. What is the column space $\mathbf{C}(P)$?
- ▶ $\mathbf{C}(P)$ is the line through \mathbf{a} .
- ▶ Is P symmetric?

$$P^\top = \left(\frac{\mathbf{a}\mathbf{a}^\top}{\mathbf{a}^\top\mathbf{a}} \right)^\top = \frac{\mathbf{a}\mathbf{a}^\top}{\mathbf{a}^\top\mathbf{a}} = P. \quad \text{Yes.}$$

- ▶ What if we project \mathbf{b} twice?

Projection onto a Line

$$P = \frac{\mathbf{a}\mathbf{a}^\top}{\mathbf{a}^\top\mathbf{a}}$$

- ▶ Column space of A : matrix-vector multiplication $A\mathbf{x} \in \mathbf{C}(A)$.
- ▶ $\mathbf{p} = P\mathbf{b}$. What is the column space $\mathbf{C}(P)$?
- ▶ $\mathbf{C}(P)$ is the line through \mathbf{a} .
- ▶ Is P symmetric?

$$P^\top = \left(\frac{\mathbf{a}\mathbf{a}^\top}{\mathbf{a}^\top\mathbf{a}} \right)^\top = \frac{\mathbf{a}\mathbf{a}^\top}{\mathbf{a}^\top\mathbf{a}} = P. \quad \text{Yes.}$$

- ▶ What if we project \mathbf{b} twice?

$$P^2 = \left(\frac{\mathbf{a}\mathbf{a}^\top}{\mathbf{a}^\top\mathbf{a}} \right) \left(\frac{\mathbf{a}\mathbf{a}^\top}{\mathbf{a}^\top\mathbf{a}} \right) = \frac{\mathbf{a}\mathbf{a}^\top}{\mathbf{a}^\top\mathbf{a}} = P$$

Projection onto a Subspace

- ▶ Why bother with projection?

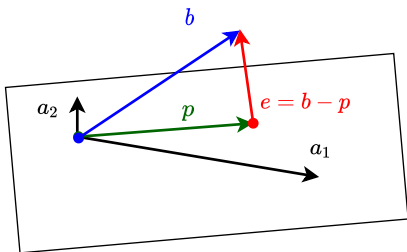
Projection onto a Subspace

- ▶ Why bother with projection?
- ▶ Because $Ax = b$ may have no solution ($m \gg n$). b might not in the column space $\mathbf{C}(A)$.

Projection onto a Subspace

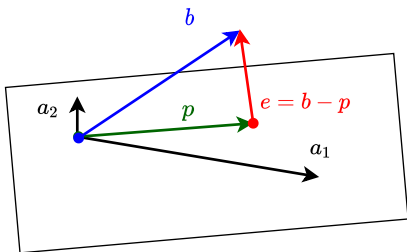
- ▶ Why bother with projection?
- ▶ Because $A\mathbf{x} = \mathbf{b}$ may have no solution ($m \gg n$). \mathbf{b} might not be in the column space $\mathbf{C}(A)$.
- ▶ Solve $A\hat{\mathbf{x}} = \mathbf{p}$ instead, where \mathbf{p} is the projection of \mathbf{b} onto the column space $\mathbf{C}(A)$.

Projection onto a Subspace



Projection onto a Subspace

- Choose two independent vectors \mathbf{a}_1 , \mathbf{a}_2 in the plane to form a basis.

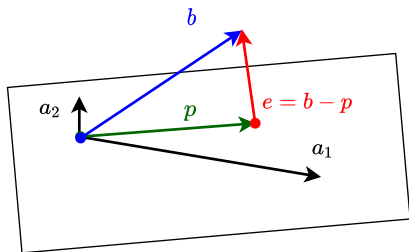


$$A = \begin{bmatrix} | & | \\ \mathbf{a}_1 & \mathbf{a}_2 \\ | & | \end{bmatrix}$$

- Plane of \mathbf{a}_1 , \mathbf{a}_2 = Column space of A .

Projection onto a Subspace

- Choose two independent vectors \mathbf{a}_1 , \mathbf{a}_2 in the plane to form a basis.



$$A = \begin{bmatrix} | & | \\ \mathbf{a}_1 & \mathbf{a}_2 \\ | & | \end{bmatrix}$$

- Plane of \mathbf{a}_1 , \mathbf{a}_2 = Column space of A .
- \mathbf{p} is a linear combination of \mathbf{a}_1 , \mathbf{a}_2 .

$$\begin{aligned} \mathbf{p} &= \hat{x}_1 \mathbf{a}_1 + \hat{x}_2 \mathbf{a}_2 \\ &= A\hat{\mathbf{x}} \end{aligned}$$

- Find $\hat{\mathbf{x}}$.

Projection onto a Subspace

► $p = A\hat{x}$. Find \hat{x} .

Projection onto a Subspace

- ▶ $\mathbf{p} = A\hat{\mathbf{x}}$. Find $\hat{\mathbf{x}}$.
- ▶ $\mathbf{e} = \mathbf{b} - \mathbf{p}$ is perpendicular to the plane.

$$\begin{bmatrix} \mathbf{a}_1^\top \\ \mathbf{a}_2^\top \end{bmatrix} \begin{bmatrix} | \\ \mathbf{e} \\ | \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$A^\top \mathbf{e} = \mathbf{0}$$

$$A^\top (\mathbf{b} - A\hat{\mathbf{x}}) = \mathbf{0}$$

$$A^\top A\hat{\mathbf{x}} = A^\top \mathbf{b}$$

$$\hat{\mathbf{x}} = (A^\top A)^{-1} A^\top \mathbf{b}$$

Projection onto a Subspace

- ▶ $\mathbf{p} = A\hat{\mathbf{x}}$. Find $\hat{\mathbf{x}}$.
- ▶ $\mathbf{e} = \mathbf{b} - \mathbf{p}$ is perpendicular to the plane.

$$\begin{bmatrix} \mathbf{a}_1^\top \\ \mathbf{a}_2^\top \end{bmatrix} \begin{bmatrix} | \\ \mathbf{e} \\ | \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$A^\top \mathbf{e} = \mathbf{0}$$

$$A^\top (\mathbf{b} - A\hat{\mathbf{x}}) = \mathbf{0}$$

$$A^\top A\hat{\mathbf{x}} = A^\top \mathbf{b}$$

$$\hat{\mathbf{x}} = (A^\top A)^{-1} A^\top \mathbf{b}$$

- ▶ We have $\mathbf{p} = A\hat{\mathbf{x}} = A(A^\top A)^{-1} A^\top \mathbf{b}$.

Projection onto a Subspace

- ▶ $\mathbf{p} = A\hat{\mathbf{x}}$. Find $\hat{\mathbf{x}}$.
- ▶ $\mathbf{e} = \mathbf{b} - \mathbf{p}$ is perpendicular to the plane.

$$\begin{bmatrix} \mathbf{a}_1^\top \\ \mathbf{a}_2^\top \end{bmatrix} \begin{bmatrix} | \\ \mathbf{e} \\ | \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$A^\top \mathbf{e} = \mathbf{0}$$

$$A^\top (\mathbf{b} - A\hat{\mathbf{x}}) = \mathbf{0}$$

$$A^\top A\hat{\mathbf{x}} = A^\top \mathbf{b}$$

$$\hat{\mathbf{x}} = (A^\top A)^{-1} A^\top \mathbf{b}$$

- ▶ We have $\mathbf{p} = A\hat{\mathbf{x}} = A(A^\top A)^{-1} A^\top \mathbf{b}$.
- ▶ The projection matrix P :

$$P = A(A^\top A)^{-1} A^\top$$

Projection onto a Subspace

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

► Is P symmetric?

Projection onto a Subspace

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

- ▶ Is P symmetric?

$$\begin{aligned} P^T &= (A(A^T A)^{-1} A^T)^T = A((A^T A)^{-1})^T A^T \\ &= A((A^T A)^T)^{-1} A^T \\ &= A(A^T A)^{-1} A^T = P \end{aligned}$$

Yes.

- ▶ Is $P^2 = P$?

Projection onto a Subspace

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

- Is P symmetric?

$$\begin{aligned} P^T &= (A(A^T A)^{-1} A^T)^T = A((A^T A)^{-1})^T A^T \\ &= A((A^T A)^T)^{-1} A^T \\ &= A(A^T A)^{-1} A^T = P \end{aligned}$$

Yes.

- Is $P^2 = P$?

$$\begin{aligned} P^2 &= A(A^T A)^{-1} A^T A(A^T A)^{-1} A^T \\ &= A(A^T A)^{-1} (A^T A) (A^T A)^{-1} A^T \\ &= A(A^T A)^{-1} A^T = P \end{aligned}$$

Yes.

Q with Orthonormal Columns

$$Q_1 = \frac{1}{3} \begin{bmatrix} 2 \\ 2 \\ -1 \end{bmatrix} \quad Q_1^T Q_1 = [1]$$

Q with Orthonormal Columns

$$Q_1 = \frac{1}{3} \begin{bmatrix} 2 \\ 2 \\ -1 \end{bmatrix} \quad Q_1^\top Q_1 = [1]$$

$$Q_2 = \frac{1}{3} \begin{bmatrix} 2 & 2 \\ 2 & -1 \\ -1 & 2 \end{bmatrix} \quad Q_2^\top Q_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Q with Orthonormal Columns

$$Q_1 = \frac{1}{3} \begin{bmatrix} 2 \\ 2 \\ -1 \end{bmatrix}$$

$$Q_1^\top Q_1 = [1]$$

$$Q_2 = \frac{1}{3} \begin{bmatrix} 2 & 2 \\ 2 & -1 \\ -1 & 2 \end{bmatrix}$$

$$Q_2^\top Q_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$Q_3 = \frac{1}{3} \begin{bmatrix} 2 & 2 & -1 \\ 2 & -1 & 2 \\ -1 & 2 & 2 \end{bmatrix}$$

$$Q_3^\top Q_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Q with Orthonormal Columns

$$Q_1 = \frac{1}{3} \begin{bmatrix} 2 \\ 2 \\ -1 \end{bmatrix} \quad Q_1^\top Q_1 = [1]$$

$$Q_2 = \frac{1}{3} \begin{bmatrix} 2 & 2 \\ 2 & -1 \\ -1 & 2 \end{bmatrix} \quad Q_2^\top Q_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$Q_3 = \frac{1}{3} \begin{bmatrix} 2 & 2 & -1 \\ 2 & -1 & 2 \\ -1 & 2 & 2 \end{bmatrix} \quad Q_3^\top Q_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

► Columns of Q 's are orthonormal.

Q with Orthonormal Columns

$$Q_1 = \frac{1}{3} \begin{bmatrix} 2 \\ 2 \\ -1 \end{bmatrix} \quad Q_1^\top Q_1 = [1]$$

$$Q_2 = \frac{1}{3} \begin{bmatrix} 2 & 2 \\ 2 & -1 \\ -1 & 2 \end{bmatrix} \quad Q_2^\top Q_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$Q_3 = \frac{1}{3} \begin{bmatrix} 2 & 2 & -1 \\ 2 & -1 & 2 \\ -1 & 2 & 2 \end{bmatrix} \quad Q_3^\top Q_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

- ▶ Columns of Q 's are orthonormal.
- ▶ Each one of those matrices has $Q^\top Q = I$.

Q with Orthonormal Columns

$$Q_1 = \frac{1}{3} \begin{bmatrix} 2 \\ 2 \\ -1 \end{bmatrix} \quad Q_1^\top Q_1 = [1]$$

$$Q_2 = \frac{1}{3} \begin{bmatrix} 2 & 2 \\ 2 & -1 \\ -1 & 2 \end{bmatrix} \quad Q_2^\top Q_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$Q_3 = \frac{1}{3} \begin{bmatrix} 2 & 2 & -1 \\ 2 & -1 & 2 \\ -1 & 2 & 2 \end{bmatrix} \quad Q_3^\top Q_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

- ▶ Columns of Q 's are orthonormal.
- ▶ Each one of those matrices has $Q^\top Q = I$.
- ▶ Q^\top is a **left inverse** of Q .

Q with Orthonormal Columns

$$Q_1 = \frac{1}{3} \begin{bmatrix} 2 \\ 2 \\ -1 \end{bmatrix} \quad Q_1^\top Q_1 = [1]$$

$$Q_2 = \frac{1}{3} \begin{bmatrix} 2 & 2 \\ 2 & -1 \\ -1 & 2 \end{bmatrix} \quad Q_2^\top Q_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$Q_3 = \frac{1}{3} \begin{bmatrix} 2 & 2 & -1 \\ 2 & -1 & 2 \\ -1 & 2 & 2 \end{bmatrix} \quad Q_3^\top Q_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

- ▶ Columns of Q 's are orthonormal.
- ▶ Each one of those matrices has $Q^\top Q = I$.
- ▶ Q^\top is a **left inverse** of Q .
- ▶ $Q_3 Q_3^\top = I$. Q_3^\top is also a **right inverse**.

Orthogonal Projection

- ▶ All the matrices $P = QQ^\top$ have $P^\top = P$.

$$P^\top = (QQ^\top)^\top = QQ^\top = P$$

Orthogonal Projection

- ▶ All the matrices $P = QQ^\top$ have $P^\top = P$.

$$P^\top = (QQ^\top)^\top = QQ^\top = P$$

- ▶ All the matrices $P = QQ^\top$ have $P^2 = P$.

$$P^2 = (QQ^\top)(QQ^\top) = Q(Q^\top Q)Q^\top = QQ^\top = P$$

Orthogonal Projection

- ▶ All the matrices $P = QQ^\top$ have $P^\top = P$.

$$P^\top = (QQ^\top)^\top = QQ^\top = P$$

- ▶ All the matrices $P = QQ^\top$ have $P^2 = P$.

$$P^2 = (QQ^\top)(QQ^\top) = Q(Q^\top Q)Q^\top = QQ^\top = P$$

- ▶ P is a **projection matrix**.

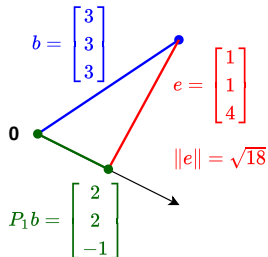
Orthogonal Projection

If $P^2 = P = P^\top$ then $P\mathbf{b}$ is the orthogonal projection of \mathbf{b} onto the column space of P .

Orthogonal Projection

- Project $\mathbf{b} = (3, 3, 3)$ on the Q_1 line. $P_1 = Q_1 Q_1^\top$

$$P_1 \mathbf{b} = \frac{1}{9} \begin{bmatrix} 2 \\ 2 \\ -1 \end{bmatrix} [2 \ 2 \ -1] \begin{bmatrix} 3 \\ 3 \\ 3 \end{bmatrix} = \frac{1}{9} \begin{bmatrix} 2 \\ 2 \\ -1 \end{bmatrix} 9 = \begin{bmatrix} 2 \\ 2 \\ -1 \end{bmatrix}$$

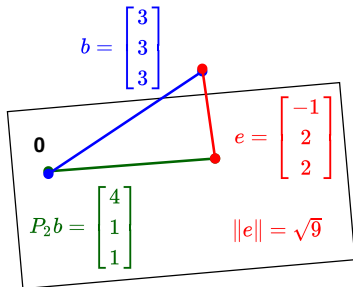


- P_1 splits \mathbf{b} in 2 perpendicular parts: projection $P_1 \mathbf{b}$ and error $\mathbf{e} = \mathbf{b} - P_1 \mathbf{b}$

Orthogonal Projection

- Project $\mathbf{b} = (3, 3, 3)$ on the Q_2 plane. $P_2 = Q_2 Q_2^\top$

$$P_2 \mathbf{b} = \frac{1}{9} \begin{bmatrix} 2 & 2 \\ 2 & -1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} 2 & 2 & -1 \\ 2 & -1 & 2 \end{bmatrix} \begin{bmatrix} 3 \\ 3 \\ 3 \end{bmatrix} = \frac{1}{9} \begin{bmatrix} 2 & 2 \\ 2 & -1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} 9 \\ 9 \\ 9 \end{bmatrix} = \begin{bmatrix} 4 \\ 1 \\ 1 \end{bmatrix}$$



- P_2 projects \mathbf{b} on the column space of Q_2 .
- The error vector $\mathbf{b} - P_2 \mathbf{b}$ is shorter than $\mathbf{b} - P_1 \mathbf{b}$.

Orthogonal Projection

$$Q_3 = \frac{1}{3} \begin{bmatrix} 2 & 2 & -1 \\ 2 & -1 & 2 \\ -1 & 2 & 2 \end{bmatrix}$$

► What is $P_3 \mathbf{b} = Q_3 Q_3^\top \mathbf{b}$?

Orthogonal Projection

$$Q_3 = \frac{1}{3} \begin{bmatrix} 2 & 2 & -1 \\ 2 & -1 & 2 \\ -1 & 2 & 2 \end{bmatrix}$$

- ▶ What is $P_3 \mathbf{b} = Q_3 Q_3^\top \mathbf{b}$?
- ▶ Project \mathbf{b} onto the whole space \mathbb{R}^3 .

Orthogonal Projection

$$Q_3 = \frac{1}{3} \begin{bmatrix} 2 & 2 & -1 \\ 2 & -1 & 2 \\ -1 & 2 & 2 \end{bmatrix}$$

- ▶ What is $P_3 \mathbf{b} = Q_3 Q_3^\top \mathbf{b}$?
- ▶ Project \mathbf{b} onto the whole space \mathbb{R}^3 .
- ▶ $P_3 = Q_3 Q_3^\top = I$. Thus, $P_3 \mathbf{b} = \mathbf{b}$. Vector \mathbf{b} is in \mathbb{R}^3 already.

Orthogonal Projection

$$Q_3 = \frac{1}{3} \begin{bmatrix} 2 & 2 & -1 \\ 2 & -1 & 2 \\ -1 & 2 & 2 \end{bmatrix}$$

- ▶ What is $P_3 \mathbf{b} = Q_3 Q_3^\top \mathbf{b}$?
- ▶ Project \mathbf{b} onto the whole space \mathbb{R}^3 .
- ▶ $P_3 = Q_3 Q_3^\top = I$. Thus, $P_3 \mathbf{b} = \mathbf{b}$. Vector \mathbf{b} is in \mathbb{R}^3 already.
- ▶ The error e is **zero!!!**

Orthogonalization

- ▶ Determine if a list of vectors $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_k$ is linearly independent.

Orthogonalization

- Determine if a list of vectors $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_k$ is linearly independent.

Gram-Smidt algorithm

given vectors $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_k$

for $i = 1, \dots, k$

- 1 Orthogonalization. $\tilde{\mathbf{q}}_i = \mathbf{a}_i - (\mathbf{q}_1^T \mathbf{a}_i) \mathbf{q}_1 - \dots - (\mathbf{q}_{i-1}^T \mathbf{a}_i) \mathbf{q}_{i-1}$
- 2 Test for linear dependence. If $\tilde{\mathbf{q}}_i = 0$, quit.
- 3 Normalization. $\mathbf{q}_i = \tilde{\mathbf{q}}_i / \|\tilde{\mathbf{q}}_i\|$

Orthogonalization

- Determine if a list of vectors $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_k$ is linearly independent.

Gram-Smidt algorithm

given vectors $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_k$

for $i = 1, \dots, k$

- 1 Orthogonalization. $\tilde{\mathbf{q}}_i = \mathbf{a}_i - (\mathbf{q}_1^T \mathbf{a}_i) \mathbf{q}_1 - \dots - (\mathbf{q}_{i-1}^T \mathbf{a}_i) \mathbf{q}_{i-1}$
- 2 Test for linear dependence. If $\tilde{\mathbf{q}}_i = 0$, quit.
- 3 Normalization. $\mathbf{q}_i = \tilde{\mathbf{q}}_i / \|\tilde{\mathbf{q}}_i\|$

- If the vectors are **linearly independent**, the Gram-Smidt algorithm produces an **orthonormal** collection of vectors $\mathbf{q}_1, \dots, \mathbf{q}_k$.

Orthogonalization

- ▶ Determine if a list of vectors $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_k$ is linearly independent.

Gram-Smidt algorithm

given vectors $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_k$

for $i = 1, \dots, k$

- 1 Orthogonalization. $\tilde{\mathbf{q}}_i = \mathbf{a}_i - (\mathbf{q}_1^T \mathbf{a}_i) \mathbf{q}_1 - \dots - (\mathbf{q}_{i-1}^T \mathbf{a}_i) \mathbf{q}_{i-1}$
- 2 Test for linear dependence. If $\tilde{\mathbf{q}}_i = 0$, quit.
- 3 Normalization. $\mathbf{q}_i = \tilde{\mathbf{q}}_i / \|\tilde{\mathbf{q}}_i\|$

- ▶ If the vectors are **linearly independent**, the Gram-Smidt algorithm produces an **orthonormal** collection of vectors $\mathbf{q}_1, \dots, \mathbf{q}_k$.
- ▶ If the vectors $\mathbf{a}_1, \dots, \mathbf{a}_{j-1}$ are linearly independent, but $\mathbf{a}_1, \dots, \mathbf{a}_j$ are linearly dependent, the algorithm detects this and terminates.

Orthogonalization: Example

$$\mathbf{a}_1 = (-1, 1, -1, 1), \quad \mathbf{a}_2 = (-1, 3, -1, 3), \quad \mathbf{a}_3 = (1, 3, 5, 7)$$

Applying the Gram-Smidt algorithm gives the following results.

► $i = 1$:

$$\tilde{\mathbf{q}}_1 = \mathbf{a}_1$$

$$\mathbf{q}_1 = \frac{1}{\|\tilde{\mathbf{q}}_1\|} \tilde{\mathbf{q}}_1 = (-1/2, 1/2, -1/2, 1/2)$$

► $i = 2$:

$$\begin{aligned} \tilde{\mathbf{q}}_2 &= \mathbf{a}_2 - (\mathbf{q}_1^T \mathbf{a}_2) \mathbf{q}_1 \\ &= (-1, 3, -1, 3) - 4(-1/2, 1/2, -1/2, 1/2) = (1, 1, 1, 1) \end{aligned}$$

$$\mathbf{q}_2 = \frac{1}{\|\tilde{\mathbf{q}}_2\|} \tilde{\mathbf{q}}_2 = (1/2, 1/2, 1/2, 1/2)$$

Orthogonalization: Example

► $i = 3$:

$$\begin{aligned}\tilde{\mathbf{q}}_3 &= \mathbf{a}_3 - (\mathbf{q}_1^T \mathbf{a}_3) \mathbf{q}_1 - (\mathbf{q}_2^T \mathbf{a}_3) \mathbf{q}_2 \\ &= \begin{bmatrix} 1 \\ 3 \\ 5 \\ 7 \end{bmatrix} - 2 \begin{bmatrix} -1/2 \\ 1/2 \\ -1/2 \\ 1/2 \end{bmatrix} - 8 \begin{bmatrix} 1/2 \\ 1/2 \\ 1/2 \\ 1/2 \end{bmatrix} = \begin{bmatrix} -2 \\ -2 \\ 2 \\ 2 \end{bmatrix} \\ \mathbf{q}_3 &= \frac{1}{\|\tilde{\mathbf{q}}_3\|} \tilde{\mathbf{q}}_3 = (-1/2, -1/2, 1/2, 1/2)\end{aligned}$$

► The completion of the Gram-Smidt algorithm without early termination indicates that the vectors $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ are linearly independent.

QR factorization: $A = QR$

$$A = QR$$

$$\begin{bmatrix} | & | & \dots & | \\ \mathbf{a}_1 & \mathbf{a}_2 & \dots & \mathbf{a}_n \\ | & | & \dots & | \end{bmatrix} = \begin{bmatrix} | & | & \dots & | \\ \mathbf{q}_1 & \mathbf{q}_2 & \dots & \mathbf{q}_n \\ | & | & \dots & | \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & \dots & r_{1n} \\ 0 & r_{22} & \dots & r_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & r_{nn} \end{bmatrix}$$

$$r_{kk} = \|\tilde{\mathbf{q}}_k\|$$

$$r_{k-1,k} = \mathbf{q}_{k-1}^\top \mathbf{a}_k$$

QR factorization: $A = QR$

$$\begin{aligned}\hat{\mathbf{x}} &= (A^T A)^{-1} A^T \mathbf{b} \\ &= ((QR)^T (QR))^{-1} (QR)^T \mathbf{b} \\ &= (R^T Q^T Q R)^{-1} R^T Q^T \mathbf{b} \\ &= (R^T R)^{-1} R^T Q^T \mathbf{b} \quad (\text{because } Q^T Q = I) \\ &= R^{-1} R^{-T} R^T Q^T \mathbf{b} \\ &= R^{-1} Q^T \mathbf{b}\end{aligned}$$

QR factorization: $A = QR$

$$\begin{aligned}\hat{\mathbf{x}} &= (A^T A)^{-1} A^T \mathbf{b} \\ &= ((QR)^T (QR))^{-1} (QR)^T \mathbf{b} \\ &= (R^T Q^T Q R)^{-1} R^T Q^T \mathbf{b} \\ &= (R^T R)^{-1} R^T Q^T \mathbf{b} \quad (\text{because } Q^T Q = I) \\ &= R^{-1} R^{-T} R^T Q^T \mathbf{b} \\ &= R^{-1} Q^T \mathbf{b}\end{aligned}$$

Solving for $\hat{\mathbf{x}}$ by solving $R\hat{\mathbf{x}} = Q^T \mathbf{b}$ with back-substitution.