Section 1.1 : Systems of Linear Equations

Chapter 1 : Linear Equations

Math 1554 Linear Algebra

Section 1.1 Systems of Linear Equations

Topics

We will cover these topics in this section.

1. Systems of Linear Equations
2. Matrix Notation
3. Elementary Row Operations
4. Questions of Existence and Uniqueness of Solutions

Objectives

For the topics covered in this section, students are expected to be able to do the following.

1. Characterize a linear system in terms of the number of solutions, and whether the system is consistent or inconsistent.
2. Apply elementary row operations to solve linear systems of equations.
3. Express a set of linear equations as an augmented matrix.

A Single Linear Equation

A linear equation has the form

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b$$

$a_1, \ldots, a_n$ and b are the **coefficients**, $x_1, \ldots, x_n$ are the **variables** or **unknowns**, and n is the **dimension**, or number of variables.

For example,

- $2x_1 + 4x_2 = 4$ is a line in two dimensions
- $3x_1 + 2x_2 + x_3 = 6$ is a plane in three dimensions

Systems of Linear Equations

When we have more than one linear equation, we have a **linear system** of equations. For example, a linear system with two equations is

$$\begin{array}{cccc} x_1 & + 1.5x_2 & + \pi x_3 & = 4 \\ 5x_1 & & + 7x_3 & = 5 \end{array}$$

Definition: Solution to a Linear System

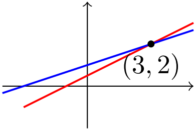
The set of all possible values of $x_1, x_2, \ldots x_n$ that satisfy all equations is the **solution** to the system.

A system can have a unique solution, no solution, or an infinite number of solutions.

Two Variables

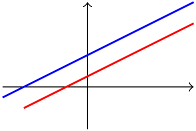
Consider the following systems. How are they different from each other?

$$\begin{aligned} x_1 - 2x_2 &= -1 \\ -x_1 + 3x_2 &= 3 \end{aligned}$$



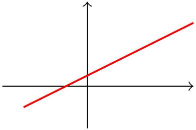
non-parallel lines

$$\begin{aligned} x_1 - 2x_2 &= -1 \\ -x_1 + 2x_2 &= 3 \end{aligned}$$



parallel lines

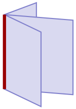
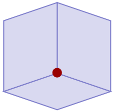
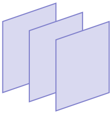
$$\begin{aligned} x_1 - 2x_2 &= -1 \\ -x_1 + 2x_2 &= 1 \end{aligned}$$



identical lines

Three-Dimensional Case

An equation $a_1x_1 + a_2x_2 + a_3x_3 = b$ defines a plane in $\mathbb{R}^3$. The **solution** to a system of **three equations** is the set of intersections of the planes.

solution set	**sketch**	**number of solutions**
line		
point		
empty		

Row Reduction by Elementary Row Operations

How can we find the solution set to a set of linear equations?
We can manipulate equations in a linear system using **row operations**.

1. (Replacement/Addition) Add a multiple of one row to another.
2. (Interchange) Interchange two rows.
3. (Scaling) Multiply a row by a non-zero scalar.

Let's use these operations to solve a system of equations.

Example 1

Identify the solution to the linear system.

$$\begin{array}{rrrl} x_1 & -2x_2 & +x_3 & = 0 \\ & 2x_2 & -8x_3 & = 8 \\ 5x_1 & & -5x_3 & = 10 \end{array}$$

Augmented Matrices

It is redundant to write x_1, x_2, x_3 again and again, so we rewrite systems using matrices. For example,

$$\begin{array}{rrrl} x_1 & -2x_2 & +x_3 & = 0 \\ & 2x_2 & -8x_3 & = 8 \\ 5x_1 & & -5x_3 & = 10 \end{array}$$

can be written as the **augmented matrix**,

$$\left[\begin{array}{ccc|c} 1 & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ 5 & 0 & -5 & 10 \end{array}\right]$$

The vertical line reminds us that the first three columns are the coefficients to our variables x_1, x_2, and x_3.

Consistent Systems and Row Equivalence

Definition (Consistent)

A linear system is **consistent** if it has at least one ___________.

Definition (Row Equivalence)

Two matrices are **row equivalent** if a sequence of ___________

______________________ transforms one matrix into the other.

Note: if the augmented matrices of two linear systems are row equivalent, then they have the same solution set.

Fundamental Questions

Two questions that we will revisit many times throughout our course.

1. Does a given linear system have a solution? In other words, is it consistent?
2. If it is consistent, is the solution unique?

Section 1.2 : Row Reduction and Echelon Forms

Chapter 1 : Linear Equations

Math 1554 Linear Algebra

Section 1.2 : Row Reductions and Echelon Forms

Topics

We will cover these topics in this section.

1. Row reduction algorithm
2. Pivots, and basic and free variables
3. Echelon forms, existence and uniqueness

Objectives

For the topics covered in this section, students are expected to be able to do the following.

1. Characterize a linear system in terms of the number of leading entries, free variables, pivots, pivot columns, pivot positions.
2. Apply the row reduction algorithm to reduce a linear system to echelon form, or reduced echelon form.
3. Apply the row reduction algorithm to compute the coefficients of a polynomial.

Definition: Echelon Form and RREF

A rectangular matrix is in **echelon form** if

1. All zero rows (if any are present) are at the bottom.
2. The first non-zero entry (or **leading entry**) of a row is to the right of any leading entries in the row above it (if any).
3. All elements below a leading entry (if any) are zero.

A matrix in echelon form is in **row reduced echelon form** (RREF) if

1. All leading entries, if any, are equal to 1.
2. Leading entries are the only nonzero entry in their respective column.

Example of a Matrix in Echelon Form

$\blacksquare$ = non-zero number, $\quad * =$ any number

$$\begin{bmatrix} 0 & \blacksquare & * & * & * & * & * & * & * & * \\ 0 & 0 & 0 & \blacksquare & * & * & * & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \blacksquare & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \blacksquare & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Example 1

Which of the following are in RREF?

$$a)\quad \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \qquad\qquad d)\quad \begin{bmatrix} 0 & 6 & 3 & 0 \end{bmatrix}$$

$$b)\quad \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \qquad\qquad e)\quad \begin{bmatrix} 1 & 17 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$c)\quad \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

Definition: Pivot Position, Pivot Column

A **pivot position** in a matrix A is a location in A that corresponds to a leading 1 in the reduced echelon form of A.

A **pivot column** is a column of A that contains a pivot position.

Example 2: Express the matrix in row reduced echelon form and identify the pivot columns.

$$\begin{bmatrix} 0 & -3 & -6 & 4 \\ -1 & -2 & -1 & 3 \\ -2 & -3 & 0 & 3 \end{bmatrix}$$

Row Reduction Algorithm

The algorithm we used in the previous example produces a matrix in RREF. Its steps can be stated as follows.

- **Step 1a** Swap the 1st row with a lower one so the leftmost nonzero entry is in the 1st row
- **Step 1b** Scale the 1st row so that its leading entry is equal to 1
- **Step 1c** Use row replacement so all entries below this 1 are 0
- **Step 2a** Swap the 2nd row with a lower one so that the leftmost nonzero entry below 1st row is in the 2nd row
- **etc.** ...

  Now the matrix is in echelon form, with leading entries equal to 1.
- **Last step** Use row replacement so all entries above each leading entry are 0, starting from the right.

Basic And Free Variables

Consider the augmented matrix

$$\left[A \mid \vec{b}\right] = \left[\begin{array}{ccccc|c} 1 & 3 & 0 & 7 & 0 & 4 \\ 0 & 0 & 1 & 4 & 0 & 5 \\ 0 & 0 & 0 & 0 & 1 & 6 \end{array}\right]$$

The leading one's are in first, third, and fifth columns. So:

- the pivot variables of the system $A\vec{x} = \vec{b}$ are x_1, x_3, and x_5.
- The free variables are x_2 and x_4. **Any choice** of the free variables leads to a solution of the system.

Note that A does not have basic variables or free variables. Systems have variables.

Existence and Uniqueness

Theorem

A linear system is consistent if and only if (exactly when) the last column of the **augmented** matrix does not have a pivot. This is the same as saying that the RREF of the augmented matrix does **not** have a row of the form

$$\left(\begin{array}{ccccc|c} 0 & 0 & 0 & \cdots & 0 & 1 \end{array}\right)$$

Moreover, if a linear system is consistent, then it has

1. a unique solution if and only if there are no free variables.
2. infinitely many solutions that are parameterized by free variables.

Section 1.3 : Vector Equations

Chapter 1 : Linear Equations

Math 1554 Linear Algebra

1.3: Vector Equations

Topics

We will cover these topics in this section.

1. Vectors in $\mathbb{R}^n$, and their basic properties
2. Linear combinations of vectors

Objectives

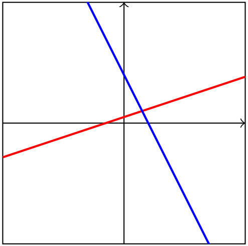
For the topics covered in this section, students are expected to be able to do the following.

1. Apply geometric and algebraic properties of vectors in $\mathbb{R}^n$ to compute vector additions and scalar multiplications.
2. Characterize a set of vectors in terms of **linear combinations**, their **span**, and how they are related to each other geometrically.

Motivation

We want to think about the **algebra** in linear algebra (systems of equations and their solution sets) in terms of **geometry** (points, lines, planes, etc).

$$x - 3y = -3$$
$$2x + y = 8$$



- This will give us better insight into the properties of systems of equations and their solution sets.
- To do this, we need to introduce n-dimensional space $\mathbb{R}^n$, and **vectors** inside it.

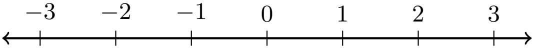
$\mathbb{R}^n$

Recall that $\mathbb{R}$ denotes the collection of all real numbers.

Let n be a positive whole number. We define

$$\mathbb{R}^n = \text{all ordered } n\text{-tuples of real numbers } (x_1, x_2, x_3, \ldots, x_n).$$

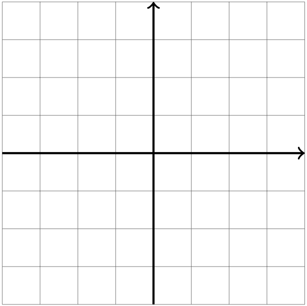
When $n = 1$, we get $\mathbb{R}$ back: $\mathbb{R}^1 = \mathbb{R}$. Geometrically, this is the **number line**.



Note that:

- when $n = 2$, we can think of $\mathbb{R}^2$ as a **plane**
- every point in this plane can be represented by an ordered pair of real numbers, its x- and y-coordinates

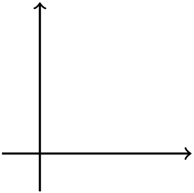
Example: Sketch the point $(3, 2)$ and the vector $\begin{pmatrix} 3 \\ 2 \end{pmatrix}$.



Vectors

In the previous slides, we were thinking of elements of $\mathbb{R}^n$ as **points**: in the line, plane, space, etc.

We can also think of them as **vectors**: arrows with a given length and direction.



For example, the vector $\begin{pmatrix} 3 \\ 2 \end{pmatrix}$ points **horizontally** in the amount of its x-coordinate, and **vertically** in the amount of its y-coordinate.

Vector Algebra

When we think of an element of $\mathbb{R}^n$ as a vector, we write it as a matrix with n rows and one column:

$$\vec{v} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$$

Suppose

$$\vec{u} = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}, \quad \vec{v} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}.$$

Vectors have the following properties.

1. **Scalar Multiple**:

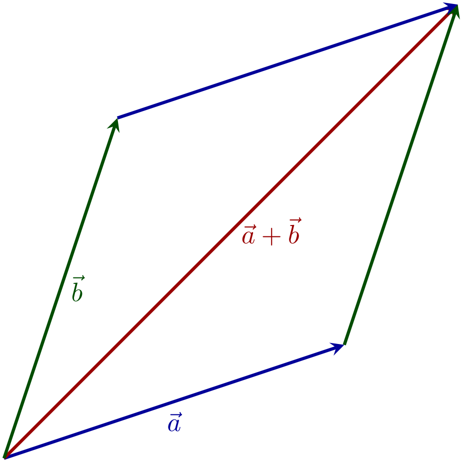
$$c\vec{u} =$$

2. **Vector Addition**:

$$\vec{u} + \vec{v} =$$

Note that vectors in higher dimensions have the same properties.

Parallelogram Rule for Vector Addition



Linear Combinations and Span

Definition

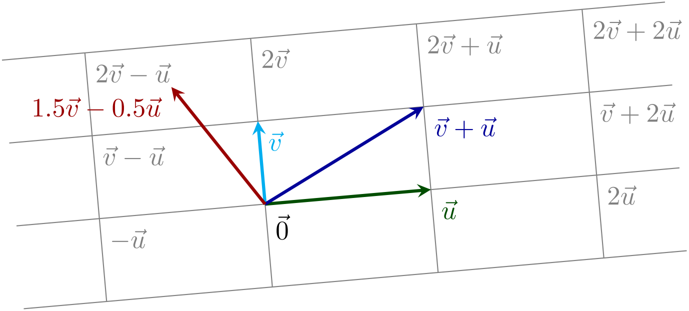
1. Given vectors $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_p \in \mathbb{R}^n$, and scalars $c_1, c_2, \dots, c_p$, the vector below

$$\vec{y} = c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_p\vec{v}_p$$

   is called a **linear combination of $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_p$ with weights $c_1, c_2, \dots, c_p$.**
2. The set of all linear combinations of $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_p$ is called the **Span** of $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_p$.

Geometric Interpretation of Linear Combinations

Note that any two vectors in $\mathbb{R}^2$ that are not scalar multiples of each other, span $\mathbb{R}^2$. In other words, any vector in $\mathbb{R}^2$ can be represented as a linear combination of two vectors that are not multiples of each other.



Example

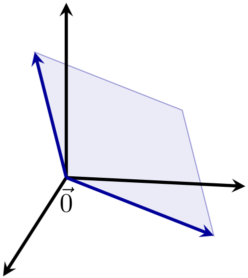
Is $\vec{y}$ in the span of vectors $\vec{v}_1$ and $\vec{v}_2$?

$$\vec{v}_1 = \begin{pmatrix} 1 \\ -2 \\ -3 \end{pmatrix}, \ \vec{v}_2 = \begin{pmatrix} 2 \\ 5 \\ 6 \end{pmatrix}, \text{ and } \vec{y} = \begin{pmatrix} 7 \\ 4 \\ 15 \end{pmatrix}.$$

The Span of Two Vectors in $\mathbb{R}^3$

In the previous example, did we find that $\vec{y}$ is in the span of $\vec{v}_1$ and $\vec{v}_2$?

In general: Any two non-parallel vectors in $\mathbb{R}^3$ span a plane that passes through the origin. Any vector in that plane is also in the span of the two vectors.



Section 1.4 : The Matrix Equation

Chapter 1 : Linear Equations

Math 1554 Linear Algebra

"Mathematics is the art of giving the same name to different things."
- H. Poincaré

In this section we introduce another way of expressing a linear system that we will use throughout this course.

1.4 : Matrix Equation $A\vec{x} = \vec{b}$

Topics

We will cover these topics in this section.

1. Matrix notation for systems of equations.
2. The matrix product $A\vec{x}$.

Objectives

For the topics covered in this section, students are expected to be able to do the following.

1. Compute matrix-vector products.
2. Express linear systems as vector equations and matrix equations.
3. Characterize linear systems and sets of vectors using the concepts of span, linear combinations, and pivots.

Notation

symbol	meaning
$\in$	belongs to
$\mathbb{R}^n$	the set of vectors with n real-valued elements
$\mathbb{R}^{m \times n}$	the set of real-valued matrices with m rows and n columns

Example: the notation $\vec{x} \in \mathbb{R}^5$ means that $\vec{x}$ is a vector with five real-valued elements.

Linear Combinations

Definition

A is a $m \times n$ matrix with columns $\vec{a}_1, \dots, \vec{a}_n$ and $x \in \mathbb{R}^n$, then the **matrix vector product** $A\vec{x}$ is a linear combination of the columns of A:

$$A\vec{x} = \begin{bmatrix} | & | & \cdots & | \\ \vec{a}_1 & \vec{a}_2 & \cdots & \vec{a}_n \\ | & | & \cdots & | \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = x_1\vec{a}_1 + x_2\vec{a}_2 + \cdots + x_n\vec{a}_n$$

Note that $A\vec{x}$ is in the span of the columns of A.

Example

The following product can be written as a linear combination of vectors:

$$\begin{bmatrix} 1 & 0 & -1 \\ 0 & -3 & 3 \end{bmatrix} \begin{bmatrix} 4 \\ 3 \\ 7 \end{bmatrix} =$$

Solution Sets

Theorem

If A is a $m \times n$ matrix with columns $\vec{a}_1, \ldots, \vec{a}_n$, and $x \in \mathbb{R}^n$ and $\vec{b} \in \mathbb{R}^m$, then the solutions to

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

has the same set of solutions as the vector equation

$$x_1\vec{a}_1 + \cdots + x_n\vec{a}_n = \vec{b}$$

which as the same set of solutions as the set of linear equations with the augmented matrix

$$\begin{bmatrix} \vec{a}_1 & \vec{a}_2 & \cdots & \vec{a}_n & \vec{b} \end{bmatrix}$$

Existence of Solutions

Theorem

The equation $A\vec{x} = \vec{b}$ has a solution if and only if $\vec{b}$ is a linear combination of the columns of A.

Example

For what vectors $\vec{b} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$ does the equation have a solution?

$$\begin{pmatrix} 1 & 3 & 4 \\ 2 & 8 & 4 \\ 0 & 1 & -2 \end{pmatrix} \vec{x} = \vec{b}$$

The Row Vector Rule for Computing $A\vec{x}$

$$\begin{bmatrix} 1 & 0 & 2 & 0 & 3 \\ 0 & 1 & 0 & 2 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} \quad \\ \quad \end{bmatrix}$$

Summary

We now have four **equivalent** ways of expressing linear systems.

1. A system of equations:

$$\begin{aligned} 2x_1 + 3x_2 &= 7 \\ x_1 - x_2 &= 5 \end{aligned}$$

2. An augmented matrix:

$$\left[\begin{array}{cc|c} 2 & 3 & 7 \\ 1 & -1 & 5 \end{array}\right]$$

3. A vector equation:

$$x_1 \begin{pmatrix} 2 \\ 1 \end{pmatrix} + x_2 \begin{pmatrix} 3 \\ -1 \end{pmatrix} = \begin{pmatrix} 7 \\ 5 \end{pmatrix}$$

4. As a matrix equation:

$$\begin{pmatrix} 2 & 3 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 7 \\ 5 \end{pmatrix}$$

Each representation gives us a different way to think about linear systems.

Section 1.5 : Solution Sets of Linear Systems

Chapter 1 : Linear Equations

Math 1554 Linear Algebra

1.5 : Solution Sets of Linear Systems

Topics

We will cover these topics in this section.

1. Homogeneous systems
2. Parametric **vector** forms of solutions to linear systems

Objectives

For the topics covered in this section, students are expected to be able to do the following.

1. Express the solution set of a linear system in parametric vector form.
2. Provide a geometric interpretation to the solution set of a linear system.
3. Characterize homogeneous linear systems using the concepts of free variables, span, pivots, linear combinations, and echelon forms.

Definition

Linear systems of the form $A\vec{x} = \vec{0}$ are **homogeneous**.

Linear systems of the form $A\vec{x} \neq \vec{0}$ are **inhomogeneous**.

Because homogeneous systems always have the **trivial solution**, $\vec{x} = \vec{0}$, the interesting question is whether they have non-trivial solutions.

Observation

$A\vec{x} = \vec{0}$ has a nontrivial solution

$\iff$ there is a free variable

$\iff$ A has a column with no pivot.

Example: a Homogeneous System

Identify the free variables, and the solution set, of the system.

$$\begin{aligned} x_1 + 3x_2 + x_3 &= 0 \\ 2x_1 - x_2 - 5x_3 &= 0 \\ x_1 - 2x_3 &= 0 \end{aligned}$$

Parametric Forms, Homogeneous Case

In the example on the previous slide we expressed the solution to a system using a vector equation. This is a **parametric form** of the solution.

In general, suppose the free variables for $A\vec{x} = \vec{0}$ are $x_k, \ldots, x_n$. Then all solutions to $A\vec{x} = \vec{0}$ can be written as

$$\vec{x} = x_k \vec{v}_k + x_{k+1} \vec{v}_{k+1} + \cdots + x_n \vec{v}_n$$

for some $\vec{v}_k, \ldots, \vec{v}_n$. This is the **parametric form** of the solution.

Example 2 (non-homogeneous system)

Write the parametric vector form of the solution, and give a geometric interpretation of the solution.

$$\begin{aligned} x_1 + 3x_2 + x_3 &= 9 \\ 2x_1 - x_2 - 5x_3 &= 11 \\ x_1 - 2x_3 &= 6 \end{aligned}$$

(Note that the left-hand side is the same as Example 1).

Section 1.7 : Linear Independence

Chapter 1 : Linear Equations

Math 1554 Linear Algebra

1.7 : Linear Independence

Topics

We will cover these topics in this section.

- Linear independence
- Geometric interpretation of linearly independent vectors

Objectives

For the topics covered in this section, students are expected to be able to do the following.

1. Characterize a set of vectors and linear systems using the concept of linear independence.
2. Construct dependence relations between linearly dependent vectors.

Motivating Question

What is the smallest number of vectors needed in a parametric solution to a linear system?

Linear Independence

A set of vectors $\{\vec{v}_1, \ldots, \vec{v}_k\}$ in $\mathbb{R}^n$ are **linearly independent** if

$$c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_k\vec{v}_k = \vec{0}$$

has only the **trivial** solution. It is **linearly dependent** otherwise.

In other words, $\{\vec{v}_1, \ldots, \vec{v}_k\}$ are linearly dependent if there are real numbers $c_1, c_2, \ldots, c_k$ **not all zero** so that

$$c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_k\vec{v}_k = \vec{0}$$

Consider the vectors:

$$\vec{v}_1, \vec{v}_2, \dots \vec{v}_k$$

To determine whether the vectors are linearly independent, we can set the linear combination to the zero vector:

$$c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_k\vec{v}_k = \begin{bmatrix} \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_k \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} = V\vec{c} \overset{??}{=} \vec{0}$$

Linear independence: There is NO non-zero solution $\vec{c}$

Linear dependence: There is a non-zero solution $\vec{c}$.

Example 1

For what values of h are the vectors linearly independent?

$$\begin{bmatrix} 1 \\ 1 \\ h \end{bmatrix}, \begin{bmatrix} 1 \\ h \\ 1 \end{bmatrix}, \begin{bmatrix} h \\ 1 \\ 1 \end{bmatrix}$$

Example 2 (One Vector)

Suppose $\vec{v} \in \mathbb{R}^n$. When is the set $\{\vec{v}\}$ linearly dependent?

Example 3 (Two Vectors)

Suppose $\vec{v}_1, \vec{v}_2 \in \mathbb{R}^n$. When is the set $\{\vec{v}_1, \vec{v}_2\}$ linearly dependent? Provide a geometric interpretation.

Two Theorems

Fact 1. Suppose $\vec{v}_1, \ldots, \vec{v}_k$ are vectors in $\mathbb{R}^n$. If $k > n$, then $\{\vec{v}_1, \ldots, \vec{v}_k\}$ is linearly dependent.

Fact 2. If any one or more of $\vec{v}_1, \ldots, \vec{v}_k$ is $\vec{0}$, then $\{\vec{v}_1, \ldots, \vec{v}_k\}$ is linearly dependent.

Section 1.8 : An Introduction to Linear Transforms

Chapter 1 : Linear Equations

Math 1554 Linear Algebra

1.8 : An Introduction to Linear Transforms

Topics

We will cover these topics in this section.

1. The definition of a linear transformation.
2. The interpretation of matrix multiplication as a linear transformation.

Objectives

For the topics covered in this section, students are expected to be able to do the following.

1. Construct and interpret linear transformations in $\mathbb{R}^n$ (for example, interpret a linear transform as a projection, or as a shear).
2. Characterize linear transforms using the concepts of
   - existence and uniqueness
   - domain, co-domain and range

From Matrices to Functions

Let A be an $m \times n$ matrix. We define a function

$$T : \mathbb{R}^n \to \mathbb{R}^m, \quad T(\vec{v}) = A\vec{v}$$

This is called a **matrix transformation**.

- The **domain** of T is $\mathbb{R}^n$.
- The **co-domain** or **target** of T is $\mathbb{R}^m$.
- The vector $T(\vec{x})$ is the **image** of $\vec{x}$ under T
- The set of all possible images $T(\vec{x})$ is the **range**.

This gives us **another** interpretation of $A\vec{x} = \vec{b}$:

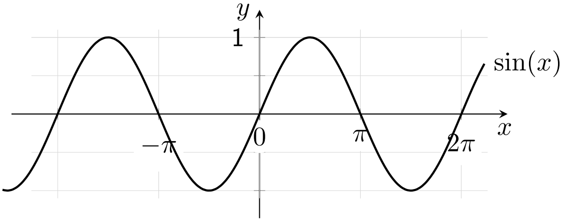
- set of equations
- augmented matrix
- matrix equation
- vector equation
- linear transformation equation

Functions from Calculus

Many of the functions we know have **domain** and **codomain** $\mathbb{R}$.We can express the **rule** that defines the function $\sin$ this way:

$$f\colon \mathbb{R} \to \mathbb{R} \qquad f(x) = \sin(x)$$

In calculus we often think of a function in terms of its graph, whose horizontal axis is the **domain**, and the vertical axis is the **codomain**.



This is ok when the domain and codomain are $\mathbb{R}$. It's hard to do when the domain is $\mathbb{R}^2$ and the codomain is $\mathbb{R}^3$. We would need five dimensions to draw that graph.

Example 1

Let $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 1 \end{bmatrix}$, $\vec{u} = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$, $\vec{b} = \begin{bmatrix} 7 \\ 5 \\ 7 \end{bmatrix}$.

a) Compute $T(\vec{u})$.

b) Calculate $\vec{v} \in \mathbb{R}^2$ so that $T(\vec{v}) = \vec{b}$

c) Give a $\vec{c} \in \mathbb{R}^3$ so there is no $\vec{v}$ with $T(\vec{v}) = \vec{c}$

or: Give a $\vec{c}$ that is not in the range of T.

or: Give a $\vec{c}$ that is not in the span of the columns of A.

Linear Transformations

A function $T : \mathbb{R}^n \to \mathbb{R}^m$ is **linear** if

- $T(\vec{u} + \vec{v}) = T(\vec{u}) + T(\vec{v})$ for all $\vec{u}, \vec{v}$ in $\mathbb{R}^n$.
- $T(c\vec{v}) = cT(\vec{v})$ for all $\vec{v} \in \mathbb{R}^n$, and c in $\mathbb{R}$.

So if T is linear, then

$$T(c_1\vec{v}_1 + \cdots + c_k\vec{v}_k) = c_1T(\vec{v}_1) + \cdots + c_kT(\vec{v}_k)$$

This is called the **principle of superposition**. The idea is that if we know $T(\vec{e}_1), \ldots, T(\vec{e}_n)$, then we know every $T(\vec{v})$.

Fact: Every matrix transformation T_A is linear.

Example 2

Suppose T is the linear transformation $T(\vec{x}) = A\vec{x}$. Give a short geometric interpretation of what $T(\vec{x})$ does to vectors in $\mathbb{R}^2$.

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

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

3) $A = \begin{bmatrix} k & 0 \\ 0 & k \end{bmatrix}$ for $k \in \mathbb{R}$

Example 3

What does T_A do to vectors in $\mathbb{R}^3$?

a) $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

b) $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

Example 4

A linear transformation $T \ : \ \mathbb{R}^2 \mapsto \mathbb{R}^3$ satisfies

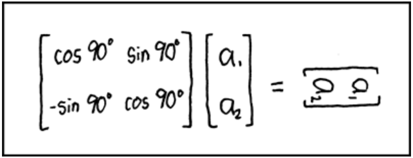
$$T\left(\begin{bmatrix}1\\0\end{bmatrix}\right) = \begin{bmatrix}5\\-7\\2\end{bmatrix}, \qquad T\left(\begin{bmatrix}0\\1\end{bmatrix}\right) = \begin{bmatrix}-3\\8\\0\end{bmatrix}$$

What is the matrix that represents T?

Section 1.9 : Linear Transforms

Chapter 1 : Linear Equations

Math 1554 Linear Algebra



https://xkcd.com/184

1.9 : Matrix of a Linear Transformation

Topics

We will cover these topics in this section.

1. The **standard vectors** and the **standard matrix**.
2. Two and three dimensional transformations in more detail.
3. **Onto** and **one-to-one** transformations.

Objectives

For the topics covered in this section, students are expected to be able to do the following.

1. Identify and construct linear transformations of a matrix.
2. Characterize linear transformations as onto and/or one-to-one.
3. Solve linear systems represented as linear transforms.
4. Express linear transforms in other forms, such as as matrix equations or as vector equations.

Definition: The Standard Vectors

The **standard vectors** in $\mathbb{R}^n$ are the vectors $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n$, where:

$$\vec{e}_1 = \qquad \vec{e}_2 = \qquad \vec{e}_n =$$

For example, in $\mathbb{R}^3$,

$$\vec{e}_1 = \qquad \vec{e}_2 = \qquad \vec{e}_3 =$$

A Property of the Standard Vectors

Note: if A is an $m \times n$ matrix with columns $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$, then

$$A\vec{e}_i = \vec{v}_i, \text{ for } i = 1, 2, \ldots, n$$

So multiplying a matrix by $\vec{e}_i$ gives column i of A.

Example

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \vec{e}_2 =$$

The Standard Matrix

Theorem

Let $T : \mathbb{R}^n \mapsto \mathbb{R}^m$ be a linear transformation. Then there is a unique matrix A such that

$$T(\vec{x}) = A\vec{x}, \qquad \vec{x} \in \mathbb{R}^n.$$

In fact, A is a $m \times n$, and its j^{th} column is the vector $T(\vec{e}_j)$.

$$A = \begin{bmatrix} T(\vec{e}_1) & T(\vec{e}_2) & \cdots & T(\vec{e}_n) \end{bmatrix}$$

The matrix A is the **standard matrix** for a linear transformation.

Rotations

Example 1

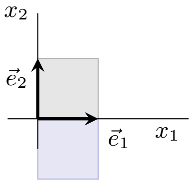
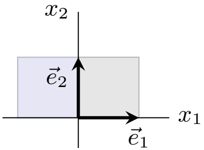
What is the linear transform $T : \mathbb{R}^2 \to \mathbb{R}^2$ defined by

$$T(\vec{x}) = \vec{x} \text{ rotated counterclockwise by angle } \theta?$$

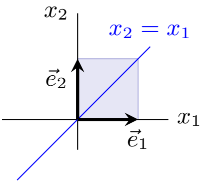
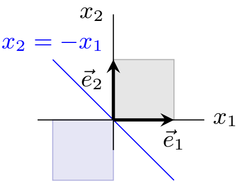
Standard Matrices in $\mathbb{R}^2$

- There is a long list of geometric transformations of $\mathbb{R}^2$ in our textbook, as well as on the next few slides (reflections, rotations, contractions and expansions, shears, projections, ...)
- Please familiarize yourself with them: you are expected to memorize them (or be able to derive them)

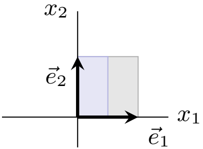
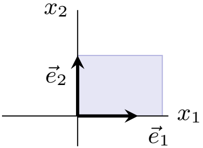
Two Dimensional Examples: Reflections

transformation	image of unit square	standard matrix
reflection through x_1-axis		$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$
reflection through x_2-axis		$\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$

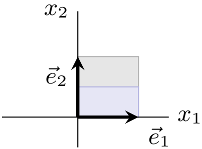
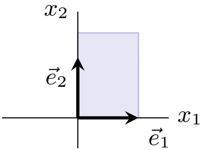
Two Dimensional Examples: Reflections

transformation	**image of unit square**	**standard matrix**
reflection through $x_2 = x_1$		$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$
reflection through $x_2 = -x_1$		$\begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}$

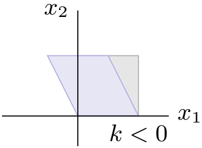
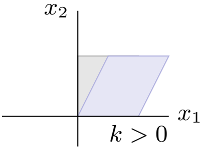
Two Dimensional Examples: Contractions and Expansions

transformation	**image of unit square**	**standard matrix**
Horizontal Contraction		$\begin{pmatrix} k & 0 \\ 0 & 1 \end{pmatrix}$. $\lvert k\rvert < 1$
Horizontal Expansion		$\begin{pmatrix} k & 0 \\ 0 & 1 \end{pmatrix}$, $k > 1$

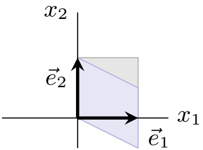
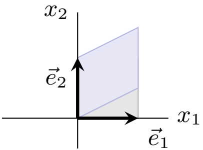
Two Dimensional Examples: Contractions and Expansions

transformation	**image of unit square**	**standard matrix**
Vertical Contraction		$\begin{pmatrix} 1 & 0 \\ 0 & k \end{pmatrix}$, $\lvert k\rvert < 1$
Vertical Expansion		$\begin{pmatrix} 1 & 0 \\ 0 & k \end{pmatrix}$, $k > 1$

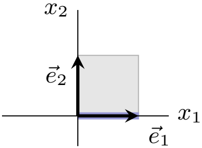
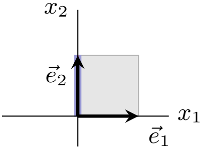
Two Dimensional Examples: Shears

transformation	image of unit square	standard matrix
Horizontal Shear(left)		$\begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix}$, $k < 0$
Horizontal Shear(right)		$\begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix}$, $k > 0$

Two Dimensional Examples: Shears

transformation	image of unit square	standard matrix
Vertical Shear(down)		$\begin{pmatrix} 1 & 0 \\ k & 1 \end{pmatrix}$, $k < 0$
Vertical Shear(up)		$\begin{pmatrix} 1 & 0 \\ k & 1 \end{pmatrix}$, $k > 0$

Two Dimensional Examples: Projections

transformation	**image of unit square**	**standard matrix**
Projection onto the x_1-axis		$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$
Projection onto the x_2-axis		$\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$

Onto

Definition

A linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ is **onto** if for all $\vec{b} \in \mathbb{R}^m$ there is a $\vec{x} \in \mathbb{R}^n$ so that $T(\vec{x}) = \vec{b}$.

Onto is an **existence property:** for any $\vec{b} \in \mathbb{R}^m$, $A\vec{x} = \vec{b}$ has a solution.

Examples

- A rotation on the plane is an onto linear transformation.
- A projection in the plane is not onto.

Useful Fact

T is onto if and only if its standard matrix has a pivot in every row.

One-to-One

Definition

A linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ is **one-to-one** if for all $\vec{b} \in \mathbb{R}^m$ there is at most one (possibly no) $\vec{x} \in \mathbb{R}^n$ so that $T(\vec{x}) = \vec{b}$.

One-to-one is a uniqueness property, it does not assert existence for all $\vec{b}$.

Examples

- A rotation on the plane is a one-to-one linear transformation.
- A projection in the plane is not one-to-one.

Useful Facts

- T is one-to-one if and only if the only solution to $T(\vec{x}) = 0$ is the zero vector, $\vec{x} = \vec{0}$.
- T is one-to-one if and only if the standard matrix A of T has no free variables.

Example

Complete the matrices below by entering numbers into the missing entries so that the properties are satisfied. **If it isn't possible to do so, state why**.

a) A is a 2×3 standard matrix for a one-to-one linear transform.
$$A = \begin{pmatrix} 1 & 0 & \\ 0 & & 1 \end{pmatrix}$$

b) B is a 3×2 standard matrix for an onto linear transform.
$$B = \begin{pmatrix} 1 & \\ & \\ & \end{pmatrix}$$

c) C is a 3×3 standard matrix of a linear transform that is one-to-one and onto.
$$C = \begin{pmatrix} 1 & 1 & 1 \\ & & \\ & & \end{pmatrix}$$

Theorem

For a linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ with standard matrix A these are equivalent statements.

1. T is onto.
2. The matrix A has columns which span $\mathbb{R}^m$.
3. The matrix A has m pivotal columns.

Theorem

For a linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ with standard matrix A these are equivalent statements.

1. T is one-to-one.
2. The unique solution to $T(\vec{x}) = \vec{0}$ is the trivial one.
3. The matrix A linearly independent columns.
4. Each column of A is pivotal.

Additional Examples

1. Construct a matrix $A \in \mathbb{R}^{2\times 2}$, such that $T(\vec{x}) = A\vec{x}$, where T is a linear transformation that rotates vectors in $\mathbb{R}^2$ counterclockwise by $\pi/2$ radians about the origin, then reflects them through the line $x_1 = x_2$.
2. Define a linear transformation by

$$T(x_1, x_2) = (3x_1 + x_2, 5x_1 + 7x_2, x_1 + 3x_2)$$

   Is T one-to-one? Is T onto?

Section 2.1 : Matrix Operations

Chapter 2 : Matrix Algebra

Math 1554 Linear Algebra

Topics and Objectives

Topics

We will cover these topics in this section.

1. Identity and zero matrices
2. Matrix algebra (sums and products, scalar multiplies, matrix powers)
3. Transpose of a matrix

Objectives

For the topics covered in this section, students are expected to be able to do the following.

1. **Apply** matrix algebra, the matrix transpose, and the zero and identity matrices, to **solve** and **analyze** matrix equations.

Definitions: Zero and Identity Matrices

1. A **zero matrix** is any matrix whose every entry is zero.

$$0_{2\times 3} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad 0_{2\times 1} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

2. The $n \times n$ **identity matrix** has ones on the main diagonal, otherwise all zeros.

$$I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Note: any matrix with dimensions $n \times n$ is **square**. Zero matrices need not be square, identity matrices must be square.

Sums and Scalar Multiples

Suppose $A \in \mathbb{R}^{m \times n}$, and $a_{i,j}$ is the element of A in row i and column j.

1. If A and B are $m \times n$ matrices, then the elements of $A + B$ are $a_{i,j} + b_{i,j}$.
2. If $c \in \mathbb{R}$, then the elements of cA are $ca_{i,j}$.

For example, if

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} + c \begin{bmatrix} 7 & 4 & 7 \\ 0 & 0 & k \end{bmatrix} = \begin{bmatrix} 15 & 10 & 17 \\ 4 & 5 & 16 \end{bmatrix}$$

What are the values of c and k?

Properties of Sums and Scalar Multiples

Scalar multiples and matrix addition have the expected properties.

If $r, s \in \mathbb{R}$ are scalars, and A, B, C are $m \times n$ matrices, then

1. $A + 0_{m\times n} = A$
2. $(A + B) + C = A + (B + C)$
3. $r(A + B) = rA + rB$
4. $(r + s)A = rA + sA$
5. $r(sA) = (rs)A$

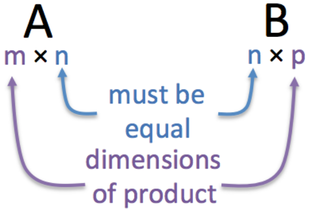
Matrix Multiplication

Definition

Let A be a $m \times n$ matrix, and B be a $n \times p$ matrix. The product is AB a $m \times p$ matrix, equal to

$$AB = A \begin{bmatrix} \vec{b}_1 & \cdots & \vec{b}_p \end{bmatrix} = \begin{bmatrix} A\vec{b}_1 & \cdots & A\vec{b}_p \end{bmatrix}$$

Note: the dimensions of A and B determine whether AB is defined, and what its dimensions will be.



Row Column Rule for Matrix Multiplication

The Row Column Rule is a convenient way to calculate the product AB that many students have encountered in pre-requisite courses.

Row Column Method

If $A \in \mathbb{R}^{m\times n}$ has rows $\vec{a}_i$, and $B \in \mathbb{R}^{n\times p}$ has columns $\vec{b}_j$, each element of the product $C = AB$ is $c_{ij} = \vec{a}_i \cdot \vec{b}_j$.

Example

Compute the following using the row-column method.

$$C = AB = \begin{pmatrix} 2 & 0 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 3 & 0 & 1 \\ 4 & 5 & 6 \end{pmatrix}$$

Properties of Matrix Multiplication

Let A, B, C be matrices of the sizes needed for the matrix multiplication to be defined, and A is a $m \times n$ matrix.

1. (Associative) $(AB)C = A(BC)$
2. (Left Distributive) $A(B + C) = AB + AC$
3. (Right Distributive) $\cdots$
4. (Identity for matrix multiplication) $I_m A = A I_n$

Warnings:

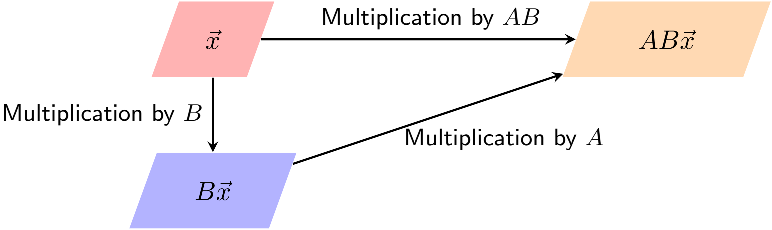
1. (non-commutative) In general, $AB \neq BA$.
2. (non-cancellation) $AB = AC$ does not mean $B = C$.
3. (Zero divisors) $AB = 0$ does not mean that either $A = 0$ or $B = 0$.

The Associative Property

The associative property is $(AB)C = A(BC)$. If $C = \vec{x}$, then

$$(AB)\vec{x} = A(B\vec{x})$$

Schematically:



The matrix product $AB\vec{x}$ can be obtained by either: multiplying by matrix AB, or by multiplying by B then by A. This means that matrix multiplication corresponds to **composition of the linear transformations**.

Proof of the Associative Law

Let A be $m \times n$, $B = \begin{bmatrix} \vec{b}_1 & \cdots & \vec{b}_p \end{bmatrix}$ a $n \times p$ and $C = \begin{bmatrix} c_1 \\ \vdots \\ c_p \end{bmatrix}$ a $p \times 1$ matrix. Then,

$$BC = \underbrace{c_1\vec{b}_1 + \cdots + c_p\vec{b}_p}_{\text{lin combin of cols of } B}$$

So

$$\begin{aligned}
A(BC) &= A(c_1\vec{b}_1 + \cdots + c_p\vec{b}_p) \\
&= c_1A\vec{b}_1 + \cdots + c_pA\vec{b}_p && \text{(multiply by } A \text{ is linear)} \\
&= \begin{bmatrix} A\vec{b}_1 & \cdots & A\vec{b}_p \end{bmatrix} \begin{bmatrix} c_1 \\ \vdots \\ c_p \end{bmatrix} && \text{(lin combin of cols of } AB\text{)} \\
&= (AB)C.
\end{aligned}$$

Example

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

Give an example of a 2×2 matrix B that is non-commutative with A.

The Transpose of a Matrix

A^T is the matrix whose columns are the rows of A.

Example

$$\begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & 1 & 0 & 2 & 0 \end{bmatrix}^T =$$

Properties of the Matrix Transpose

1. $(A^T)^T =$

2. $(A + B)^T =$

3. $(rA)^T =$

4. $(AB)^T =$

Matrix Powers

For any $n \times n$ matrix and positive integer k, A^k is the product of k copies of A.

$$A^k = AA \dots A$$

Example: Compute C^8.

$$C = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

Example

Define

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 8 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

Which of these operations are defined, and what are the dimensions of the result?

1. $A + 3C$

2. $A(AB)^T$

3. $A + ABCB^T$

Additional Examples

True or false:

1. For any I_n and any $A \in \mathbb{R}^{n\times n}$, $(I_n + A)(I_n - A) = I_n - A^2$.

2. For any A and B in $\mathbb{R}^{n\times n}$, $(A+B)^2 = A^2 + B^2 + 2AB$.

Section 2.2 : Inverse of a Matrix

Chapter 2 : Matrix Algebra

Math 1554 Linear Algebra

"Your scientists were so preoccupied with whether or not they could, they didn't stop to think if they should."

- Spielberg and Crichton, Jurassic Park, 1993 film

The algorithm we introduce in this section **could** be used to compute an inverse of an $n \times n$ matrix. At the end of the lecture we'll discuss some of the problems with our algorithm and why it can be difficult to compute a matrix inverse.

Topics and Objectives

Topics

We will cover these topics in this section.

1. Inverse of a matrix, its algebraic properties, and its relation to solving systems of linear equations.
2. Elementary matrices and their role in calculating the matrix inverse.

Objectives

For the topics covered in this section, students are expected to be able to do the following.

1. Apply the formal definition of an inverse, and its algebraic properties, to solve and analyze linear systems.
2. Compute the inverse of an $n \times n$ matrix, and use it to solve linear systems.
3. Construct elementary matrices.

Motivating Question

Is there a matrix, A, such that $\begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} A = I_3$?

The Matrix Inverse

Definition

$A \in \mathbb{R}^{n\times n}$ is **invertible** (or **non-singular**) if there is a $C \in \mathbb{R}^{n\times n}$ so that

$$AC = CA = I_n.$$

If there is, we write $C = A^{-1}$.

The Inverse of a 2×2 Matrix

There's a formula for computing the inverse of a 2×2 matrix.

Theorem

The 2×2 matrix $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is non-singular if and only if $ad - bc \neq 0$, and then

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

Example

State the inverse of the matrix below.

$$\begin{bmatrix} 2 & 5 \\ -3 & -7 \end{bmatrix}$$

The Matrix Inverse

Theorem

$A \in \mathbb{R}^{n\times n}$ has an inverse if and only if for all $\vec{b} \in \mathbb{R}^n$, $A\vec{x} = \vec{b}$ has a unique solution. And, in this case, $\vec{x} = A^{-1}\vec{b}$.

Example

Solve the linear system.

$$\begin{aligned} 3x_1 + 4x_2 &= 7 \\ 5x_1 + 6x_2 &= 7 \end{aligned}$$

Properties of the Matrix Inverse

A and B are invertible $n \times n$ matrices.

1. $(A^{-1})^{-1} = A$
2. $(AB)^{-1} = B^{-1}A^{-1}$ (Non-commutative!)
3. $(A^T)^{-1} = (A^{-1})^T$

Example

True or false: $(ABC)^{-1} = C^{-1}B^{-1}A^{-1}$.

An Algorithm for Computing A^{-1}

If $A \in \mathbb{R}^{n\times n}$, and $n > 2$, how do we calculate A^{-1}? Here's an algorithm we can use:

1. Row reduce the augmented matrix $(A \,|\, I_n)$
2. If reduction has form $(I_n \,|\, B)$ then A is invertible and $B = A^{-1}$. Otherwise, A is not invertible.

Example

Compute the inverse of $A = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 0 & 3 \\ 0 & 0 & 1 \end{bmatrix}$.

Why Does This Work?

We can think of our algorithm as simultaneously solving n linear systems:

$$\begin{aligned} A\vec{x}_1 &= \vec{e}_1 \\ A\vec{x}_2 &= \vec{e}_2 \\ &\vdots \\ A\vec{x}_n &= \vec{e}_n \end{aligned}$$

Each column of A^{-1} is $A^{-1}\vec{e}_i = \vec{x}_i$.

Over the next few slides we explore another explanation for how our algorithm works. This other explanation uses elementary matrices.

Elementary Matrices

An elementary matrix, E, is one that differs by I_n by one row operation. Recall our elementary row operations:

1. swap rows
2. multiply a row by a non-zero scalar
3. add a multiple of one row to another

We can represent each operation by a matrix multiplication with an **elementary matrix**.

Example

Suppose

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

By inspection, what is E? How does it compare to I_3?

Theorem

Returning to understanding why our algorithm works, we apply a sequence of row operations to A to obtain I_n:

$$(E_k \cdots E_3 E_2 E_1)A = I_n$$

Thus, $E_k \cdots E_3 E_2 E_1$ is the inverse matrix we seek.

Our algorithm for calculating the inverse of a matrix is the result of the following theorem.

Theorem

Matrix A is invertible if and only if it is row equivalent to the identity. In this case, the any sequence of elementary row operations that transforms A into I, applied to I, generates A^{-1}.

Using The Inverse to Solve a Linear System

- We could use A^{-1} to solve a linear system,

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

  We would calculate A^{-1} and then:

- As our textbook points out, A^{-1} is seldom used: computing it can take a very long time, and is prone to numerical error.
- So why did we learn how to compute A^{-1}? Later on in this course, we use elementary matrices and properties of A^{-1} to derive results.
- A recurring theme of this course: just because we **can** do something a certain way, doesn't that we **should**.

Section 2.3 : Invertible Matrices

Chapter 2 : Matrix Algebra

Math 1554 Linear Algebra

"A synonym is a word you use when you can't spell the other one."
- Baltasar Gracián

The theorem we introduce in this section of the course gives us many ways of saying the same thing. Depending on the context, some will be more convenient than others.

Topics and Objectives

Topics
We will cover these topics in this section.

1. The invertible matrix theorem, which is a review/synthesis of many of the concepts we have introduced.

Objectives
For the topics covered in this section, students are expected to be able to do the following.

1. Characterize the invertibility of a matrix using the Invertible Matrix Theorem.
2. Construct and give examples of matrices that are/are not invertible.

Motivating Question
When is a square matrix invertible? Let me count the ways!

The Invertible Matrix Theorem

Invertible matrices enjoy a rich set of equivalent descriptions.

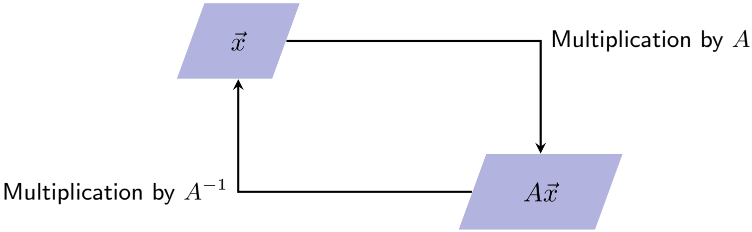
Theorem

Let A be an $n \times n$ matrix. These statements are all equivalent.

a) *A is invertible.*
b) *A is row equivalent to I_n.*
c) *A has n pivotal columns. (All columns are pivotal.)*
d) *$A\vec{x} = \vec{0}$ has only the trivial solution.*
e) *The columns of A are linearly independent.*
f) *The linear transformation $\vec{x} \mapsto A\vec{x}$ is one-to-one.*
g) *The equation $A\vec{x} = \vec{b}$ has a solution for all $\vec{b} \in \mathbb{R}^n$.*
h) *The columns of A span $\mathbb{R}^n$.*
i) *The linear transformation $\vec{x} \mapsto A\vec{x}$ is onto.*
j) *There is a $n \times n$ matrix C so that $CA = I_n$. (A has a left inverse.)*
k) *There is a $n \times n$ matrix D so that $AD = I_n$. (A has a right inverse.)*
l) *A^T is invertible.*

Invertibility and Composition

The diagram below gives us another perspective on the role of A^{-1}.



The matrix inverse A^{-1} transforms Ax back to $\vec{x}$. This is because:

$$A^{-1}(A\vec{x}) = (A^{-1}A)\vec{x} =$$

The Invertible Matrix Theorem: Final Notes

- Items j and k of the invertible matrix theorem (IMT) lead us directly to the following theorem.

Theorem

If A and B are $n \times n$ matrices and $AB = I$, then A and B are invertible, and $B = A^{-1}$ and $A = B^{-1}$.

- The IMT is a set of equivalent statements. They divide the set of all square matrices into two separate classes: invertible, and non-invertible.
- As we progress through this course, we will be able to add additional equivalent statements to the IMT (that deal with determinants, eigenvalues, etc).

Example 1

Is this matrix invertible?

$$\begin{bmatrix} 1 & 0 & -2 \\ 3 & 1 & -2 \\ -5 & -1 & 9 \end{bmatrix}$$

Example 2

If possible, fill in the missing elements of the matrices below with numbers so that each of the matrices are singular. If it is not possible to do so, state why.

$$\begin{pmatrix} 1 & 0 & 1 \\ 1 & & 1 \\ 0 & 0 & 1 \end{pmatrix}, \qquad \begin{pmatrix} 1 & & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}, \qquad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & & 1 \end{pmatrix}$$

Matrix Completion Problems

- The previous example is an example of a matrix completion problem (MCP).
- MCPs are great questions for recitations, midterms, exams.
- the **Netflix Problem** is another example of an MCP.

> Given a **ratings matrix** in which each entry (i, j) represents the rating of movie j by customer i if customer i has watched movie j, and is otherwise missing, predict the remaining matrix entries in order to make recommendations to customers on what to watch next.

Students are not expected to be familiar with this material. It's presented to motivate matrix completion.

Section 2.4 : Partitioned Matrices

Chapter 2 : Matrix Algebra

Math 1554 Linear Algebra

"Mathematics is not about numbers, equations, computations, or algorithms. Mathematics is about understanding."
- William Paul Thurston

Multiple perspectives of the same concept is a theme of this course; each perspective deepens our understanding. In this section we explore another way of representing matrices and their algebra that gives us another way of thinking about them.

Topics and Objectives

Topics

We will cover these topics in this section.

1. Partitioned matrices (or block matrices)

Objectives

For the topics covered in this section, students are expected to be able to do the following.

1. Apply partitioned matrices to solve problems regarding matrix invertibility and matrix multiplication.

What is a Partitioned Matrix?

Example

This matrix:

$$\begin{bmatrix} 3 & 1 & 4 & 1 & 0 \\ 1 & 6 & 1 & 0 & 1 \\ 0 & 0 & 0 & 4 & 2 \end{bmatrix}$$

can also be written as:

$$\begin{bmatrix} \begin{bmatrix} 3 & 1 & 4 \\ 1 & 6 & 1 \end{bmatrix} & \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \\ \\ \begin{bmatrix} 0 & 0 & 0 \end{bmatrix} & \begin{bmatrix} 4 & 2 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} A_{1,1} & A_{1,2} \\ A_{2,1} & A_{2,2} \end{bmatrix}$$

We partitioned our matrix into four **blocks**, each of which has different dimensions.

Another Example of a Partitioned Matrix

Example: The reduced echelon form of a matrix. We can use a partitioned matrix to

$$\begin{bmatrix} 1 & 0 & 0 & 0 & * & \cdots & * \\ 0 & 1 & 0 & 0 & * & \cdots & * \\ 0 & 0 & 1 & 0 & * & \cdots & * \\ 0 & 0 & 0 & 1 & * & \cdots & * \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 \end{bmatrix} = \begin{bmatrix} I_4 & F \\ 0 & 0 \end{bmatrix}$$

This is useful when studying the **null space** of A, as we will see later in this course.

Row Column Method

Recall that a row vector times a column vector (of the right dimensions) is a scalar. For example,

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

This is the **row column** matrix multiplication method from Section 2.1.

Theorem

Let A be $m \times n$ and B be $n \times p$ matrix. Then, the (i, j) entry of AB is

$$\operatorname{row}_i A \cdot \operatorname{col}_j B.$$

This is the **Row Column Method** for matrix multiplication.

Partitioned matrices can be multiplied using this method, as if each block were a scalar (provided each block has appropriate dimensions).

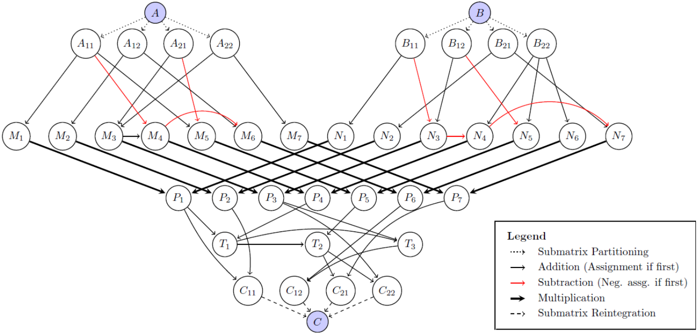
Example of Row Column Method

Recall, using our formula for a 2×2 matrix, $\begin{bmatrix} a & b \\ 0 & c \end{bmatrix}^{-1} = \frac{1}{ac} \begin{bmatrix} c & -b \\ 0 & a \end{bmatrix}$.

Example: Suppose $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times n}$, and $C \in \mathbb{R}^{n \times n}$ are invertible matrices. Construct the inverse of $\begin{bmatrix} A & B \\ 0 & C \end{bmatrix}$.

The Strassen Algorithm: An impressive use of partitioned matrices

Naive Multiplication of two $n \times n$ matrices A and B requires n^3 arithmetic steps. Strassen's algorithm **partitions** the matrices, makes a very clever sequence of multiplications, additions, to reduce the computation to $n^{2.803\ldots}$ steps.

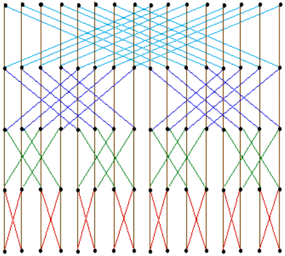


Students aren't expected to be familiar with this material. It's presented to motivate matrix partitioning.

The Fast Fourier Transform (FFT)

The FFT is an essential algorithm of modern technology that uses partitioned matrices recursively.

$$G_0 = \left[1\right], \qquad G_{n+1} = \begin{bmatrix} G_n & -G_n \\ G_n & G_n \end{bmatrix}$$



The recursive structure of the matrix means that it can be computed in nearly **linear** time. This is an incredible saving over the general complexity of n^3. It means that we can compute $G_n x$, and G_n^{-1} very quickly.

Students aren't expected to be familiar with this material. It is presented to motivate matrix partitioning.

Section 2.5 : Matrix Factorizations

Chapter 2 : Matrix Algebra

Math 1554 Linear Algebra

"Mathematical reasoning may be regarded rather schematically as the exercise of a combination of two facilities, which we may call intuition and ingenuity." - Alan Turing

The use of the LU Decomposition to solve linear systems was one of the areas of mathematics that Turing helped develop.

Topics and Objectives

Topics

We will cover these topics in this section.

1. The LU factorization of a matrix
2. Using the LU factorization to solve a system
3. Why the LU factorization works

Objectives

For the topics covered in this section, students are expected to be able to do the following.

1. Compute an LU factorization of a matrix.
2. Apply the LU factorization to solve systems of equations.
3. Determine whether a matrix has an LU factorization.

Motivation

- Recall that we **could** solve $A\vec{x} = \vec{b}$ by using

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

- This requires computation of the inverse of an $n \times n$ matrix, which is especially difficult for large n.
- Instead we could solve $A\vec{x} = \vec{b}$ with Gaussian Elimination, but this is not efficient for large n
- There are more efficient and accurate methods for solving linear systems that rely on matrix factorizations.

Matrix Factorizations

- A **matrix factorization**, or **matrix decomposition** is a factorization of a matrix into a product of matrices.
- Factorizations can be useful for solving $A\vec{x} = \vec{b}$, or understanding the properties of a matrix.
- We explore a few matrix factorizations throughout this course.
- In this section, we factor a matrix into **lower** and into **upper** triangular matrices.

Triangular Matrices

- A rectangular matrix A is **upper triangular** if $a_{i,j} = 0$ for $i > j$. Examples:

$$\begin{pmatrix} 1 & 5 & 0 \\ 0 & 2 & 4 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 2 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

- A rectangular matrix A is **lower triangular** if $a_{i,j} = 0$ for $i < j$. Examples:

$$\begin{pmatrix} 1 & 0 & 0 \\ 3 & 2 & 0 \end{pmatrix}, \quad \begin{pmatrix} 3 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 2 & 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ 2 \\ 1 \\ 2 \end{pmatrix}$$

Ask: Can you name a matrix that is both upper and lower triangular?

The LU Factorization

Theorem

If A is an $m \times n$ matrix that can be row reduced to echelon form without row exchanges, then $A = LU$. L is a lower triangular $m \times m$ matrix with 1's on the diagonal, U is an **echelon** form of A.

Example: If $A \in \mathbb{R}^{3\times 2}$, the LU factorization has the form:

$$A = LU = \begin{pmatrix} 1 & 0 & 0 \\ * & 1 & 0 \\ * & * & 1 \end{pmatrix} \begin{pmatrix} * & * \\ 0 & * \\ 0 & 0 \end{pmatrix}$$

Why We Can Compute the LU Factorization

Suppose A can be row reduced to echelon form U without interchanging rows. Then,

$$E_p \cdots E_1 A = U$$

where the E_j are matrices that perform elementary row operations. They happen to be lower triangular and invertible, e.g.

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

Therefore,

$$A = \underbrace{E_1^{-1} \cdots E_p^{-1}}_{=L} U = LU.$$

Using the LU Decomposition

Goal: given A and $\vec{b}$, solve $A\vec{x} = \vec{b}$ for $\vec{x}$.

Algorithm: construct $A = LU$, solve $A\vec{x} = LU\vec{x} = \vec{b}$ by:

1. Forward solve for $\vec{y}$ in $L\vec{y} = \vec{b}$.
2. Backwards solve for x in $U\vec{x} = \vec{y}$.

Example: Solve the linear system whose LU decomposition is given.

$$A = LU = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}, \quad \vec{b} = \begin{pmatrix} 2 \\ 3 \\ 2 \\ 0 \end{pmatrix}$$

An Algorithm for Computing LU

To compute the LU decomposition:

1. Reduce A to an echelon form U by a sequence of row replacement operations, if possible.
2. Place entries in L such that the same sequence of row operations reduces L to I.

Note that

- In MATH 1554, the only row replacement operation we can use is to *replace a row with a multiple of a row above it*.
- More advanced linear algebra courses address this limitation.

Example: Compute the LU factorization of A.

$$A = \begin{pmatrix} 4 & -3 & -1 & 5 \\ -16 & 12 & 2 & -17 \\ 8 & -6 & -12 & 22 \end{pmatrix}$$

Summary

- To solve $A\vec{x} = LU\vec{x} = \vec{b}$,
  1. Forward solve for $\vec{y}$ in $L\vec{y} = \vec{b}$.
  2. Backwards solve for $\vec{x}$ in $U\vec{x} = \vec{y}$.
- To compute the LU decomposition:
  1. Reduce A to an echelon form U by a sequence of row replacement operations, if possible.
  2. Place entries in L such that the same sequence of row operations reduces L to I.
- The textbook offers a different explanation of how to construct the LU decomposition that students may find helpful.
- Another explanation on how to calculate the LU decomposition that students may find helpful is available from MIT OpenCourseWare: www.youtube.com/watch?v=rhNKncraJMk

Section 2.6 : The Leontif Input-Output Model

Chapter 2 : Matrix Algebra

Math 1554 Linear Algebra

"Computers and robots replace humans in the exercise of mental functions in the same way as mechanical power replaced them in the performance of physical tasks." - Wassily Leontif, 1983

Students in this course are of course required to demonstrate an understanding of underlying concepts behind procedures and algorithms. This is in part because computers are continuing to take on a much larger role in performing calculations.

Topics and Objectives

Topics

We will cover these topics in this section.

1. The Leontief Input-Output model, as a simple example of a model of an economy.

Objectives

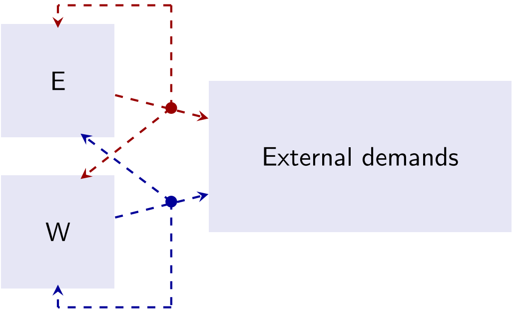
For the topics covered in this section, students are expected to be able to do the following.

1. Apply matrix algebra and inverses to solve and analyze Leontif Input-Output problems.

Motivating Question

An economy consisting of 3 sectors: agriculture, manufacturing, and energy. The output of one sector is absorbed by all the sectors. If there is an increase in demand for energy, how does this impact the economy?

Example: An Economy with Two Sectors



This economy contains two sectors.

1. electricity (E)
2. water (W)

The “external demands” is another part of the economy, which does not produce E and W.

How might we represent this economy with a set of linear equations?

The Leontif Model: Internal Consumption

Suppose economy has N sectors, with outputs measured by $\vec{x} \in \mathbb{R}^N$.

$$\begin{aligned} \vec{x} &= \text{output vector} \\ x_i &= \text{element } i \text{ of vector } \vec{x} = \text{number of units produced by sector } i \end{aligned}$$

The **consumption matrix**, C, describes how units are consumed by sectors to produce output. Two equivalent ways of defining C:

- Sector j requires a proportion of the units created by sector i. Call that $c_{i,j}x_i$
- Sector i sends a proportion of its units to sector j. Call that $c_{i,j}x_i$

Elements of C are $c_{i,j}$, with $c_{i,j} \in [0, 1]$, and

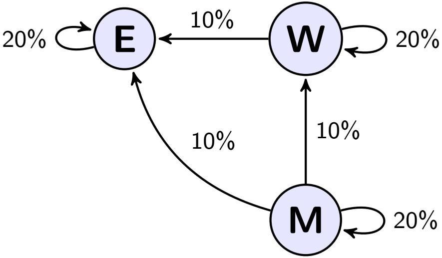
$$\begin{aligned} C\vec{x} &= \text{units consumed} \\ \vec{x} - C\vec{x} &= \text{units left after internal consumption} \end{aligned}$$

Example 1

An economy contains three sectors, E, W, M. For every 100 units of output,

- E requires 20 units from E, 10 units from W, and 10 units from M
- W requires 0 units from E, 20 units from W, and 10 units from M
- M requires 0 units from E, 0 units from W, and 20 units from M

Construct the consumption matrix for this economy.



Solution: Creating C

Our consumption matrix is

$$C = \frac{1}{10} \begin{pmatrix} 2 & 0 & 0 \\ 1 & 2 & 0 \\ 1 & 1 & 2 \end{pmatrix}$$

Note:

- total output for each sector is the sum along the outgoing edges for each sector, which generates rows of C
- elements of C represent percentages with no units, they have values between 0 and 1
- our output vector has units

The Leontif Model: Demand

There is also an external demand given by $\vec{d} \in \mathbb{R}^N$. We ask if there is an $\vec{x}$ such that

$$\vec{x} - C\vec{x} = \vec{d}$$

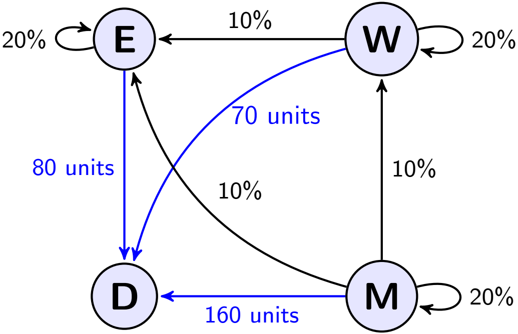
Solving for $\vec{x}$ yields

$$\vec{x} = (I - C)^{-1}\vec{d}$$

This is the **Leontief Input-Output Model**.

Example 1 Revisited

Now suppose there is an external demand: what production level is required to satisfy a final demand of 80 units of E, 70 units of W, and 160 unites of M?



Solution

The production level would be found by solving:

$$\vec{x} - C\vec{x} = \vec{d}$$

$$(I - C)\vec{x} = \vec{d}$$

$$\frac{1}{10}\begin{pmatrix} 8 & 0 & 0 \\ -1 & 8 & 0 \\ -1 & -1 & 8 \end{pmatrix}\vec{x} = \begin{pmatrix} 80 \\ 70 \\ 160 \end{pmatrix}$$

$$\begin{aligned} 8x_1 = 800 \quad &\Rightarrow \quad x_1 = 100 \\ -x_1 + 8x_2 = 700 \quad &\Rightarrow \quad x_2 = 100 \\ -x_1 - x_2 + 8x_3 = 1600 \quad &\Rightarrow \quad x_3 = 1800/8 = 225 \end{aligned}$$

The output that balances demand with internal consumption is

$$\vec{x} = \begin{pmatrix} 100 \\ 100 \\ 225 \end{pmatrix}.$$

The Importance of $(I - C)^{-1}$

For the example above

$$(I-C)^{-1} \approx \begin{pmatrix} 1.25 & 0 & 0 \\ 0.15 & 1.25 & 0 \\ 0.18 & 0.17 & 1.25 \end{pmatrix}$$

The entries of $(I - C)^{-1} = B$ have this meaning: if the final demand vector $\vec{d}$ increases by one unit in the j^{th} place, the column vector b_j is the additional output required from other sectors.

So to meet an increase in demand for M by one unit, requires 1.25 of one additional units from M to meet internal consumption.

Section 2.7 : Computer Graphics

Chapter 2 : Matrix Algebra

Math 1554 Linear Algebra

Topics and Objectives

Topics

We will cover these topics in this section.

1. Homogeneous coordinates in 2D and 3D
2. Translations and composite transforms in 2D and 3D

Objectives

For the topics covered in this section, students are expected to be able to do the following.

1. Construct a data matrix to represent points in $\mathbb{R}^2$ and $\mathbb{R}^3$ using homogeneous coordinates.
2. Construct transformation matrices to represent composite transforms in 2D and 3D using homogeneous coordinates.
3. Apply composite transforms and data matrices to transform points in $\mathbb{R}^3$

In the interest of time, students are not expected to be familiar with perspective projections.

Motivating Question

How can we represent translations using linear transforms?

Homogenous Coordinates

Translations of points in $\mathbb{R}^n$ does not correspond directly to a linear transform. **Homogeneous coordinates** are used model translations using matrix multiplication.

Homogeneous Coordinates in $\mathbb{R}^2$

Each point (x, y) in $\mathbb{R}^2$ can be identified with the point (x, y, H), $H \neq 0$, on the plane in $\mathbb{R}^3$ that lies H units above the xy-plane.

Note: we often we set $H = 1$.

Example: A translation of the form $(x, y) \to (x + h, y + k)$ can be represented as a matrix multiplication with homogeneous coordinates:

$$\begin{pmatrix} 1 & 0 & h \\ 0 & 1 & k \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \begin{pmatrix} x + h \\ y + k \\ 1 \end{pmatrix}$$

A Composite Transform with Homogeneous Coordinates

Triangle S is determined by three data points, $(1, 1), (2, 4), (3, 1)$.

Transform T rotates points by $\pi/2$ radians counterclockwise about the point $(0, 1)$.

a) Represent the data with a matrix, D. Use homogeneous coordinates.
b) Use matrix multiplication to determine the image of S under T.
c) Sketch S and its image under T.

3D Homogeneous Coordinates

Homogeneous coordinates in 3D are analogous to our 2D coordinates.

Homogeneous Coordinates in $\mathbb{R}^3$

(X, Y, Z, H) are homogeneous coordinates for (x, y, z) in $\mathbb{R}^3$, $H \neq 0$, and

$$x = \frac{X}{H}, \quad y = \frac{Y}{H}, \quad z = \frac{Z}{H}$$

For example, $(a, b, c, 1)$ and $(3a, 3b, 3c, 3)$ are both homogeneous coordinates for the point (a, b, c).

3D Transformation Matrices

Construct matrices for the following transformations.

a) A rotation in $\mathbb{R}^3$ about the y-axis by π radians.

b) A translation specified by the vector $\vec{p} = \begin{pmatrix} -2 \\ 3 \\ 4 \end{pmatrix}$.

Section 2.8 : Subspaces of $\mathbb{R}^n$

Chapter 2 : Matrix Algebra

Math 1554 Linear Algebra

Topics and Objectives

Topics

We will cover these topics in this section.

1. Subspaces, Column space, and Null spaces
2. A basis for a subspace.

Objectives

For the topics covered in this section, students are expected to be able to do the following.

1. Determine whether a set is a subspace.
2. Determine whether a vector is in a particular subspace, or find a vector in that subspace.
3. Construct a basis for a subspace (for example, a basis for Col(A))

Motivating Question

Given a matrix A, what is the set of vectors $\vec{b}$ for which we can solve $A\vec{x} = \vec{b}$?

Subsets of $\mathbb{R}^n$

Definition

A **subset of** $\mathbb{R}^n$ is any collection of vectors that are in $\mathbb{R}^n$.

Subspaces in $\mathbb{R}^n$

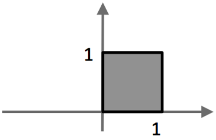
Definition

A subset H of $\mathbb{R}^n$ is a **subspace** if it is closed under scalar multiplies and vector addition. That is: for any $c \in \mathbb{R}$ and for $\vec{u}, \vec{v} \in H$,

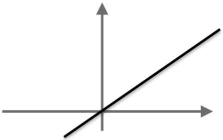
1. $c\,\vec{u} \in H$
2. $\vec{u} + \vec{v} \in H$

Note that condition 1 implies that the zero vector must be in H.

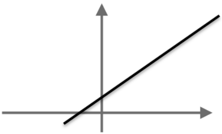
Example 1: Which of the following subsets could be a subspace of $\mathbb{R}^2$?



a) the unit square



b) a line passing through the origin



c) a line that doesn't pass through the origin

The Column Space and the Null Space of a Matrix

Recall: for $\vec{v}_1, \ldots, \vec{v}_p \in \mathbb{R}^n$, that $\mathrm{Span}\{\vec{v}_1, \ldots, \vec{v}_p\}$ is:

This is a **subspace**, spanned by $\vec{v}_1, \ldots, \vec{v}_p$.

Definition

Given an $m \times n$ matrix $A = \begin{bmatrix} \vec{a}_1 & \cdots & \vec{a}_n \end{bmatrix}$

1. The **column space of** A, $\mathrm{Col}\, A$, is the subspace of $\mathbb{R}^m$ spanned by $\vec{a}_1, \ldots, \vec{a}_n$.
2. The **null space of** A, $\mathrm{Null}\, A$, is the subspace of $\mathbb{R}^n$ spanned by the set of all vectors $\vec{x}$ that solve $A\vec{x} = \vec{0}$.

Example

Is $\vec{b}$ in the column space of A?

$$A = \begin{bmatrix} 1 & -3 & -4 \\ -4 & 6 & -2 \\ -3 & 7 & 6 \end{bmatrix} \sim \begin{bmatrix} 1 & -3 & -4 \\ 0 & -6 & -18 \\ 0 & 0 & 0 \end{bmatrix}, \quad \vec{b} = \begin{pmatrix} 3 \\ 3 \\ -4 \end{pmatrix}$$

Example 2 (continued)

Using the matrix on the previous slide: is $\vec{v}$ in the null space of A?

$$\vec{v} = \begin{pmatrix} -5\lambda \\ -3\lambda \\ \lambda \end{pmatrix}, \quad \lambda \in \mathbb{R}$$

Basis

Definition

A **basis** for a subspace H is a set of linearly independent vectors in H that span H.

Example

The set $H = \{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} \in \mathbb{R}^4 \mid x_1 + 2x_2 + x_3 + 5x_4 = 0\}$ is a subspace.

a) H is a null space for what matrix A?

b) Construct a basis for H.

Example

Construct a basis for Null$A$ and a basis for Col$A$.

$$A = \begin{bmatrix} -3 & 6 & -1 & 0 \\ 1 & -2 & 2 & 0 \\ 2 & -4 & 5 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & -2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Additional Example

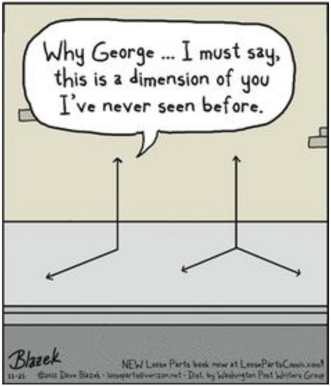
Let $V = \left\{ \begin{pmatrix} a \\ b \end{pmatrix} \in \mathbb{R}^2 \mid ab = 0 \right\}$.

1. Give an example of a vector that is in V.
2. Give an example of a vector that is not in V.
3. Is the zero vector in V?
4. Is V a subspace?

Section 2.9 : Dimension and Rank

Chapter 2 : Matrix Algebra

Math 1554 Linear Algebra



Topics and Objectives

Topics

We will cover these topics in this section.

1. Coordinates, relative to a basis.
2. Dimension of a subspace.
3. The Rank of a matrix

Objectives

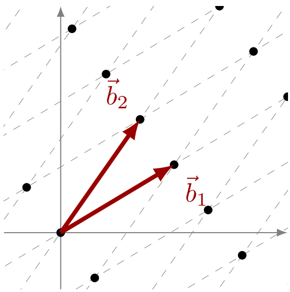
For the topics covered in this section, students are expected to be able to do the following.

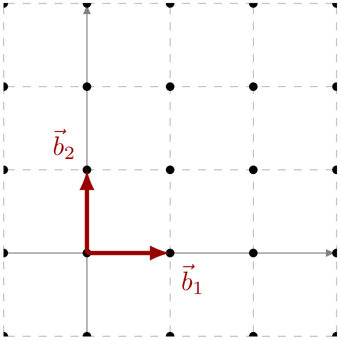
1. Calculate the coordinates of a vector in a given basis.
2. Characterize a subspace using the concept of dimension (or cardinality).
3. Characterize a matrix using the concepts of rank, column space, null space.
4. Apply the Rank, Basis, and Matrix Invertibility theorems to describe matrices and subspaces.

Choice of Basis

Key idea: There are many possible choices of basis for a subspace. Our choice can give us dramatically different properties.

Example: sketch $\vec{b}_1 + \vec{b}_2$ for the two different coordinate systems below.





Coordinates

Definition

Let $\mathcal{B} = \{\vec{b}_1, \ldots, \vec{b}_p\}$ be a basis for a subspace H. If $\vec{x}$ is in H, then **coordinates of $\vec{x}$ relative $\mathcal{B}$** are the weights (scalars) $c_1, \ldots, c_p$ so that

$$\vec{x} = c_1\vec{b}_1 + \cdots + c_p\vec{b}_p$$

And

$$[x]_{\mathcal{B}} = \begin{bmatrix} c_1 \\ \vdots \\ c_p \end{bmatrix}$$

is the **coordinate vector of $\vec{x}$ relative to** $\mathcal{B}$, or the $\mathcal{B}$**-coordinate vector of $\vec{x}$**

Example 1

Let $\vec{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$, $\vec{v}_2 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$, and $\vec{x} = \begin{bmatrix} 5 \\ 3 \\ 5 \end{bmatrix}$. Verify that $\vec{x}$ is in the span of $\mathcal{B} = \{\vec{v}_1, \vec{v}_2\}$, and calculate $[\vec{x}]_{\mathcal{B}}$.

Dimension

Definition

The **dimension** (or cardinality) of a non-zero subspace H, $\dim H$, is the number of vectors in a basis of H. We define $\dim\{0\} = 0$.

Theorem

Any two choices of bases $\mathcal{B}_1$ and $\mathcal{B}_2$ of a non-zero subspace H have the same dimension.

Examples:

1. $\dim \mathbb{R}^n =$

2. $H = \{(x_1, \dots, x_n) \, : \, x_1 + \cdots + x_n = 0\}$ has dimension

3. $\dim(\mathrm{Null}\, A)$ is the number of

4. $\dim(\mathrm{Col}\, A)$ is the number of

Rank

Definition

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

Example 2: Compute rank(A) and dim(Nul(A)).

$$\begin{bmatrix} 2 & 5 & -3 & -4 & 8 \\ 4 & 7 & -4 & -3 & 9 \\ 6 & 9 & -5 & 2 & 4 \\ 0 & -9 & 6 & 5 & -6 \end{bmatrix} \sim \cdots \sim \begin{bmatrix} 2 & 5 & -3 & -4 & 8 \\ 0 & -3 & 2 & 5 & -7 \\ 0 & 0 & 0 & 4 & -6 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Rank, Basis, and Invertibility Theorems

Theorem (Rank Theorem)

If a matrix A has n columns, then $\operatorname{Rank} A + \dim(\operatorname{Nul} A) = n$.

Theorem (Basis Theorem)

Any two bases for a subspace have the same dimension.

Theorem (Invertibility Theorem)

Let A be a $n \times n$ matrix. These conditions are equivalent.

1. *A is invertible.*
2. *The columns of A are a basis for $\mathbb{R}^n$.*
3. $\operatorname{Col} A = \mathbb{R}^n$.
4. $\operatorname{rank} A = \dim(\operatorname{Col} A) = n$.
5. $\operatorname{Null} A = \{0\}$.

Examples

If possible give an example of a 2×3 matrix A, that is in RREF and has the given properties.

a) rank$(A) = 3$

b) rank$(A) = 2$

c) dim(Null$(A)) = 2$

d) Null$A = \{0\}$

Section 3.1 : Introduction to Determinants

Chapter 3 : Determinants

Math 1554 Linear Algebra

Topics and Objectives

Topics

We will cover these topics in this section.

1. The definition and computation of a determinant
2. The determinant of triangular matrices

Objectives

For the topics covered in this section, students are expected to be able to do the following.

1. Compute determinants of $n \times n$ matrices using a cofactor expansion.
2. Apply theorems to compute determinants of matrices that have particular structures.

A Definition of the Determinant

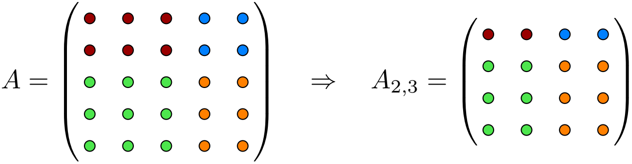
Suppose A is $n \times n$ and has elements a_{ij}.

1. If $n = 1$, $A = [a_{11}]$, and has determinant $\det A = a_{11}$.
2. Inductive case: for $n > 1$,

$$\det A = a_{11} \det A_{11} - a_{12} \det A_{12} + \cdots + (-1)^{1+n} a_{1n} \det A_{1n}$$

where A_{ij} is the submatrix obtained by eliminating row i and column j of A.

Example



Example 1

Compute $\det \begin{bmatrix} a & b \\ c & d \end{bmatrix}$.

Example 2

Compute $\det \begin{bmatrix} 1 & -5 & 0 \\ 2 & 4 & -1 \\ 0 & 2 & 0 \end{bmatrix} = \begin{vmatrix} 1 & -5 & 0 \\ 2 & 4 & -1 \\ 0 & 2 & 0 \end{vmatrix}$.

Cofactors

Cofactors give us a more convenient notation for determinants.

Definition: Cofactor

The (i,j) cofactor of an $n \times n$ matrix A is

$$C_{ij} = (-1)^{i+j} \det A_{ij}$$

The pattern for the negative signs is

$$\begin{pmatrix} + & - & + & - & \dots \\ - & + & - & + & \dots \\ + & - & + & - & \dots \\ - & + & - & + & \dots \\ \vdots & \vdots & \vdots & \vdots & \end{pmatrix}$$

Theorem

The determinant of a matrix A can be computed down any row or column of the matrix. For instance, down the j^{th} column, the determinant is

$$\det A = a_{1j}C_{1j} + a_{2j}C_{2j} + \cdots + a_{nj}C_{nj}.$$

This gives us a way to calculate determinants more efficiently.

Example 3

Compute the determinant of $\begin{bmatrix} 5 & 4 & 3 & 2 \\ 0 & 1 & 2 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 1 & 1 & 3 \end{bmatrix}$.

Triangular Matrices

Theorem

If A is a triangular matrix then

$$\det A = a_{11}a_{22}a_{33}\cdots a_{nn}.$$

Example 4

Compute the determinant of the matrix. Empty elements are zero.

$$\begin{bmatrix} 2 & 1 & & & & & \\ & 2 & 1 & & & & \\ & & 2 & 1 & & & \\ & & & 2 & 1 & & \\ & & & & 2 & 1 & \\ & & & & & 2 & 1 \\ & & & & & & 2 \end{bmatrix}$$

Computational Efficiency

Note that computation of a co-factor expansion for an $N \times N$ matrix requires roughly $N!$ multiplications.

- A 10×10 matrix requires roughly $10! = 3.6$ million multiplications
- A 20×20 matrix requires $20! \approx 2.4 \times 10^{18}$ multiplications

Co-factor expansions may not be practical, but determinants are still useful.

- We will explore other methods for computing determinants that are more efficient.
- Determinants are very useful in multivariable calculus for solving certain integration problems.

Section 3.2 : Properties of the Determinant

Chapter 3 : Determinants

Math 1554 Linear Algebra

"A problem isn't finished just because you've found the right answer."
- Yōko Ogawa

We have a method for computing determinants, but without some of the strategies we explore in this section, the algorithm can be very inefficient.

Topics and Objectives

Topics
We will cover these topics in this section.

- The relationships between row reductions, the invertibility of a matrix, and determinants.

Objectives
For the topics covered in this section, students are expected to be able to do the following.

1. Apply properties of determinants (related to row reductions, transpose, and matrix products) to compute determinants.
2. Use determinants to determine whether a square matrix is invertible.

Row Operations

- We saw how determinants are difficult or impossible to compute with a cofactor expansion for large N.
- Row operations give us a more efficient way to compute determinants.

Theorem: Row Operations and the Determinant

Let A be a square matrix.

1. If a multiple of a row of A is added to another row to produce B, then $\det B = \det A$.
2. If two rows are interchanged to produce B, then $\det B = -\det A$.
3. If one row of A is multiplied by a scalar k to produce B, then $\det B = k \det A$.

Example 1 Compute $\begin{vmatrix} 1 & -4 & 2 \\ -2 & 8 & -9 \\ -1 & 7 & 0 \end{vmatrix}$

Invertibility

Important practical implication: If A is reduced to echelon form, by r interchanges of rows and columns, then

$$|A| = \begin{cases} (-1)^{r} \times \text{(product of pivots)}, & \text{when } A \text{ is invertible} \\ 0, & \text{when } A \text{ is singular.} \end{cases}$$

Example 2 Compute the determinant

$$\begin{vmatrix} 0 & 1 & 2 & -1 \\ 2 & 5 & -7 & 3 \\ 0 & 3 & 6 & 2 \\ -2 & -5 & 4 & 2 \end{vmatrix}$$

Properties of the Determinant

For any square matrices A and B, we can show the following.

1. $\det A = \det A^T$.
2. A is invertible if and only if $\det A \neq 0$.
3. $\det(AB) = \det A \cdot \det B$.

Additional Example (if time permits)

Use a determinant to find all values of λ such that matrix C is not invertible.

$$C = \begin{pmatrix} 5 & 0 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} - \lambda I_3$$

Additional Example (if time permits)

Determine the value of

$$\det A = \det\left(\begin{pmatrix} 0 & 2 & 0 \\ 1 & 1 & 2 \\ 1 & 1 & 3 \end{pmatrix}^8\right).$$

Section 3.3 : Volume, Linear Transformations

Chapter 3 : Determinants

Math 1554 Linear Algebra

Topics and Objectives

Topics

We will cover these topics in this section.

1. Relationships between area, volume, determinants, and linear transformations.

Objectives

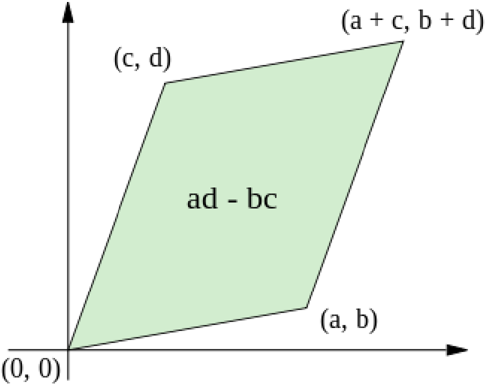
For the topics covered in this section, students are expected to be able to do the following.

1. Use determinants to compute the area of a parallelogram, or the volume of a parallelepiped, possibly under a given linear transformation.

Students are not expected to be familiar with Cramer's rule.

Determinants, Area and Volume

In $\mathbb{R}^2$, determinants give us the area of a parallelogram.



area of parallelogram $= \det\begin{pmatrix} a & c \\ b & d \end{pmatrix} = ad - bc$.

Determinants as Area, or Volume

Theorem

The volume of the parallelpiped spanned by the columns of an $n \times n$ matrix A is $|\det A|$.

Key Geometric Fact (which works in any dimension). The area of the parallelogram spanned by two vectors $\vec{a}, \vec{b}$ is equal to the area spanned by $\vec{a}, c\vec{a} + \vec{b}$, for any scalar c.

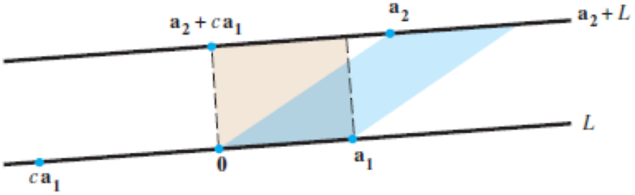


FIGURE 2 Two parallelograms of equal area.

Example 1

Calculate the area of the parallelogram determined by the points $(-2,-2),(0,3),(4,-1),(6,4)$

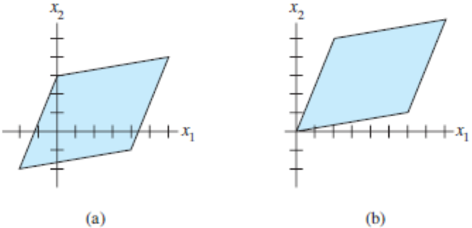


FIGURE 5 Translating a parallelogram does not change its area.

Linear Transformations

Theorem

If $T_A \ : \mathbb{R}^n \mapsto \mathbb{R}^n$, and S is some parallelogram in $\mathbb{R}^n$, then

$$\text{volume}\left(T_A(S)\right) = |\det(A)| \cdot \text{volume}(S)$$

An example that applies this theorem is given in this week's worksheets.

Section 4.9 : Applications to Markov Chains

Chapter 4 : Vector Spaces

Math 1554 Linear Algebra

Topics and Objectives

Topics
We will cover these topics in this section.

1. Markov chains
2. Steady-state vectors
3. Convergence

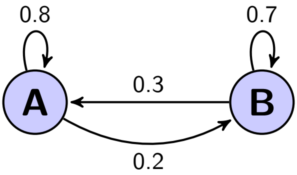
Objectives
For the topics covered in this section, students are expected to be able to do the following.

1. Construct stochastic matrices and probability vectors.
2. Model and solve real-world problems using Markov chains (e.g. - find a steady-state vector for a Markov chain)
3. Determine whether a stochastic matrix is regular.

Example 1

- A small town has two libraries, A and B.
- After 1 month, among the books checked out of A,
  - 80% returned to A
  - 20% returned to B
- After 1 month, among the books checked out of B,
  - 30% returned to A
  - 70% returned to B

If both libraries have 1000 books today, how many books does each library have after 1 month? After one year? After n months? A place to simulate this is `http://setosa.io/markov/index.html`



Example 1 Continued

The books are equally divided by between the two branches, denoted by $\vec{x}_0 = \begin{bmatrix} .5 \\ .5 \end{bmatrix}$. What is the distribution after 1 month, call it $\vec{x}_1$? After two months?

After k months, the distribution is $\vec{x}_k$, which is what in terms of $\vec{x}_0$?

Markov Chains

A few definitions:

- A **probability vector** is a vector, $\vec{x}$, with non-negative elements that sum to 1.
- A **stochastic matrix** is a square matrix, P, whose columns are probability vectors.
- A **Markov chain** is a sequence of probability vectors $\vec{x}_k$, and a stochastic matrix P, such that:

$$\vec{x}_{k+1} = P\vec{x}_k, \quad k = 0, 1, 2, \ldots$$

- A **steady-state vector** for P is a vector $\vec{q}$ such that $P\vec{q} = \vec{q}$.

Example 2

Determine a steady-state vector for the stochastic matrix

$$\begin{pmatrix} .8 & .3 \\ .2 & .7 \end{pmatrix}$$

Convergence

We often want to know what happens to a process,

$$\vec{x}_{k+1} = P\vec{x}_k, \quad k = 0, 1, 2, \dots$$

as $k \to \infty$.

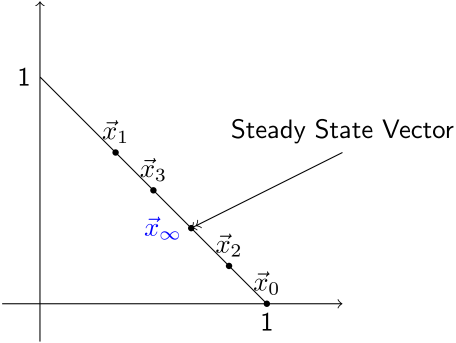
Definition: a stochastic matrix P is **regular** if there is some k such that P^k only contains strictly positive entries.

Theorem

If P is a regular stochastic matrix, then P has a unique steady-state vector $\vec{q}$, and $\vec{x}_{k+1} = P\vec{x}_k$ converges to $\vec{q}$ as $k \to \infty$.

Stochastic Vectors in the Plane

The stochastic vectors in the plane are the line segment below, and a stochastic matrix maps stochastic vectors to themselves. Iterates $P^k\vec{x}_0$ converge to the steady state.



$$P^k \longrightarrow \begin{bmatrix} \vec{x}_\infty & \vec{x}_\infty \end{bmatrix}$$

Example 3

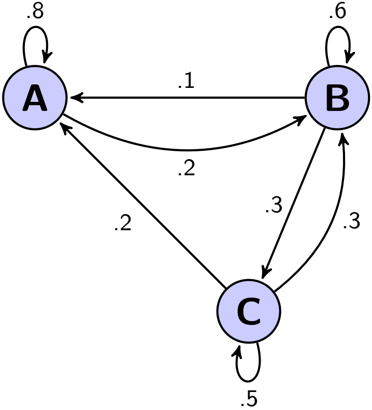
A car rental company has 3 rental locations, A, B, and C. Cars can be returned at any location. The table below gives the pattern of rental and returns for a given week.

		rented from		
		A	B	C
	A	.8	.1	.2
returned to	B	.2	.6	.3
	C	.0	.3	.5

There are 10 cars at each location today.

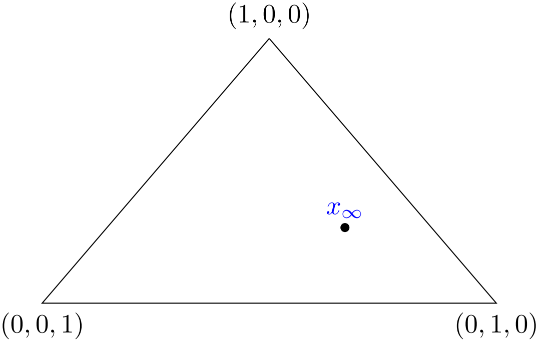
a) Construct a stochastic matrix, P, for this problem.

b) What happens to the distribution of cars after a long time? You may assume that P is regular.



$$P = \begin{bmatrix} .8 & .1 & .2 \\ .2 & .6 & .3 \\ .0 & .3 & .5 \end{bmatrix}$$

The Stochastic vectors in $\mathbb{R}^3$, are vectors $\begin{bmatrix} s \\ t \\ 1-s-t \end{bmatrix}$, where $0 \leq s, t \leq 1$ and $s + t \leq 1$. P 'contracts' stochastic vectors to x_∞.



Section 5.1 : Eigenvectors and Eigenvalues

Chapter 5 : Eigenvalues and Eigenvectors

Math 1554 Linear Algebra

Topics and Objectives

Topics

We will cover these topics in this section.

1. Eigenvectors, eigenvalues, eigenspaces
2. Eigenvalue theorems

Objectives

For the topics covered in this section, students are expected to be able to do the following.

1. Verify that a given vector is an eigenvector of a matrix.
2. Verify that a scalar is an eigenvalue of a matrix.
3. Construct an eigenspace for a matrix.
4. Apply theorems related to eigenvalues (for example, to characterize the invertibility of a matrix).

Eigenvectors and Eigenvalues

If $A \in \mathbb{R}^{n \times n}$, and there is a $\vec{v} \neq \vec{0}$ in $\mathbb{R}^n$, and

$$A\vec{v} = \lambda\vec{v}$$

then $\vec{v}$ is an **eigenvector** for A, and $\lambda \in \mathbb{C}$ is the corresponding **eigenvalue**.

Note that

- We will only consider square matrices.
- If $\lambda \in \mathbb{R}$, then
  - when $\lambda > 0$, $A\vec{v}$ and $\vec{v}$ point in the same direction
  - when $\lambda < 0$, $A\vec{v}$ and $\vec{v}$ point in opposite directions
- Even when all entries of A and $\vec{v}$ are real, λ can be complex (a rotation of the plane has no **real** eigenvalues.)
- We explore complex eigenvalues in Section 5.5.

Example 1

Which of the following are eigenvectors of $A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$? What are the corresponding eigenvalues?

a) $\vec{v}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$

b) $\vec{v}_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$

c) $\vec{v}_3 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

Example 2

Confirm that $\lambda = 3$ is an eigenvalue of $A = \begin{pmatrix} 2 & -4 \\ -1 & -1 \end{pmatrix}$.

Eigenspace

Definition

Suppose $A \in \mathbb{R}^{n\times n}$. The eigenvectors for a given λ span a subspace of $\mathbb{R}^n$ called the λ-**eigenspace** of A.

Note: the λ-eigenspace for matrix A is $\mathrm{Nul}(A - \lambda I)$.

Example 3

Construct a basis for the eigenspaces for the matrix whose eigenvalues are given, and sketch the eigenvectors.

$$\begin{pmatrix} 5 & -6 \\ 3 & -4 \end{pmatrix}, \quad \lambda = -1, 2$$

Theorems

Proofs for the most these theorems are in Section 5.1. If time permits, we will explain or prove all/most of these theorems in lecture.

1. The diagonal elements of a triangular matrix are its eigenvalues.

2. A invertible $\Leftrightarrow 0$ is not an eigenvalue of A.

3. Stochastic matrices have an eigenvalue equal to 1.

4. If $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k$ are eigenvectors that correspond to distinct eigenvalues, then $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k$ are linearly independent.

Warning!

We can't determine the eigenvalues of a matrix from its reduced form.

Row reductions change the eigenvalues of a matrix.

Example: suppose $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$. The eigenvalues are $\lambda = 2, 0$, because

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

- But the reduced echelon form of A is:
- The reduced echelon form is triangular, and its eigenvalues are:

Section 5.2 : The Characteristic Equation

Chapter 5 : Eigenvalues and Eigenvectors

Math 1554 Linear Algebra

Topics and Objectives

Topics

We will cover these topics in this section.

1. The characteristic polynomial of a matrix
2. Algebraic and geometric multiplicity of eigenvalues
3. Similar matrices

Objectives

For the topics covered in this section, students are expected to be able to do the following.

1. Construct the characteristic polynomial of a matrix and use it to identify eigenvalues and their multiplicities.
2. Characterize the long-term behaviour of dynamical systems using eigenvalue decompositions.

The Characteristic Polynomial

Recall:

λ is an eigenvalue of $A \iff (A - \lambda I)$ is not ________________

Therefore, to calculate the eigenvalues of A, we can solve

$$\det(A - \lambda I) =$$

The quantity $\det(A - \lambda I)$ is the **characteristic polynomial** of A.

The quantity $\det(A - \lambda I) = 0$ is the **characteristic equation** of A.

The roots of the characteristic polynomial are the _____________ of A.

Example

The characteristic polynomial of $A = \begin{pmatrix} 5 & 2 \\ 2 & 1 \end{pmatrix}$ is:

So the eigenvalues of A are:

Characteristic Polynomial of 2×2 Matrices

Express the characteristic equation of

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

in terms of its determinant. What is the equation when M is singular?

Algebraic Multiplicity

Definition

The **algebraic multiplicity** of an eigenvalue is its multiplicity as a root of the characteristic polynomial.

Example

Compute the algebraic multiplicities of the eigenvalues for the matrix

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Geometric Multiplicity

Definition

The **geometric multiplicity** of an eigenvalue λ is the dimension of $\mathrm{Null}(A - \lambda I)$.

1. Geometric multiplicity is always at least 1. It can be smaller than algebraic multiplicity.
2. Here is the basic example:

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

$\lambda = 0$ is the only eigenvalue. Its algebraic multiplicity is 2, but the geometric multiplicity is 1.

Example

Give an example of a 4×4 matrix with $\lambda = 0$ the only eigenvalue, but the geometric multiplicity of $\lambda = 0$ is one.

Recall: Long-Term Behavior of Markov Chains

Recall:

- We often want to know what happens to a Markov Chain

$$\vec{x}_{k+1} = P\vec{x}_k, \quad k = 0, 1, 2, \ldots$$

  as $k \to \infty$.
- If P is regular, then there is a ________________________

Now lets ask:

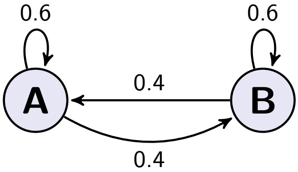
- If we don't know whether P is regular, what else might we do to describe the long-term behavior of the system?
- What can eigenvalues tell us about the behavior of these systems?

Example: Eigenvalues and Markov Chains

Consider the Markov Chain:

$$\vec{x}_{k+1} = P\vec{x}_k = \begin{pmatrix} 0.6 & 0.4 \\ 0.4 & 0.6 \end{pmatrix} \vec{x}_k, \quad k = 0, 1, 2, 3, \ldots, \quad \vec{x}_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

This system can be represented schematically with two nodes, A and B:



Goal: use eigenvalues to describe the long-term behavior of our system.

What are the eigenvalues of P?

What are the corresponding eigenvectors of P?

Use the eigenvalues and eigenvectors of P to analyze the long-term behaviour of the system. In other words, determine what $\vec{x}_k$ tends to as $k \to \infty$.

Similar Matrices

Definition

Two $n \times n$ matrices A and B are **similar** if there is a matrix P so that $A = PBP^{-1}$.

Theorem

If A and B similar, then they have the same characteristic polynomial.

If time permits, we will explain or prove this theorem in lecture. Note:

- Our textbook introduces similar matrices in Section 5.2, but doesn't have exercises on this concept until 5.3.
- Two matrices, A and B, do not need to be similar to have the same eigenvalues. For example,

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

Additional Examples (if time permits)

1. True or false.
   a) If A is similar to the identity matrix, then A is equal to the identity matrix.
   b) A row replacement operation on a matrix does not change its eigenvalues.
2. For what values of k does the matrix have one real eigenvalue with algebraic multiplicity 2?

$$\begin{pmatrix} -3 & k \\ 2 & -6 \end{pmatrix}$$

Section 5.3 : Diagonalization

Chapter 5 : Eigenvalues and Eigenvectors

Math 1554 Linear Algebra

Motivation: it can be useful to take large powers of matrices, for example A^k, for large k.

But: multiplying two $n \times n$ matrices requires roughly n^3 computations. Is there a more efficient way to compute A^k?

Topics and Objectives

Topics

1. Diagonal, similar, and diagonalizable matrices
2. Diagonalizing matrices

Learning Objectives

For the topics covered in this section, students are expected to be able to do the following.

1. Determine whether a matrix can be diagonalized, and if possible diagonalize a square matrix.
2. Apply diagonalization to compute matrix powers.

Diagonal Matrices

A matrix is **diagonal** if the only non-zero elements, if any, are on the main diagonal.

The following are all diagonal matrices.

$$\begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}, \quad \begin{bmatrix} 2 \end{bmatrix}, \quad I_n, \quad \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

We'll only be working with diagonal square matrices in this course.

Powers of Diagonal Matrices

If A is diagonal, then A^k is easy to compute. For example,

$$A = \begin{pmatrix} 3 & 0 \\ 0 & 0.5 \end{pmatrix}$$

$$A^2 =$$

$$A^k =$$

But what if A is not diagonal?

Diagonalization

Suppose $A \in \mathbb{R}^{n \times n}$. We say that A is **diagonalizable** if it is similar to a diagonal matrix, D. That is, we can write

$$A = PDP^{-1}$$

Diagonalization

Theorem

If A is diagonalizable $\Leftrightarrow$ A has n linearly independent eigenvectors.

Note: the symbol $\Leftrightarrow$ means " if and only if ".

Also note that $A = PDP^{-1}$ if and only if

$$A = [\vec{v}_1 \ \vec{v}_2 \cdots \vec{v}_n] \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix} [\vec{v}_1 \ \vec{v}_2 \cdots \vec{v}_n]^{-1}$$

where $\vec{v}_1, \ldots, \vec{v}_n$ are linearly independent eigenvectors, and $\lambda_1, \ldots, \lambda_n$ are the corresponding eigenvalues (**in order**).

Example 1

Diagonalize if possible.

$$\begin{pmatrix} 2 & 6 \\ 0 & -1 \end{pmatrix}$$

Example 2

Diagonalize if possible.

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

Distinct Eigenvalues

Theorem

If A is $n \times n$ and has n distinct eigenvalues, then A is diagonalizable.

Why does this theorem hold?

Is it necessary for an $n \times n$ matrix to have n distinct eigenvalues for it to be diagonalizable?

Non-Distinct Eigenvalues

Theorem. Suppose

- A is $n \times n$
- A has distinct eigenvalues $\lambda_1, \ldots, \lambda_k$, $k \leq n$
- $a_i =$ algebraic multiplicity of λ_i
- $d_i =$ dimension of λ_i eigenspace ("geometric multiplicity")

Then

1. $d_i \leq a_i$ for all i
2. A is diagonalizable $\Leftrightarrow \Sigma d_i = n \Leftrightarrow d_i = a_i$ for all i
3. A is diagonalizable $\Leftrightarrow$ the eigenvectors, for all eigenvalues, together form a basis for $\mathbb{R}^n$.

Example 3

The eigenvalues of A are $\lambda = 3, 1$. If possible, construct P and D such that $AP = PD$.

$$A = \begin{pmatrix} 7 & 4 & 16 \\ 2 & 5 & 8 \\ -2 & -2 & -5 \end{pmatrix}$$

Additional Example (if time permits)

Note that

$$\vec{x}_k = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \vec{x}_{k-1}, \quad \vec{x}_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad k = 1, 2, 3, \ldots$$

generates a well-known sequence of numbers.

Use a diagonalization to find a matrix equation that gives the n^{th} number in this sequence.

Example

$T = Ax$ is the linear transform that:

1. scales vectors in $\mathbb{R}^2$ by a factor of 2, then
2. rotates vectors by $\pi/6$ radians counter-clockwise

Construct the standard matrix for the transformation, A, compute the eiegnvalues of A, and express them in polar form.

Chapter 10 : Finite-State Markov Chains

10.2 : The Steady-State Vector and Page Rank

Topics and Objectives

Topics

1. Review of Markov chains
2. Theorem describing the steady state of a Markov chain
3. Applying Markov chains to model website usage.
4. Calculating the PageRank of a web.

Learning Objectives

1. Determine whether a stochastic matrix is regular.
2. Apply matrix powers and theorems to characterize the long-term behaviour of a Markov chain.
3. Construct a transition matrix, a Markov Chain, and a Google Matrix for a given web, and compute the PageRank of the web.

Where is Chapter 10?

- The material for this part of the course is covered in Section 10.2
- Chapter 10 is not included in the **print** version of the book, but it is in the **on-line version**.
- If you read 10.2, and I recommend that you do, you will find that it requires an understanding of 10.1.
- You are not required to understand the material in 10.1.

Steady State Vectors

Recall the car rental problem from our Section 4.9 lecture.

Problem

A car rental company has 3 rental locations, A, B, and C.

		rented from		
		A	B	C
returned to	A	.8	.1	.2
	B	.2	.6	.3
	C	.0	.3	.5

There are 10 cars at each location today, what happens to the distribution of cars after a long time?

Long Term Behaviour

Can use the transition matrix, P, to find the distribution of cars after 1 week:

$$\vec{x}_1 = P\vec{x}_0$$

The distribution of cars after 2 weeks is:

$$\vec{x}_2 = P\vec{x}_1 = PP\vec{x}_0$$

The distribution of cars after n weeks is:

Long Term Behaviour

To investigate the long-term behaviour of a system that has a regular transition matrix P, we could:

1. compute the **steady-state vector**, $\vec{q}$, by solving $\vec{q} = P\vec{q}$.
2. compute $P^n\vec{x}_0$ for large n.
3. compute P^n for large n, each column of the resulting matrix is the steady-state

Theorem 1

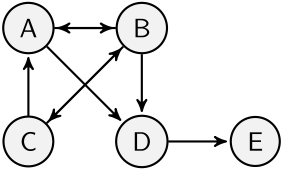
If P is a regular $m \times m$ transition matrix with $m \geq 2$, then the following statements are all true.

1. There is a stochastic matrix Π such that
$$\lim_{n\to\infty} P^n = \Pi$$
2. Each column of Π is the same probability vector $\vec{q}$.
3. For any initial probability vector $\vec{x}_0$,
$$\lim_{n\to\infty} P^n \vec{x}_0 = \vec{q}$$
4. P has a unique eigenvector, $\vec{q}$, which has eigenvalue $\lambda = 1$.
5. The eigenvalues of P satisfy $|\lambda| \leq 1$.

We will apply this theorem when solving PageRank problems.

Example 1

A set of web pages link to each other according to this diagram.



Page A has links to pages ___________ .

Page B has links to pages ___________ .

We make two assumptions:

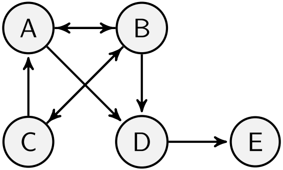
a) A user on a page in this web is equally likely to go to any of the pages that their page links to.

b) If a user is on a page that does not link to other pages, the user stays at that page.

Use these assumptions to construct a Markov chain that represents how users navigate the above web.

Solution

Use the assumptions on the previous slide to construct a Markov chain that represents how users navigate the web.



Transition Matrix, Importance, and PageRank

- The square matrix we constructed in the previous example is a **transition matrix**. It describes how users transition between pages in the web.
- The steady-state vector, $\vec{q}$, for the Markov-chain, can characterize the long-term behavior of users in a given web.
- If $\vec{q}$ is unique, the **importance** of a page in a web is given by its corresponding entry in $\vec{q}$.
- The **PageRank** is the ranking assigned to each page based on its importance. The highest ranked page has PageRank 1, the second PageRank 2, and so on.
- Two pages with same importance receive the same PageRank (some other method would be needed to resolve ties)

Is the transition matrix in Example 1 a regular matrix?

Adjustment 1

Adjustment 1

If a user reaches a page that does not link to other pages, the user will choose any page in the web, with equal probability, and move to that page.

Let's denote this modified transition matrix as P_*. Our transition matrix in Example 1 becomes:

Adjustment 2

Adjustment 2

A user at any page will navigate to any page among those that their page links to with equal probability p, and to any page in the web with equal probability $1-p$. The transition matrix becomes

$$G = pP_* + (1-p)K$$

All the elements of the $n \times n$ matrix K are equal to $1/n$.

p is referred to as the **damping factor**, Google is said to use $p = 0.85$.

With adjustments 1 and 2, our the Google matrix is:

Computing Page Rank

- Because G is stochastic, for any initial probability vector $\vec{x}_0$,

$$\lim_{n \to \infty} G^n \vec{x}_0 = \vec{q}$$

- We can obtain steady-state evaluating $G^n \vec{x}_0$ for large n, by solving $G\vec{q} = \vec{q}$, or by evaluating $\vec{x}_n = G\vec{x}_{n-1}$ for large n.
- Elements of the steady-state vector give the importance of each page in the web, which can be used to determine PageRank.
- Largest element in steady-state vector corresponds to page with PageRank 1, second largest with PageRank 2, and so on.

On an exam,

- problems that require a calculator will not be on your exam
- you may construct your G matrix using factions instead of decimal expansions

There is (of course) Much More to PageRank

The PageRank Algorithm currently used by Google is under constant development, and tailored to individual users.

- When PageRank was devised, in 1996, Yahoo! used humans to provide a "index for the Internet, " which was 10 million pages.
- The PageRank algorithm was produced as a competing method. The patent was awarded to Stanford University, and exclusively licensed to the newly formed Google corporation.
- Brin and Page combined the PageRank algorithm with a webcrawler to provide regular updates to the transition matrix for the web.
- The explosive growth of the web soon overwhelmed human based approaches to searching the internet.

WolframAlpha and MATLAB/Octave Syntax

Suppose we want to compute

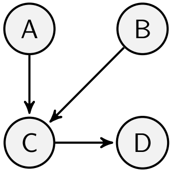
$$\begin{pmatrix} .8 & .1 & .2 \\ .2 & .6 & .3 \\ .0 & .3 & .5 \end{pmatrix}^{10}$$

- At wolframalpha.com, we can use the syntax:
  `MatrixPower[{{.8,.1,.2},{.2,.6,.3},{.0,.3,.5}},10]`
- In MATLAB, we can use the syntax
  `[.8 .1 .2 ;.2 .6 .3;.0 .3 .5]^10`
- Octave uses the same syntax as MATLAB, and there are several free, online, Octave compilers. For example: https://octave-online.net.

You will need to compute a few matrix powers in your homework, and in your future courses, depending on what courses you end up taking.

Example 2 (if time permits)

Construct the Google Matrix for the web below. Which page do you think will have the highest PageRank? How would your result depend on the damping factor p? Use software to explore these questions.



Section 6.1 : Inner Product, Length, and Orthogonality

Chapter 6: Orthogonality and Least Squares

Math 1554 Linear Algebra

Topics and Objectives

Topics

1. Dot product of vectors
2. Magnitude of vectors, and distances in $\mathbb{R}^n$
3. Orthogonal vectors and complements
4. Angles between vectors

Learning Objectives

1. Compute (a) dot product of two vectors, (b) length (or magnitude) of a vector, (c) distance between two points in $\mathbb{R}^n$, and (d) angles between vectors.
2. Apply theorems related to orthogonal complements, and their relationships to Row and Null space, to characterize vectors and linear systems.

Motivating Question

For a matrix A, which vectors are orthogonal to all the rows of A? To the columns of A?

The Dot Product

The dot product between two vectors, $\vec{u}$ and $\vec{v}$ in $\mathbb{R}^n$, is defined as

$$\vec{u} \cdot \vec{v} = \vec{u}^{\,T} \vec{v} = \begin{bmatrix} u_1 & u_2 & \cdots & u_n \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n.$$

Example 1: For what values of k is $\vec{u} \cdot \vec{v} = 0$?

$$\vec{u} = \begin{pmatrix} -1 \\ 3 \\ k \\ 2 \end{pmatrix}, \qquad \vec{v} = \begin{pmatrix} 4 \\ 2 \\ 1 \\ -3 \end{pmatrix}$$

Properties of the Dot Product

The dot product is a special form of matrix multiplication, so it inherits linear properties.

Theorem (Basic Identities of Dot Product)

Let $\vec{u}, \vec{v}, \vec{w}$ be three vectors in $\mathbb{R}^n$, and $c \in \mathbb{R}$.

1. (Symmetry) $\vec{u} \cdot \vec{w} =$ ___________
2. (Linear in each vector) $(\vec{v} + \vec{w}) \cdot \vec{u} =$ ___________
3. (Scalars) $(c\vec{u}) \cdot \vec{w} =$ ___________
4. (Positivity) $\vec{u} \cdot \vec{u} \geq 0$, and the dot product equals ___________

The Length of a Vector

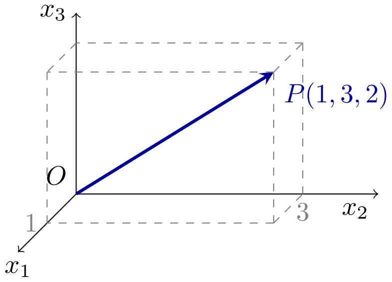
Definition

The **length** of a vector $\vec{u} \in \mathbb{R}^n$ is

$$\|\vec{u}\| = \sqrt{\vec{u} \cdot \vec{u}} = \sqrt{u_1^2 + u_2^2 + \cdots + u_n^2}$$

Example: the length of the vector $\overrightarrow{OP}$ is

$$\sqrt{1^2 + 3^2 + 2^2} = \sqrt{14}$$



Example

Let $\vec{u}, \vec{v}$ be two vectors in $\mathbb{R}^n$ with $\|\vec{u}\| = 5$, $\|\vec{v}\| = \sqrt{3}$, and $\vec{u} \cdot \vec{v} = -1$. Compute the value of $\|\vec{u} + \vec{v}\|$.

Length of Vectors and Unit Vectors

Note: for any vector $\vec{v}$ and scalar c, the length of $c\vec{v}$ is

$$\|c\vec{v}\| = |c| \, ||\vec{v}||$$

Definition

If $\vec{v} \in \mathbb{R}^n$ has length one, we say that it is a **unit vector.**

For example, each of the following vectors are unit vectors.

$$\vec{e}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \vec{y} = \frac{1}{\sqrt{5}} \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \quad \vec{v} = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 0 \\ 1 \\ 1 \end{pmatrix}$$

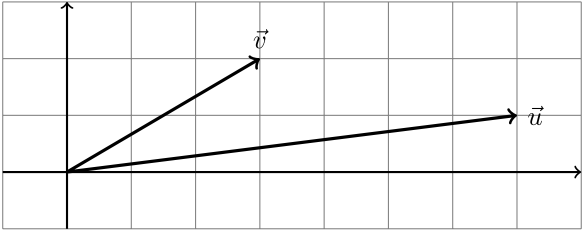
Distance in $\mathbb{R}^n$

Definition

For $\vec{u}, \vec{v} \in \mathbb{R}^n$, the **distance** between $\vec{u}$ and $\vec{v}$ is given by the formula

$$\boxed{}$$

Example: Compute the distance from $\vec{u} = \begin{pmatrix} 7 \\ 1 \end{pmatrix}$ and $\vec{v} = \begin{pmatrix} 3 \\ 2 \end{pmatrix}$.



Orthogonality

Definition (Orthogonal Vectors)

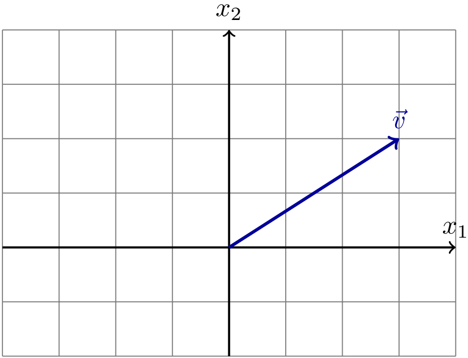
Two vectors $\vec{u}$ and $\vec{w}$ are **orthogonal** if $\vec{u} \cdot \vec{w} = 0$. This is equivalent to:

$$\|\vec{u} + \vec{w}\|^2 =$$

Note: The zero vector in $\mathbb{R}^n$ is orthogonal to every vector in $\mathbb{R}^n$. But we usually only mean non-zero vectors.

Example

Sketch the subspace spanned by the set of all vectors $\vec{u}$ that are orthogonal to $\vec{v} = \begin{pmatrix} 3 \\ 2 \end{pmatrix}$.



Orthogonal Compliments

Definitions

Let W be a subspace of $\mathbb{R}^n$. Vector $\vec{z} \in \mathbb{R}^n$ is **orthogonal** to W if $\vec{z}$ is orthogonal to every vector in W.

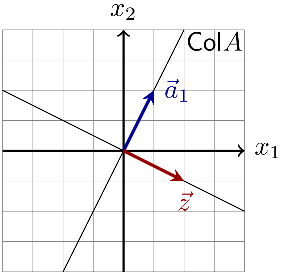
The set of all vectors orthogonal to W is a subspace, the **orthogonal compliment** of W, or $W^\perp$ or 'W perp.'

$$W^\perp = \{\vec{z} \in \mathbb{R}^n \ : \ \vec{z} \cdot \vec{w} = 0 \text{ for all } \vec{w} \in W\}$$

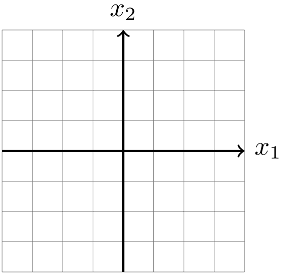
Example

Example: suppose $A = \begin{pmatrix} 1 & 3 \\ 2 & 6 \end{pmatrix}$.

- Col$A$ is the span of $\vec{a}_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$
- Col$A^\perp$ is the span of $\vec{z} = \begin{pmatrix} 2 \\ -1 \end{pmatrix}$

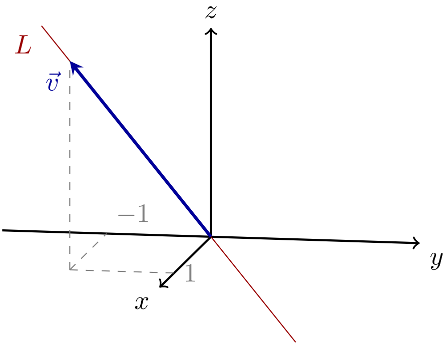


Sketch Null$A$ and Null$A^\perp$ on the grid below.



Example

Line L is a subspace of $\mathbb{R}^3$ spanned by $\vec{v} = \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}$. Then the space $L^\perp$ is a plane. Construct an equation of the plane $L^\perp$.



Can also visualise line and plane with CalcPlot3D: web.monroecc.edu/calcNSF

Row$A$

Definition

Row$A$ is the space spanned by the rows of matrix A.

We can show that

- $\dim(\text{Row}(A)) = \dim(\text{Col}(A))$
- a basis for Row$A$ is the pivot rows of A

Note that $\text{Row}(A) = \text{Col}(A^T)$, but in general Row$A$ and Col$A$ are not related to each other

Example 3

Describe the Null(A) in terms of an orthogonal subspace.

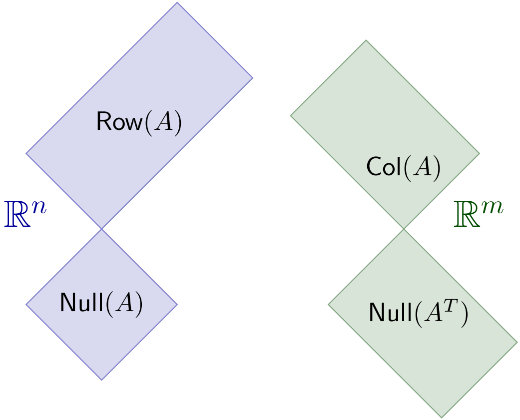
A vector $\vec{x}$ is in $\mathrm{Null}\, A$ if and only if

1. $A\vec{x} =$
2. This means that $\vec{x}$ is $\boxed{}$ to each row of A.
3. $\mathrm{Row}\, A$ is $\boxed{}$ to $\mathrm{Null}\, A$.
4. The dimension of $\mathrm{Row}\, A$ plus the dimension of $\mathrm{Null}\, A$ equals $\boxed{}$

Theorem (The Four Subspaces)

For any $A \in \mathbb{R}^{m\times n}$, the orthogonal complement of $\mathrm{Row}\,A$ is $\mathrm{Null}\,A$, and the orthogonal complement of $\mathrm{Col}\,A$ is $\mathrm{Null}\,A^T$.

The idea behind this theorem is described in the diagram below.



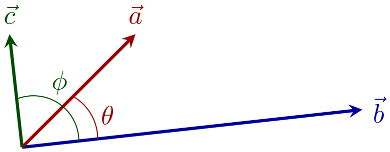
Angles

Theorem

$\vec{a} \cdot \vec{b} = |\vec{a}|\,|\vec{b}| \cos\theta$. Thus, if $\vec{a} \cdot \vec{b} = 0$, then:

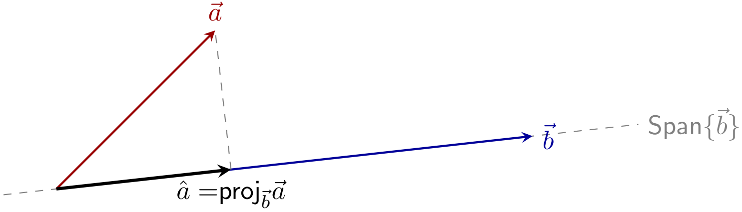
- $\vec{a}$ and/or $\vec{b}$ are ________ vectors, or
- $\vec{a}$ and $\vec{b}$ are ________.

For example, consider the vectors below.



Looking Ahead - Projections

Suppose we want to find the closed vector in Span$\{\vec{b}\}$ to $\vec{a}$.



- Later in this Chapter, we will make connections between dot products and **projections**.
- Projections are also used throughout multivariable calculus courses.

Section 6.2 : Orthogonal Sets

Chapter 6 : Orthogonality and Least Squares

Math 1554 Linear Algebra

Topics and Objectives

Topics

1. Orthogonal Sets of Vectors
2. Orthogonal Bases and Projections.

Learning Objectives

1. Apply the concepts of orthogonality to
   a) compute orthogonal projections and distances,
   b) express a vector as a linear combination of orthogonal vectors,
   c) characterize bases for subspaces of $\mathbb{R}^n$, and
   d) construct orthonormal bases.

Motivating Question

What are the special properties of this basis for $\mathbb{R}^3$?

$$\begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix} / \sqrt{11}, \quad \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix} / \sqrt{6}, \quad \begin{bmatrix} -1 \\ -4 \\ 7 \end{bmatrix} / \sqrt{66}$$

Orthogonal Vector Sets

Definition

A set of vectors $\{\vec{u}_1, \ldots, \vec{u}_p\}$ are an **orthogonal set** of vectors if for each $j \neq k$, $\vec{u}_j \perp \vec{u}_k$.

Example: Fill in the missing entries to make $\{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$ an orthogonal set of vectors.

$$\vec{u}_1 = \begin{bmatrix} 4 \\ 0 \\ 1 \end{bmatrix}, \quad \vec{u}_2 = \begin{bmatrix} \quad -2 \quad \\ 0 \\ \ \end{bmatrix}, \quad \vec{u}_3 = \begin{bmatrix} \quad 0 \quad \\ \ \\ \ \end{bmatrix}$$

Linear Independence

Theorem (Linear Independence for Orthogonal Sets)

Let $\{\vec{u}_1, \ldots, \vec{u}_p\}$ be an orthogonal set of vectors. Then, for scalars $c_1, \ldots, c_p$,

$$\left\|c_1\vec{u}_1 + \cdots + c_p\vec{u}_p\right\|^2 = c_1^2\|\vec{u}_1\|^2 + \cdots + c_p^2\|\vec{u}_p\|^2.$$

In particular, if all the vectors $\vec{u}_r$ are non-zero, the set of vectors $\{\vec{u}_1, \ldots, \vec{u}_p\}$ are linearly independent.

Orthogonal Bases

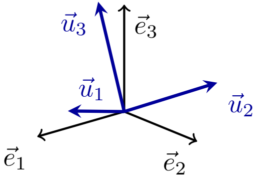
Theorem (Expansion in Orthogonal Basis)

Let $\{\vec{u}_1, \ldots, \vec{u}_p\}$ be an orthogonal basis for a subspace W of $\mathbb{R}^n$. Then, for any vector $\vec{w} \in W$,

$$\vec{w} = c_1\vec{u}_1 + \cdots + c_p\vec{u}_p.$$

Above, the scalars are $c_q = \dfrac{\vec{w} \cdot \vec{u}_q}{\vec{u}_q \cdot \vec{u}_q}$.

For example, any vector $\vec{w} \in \mathbb{R}^3$ can be written as a linear combination of $\{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$, or some other orthogonal basis $\{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$.



Example

$$\vec{x} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad \vec{u} = \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}, \quad \vec{v} = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}, \quad \vec{s} = \begin{pmatrix} 3 \\ -4 \\ 1 \end{pmatrix}$$

Let W be the subspace of $\mathbb{R}^3$ that is orthogonal to $\vec{x}$.

a) Check that an orthogonal basis for W is given by $\vec{u}$ and $\vec{v}$.

b) Compute the expansion of $\vec{s}$ in basis W.

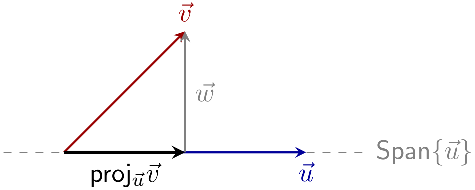
Projections

Let $\vec{u}$ be a non-zero vector, and let $\vec{v}$ be some other vector. The **orthogonal projection of $\vec{v}$ onto the direction of $\vec{u}$** is the vector in the span of $\vec{u}$ that is closest to $\vec{v}$.

$$\boxed{\text{proj}_{\vec{u}}\vec{v} = \frac{\vec{v}\cdot\vec{u}}{\vec{u}\cdot\vec{u}}\vec{u}.}$$

The vector $\vec{w} = \vec{v} - \text{proj}_{\vec{u}}\vec{v}$ is orthogonal to $\vec{u}$, so that

$$\vec{v} = \text{proj}_{\vec{u}}\vec{v} + \vec{w}$$
$$\|\vec{v}\|^2 = \|\text{proj}_{\vec{u}}\vec{v}\|^2 + \|\vec{w}\|^2$$



Example

Let L be spanned by $\vec{u} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$.

1. Calculate the projection of $\vec{y} = (-3, 5, 6, -4)$ onto line L.
2. How close is $\vec{y}$ to the line L?

Definition

Definition (Orthonormal Basis)

An **orthonormal basis** for a subspace W is an orthogonal basis $\{\vec{u}_1, \dots, \vec{u}_p\}$ in which every vector $\vec{u}_q$ has unit length. In this case, for each $\vec{w} \in W$,

$$\vec{w} = (\vec{w} \cdot \vec{u}_1)\vec{u}_1 + \cdots + (\vec{w} \cdot \vec{u}_p)\vec{u}_p$$

$$\|\vec{w}\| = \sqrt{(\vec{w} \cdot \vec{u}_1)^2 + \cdots + (\vec{w} \cdot \vec{u}_p)^2}$$

Example

The subspace W is a subspace of $\mathbb{R}^3$ perpendicular to $x = (1, 1, 1)$. Calculate the missing coefficients in the orthonormal basis for W.

$$u = \frac{1}{\sqrt{\quad}} \begin{bmatrix} 1 \\ 0 \\ \; \end{bmatrix} \qquad v = \frac{1}{\sqrt{\quad}} \begin{bmatrix} \; \\ \; \\ \; \end{bmatrix}$$

Orthogonal Matrices

An **orthogonal matrix** is a square matrix whose columns are orthonormal.

Theorem

An $m \times n$ matrix U has orthonormal columns if and only if $U^T U = I_n$.

Can U have orthonormal columns if $n > m$?

Theorem

Theorem (Mapping Properties of Orthogonal Matrices)

Assume $m \times m$ matrix U has orthonormal columns. Then

1. (Preserves length) $\|U\vec{x}\| = \boxed{}$
2. (Preserves angles) $(U\vec{x}) \cdot (U\vec{y}) = \boxed{}$
3. (Preserves orthogonality)

Example

Compute the length of the vector below.

$$\begin{bmatrix} 1/2 & 2/\sqrt{14} \\ 1/2 & 1/\sqrt{14} \\ 1/2 & -3/\sqrt{14} \\ 1/2 & 0 \end{bmatrix} \begin{bmatrix} \sqrt{2} \\ -3 \end{bmatrix}$$

Section 6.3 : Orthogonal Projections

Chapter 6 : Orthogonality and Least Squares

Math 1554 Linear Algebra

Vectors $\vec{e}_1$ and $\vec{e}_2$ form an orthonormal basis for subspace W.
Vector $\vec{y}$ is not in W.
The orthogonal projection of $\vec{y}$ onto W =Span$\{\vec{e}_1, \vec{e}_2\}$ is $\hat{y}$.

Topics and Objectives

Topics

1. Orthogonal projections and their basic properties
2. Best approximations

Learning Objectives

1. Apply concepts of orthogonality and projections to
   a) compute orthogonal projections and distances,
   b) express a vector as a linear combination of orthogonal vectors,
   c) construct vector approximations using projections,
   d) characterize bases for subspaces of $\mathbb{R}^n$, and
   e) construct orthonormal bases.

Motivating Question For the matrix A and vector $\vec{b}$, which vector $\widehat{b}$ in column space of A, is closest to $\vec{b}$?

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 0 \\ -4 & -2 \end{bmatrix}, \qquad \vec{b} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

Example 1

Let $\vec{u}_1, \dots, \vec{u}_5$ be an orthonormal basis for $\mathbb{R}^5$. Let $W = \mathrm{Span}\{\vec{u}_1, \vec{u}_2\}$. For a vector $\vec{y} \in \mathbb{R}^5$, write $\vec{y} = \widehat{y} + w^\perp$, where $\widehat{y} \in W$ and $w^\perp \in W^\perp$.

Orthogonal Decomposition Theorem

Theorem

Let W be a subspace of $\mathbb{R}^n$. Then, each vector $\vec{y} \in \mathbb{R}^n$ has the **unique** decomposition

$$\vec{y} = \widehat{y} + w^{\perp}, \quad \widehat{y} \in W, \quad w^{\perp} \in W^{\perp}.$$

And, if $\vec{u}_1, \ldots, \vec{u}_p$ is any orthogonal basis for W,

$$\hat{y} = \frac{\vec{y} \cdot \vec{u}_1}{\vec{u}_1 \cdot \vec{u}_1} \vec{u}_1 + \cdots + \frac{\vec{y} \cdot \vec{u}_p}{\vec{u}_p \cdot \vec{u}_p} \vec{u}_p.$$

We say that $\widehat{y}$ is the **orthogonal projection of $\vec{y}$ onto W.**

If time permits, we will explain some of this theorem on the next slide.

Explanation (if time permits)

We can write

$$\widehat{y} =$$

Then, $w^{\perp} = \vec{y} - \widehat{y}$ is in $W^{\perp}$ because

Example 2a

$$\vec{y} = \begin{pmatrix} 4 \\ 0 \\ 3 \end{pmatrix}, \quad \vec{u}_1 = \begin{pmatrix} 2 \\ 2 \\ 0 \end{pmatrix}, \quad \vec{u}_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

Construct the decomposition $\vec{y} = \widehat{y} + w^{\perp}$, where $\widehat{y}$ is the orthogonal projection of $\vec{y}$ onto the subspace $W = \mathrm{Span}\{\vec{u}_1, \vec{u}_2\}$.

Best Approximation Theorem

Theorem

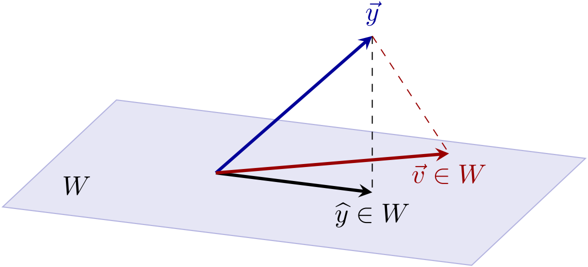
Let W be a subspace of $\mathbb{R}^n$, $\vec{y} \in \mathbb{R}^n$, and $\widehat{y}$ is the orthogonal projection of $\vec{y}$ onto W. Then for **any** $\vec{w} \neq \hat{y} \in W$, we have

$$\|\vec{y} - \widehat{y}\| < \|\vec{y} - \vec{w}\|$$

That is, $\widehat{y}$ is the unique vector in W that is closest to $\vec{y}$.

Proof (if time permits)

The orthogonal projection of $\vec{y}$ onto W is the closest point in W to $\vec{y}$.



Example 2b

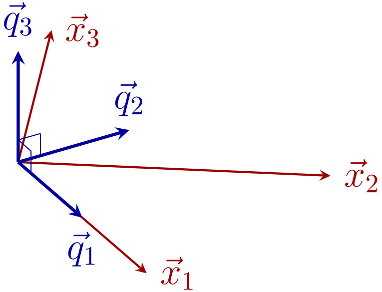
$$\vec{y} = \begin{pmatrix} 4 \\ 0 \\ 3 \end{pmatrix}, \quad \vec{u}_1 = \begin{pmatrix} 2 \\ 2 \\ 0 \end{pmatrix}, \quad \vec{u}_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

What is the distance between $\vec{y}$ and subspace $W = \mathrm{Span}\{\vec{u}_1, \vec{u}_2\}$? Note that these vectors are the same vectors that we used in Example 2a.

Section 6.4 : The Gram-Schmidt Process

Chapter 6 : Orthogonality and Least Squares

Math 1554 Linear Algebra



Vectors $\vec{x}_1, \vec{x}_2, \vec{x}_3$ are given linearly independent vectors. We wish to construct an orthonormal basis $\{\vec{q}_1, \vec{q}_2, \vec{q}_3\}$ for the space that they span.

Topics and Objectives

Topics

1. Gram Schmidt Process
2. The QR decomposition of matrices and its properties

Learning Objectives

1. Apply the iterative Gram Schmidt Process, and the QR decomposition, to construct an orthogonal basis.
2. Compute the QR factorization of a matrix.

Motivating Question The vectors below span a subspace W of $\mathbb{R}^4$. Identify an orthogonal basis for W.

$$\vec{x}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \quad \vec{x}_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \quad \vec{x}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}.$$

Example

The vectors below span a subspace W of $\mathbb{R}^4$. Construct an orthogonal basis for W.

$$\vec{x}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \quad \vec{x}_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \quad \vec{x}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}.$$

The Gram-Schmidt Process

Given a basis $\{\vec{x}_1, \ldots, \vec{x}_p\}$ for a subspace W of $\mathbb{R}^n$, iteratively define

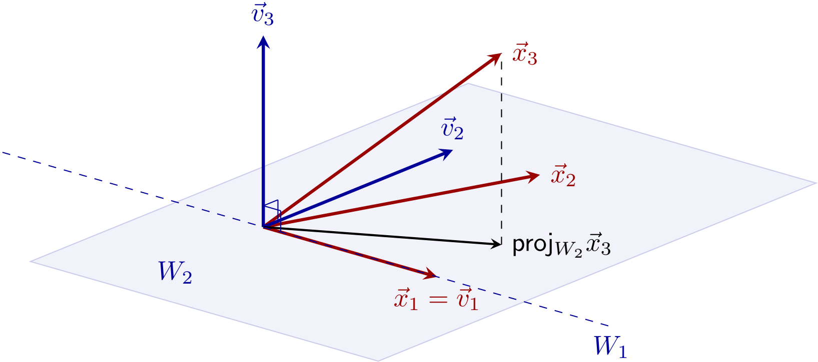
$$\begin{aligned}
\vec{v}_1 &= \vec{x}_1 \\
\vec{v}_2 &= \vec{x}_2 - \frac{\vec{x}_2 \cdot \vec{v}_1}{\vec{v}_1 \cdot \vec{v}_1}\vec{v}_1 \\
\vec{v}_3 &= \vec{x}_3 - \frac{\vec{x}_3 \cdot \vec{v}_1}{\vec{v}_1 \cdot \vec{v}_1}\vec{v}_1 - \frac{\vec{x}_3 \cdot \vec{v}_2}{\vec{v}_2 \cdot \vec{v}_2}\vec{v}_2 \\
&\vdots \\
\vec{v}_p &= \vec{x}_p - \frac{\vec{x}_p \cdot \vec{v}_1}{\vec{v}_1 \cdot \vec{v}_1}\vec{v}_1 - \cdots - \frac{\vec{x}_p \cdot \vec{v}_{p-1}}{\vec{v}_{p-1} \cdot \vec{v}_{p-1}}\vec{v}_{p-1}
\end{aligned}$$

Then, $\{\vec{v}_1, \ldots, \vec{v}_p\}$ is an orthogonal basis for W.

Proof

Geometric Interpretation

Suppose $\vec{x}_1, \vec{x}_2, \vec{x}_3$ are linearly independent vectors in $\mathbb{R}^3$. We wish to construct an orthogonal basis for the space that they span.



We construct vectors $\vec{v}_1, \vec{v}_2, \vec{v}_3$, which form our **orthogonal** basis.
$W_1 = \text{Span}\{\vec{v}_1\}$, $W_2 = \text{Span}\{\vec{v}_1, \vec{v}_2\}$.

Orthonormal Bases

Definition

A set of vectors form an **orthonormal basis** if the vectors are mutually orthogonal and have unit length.

Example

The two vectors below form an orthogonal basis for a subspace W. Obtain an orthonormal basis for W.

$$\vec{v}_1 = \begin{bmatrix} 3 \\ 2 \\ 0 \end{bmatrix}, \quad \vec{v}_2 = \begin{bmatrix} -2 \\ 3 \\ 1 \end{bmatrix}.$$

QR Factorization

Theorem

Any $m \times n$ matrix A with linearly independent columns has the **QR factorization**

$$A = QR$$

where

1. Q is $m \times n$, its columns are an orthonormal basis for $\operatorname{Col} A$.
2. R is $n \times n$, upper triangular, with positive entries on its diagonal, and the length of the j^{th} column of R is equal to the length of the j^{th} column of A.

In the interest of time:

- we will not consider the case where A has linearly dependent columns
- students are not expected to know the conditions for which A has a QR factorization

Proof

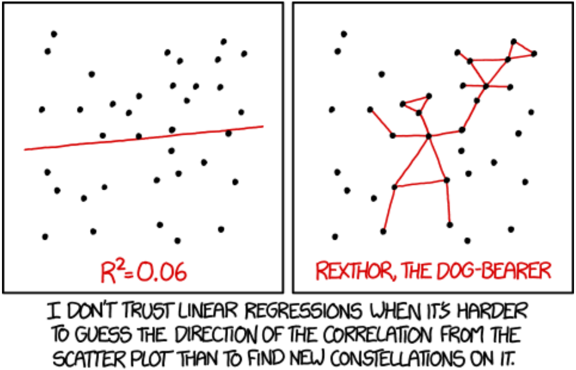
Example

Construct the QR decomposition for $A = \begin{bmatrix} 3 & -2 \\ 2 & 3 \\ 0 & 1 \end{bmatrix}$.

Section 6.5 : Least-Squares Problems

Chapter 6 : Orthogonality and Least Squares

Math 1554 Linear Algebra



https://xkcd.com/1725

Topics and Objectives

Topics

1. Least Squares Problems
2. Different methods to solve Least Squares Problems

Learning Objectives

1. Compute general solutions, and least squares errors, to least squares problems using the normal equations and the QR decomposition.

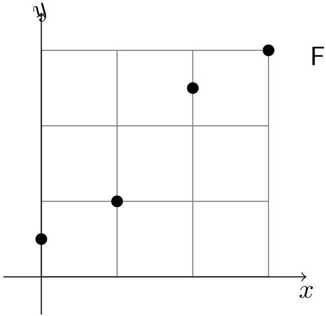
Motivating Question A series of measurements are corrupted by random errors. How can the dominant trend be extracted from the measurements with random error?

Inconsistent Systems

Suppose we want to construct a line of the form

$$y = mx + b$$

that best fits the data below.



From the data, we can construct the system:

$$\begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} b \\ m \end{bmatrix} = \begin{bmatrix} 0.5 \\ 1 \\ 2.5 \\ 3 \end{bmatrix}$$

Can we 'solve' this inconsistent system?

The Least Squares Solution to a Linear System

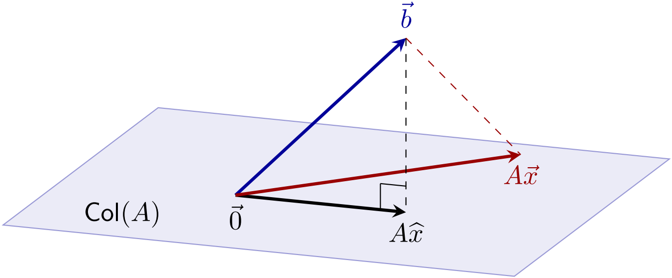
Definition: Least Squares Solution

Let A be a $m \times n$ matrix. A **least squares solution to** $A\vec{x} = \vec{b}$ is the solution $\widehat{x}$ for which

$$\| \vec{b} - A\widehat{x} \| \leq \| \vec{b} - A\vec{x} \|$$

for all $\vec{x} \in \mathbb{R}^n$.

A Geometric Interpretation

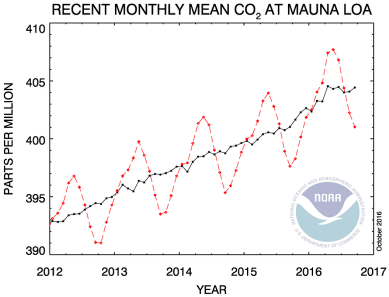


The vector $\vec{b}$ is closer to $A\hat{x}$ than to $A\vec{x}$ for all other $\vec{x} \in$ Col$A$.

1. If $\vec{b} \in \operatorname{Col} A$, then $\widehat{x}$ is ...
2. Seek $\widehat{x}$ so that $A\widehat{x}$ is as close to $\vec{b}$ as possible. That is, $\widehat{x}$ should solve $A\widehat{x} = \widehat{b}$ where $\widehat{b}$ is ...

Important Examples: Overdetermined Systems (Tall/Thin Matrices)

A variety of factors impact the measured quantity.



In the above figure, the dashed red line with diamond symbols represents the monthly mean values, centered on the middle of each month. The black line with the square symbols represents the same, after correction for the average seasonal cycle. (NOAA graph.)

Previous data is the important time series of mean CO_2 in the atmosphere. The data is collected at the Mauna Loa observatory on the island of Hawaii (The Big Island). One of the most important observatories in the world, it is located at the top of the Mauna Kea volcano, 4,205 meters altitude.

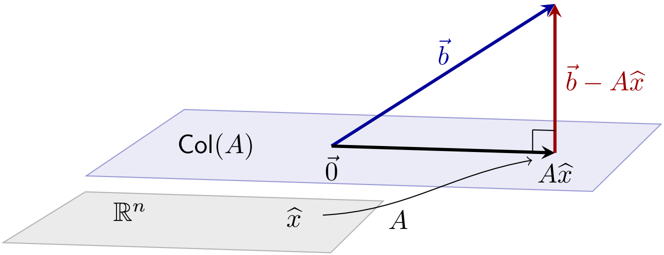
The Normal Equations

Theorem (Normal Equations for Least Squares)

The least squares solutions to $A\vec{x} = \vec{b}$ coincide with the solutions to

$$\underbrace{A^T A\vec{x} = A^T\vec{b}}_{\text{Normal Equations}}$$

Derivation



The least-squares solution $\hat{x}$ is in $\mathbb{R}^n$.

1. $\widehat{x}$ is the least squares solution, is equivalent to $\vec{b} - A\widehat{x}$ is orthogonal to $\boxed{}A$.
2. A vector $\vec{v}$ is in $\mathrm{Null}\, A^T$ if and only if $\boxed{}\vec{v} = \vec{0}$.
3. So we obtain the Normal Equations:

Example

Compute the least squares solution to $A\vec{x} = \vec{b}$, where

$$A = \begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{bmatrix}, \qquad \vec{b} = \begin{bmatrix} 2 \\ 0 \\ 11 \end{bmatrix}$$

Solution:

$$A^T A = \begin{bmatrix} 4 & 0 & 1 \\ 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{bmatrix} =$$

$$A^T \vec{b} = \begin{bmatrix} 4 & 0 & 1 \\ 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 0 \\ 11 \end{bmatrix} =$$

The normal equations $A^T A\vec{x} = A^T\vec{b}$ become:

Theorem

Theorem (Unique Solutions for Least Squares)

Let A be any $m \times n$ matrix. These statements are equivalent.

1. The equation $A\vec{x} = \vec{b}$ has a unique least-squares solution for each $\vec{b} \in \mathbb{R}^m$.
2. The columns of A are linearly independent.
3. The matrix $A^T A$ is invertible.

And, if these statements hold, the least square solution is

$$\widehat{x} = (A^T A)^{-1} A^T \vec{b}.$$

Useful heuristic: $A^T A$ plays the role of 'length-squared' of the matrix A. (See the sections on symmetric matrices and singular value decomposition.)

Example

Compute the least squares solution to $A\vec{x} = \vec{b}$, where

$$A = \begin{bmatrix} 1 & -6 \\ 1 & -2 \\ 1 & 1 \\ 1 & 7 \end{bmatrix}, \qquad \vec{b} = \begin{bmatrix} -1 \\ 2 \\ 1 \\ 6 \end{bmatrix}$$

Hint: the columns of A are orthogonal.

Theorem (Least Squares and QR)

Let $m \times n$ matrix A have a QR decomposition. Then for each $\vec{b} \in \mathbb{R}^m$ the equation $A\vec{x} = \vec{b}$ has the unique least squares solution

$$R\widehat{x} = Q^T \vec{b}.$$

(Remember, R is upper triangular, so the equation above is solved by back-substitution.)

Example 3. Compute the least squares solution to $A\vec{x} = \vec{b}$, where

$$A = \begin{bmatrix} 1 & 3 & 5 \\ 1 & 1 & 0 \\ 1 & 1 & 2 \\ 1 & 3 & 3 \end{bmatrix}, \qquad \vec{b} = \begin{bmatrix} 3 \\ 5 \\ 7 \\ -3 \end{bmatrix}$$

Solution. The QR decomposition of A is

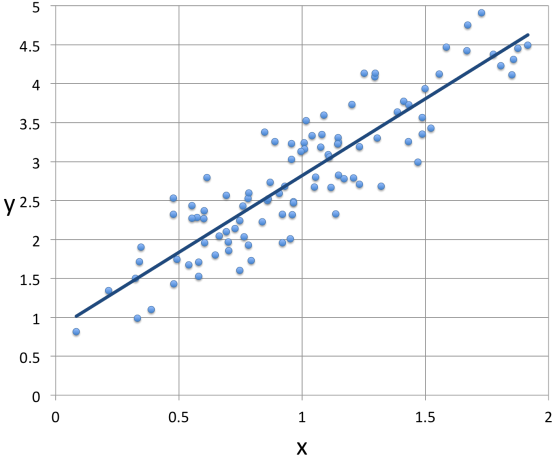
$$A = QR = \frac{1}{2} \begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & -1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \end{bmatrix} \begin{bmatrix} 2 & 4 & 5 \\ 0 & 2 & 3 \\ 0 & 0 & 2 \end{bmatrix}$$

$$Q^T\vec{b} = \tfrac{1}{2}\begin{bmatrix}1 & 1 & 1 & 1\\ 1 & -1 & -1 & 1\\ 1 & -1 & 1 & -1\end{bmatrix}\begin{bmatrix}3\\ 5\\ 7\\ -3\end{bmatrix} = \begin{bmatrix} \\ -6\\ 4\end{bmatrix}$$

And then we solve by backwards substitution $R\vec{x} = Q^T\vec{b}$

$$\begin{bmatrix}2 & 4 & 5\\ 0 & 2 & 3\\ 0 & 0 & 2\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix} = \begin{bmatrix} \\ -6\\ 4\end{bmatrix}$$

Chapter 6 : Orthogonality and Least Squares
6.6 : Applications to Linear Models



Topics and Objectives

Topics

1. Least Squares Lines
2. Linear and more complicated models

Learning Objectives

For the topics covered in this section, students are expected to be able to do the following.

1. Apply least-squares and multiple regression to construct a linear model from a set of data points.
2. Apply least-squares to fit polynomials and other curves to data.

Motivating Question

Compute the equation of the line $y = \beta_0 + \beta_1 x$ that best fits the data

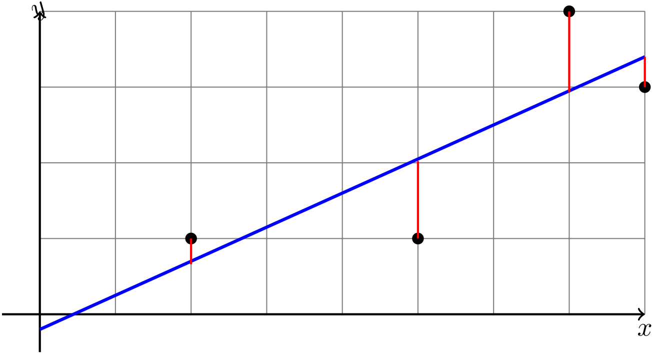
x	2	5	7	8
y	1	1	4	3

The Least Squares Line

Graph below gives an approximate linear relationship between x and y.

1. Black circles are data.
2. Blue line is the **least squares** line.
3. Lengths of red lines are the ___________.

The least squares line minimizes the sum of squares of the ___________.



Example 1 Compute the least squares line $y = \beta_0 + \beta_1 x$ that best fits the data

x	2	5	7	8
y	1	1	4	3

We want to solve

$$\begin{bmatrix} 1 & 2 \\ 1 & 5 \\ 1 & 7 \\ 1 & 8 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 4 \\ 3 \end{bmatrix}$$

This is a least-squares problem : $X\vec{\beta} = \vec{y}$.

The normal equations are

$$X^T X = \begin{bmatrix} 1 & 1 & 1 & 1 \\ & & & \end{bmatrix} \begin{bmatrix} 1 & \\ 1 & \\ 1 & \\ 1 & \end{bmatrix} = \begin{bmatrix} 4 & 22 \\ 22 & 142 \end{bmatrix}$$

$$X^T \vec{y} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ & & & \end{bmatrix} \begin{bmatrix} \\ \\ \\ \\ \end{bmatrix} = \begin{bmatrix} 9 \\ 59 \end{bmatrix}$$

So the least-squares solution is given by

$$\begin{bmatrix} 4 & 22 \\ 22 & 142 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \begin{bmatrix} 9 \\ 59 \end{bmatrix}$$

$$y = \beta_0 + \beta_1 x = \frac{-5}{21} + \frac{19}{42} x$$

As we may have guessed, β_0 is negative, and β_1 is positive.

Least Squares Fitting for Other Curves

We can consider least squares fitting for the form

$$y = c_0 + c_1 f_1(x) + c_2 f_2(x) + \cdots + c_k f_k(x).$$

If functions f_i are known, this is a linear problem in the c_i variables.

Example

Consider the data in the table below.

x	-1	0	0	1
y	2	1	0	6

Determine the coefficients c_1 and c_2 for the curve $y = c_1 x + c_2 x^2$ that best fits the data.

WolframAlpha and Mathematica Syntax

Least squares problems can be computed with WolframAlpha, Mathematica, and many other software.

WolframAlpha

$$\texttt{linear fit } \{\{x_1, y_1\}, \{x_2, y_2\}, \ldots, \{x_n, y_n\}\}$$

Mathematica

$$\texttt{LeastSquares}[\{\{x_1, x_1, y_1\}, \{x_2, x_2, y_2\}, \ldots, \{x_n, x_n, y_n\}\}]$$

Almost any spreadsheet program does this as a function as well.

Section 7.1 : Diagonalization of Symmetric Matrices

Chapter 7: Orthogonality and Least Squares

Math 1554 Linear Algebra

Topics and Objectives

Topics

1. Symmetric matrices
2. Orthogonal diagonalization
3. Spectral decomposition

Learning Objectives

1. Construct an orthogonal diagonalization of a symmetric matrix, $A = PDP^T$.
2. Construct a spectral decomposition of a matrix.

Symmetric Matrices

Definition

Matrix A is **symmetric** if $A^T = A$.

Example. Which of the following matrices are symmetric? Symbols $*$ and $\star$ represent real numbers.

$$A = [*] \qquad B = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \qquad C = \begin{bmatrix} 4 & 0 \\ 0 & 0 \end{bmatrix}$$

$$D = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \qquad E = \begin{bmatrix} 4 & 2 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \qquad F = \begin{bmatrix} 4 & 2 & 0 & 1 \\ 2 & 0 & 7 & 4 \\ 0 & 7 & 6 & 0 \\ 1 & 4 & 0 & 3 \end{bmatrix}$$

A^TA is Symmetric

A very common example: For **any** matrix A with columns $a_1, \ldots, a_n$,

$$A^TA = \begin{bmatrix} -- & a_1^T & -- \\ -- & a_2^T & -- \\ \vdots & \vdots & \vdots \\ -- & a_n^T & -- \end{bmatrix} \begin{bmatrix} | & | & \cdots & | \\ a_1 & a_2 & \cdots & a_n \\ | & | & \cdots & | \end{bmatrix}$$

$$= \underbrace{\begin{bmatrix} a_1^Ta_1 & a_1^Ta_2 & \cdots & a_1^Ta_n \\ a_2^Ta_1 & a_2^Ta_2 & \cdots & a_2^Ta_n \\ \vdots & \vdots & \ddots & \vdots \\ a_n^Ta_1 & a_n^Ta_2 & \cdots & a_n^Ta_n \end{bmatrix}}_{\text{Entries are the dot products of columns of } A}$$

Symmetric Matrices and their Eigenspaces

Theorem

A is a symmetric matrix, with eigenvectors $\vec{v}_1$ and $\vec{v}_2$ corresponding to two distinct eigenvalues. Then $\vec{v}_1$ and $\vec{v}_2$ are orthogonal.

More generally, eigenspaces associated to distinct eigenvalues are orthogonal subspaces.

Proof:

Example 1

Diagonalize A using an orthogonal matrix. Eigenvalues of A are given.

$$A = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad \lambda = -1, 1$$

Hint: Gram-Schmidt

Spectral Theorem

Recall: If P is an orthogonal $n \times n$ matrix, then $P^{-1} = P^T$, which implies $A = PDP^T$ is diagonalizable and symmetric.

Theorem: Spectral Theorem

An $n \times n$ symmetric matrix A has the following properties.

1. All eigenvalues of A are real.
2. The dimenison of each eigenspace is full, that it's dimension is **equal to** it's algebraic multiplicity.
3. The eigenspaces are mutually orthogonal.
4. A can be diagonalized: $A = PDP^T$, where D is diagonal and P is orthogonal.

Proof (if time permits):

Spectral Decomposition of a Matrix

Spectral Decomposition

Suppose A can be orthogonally diagonalized as

$$A = PDP^T = \begin{bmatrix} \vec{u}_1 & \cdots & \vec{u}_n \end{bmatrix} \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix} \begin{bmatrix} \vec{u}_1^T \\ \vdots \\ \vec{u}_n^T \end{bmatrix}$$

Then A has the decomposition

$$A = \lambda_1 \vec{u}_1 \vec{u}_1^T + \cdots + \lambda_n \vec{u}_n \vec{u}_n^T = \sum_{i=1}^{n} \lambda_i \vec{u}_i \vec{u}_i^T$$

Each term in the sum, $\lambda_i \vec{u}_i \vec{u}_i^T$, is an $n \times n$ matrix with rank 1.

Example 2

Construct a spectral decomposition for A whose orthogonal diagonalization is given.

$$
\begin{aligned}
A = \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix} &= PDP^T \\
&= \begin{pmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix} \begin{pmatrix} 4 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix}
\end{aligned}
$$

Section 7.2 : Quadratic Forms

Chapter 7: Orthogonality and Least Squares

Math 1554 Linear Algebra

Topics and Objectives

Topics

1. Quadratic forms
2. Change of variables
3. Principle axes theorem
4. Classifying quadratic forms

Learning Objectives

1. Characterize and classify quadratic forms using eigenvalues and eigenvectors.
2. Express quadratic forms in the form $Q(\vec{x}) = \vec{x}^T A\vec{x}$.
3. Apply the principle axes theorem to express quadratic forms with no cross-product terms.

Motivating Question Does this inequality hold for all x, y?

$$x^2 - 6xy + 9y^2 \geq 0$$

Quadratic Forms

Definition

A **quadratic form** is a function $Q : \mathbb{R}^n \to \mathbb{R}$, given by

$$Q(\vec{x}) = \vec{x}^T A\vec{x} = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{12} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{bmatrix}$$

Matrix A is $n \times n$ and symmetric.

In the above, $\vec{x}$ is a vector of variables.

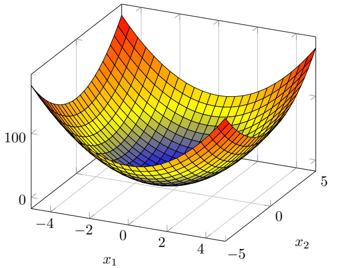
Example 1

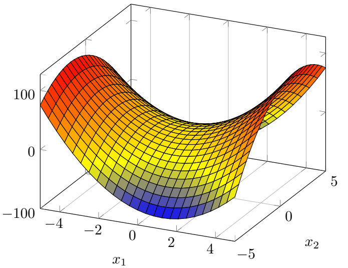
Compute the quadratic form $\vec{x}^T A \vec{x}$ for the matrices below.

$$A = \begin{bmatrix} 4 & 0 \\ 0 & 3 \end{bmatrix}, \qquad B = \begin{bmatrix} 4 & 1 \\ 1 & -3 \end{bmatrix}$$

Example 1 - Surface Plots

The surfaces for Example 1 are shown below.





Students are not expected to be able to sketch quadratic surfaces, but it is helpful to see what they look like.

Example 2

Write Q in the form $\vec{x}^T A\vec{x}$ for $\vec{x} \in \mathbb{R}^3$.

$$Q(x) = 5x_1^2 - x_2^2 + 3x_3^2 + 6x_1x_3 - 12x_2x_3$$

Change of Variable

If $\vec{x}$ is a variable vector in $\mathbb{R}^n$, then a **change of variable** can be represented as

$$\vec{x} = P\vec{y}, \quad \text{or} \quad \vec{y} = P^{-1}\vec{x}$$

With this change of variable, the quadratic form $\vec{x}^T A\vec{x}$ becomes:

Example 3

Make a change of variable $\vec{x} = P\vec{y}$ that transforms $Q = \vec{x}^T A \vec{x}$ so that it does not have cross terms. The orthogonal decomposition of A is given.

$$A = \begin{pmatrix} 3 & 2 \\ 2 & 6 \end{pmatrix} = PDP^T$$

$$P = \frac{1}{\sqrt{5}} \begin{pmatrix} 2 & 1 \\ -1 & 2 \end{pmatrix}$$

$$D = \begin{pmatrix} 2 & 0 \\ 0 & 7 \end{pmatrix}$$

Geometry

Suppose $Q(\vec{x}) = \vec{x}^T A\vec{x}$, where $A \in \mathbb{R}^{n\times n}$ is symmetric. Then the set of $\vec{x}$ that satisfies

$$C = \vec{x}^T A\vec{x}$$

defines a curve or surface in $\mathbb{R}^n$.

Principle Axes Theorem

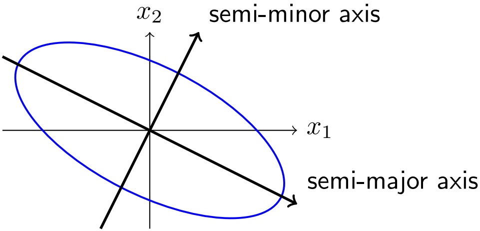
Theorem

If A is a ____________________________ matrix then there exists an orthogonal change of variable $\vec{x} = P\vec{y}$ that transforms $\vec{x}^T A\vec{x}$ to $\vec{x}^T D\vec{x}$ with no cross-product terms.

Proof (if time permits):

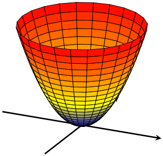
Example 5

Compute the quadratic form $Q = \vec{x}^T A \vec{x}$ for $A = \begin{pmatrix} 5 & 2 \\ 2 & 8 \end{pmatrix}$, and find a change of variable that removes the cross-product term. A sketch of Q is below.

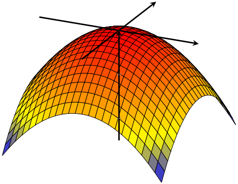


Classifying Quadratic Forms

$Q = x_1^2 + x_2^2$



$Q = -x_1^2 - x_2^2$



Definition

A quadratic form Q is

1. **positive definite** if ____________ for all $\vec{x} \neq \vec{0}$.
2. **negative definite** if ____________ for all $\vec{x} \neq \vec{0}$.
3. **positive semidefinite** if ____________ for all $\vec{x}$.
4. **negative semidefinite** if ____________ for all $\vec{x}$.
5. **indefinite** if ________________________

Quadratic Forms and Eigenvalues

Theorem

If A is a ____________________ matrix with eigenvalues λ_i, then $Q = \vec{x}^T A \vec{x}$ is

1. **positive definite** iff λ_i __________
2. **negative definite** iff λ_i __________
3. **indefinite** iff λ_i __________

Proof (if time permits):

Example 6

We can now return to our motivating question (from first slide): does this inequality hold for all x, y?

$$x^2 - 6xy + 9y^2 \geq 0$$

Section 7.3 : Constrained Optimization

Chapter 7: Orthogonality and Least Squares

Math 1554 Linear Algebra

Topics and Objectives

Topics

1. Constrained optimization as an eigenvalue problem
2. Distance and orthogonality constraints

Learning Objectives

1. Apply eigenvalues and eigenvectors to solve optimization problems that are subject to distance and orthogonality constraints.

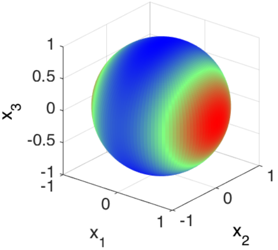
Example 1

The surface of a unit sphere in $\mathbb{R}^3$ is given by

$$1 = x_1^2 + x_2^2 + x_3^2 = ||\vec{x}||^2$$

Q is a quantity we want to optimize

$$Q(\vec{x}) = 9x_1^2 + 4x_2^2 + 3x_3^2$$



Find the largest and smallest values of Q on the surface of the sphere.

A Constrained Optimization Problem

Suppose we wish to find the maximum or minimum values of

$$Q(\vec{x}) = \vec{x}^T A \vec{x}$$

subject to

$$||\vec{x}|| = 1$$

That is, we want to find

$$\begin{aligned} m &= \min\{Q(\vec{x}) : ||\vec{x}|| = 1\} \\ M &= \max\{Q(\vec{x}) : ||\vec{x}|| = 1\} \end{aligned}$$

This is an example of a **constrained optimization** problem. Note that we may also want to know where these extreme values are obtained.

Constrained Optimization and Eigenvalues

Theorem

If $Q = \vec{x}^T A \vec{x}$, A is a real $n \times n$ symmetric matrix, with eigenvalues

$$\lambda_1 \geq \lambda_2 \ldots \geq \lambda_n$$

and associated normalized eigenvectors

$$\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_n$$

Then, subject to the constraint $||\vec{x}|| = 1$,

- the **maximum** value of $Q(\vec{x}) = \lambda_1$, attained at $\vec{x} = \pm\,\vec{u}_1$.
- the **minimum** value of $Q(\vec{x}) = \lambda_n$, attained at $\vec{x} = \pm\,\vec{u}_n$.

Proof:

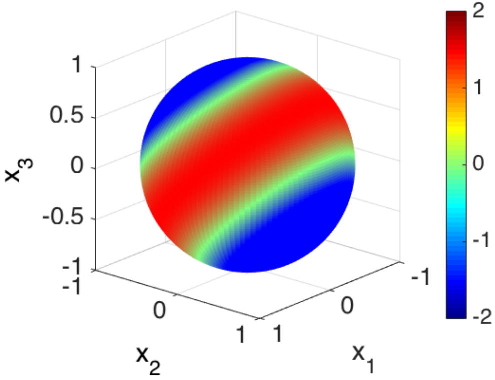
Example 2

Calculate the maximum and minimum values of $Q(\vec{x}) = \vec{x}^{\,T} A\vec{x}$, $\vec{x} \in \mathbb{R}^3$, subject to $||\vec{x}|| = 1$, and identify points where these values are obtained.

$$Q(\vec{x}) = x_1^2 + 2x_2x_3$$

Example 2

The image below is the unit sphere whose surface is colored according to the quadratic from the previous example. Notice the agreement between our solution and the image.



An Orthogonality Constraint

Theorem

Suppose $Q = \vec{x}^T A\vec{x}$, A is a real $n \times n$ symmetric matrix, with eigenvalues

$$\lambda_1 \geq \lambda_2 \ldots \geq \lambda_n$$

and associated eigenvectors

$$\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_n$$

Subject to the constraints $||\vec{x}|| = 1$ and $\vec{x} \cdot \vec{u}_1 = 0$,

- The maximum value of $Q(\vec{x}) = \lambda_2$, attained at $\vec{x} = \vec{u}_*$.
- The minimum value of $Q(\vec{x}) = \lambda_n$, attained at $\vec{x} = \vec{u}_n$.

Note that λ_2 is the second largest eigenvalue of A.

Example 3

Calculate the maximum value of $Q(\vec{x}) = \vec{x}^T A \vec{x}$, $\vec{x} \in \mathbb{R}^3$, subject to $||\vec{x}|| = 1$ and to $\vec{x} \cdot \vec{u}_1 = 0$, and identify a point where this maximum is obtained.

$$Q(\vec{x}) = x_1^2 + 2x_2x_3, \qquad \vec{u}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

Example 4 (if time permits)

Calculate the maximum value of $Q(\vec{x}) = \vec{x}^{\,T} A \vec{x}$, $\vec{x} \in \mathbb{R}^3$, subject to $||\vec{x}|| = 5$, and identify a point where this maximum is obtained.

$$Q(\vec{x}) = x_1^2 + 2x_2x_3$$

Section 7.4 : The Singular Value Decomposition

Chapter 7: Orthogonality and Least Squares

Math 1554 Linear Algebra

Topics and Objectives

Topics

1. The Singular Value Decomposition (SVD) and some of its applications.

Learning Objectives

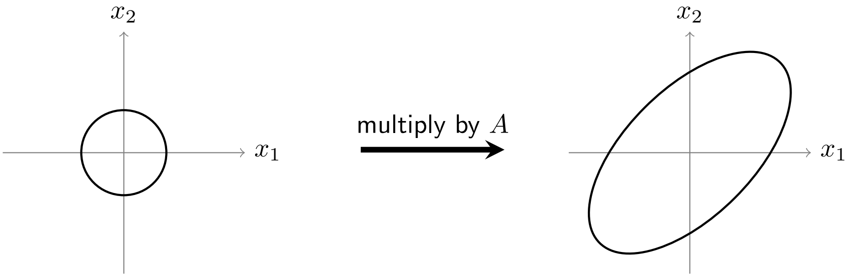
1. Compute the SVD for a rectangular matrix.
2. Apply the SVD to
   - estimate the rank and condition number of a matrix,
   - construct a basis for the four fundamental spaces of a matrix, and
   - construct a spectral decomposition of a matrix.

Example 1

The linear transform whose standard matrix is

$$A = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 2\sqrt{2} & 0 \\ 0 & \sqrt{2} \end{pmatrix} = \begin{pmatrix} 2 & -1 \\ 2 & 1 \end{pmatrix}$$

maps the unit circle in $\mathbb{R}^2$ to an ellipse, as shown below. Identify the unit vector $\vec{x}$ in which $||A\vec{x}||$ is maximized and compute this length.



Example 1 - Solution

Singular Values

The matrix $A^T A$ is always symmetric, with non-negative eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$. Let $\{\vec{v}_1, \ldots, \vec{v}_n\}$ be the associated orthonormal eigenvectors. Then

$$\|A\vec{v}_j\|^2 =$$

If the A has rank r, then $\{A\vec{v}_1, \ldots, A\vec{v}_r\}$ is an orthogonal basis for Col$A$: For $1 \leq j < k \leq r$:

$$(A\vec{v}_j)^T A\vec{v}_k =$$

Definition: $\sigma_1 = \sqrt{\lambda_1} \geq \sigma_2 = \sqrt{\lambda_2} \cdots \geq \sigma_n = \sqrt{\lambda_n}$ are the singular values of A.

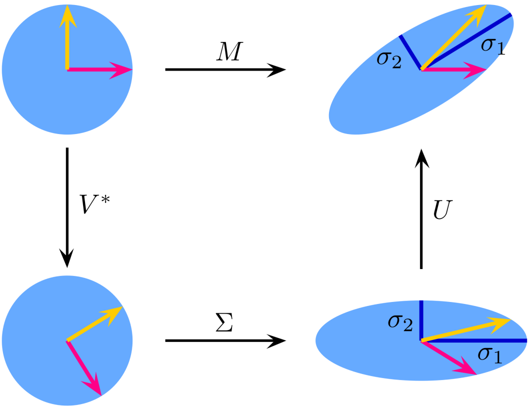
The SVD

Theorem: Singular Value Decomposition

A $m \times n$ matrix with rank r and non-zero singular values $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r$ has a decomposition $U\Sigma V^T$ where

$$\Sigma = \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix}_{m\times n} = \begin{bmatrix} \sigma_1 & 0 & \dots & 0 & \\ 0 & \sigma_2 & \dots & \vdots & \mathbf{0} \\ \vdots & \vdots & \ddots & & \\ 0 & 0 & \dots & \sigma_r & \\ & \mathbf{0} & & & \mathbf{0} \end{bmatrix}$$

U is a $m \times m$ orthogonal matrix, and V is a $n \times n$ orthogonal matrix.



$$M = U \cdot \Sigma \cdot V^*$$

Algorithm to find the SVD of A

Suppose A is $m \times n$ and has rank $r \leq n$.

1. Compute the squared singular values of A^TA, σ_i^2, and construct Σ.

2. Compute the unit singular vectors of A^TA, $\vec{v}_i$, use them to form V.

3. Compute an orthonormal basis for Col$A$ using

$$\vec{u}_i = \frac{1}{\sigma_i} A\vec{v}_i, \quad i = 1, 2, \ldots r$$

   Extend the set $\{\vec{u}_i\}$ to form an orthonomal basis for $\mathbb{R}^m$, use the basis for form U.

Example 2: Write down the singular value decomposition for

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

Example 3: Construct the singular value decomposition of

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

(It has rank 1.)

Applications of the SVD

The SVD has been applied to many modern applications in CS, engineering, and mathematics (our textbook mentions the first four).

- Estimating the rank and condition number of a matrix
- Constructing bases for the four fundamental spaces
- Computing the pseudoinverse of a matrix
- Linear least squares problems
- Non-linear least-squares
  https://en.wikipedia.org/wiki/Non-linear_least_squares
- Machine learning and data mining
  https://en.wikipedia.org/wiki/K-SVD
- Facial recognition
  https://en.wikipedia.org/wiki/Eigenface
- Principle component analysis
  https://en.wikipedia.org/wiki/Principal_component_analysis
- Image compression

Students are expected to be familiar with the 1st two items in the list.

The Condition Number of a Matrix

If A is an invertible $n \times n$ matrix, the ratio

$$\frac{\sigma_1}{\sigma_n}$$

is the **condition number** of A.

Note that:

- The condition number of a matrix describes the sensitivity of a solution to $A\vec{x} = \vec{b}$ is to errors in A.
- We could define the condition number for a rectangular matrix, but that would go beyond the scope of this course.

Example 4

For $A = U\Sigma V^*$, determine the rank of A, and orthonormal bases for Null$A$ and $(\text{Col}A)^\perp$.

$$\mathbf{U} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \end{bmatrix}$$

$$\mathbf{\Sigma} = \begin{bmatrix} 4 & 0 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 \\ 0 & 0 & \sqrt{5} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\mathbf{V}^* = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ \sqrt{0.2} & 0 & 0 & 0 & \sqrt{0.8} \\ 0 & 0 & 0 & 1 & 0 \\ -\sqrt{0.8} & 0 & 0 & 0 & \sqrt{0.2} \end{bmatrix}$$

Example 4 - Solution

The Four Fundamental Spaces

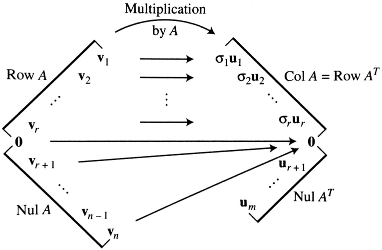


FIGURE 4 The four fundamental subspaces and the action of A.

1. $A\vec{v}_s = \sigma_s\vec{u}_s$.
2. $\vec{v}_1, \ldots, \vec{v}_r$ is an orthonormal basis for Row$A$.
3. $\vec{u}_1, \ldots, \vec{u}_r$ is an orthonormal basis for Col$A$.
4. $\vec{v}_{r+1}, \ldots, \vec{v}_n$ is an orthonormal basis for Null$A$.
5. $\vec{u}_{r+1}, \ldots, \vec{u}_n$ is an orthonormal basis for NullA^T.

The Spectral Decomposition of a Matrix

The SVD can also be used to construct the spectral decomposition for any matrix with rank r

$$A = \sum_{s=1}^{r} \sigma_s \vec{u}_s \vec{v}_s^T,$$

where $\vec{u}_s, \vec{v}_s$ are the s^{th} columns of U and V respectively.

For the case when $A = A^T$, we obtain the same spectral decomposition that we encountered in Section 7.2.