Linear. Algebra.

What is Linear Algebra?

Linear

Algebra

- from al-jebr (Arabic), meaning reunion of broken parts
- 9^{th} century Abu Ja'far Muhammad ibn Muso al-Khwarizmi

Why a whole course?

Engineers need to solve *lots* of equations in *lots* of variables.

$$\begin{aligned}
3x_1 + 4x_2 + 10x_3 + 19x_4 - 2x_5 - 3x_6 &= 141 \\
7x_1 + 2x_2 - 13x_3 - 7x_4 + 21x_5 + 8x_6 &= 2567 \\
-x_1 + 9x_2 + \tfrac{3}{2}x_3 + x_4 + 14x_5 + 27x_6 &= 26 \\
\tfrac{1}{2}x_1 + 4x_2 + 10x_3 + 11x_4 + 2x_5 + x_6 &= -15
\end{aligned}$$

Often, it's enough to know some information about the set of solutions without having to solve the equations at all!

In real life, the difficult part is often in recognizing that a problem can be solved using linear algebra in the first place: need *conceptual* understanding.

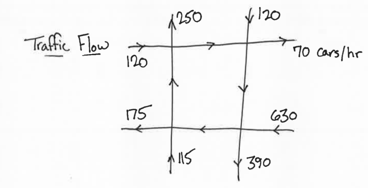
Linear Algebra in Engineering

Almost every engineering problem, no matter how huge, can be reduced to linear algebra:

$$Ax = b \qquad \text{or}$$
$$Ax = \lambda x \qquad \text{or}$$
$$Ax \approx x$$

Applications of Linear Algebra

Civil Engineering: How much traffic lies in the four unlabeled segments?



Applications of Linear Algebra

Chemistry: Balancing reaction equations

$$\underline{\ \ x\ \ }\ C_2H_6 + \underline{\ y\ }\ O_2 \to \underline{\ \ z\ \ }\ CO_2 + \underline{\ \ w\ \ }\ H_2O$$

Applications of Linear Algebra

Biology: In a population of rabbits...

- half of the new born rabbits survive their first year
- of those, half survive their second year
- the maximum life span is three years
- rabbits produce 0, 6, 8 rabbits in their first, second, and third years

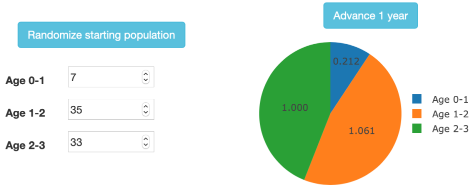
If I know the population in 2016 (in terms of the number of first, second, and third year rabbits), then what is the population in 2017?

Say the numbers of first, second, and third year rabbits in year n are:

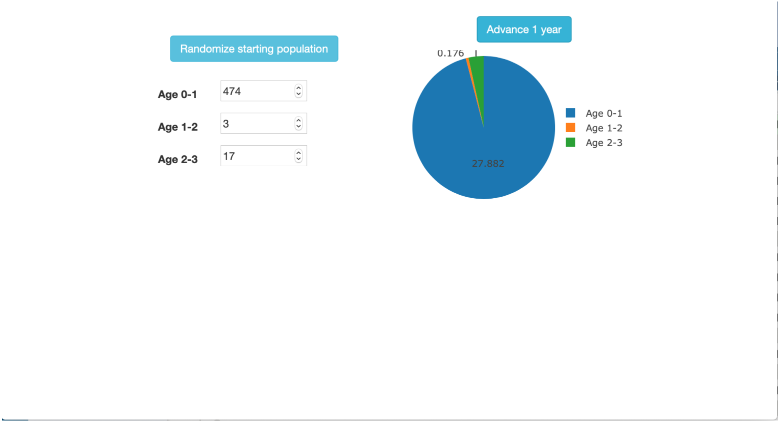
$$F_n, S_n, T_n$$

Demo

Rabbit populations



Rabbit populations



Rabbit populations

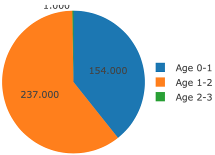
Randomize starting population

Age 0-1 154

Age 1-2 237

Age 2-3 1

Advance 1 year

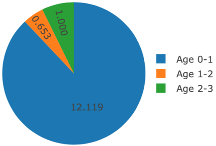


Rabbit populations

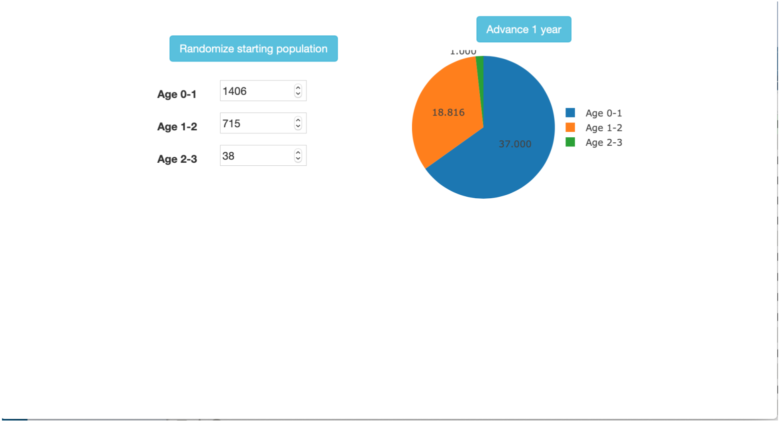
Randomize starting population

Age 0-1	1430
Age 1-2	77
Age 2-3	118

Advance 1 year



Rabbit populations

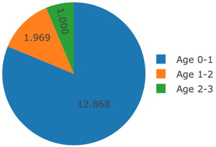


Rabbit populations

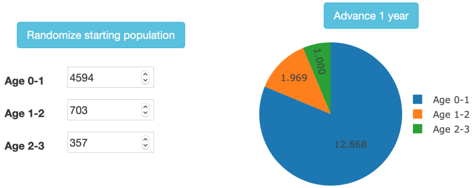
Randomize starting population

Age 0-1	4594
Age 1-2	703
Age 2-3	357

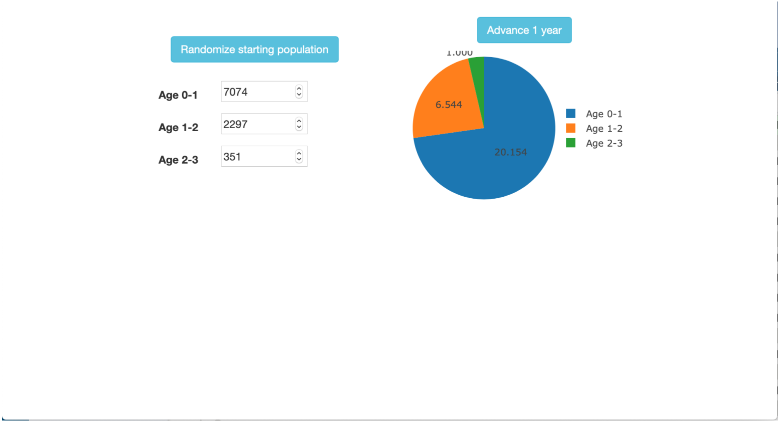
Advance 1 year



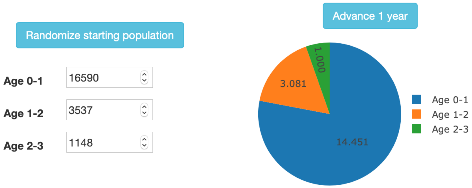
Rabbit populations



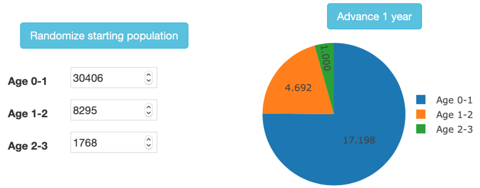
Rabbit populations



Rabbit populations



Rabbit populations

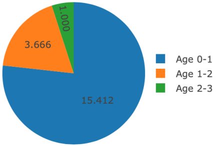


Rabbit populations

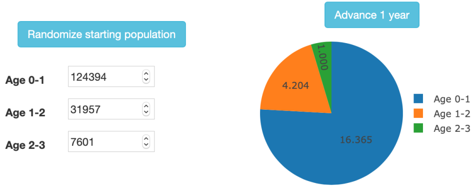
Randomize starting population

Age 0-1	63914
Age 1-2	15203
Age 2-3	4147

Advance 1 year



Rabbit populations

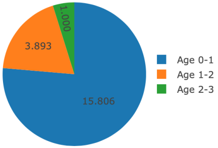


Rabbit populations

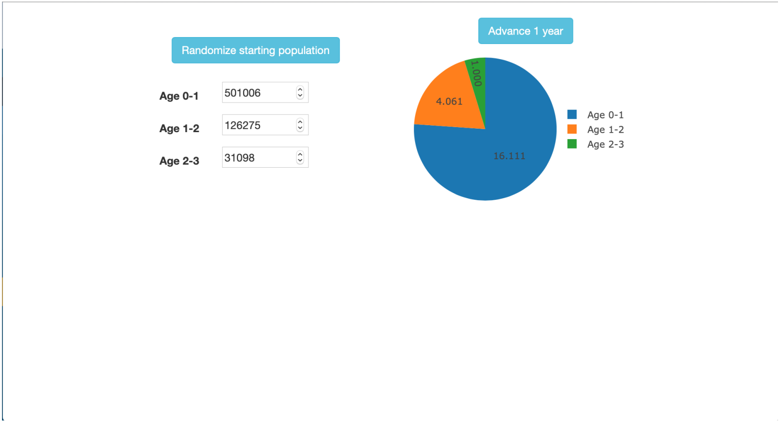
Randomize starting population

Age 0-1	252550
Age 1-2	62197
Age 2-3	15978

Advance 1 year



Rabbit populations

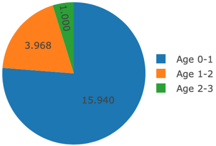


Rabbit populations

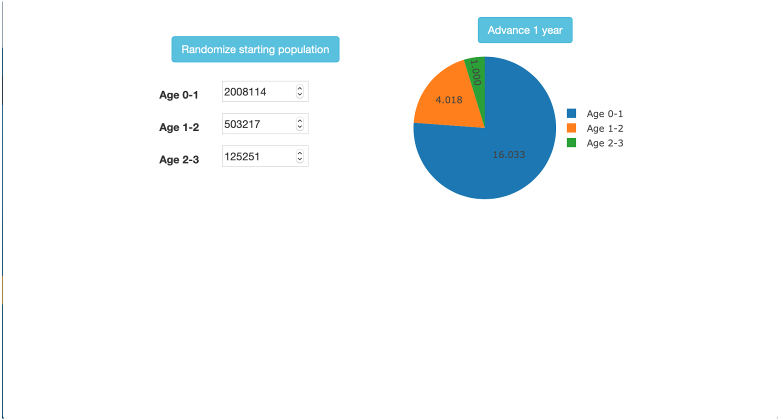
Randomize starting population

Age 0-1	1006434
Age 1-2	250503
Age 2-3	63137

Advance 1 year



Rabbit populations



Demo

Applications of Linear Algebra

Geometry and Astronomy: Find the equation of a circle passing through 3 given points, say (1,0), (0,1), and (1,1). The general form of a circle is $a(x^2 + y^2) + bx + cy + d = 0 \rightsquigarrow$ system of linear equations.

Very similar to: compute the orbit of a planet: $a(x^2 + y^2) + bx + cy + d = 0$

Applications of Linear Algebra

Google: "The 25 billion dollar eigenvector." Each web page has some importance, which it shares via outgoing links to other pages $\rightsquigarrow$ system of linear equations. Stay tuned!

Overview of the course

- Solve the matrix equation $Ax = b$

Overview of the course

- Solve the matrix equation $Ax = b$
  - Solve systems of linear equations using matrices and row reduction, and inverses

Overview of the course

- Solve the matrix equation $Ax = b$
  - Solve systems of linear equations using matrices and row reduction, and inverses
  - Solve systems of linear equations with varying parameters using parametric forms for solutions, the geometry of linear transformations, the characterizations of invertible matrices, and determinants
- Solve the matrix equation $Ax = \lambda x$

Overview of the course

- Solve the matrix equation $Ax = b$
  - Solve systems of linear equations using matrices and row reduction, and inverses
  - Solve systems of linear equations with varying parameters using parametric forms for solutions, the geometry of linear transformations, the characterizations of invertible matrices, and determinants
- Solve the matrix equation $Ax = \lambda x$
  - Solve eigenvalue problems through the use of the characteristic polynomial

Overview of the course

- Solve the matrix equation $Ax = b$
  - Solve systems of linear equations using matrices and row reduction, and inverses
  - Solve systems of linear equations with varying parameters using parametric forms for solutions, the geometry of linear transformations, the characterizations of invertible matrices, and determinants
- Solve the matrix equation $Ax = \lambda x$
  - Solve eigenvalue problems through the use of the characteristic polynomial
  - Understand the dynamics of a linear transformation via the computation of eigenvalues, eigenvectors, and diagonalization
- Almost solve the equation $Ax = b$

Overview of the course

- Solve the matrix equation $Ax = b$
  - Solve systems of linear equations using matrices and row reduction, and inverses
  - Solve systems of linear equations with varying parameters using parametric forms for solutions, the geometry of linear transformations, the characterizations of invertible matrices, and determinants
- Solve the matrix equation $Ax = \lambda x$
  - Solve eigenvalue problems through the use of the characteristic polynomial
  - Understand the dynamics of a linear transformation via the computation of eigenvalues, eigenvectors, and diagonalization
- Almost solve the equation $Ax = b$
  - Find best-fit solutions to systems of linear equations that have no actual solution using least squares approximations

Section 1.1

Solving systems of equations

Outline of Section 1.1

- Learn what it means to solve a system of linear equations
- Describe the solutions as points in $\mathbb{R}^n$
- Learn what it means for a system of linear equations to be inconsistent

Solving equations

Solving equations

What does it mean to solve an equation?

$2x = 10$

$x + y = 1$

$x + y + z = 0$

Find one solution to each. Can you find all of them?

A solution is a *list* of numbers. For example $(3, -4, 1)$.

Solving equations

What does it mean to solve a system of equations?

$$\begin{aligned} x + y &= 2 \\ y &= 1 \end{aligned}$$

What about...

$$\begin{aligned} x + y + z &= 3 \\ x + y - z &= 1 \\ x - y + z &= 1 \end{aligned}$$

Is $(1, 1, 1)$ a solution? Is $(2, 0, 1)$ a solution? What are all the solutions?

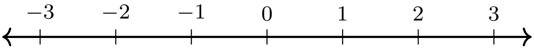
Soon, you will be able to see just by looking that there is exactly one solution.

$$\mathbb{R}^n$$

$\mathbb{R}^n$

$\mathbb{R} =$ denotes the set of all real numbers

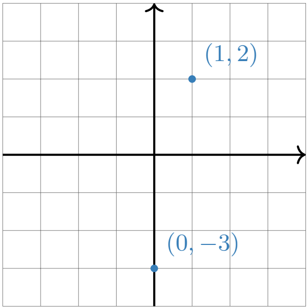
Geometrically, this is the *number line*.



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

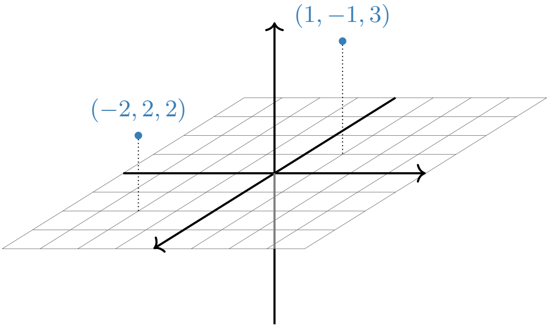
Solutions to systems of equations are exactly points in $\mathbb{R}^n$. In other words, $\mathbb{R}^n$ is where our solutions will live (the n depends on the system of equations).

When $n = 2$, we can visualize of $\mathbb{R}^2$ as the *plane*.



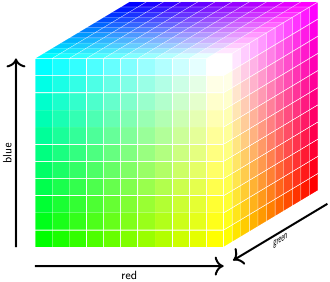
$\mathbb{R}^n$

When $n = 3$, we can visualize $\mathbb{R}^3$ as the *space* we (appear to) live in.



$\mathbb{R}^n$

We can think of the space of all *colors* as (a subset of) $\mathbb{R}^3$:



So what is $\mathbb{R}^4$? or $\mathbb{R}^5$? or $\mathbb{R}^n$?

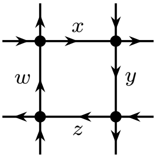
…go back to the *definition*: ordered n-tuples of real numbers

$$(x_1, x_2, x_3, \ldots, x_n).$$

They're still "geometric" spaces, in the sense that our intuition for $\mathbb{R}^2$ and $\mathbb{R}^3$ sometimes extends to $\mathbb{R}^n$, but they're harder to visualize.

Last time we could have used $\mathbb{R}^3$ to describe a rabbit population in a given year: (first year, second year, third year).

Similarly, we could have used $\mathbb{R}^4$ to label the amount of traffic (x, y, z, w) passing through four streets.



We'll make definitions and state theorems that apply to any $\mathbb{R}^n$, but we'll only draw pictures in $\mathbb{R}^2$ and $\mathbb{R}^3$.

$\mathbb{R}^n$ and QR codes

This is a 21×21 QR code. We can also think of this as an element of $\mathbb{R}^n$.

How? Which n?

What about a greyscale image?

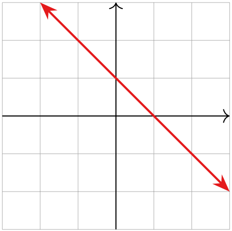
This is a powerful idea: instead of thinking of a QR code as 441 pieces of information, we think of it as one piece of information.

Visualizing solutions: a preview

One Linear Equation

What does the solution set of a linear equation look like?

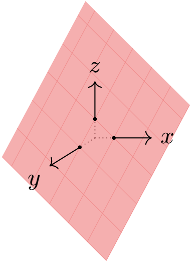
$x + y = 1 \rightsquigarrow$ a line in the plane: $y = 1 - x$



One Linear Equation

What does the solution set of a linear equation look like?

$x + y + z = 1 \rightsquigarrow$ a plane in space:



One Linear Equation

Continued

What does the solution set of a linear equation look like?

$x + y + z + w = 1 \rightsquigarrow$ a "3-plane" in "4-space"...

Systems of Linear Equations

What does the solution set of a *system* of more than one linear equation look like?

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

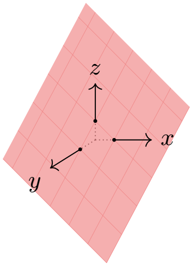
What are the other possibilities for two equations with two variables?

What if there are more variables? More equations?

Poll

Is the plane in $\mathbb{R}^3$ from the previous example equal to $\mathbb{R}^2$? What about the xy-plane in $\mathbb{R}^3$?

1. yes + yes
2. yes + no
3. no + yes
4. no + no



Consistent versus Inconsistent

We say that a system of linear equations is consistent if it has a solution and inconsistent otherwise.

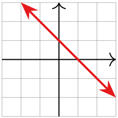
$$\begin{aligned} x + y &= 1 \\ x + y &= 2 \end{aligned}$$

Why is this inconsistent?

What are other examples of inconsistent systems of linear equations?

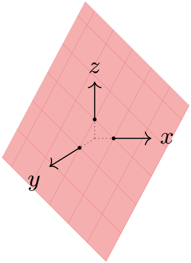
Parametric form

The equation $y = 1 - x$ is an implicit equation for the line in the picture.



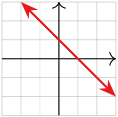
It also has a parametric form: $(x, 1 - x)$

Similarly the equation $x + y + z = 1$ is an implicit equation. One parametric form is: $(x, y, 1 - x - y)$.



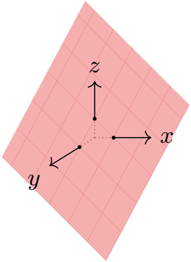
Parametric form

The equation $y = 1 - x$ is an implicit equation for the line in the picture.



It also has a parametric form: $(x, 1 - x)$

Similarly the equation $x + y + z = 1$ is an implicit equation. One parametric form is: $(x, y, 1 - x - y)$.



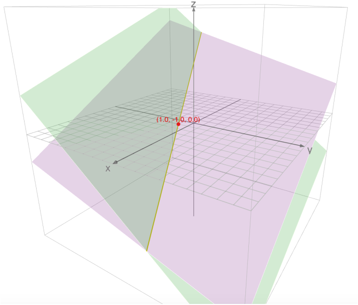
What is an implicit equation and a parametric form for the xy-plane in $\mathbb{R}^3$?

Parametric form

The system of equations

$$\begin{aligned} 2x + y + 12z &= 1 \\ x + 2y + 9z &= -1 \end{aligned}$$

is the implicit form for the line of intersection in the picture.



The line of intersection also has a parametric form: $(1 - 5z, -1 - 2z, z)$

We think of the former as being the problem and the latter as being the explicit solution. One of our first tasks this semester is to learn how to go from the implicit form to the parametric form.

Summary of Section 1.1

- A solution to a system of linear equations in n variables is a point in $\mathbb{R}^n$.
- The set of all solutions to a single equation in n variables is an $(n-1)$-dimensional plane in $\mathbb{R}^n$
- The set of solutions to a system of m linear equations in n variables is the intersection of m of these $(n-1)$-dimensional planes in $\mathbb{R}^n$.
- A system of equations with no solutions is said to be inconsistent.
- Line and planes have implicit equations and parametric forms.

Section 1.2

Row reduction

Outline of Section 1.2

- Solve systems of linear equations via elimination
- Solve systems of linear equations via matrices and row reduction
- Learn about row echelon form and reduced row echelon form of a matrix
- Learn the algorithm for finding the (reduced) row echelon form of a matrix
- Determine from the row echelon form of a matrix if the corresponding system of linear equations is consistent or not.

Solving systems of linear equations by elimination

Example

Solve:

$$\begin{aligned} -y + 8z &= 10 \\ 5y + 10z &= 0 \end{aligned}$$

How many ways can you do it?

Example

Solve:

$$\begin{aligned} -x + y + 3z &= -2 \\ 2x - 3y + 2z &= 14 \\ 3x + 2y + z &= 6 \end{aligned}$$

Hint: Eliminate x!

Solving systems of linear equations with matrices

Example

Solve:

$$\begin{aligned} -y + 8z &= 10 \\ 5y + 10z &= 0 \end{aligned}$$

It is redundant to write x and y again and again, so we rewrite using (augmented) *matrices*. In other words, just keep track of the coefficients, drop the $+$ and $=$ signs. We put a vertical line where the equals sign is.

$$\left(\begin{array}{rr|r} -1 & 8 & 10 \\ 5 & 10 & 0 \end{array} \right) \rightsquigarrow$$

Example

Solve:

$$
\begin{aligned}
-x + y + 3z &= -2 \\
2x - 3y + 2z &= 14 \\
3x + 2y + z &= 6
\end{aligned}
$$

Again we rewrite using augmented matrices...

Row operations

Our manipulations of matrices are called row operations:

row swap, row scale, row replacement

If two matrices differ by a sequence of these three row operations, we say they are row equivalent.

Goal: Produce a system of equations like:

$$\begin{aligned} x \qquad\qquad &= 2 \\ y \qquad &= 1 \\ z &= 5 \end{aligned}$$

What does this look like in matrix form?

Row operations

Why do row operations not change the solution?

Solve:

$$\begin{aligned} x + y &= 2 \\ -2x + y &= -1 \end{aligned}$$

System has one solution, $x = 1, y = 1$.

What happens to the two lines as you do row operations?

$$\left(\begin{array}{cc|c} 1 & 1 & 2 \\ -2 & 1 & -1 \end{array}\right) \rightsquigarrow$$

They pivot around the solution!

Row Reduction and Echelon Forms

Row echelon form

Remember our goal.

Goal: Produce a system of equations like

$$\begin{aligned} x \qquad\qquad &= 2 \\ y \qquad &= 1 \\ z &= 5 \end{aligned}$$

Or at least...

Easier goal: Produce a system of equations like

$$\begin{aligned} x + 5y - 3z &= 2 \\ y + 7z &= 1 \\ z &= 5 \end{aligned}$$

Row Reduction and Echelon Forms

A matrix is in row echelon form if

1. all zero rows are at the bottom, and
2. each leading (nonzero) entry of a row is to the right of the leading entry of the row above.

$$\begin{pmatrix} \boxed{\star} & \star & \star & \star & \star \\ 0 & \boxed{\star} & \star & \star & \star \\ 0 & 0 & 0 & \boxed{\star} & \star \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

This system is easy to solve using back substitution.

The pivot positions are the leading entries in each row.

Reduced Row Echelon Form

A system is in reduced row echelon form if also:

4. the leading entry in each nonzero row is 1
5. each leading entry of a row is the only nonzero entry in its column

For example:

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

This system is even easier to solve.

Important. In any discussion of row echelon form, we ignore any vertical lines!

Can every matrix be put in reduced row echelon form?

Reduced Row Echelon Form

Poll

Which are in reduced row echelon form?

$$\left(\begin{array}{c|c} 1 & 0 \\ 0 & 2 \end{array}\right) \qquad \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$\begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \quad \begin{pmatrix} 0 & 1 & 0 & 0 \end{pmatrix} \quad \begin{pmatrix} 0 & 1 & 8 & 0 \end{pmatrix}$$

$$\left(\begin{array}{cc|c} 1 & 17 & 0 \\ 0 & 0 & 1 \end{array}\right) \qquad \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

REF:

1. all zero rows are at the bottom, and
2. each leading (nonzero) entry of a row is to the right of the leading entry of the row above.

RREF:

4. the leading entry in each nonzero row is 1
5. each leading entry of a row is the only nonzero entry in its column

Row Reduction

Theorem. Each matrix is row equivalent to one and only one matrix in reduced row echelon form.

We'll give an algorithm. That shows a matrix is equivalent to at least one matrix in reduced row echelon form.

Row Reduction Algorithm

To find row echelon form:

Step 1 Swap rows so a leftmost nonzero entry is in 1st row (if needed)

Step 2 Scale 1st row so that its leading entry is equal to 1

Step 3 Use row replacement so all entries below this 1 (or, pivot) are 0

Then cover the first row and repeat the three steps.

To then find reduced row echelon form:

- Use row replacement so that all entries above the pivots are 0.

Examples.

$$\left(\begin{array}{rrr|r} 1 & 2 & 3 & 9 \\ 2 & -1 & 1 & 8 \\ 3 & 0 & -1 & 3 \end{array}\right) \left(\begin{array}{rrr|r} 0 & 7 & -4 & 2 \\ 2 & 4 & 6 & 12 \\ 3 & 1 & -1 & -2 \end{array}\right) \left(\begin{array}{rrr|r} 4 & -5 & 3 & 2 \\ 1 & -1 & -2 & -6 \\ 4 & -4 & -14 & 18 \end{array}\right)$$

Interactive Row Reducer

Solutions of Linear Systems

We want to go from reduced row echelon forms to solutions of linear systems.

Solve the linear system associated to:

$$\left(\begin{array}{cc|c} 1 & 0 & 5 \\ 0 & 1 & 2 \end{array}\right)$$

What are the solutions? Say the variables are x and y.

Solutions of Linear Systems: Consistency

Solve the linear system associated to:

$$\left(\begin{array}{ccc|c} 1 & 0 & 5 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right)$$

Say the variables are x, y, and z.

A system of equations is inconsistent exactly when the corresponding augmented matrix has a pivot in the last column.

Example with a parameter

For which values of h does the following system have a solution?

$$\begin{aligned} x + y &= 1 \\ 2x + 2y &= h \end{aligned}$$

Solve this by row reduction and also solve it by thinking geometrically.

Summary of Section 1.2

- To solve a system of linear equations we can use the method of elimination.
- We can more easily do elimination with matrices. The allowable moves are row swaps, row scales, and row replacements. This is called row reduction.
- A matrix in row echelon form corresponds to a system of linear equations that we can easily solve by back substitution.
- A matrix in reduced row echelon form corresponds to a system of linear equations that we can easily solve just by looking.
- We have an algorithm for row reducing a matrix to row echelon form.
- The reduced row echelon form of a matrix is unique.
- Two matrices that differ by row operations are called row equivalent.
- A system of equations is inconsistent exactly when the corresponding augmented matrix has a pivot in the last column.

1.3 Parametric Form

Outline of Section 1.3

- Find the parametric form for the solutions to a system of linear equations.
- Describe the geometric picture of the set of solutions.

Free Variables

We know how to understand the solution to a system of linear equations when every column to the left of the vertical line has a pivot. For instance:

$$\left(\begin{array}{cc|c} 1 & 0 & 5 \\ 0 & 1 & 2 \end{array}\right)$$

If the variables are x and y what are the solutions?

Free Variables

How do we solve a system of linear equations if the row reduced matrix has a column without a pivot? For instance:

$$\left(\begin{array}{ccc|c} 1 & 0 & 5 & 0 \\ 0 & 1 & 2 & 1 \end{array}\right)$$

represents two equations:

$$\begin{aligned} x_1 \qquad + 5x_3 &= 0 \\ x_2 + 2x_3 &= 1 \end{aligned}$$

There is one free variable x_3, corresponding to the non-pivot column.
To solve, we move the free variable to the right:

$$\begin{aligned} x_1 \qquad\qquad &= -5x_3 \\ x_2 \qquad &= 1 - 2x_3 \\ x_3 &= x_3 \text{ (free; any real number)} \end{aligned}$$

This is the parametric solution. We can also write the solution as:

$$(-5x_3, 1 - 2x_3, x_3)$$

What is one particular solution? What does the set of solutions look like?

Free Variables

Solve the system of linear equations in x_1, x_2, x_3, x_4:

$$\begin{aligned} x_1 \quad + 5x_3 \quad &= 0 \\ x_4 &= 0 \end{aligned}$$

So the associated matrix is:

$$\left(\begin{array}{cccc|c} 1 & 0 & 5 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{array}\right)$$

To solve, we move the free variable to the right:

$$\begin{aligned} x_1 &= -5x_3 \\ x_2 &= \quad x_2 \quad \text{(free)} \\ x_3 &= \quad x_3 \quad \text{(free)} \\ x_4 &= 0 \end{aligned}$$

Or: $(-5x_3, x_2, x_3, 0)$. This is a plane in $\mathbb{R}^4$.

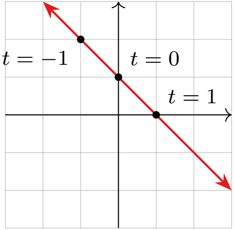
The original equations are the implicit equations for the solution. The answer to this question is the parametric solution.

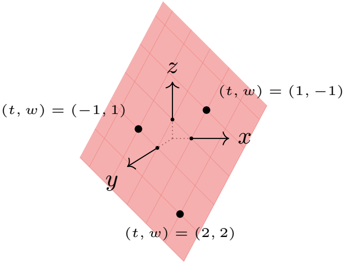
Free variables

Geometry

If we have a consistent system of linear equations, with n variables and k free variables, then the set of solutions is a k-dimensional plane in $\mathbb{R}^n$.

Why does this make sense?





Poll

A linear system has 4 variables and 3 equations. What are the possible solution sets?

1. nothing
2. point
3. two points
4. line
5. plane
6. 3-dimensional plane
7. 4-dimensional plane

Implicit versus parametric equations of planes

Find a parametric description of the plane

$$x + y + z = 1$$

The original version is the implicit equation for the plane. The answer to this problem is the parametric description.

Summary

There are *three possibilities* for the reduced row echelon form of the augmented matrix of system of linear equations.

1. The last column is a pivot column.
   $\rightsquigarrow$ the system is *inconsistent*.

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

2. Every column except the last column is a pivot column.
   $\rightsquigarrow$ the system has a *unique solution*.

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

3. The last column is not a pivot column, and some other column isn't either.
   $\rightsquigarrow$ the system has *infinitely many* solutions; free variables correspond to columns without pivots.

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

Typical exam questions

True/False: If a system of equations has 100 variables and 70 equations, then there must be infinitely many solutions.

True/False: If a system of equations has 70 variables and 100 equations, then it must be inconsistent.

How can we tell if an augmented matrix corresponds to a consistent system of linear equations?

If a system of linear equations has finitely many solutions, what are the possible numbers of solutions?

Chapter 2

System of Linear Equations: Geometry

Where are we?

In Chapter 1 we learned to solve any system of linear equations in any number of variables. The answer is row reduction, which gives an algebraic solution. In Chapter 2 we put some geometry behind the algebra. It is the geometry that gives us intuition and deeper meaning. There are three main points:

Sec 2.3: $Ax = b$ is consistent $\Leftrightarrow$ b is in the span of the columns of A.

Sec 2.4: The solutions to $Ax = b$ are parallel to the solutions to $Ax = 0$.

Sec 2.9: The dim's of $\{b : Ax = b$ is consistent$\}$ and $\{$solutions to $Ax = b\}$ add up to the number of columns of A.

Section 2.1

Vectors

Outline

- Think of points in $\mathbb{R}^n$ as vectors.
- Learn how to add vectors and multiply them by a scalar
- Understand the geometry of adding vectors and multiplying them by a scalar
- Understand linear combinations algebraically and geometrically

Vectors

A vector is a matrix with one row or one column. We can think of a vector with n rows as:

- a point in $\mathbb{R}^n$
- an arrow in $\mathbb{R}^n$

To go from an arrow to a point in $\mathbb{R}^n$, we subtract the tip of the arrow from the starting point. Note that there are many arrows representing the same vector.

Adding vectors / parallelogram rule ▸ Demo

Scaling vectors ▸ Demo

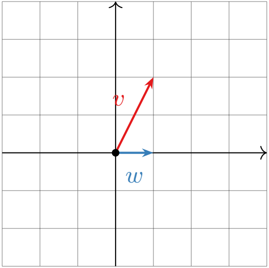
A scalar is just a real number. We use this term to indicate that we are scaling a vector by this number.

Linear Combinations

A linear combination of the vectors $v_1, \ldots, v_k$ is any vector

$$c_1 v_1 + c_2 v_2 + \cdots + c_k v_k$$

where $c_1, \ldots, c_k$ are real numbers.



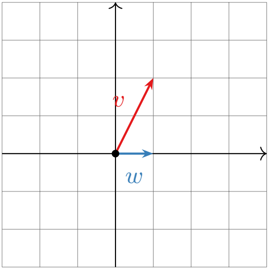
Let $v = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ and $w = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$.

What are some linear combinations of v and w?

Poll

Is there a vector in $\mathbb{R}^2$ that is not a linear combination of v and w?

- yes
- no



Linear Combinations

What are some linear combinations of $(1, 1)$?

What are some linear combinations of $(1, 1)$ and $(2, 2)$?

What are some linear combinations of $(0, 0)$?

Summary of Section 2.1

- A vector is a point/arrow in $\mathbb{R}^n$
- We can add/scale vectors algebraically & geometrically (parallelogram rule)
- A linear combination of vectors $v_1, \ldots, v_k$ is a vector

$$c_1 v_1 + \cdots + c_k v_k$$

where $c_1, \ldots, c_k$ are real numbers.

Typical exam questions

True/False: For any collection of vectors $v_1, \dots, v_k$ in $\mathbb{R}^n$, the zero vector in $\mathbb{R}^n$ is a linear combination of $v_1, \dots, v_k$.

True/False: The vector $(1, 1)$ can be written as a linear combination of $(2, 2)$ and $(-2, -2)$ in infinitely many ways.

Suppose that v is a vector in $\mathbb{R}^n$, and consider the set of all linear combinations of v. What geometric shape is this?

Section 2.2

Vector Equations and Spans

Outline of Section 2.2

- Learn the equivalences:

$$\text{vector equations} \leftrightarrow \text{augmented matrices} \leftrightarrow \text{linear systems}$$

- Learn the definition of span
- Learn the relationship between spans and consistency

Linear Combinations

Is $\begin{pmatrix} 8 \\ 16 \\ 3 \end{pmatrix}$ a linear combination of $\begin{pmatrix} 1 \\ 2 \\ 6 \end{pmatrix}$ and $\begin{pmatrix} -1 \\ -2 \\ -1 \end{pmatrix}$?

Write down an equation in order to solve this problem. This is called a vector equation.

Notice that the vector equation can be rewritten as a system of linear equations. Solve it!

Linear combinations, vector equations, and linear systems

In general, asking:

Is b a linear combination of $v_1, \ldots, v_k$?

is the same as asking if the vector equation

$$x_1 v_1 + \cdots + x_k v_k = b$$

is consistent, which is the same as asking if the system of linear equations corresponding to the augmented matrix

$$\left(\begin{array}{cccc|c} | & | & & | & | \\ v_1 & v_2 & \cdots & v_k & b \\ | & | & & | & | \end{array}\right),$$

is consistent.

Compare with the previous slide! Make sure you are comfortable going back and forth between the specific case (last slide) and the general case (this slide).

Span

Essential vocabulary word!

$$\begin{aligned}\text{Span}\{v_1, v_2, \ldots, v_k\} &= \{x_1v_1 + x_2v_2 + \cdots + x_kv_k \mid x_i \text{ in } \mathbb{R}\} \leftarrow \text{(set builder notation)} \\ &= \text{the set of all linear combinations of vectors } v_1, v_2, \ldots, v_k \\ &= \text{plane through the origin and } v_1, v_2, \ldots, v_k.\end{aligned}$$

What are the possibilities for the span of two vectors in $\mathbb{R}^2$?

Demo

What are the possibilities for the span of three vectors in $\mathbb{R}^3$?

Demo

Conclusion: Spans are planes (of some dimension) through the origin, and the dimension of the plane is at most the number of vectors you started with and is at most the dimension of the space they're in.

Span

Essential vocabulary word!

$$\begin{aligned}\mathrm{Span}\{v_1, v_2, \ldots, v_k\} &= \{x_1 v_1 + x_2 v_2 + \cdots x_k v_k \mid x_i \text{ in } \mathbb{R}\} \\ &= \text{the set of all linear combinations of vectors } v_1, v_2, \ldots, v_k \\ &= \text{plane through the origin and } v_1, v_2, \ldots, v_k.\end{aligned}$$

Four ways of saying the same thing:

- b is in Span$\{v_1, v_2, \ldots, v_k\}$ $\leftarrow$ geometry
- b is a linear combination of $v_1, \ldots, v_k$
- the vector equation $x_1 v_1 + \cdots + x_k v_k = b$ has a solution $\leftarrow$ algebra
- the system of linear equations corresponding to

$$\left(\begin{array}{cccc|c} | & | & & | & | \\ v_1 & v_2 & \cdots & v_k & b \\ | & | & & | & | \end{array}\right),$$

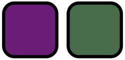
is consistent.

Demo Demo

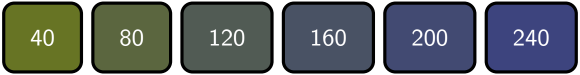
Application: Additive Color Theory

Consider now the two colors

$$\begin{pmatrix} 180 \\ 50 \\ 200 \end{pmatrix}, \begin{pmatrix} 100 \\ 150 \\ 100 \end{pmatrix}$$



For which h is $(116, 130, h)$ in the span of those two colors?



Summary of Section 2.2

- vector equations $\leftrightarrow$ augmented matrices $\leftrightarrow$ linear systems
- Checking if a linear system is consistent is the same as asking if the column vector on the end of an augmented matrix is in the span of the other column vectors.
- Spans are planes, and the dimension of the plane is at most the number of vectors you started with.

Typical exam questions

Is $\begin{pmatrix} 8 \\ 16 \\ 1 \end{pmatrix}$ in the span of $\begin{pmatrix} 1 \\ 2 \\ 6 \end{pmatrix}$ and $\begin{pmatrix} -1 \\ -2 \\ -1 \end{pmatrix}$?

Write down the vector equation for the previous problem.

True/False: The vector equation $x_1 v_1 + \cdots + x_k v_k = 0$ is always consistent.

True/False: It is possible for the span of 3 vectors in $\mathbb{R}^3$ to be a line.

True/False: the plane $z = 1$ in $\mathbb{R}^3$ is a span.

Section 2.3

Matrix equations

Outline Section 2.3

- Understand the equivalences:

  linear system $\leftrightarrow$ augmented matrix $\leftrightarrow$ vector equation $\leftrightarrow$ matrix equation

- Understand the equivalence:

  $$Ax = b \text{ is consistent} \longleftrightarrow b \text{ is in the span of the columns of } A$$

  (also: what does this mean geometrically)

- Learn for which A the equation $Ax = b$ is always consistent
- Learn to multiply a vector by a matrix

Multiplying Matrices

$$\text{matrix} \times \text{column} : \begin{pmatrix} | & | & & | \\ x_1 & x_2 & \cdots & x_n \\ | & | & & | \end{pmatrix} \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = \begin{pmatrix} | & | & & | \\ b_1 x_1 & b_2 x_2 & \cdots & b_n x_n \\ | & | & & | \end{pmatrix}$$

Read this as: b_1 times the first column x_1 is the first column of the answer, b_2 times x_2 is the second column of the answer...

Example:

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

Multiplying Matrices

Another way to multiply

$$\text{row vector} \times \text{column vector} : \begin{pmatrix} a_1 & \cdots & a_{\color{red}n} \end{pmatrix} \begin{pmatrix} b_1 \\ \vdots \\ b_{\color{red}n} \end{pmatrix} = a_1 b_1 + \cdots + a_n b_n$$

$$\text{matrix} \times \text{column vector} : \begin{pmatrix} r_1 \\ \vdots \\ r_m \end{pmatrix} b = \begin{pmatrix} r_1 b \\ \vdots \\ r_m b \end{pmatrix}$$

Example:

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

Linear Systems vs Augmented Matrices vs Matrix Equations vs Vector Equations

A matrix equation is an equation $Ax = b$ where A is a matrix and b is a vector. So x is a vector of variables.

A is an $m \times n$ matrix if it has m rows and n columns.
What sizes must x and b be?

Example:

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 9 \\ 10 \\ 11 \end{pmatrix}$$

Rewrite this equation as a vector equation, a system of linear equations, and an augmented matrix.

We will go back and forth between these four points of view over and over again. You need to get comfortable with this.

Solving matrix equations

Solve the matrix equation

$$\begin{pmatrix} 0 & 6 & 8 \\ 1/2 & 0 & 0 \\ 0 & 1/2 & 0 \end{pmatrix} \begin{pmatrix} f \\ s \\ t \end{pmatrix} = \begin{pmatrix} 20 \\ 1 \\ 1 \end{pmatrix}$$

What does this mean about rabbits?

Solutions to Linear Systems vs Spans

Say that

$$A = \begin{pmatrix} | & | & & | \\ v_1 & v_2 & \cdots & v_n \\ | & | & & | \end{pmatrix}.$$

Fact. $Ax = b$ has a solution $\iff$ b is in the span of columns of A
algebra $\iff$ geometry

Why?

Again this is a basic fact we will use over and over and over.

Solutions to Linear Systems vs Spans

Fact. $Ax = b$ has a solution $\iff$ b is in the span of columns of A

Examples:

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \\ 5 \end{pmatrix} \qquad \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \\ 0 \end{pmatrix}$$

Is a given vector in the span?

Fact. $Ax = b$ has a solution $\iff$ b is in the span of columns of A

algebra $\iff$ geometry

Is $(9, 10, 11)$ in the span of $(1, 3, 5)$ and $(2, 4, 6)$?

Is a given vector in the span?

Poll

Which of the following true statements can you verify without row reduction?

1. $(0, 1, 2)$ is in the span of $(3, 3, 4)$, $(0, 10, 20)$, $(0, -1, -2)$
2. $(0, 1, 2)$ is in the span of $(3, 3, 4)$, $(0, 1, 0)$, $(0, 0, \sqrt{2})$
3. $(0, 1, 2)$ is in the span of $(3, 3, 4)$, $(0, 5, 7)$, $(0, 6, 8)$
4. $(0, 1, 2)$ is in the span of $(5, 7, 0)$, $(6, 8, 0)$, $(3, 3, 4)$

Pivots vs Solutions

Theorem. Let A be an $m \times n$ matrix. The following are equivalent.

1. $Ax = b$ has a solution for all b
2. The span of the columns of A is $\mathbb{R}^m$
3. A has a pivot in each row

Why?

More generally, if you have some vectors and you want to know the dimension of the span, you should row reduce and count the number of pivots.

Properties of the Matrix Product Ax

$c =$ real number, $u, v =$ vectors,

- $A(u + v) = Au + Av$
- $A(cv) = cAv$

Application. If u and v are solutions to $Ax = 0$ then so is every element of $\mathrm{Span}\{u, v\}$.

Guiding questions

Here are the guiding questions for the rest of the chapter:

1. What are the solutions to $Ax = 0$?
2. For which b is $Ax = b$ consistent?

These are two separate questions!

Summary of Section 2.3

- Two ways to multiply a matrix times a column vector:

$$\begin{pmatrix} r_1 \\ \vdots \\ r_m \end{pmatrix} b = \begin{pmatrix} r_1 b \\ \vdots \\ r_m b \end{pmatrix}$$

OR

$$\begin{pmatrix} | & | & & | \\ x_1 & x_2 & \cdots & x_n \\ | & | & & | \end{pmatrix} \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = \begin{pmatrix} | & & | \\ b_1 x_1 & \cdots & b_n x_n \\ | & & | \end{pmatrix}$$

- Linear systems, augmented matrices, vector equations, and matrix equations are all equivalent.
- Fact. $Ax = b$ has a solution $\Leftrightarrow$ b is in the span of columns of A
- Theorem. Let A be an $m \times n$ matrix. The following are equivalent.
  1. $Ax = b$ has a solution for all b
  2. The span of the columns of A is $\mathbb{R}^m$
  3. A has a pivot in each row

Typical exam questions

- If A is a 3×5 matrix, and the product Ax makes sense, then which $\mathbb{R}^n$ does x lie in?

- Rewrite the following linear system as a matrix equation and a vector equation:

$$x + y + z = 1$$

- Multiply:

$$\begin{pmatrix} 0 & 2 \\ 0 & 4 \\ 5 & 0 \end{pmatrix} \begin{pmatrix} 3 \\ 2 \end{pmatrix}$$

- Which of the following matrix equations are consistent?

$$\begin{pmatrix} 1 & 1 \\ 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \\ 2 \end{pmatrix} \qquad \begin{pmatrix} 1 & 1 \\ 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \\ 3 \end{pmatrix}$$

(And can you do it without row reducing?)

Section 2.4

Solution Sets

Outline

- Understand the geometric relationship between the solutions to $Ax = b$ and $Ax = 0$
- Understand the relationship between solutions to $Ax = b$ and spans
- Learn the parametric vector form for solutions to $Ax = b$

Homogeneous systems

Solving $Ax = b$ is easiest when $b = 0$. Such equations are called homogeneous.

Homogenous systems are always consistent. *Why?*

When does $Ax = 0$ have a nonzero/nontrivial solution?

If there are k-free variables and n total variables, then the solution is a k-dimensional plane through the origin in $\mathbb{R}^n$. In particular it is a span.

Parametric Vector Forms for Solutions

Homogeneous case

Solve the matrix equation $Ax = 0$ where

$$A = \begin{pmatrix} 1 & 2 & 0 & -1 \\ -2 & -3 & 4 & 5 \\ 2 & 4 & 0 & -2 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & -8 & -7 \\ 0 & 1 & 4 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

We already know the parametric form:

$$\begin{aligned} x_1 &= 8x_3 + 7x_4 \\ x_2 &= -4x_3 - 3x_4 \\ x_3 &= x_3 \qquad\qquad \text{(free)} \\ x_4 &= x_4 \quad \text{(free)} \end{aligned}$$

We can also write this in parametric vector form:

$$x_3 \begin{pmatrix} 8 \\ -4 \\ 1 \\ 0 \end{pmatrix} + x_4 \begin{pmatrix} 7 \\ -3 \\ 0 \\ 1 \end{pmatrix}$$

Or we can write the solution as a span: $\mathrm{Span}\{(8, -4, 1, 0), (7, -3, 0, 1)\}$.

Parametric Vector Forms for Solutions

Homogeneous case

Find the parametric vector form of the solution to $Ax = 0$ where

$$A = \begin{pmatrix} 1 & 1 & 1 & 1 \end{pmatrix}$$

Variables, equations, and dimension

Poll

For $b \neq 0$, the solutions to $Ax = b$ are...

1. always a span
2. sometimes a span
3. never a span

Nonhomogeneous Systems

Suppose $Ax = b$ and $b \neq 0$.

As before, we can find the parametric vector form for the solution in terms of free variables.

What is the difference?

Parametric Vector Forms for Solutions

Nonhomogeneous case

Find the parametric vector form of the solution to $Ax = b$ where:

$$(A|b) = \left(\begin{array}{cccc|c} 1 & 2 & 0 & -1 & 3 \\ -2 & -3 & 4 & 5 & 2 \\ 2 & 4 & 0 & -2 & 6 \end{array}\right) \rightsquigarrow \left(\begin{array}{cccc|c} 1 & 0 & -8 & -7 & -13 \\ 0 & 1 & 4 & 3 & 8 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

We already know the parametric form:

$$\begin{aligned} x_1 &= -13 + 8x_3 + 7x_4 \\ x_2 &= 8 - 4x_3 - 3x_4 \\ x_3 &= x_3 \quad \text{(free)} \\ x_4 &= x_4 \quad \text{(free)} \end{aligned}$$

We can also write this in parametric vector form:

$$\begin{pmatrix} -13 \\ 8 \\ 0 \\ 0 \end{pmatrix} + x_3 \begin{pmatrix} 8 \\ -4 \\ 1 \\ 0 \end{pmatrix} + x_4 \begin{pmatrix} 7 \\ -3 \\ 0 \\ 1 \end{pmatrix}$$

This is a translate of a span: $(-13, 8, 0, 0) + \text{Span}\{(8, -4, 1, 0), (7, -3, 0, 1)\}$.

Parametric Vector Forms for Solutions

Nonhomogeneous case

Find the parametric vector form for the solution to $Ax = (9)$ where

$$A = \begin{pmatrix} 1 & 1 & 1 & 1 \end{pmatrix}$$

$$\left(\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 9 \end{array}\right)$$

Homogeneous vs. Nonhomogeneous Systems

Key realization. Set of solutions to $Ax = b$ obtained by taking one solution and adding all possible solutions to $Ax = 0$.

$$Ax = 0 \text{ solutions} \rightsquigarrow Ax = b \text{ solutions}$$

$$x_k v_k + \cdots + x_n v_n \rightsquigarrow p + x_k v_k + \cdots + x_n v_n$$

So: set of solutions to $Ax = b$ is parallel to the set of solutions to $Ax = 0$. It is a translate of a plane through the origin. (Again, we are using geometry to understand algebra!)

So by understanding $Ax = 0$ we gain understanding of $Ax = b$ for all b. This gives structure to the set of equations $Ax = b$ for all b.

Demo

Demo

Parametric Vector Forms for Solutions

Nonhomogeneous case

Find the parametric vector forms for $\begin{pmatrix} 1 & -3 \\ 2 & -6 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

...and $\begin{pmatrix} 1 & -3 \\ 2 & -6 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 3 \\ 6 \end{pmatrix}$.

Solving matrix equations

The matrix equation

$$\begin{pmatrix} 0 & 6 & 8 \\ 1/2 & 0 & 0 \\ 0 & 1/2 & 0 \end{pmatrix} \begin{pmatrix} f \\ s \\ t \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

has only the trivial solution.

What does this mean about the matrix equation

$$\begin{pmatrix} 0 & 6 & 8 \\ 1/2 & 0 & 0 \\ 0 & 1/2 & 0 \end{pmatrix} \begin{pmatrix} f \\ s \\ t \end{pmatrix} = \begin{pmatrix} 20 \\ 1 \\ 1 \end{pmatrix}?$$

What does this mean about rabbits?

Two different things

Suppose A is an $m \times n$ matrix. Notice that if $Ax = b$ is a matrix equation then x is in $\mathbb{R}^n$ and b is in $\mathbb{R}^m$. There are two different problems to solve.

1. If we are given a specific b, then we can solve $Ax = b$. This means we find all x in $\mathbb{R}^n$ so that $Ax = b$. We do this by row reducing, taking free variables for the columns without pivots, and writing the (parametric) vector form for the solution.

2. We can also ask for which b in $\mathbb{R}^m$ does $Ax = b$ have a solution? The answer is: when b is in the span of the columns of A. So the answer is "all b in $\mathbb{R}^m$" exactly when the span of the columns is $\mathbb{R}^m$ which is exactly when A has m pivots.

If you go back to the ▸ Demo from earlier in this section, the first question is happening on the left and the second question on the right.

Example. Say that $A = \begin{pmatrix} 1 & -3 \\ 2 & -6 \end{pmatrix}$. We can ask: (1) Does $Ax = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ have a solution? and (2) For which b does $Ax = b$ have a solution?

Summary of Section 2.4

- The solutions to $Ax = 0$ form a plane through the origin (span)
- The solutions to $Ax = b$ form a plane not through the origin
- The set of solutions to $Ax = b$ is parallel to the one for $Ax = 0$
- In either case we can write the parametric vector form. The parametric vector form for the solution to $Ax = 0$ is obtained from the one for $Ax = b$ by deleting the constant vector. And conversely the parametric vector form for $Ax = b$ is obtained from the one for $Ax = 0$ by adding a constant vector. This vector translates the solution set.

Typical exam questions

- Suppose that the set of solutions to $Ax = b$ is the plane $z = 1$ in $\mathbb{R}^3$. What is the set of solutions to $Ax = 0$?
- Suppose that the set of solutions to $Ax = 0$ is the line $y = x$ in $\mathbb{R}^2$. Is it possible that there is a b so that the set of solutions to $Ax = b$ is the line $x + y = 1$?
- Suppose that the set of solutions to $Ax = b$ is the plane $x + y = 1$ in $\mathbb{R}^3$. Is is possible that there is a b so that the set of solutions to $Ax = b$ is the z-axis?
- Suppose that the set of solutions to $Ax = 0$ is the plane $x + 2y - 3z = 0$ in $\mathbb{R}^3$ and that the vector $(1, 3, 5)$ is a solution to $Ax = b$. Find one other solution to $Ax = b$. Find all of them.
- Is there a 2×2 matrix so that the set of solutions to $Ax = \left(\begin{smallmatrix} 1 \\ 2 \end{smallmatrix}\right)$ is the line $y = x + 1$? If so, find such an A. If not, explain why not.

Section 2.5

Linear Independence

Section 2.5 Outline

- Understand what is means for a set of vectors to be linearly independent
- Understand how to check if a set of vectors is linearly independent

Linear Independence

The idea of linear independence: a collection of vectors $v_1, \ldots, v_k$ is linearly independent if they are all pointing in truly different directions. This means that none of the v_i is in the span of the others.

For example, $(1, 0, 0)$, $(0, 1, 0)$ and $(0, 0, 1)$ are linearly independent.

Also, $(1, 0, 0)$, $(0, 1, 0)$ and $(1, 1, 0)$ are linearly dependent.

What is this good for? A basic question we can ask about solving linear equations is: What is the smallest number of vectors needed in the parametric solution to a linear system? We need linear independence to answer this question. See the last slide in this section.

Linear Independence

A set of vectors $\{v_1, \ldots, v_k\}$ in $\mathbb{R}^n$ is linearly independent if the vector equation

$$x_1 v_1 + x_2 v_2 + \cdots + x_k v_k = 0$$

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

So, linearly dependent means there are $x_1, x_2, \ldots, x_k$ not all zero so that

$$x_1 v_1 + x_2 v_2 + \cdots + x_k v_k = 0$$

This is a *linear dependence* relation.

Linear Independence

A set of vectors $\{v_1, \ldots, v_k\}$ in $\mathbb{R}^n$ is linearly independent if the vector equation

$$x_1 v_1 + x_2 v_2 + \cdots + x_k v_k = 0$$

has only the trivial solution.

Fact. The columns of A are linearly independent
$\Leftrightarrow$ $Ax = 0$ has only the trivial solution.
$\Leftrightarrow$ A has a pivot in each column

Why?

Linear Independence

Is $\left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}, \begin{pmatrix} 3 \\ 1 \\ 4 \end{pmatrix} \right\}$ linearly independent?

Is $\left\{ \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}, \begin{pmatrix} 3 \\ 1 \\ 4 \end{pmatrix} \right\}$ linearly independent?

Linear Independence

When is $\{v\}$ is linearly dependent?

When is $\{v_1, v_2\}$ is linearly dependent?

When is the set $\{v_1, v_2, \ldots, v_k\}$ linearly dependent?

Fact. The set $\{v_1, v_2, \ldots, v_k\}$ is linearly independent if and only if they span a k-dimensional plane. (algebra $\leftrightarrow$ geometry)

Fact. The set $\{v_1, v_2, \ldots, v_k\}$ is linearly dependent if and only if we can remove a vector from the set without changing (the dimension of) the span.

Fact. The set $\{v_1, v_2, \ldots, v_k\}$ is linearly dependent if and only if some v_i lies in the span of $v_1, \ldots, v_{i-1}$.

Demo

Span and Linear Independence

Is $\left\{ \begin{pmatrix} 5 \\ 7 \\ 0 \end{pmatrix}, \begin{pmatrix} -5 \\ 7 \\ 0 \end{pmatrix}, \begin{pmatrix} 3 \\ 1 \\ 4 \end{pmatrix} \right\}$ linearly independent?

Try using the last fact: the set $\{v_1, v_2, \ldots, v_k\}$ is linearly dependent if and only if some v_i lies in the span of $v_1, \ldots, v_{i-1}$.

Linear independence and free variables

Theorem. Let $v_1, \ldots, v_k$ be vectors in $\mathbb{R}^n$ and consider the vector equation

$$x_1 v_1 + \cdots + x_k v_k = 0.$$

The set of vectors corresponding to non-free variables are linearly independent.

So, given a bunch of vectors $v_1, \ldots, v_k$, if you want to find a collection of v_i that are linearly independent, you put them in the columns of a matrix, row reduce, find the pivots, and then take the original v_i corresponding to those columns.

Example. Try this with $(1, 1, 1)$, $(2, 2, 2)$, and $(1, 2, 3)$.

Linear independence and coordinates

Fact. If $v_1, \ldots, v_k$ are linearly independent vectors then we can write each element of

$$\mathrm{Span}\{v_1, \ldots, v_k\}$$

in exactly one way as a linear combination of $v_1, \ldots, v_k$.

Span and Linear Independence

Two More Facts

Fact 1. Say $v_1, \ldots, v_k$ are in $\mathbb{R}^n$. If $k > n$, then $\{v_1, \ldots, v_k\}$ is linearly dependent.

Fact 2. If one of $v_1, \ldots, v_k$ is 0, then $\{v_1, \ldots, v_k\}$ is linearly dependent.

Parametric vector form and linear independence

Poll

Say you find the parametric vector form for a homogeneous system of linear equations, and you find that the set of solutions is the span of certain vectors. Then those vectors are...

1. always linearly independent
2. sometimes linearly independent
3. never linearly independent

Example. In Section 2.4 we solved the matrix equation $Ax = 0$ where

$$A = \begin{pmatrix} 1 & 2 & 0 & -1 \\ -2 & -3 & 4 & 5 \\ 2 & 4 & 0 & -2 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & -8 & -7 \\ 0 & 1 & 4 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

In parametric vector form, the solution is:

$$x_3 \begin{pmatrix} 8 \\ -4 \\ 1 \\ 0 \end{pmatrix} + x_4 \begin{pmatrix} 7 \\ -3 \\ 0 \\ 1 \end{pmatrix}$$

Parametric Vector Forms and Linear Independence

In Section 2.4 we solved the matrix equation $Ax = 0$ where

$$A = \begin{pmatrix} 1 & 2 & 0 & -1 \\ -2 & -3 & 4 & 5 \\ 2 & 4 & 0 & -2 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & -8 & -7 \\ 0 & 1 & 4 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

In parametric vector form, the solution is:

$$x_3 \begin{pmatrix} 8 \\ -4 \\ 1 \\ 0 \end{pmatrix} + x_4 \begin{pmatrix} 7 \\ -3 \\ 0 \\ 1 \end{pmatrix}$$

The two vectors that appear are linearly independent (why?). This means that we can't write the solution with fewer than two vectors (why?). This also means that this way of writing the solution set is efficient: for each solution, there is only one choice of x_3 and x_4 that gives that solution.

Summary of Section 2.5

- A set of vectors $\{v_1, \ldots, v_k\}$ in $\mathbb{R}^n$ is linearly independent if the vector equation

$$x_1 v_1 + x_2 v_2 + \cdots + x_k v_k = 0$$

  has only the trivial solution. It is linearly dependent otherwise.
- The cols of A are linearly independent
  $\Leftrightarrow$ $Ax = 0$ has only the trivial solution.
  $\Leftrightarrow$ A has a pivot in each column
- The number of pivots of A equals the dimension of the span of the columns of A
- The set $\{v_1, \ldots, v_k\}$ is linearly independent $\Leftrightarrow$ they span a k-dimensional plane
- The set $\{v_1, \ldots, v_k\}$ is linearly dependent $\Leftrightarrow$ some v_i lies in the span of $v_1, \ldots, v_{i-1}$.
- To find a collection of linearly independent vectors among the $\{v_1, \ldots, v_k\}$, row reduce and take the (original) v_i corresponding to pivots.

Typical exam questions

- State the definition of linear independence.
- *Always/sometimes/never.* A collection of 99 vectors in $\mathbb{R}^{100}$ is linearly dependent.
- *Always/sometimes/never.* A collection of 100 vectors in $\mathbb{R}^{99}$ is linearly dependent.
- Find all values of h so that the following vectors are linearly independent:

$$\left\{ \begin{pmatrix} 5 \\ 7 \\ 1 \end{pmatrix}, \begin{pmatrix} -5 \\ 7 \\ 0 \end{pmatrix}, \begin{pmatrix} 10 \\ 0 \\ h \end{pmatrix} \right\}$$

- *True/false.* If A has a pivot in each column, then the rows of A are linearly independent.
- *True/false.* If u and v are vectors in $\mathbb{R}^5$ then $\{u, v, \sqrt{2}u - \pi v\}$ is linearly independent.
- If you have a set of linearly independent vectors, and their span is a line, how many vectors are in the set?

Section 2.6

Subspaces

Outline of Section 2.6

- Definition of subspace
- Examples and non-examples of subspaces
- Spoiler alert: Subspaces are the same as spans
- Spanning sets for subspaces
- Two important subspaces for a matrix: $\mathrm{Col}(A)$ and $\mathrm{Nul}(A)$

Subspaces

A subspace of $\mathbb{R}^n$ is a subset V of $\mathbb{R}^n$ with:

1. The zero vector is in V.
2. If u and v are in V, then $u + v$ is also in V.
3. If u is in V and c is a scalar, then cu is in V.

The second and third properties are called "closure under addition" and "closure under scalar multiplication."

Together, the second and third properties could together be rephrased as: closure under linear combinations.

Which are subspaces?

1. the unit circle in $\mathbb{R}^2$
2. the point $(1, 2, 3)$ in $\mathbb{R}^3$
3. the xy-plane in $\mathbb{R}^3$
4. the xy-plane together with the z-axis in $\mathbb{R}^3$

Which are subspaces?

Poll

Is the first quadrant of $\mathbb{R}^2$ a subspace?

1. yes
2. no

Which are subspaces?

1. $\left\{ \begin{pmatrix} a \\ b \end{pmatrix} \text{ in } \mathbb{R}^2 \mid a+b=0 \right\}$

2. $\left\{ \begin{pmatrix} a \\ b \end{pmatrix} \text{ in } \mathbb{R}^2 \mid a+b=1 \right\}$

3. $\left\{ \begin{pmatrix} a \\ b \end{pmatrix} \text{ in } \mathbb{R}^2 \mid ab \neq 0 \right\}$

4. $\left\{ \begin{pmatrix} a \\ b \end{pmatrix} \text{ in } \mathbb{R}^2 \mid a, b \text{ rational} \right\}$

Spans and subspaces

Fact. Any $\mathrm{Span}\{v_1, \ldots, v_k\}$ is a subspace.

Why?

Fact. Every subspace V is a span.

Why?

So now we know that three things are the same:

- subspaces
- spans
- planes through 0

So why bother with the word "subspace"? Sometimes easier to check a subset is a subspace than to check it is a span (see null spaces, eigenspaces). Also, it makes sense (and is often useful) to think of a subspace *without a particular spanning set in mind*. Try thinking of other examples where it is useful to have two names for the same thing, like: water / $\mathrm{H_2O}$ or free throw / foul shot.

Column Space and Null Space

$A = m \times n$ matrix.

$\mathrm{Col}(A) =$ column space of $A =$ span of the columns of A

$\mathrm{Nul}(A) =$ null space of $A =$ (set of solutions to $Ax = 0$)

Example. $A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \end{pmatrix}$

$\mathrm{Col}(A) =$ subspace of $\mathbb{R}^m$

$\mathrm{Nul}(A) =$ subspace of $\mathbb{R}^n$

We have already been interested in both. We have been computing null spaces all semester. Also, we have seen that $Ax = b$ is consistent exactly when b is in the span of the columns of A, or, b is in $\mathrm{Col}(A)$.

Spanning sets for $\mathrm{Nul}(A)$ and $\mathrm{Col}(A)$

Find spanning sets for $\mathrm{Nul}(A)$ and $\mathrm{Col}(A)$

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

Spanning sets for $\mathrm{Nul}(A)$ and $\mathrm{Col}(A)$

In general:

- our usual parametric solution for $Ax = 0$ gives a spanning set for $\mathrm{Nul}(A)$
- the pivot columns of A form a spanning set for $\mathrm{Col}(A)$

Warning! Not the pivot columns of the reduced matrix.

Notice that the columns of A form a (possibly larger) spanning set. We'll see later that the above recipe is the smallest spanning set.

Spanning sets

Find a spanning set for the plane $2x + 3y + z = 0$ in $\mathbb{R}^3$.

Subspaces and Null spaces

Fact. Every subspace is a null space.

Why? Given a spanning set, you can reverse engineer the A...

It's actually a little tricky to do this. Given the spanning set, you make those vectors the rows of a matrix, then row reduce and find vector parametric form, and then make those vectors the rows of a new matrix. Why does this work? Try an example!

Example. Find a matrix A whose null space is the span of $(1, 1, 1)$ and $(1, 2, 3)$. You should get the matrix $A = (1 \quad -2 \quad 1)$.

So now we know that four things are the same:

- subspaces
- spans
- planes through 0
- solutions to $Ax = 0$

(Make sure you understand what we mean when we say these are all the same!)

So why learn about subspaces?

If subspaces are the same as spans, planes through the origin, and solutions to $Ax = 0$, why bother with this new vocabulary word?

The point is that we have been throwing around terms like "3-dimensional plane in $\mathbb{R}^4$" all semester, but we never said what "dimension" and "plane" are. Subspaces give the proper way to define a plane. Soon we will learn the meaning of a dimension of a subspace.

All the ways

Here are all the ways we know to describe a subspace:

1. As span:

$$\mathrm{Span}\left\{\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}\right\}$$

2. As a column space:

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

3. As a null space:

$$\mathrm{Nul}\begin{pmatrix} 1 & 1 & 1 \end{pmatrix}$$

4. As the set of solutions to a homogeneous linear system:

$$x + y + z = 0$$

5. Same, but in set builder notation:

$$\left\{\begin{pmatrix} a \\ b \\ c \end{pmatrix} : a + b + c = 0\right\}$$

Section 2.6 Summary

- A subspace of $\mathbb{R}^n$ is a subset V with:
  1. The zero vector is in V.
  2. If u and v are in V, then $u + v$ is also in V.
  3. If u is in V and c is in $\mathbb{R}$, then $cu \in V$.
- Two important subspaces: $\mathrm{Nul}(A)$ and $\mathrm{Col}(A)$
- Find a spanning set for $\mathrm{Nul}(A)$ by solving $Ax = 0$ in vector parametric form
- Find a spanning set for $\mathrm{Col}(A)$ by taking pivot columns of A (not reduced A)
- Four things are the same: subspaces, spans, planes through 0, null spaces

Typical exam questions

- Consider the set $\{(x, y) \in \mathbb{R}^2 \mid xy \geq 0\}$. Is it a subspace? If not, which properties does it fail?
- Consider the x-axis in $\mathbb{R}^3$. Is it a subspace? If not, which properties does it fail?
- Consider the set $\{(x, y, z, w) \in \mathbb{R}^4 \mid x + y - z + w = 0\}$. Is it a subspace? If not, which properties does it fail?
- Find spanning sets for the column space and the null space of

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

- True/False: The set of solutions to a matrix equation is always a subspace.
- True/False: The zero vector is a subspace.

Section 2.7

Bases

Bases

V = subspace of $\mathbb{R}^n$

A basis for V is a set of vectors $\{v_1, v_2, \ldots, v_k\}$ such that

1. $V = \text{Span}\{v_1, \ldots, v_k\}$
2. $v_1, \ldots, v_k$ are linearly independent

Equivalently, a basis is a *minimal spanning set*, that is, a spanning set where if you remove any one of the vectors you no longer have a spanning set.

Q. What is one basis for $\mathbb{R}^2$? $\mathbb{R}^n$? How many bases are there?

Dimension

$V =$ subspace of $\mathbb{R}^n$

$\dim(V) =$ dimension of $V = k =$the number of vectors in the basis

(What is the problem with this definition of dimension?)

Basis example

Find a basis for the xy-plane in $\mathbb{R}^3$? Find all bases for the xy-plane in R^3. (Remember: a basis is a set of vectors in the subspace that span the subspace and are linearly independent.)

Bases for $\mathbb{R}^n$

Let us consider the special case where V is equal to all of $\mathbb{R}^n$.

What are all bases for $V = \mathbb{R}^n$? Or, if we have a set of vectors $\{v_1, \ldots, v_k\}$, how do we check if they form a basis for $\mathbb{R}^n$? First, we make them the columns of a matrix....

- For the vectors to be linearly independent we need a pivot in every column.
- For the vectors to span $\mathbb{R}^n$ we need a pivot in every row.

Conclusion: $k = n$ and the matrix has n pivots.

The standard basis for $\mathbb{R}^n$

We have the standard basis vectors for $\mathbb{R}^n$:

$$\begin{aligned} e_1 &= (1, 0, 0, \dots, 0) \\ e_2 &= (0, 1, 0, \dots, 0) \\ &\vdots \end{aligned}$$

Who cares about bases?

A basis $\{v_1, \ldots, v_k\}$ for a subspace V of $\mathbb{R}^n$ is useful because:

Every vector v in V can be written in exactly one way:

$$v = c_1 v_1 + \cdots + c_k v_k$$

So a basis gives coordinates for V, like latitude and longitude. See Section 2.8.

Bases for $\mathrm{Nul}(A)$ and $\mathrm{Col}(A)$

Find bases for $\mathrm{Nul}(A)$ and $\mathrm{Col}(A)$

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

Bases for $\mathrm{Nul}(A)$ and $\mathrm{Col}(A)$

Find bases for $\mathrm{Nul}(A)$ and $\mathrm{Col}(A)$

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

Bases for $\mathrm{Nul}(A)$ and $\mathrm{Col}(A)$

Find bases for $\mathrm{Nul}(A)$ and $\mathrm{Col}(A)$

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{pmatrix}$$

Bases for $\mathrm{Nul}(A)$ and $\mathrm{Col}(A)$

In general:

- our usual parametric solution for $Ax = 0$ gives a basis for $\mathrm{Nul}(A)$
- the pivot columns of A form a basis for $\mathrm{Col}(A)$

Warning! Not the pivot columns of the reduced matrix.

What should you do if you are asked to find a basis for Span$\{v_1, \ldots, v_k\}$?

Bases for planes

Find a basis for the plane $2x + 3y + z = 0$ in $\mathbb{R}^3$.

Basis theorem

Basis Theorem

If V is a k-dimensional subspace of $\mathbb{R}^n$, then

- any k linearly independent vectors of V form a basis for V
- any k vectors that span V form a basis for V

In other words if a set has two of these three properties, it is a basis:

spans V, linearly independent, k vectors

We are skipping Section 2.8 this semester. But remember: the whole point of a basis is that it gives coordinates (like latitude and longitude) for a subspace. Every point has a unique address.

Section 2.7 Summary

- A basis for a subspace V is a set of vectors $\{v_1, v_2, \ldots, v_k\}$ such that
  1. $V = \text{Span}\{v_1, \ldots, v_k\}$
  2. $v_1, \ldots, v_k$ are linearly independent
- The number of vectors in a basis for a subspace is the dimension.
- Find a basis for $\mathrm{Nul}(A)$ by solving $Ax = 0$ in vector parametric form
- Find a basis for $\mathrm{Col}(A)$ by taking pivot columns of A (not reduced A)
- Basis Theorem. Suppose V is a k-dimensional subspace of $\mathbb{R}^n$. Then
  - Any k linearly independent vectors in V form a basis for V.
  - Any k vectors in V that span V form a basis.

Typical exam questions

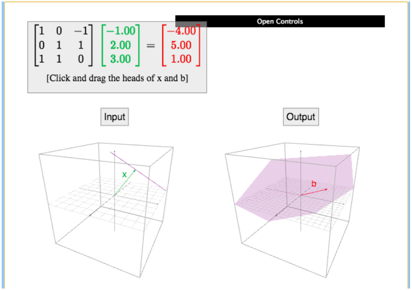
- Find a basis for the yz-plane in $\mathbb{R}^3$
- Find a basis for $\mathbb{R}^3$ where no vector has a zero
- How many vectors are there in a basis for a line in R^7?
- True/false: every basis for a plane in $\mathbb{R}^3$ has exactly two vectors.
- True/false: if two vectors lie in a plane through the origin in $\mathbb{R}^3$ and they are not collinear then they form a basis for the plane.
- True/false: The dimension of the null space of A is the number of pivots of A.
- True/false: If b lies in the column space of A, and the columns of A are linearly independent, then $Ax = b$ has infinitely many solutions.
- True/false: Any three vectors that span R^3 must be linearly independent.

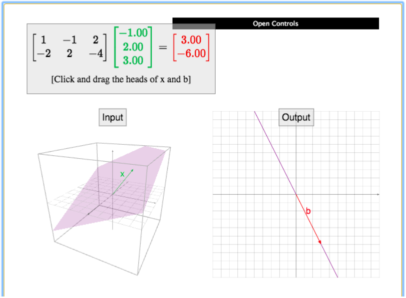
Section 2.9

The rank theorem

Rank Theorem

On the left are solutions to $Ax = 0$, on the right is $\mathrm{Col}(A)$:





Rank Theorem

$$\text{rank}(A) = \dim \text{Col}(A) = \#\text{ pivot columns}$$
$$\text{nullity}(A) = \dim \text{Nul}(A) = \#\text{ nonpivot columns}$$

Rank Theorem. $\text{rank}(A) + \text{nullity}(A) = \#\text{cols}(A)$

This ties together everything in the whole chapter: rank A describes the b's so that $Ax = b$ is consistent and the nullity describes the solutions to $Ax = 0$. So more flexibility with b means less flexibility with x, and vice versa.

Example. $A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$

About names

Again, why did we need all these vocabulary words? One answer is that the rank theorem would be harder to understand if it was:

The size of a minimal spanning set for the set of solutions to $Ax = 0$ plus the size of a minimal spanning set for the set of b so that $Ax = b$ has a solution is equal to the number of columns of A.

Compare to: $\text{rank}(A) + \text{nullity}(A) = n$

"A common concept in history is that knowing the name of something or someone gives one power over that thing or person." –Loren Graham
`http://philoctetes.org/news/the_power_of_names_religion_mathematics`

Section 2.9 Summary

- Rank Theorem. $\operatorname{rank}(A) + \dim \operatorname{Nul}(A) = \#\text{cols}(A)$

Typical exam questions

- Suppose that A is a 5×7 matrix, and that the column space of A is a line in $\mathbb{R}^5$. Describe the set of solutions to $Ax = 0$.
- Suppose that A is a 5×7 matrix, and that the column space of A is $\mathbb{R}^5$. Describe the set of solutions to $Ax = 0$.
- Suppose that A is a 5×7 matrix, and that the null space is a plane. Is $Ax = b$ consistent, where $b = (1, 2, 3, 4, 5)$?
- True/false. There is a 3×2 matrix so that the column space and the null space are both lines.
- True/false. There is a 2×3 matrix so that the column space and the null space are both lines.
- True/false. Suppose that A is a 6×2 matrix and that the column space of A is 2-dimensional. Is it possible for $(1, 0)$ and $(1, 1)$ to be solutions to $Ax = b$ for some b in $\mathbb{R}^6$?

Chapter 3

Linear Transformations and Matrix Algebra

Where are we?

In Chapter 1 we learned to solve all linear systems algebraically.

In Chapter 2 we learned to think about the solutions geometrically.

In Chapter 3 we continue with the algebraic abstraction. We learn to think about solving linear systems in terms of inputs and outputs. This is similar to control systems in AE, objects in computer programming, or hot pockets in a microwave.

More specifically, we think of a matrix as giving rise to a function with inputs and outputs. Solving a linear system means finding an input that produces a desired output. We will see that sometimes these functions are invertible, which means that you can reverse the function, inputting the outputs and outputting the inputs.

The invertible matrix theorem is the highlight of the chapter; it tells us when we can reverse the function. As we will see, it ties together everything in the course.

Sections 3.1

Matrix Transformations

Section 3.1 Outline

- Learn to think of matrices as functions, called matrix transformations
- Learn the associated terminology: domain, codomain, range
- Understand what certain matrices do to $\mathbb{R}^n$

From matrices to functions

Let A be an $m \times n$ matrix.

We define a function

$$T : \mathbb{R}^n \to \mathbb{R}^m$$
$$T(v) = Av$$

This is called a matrix transformation.

The domain of T is $\mathbb{R}^n$.

The co-domain of T is $\mathbb{R}^m$.

The range of T is the set of outputs: $\mathrm{Col}(A)$

This gives us a*nother* point of view of $Ax = b$

Demo

Example

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

What is $T(u)$?

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

Find a vector in $\mathbb{R}^3$ that is not in the range of T.

Square matrices

For a square matrix we can think of the associated matrix transformation

$$T : \mathbb{R}^n \to \mathbb{R}^n$$

as doing something to $\mathbb{R}^n$.

Example. The matrix transformation T for

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

What does T do to $\mathbb{R}^2$?

Square matrices

What does each matrix do to $\mathbb{R}^2$?

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

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

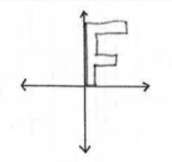
$$\begin{pmatrix} 3 & 0 \\ 0 & 3 \end{pmatrix}$$

What is the range in each case?

Poll

Poll

What does $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ do to this letter F?



Square matrices

What does each matrix do to $\mathbb{R}^2$?

Hint: if you can't see it all at once, see what happens to the x- and y-axes.

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

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

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

Examples in $\mathbb{R}^3$

What does each matrix do to $\mathbb{R}^3$?

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

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

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

Why are we learning about matrix transformations?

Sample applications:

- Cryptography (Hill cypher)
- Computer graphics (Perspective projection is a linear map!)
- Aerospace (Control systems - input/output)
- Biology
- Many more!

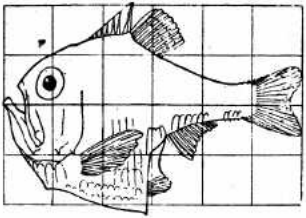


Fig. 517. *Argyropelecus Olfersi.*

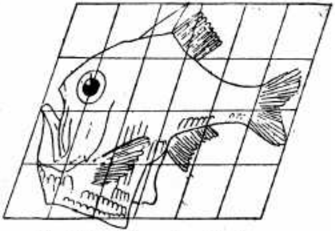


Fig. 518. *Sternoptyx diaphana.*

Applications of Linear Algebra

Biology: In a population of rabbits...

- half of the new born rabbits survive their first year
- of those, half survive their second year
- the maximum life span is three years
- rabbits produce 0, 6, 8 rabbits in their first, second, and third years

If I know the population in 2016 (in terms of the number of first, second, and third year rabbits), then what is the population in 2017?

These relations can be represented using a matrix.

$$\begin{pmatrix} 0 & 6 & 8 \\ \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & 0 \end{pmatrix}$$

How does this relate to matrix transformations?

▸ Demo

Section 3.1 Summary

- If A is an $m \times n$ matrix, then the associated matrix transformation T is given by $T(v) = Av$. This is a function with domain $\mathbb{R}^n$ and codomain $\mathbb{R}^m$ and range $\mathrm{Col}(A)$.
- If A is $n \times n$ then T does something to $\mathbb{R}^n$; basic examples: reflection, projection, scaling, shear, rotation

Typical exam questions

- What does the matrix $\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$ do to $\mathbb{R}^2$?
- What does the matrix $\begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix}$ do to $\mathbb{R}^2$?
- What does the matrix $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ do to $\mathbb{R}^3$?
- What does the matrix $\begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}$ do to $\mathbb{R}^2$?
- True/false. If A is a matrix and T is the associated matrix transformation, then the statement $Ax = b$ is consistent is equivalent to the statement that b is in the range of T.
- True/false. There is a matrix A so that the domain of the associated matrix transformation is a line in $\mathbb{R}^3$.

Section 3.2

One-to-one and onto transformations

Section 3.2 Outline

- Learn the definitions of one-to-one and onto functions
- Determine if a given matrix transformation is one-to-one and/or onto

One-to-one and onto in calculus

What do one-to-one and onto mean for a function $f : \mathbb{R} \to \mathbb{R}$?

One-to-one

A matrix transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ is one-to-one if each b in $\mathbb{R}^m$ is the output for at most one v in $\mathbb{R}^n$.

In other words: different inputs have different outputs.

Do not confuse this with the definition of a function, which says that for each input x in $\mathbb{R}^n$ there is at most one output b in $\mathbb{R}^m$.

One-to-one

$T : \mathbb{R}^n \to \mathbb{R}^m$ is one-to-one if each b in $\mathbb{R}^m$ is the output for at most one v in $\mathbb{R}^n$.

Theorem. Suppose $T : \mathbb{R}^n \to \mathbb{R}^m$ is a matrix transformation with matrix A. Then the following are all equivalent:

- T is one-to-one
- the columns of A are linearly independent
- $Ax = 0$ has only the trivial solution
- A has a pivot in each column
- the range of T has dimension n

What can we say about the relative sizes of m and n if T is one-to-one?

Draw a picture of the range of a one-to-one matrix transformation $\mathbb{R} \to \mathbb{R}^3$.

Onto

A matrix transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ is onto if the range of T equals the codomain $\mathbb{R}^m$, that is, each b in $\mathbb{R}^m$ is the output for at least one input v in $\mathbb{R}^m$.

Onto

$T : \mathbb{R}^n \to \mathbb{R}^m$ is onto if the range of T equals the codomain $\mathbb{R}^m$, that is, each b in $\mathbb{R}^m$ is the output for at least one input v in $\mathbb{R}^m$.

Theorem. Suppose $T : \mathbb{R}^n \to \mathbb{R}^m$ is a matrix transformation with matrix A. Then the following are all equivalent:

- T is onto
- the columns of A span $\mathbb{R}^m$
- A has a pivot in each row
- $Ax = b$ is consistent for all b in $\mathbb{R}^m$
- the range of T has dimension m

What can we say about the relative sizes of m and n if T is onto?

Give an example of an onto matrix transformation $\mathbb{R}^3 \to \mathbb{R}$.

One-to-one and Onto

Do the following give matrix transformations that are one-to-one? onto?

$$\begin{pmatrix} 1 & 0 & 7 \\ 0 & 1 & 2 \\ 0 & 0 & 9 \end{pmatrix} \quad \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 2 & 1 \end{pmatrix} \quad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \quad \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$$

One-to-one and Onto

Which of the previously-studied matrix transformations of $\mathbb{R}^2$ are one-to-one? Onto?

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

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

$$\begin{pmatrix} 3 & 0 \\ 0 & 3 \end{pmatrix} \quad \text{scaling}$$

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

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \quad \text{rotation}$$

Which are one to one / onto?

Poll

Which give one to one-to-one / onto matrix transformations?

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

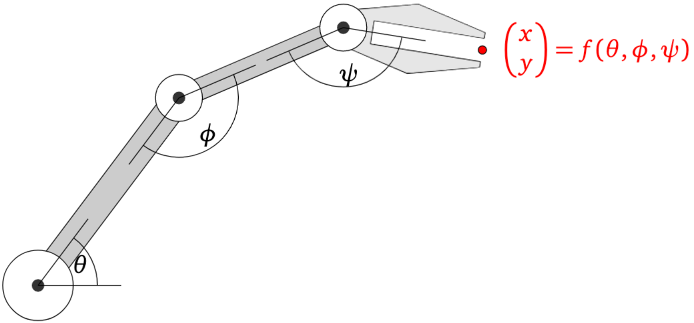
Demo

Demo

Demo

Robot arm

Consider the robot arm example from the book.



There is a natural function f here (not a matrix transformation). The input is a set of three angles and the co-domain is $\mathbb{R}^2$. Is this function one-to-one? Onto?

The geometry

Say that $T : \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation.

The geometry of one-to-one:

The range has dimension n (and the null space is a point).

The geometry of onto:

The range has dimension m, so it is all of $\mathbb{R}^m$ (and the null space has dimension $n - m$).

Summary of Section 3.2

- $T : \mathbb{R}^n \to \mathbb{R}^m$ is one-to-one if each b in $\mathbb{R}^m$ is the output for at most one v in $\mathbb{R}^n$.
- **Theorem.** Suppose $T : \mathbb{R}^n \to \mathbb{R}^m$ is a matrix transformation with matrix A. Then the following are all equivalent:
  - T is one-to-one
  - the columns of A are linearly independent
  - $Ax = 0$ has only the trivial solution
  - A has a pivot in each column
  - the range has dimension n
- $T : \mathbb{R}^n \to \mathbb{R}^m$ is onto if the range of T equals the codomain $\mathbb{R}^m$, that is, each b in $\mathbb{R}^m$ is the output for at least one input v in $\mathbb{R}^m$.
- **Theorem.** Suppose $T : \mathbb{R}^n \to \mathbb{R}^m$ is a matrix transformation with matrix A. Then the following are all equivalent:
  - T is onto
  - the columns of A span $\mathbb{R}^m$
  - A has a pivot in each row
  - $Ax = b$ is consistent for all b in $\mathbb{R}^m$.
  - the range of T has dimension m

Typical exam questions

- True/False. It is possible for the matrix transformation for a 5×6 matrix to be both one-to-one and onto.
- True/False. The matrix transformation $T : \mathbb{R}^3 \to \mathbb{R}^3$ given by projection to the yz-plane is onto.
- True/False. The matrix transformation $T : \mathbb{R}^2 \to \mathbb{R}^2$ given by rotation by π is onto.
- Is there an onto matrix transformation $\mathbb{R}^2 \to \mathbb{R}^3$? If so, write one down, if not explain why not.
- Is there an one-to-one matrix transformation $\mathbb{R}^2 \to \mathbb{R}^3$? If so, write one down, if not explain why not.

Section 3.3

Linear Transformations

Section 3.3 Outline

- Understand the definition of a linear transformation
- Linear transformations are the same as matrix transformations
- Find the matrix for a linear transformation

Linear transformations

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

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

First examples: matrix transformations.

Linear transformations

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

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

Notice that $T(0) = 0$. *Why?*

We have the standard basis vectors for $\mathbb{R}^n$:

$$\begin{aligned} e_1 &= (1, 0, 0, \ldots, 0) \\ e_2 &= (0, 1, 0, \ldots, 0) \\ &\vdots \end{aligned}$$

If we know $T(e_1), \ldots, T(e_n)$, then we know every $T(v)$. *Why?*

In engineering, this is called the principle of superposition.

Which are linear transformations?

And why?

$$T\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x+y \\ y \\ x-y \end{pmatrix}$$

$$T\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x+y+1 \\ y \\ x-y \end{pmatrix}$$

$$T\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} xy \\ y \\ x-y \end{pmatrix}$$

A function $\mathbb{R}^n \to \mathbb{R}^m$ is linear exactly when the coordinates are linear (linear combinations of the variables, no constant terms).

Linear transformations

Which properties of a linear transformation fail for this function $T : \mathbb{R}^2 \to \mathbb{R}^2$?

$$T\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ |y| \end{pmatrix}$$

Linear transformations are matrix transformations

Theorem. Every linear transformation is a matrix transformation.

This means that for any linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ there is an $m \times n$ matrix A so that

$$T(v) = Av$$

for all v in $\mathbb{R}^n$.

The matrix for a linear transformation is called the standard matrix.

Linear transformations are matrix transformations

Theorem. Every linear transformation is a matrix transformation.

Given a linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ the standard matrix is:

$$A = \begin{pmatrix} | & | & & | \\ T(e_1) & T(e_2) & \cdots & T(e_n) \\ | & | & & | \end{pmatrix}$$

Why? Notice that $Ae_i = T(e_i)$ for all i. Then it follows from linearity that $T(v) = Av$ for all v.

The identity

The identity linear transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ is

$$T(v) = v$$

What is the standard matrix?

This standard matrix is called I_n or I.

Linear transformations are matrix transformations

Suppose $T : \mathbb{R}^2 \to \mathbb{R}^3$ is the function given by:

$$T\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x+y \\ y \\ x-y \end{pmatrix}$$

What is the standard matrix for T?

Linear transformations are matrix transformations

Find the standard matrix for the linear transformation of $\mathbb{R}^2$ that stretches by 2 in the x-direction and 3 in the y-direction, and then reflects over the line $y = x$.

Linear transformations are matrix transformations

Find the standard matrix for the linear transformation of $\mathbb{R}^2$ that projects onto the y-axis and then rotates counterclockwise by $\pi/2$.

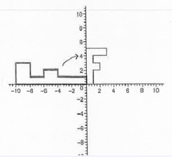
Linear transformations are matrix transformations

Find the standard matrix for the linear transformation of $\mathbb{R}^3$ that reflects through the xy-plane and then projects onto the yz-plane.

Discussion

Discussion Question

Find a matrix that does this.



Summary of Section 3.3

- A function $T : \mathbb{R}^n \to \mathbb{R}^m$ is linear if
  - $T(u+v) = T(u) + T(v)$ for all u, v in $\mathbb{R}^n$.
  - $T(cv) = cT(v)$ for all $v \in \mathbb{R}^n$ and c in $\mathbb{R}$.
- **Theorem.** Every linear transformation is a matrix transformation (and vice versa).
- The standard matrix for a linear transformation has its ith column equal to $T(e_i)$.

Typical Exam Questions Section 3.3

- Is the function $T : \mathbb{R} \to \mathbb{R}$ given by $T(x) = x + 1$ a linear transformation?
- Suppose that $T : \mathbb{R}^2 \to \mathbb{R}^3$ is a linear transformation and that

$$T\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 3 \\ 1 \end{pmatrix} \quad \text{and} \quad T\begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix}$$

  What is

$$T\begin{pmatrix} 1 \\ 0 \end{pmatrix}?$$

- Find the matrix for the linear transformation $T : \mathbb{R}^3 \to \mathbb{R}^3$ that rotates about the z-axis by π and then scales by 2.
- Suppose $T : \mathbb{R}^3 \to \mathbb{R}^3$ is the function given by:

$$T\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} z \\ 0 \\ x \end{pmatrix}$$

  Is this a linear transformation? If so, what is the standard matrix for T?
- Is the identity transformation one-to-one?

Chapter 4

Determinants

Where are we?

- We have studied the problem $Ax = b$
- We next want to study $Ax = \lambda x$
- At the end of the course we want to almost solve $Ax = b$

We need determinants for the second item.

Section 4.1

The definition of the determinant

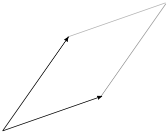
Outline of Sections 4.1 and 4.3

- Volume and invertibility
- A definition of determinant in terms of row operations
- Using the definition of determinant to compute the determinant
- Determinants of products: $\det(AB)$
- Determinants and linear transformations and volumes

Invertibility and volume

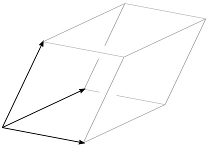
When is a 2×2 matrix invertible? $\leftarrow$ Algebra

When the rows (or columns) don't lie on a line $\Leftrightarrow$ the corresponding parallelogram has non-zero area. $\leftarrow$ Geometry



When is a 3×3 matrix invertible?

When the rows (or columns) don't lie on a plane $\Leftrightarrow$ the corresponding parallelepiped (3D parallelogram) has non-zero volume



Same for $n \times n$!

The definition of determinant

The determinant of a *square* matrix is a number so that

1. If we do a row replacement on a matrix, the determinant is unchanged
2. If we swap two rows of a matrix, the determinant scales by -1
3. If we scale a row of a matrix by k, the determinant scales by k
4. $\det(I_n) = 1$

Why would we think of this? *Answer: This is exactly how volume works.*

Try it out for 2×2 matrices.

The definition of determinant

The determinant of a *square* matrix is a number so that

1. If we do a row replacement on a matrix, the determinant is unchanged
2. If we swap two rows of a matrix, the determinant scales by -1
3. If we scale a row of a matrix by k, the determinant scales by k
4. $\det(I_n) = 1$

Problem. Just using these rules, compute the determinants:

$$\begin{pmatrix} 1 & 0 & 8 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \quad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 17 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \begin{pmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{pmatrix}$$

A basic fact about determinants

Fact. If A has a zero row, then $\det(A) = 0$.

Fact. If A is a diagonal matrix then $\det(A)$ is the product of the diagonal entries.

Fact. If A is in row echelon form then $\det(A)$ is the product of the diagonal entries.

Why do these follow from the definition?

A first formula for the determinant

Fact. Suppose we row reduce A. Then

$$\det A = (-1)^{\#\text{row swaps used}} \left(\frac{\text{product of diagonal entries of row reduced matrix}}{\text{product of scalings used}} \right)$$

Use the fact to get a formula for the determinant of any 2×2 matrix.

Consequence of the above fact:

Fact. $\det A \neq 0 \Leftrightarrow A$ invertible

Computing determinants

...using the definition in terms of row operations

$$\det\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 5 & 7 & -4 \end{pmatrix} =$$

Computing determinants

...using the definition in terms of row operations

$$\det \begin{pmatrix} 0 & 6 & 8 \\ 1/2 & 0 & 0 \\ 0 & 1/2 & 0 \end{pmatrix} =$$

A Mathematical Conundrum

We have this definition of a determinant, and it gives us a way to compute it.

But: we don't know that such a determinant function exists.

More specifically, we haven't ruled out the possibility that two different row reductions might gives us two different answers for the determinant.

Don't worry! It is all okay.

We already gave the key idea: that determinant is just the volume of the corresponding parallelepiped. You can read the proof in the book if you want.

Fact 1. There is such a number $\det$ and it is unique.

Properties of the determinant

Fact 1. There is such a number $\det$ and it is unique.

Fact 2. A is invertible $\Leftrightarrow \det(A) \neq 0$ important!

Fact 3. $\det A = (-1)^{\#\text{row swaps used}} \left(\frac{\text{product of diagonal entries of row reduced matrix}}{\text{product of scalings used}} \right)$

Fact 4. The function can be computed by any of the $2n$ cofactor expansions.

Fact 5. $\det(AB) = \det(A)\det(B)$ important!

Fact 6. $\det(A^T) = \det(A)$ ok, now we need to say what transpose is

Fact 7. $\det(A)$ is signed volume of the parallelepiped spanned by cols of A.

If you want the proofs, see the book. Actually Fact 1 is the hardest!

Powers

Fact 5. $\det(AB) = \det(A)\det(B)$

Use this fact to compute

$$\det\left(\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 5 & 7 & -4 \end{pmatrix}^5\right)$$

What is $\det(A^{-1})$?

Powers

Poll

Suppose we know A^5 is invertible. Is A invertible?

1. yes
2. no
3. maybe

Section 4.3

The determinant and volumes

Areas of triangles

What is the area of the triangle in $\mathbb{R}^2$ with vertices $(1, 2)$, $(4, 3)$, and $(2, 5)$?

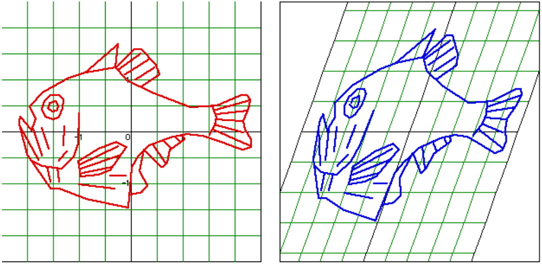
What is the area of the parallelogram in $\mathbb{R}^2$ with vertices $(1, 2)$, $(4, 3)$, $(2, 5)$, and $(5, 6)$?

Determinants and linear transformations

Say A is an $n \times n$ matrix and $T(v) = Av$.

Fact 8. If S is some subset of $\mathbb{R}^n$, then $\mathrm{vol}(T(S)) = |\det(A)| \cdot \mathrm{vol}(S)$.

This works even if S is curvy, like a circle or an ellipse, or:



Why? First check it for little squares/cubes (Fact 7). Then: Calculus!

Summary of Sections 4.1 and 4.3

Say det is a function $\det : \{\text{matrices}\} \to \mathbb{R}$ with:

1. $\det(I_n) = 1$
2. If we do a row replacement on a matrix, the determinant is unchanged
3. If we swap two rows of a matrix, the determinant scales by -1
4. If we scale a row of a matrix by k, the determinant scales by k

Fact 1. There is such a function det and it is unique.

Fact 2. A is invertible $\Leftrightarrow \det(A) \neq 0$ important!

Fact 3. $\det A = (-1)^{\#\text{row swaps used}} \left(\dfrac{\text{product of diagonal entries of row reduced matrix}}{\text{product of scalings used}} \right)$

Fact 4. The function can be computed by any of the $2n$ cofactor expansions.

Fact 5. $\det(AB) = \det(A)\det(B)$ important!

Fact 6. $\det(A^T) = \det(A)$

Fact 7. $\det(A)$ is signed volume of the parallelepiped spanned by cols of A.

Fact 8. If S is some subset of $\mathbb{R}^n$, then $\mathrm{vol}(T(S)) = |\det(A)| \cdot \mathrm{vol}(S)$.

Typical Exam Questions 4.1 and 4.3

- Find the value of h that makes the determinant 0:

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

- If the matrix on the left has determinant 5, what is the determinant of the matrix on the right?

$$\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} \qquad \begin{pmatrix} g & h & i \\ d & e & f \\ a-d & b-e & c-f \end{pmatrix}$$

- If the area of a fish (in a photo) is 7 square inches, and we apply a shear, what is the new area?
- Suppose that T is a linear transformation with the property that $T \circ T = T$. What is the determinant of the standard matrix for T?
- Suppose that T is a linear transformation with the property that $T \circ T = \text{identity}$. What is the determinant of the standard matrix for T?
- Find the volume of the triangular pyramid with vertices $(0,0,0)$, $(0,0,1)$, $(1,0,0)$, and $(1,2,3)$.

Section 4.2

Cofactor expansions

Outline of Section 4.2

- We will give a recursive formula for the determinant of a square matrix.

A formula for the determinant

We will give a recursive formula.

First some terminology:

$$\begin{aligned} A_{ij} &= ij\text{th minor of } A \\ &= (n-1)\times(n-1) \text{ matrix obtained by deleting the } i\text{th row and } j\text{th column} \end{aligned}$$

$$\begin{aligned} C_{ij} &= (-1)^{i+j}\det(A_{ij}) \\ &= ij\text{th cofactor of } A \end{aligned}$$

Finally:

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

Or:

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

So we find the determinant of a 3×3 matrix in terms of the determinants of 2×2 matrices, etc.

Determinants

Consider

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

Compute the following:

$a_{11} =$ $\qquad$ $a_{12} =$ $\qquad$ $a_{13} =$

$A_{11} =$ $\qquad$ $A_{12} =$ $\qquad$ $A_{13} =$

$\det A_{11} =$ $\qquad$ $\det A_{12} =$ $\qquad$ $\det A_{13} =$

$C_{11} =$ $\qquad$ $C_{12} =$ $\qquad$ $C_{13} =$

$$\det A = a_{11}C_{11} + a_{12}C_{12} + a_{13}C_{13}$$

A formula for the determinant

We can take the recursive formula further....

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

Say that....

1×1 matrices

$$\det(a_{11}) = a_{11}$$

Now apply the formula to...

2×2 matrices

$$\det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} =$$

(Could also go really nuts and define the determinant of a 0×0 matrix to be 1 and use the formula to get the formula for 1×1 matrices...)

A formula for the determinant

3×3 matrices

$$\det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \cdots$$

You can write this out. And it is a good exercise. But you won't want to memorize it.

A formula for the determinant

Another formula for 3×3 matrices

$$\det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32}$$
$$- a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33} - a_{11}a_{23}a_{32}$$

(Check this is gives the same answer as before. It is a small miracle!)

Use this formula to compute

$$\det \begin{pmatrix} 5 & 1 & 0 \\ -1 & 3 & 2 \\ 4 & 0 & -1 \end{pmatrix}$$

Expanding across other rows and columns

The formula we gave for $\det(A)$ is the expansion across the first row. It turns out you can compute the determinant by expanding across any row or column:

$$\det(A) = a_{i1}C_{i1} + \cdots + a_{in}C_{in} \text{ for any fixed } i$$
$$\det(A) = a_{1j}C_{1j} + \cdots + a_{nj}C_{nj} \text{ for any fixed } j$$

Or for odd rows and columns:

$$\det(A) = a_{i1}(\det(A_{i1})) - a_{i2}(\det(A_{i2})) + \cdots \pm a_{in}(\det(A_{in}))$$
$$\det(A) = a_{1j}(\det(A_{1j})) - a_{2j}(\det(A_{2j})) + \cdots \pm a_{nj}(\det(A_{nj}))$$

and for even rows and columns:

$$\det(A) = -a_{i1}(\det(A_{i1})) + a_{i2}(\det(A_{i2})) + \cdots \mp a_{in}(\det(A_{in}))$$
$$\det(A) = -a_{1j}(\det(A_{1j})) + a_{2j}(\det(A_{2j})) + \cdots \mp a_{nj}(\det(A_{nj}))$$

Compute:

$$\det \begin{pmatrix} 2 & 1 & 0 \\ 1 & 1 & 0 \\ 5 & 9 & 1 \end{pmatrix}$$

Determinants of triangular matrices

If A is upper (or lower) triangular, $\det(A)$ is easy to compute with cofactor expansions (it was also easy using the definition of the determinant):

$$\det\begin{pmatrix} 2 & 1 & 5 & -2 \\ 0 & 1 & 2 & -3 \\ 0 & 0 & 5 & 9 \\ 0 & 0 & 0 & 10 \end{pmatrix}$$

Determinants

Poll

What is the determinant?

$$\det \begin{pmatrix} 4 & 7 & 0 & 9 & 3 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 3 \\ 5 & 9 & 2 & 10 & 2 \\ 0 & 0 & 0 & 2 & 0 \end{pmatrix}$$

A formula for the inverse

(from Section 3.3)

2×2 matrices

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad \rightsquigarrow \quad A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

$n \times n$ matrices

$$\begin{aligned} A^{-1} &= \frac{1}{\det(A)} \begin{pmatrix} C_{11} & \cdots & C_{n1} \\ \vdots & \ddots & \vdots \\ C_{1n} & \ldots & C_{nn} \end{pmatrix} \\ &= \frac{1}{\det(A)} (C_{ij})^T \end{aligned}$$

Check that these agree!

The proof uses Cramer's rule (see the notes on the course home page. We're not testing on this - it's just for your information.)

Summary of Section 4.2

- There is a recursive formula for the determinant of a square matrix:

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

- We can use the same formula along any row/column.
- There are special formulas for the 2×2 and 3×3 cases.

Typical Exam Questions 4.2

- True or false. The cofactor expansion across the first row gives the negative of the cofactor expansion across the second row.
- Find the determinant of the following matrix using one of the formulas from this section:
$$\begin{pmatrix} 1 & 0 & -2 \\ 3 & 1 & -2 \\ -5 & 0 & 9 \end{pmatrix}$$
- Find the determinant of the following matrix using one of the formulas from this section:
$$\begin{pmatrix} 1 & 0 & -2 \\ 3 & 1 & -2 \\ -5 & -1 & 9 \end{pmatrix}$$
- Find the cofactor matrix for the above matrix and use it to find the inverse.

Chapter 5

Eigenvectors and eigenvalues

Where are we?

Remember:

> Almost every engineering problem, no matter how huge, can be reduced to linear algebra:
>
> $$Ax = b \quad \text{or}$$
> $$Ax = \lambda x$$

A few examples of the second: column buckling, control theory, image compression, exploring for oil, materials, natural frequency (bridges and car stereos), fluid mixing, RLC circuits, clustering (data analysis), principal component analysis, Google, Netflix (collaborative prediction), infectious disease models, special relativity, and many more!

We have said most of what we are going to say about the first problem. We now begin in earnest on the second problem.

A Question from Biology

In a population of rabbits...

- half of the new born rabbits survive their first year
- of those, half survive their second year
- the maximum life span is three years
- rabbits produce 0, 6, 8 rabbits in their first, second, and third years

If I know the population one year - think of it as a vector (f, s, t) - what is the population the next year?

$$\begin{pmatrix} 0 & 6 & 8 \\ \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & 0 \end{pmatrix} \begin{pmatrix} f \\ s \\ t \end{pmatrix}$$

Now choose some starting population vector $u = (f, s, t)$ and choose some number of years N. What is the new population after N years?

Demo

Section 5.1

Eigenvectors and eigenvalues

Eigenvectors and Eigenvalues

Suppose A is an $n \times n$ matrix and there is a $v \neq 0$ in $\mathbb{R}^n$ and λ in $\mathbb{R}$ so that

$$Av = \lambda v$$

then v is called an eigenvector for A, and λ is the corresponding eigenvalue.

In simpler terms: Av is a scalar multiple of v.

In other words: Av points in the same direction as v.

Think of this in terms of inputs and outputs!

eigen $=$ *characteristic (or: self)*

This the the most important definition in the course.

Demo

Eigenvectors and Eigenvalues

Suppose A is an $n \times n$ matrix and there is a $v \neq 0$ in $\mathbb{R}^n$ and λ in $\mathbb{R}$ so that

$$Av = \lambda v$$

then v is called an eigenvector for A, and λ is the corresponding eigenvalue.

Can you find any eigenvectors/eigenvalues for the following matrix?

$$A = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$$

What happens when you apply larger and larger powers of A to a vector?

Rabbits

What's up with them?

Eigenvectors and Eigenvalues

When we apply large powers of the matrix

$$A = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$$

to a vector v not on the x-axis, we see that $A^n v$ gets closer and closer to the y-axis, and it's length gets approximately tripled each time. This is because the largest eigenvalue is 3 and its eigenspace is the y-axis.

For the rabbit matrix

$$\begin{pmatrix} 0 & 6 & 8 \\ \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & 0 \end{pmatrix}$$

We will see that 2 is the largest eigenvalue, and its eigenspace is the span of the vector $(16, 4, 1)$. That's why all populations of rabbits tend towards the ratio 16:4:1 and why the population approximately doubles each year.

$$A = \begin{pmatrix} 0 & 6 & 8 \\ 1/2 & 0 & 0 \\ 0 & 1/2 & 0 \end{pmatrix}, \quad v = \begin{pmatrix} 16 \\ 4 \\ 1 \end{pmatrix}, \quad \lambda = 2$$

$$A = \begin{pmatrix} 2 & 2 \\ -4 & 8 \end{pmatrix}, \quad v = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \lambda = 4$$

How do you check?

Eigenvectors and Eigenvalues

Confirming eigenvectors

Poll

Which of $\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ are eigenvectors of

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

What are the eigenvalues?

Eigenvectors and Eigenvalues

Confirming eigenvalues

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

(Any eigenvector you find is called a 3-eigenvector.)

What is a general procedure for finding eigenvalues?

The following are equivalent:

- λ is an eigenvalue of A
- $\mathrm{Nul}(A - \lambda I)$ is nontrivial

So the recipe for checking if λ is an eigenvalue of A is:

- subtract λ from the diagonal entries of A
- row reduce
- check if there are fewer than n pivots

Confirm that $\lambda = 1$ is not an eigenvalue of $A = \begin{pmatrix} 2 & -4 \\ -1 & -1 \end{pmatrix}$.

Eigenspaces

Let A be an $n \times n$ matrix. The set of eigenvectors for a given eigenvalue λ of A (plus the zero vector) is a subspace of $\mathbb{R}^n$ called the λ-eigenspace of A.

Why is this a subspace?

Fact. λ-eigenspace for $A = \mathrm{Nul}(A - \lambda I)$

Example. Find the eigenspaces for $\lambda = 2$ and $\lambda = -1$ and sketch.

$$\begin{pmatrix} 5 & -6 \\ 3 & -4 \end{pmatrix}$$

Find a basis for the 2–eigenspace:

$$\begin{pmatrix} 4 & -1 & 6 \\ 2 & 1 & 6 \\ 2 & -1 & 8 \end{pmatrix}$$

Find a basis for the 2–eigenspace:

$$\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

Find a basis for the 2–eigenspace:

$$\begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Eigenspaces

Bases

Find a basis for the 2–eigenspace:

$$\begin{pmatrix} 1 & 5 \\ 0 & 1 \end{pmatrix}$$

Eigenvalues

And invertibility

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

Why?

Fact. The eigenvalues of a triangular matrix are the diagonal entries.

Why?

Important! You can not find the eigenvalues by row reducing first! After you find the eigenvalues, you row reduce $A - \lambda I$ to find the eigenspaces. But once you start row reducing the original matrix, you change the eigenvalues.

Eigenvalues

Distinct eigenvalues

Fact. If $v_1 \dots v_k$ are eigenvectors that correspond to distinct eigenvalues $\lambda_1, \dots \lambda_k$, then $\{v_1, \dots, v_k\}$ are linearly independent.

Why?

Consequence. An $n \times n$ matrix has at most n distinct eigenvalues.

Eigenvalues geometrically

If v is an eigenvector of A then that means v and Av are scalar multiples, i.e. they lie on a line.

Without doing any calculations, find the eigenvectors and eigenvalues of the matrices corresponding to the following linear transformations:

- Reflection about a line in $\mathbb{R}^2$ (doesn't matter which line!)
- Orthogonal projection onto a line in $\mathbb{R}^2$ (doesn't matter which line!)
- Scaling of $\mathbb{R}^2$ by 3
- (Standard) shear of $\mathbb{R}^2$
- Orthogonal projection to a plane in $\mathbb{R}^3$ (doesn't matter which plane!)

▸ Demo

Eigenvalues for rotations?

If v is an eigenvector of A then that means v and Av are scalar multiples, i.e. they lie on a line.

What are the eigenvectors and eigenvalues for rotation of $\mathbb{R}^2$ by $\pi/2$ (counterclockwise)?

Demo

Summary of Section 5.1

- If $v \neq 0$ and $Av = \lambda v$ then λ is an eigenvector of A with eigenvalue λ
- Given a matrix A and a vector v, we can check if v is an eigenvector for A: just multiply
- Recipe: The λ-eigenspace of A is the solution to $(A - \lambda I)x = 0$
- Fact. A invertible $\Leftrightarrow$ 0 is not an eigenvalue of A
- Fact. If $v_1 \dots v_k$ are distinct eigenvectors that correspond to distinct eigenvalues $\lambda_1, \dots \lambda_k$, then $\{v_1, \dots, v_k\}$ are linearly independent.
- We can often see eigenvectors and eigenvalues without doing calculations

Typical exam questions 5.1

- Find the 2-eigenvectors for the matrix

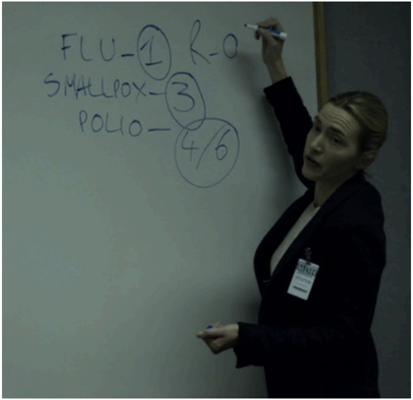
$$\begin{pmatrix} 0 & 13 & 12 \\ 1/4 & 0 & 0 \\ 0 & 1/2 & 0 \end{pmatrix}$$

- True or false: The zero vector is an eigenvector for every matrix.
- How many different eigenvalues can there be for an $n \times n$ matrix?
- Consider the reflection of $\mathbb{R}^2$ about the line $y = 7x$. What are the eigenvalues (of the standard matrix)?
- Consider the $\pi/2$ rotation of $\mathbb{R}^3$ about the z-axis. What are the eigenvalues (of the standard matrix)?

$$R_0$$

R_0

For a given virus, R_0 is the average number of people that each infected person infects. If R_0 is large, that is bad. Patient zero infects R_0 people, who then infect R_0^2 people, who then infect R_0^3 people. That is exponential growth. (If R_0 is less than 1, then that's good.)



R_0

For a given virus, R_0 is the average number of people that each infected person infects. If R_0 is large, that is bad. Patient zero infects R_0 people, who then infect R_0^2 people, who then infect R_0^3 people. That is exponential growth.

Whenever we see an exponential growth rate, we should think: eigenvalue.

It turns out that R_0 is an eigenvalue. The rough idea is very similar to our rabbit example: split the population into compartments, figure out how often each compartment infects each other compartment. That's a matrix. The largest eigenvalue is R_0.

R_0 is an eigenvalue

It turns out that R_0 is an eigenvalue. The rough idea is very similar to our rabbit example: split the population into compartments, figure out how often each compartment infects each other compartment.

For malaria, the compartments might be mosquitoes and humans.

For a sexually transmitted disease in a heterosexual population, the compartments might be males and females.

R_0 is an eigenvalue

It turns out that R_0 is an eigenvalue. The rough idea is very similar to our rabbit example: split the population into compartments, figure out how often each compartment infects each other compartment.

The SIR model has compartments for Susceptible, Infected, and Recovered.

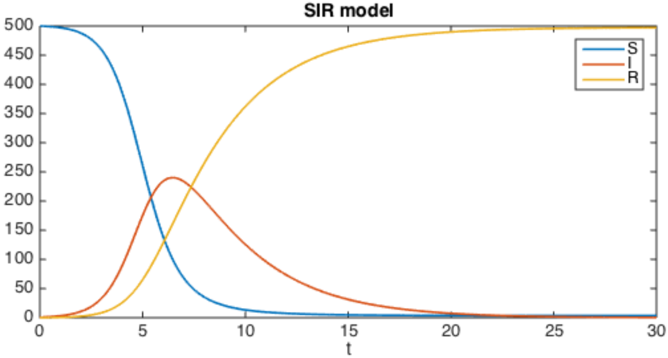
The arrows are governed by differential equations (Math 2552). Why do the labels on the arrows make sense? (The greek letters are constants).

There is a nice discussion of this by James Holland Jones (Stanford).

Paper

Bell curves

The growth rate of infection does not stay exponential forever, because the recovered population has immunity. That's where you get these bell curves.



Section 5.2

The characteristic polynomial

Outline of Section 5.2

- How to find the eigenvalues, via the characteristic polynomial
- Techniques for the 3×3 case

Characteristic polynomial

Recall:

$$\lambda \text{ is an eigenvalue of } A \iff A - \lambda I \text{ is not invertible}$$

So to find eigenvalues of A we solve

$$\det(A - \lambda I) = 0$$

The left hand side is a polynomial, the characteristic polynomial of A.

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

The eigenrecipe

Say you are given a square matrix A.

Step 1. Find the eigenvalues of A by solving

$$\det(A - \lambda I) = 0$$

Step 2. For each eigenvalue λ_i the λ_i-eigenspace is the solution to

$$(A - \lambda_i I)x = 0$$

To find a basis, find the vector parametric solution, as usual.

Characteristic polynomial

Find the characteristic polynomial and eigenvalues of

$$\begin{pmatrix} 5 & 2 \\ 2 & 1 \end{pmatrix}$$

Two shortcuts for 2×2 eigenvectors

Find the eigenspaces for the eigenvalues on the last page. Two tricks.

(1) We do not need to row reduce $A - \lambda I$ by hand; we know the bottom row will become zero.

(2) Then if the reduced matrix is:

$$A = \begin{pmatrix} x & y \\ 0 & 0 \end{pmatrix}$$

the eigenvector is

$$A = \begin{pmatrix} -y \\ x \end{pmatrix}$$

3×3 matrices

The 3×3 case is harder. There is a version of the quadratic formula for cubic polynomials, called Cardano's formula. But it is more complicated. It looks something like this:

$$
\begin{aligned}
x \;=\; & \sqrt[3]{\left(\frac{-b^3}{27a^3}+\frac{bc}{6a^2}-\frac{d}{2a}\right)+\sqrt{\left(\frac{-b^3}{27a^3}+\frac{bc}{6a^2}-\frac{d}{2a}\right)^2+\left(\frac{c}{3a}-\frac{b^2}{9a^2}\right)^3}} \\
+ \; & \sqrt[3]{\left(\frac{-b^3}{27a^3}+\frac{bc}{6a^2}-\frac{d}{2a}\right)-\sqrt{\left(\frac{-b^3}{27a^3}+\frac{bc}{6a^2}-\frac{d}{2a}\right)^2+\left(\frac{c}{3a}-\frac{b^2}{9a^2}\right)^3}} \;-\; \frac{b}{3a}.
\end{aligned}
$$

There is an even more complicated formula for quartic polynomials.

One of the most celebrated theorems in math, the Abel–Ruffini theorem, says that there is no such formula for quintic polynomials.

Characteristic polynomials

3×3 matrices

Find the characteristic polynomial of the following matrix.

$$\begin{pmatrix} 7 & 0 & 3 \\ -3 & 2 & -3 \\ -3 & 0 & -1 \end{pmatrix}$$

What are the eigenvalues? Hint: Don't multiply everything out!

Characteristic polynomials

3×3 matrices

Find the characteristic polynomial of the following matrix.

$$\begin{pmatrix} 7 & 0 & 3 \\ -3 & 2 & -3 \\ 4 & 2 & 0 \end{pmatrix}$$

Answer: $-\lambda^3 + 9\lambda^2 - 8\lambda$

What are the eigenvalues?

Characteristic polynomials

3×3 matrices

Find the characteristic polynomial of the rabbit population matrix.

$$\begin{pmatrix} 0 & 6 & 8 \\ \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & 0 \end{pmatrix}$$

Answer:

$$-\lambda^3 + 3\lambda + 2$$

What are the eigenvalues?

Hint: We already know one eigenvalue! Polynomial long division $\rightsquigarrow$

$$(\lambda - 2)(-\lambda^2 - 2\lambda - 1)$$

Don't really need long division: the first and last terms of the quadratic are easy to find; can guess and check the other term.

Characteristic polynomials

3×3 matrices

Find the characteristic polynomial and eigenvalues.

$$\begin{pmatrix} 5 & -2 & 2 \\ 4 & -3 & 4 \\ 4 & -6 & 7 \end{pmatrix}$$

Characteristic polynomial: $-\lambda^3 + 9\lambda^2 - 23\lambda + 15$

This time we don't know any of the roots! We can use the rational root theorem: any integer root of a polynomial with leading coefficient ± 1 divides the constant term.

So we plug in ± 1, ± 3, ± 5, ± 15 into the polynomial and hope for the best. Luckily we find that 1, 3, and 5 are all roots, so we found all the eigenvalues!

If we were less lucky and found only one eigenvalue, we could again use long division like on the last slide.

Eigenvalues

Triangular matrices

Fact. The eigenvalues of a triangular matrix are the diagonal entries.

Why?

Warning! You cannot find eigenvalues by row reducing and then using this fact. You need to work with the original matrix. Finding eigenspaces involves row reducing $A - \lambda I$, but there is no row reduction in finding eigenvalues.

Characteristic polynomials, trace, and determinant

The trace of a matrix is the sum of the diagonal entries.

The characteristic polynomial always looks like:

$$(-1)^n\lambda^n + (-1)^{n-1}\boxed{\text{trace(A)}}\lambda^{n-1} + \boxed{???}\lambda^{n-2} + \cdots \boxed{???}\lambda + \boxed{\det\text{(A)}}$$

So for a 2×2 matrix:

$$\lambda^2 - \operatorname{trace}(A)\lambda + \det(A)$$

And for a 3×3 matrix:

$$-\lambda^3 + \operatorname{trace}(A)\lambda^2 - \boxed{???}\lambda + \det(A)$$

Characteristic polynomials, trace, and determinant

The trace of a matrix is the sum of the diagonal entries.

The characteristic polynomial always looks like:

$$(-1)^n\lambda^n + (-1)^{n-1}\boxed{\text{trace(A)}}\lambda^{n-1} + \boxed{???}\lambda^{n-2} + \cdots \boxed{???}\lambda + \boxed{\det(\text{A})}$$

Consequence 1. The constant term is zero $\Leftrightarrow$ A is not invertible

Consequence 2. The determinant is the product of the eigenvalues.

Algebraic multiplicity

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

Example. Find the algebraic multiplicities of the eigenvalues for

$$\begin{pmatrix} 5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 5 \end{pmatrix}$$

Fact. The sum of the algebraic multiplicities of the (real) eigenvalues of an $n \times n$ matrix is at most n.

Summary of Section 5.2

- The characteristic polynomial of A is $\det(A - \lambda I)$
- The roots of the characteristic polynomial for A are the eigenvalues
- Techniques for 3×3 matrices:
  - Don't multiply out if there is a common factor
  - If there is no constant term then factor out λ
  - If the matrix is triangular, the eigenvalues are the diagonal entries
  - Guess one eigenvalue using the rational root theorem, reverse engineer the rest (or use long division)
  - Use the geometry to determine an eigenvalue
- Given an square matrix A:
  - The eigenvalues are the solutions to $\det(A - \lambda I) = 0$
  - Each λ_i-eigenspace is the solution to $(A - \lambda_i I)x = 0$

Typical Exam Questions 5.2

- True or false: Every $n \times n$ matrix has an eigenvalue.
- True or false: Every $n \times n$ matrix has n distinct eigenvalues.
- True or false: The nullity of $A - \lambda I$ is the dimension of the λ-eigenspace.
- What are the eigenvalues for the standard matrix for a reflection?
- What are the eigenvalues and eigenvectors for the $n \times n$ zero matrix?
- Find the eigenvalues of the following matrix.

$$\begin{pmatrix} 1 & 2 & 1 \\ 0 & -5 & 0 \\ 1 & 8 & 0 \end{pmatrix}$$

- Find the eigenvalues of the following matrix.

$$\begin{pmatrix} 5 & 6 & 2 \\ 0 & -1 & -8 \\ 1 & 0 & 2 \end{pmatrix}$$

Hint: All of the eigenvalues are integers. Use the rational root theorem to guess one of the eigenvalues, and then factor out a linear term.

Eigenvalues in Structural Engineering

Watch this video about the Tacoma Narrows bridge. ▸ Watch

Here are some toy models. ▸ Check it out

The masses move the most at their natural frequencies ω. To find those, use the spring equation: $mx'' = -kx \quad \rightsquigarrow \quad \sin(\omega t)$.

With 3 springs and 2 equal masses, we get:

$$\begin{aligned} mx_1'' &= -kx_1 + k(x_2 - x_1) \\ mx_2'' &= -kx_2 + k(x_1 - x_2) \end{aligned}$$

Guess a solution $x_1(t) = A_1(\cos(\omega t) + i\sin(\omega t))$ and similar for x_2. Finding ω reduces to finding eigenvalues of $\begin{pmatrix} -2k & k \\ k & -2k \end{pmatrix}$.

Eigenvectors: $(1,1)$ & $(1,-1)$ (in/out of phase) ▸ Details

Section 5.4

Diagonalization

Section 5.4 Outline

- Diagonalization
- Using diagonalization to take powers
- Algebraic versus geometric dimension

We understand diagonal matrices

We completely understand what diagonal matrices do to $\mathbb{R}^n$. For example:

$$\begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$$

We have seen that it is useful to take powers of matrices: for instance in computing rabbit populations.

If A is diagonal, powers of A are easy to compute. For example:

$$\begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}^{10} =$$

Powers of matrices that are similar to diagonal ones

What if A is not diagonal? Suppose want to understand the matrix

$$A = \begin{pmatrix} 5/4 & 3/4 \\ 3/4 & 5/4 \end{pmatrix}$$

geometrically? Or take it's 10th power? What would we do?

What if I give you the following equality:

$$\begin{array}{ccccccc} \begin{pmatrix} 5/4 & 3/4 \\ 3/4 & 5/4 \end{pmatrix} & = & \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} & \begin{pmatrix} 2 & 0 \\ 0 & 1/2 \end{pmatrix} & \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}^{-1} \\ A & = & C & D & C^{-1} \end{array}$$

This is called diagonalization.

How does this help us understand A? Or find A^{10}?

Powers of matrices that are similar to diagonal ones

What if I give you the following equality:

$$\begin{array}{ccccccc} \begin{pmatrix} 5/4 & 3/4 \\ 3/4 & 5/4 \end{pmatrix} & = & \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} & \begin{pmatrix} 2 & 0 \\ 0 & 1/2 \end{pmatrix} & \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}^{-1} \\ A & = & C & D & C^{-1} \end{array}$$

This is called diagonalization.

How does this help us understand A? Or find A^{10}? ▸ Demo

Diagonalization

Suppose A is $n \times n$. We say that A is diagonalizable if we can write:

$$A = CDC^{-1} \qquad\qquad D = \text{diagonal}$$

We say that A is similar to D.

How does this factorization of A help describe what A does to $\mathbb{R}^n$? How does this help us take powers of A?

Understanding the rabbit example: since 2 is the largest eigenvalue, (almost) all other vectors get pulled towards that eigenvector. Compare with the example from the last slide.

Diagonalization

The recipe

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

In this case

$$
\begin{aligned}
A &= \begin{pmatrix} | & | & & | \\ v_1 & v_2 & \cdots & v_n \\ | & | & & | \end{pmatrix} \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} \begin{pmatrix} | & | & & | \\ v_1 & v_2 & \cdots & v_n \\ | & | & & | \end{pmatrix}^{-1} \\
&= \qquad\qquad C \qquad\qquad\qquad\qquad D \qquad\qquad\qquad\qquad C^{-1}
\end{aligned}
$$

where $v_1, \ldots, v_n$ are linearly independent eigenvectors and $\lambda_1, \ldots, \lambda_n$ are the corresponding eigenvalues, with multiplicity, in order.

Why?

Example

Diagonalize if possible.

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

Example

Diagonalize if possible.

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

Example

Diagonalize if possible.

$$\begin{pmatrix} 3/4 & 1/4 \\ 1/4 & 3/4 \end{pmatrix}$$

▸ Demo

Hint: the eigenvalues are 1 and $1/2$

More Examples

Diagonalize if possible.

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

Hint: the eigenvalues (with multiplicity) are $3, -1, 1$ and $2, 2, 1$

Poll

Poll

Which are diagonalizable?

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

Distinct Eigenvalues

Fact. If A has n distinct eigenvalues, then A is diagonalizable.

Why?

Non-Distinct Eigenvalues

Theorem. Suppose

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

Then

1. $1 \leq d_i \leq a_i$ for all i
2. A is diagonalizable $\Leftrightarrow \Sigma d_i = n$
   $\Leftrightarrow \Sigma a_i = n$ and $d_i = a_i$ for all i

So the recipe for checking diagonalizability is:

- If there are not n eigenvalues with multiplicity, then stop.
- For each eigenvalue with alg. mult. greater than 1, check if the geometric multiplicity is equal to the algebraic multiplicity. If any of them are smaller, the matrix is not diagonalizable.
- Otherwise, the matrix is diagonalizable.

More rabbits

Which ones are diagonalizable?

$$\begin{pmatrix} 0 & 6 & 8 \\ \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & 0 \end{pmatrix} \qquad \begin{pmatrix} 0 & 4 & 4 \\ \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & 0 \end{pmatrix}$$

Hint: the characteristic polynomials are $-\lambda^3 + 3\lambda + 2$ *and* $-\lambda^3 + 2\lambda + 1$ *and both have rational roots.*

Summary of Section 5.4

- A is diagonalizable if $A = CDC^{-1}$ where D is diagonal
- A diagonal matrix stretches along its eigenvectors by the eigenvalues, similar to a diagonal matrix
- If $A = CDC^{-1}$ then $A^k = CD^kC^{-1}$
- A is diagonalizable $\Leftrightarrow$ A has n linearly independent eigenvectors $\Leftrightarrow$ the sum of the geometric dimensions of the eigenspaces in n
- If A has n distinct eigenvalues it is diagonalizable

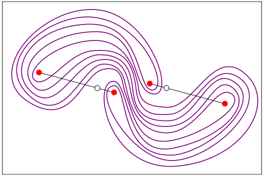
Typical Exam Questions 5.4

- True or False. If A is a 3×3 matrix with eigenvalues 0, 1, and 2, then A is diagonalizable.
- True or False. It is possible for an eigenspace to be 0-dimensional.
- True or False. Diagonalizable matrices are invertible.
- True or False. Diagonal matrices are diagonalizable.
- True or False. Upper triangular matrices are diagonalizable.
- Find the 100th power of $\begin{pmatrix} 2 & 6 \\ 0 & -1 \end{pmatrix}$.
- For each of the following matrices, diagonalize or show they are not diagonalizable:

$$\begin{pmatrix} 2 & 0 & 0 \\ 1 & 2 & 1 \\ -1 & 0 & 1 \end{pmatrix} \qquad \begin{pmatrix} 2 & 4 & 6 \\ 0 & 2 & 2 \\ 0 & 0 & 4 \end{pmatrix}$$

Taffy pullers

How efficient is this taffy puller?



If you run the taffy puller, the taffy starts to look like the shape on the right. Every rotation of the machine changes the number of strands of taffy by a matrix:

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

The largest eigenvalue λ of this matrix describes the efficiency of the taffy puller. With every rotation, the number of strands multiplies by λ.

Section 5.5

Complex Eigenvalues

Outline of Section 5.5

- Rotation matrices have no eigenvectors
- Crash course in complex numbers
- Finding complex eigenvectors and eigenvalues
- Complex eigenvalues correspond to rotations + dilations

Demo Demo

A matrix without an eigenvector

Recall that rotation matrices like

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

and

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

have no eigenvectors. Why?

Imaginary numbers

Problem. When solving polynomial equations, we often run up against the issue that we can't take the square root of a negative number:

$$x^2 + 1 = 0$$

Solution. Take square roots of negative numbers:

$$x = \pm\sqrt{-1}$$

We usually write $\sqrt{-1}$ as i (for "imaginary"), so $x = \pm i$.

Now try solving these:

$$x^2 + 3 = 0$$

$$x^2 - x + 1 = 0$$

Complex numbers

We can add/multiply (and divide!) complex numbers:

$(2 - 3i) + (-1 + i) =$

$(2 - 3i)(-1 + i) =$

Complex numbers

The complex numbers are the numbers

$$\mathbb{C} = \{a + bi \mid a, b \text{ in } \mathbb{R}\}$$

We can conjugate complex numbers: $\overline{a + bi} = a - bi$

Complex numbers and polynomials

Fundamental theorem of algebra. Every polynomial of degree n has exactly n complex roots (counted with multiplicity).

Fact. If z is a root of a real polynomial then $\overline{z}$ is also a root.

So what are the possibilities for degree 2, 3 polynomials?

What does this have to do with eigenvalues of matrices?

Complex eigenvalues

Say A is a square matrix with real entries.

We can now find complex eigenvectors and eigenvalues.

Fact. If λ is an eigenvalue of A with eigenvector v then $\overline{\lambda}$ is an eigenvalue of A with eigenvector $\overline{v}$.

Why?

Trace and determinant

Now that we have complex eigenvalues, we have the following fact.

Fact. The sum of the eigenvalues of A (with multiplicity) is the trace of A and the product of the eigenvalues of A (with multiplicity) is the determinant.

Indeed, by the fundamental theorem of algebra, the characteristic polynomial factors as:

$$(x_1 - \lambda)(x_2 - \lambda) \cdots (x_n - \lambda).$$

From this we see that the product of the eigenvalues $x_1 x_2 \cdots x_n$ is the constant term, which we said was the determinant, and the sum $x_1 + x_2 + \cdots + x_n$ is $(-1)^{n-1}$ times the λ^{n-1} term, which we said was the trace.

Complex eigenvalues

Find the complex eigenvalues and eigenvectors for

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

Three shortcuts for complex eigenvectors

Suppose we have a 2×2 matrix with complex eigenvalue λ.

(1) We do not need to row reduce $A - \lambda I$ by hand; we know the bottom row will become zero.

(2) Then if the reduced matrix is:

$$A = \begin{pmatrix} x & y \\ 0 & 0 \end{pmatrix}$$

the eigenvector is

$$A = \begin{pmatrix} -y \\ x \end{pmatrix}$$

(3) Also, we get the other eigenvalue/eigenvector pair for free: conjugation.

Complex eigenvalues

Find the complex eigenvalues and eigenvectors for

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

Summary of Section 5.5

- Complex numbers allow us to solve all polynomials completely, and find n eigenvalues for an $n \times n$ matrix, counting multiplicity
- If λ is an eigenvalue with eigenvector v then $\overline{\lambda}$ is an eigenvalue with eigenvector $\overline{v}$

Typical Exam Questions 5.5

- True/False. If v is an eigenvector for A with complex entries then $i \cdot v$ is also an eigenvector for A.
- True/False. If $(i, 1)$ is an eigenvector for A then $(i, -1)$ is also an eigenvector for A.
- If A is a 4×4 matrix with real entries, what are the possibilities for the number of non-real eigenvalues of A?
- Find the eigenvalues and eigenvectors for the following matrices.

$$\begin{pmatrix} 3 & -2 \\ 4 & -1 \end{pmatrix} \qquad \begin{pmatrix} -1 & -4 \\ 1 & -1 \end{pmatrix} \qquad \begin{pmatrix} 1 & 0 & -2 \\ 1 & 3 & 1 \\ 2 & 0 & 1 \end{pmatrix}$$

Section 5.6

Stochastic Matrices (and Google!)

Outline of Section 5.6

- Stochastic matrices and applications
- The steady state of a stochastic matrix
- Important web pages

Stochastic matrices

A stochastic matrix is a non-negative square matrix where all of the columns add up to 1.

Examples:

$$\begin{pmatrix} 1/4 & 3/5 \\ 3/4 & 2/5 \end{pmatrix} \quad \begin{pmatrix} .3 & .4 & .5 \\ .3 & .4 & .3 \\ .4 & .2 & .2 \end{pmatrix} \quad \begin{pmatrix} 1/2 & 1 & 1/2 \\ 1/2 & 0 & 1/4 \\ 0 & 0 & 1/4 \end{pmatrix}$$

Application: Rental Cars (or Redbox...)

Say your car rental company has 3 locations. Make a matrix whose ij entry is the fraction of cars at location j that end up at location i. For example,

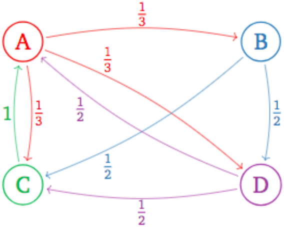
$$\begin{pmatrix} .3 & .4 & .5 \\ .3 & .4 & .3 \\ .4 & .2 & .2 \end{pmatrix}$$

Note the columns sum to 1. Why?

If there are 100 cars at each location on the first day, and every car gets rented, how many cars are at each location on the second day? third day? nth day?

Application: Web pages

Make a matrix whose ij entry is the fraction of (randomly surfing) web surfers at page j that end up at page i. If page i has N links then the ij-entry is either 0 or $1/N$.



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

Which web page seems most important?

Properties of stochastic matrices

Let A be a stochastic matrix.

Fact 1. One of the eigenvalues of A is 1 and all other eigenvalues have absolute value at most 1.

Why?

Positive stochastic matrices

Let A be a positive stochastic matrix, meaning all entries are positive.

Fact 2. The 1-eigenspace of A is 1-dimensional; it has a positive eigenvector.

The unique such eigenvector with entries adding to 1 is called the steady state vector.

Example. If $\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ is a 1-eigenvector, what's the steady state vector?

Example

Find the steady state vector.

$$A = \begin{pmatrix} 1/4 & 3/4 \\ 3/4 & 1/4 \end{pmatrix}$$

More about positive stochastic matrices

Let A be a positive stochastic matrix, meaning all entries are positive.

Fact 3. Under iteration, all nonzero vectors approach a multiple of the steady state vector. The multiple is the sum of the entries of the original vector.

Demo

The last fact tells us how to distribute rental cars, and also tells us the importance of web pages!

Example

To what vector does $A^n \left(\begin{smallmatrix} 1 \\ 9 \end{smallmatrix} \right)$ approach as $n \to \infty$

$$A = \begin{pmatrix} 1/4 & 3/4 \\ 3/4 & 1/4 \end{pmatrix}$$

Application: Rental Cars

The rental car matrix is:

$$\begin{pmatrix} .3 & .4 & .5 \\ .3 & .4 & .3 \\ .4 & .2 & .2 \end{pmatrix}$$

Its steady state vector is:

$$\begin{pmatrix} 7/18 \\ 6/18 \\ 5/18 \end{pmatrix} \approx \begin{pmatrix} .39 \\ .33 \\ .28 \end{pmatrix}$$

Application: Web pages

The web page matrix is:

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

Its steady state vector is approximately

$$\begin{pmatrix} .39 \\ .13 \\ .29 \\ .19 \end{pmatrix}$$

and so the first web page is the most important.

Fine print

There are a couple of problems with the web page matrix as given:

- What happens if there is a web page with no links?
- What if the internet graph is not connected?
- How do you find eigenvectors for a huge matrix?

Here are the solutions:

- Make a column with $1/n$ in each entry (the surfer goes to a new page randomly).
- Let B be the matrix with all entries equal to 1, replace A with

$$.85 * A + .15 * B$$

- Approximate via iteration!

Summary of Section 5.6

- A stochastic matrix is a non-negative square matrix where all of the columns add up to 1.
- Every stochastic matrix has 1 as an eigenvalue, and all other eigenvalues have absolute value at most 1.
- A positive stochastic matrix has 1-dimensional eigenspace and has a positive eigenvector. A positive 1-eigenvector with entries adding to 1 is called a steady state vector.
- For a positive stochastic matrix, all nonzero vectors approach the steady state vector under iteration.
- Steady state vectors tell us the importance of web pages (for example).

Typical Exam Questions 5.6

- Is there a stochastic matrix where the 1-eigenspace has dimension greater than 1?
- Find the steady state vector for this matrix:

$$A = \begin{pmatrix} 1/2 & 1/3 \\ 1/2 & 1/3 \end{pmatrix}$$

  To what vector does $A^n \left(\begin{smallmatrix} 5 \\ 7 \end{smallmatrix} \right)$ approach as $n \to \infty$?
- Find the steady state vector for this matrix:

$$A = \begin{pmatrix} 1/3 & 1/5 & 1/4 \\ 1/3 & 2/5 & 1/2 \\ 1/3 & 2/5 & 1/4 \end{pmatrix}$$

- Make your own internet and see if you can guess which web page is the most important. Check your answer using the method described in this section.

Chapter 6

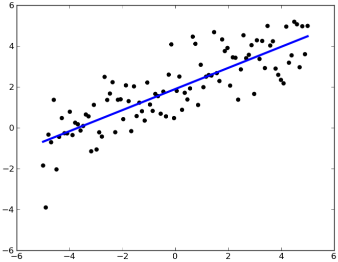
Orthogonality

Where are we?

We have learned to solve $Ax = b$ and $Av = \lambda v$.

We have one more main goal.

What if we can't solve $Ax = b$? How can we solve it as closely as possible?



The answer relies on orthogonality.

Section 6.1

Dot products and Orthogonality

Outline

- Dot products
- Length and distance
- Orthogonality

Dot product

Say $u = (u_1, \ldots, u_n)$ and $v = (v_1, \ldots, v_n)$ are vectors in $\mathbb{R}^n$

$$\begin{aligned} u \cdot v &= \sum_{i=1}^{n} u_i v_i \\ &= u_1 v_1 + \cdots + u_n v_n \\ &= u^T v \end{aligned}$$

Example. Find $(1, 2, 3) \cdot (4, 5, 6)$.

Dot product

Some properties of the dot product

- $u \cdot v = v \cdot u$
- $(u + v) \cdot w = u \cdot w + v \cdot w$
- $(cu) \cdot v = c(u \cdot v)$
- $u \cdot u \geq 0$
- $u \cdot u = 0 \Leftrightarrow u = 0$

Length

Let v be a vector in $\mathbb{R}^n$

$$\begin{aligned}\|v\| &= \sqrt{v \cdot v} \\ &= \text{length of } v\end{aligned}$$

Why? Pythagorean Theorem

Fact. $\|cv\| = |c|\|v\|$

v is a unit vector of $\|v\| = 1$

Problem. Find the unit vector in the direction of $(1, 2, 3, 4)$.

Distance

The distance between v and w is the length of $v - w$ (or $w - v$!).

Problem. Find the distance between $(1, 1, 1)$ and $(1, 4, -3)$.

Orthogonality

Fact. $u \perp v \Leftrightarrow u \cdot v = 0$

Why? Pythagorean theorem again!

$$\begin{aligned} u \perp v &\Leftrightarrow \|u\|^2 + \|v\|^2 = \|u - v\|^2 \\ &\Leftrightarrow u \cdot u + v \cdot v = u \cdot u - 2u \cdot v + v \cdot v \\ &\Leftrightarrow u \cdot v = 0 \end{aligned}$$

Problem. Find a vector in $\mathbb{R}^3$ orthogonal to $(1, 2, 3)$.

Summary of Section 6.1

- $u \cdot v = \sum u_i v_i$
- $u \cdot u = \|u\|^2$ (length of u squared)
- The unit vector in the direction of v is $v/\|v\|$.
- The distance from u to v is $\|u - v\|$
- $u \cdot v = 0 \Leftrightarrow u \perp v$

Section 6.2

Orthogonal complements

Outline of Section 6.2

- Orthogonal complements
- Computing orthogonal complements

Orthogonal complements

W = subspace of $\mathbb{R}^n$
$W^\perp = \{v \text{ in } \mathbb{R}^n \mid v \perp w \text{ for all } w \text{ in } W\}$

Question. What is the orthogonal complement of a line in $\mathbb{R}^3$?
What about the orthogonal complement of a plane in $\mathbb{R}^3$?

▸ Demo ▸ Demo

Orthogonal complements

W = subspace of $\mathbb{R}^n$
$W^\perp = \{v \text{ in } \mathbb{R}^n \mid v \perp w \text{ for all } w \text{ in } W\}$

Facts.

1. $W^\perp$ is a subspace of $\mathbb{R}^n$ (it's a null space!)
2. $(W^\perp)^\perp = W$
3. $\dim W + \dim W^\perp = n$ (rank-nullity theorem!)
4. If $W = \mathrm{Span}\{w_1, \dots, w_k\}$ then
   $W^\perp = \{v \text{ in } \mathbb{R}^n \mid v \perp w_i \text{ for all } i\}$
5. The intersection of W and $W^\perp$ is $\{0\}$.

For items 1 and 3, which linear transformation do we use?

Orthogonal complements

Finding them

Problem. Let $W = \mathrm{Span}\{(1, 1, -1)\}$. Find the equation of the plane $W^{\perp}$.

Find a basis for $W^{\perp}$.

Orthogonal complements

Finding them

Problem. Let $W = \mathrm{Span}\{(1,1,-1),(-1,2,1)\}$. Find a system of equations describing the line $W^\perp$.

Find a basis for $W^\perp$.

Orthogonal complements

Finding them

Recipe. To find (basis for) $W^{\perp}$, find a basis for W, make those vectors the rows of a matrix, and find (a basis for) the null space.

Why? $Ax = 0 \Leftrightarrow x$ is orthogonal to each row of A

Orthogonal complements

Finding them

Recipe. To find (basis for) $W^\perp$, find a basis for W, make those vectors the rows of a matrix, and find (a basis for) the null space.

Why? $Ax = 0 \Leftrightarrow x$ is orthogonal to each row of A

In other words:

Theorem. $A = m \times n$ matrix

$$(\mathrm{Row} A)^\perp = \mathrm{Nul}\, A$$

Geometry $\leftrightarrow$ Algebra

(The row space of A is the span of the rows of A.)

Orthogonal decomposition

Fact. Say W is a subspace of $\mathbb{R}^n$. Then any vector v in $\mathbb{R}^n$ can be written uniquely as

$$v = v_W + v_{W^\perp}$$

where v_W is in W and $v_{W^\perp}$ is in $W^\perp$.

Why?

Demo

Demo

Next time: Find v_W and $v_{W^\perp}$.

Orthogonal decomposition

Fact. Say W is a subspace of $\mathbb{R}^n$. Then any vector v in $\mathbb{R}^n$ can be written uniquely as

$$v = v_W + v_{W^\perp}$$

where v_W is in W and $v_{W^\perp}$ is in $W^\perp$.

Why? Say that $w_1 + w_1' = w_2 + w_2'$ where w_1 and w_2 are in W and w_1' and w_2' are in $W^\perp$. Then $w_1 - w_2 = w_2' - w_1'$. But the former is in W and the latter is in $W^\perp$, so they must both be equal to 0.

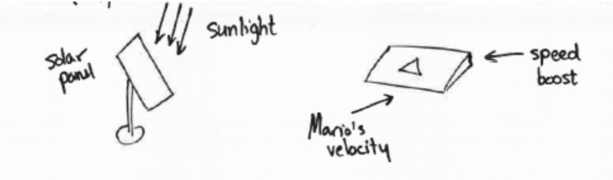
Demo

Demo

Next time: Find v_W and $v_{W^\perp}$.

Orthogonal Projections

Many applications, including:



Summary of Section 6.2

- $W^\perp = \{v \text{ in } \mathbb{R}^n \mid v \perp w \text{ for all } w \text{ in } W\}$
- Facts:
  1. $W^\perp$ is a subspace of $\mathbb{R}^n$
  2. $(W^\perp)^\perp = W$
  3. $\dim W + \dim W^\perp = n$
  4. If $W = \mathrm{Span}\{w_1, \ldots, w_k\}$ then
     $W^\perp = \{v \text{ in } \mathbb{R}^n \mid v \perp w_i \text{ for all } i\}$
  5. The intersection of W and $W^\perp$ is $\{0\}$.
- To find $W^\perp$, find a basis for W, make those vectors the rows of a matrix, and find the null space.
- Every vector v can be written uniquely as $v = v_W + v_{W^\perp}$ with v_W in W and $v_{W^\perp}$ in $W^\perp$

Typical Exam Questions 6.2

- What is the dimension of $W^\perp$ if W is a line in $\mathbb{R}^{10}$?
- What is $W^\perp$ if W is the line $y = mx$ in $\mathbb{R}^2$?
- If W is the x-axis in $\mathbb{R}^2$, and $v = \begin{pmatrix} 7 \\ -3 \end{pmatrix}$, write v as $v_W + v_{W^\perp}$.
- If W is the line $y = x$ in $\mathbb{R}^2$, and $v = \begin{pmatrix} 7 \\ -3 \end{pmatrix}$, write v as $v_W + v_{W^\perp}$.
- Find a basis for the orthogonal complement of the line through $\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ in $\mathbb{R}^3$.
- Find a basis for the orthogonal complement of the line through $\begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix}$ in $\mathbb{R}^4$.
- What is the orthogonal complement of x_1x_2-plane in $\mathbb{R}^4$?

Section 6.3

Orthogonal projection

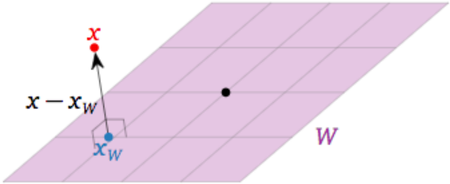
Outline of Section 6.3

- Orthogonal projections and distance
- A formula for projecting onto any subspace
- A special formula for projecting onto a line
- Matrices for projections
- Properties of projections

Orthogonal Projections

Let b be a vector in $\mathbb{R}^n$ and W a subspace of $\mathbb{R}^n$.

The orthogonal projection of b onto W the vector obtained by drawing a line segment from b to W that is perpendicular to W.



Fact. The following three things are all the same:

- The orthogonal projection of b onto W
- The vector b_W (the W-part of b) algebra!
- The closest vector in W to b geometry!

Orthogonal Projections

Theorem. Let $W = \mathrm{Col}(A)$. For any vector b in $\mathbb{R}^n$, the equation

$$A^T A x = A^T b$$

is consistent and the orthogonal projection b_W is equal to Ax where x is any solution.

Orthogonal Projections

Theorem. Let $W = \mathrm{Col}(A)$. For any vector b in $\mathbb{R}^n$, the equation

$$A^T Ax = A^T b$$

is consistent and the orthogonal projection b_W is equal to Ax where x is any solution.

Why? Choose $\widehat{x}$ so that $A\widehat{x} = b_W$. We know $b - b_W = b - A\widehat{x}$ is in $W^\perp = \mathrm{Nul}(A^T)$ and so

$$0 = A^T(b - A\widehat{x}) = A^T b - A^T A\widehat{x}$$

$$\rightsquigarrow A^T A\widehat{x} = A^T b$$

Orthogonal Projections

Theorem. Let $W = \mathrm{Col}(A)$. For any vector b in $\mathbb{R}^n$, the equation

$$A^T A x = A^T b$$

is consistent and the orthogonal projection b_W is equal to Ax where x is any solution.

What does the theorem give when $W = \mathrm{Span}\{u\}$ is a line?

Orthogonal Projection onto a line

Special case. Let $L = \mathrm{Span}\{u\}$. For any vector b in $\mathbb{R}^n$ we have:

$$b_L = \frac{u \cdot b}{u \cdot u} u$$

Find b_L and $b_{L^\perp}$ if $b = \begin{pmatrix} -2 \\ -3 \\ -1 \end{pmatrix}$ and $u = \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}$.

Orthogonal Projections

Theorem. Let $W = \mathrm{Col}(A)$. For any vector b in $\mathbb{R}^n$, the equation

$$A^T Ax = A^T b$$

is consistent and the orthogonal projection b_W is equal to Ax where x is any solution.

Example. Find b_W if $b = \begin{pmatrix} 6 \\ 5 \\ 4 \end{pmatrix}$, $W = \mathrm{Span}\left\{ \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \right\}$

Steps. Find $A^T A$ and $A^T b$, then solve for x, then compute Ax.

Question. How far is b from W?

Orthogonal Projections

Example. Find b_W if $b = \begin{pmatrix} 6 \\ 5 \\ 4 \end{pmatrix}$, $W = \mathrm{Span}\left\{ \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \right\}$

Steps. Find $A^T A$ and $A^T b$, then solve for x, then compute Ax.

Question. How far is b from W?

Orthogonal Projections

Theorem. Let $W = \mathrm{Col}(A)$. For any vector b in $\mathbb{R}^n$, the equation

$$A^T A x = A^T b$$

is consistent and the orthogonal projection b_W is equal to Ax where x is any solution.

Special case. If the columns of A are independent then $A^T A$ is invertible, and so

$$b_W = A(A^T A)^{-1} A^T b.$$

Why? The x we find tells us which linear combination of the columns of A gives us b_W. If the columns of A are independent, there's only one linear combination.

Matrices for projections

Fact. If the columns of A are independent and $W = \mathrm{Col}(A)$ and $T : \mathbb{R}^n \to \mathbb{R}^n$ is orthogonal projection onto W then the standard matrix for T is:

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

Why?

Example. Find the standard matrix for orthogonal projection of $\mathbb{R}^3$ onto $W = \mathrm{Span}\left\{ \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \right\}$

Example. Find the standard matrix for orthogonal projection of $\mathbb{R}^3$ onto $W = \mathrm{Span}\left\{ \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \right\}$

Projections as linear transformations

Let W be a subspace of $\mathbb{R}^n$ and let $T : \mathbb{R}^n \to \mathbb{R}^n$ be the function given by $T(b) = b_W$ (orthogonal projection). Then

- T is a linear transformation
- $T(b) = b$ if and only if b is in W
- $T(b) = 0$ if and only if b is in $W^\perp$
- $T \circ T = T$
- The range of T is W

Properties of projection matrices

Let W be a subspace of $\mathbb{R}^n$ and let $T : \mathbb{R}^n \to \mathbb{R}^n$ be the function given by $T(b) = b_W$ (orthogonal projection). Let A be the standard matrix for T. Then

- The 1–eigenspace of A is W (unless $W = 0$)
- The 0–eigenspace of A is $W^\perp$ (unless $W = \mathbb{R}^n$)
- $A^2 = A$
- $\mathrm{Col}(A) = W$
- $\mathrm{Nul}(A) = W^\perp$
- A is diagonalizable; its diagonal matrix has m 1's & $n - m$ 0's where $m = \dim W$ (this gives another way to find A)

You can check these properties for the matrix in the last example. It would be very hard to prove these facts without any theory. But they are all easy once you know about linear transformations!

Summary of Section 6.3

- The orthogonal projection of b onto W is b_W
- b_W is the closest point in W to b.
- The distance from b to W is $\|b_{W^\perp}\|$.
- Theorem. Let $W = \mathrm{Col}(A)$. For any b, the equation $A^TAx = A^Tb$ is consistent and b_W is equal to Ax where x is any solution.
- Special case. If $L = \mathrm{Span}\{u\}$ then $b_L = \frac{u \cdot b}{u \cdot u} u$
- Special case. If the columns of A are independent then A^TA is invertible, and so $b_W = A(A^TA)^{-1}A^Tb$
- When the columns of A are independent, the standard matrix for orthogonal projection to $\mathrm{Col}(A)$ is $A(A^TA)^{-1}A^T$
- Let W be a subspace of $\mathbb{R}^n$ and let $T : \mathbb{R}^n \to \mathbb{R}^n$ be the function given by $T(b) = b_W$. Then
  - T is a linear transformation
  - etc.
- If A is the standard matrix then
  - The 1–eigenspace of A is W (unless $W = 0$)
  - etc.

Typical Exam Questions 6.3

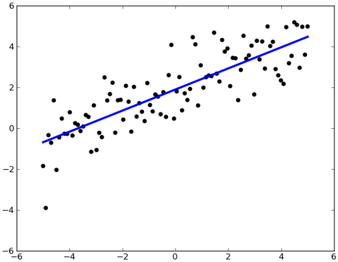
- True/false. The solution to $A^T Ax = A^T b$ is the point in $\mathrm{Col}(A)$ that is closest to b.
- True/false. If v and w are both solutions to $A^T Ax = A^T b$ then $v - w$ is in the null space of A.
- True/false. If A has two equal columns then $A^T Ax = A^T b$ has infinitely many solutions for every b.
- Find b_L and $b_{L^\perp}$ if $b = (1, 2, 3)$ and L is the span of $(1, 2, 1)$.
- Find b_W if $b = (1, 2, 3)$ and W is the span of $(1, 2, 1)$ and $(1, 0, 1)$. Find the distance from b to W.
- Find the matrix A for orthogonal projection to the span of $(1, 2, 1)$ and $(1, 0, 1)$. What are the eigenvalues of A? What is A^{100}?

Section 6.5

Least Squares Problems

Least Squares problems

What if we can't solve $Ax = b$? How can we solve it as closely as possible?



To solve $Ax = b$ as closely as possible, we orthogonally project b onto $\mathrm{Col}(A)$; call the result $\widehat{b}$. Then solve $Ax = \widehat{b}$. This is the *least squares solution* to $Ax = b$.

Outline of Section 6.5

- The method of least squares
- Application to best fit lines/planes
- Application to best fit curves

Least squares solutions

$A = m \times n$ matrix.

A least squares solution to $Ax = b$ is an $\widehat{x}$ in $\mathbb{R}^n$ so that $A\widehat{x}$ is as close as possible to b.

The error is $\|A\widehat{x} - b\|$.

Demo

Least squares solutions

A least squares solution to $Ax = b$ is an $\widehat{x}$ in $\mathbb{R}^n$ so that $A\widehat{x}$ is as close as possible to b.

The error is $\|A\widehat{x} - b\|$.

Theorem. The least squares solutions to $Ax = b$ are the solutions to

$$(A^T A)x = (A^T b)$$

So this is just like what we did before when we were finding the projection of b onto $\mathrm{Col}(A)$. But now we just solve and don't (necessarily) multiply the solution by A.

Least squares solutions

Example

Theorem. The least squares solutions to $Ax = b$ are the solutions to

$$(A^T A)x = (A^T b)$$

Find the least squares solutions to $Ax = b$ for this A and b:

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 1 \\ 2 & 1 \end{pmatrix} \qquad b = \begin{pmatrix} 6 \\ 0 \\ 0 \end{pmatrix}$$

What is the error?

Least squares solutions

Example

Formula: $(A^T A)x = (A^T b)$

Find the least squares solution/error to $Ax = b$:

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 1 \\ 2 & 1 \end{pmatrix} \qquad b = \begin{pmatrix} 6 \\ 0 \\ 0 \end{pmatrix}$$

Least squares solutions

Theorem. Let A be an $m \times n$ matrix. The following are equivalent:

1. $Ax = b$ has a unique least squares solution for all b in $\mathbb{R}^n$
2. The columns of A are linearly independent
3. $A^T A$ is invertible

In this case the least squares solution is $(A^T A)^{-1}(A^T b)$.

Application

Best fit lines

Problem. Find the best-fit line through $(0, 6)$, $(1, 0)$, and $(2, 0)$.

Demo

Best fit lines

Poll

What does the best fit line minimize?

1. the sum of the squares of the distances from the data points to the line
2. the sum of the squares of the vertical distances from the data points to the line
3. the sum of the squares of the horizontal distances from the data points to the line
4. the maximal distance from the data points to the line

Least Squares Problems

More applications

Determine the least squares problem $Ax = b$ to find the best parabola $y = Cx^2 + Dx + E$ for the points:

$$(0,0),(2,0),(3,0),(0,1)$$

Demo

Least Squares Problems

More applications

Determine the least squares problem $Ax = b$ to find the best fit ellipse $Cx^2 + Dxy + Ey^2 + Fx + Gy + H = 0$ for the points:

$$(0, 0), (2, 0), (3, 0), (0, 1)$$

Gauss invented the method of least squares to predict the orbit of the asteroid Ceres as it passed behind the sun in 1801.

Least Squares Problems

Best fit plane

Determine the least squares problem $Ax = b$ to find the best fit linear function $f(x, y) = Cx + Dy + E$

x	y	$f(x,y)$
1	0	0
0	1	1
-1	0	3
0	-1	4

Summary of Section 6.5

- A least squares solution to $Ax = b$ is an $\widehat{x}$ in $\mathbb{R}^n$ so that $A\widehat{x}$ is as close as possible to b.
- The error is $\|A\widehat{x} - b\|$.
- The least squares solutions to $Ax = b$ are the solutions to $(A^T A)x = (A^T b)$.
- To find a best fit line/parabola/etc. write the general form of the line/parabola/etc. with unknown coefficients and plug in the given points to get a system of linear equations in the unknown coefficients.

Typical Exam Questions 6.5

- Find the best fit line through $(1, 0)$, $(2, 1)$, and $(3, 1)$. What is the error?
- Find the best fit parabola through $(1, 0)$, $(2, 1)$, $(3, 1)$, and $(3, 0)$. What is the error?
- True/false. For every set of three points in $\mathbb{R}^2$ there is a unique best fit line.
- True/false. If $\widehat{x}$ is the least squares solution to $Ax = b$ for an $m \times n$ matrix A, then $\widehat{x}$ is the closest point in $\mathbb{R}^n$ to b.
- True/false. If $\widehat{x}$ and $\widehat{y}$ are both least squares solutions to $Ax = b$ then $\widehat{x} - \widehat{y}$ is in the null space of A.