MATH 257

Linear Algebra with Computational Applications

ILLINOIS

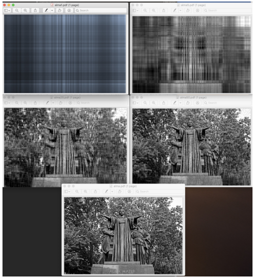
Department of Mathematics

Course description

- *Linear algebra is used in most sciences and engineering areas, because it allows modeling many natural phenomena, and efficiently computing with such models.*
- *Mathematics is the language of science and engineering, but coding is believing it.*
- *But what is data science, if not linear algebra persevering?*

This is a first course in linear algebra. This covers basic definitions and algorithms of the subject needed in the higher level (engineering, science and economics) courses and more sophisticated mathematical techniques such as the Singular Value Decomposition.

In this course you learn the mathematical theory and how to implement it in Python. You will discover many of the striking modern applications of linear algebra, such as Google's PageRank algorithm, image and audio compression schemes such as JPEG and MP3, automatic face recognition and other data science and machine learning algorithms.



Three disclaimers

- ⚠ This is not a course that only teaches you how to compute stuff. Computer will always be faster. Modern applications of linear algebra require a sophisticated understanding of theory and methods of linear algebra, and learning these is the purpose of this course. Some of it might look like *abstract* linear algebra. However, through the applications we cover in the labs, you realizes that this indeed is *applied* linear algebra.
- ⚠ If you already know some linear algebra, this course might look easy at the beginning. Don't be fooled into thinking it will stay like that. Even the material familiar to you will be covered in more depth here. Furthermore, the exams will require a deeper understanding of the concepts you already know something about. So it is a good idea to take this course seriously from the beginning.
- ⚠ MATH 257 is a new course. A significant part of the material for this new course is - not surprisingly - new. So if you find a typo or an error in any part of this course, please let us know by sending an email to the instructors. We appreciate your help, and are also happy to hear any further comments or suggestions. Thank you!

Module videos Module videos are available at

- https://mediaspace.illinois.edu/playlist/0_w39larp3/

Further resources

Books

- Philip N. Klein, *Coding the Matrix: Linear Algebra through Applications to Computer Science*, first edition, Newtonian Press
- Feryal Alayont, Steven Schlicker, *Linear Algebra and Applications:An Inquiry-Based Approach*, scholarworks.gvsu.edu/books/21/
- David Cherney, Tom Denton, Rohit Thomas, Andrew Waldron, *Linear Algebra*, www.math.ucdavis.edu/$\sim$ linear/
- Gilbert Strang, *Linear Algebra and its Applications*, fourth edition, Cengage.

Videos

- *Essence of Linear Algebra* by 3Blue1Brown, on Youtube (highly recommended!)
- MIT lectures by Gilbert Strang, MIT Open Courseware
- *Coding the Matrix* videos by Philip Klein, on Youtube

Table of Contents.

Table of Contents (ctd).

LINEAR ALGEBRA

Introduction to Linear Systems

ILLINOIS

Department of Mathematics

Definition. A **linear equation** is a equation of the form

$$a_1x_1 + \ldots + a_nx_n = b$$

where $a_1, \ldots, a_n, b$ are numbers and $x_1, \ldots, x_n$ are variables.

Example. Which of the following equations are linear equations (or can be rearranged to become linear equations)?

Solution.

$4x_1 - 5x_2 + 2 = x_1$	$3x_1 - 5x_2 = -2$	Linear
$x_2 = 2(\sqrt{6} - x_1) + x_3$	$2x_1 + x_2 - x_3 = 2\sqrt{6}$	Linear
$4x_1 - 6x_2 = x_1x_2$	$4x_1 - 6x_2 = x_1x_2$	Not linear
$x_2 = 2\sqrt{x_1} - 7$	$x_2 = 2\sqrt{x_1} - 7$	Not linear

Definition. A **linear system** is a collection of one or more linear equations involving the same set of variables, say, $x_1, x_2, \ldots, x_n$.
A **solution** of a linear system is a list $(s_1, s_2, \ldots, s_n)$ of numbers that makes each equation in the system true when the values $s_1, s_2, \ldots, s_n$ are substituted for $x_1, x_2, \ldots, x_n$, respectively.

Example. Two equations in two variables:

$$x_1 + x_2 = 1 \tag{I}$$
$$-x_1 + x_2 = 0. \tag{II}$$

What is a solution for this system of linear equations?

Solution.

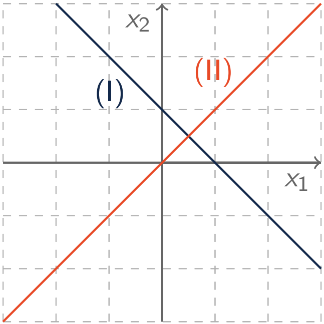
Add (I) + (II):

$$2x_2 = 1 \qquad (\star)$$

Plug into (I):

$$x_1 + \frac{1}{2} = 1 \quad \Rightarrow \quad x_1 = \frac{1}{2}$$

Thus $(x_1, x_2) = (\frac{1}{2}, \frac{1}{2})$ is the only solution.



Example. Does every system of linear equation have a solution?

$$x_1 - 2x_2 = -3 \tag{III}$$
$$2x_1 - 4x_2 = 8. \tag{IV}$$

Solution.

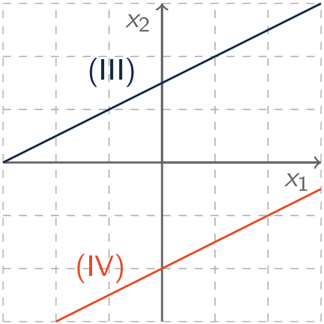
Multiply (III) by 2:

$$2x_1 - 4x_2 = -6. \tag{†}$$

Subtract (†) from (IV):

$$0 = 14.$$

The equation $0 = 14$ is always false, so no solutions exist.



Example. How many solutions are there to the following system?

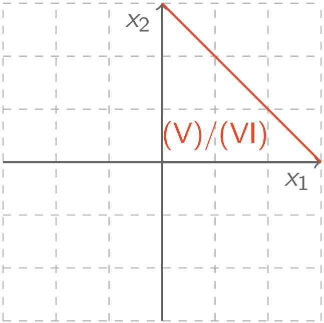
$$x_1 + x_2 = 3 \tag{V}$$
$$-2x_1 - 2x_2 = -6 \tag{VI}$$

Solution.

Multiply (V) by -2:

$$-2x_1 - 2x_2 = -6$$

(V) and (VI) has the same solutions. *Any* value of x_1 works, just set $x_2 := 3 - x_1$. Equation has infinitely many solutions.



Theorem 1. *A linear system has either*

one unique solution *or* **no solution** *or* **infinitely many solutions**.

Definition. The **solution set** of a linear system is the set of all solutions of the linear system. Two linear systems are **equivalent** if they have the same solution set.

The general strategy is to replace one system with an equivalent system that is easier to solve.

Example. Consider

$$\begin{aligned} x_1 - 3x_2 &= 1 && \text{(VII)} \\ -x_1 + 5x_2 &= 3 && \text{(VIII)} \end{aligned}$$

Transform this linear system into another easier equivalent system.

Solution.

Add (VII) to (VIII):

$$\begin{aligned} x_1 - 3x_2 &= 1 && \text{(VII)} \\ 2x_2 &= 4 && \text{(IX)} \end{aligned}$$

Thus $x_2 = 2$ and $x_1 = 7$.

LINEAR ALGEBRA

Matrices and Linear Systems

ILLINOIS

Department of Mathematics

Definition. An $m \times n$ **matrix** is a rectangular array of numbers with m rows and n columns.

Example. Let's give a few examples.

Solution.

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \begin{bmatrix} 1 & 2 & 3 \\ 2 & 1 & 4 \end{bmatrix}, \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}, \begin{bmatrix} 1+\sqrt{5} \end{bmatrix}, \begin{bmatrix} 1 & -2 & 3 & 4 \\ 5 & -6 & 7 & 8 \\ -9 & -10 & 11 & 12 \end{bmatrix}$$

In terms of the entries of A:

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \qquad a_{ij} \text{ is in the } i\text{th row and } j\text{th column}$$

Definition. For a linear system, we define the **coefficient** and **augmented** matrix as follows:

linear system

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{aligned}$$

coefficient matrix

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

augmented matrix

$$\left[\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array}\right]$$

Example. Determine the coefficient matrix and augmented matrix of the linear system

$$\begin{aligned} x_1 - 3x_2 &= 1 \\ -x_1 + 5x_2 &= 3. \end{aligned}$$

Solution.

Coefficient matrix: $\begin{bmatrix} 1 & -3 \\ -1 & 5 \end{bmatrix}$

Augmented matrix: $\left[\begin{array}{cc|c} 1 & -3 & 1 \\ -1 & 5 & 3 \end{array}\right]$

Definition. An **elementary row operation** is one of the following

(Replacement) Add a multiple of one row to another row: $R_i \to R_i + cR_j$, where $i \neq j$,

(Interchange) Interchange two rows: $R_i \leftrightarrow R_j$,

(Scaling) Multiply all entries in a row by a nonzero constant: $R_i \to cR_i$, where $c \neq 0$.

Example. Give several examples of elementary row operations.

Solution.

Replacement:

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \xrightarrow[R_2 \to R_2 + 3R_1]{} \begin{bmatrix} 1 & 0 \\ 3 & 1 \end{bmatrix}.$$

Interchange:

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 2 & 3 \end{bmatrix} \xrightarrow[R_1 \leftrightarrow R_3]{} \begin{bmatrix} 2 & 3 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}$$

Scaling:

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \xrightarrow[R_2 \to 3R_2]{} \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}.$$

Example. Consider the elementary row operation $R_3 \to R_3 + 3R_1$. Is there an elementary row operation that reverse this row operation?

Solution.

Yes. The operation $R_3 \to R_3 - 3R_1$ reverses the row operations $R_3 \to R_3 + 3R_1$. For example,

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \underset{R_3 \to R_3 + 3R_1}{\to} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 3 & 0 & 1 \end{bmatrix} \underset{R_3 \to R_3 - 3R_1}{\to} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Remark. Indeed, every row operation is reversible. We already saw how to reverse the replacement operator. The scaling operatior $R_2 \to cR_2$ is reversed by the scaling operator $R_2 \to \frac{1}{c}R_2$. Row interchange $R_1 \leftrightarrow R_2$ is reversible by performing it twice.

Definition. Two matrices are **row equivalent**, if one matrix can be transformed into the other matrix by a sequence of elementary row operations.

Theorem 2. *If the augmented matrices of two linear systems are row equivalent, then the two systems have the same solution set.*

LINEAR ALGEBRA

Echelon forms of matrices

ILLINOIS

Department of Mathematics

Definition. A matrix is in **echelon form** (or **row echelon form**) if

1. All *nonzero rows* (rows with at least one nonzero element) are above any rows of all zeros.
2. the *leading entry* (the first nonzero number from the left) of a nonzero row is always strictly to the right of the leading entry of the row above it.

Example. Are the following matrices in echelon form? Circle the leading entries.

a) $\begin{bmatrix} 3 & 1 & 2 & 0 & 5 \\ 0 & 2 & 0 & 1 & 4 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

b) $\begin{bmatrix} 0 & 2 & 0 & 1 & 4 \\ 3 & 1 & 2 & 0 & 5 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

c) $\begin{bmatrix} 2 & -2 & 3 \\ 0 & 5 & 0 \\ 0 & 0 & \frac{5}{2} \end{bmatrix}$

d) $\begin{bmatrix} 0 & 1 & 3 \\ 0 & 0 & \sqrt{2} \\ 0 & 0 & 0 \end{bmatrix}$

Definition. A matrix is in **row reduced echelon form** (or: **reduced echelon form**, or: **RREF**) if it is in echelon form and

3. The leading entry in each nonzero row is 1.
4. Each leading entry is the only nonzero entry in its column.

Example. Are the following matrices in reduced echelon form?

a) $\begin{bmatrix} 0 & 1 & 3 & 0 & 0 & 2 & 5 & 0 & 0 & 6 \\ 0 & 0 & 0 & 1 & 0 & 1 & \frac{1}{2} & 0 & 0 & -2 \\ 0 & 0 & 0 & 0 & 1 & -3 & 4 & 0 & 0 & 5 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix}$ b) $\begin{bmatrix} 1 & 0 & -2 & 3 & 2 & -24 \\ 0 & 1 & -2 & 2 & 0 & -7 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix}$

Theorem 3. *Each matrix is row-equivalent to one and only one matrix in reduced echelon form.*

Definition. We say a matrix B is **the reduced echelon form** (or: the **RREF**) of a matrix A if A and B are row-equivalent and B is in reduced echelon form.

Question. Is each matrix also row-equivalent to one and only one matrix in echelon from?

Solution.

No. For example, $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ and $\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$ are row-equivalent and both in echelon form.

Example. Find the reduced echelon form of the matrix

$$\begin{bmatrix} 3 & -9 & 12 & -9 & 6 & 15 \\ 3 & -7 & 8 & -5 & 8 & 9 \end{bmatrix}$$

Solution.

$$\underset{R_2 \to R_2 - R_1}{\rightsquigarrow} \begin{bmatrix} 3 & -9 & 12 & -9 & 6 & 15 \\ 0 & 2 & -4 & 4 & 2 & -6 \end{bmatrix} \underset{\substack{R_1 \to \frac{1}{3} R_1 \\ R_2 \to \frac{1}{2} R_2}}{\rightsquigarrow} \begin{bmatrix} 1 & -3 & 4 & -3 & 2 & 5 \\ 0 & 1 & -2 & 2 & 1 & -3 \end{bmatrix}$$

$$\underset{R_1 \to R_1 + 3R_2}{\rightsquigarrow} \begin{bmatrix} 1 & 0 & -2 & 3 & 5 & -4 \\ 0 & 1 & -2 & 2 & 1 & -3 \end{bmatrix}$$

Definition. A **pivot position** is the position of a leading entry in an echelon form of a matrix. A **pivot column** is a column that contains a pivot position.

Example. Locate the pivot columns of the following matrix.

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

Solution.

$$\underset{R_1 \leftrightarrow R_3}{\rightsquigarrow} \begin{bmatrix} 1 & 4 & 5 & -9 & -7 \\ -1 & -2 & -1 & 3 & 1 \\ 0 & -3 & -6 & 4 & 9 \end{bmatrix} \underset{R_2 \to R_2 + R_1}{\rightsquigarrow} \begin{bmatrix} 1 & 4 & 5 & -9 & -7 \\ 0 & 2 & 4 & -6 & -6 \\ 0 & -3 & -6 & 4 & 9 \end{bmatrix}$$

$$\underset{R_3 \to R_3 + 1.5R_2}{\rightsquigarrow} \begin{bmatrix} 1 & 4 & 5 & -9 & -7 \\ 0 & 2 & 4 & -6 & -6 \\ 0 & 0 & 0 & -5 & 0 \end{bmatrix}$$

Thus columns $1, 2, 4$ of A are the pivot columns of A.

Definition. A **basic variable** (or **pivot variable**) is a variable that corresponds to a pivot column in the coefficient matrix of a linear system. A **free variable** is variable that is *not* a pivot variable.

Example. Consider the following system of linear equations:

$$\left[\begin{array}{ccccc|c} 1 & 6 & 0 & 3 & 0 & 0 \\ 0 & 0 & 1 & -8 & 0 & 5 \\ 0 & 0 & 0 & 0 & 1 & 7 \end{array}\right] \qquad \begin{array}{rrrrrl} x_1 & +6x_2 & & +3x_4 & & = 0 \\ & & x_3 & -8x_4 & & = 5 \\ & & & & x_5 & = 7 \end{array}$$

Determine the basic variables and free variables of this system.

Solution.

The first, third and fifth column of the coefficient matrix are pivot columns.

Thus x_1, x_3, x_5 are basic variables and x_2, x_4 are free variables.

LINEAR ALGEBRA

Gaussian Elimination

ILLINOIS

Department of Mathematics

Goal. Solve linear systems for the pivot variables in terms of the free variables (if any) in the equation.

Algorithm. (Gaussian Elimination) Given a linear system,

(1) Write down the augmented matrix.
(2) Find the reduced echelon form of the matrix.
(3) Write down the equations corresponding to the reduced echelon form.
(4) Express pivot variables in terms of free variables.

C.F. Gauß (1777–1855)

Example. Find the general solution of

$$\begin{array}{rrrrrl} 3x_1 & -7x_2 & +8x_3 & -5x_4 & +8x_5 & = 9 \\ 3x_1 & -9x_2 & +12x_3 & -9x_4 & +6x_5 & = 15 \end{array}$$

Solution.

Augmented matrix: $\left[\begin{array}{ccccc|c} 3 & -7 & 8 & -5 & 8 & 9 \\ 3 & -9 & 12 & -9 & 6 & 15 \end{array}\right] \underset{RREF}{\rightsquigarrow} \left[\begin{array}{ccccc|c} 1 & 0 & -2 & 3 & 5 & -4 \\ 0 & 1 & -2 & 2 & 1 & -3 \end{array}\right]$

Solution.

RREF of the augmented matrix:
$$\left[\begin{array}{ccccc|c} 1 & 0 & -2 & 3 & 5 & -4 \\ 0 & 1 & -2 & 2 & 1 & -3 \end{array}\right]$$

Write down the equations corresponding to the reduced echelon form:

$$\begin{array}{cccccc} x_1 & & -2x_3 & +3x_4 & +5x_5 & = -4 \\ & x_2 & -2x_3 & +2x_4 & +x_5 & = -3 \end{array}$$

Express pivot variables in terms of free variables:

$$\begin{aligned} x_1 &= 2x_3 - 3x_4 - 5x_5 - 4 \\ x_2 &= 2x_3 - 2x_4 - x_5 - 3 \\ x_3 &= \text{ free} \\ x_4 &= \text{ free} \\ x_5 &= \text{ free} \end{aligned}$$

Example. Is the following linear system *consistent*? Is the solution *unique*?

$$\begin{aligned} 3x_2 - 6x_3 + 6x_4 + 4x_5 &= -5 \\ 3x_1 - 7x_2 + 8x_3 - 5x_4 + 8x_5 &= 9 \\ 3x_1 - 9x_2 + 12x_3 - 9x_4 + 6x_5 &= 15 \end{aligned}$$

Solution.

$$\left[\begin{array}{ccccc|c} 0 & 3 & -6 & 6 & 4 & -5 \\ 3 & -7 & 8 & -5 & 8 & 9 \\ 3 & -9 & 12 & -9 & 6 & 15 \end{array}\right] \underset{EF}{\rightsquigarrow} \left[\begin{array}{ccccc|c} 3 & -9 & 12 & -9 & 6 & 15 \\ 0 & 2 & -4 & 4 & 2 & -6 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{array}\right]$$

$$\begin{aligned} 3x_1 - 9x_2 + 12x_3 - 9x_4 + 6x_5 &= 15 \\ 2x_2 - 48x_3 + 4x_4 + 2x_5 &= -6 \\ x_5 &= 4 \end{aligned}$$

Free variables: x_3, x_4. Set $x_3 = x_4 = 0$. Solve the third equation for x_5, the second for x_2 and the first for x_1. This is possible! Thus the system is consistent. Indeed, it has infinitely many solutions (choose different values for x_3, x_4).

Only a row of the form $\left[\begin{array}{ccccc|c} 0 & 0 & 0 & 0 & 0 & b \end{array}\right]$, where $b \neq 0$, would be a problem!

Theorem 4. *A linear system is* **consistent** *if and only if an echelon form of the augmented matrix has* **no** *row of the form* $\left[\begin{array}{ccc|c} 0 & \dots & 0 & b \end{array}\right]$, *where b is nonzero. If a linear system is consistent, then the linear system has*

- *a unique solution (when there are no free variables) or*
- *infinitely many solutions (when there is at least one free variable).*

Example. Consider the linear system whose augmented matrix is $\left[\begin{array}{cc|c} 3 & 4 & -3 \\ 3 & 4 & -3 \\ 6 & 8 & -5 \end{array}\right]$. What can you say about the number of solutions of this system?

Solution.

$$\left[\begin{array}{cc|c} 3 & 4 & -3 \\ 3 & 4 & -3 \\ 6 & 8 & -5 \end{array}\right] \underset{\substack{R_2 \to R_2 - R_1 \\ R_3 \to R_3 - 2R_1}}{\leadsto} \left[\begin{array}{cc|c} 3 & 4 & -3 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{array}\right] \underset{R_2 \leftrightarrow R_3}{\leadsto} \left[\begin{array}{cc|c} 3 & 4 & -3 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{array}\right].$$

No solution, because of the row $\left[\begin{array}{cc|c} 0 & 0 & 1 \end{array}\right]$.

Because $0x_1 + 0x_2 = 1$ has no solution!

LINEAR ALGEBRA

Linear combinations

ILLINOIS

Department of Mathematics

Definition. Consider $m \times n$-matrices $A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$, and $B = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{bmatrix}$.

a) **The sum of** $A + B$ is

$$\begin{bmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \cdots & a_{1n} + b_{1n} \\ a_{21} + b_{21} & a_{22} + b_{22} & \cdots & a_{2n} + b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} + b_{m1} & a_{m2} + b_{m2} & \cdots & a_{mn} + b_{mn} \end{bmatrix}$$

b) **The product** cA for a scalar c is

$$\begin{bmatrix} ca_{11} & ca_{12} & \cdots & ca_{1n} \\ ca_{21} & ca_{22} & \cdots & ca_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ ca_{m1} & ca_{m2} & \cdots & ca_{mn} \end{bmatrix}$$

Example. Calculate

$$\begin{bmatrix} 1 & 0 \\ 5 & 2 \end{bmatrix} + \begin{bmatrix} 2 & 3 \\ 3 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 3 \\ 8 & 3 \end{bmatrix}$$

$$5 \begin{bmatrix} 2 & 1 & 0 \\ 3 & 1 & -1 \end{bmatrix} = \begin{bmatrix} 10 & 5 & 0 \\ 15 & 5 & -5 \end{bmatrix}$$

Warning. Addition is only defined if A has the same number of columns as B and the same number of rows as B.

Definition. A **column vector** is an $m \times 1$-matrix. A **row vector** is a $1 \times n$-matrix.

Example. Give a few examples of column and row vectors.

$$\begin{bmatrix} 1 \\ 2 \end{bmatrix}, \quad \begin{bmatrix} 1 & 2 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ -2 \\ 3 \end{bmatrix}, \quad \begin{bmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \end{bmatrix}$$

Definition. If A is $m \times n$, the **transpose** of A is the $n \times m$ matrix, denoted by A^T, whose columns are formed from the corresponding rows of A. In terms of matrix elements: $(A^T)_{ij} = A_{ji}$.

Example. What is the transpose of $\begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}$?

$$\begin{bmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{bmatrix}$$

Remark. The transpose of a column vector is a row vector and vice versa.

Definition. The **linear combination** of $m \times n$-matrices $A_1, A_2, \dots, A_p$ with **coefficients** $c_1, c_2, \dots, c_p$ is defined as

$$c_1 A_1 + c_2 A_2 + \cdots + c_p A_p.$$

Example. Consider $m \times n$-matrices A_1 and A_2. Give examples of linear combinations of these two matrices.

Solution.

$$3A_1 + 2A_2, \quad A_1 - 2A_2, \quad \tfrac{1}{3}A_1 = \tfrac{1}{3}A_1 + 0A_2, \quad \mathbf{0} = 0A_1 + 0A_2$$

Definition. The span $(A_1, \dots, A_p)$ is defined as the set of all linear combinations of $A_1, \dots, A_p$. *Stated another way:*

$$\operatorname{span}(A_1, \dots, A_p) := \{c_1 A_1 + c_2 A_2 + \cdots + c_p A_p \ : \ c_1, \dots, c_p \text{ scalars}\}.$$

Definition. We denote the set of all column vectors of length m by $\mathbb{R}^m$.

Example. Let $\mathbf{a}_1 = \begin{bmatrix}1\\0\\3\end{bmatrix}$, $\mathbf{a}_2 = \begin{bmatrix}4\\2\\14\end{bmatrix}$ and $\mathbf{b} = \begin{bmatrix}-1\\8\\-5\end{bmatrix}$. Is $\mathbf{b}$ a linear combination of $\mathbf{a}_1, \mathbf{a}_2$?

Solution.

Find x_1, x_2 such that

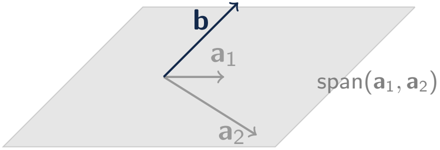
$$x_1 \begin{bmatrix}1\\0\\3\end{bmatrix} + x_2 \begin{bmatrix}4\\2\\14\end{bmatrix} = \begin{bmatrix}-1\\8\\-5\end{bmatrix}$$

$$\begin{aligned} 1x_1 + 4x_2 &= -1 \\ 0x_1 + 2x_2 &= 8 \\ 3x_1 + 14x_2 &= -5. \end{aligned}$$

$$\left[\begin{array}{cc|c} 1 & 4 & -1 \\ 0 & 2 & 8 \\ 3 & 14 & -5 \end{array}\right] \underset{R_3 \to R_3 - 3R_1}{\rightsquigarrow} \left[\begin{array}{cc|c} 1 & 4 & -1 \\ 0 & 2 & 8 \\ 0 & 2 & -2 \end{array}\right]$$

$$\underset{R_3 \to R_3 - R_2}{\rightsquigarrow} \left[\begin{array}{cc|c} 1 & 4 & -1 \\ 0 & 2 & 8 \\ 0 & 0 & -10 \end{array}\right]$$

System is inconsistent. Thus $\mathbf{b}$ is not a linear combination of $\mathbf{a}_1, \mathbf{a}_2$.



Key observation from the previous example: Solving linear systems is the same as finding linear combinations!

Theorem 5. *A vector equation*

$$x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n = \mathbf{b}$$

has the same solution set as the linear system whose augmented matrix is

$$\left[\begin{array}{cccc|c} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n & \mathbf{b} \end{array}\right]$$

In particular, $\mathbf{b}$ *can be generated by a linear combination of* $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$ *if and only if there is a solution to the linear system corresponding to the augmented matrix.*

Notation. We often define a matrix in terms of its columns or its rows:

$$A := \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{bmatrix} \qquad \textit{or} \qquad A := \begin{bmatrix} \mathbf{R}_1 \\ \mathbf{R}_2 \\ \vdots \\ \mathbf{R}_m \end{bmatrix}$$

LINEAR ALGEBRA

Matrix vector multiplication

ILLINOIS

Department of Mathematics

Definition. Let $\mathbf{x}$ be a vector in $\mathbb{R}^n$ and $A = \begin{bmatrix} \mathbf{a}_1 & \dots & \mathbf{a}_n \end{bmatrix}$ an $m \times n$-matrix. We define the product $A\mathbf{x}$ by

$$A\mathbf{x} = x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \ldots + x_n\mathbf{a}_n.$$

Remark.

- $A\mathbf{x}$ is a linear combination of the columns of A using the entries in $\mathbf{x}$ as coefficients.
- $A\mathbf{x}$ is only defined if the number of entries of $\mathbf{x}$ is equal to the number of columns of A.

Example. Consider $A = \begin{bmatrix} 2 & 0 \\ 1 & 1 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 2 \\ 0 & 1 \\ 3 & 5 \end{bmatrix}$, $\mathbf{x} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$. Determine $A\mathbf{x}$ and $B\mathbf{x}$.

Solution.

$$A\mathbf{x} = \begin{bmatrix} 2 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \end{bmatrix} = 2\begin{bmatrix} 2 \\ 1 \end{bmatrix} + 3\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 5 \end{bmatrix}$$

$$B\mathbf{x} = \begin{bmatrix} 1 & 2 \\ 0 & 1 \\ 3 & 5 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \end{bmatrix} = 2\begin{bmatrix} 1 \\ 0 \\ 3 \end{bmatrix} + 3\begin{bmatrix} 2 \\ 1 \\ 5 \end{bmatrix} = \begin{bmatrix} 8 \\ 3 \\ 21 \end{bmatrix}$$

Example. Consider the following vector equation

$$x_1 \begin{bmatrix} 1 \\ 2 \end{bmatrix} + x_2 \begin{bmatrix} 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 0 \\ 2 \end{bmatrix}.$$

Find a 2×2 matrix A such that (x_1, x_2) is a solution to the above equation if and only if

$$A \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 2 \end{bmatrix}?$$

Solution.

$$\text{Take } A = \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix} \qquad \rightsquigarrow Ax = \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ 2 \end{bmatrix} + x_2 \begin{bmatrix} 3 \\ 4 \end{bmatrix}.$$

$$\rightsquigarrow A \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 2 \end{bmatrix} \iff x_1 \begin{bmatrix} 1 \\ 2 \end{bmatrix} + x_2 \begin{bmatrix} 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 0 \\ 2 \end{bmatrix}$$

Theorem 6. *Let $A = \begin{bmatrix} \mathbf{a}_1 & \dots & \mathbf{a}_n \end{bmatrix}$ be an $m \times n$-matrix and $\mathbf{b}$ in $\mathbb{R}^m$. Then the following are equivalent*

- $(x_1, x_2, \dots, x_n)$ *is a solution of the vector equation* $x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n = \mathbf{b}$
- $\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ *is a solution to the matrix equation* $A\mathbf{x} = \mathbf{b}$.
- $(x_1, x_2, \dots, x_n)$ *is a solution of the system with augmented matrix* $\left[\begin{array}{c|c} A & \mathbf{b} \end{array}\right]$

Notation. We will write $A\mathbf{x} = \mathbf{b}$ for the system of equations with augmented matrix $\left[\begin{array}{c|c} A & \mathbf{b} \end{array}\right]$.

A matrix as machine. Let A be a $m \times n$ matrix.

- Input: n-component vector $\mathbf{x} \in \mathbb{R}^n$.
- Output: m-component vector $\mathbf{b} = A\mathbf{x} \in \mathbb{R}^m$.

$$\xrightarrow{\mathbf{x}} \boxed{A} \xrightarrow{A\mathbf{x}}$$

Example. Consider the matrix $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. What does this machine do?

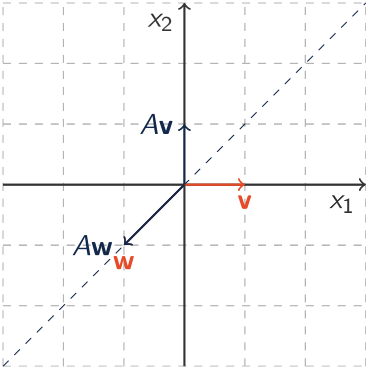
Solution.

Let $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ be our input.

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

So the machine A switches the entries of the vector $\mathbf{x}$.

Geometrically speaking, this machine reflects across the $x_1 = x_2$-line.



Example. Consider the matrix $B = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$. What does this machine do?

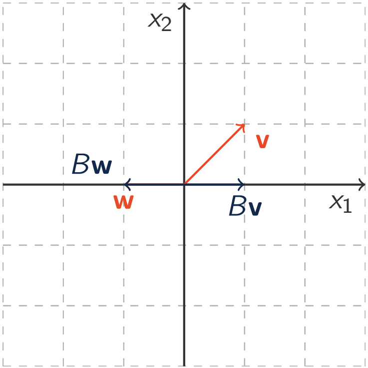
Solution.

Let $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ be our input.

$$B\mathbf{x} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} x_1 \\ 0 \end{bmatrix}.$$

So the machine B replaces the second entry of the vector $\mathbf{x}$ by 0.

Geometrically speaking, this machine projects a vector onto the x_1-axis.



Composition of machines. Let A be an $m \times n$ matrix and B be an $k \times l$ matrix.Now we can compose the two machines:

$$\xrightarrow{\mathbf{x}} \boxed{A} \xrightarrow{A\mathbf{x}} \boxed{B} \xrightarrow{B(A\mathbf{x})}$$

Question. This composition only works for some k, l, m, n. For which?

Solution.

If A is an $m \times n$-matrix and $\mathbf{x}$ in $\mathbb{R}^n$, then $A\mathbf{x}$ is in $\mathbb{R}^m$.

In order to calculate $B(A\mathbf{x})$, we then need B to have m columns.

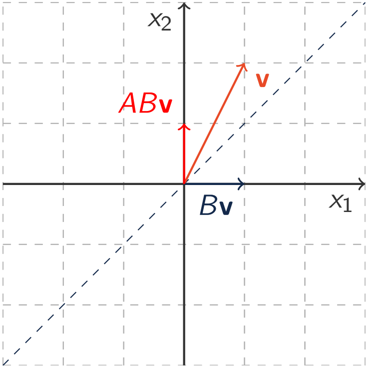
So we need $l = m$. Both n and k can be arbitrary.

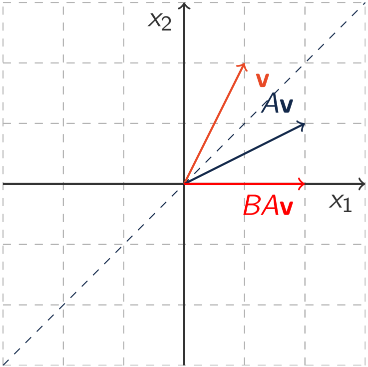
Example. Let $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ be as before. Is $A(B\mathbf{x}) = B(A\mathbf{x})$?

Solution.

No, projection and reflection do not commute!

$$A\left(B\begin{bmatrix} 1 \\ 2 \end{bmatrix}\right) = A\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \qquad B\left(A\begin{bmatrix} 1 \\ 2 \end{bmatrix}\right) = B\begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$$





LINEAR ALGEBRA

Matrix multiplication

ILLINOIS

Department of Mathematics

Definition. Let A be an $m \times n$-matrix and let $B = \begin{bmatrix} \mathbf{b}_1 & \dots & \mathbf{b}_p \end{bmatrix}$ be an $n \times p$-matrix. We define

$$AB := [A\mathbf{b}_1 \quad A\mathbf{b}_2 \quad \cdots \quad A\mathbf{b}_p]$$

Example. Compute AB where $A = \begin{bmatrix} 4 & -2 \\ 3 & -5 \\ 0 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} 2 & -3 \\ 6 & -7 \end{bmatrix}$.

Solution.

$$A\mathbf{b}_1 = \begin{bmatrix} 4 & -2 \\ 3 & -5 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 6 \end{bmatrix} = 2 \begin{bmatrix} 4 \\ 3 \\ 0 \end{bmatrix} + 6 \begin{bmatrix} -2 \\ -5 \\ 1 \end{bmatrix} = \begin{bmatrix} -4 \\ -24 \\ 6 \end{bmatrix}$$

$$A\mathbf{b}_2 = \begin{bmatrix} 4 & -2 \\ 3 & -5 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} -3 \\ -7 \end{bmatrix} = -3 \begin{bmatrix} 4 \\ 3 \\ 0 \end{bmatrix} - 7 \begin{bmatrix} -2 \\ -5 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 26 \\ -7 \end{bmatrix} \rightsquigarrow AB = \begin{bmatrix} A\mathbf{b}_1 & A\mathbf{b}_2 \end{bmatrix} = \begin{bmatrix} -4 & 2 \\ -24 & 26 \\ 6 & -7 \end{bmatrix}$$

Remark. $A\mathbf{b}_1$ is a **linear combination** of the columns of A and $A\mathbf{b}_2$ is a **linear combination** of the columns of A. Each column of AB is a **linear combination** of the columns of A using coefficients from the corresponding columns of B.

Definition. Let A be an $m \times n$-matrix and let B be an $n \times p$-matrix. We define

$$AB := [A\mathbf{b}_1 \quad A\mathbf{b}_2 \quad \cdots \quad A\mathbf{b}_p]$$

Question. If C is 4×3 and D is 3×2, are CD and DC defined? What are their sizes?

Solution.

The product AB is only defined if B has as many rows as A has columns.

If this is the case, then AB has as many rows as A and as many columns as B.

Thus CD is defined and 4×2. DC is not defined.

We already learnt that we can think of matrices as machines.

- Let B be $n \times p$: input $\mathbf{x} \in \mathbb{R}^p$, output $\mathbf{c} = B\mathbf{x} \in \mathbb{R}^n$.
- Let A be $m \times n$: input $\mathbf{y} \in \mathbb{R}^n$, output $\mathbf{b} = A\mathbf{y} \in \mathbb{R}^m$.

We already saw that these machines can be composed.

$$\xrightarrow{\ \mathbf{x}\ } \boxed{B} \xrightarrow{\ B\mathbf{x}\ } \boxed{A} \xrightarrow{\ A(B\mathbf{x})\ }$$

Question. How does that compare to the machine given by AB?

$$\xrightarrow{\ \mathbf{x}\ } \boxed{AB} \xrightarrow{\ (AB)\mathbf{x}\ }$$

Example. Consider

$$A = \begin{bmatrix} 2 & 0 \\ 1 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Compute $(AB)\mathbf{x}$ and $A(B(\mathbf{x}))$. Are these the same?

Solution.

$$B\mathbf{x} = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} x_1 + 2x_2 \\ x_2 \end{bmatrix}$$

$$A(B\mathbf{x}) = \begin{bmatrix} 2 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 + 2x_2 \\ x_2 \end{bmatrix} = (x_1 + 2x_2) \begin{bmatrix} 2 \\ 1 \end{bmatrix} + (x_2) \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 2x_1 + 4x_2 \\ x_1 + 3x_2 \end{bmatrix}$$

$$(AB)\mathbf{x} = \begin{bmatrix} A\mathbf{b}_1 & A\mathbf{b}_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1 A\mathbf{b}_1 + x_2 A\mathbf{b}_2 = x_1 (1 \begin{bmatrix} 2 \\ 1 \end{bmatrix} + 0 \begin{bmatrix} 0 \\ 1 \end{bmatrix}) + x_2 (2 \begin{bmatrix} 2 \\ 1 \end{bmatrix} + 1 \begin{bmatrix} 0 \\ 1 \end{bmatrix})$$

$$= x_1 \begin{bmatrix} 2 \\ 1 \end{bmatrix} + x_2 \begin{bmatrix} 4 \\ 3 \end{bmatrix} = \begin{bmatrix} 2x_1 + 4x_2 \\ x_1 + 3x_2 \end{bmatrix}$$

Theorem 7. *Let A be an $m \times n$ matrix and B be an $n \times p$ matrix. Then for every $\mathbf{x} \in \mathbb{R}^p$*

$$A(B\mathbf{x}) = (AB)\mathbf{x}.$$

Row-Column Rule for computing AB. Let A be $m \times n$ and B be $n \times p$ such that

$$A = \begin{bmatrix} \mathbf{R}_1 \\ \vdots \\ \mathbf{R}_m \end{bmatrix}, \quad \text{and} \quad B = \begin{bmatrix} \mathbf{C}_1 & \dots & \mathbf{C}_p \end{bmatrix}.$$

Then

$$AB = \begin{bmatrix} \mathbf{R}_1\mathbf{C}_1 & \dots & \mathbf{R}_1\mathbf{C}_p \\ \mathbf{R}_2\mathbf{C}_1 & \dots & \mathbf{R}_2\mathbf{C}_p \\ \mathbf{R}_m\mathbf{C}_1 & \dots & \mathbf{R}_m\mathbf{C}_p \end{bmatrix} \quad \text{and} \quad (AB)_{ij} = \mathbf{R}_i\mathbf{C}_j = a_{i1}b_{1j} + a_{i2}b_{2j} + \dots + a_{in}b_{nj}.$$

Example. $A = \begin{bmatrix} 2 & 3 & 6 \\ -1 & 0 & 1 \end{bmatrix}$, $B = \begin{bmatrix} 2 & -3 \\ 0 & 1 \\ 4 & -7 \end{bmatrix}$. Compute AB, if it is defined.

Solution.

$$AB = \begin{bmatrix} 2\cdot 2 + 3\cdot 0 + 6\cdot 4 & 2\cdot(-3) + 3\cdot 1 + 6\cdot(-7) \\ -1\cdot 2 + 0\cdot 0 + 1\cdot 4 & -1\cdot(-3) + 0\cdot 1 + 1\cdot(-7) \end{bmatrix} = \begin{bmatrix} 28 & -45 \\ 2 & -4 \end{bmatrix}$$

Outer Product Rule for computing AB. Let A be $m \times n$ and B be $n \times p$ such that

$$A = \begin{bmatrix} \mathbf{C}_1 & \cdots & \mathbf{C}_n \end{bmatrix}, \quad \text{and} \quad B = \begin{bmatrix} \mathbf{R}_1 \\ \vdots \\ \mathbf{R}_n \end{bmatrix}.$$

Then

$$AB = \mathbf{C}_1\mathbf{R}_1 + \cdots + \cdots + \mathbf{C}_n\mathbf{R}_n$$

Example. $A = \begin{bmatrix} 2 & 3 & 6 \\ -1 & 0 & 1 \end{bmatrix}$, $B = \begin{bmatrix} 2 & -3 \\ 0 & 1 \\ 4 & -7 \end{bmatrix}$. Compute AB, if it is defined.

Solution.

$$\begin{bmatrix} 2 & 3 & 6 \\ -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & -3 \\ 0 & 1 \\ 4 & -7 \end{bmatrix} = \begin{bmatrix} 2 \\ -1 \end{bmatrix} \begin{bmatrix} 2 & -3 \end{bmatrix} + \begin{bmatrix} 3 \\ 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \end{bmatrix} + \begin{bmatrix} 6 \\ 1 \end{bmatrix} \begin{bmatrix} 4 & -7 \end{bmatrix}$$

$$= \begin{bmatrix} 4 & -6 \\ -2 & 3 \end{bmatrix} + \begin{bmatrix} 0 & 3 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 24 & -42 \\ 4 & -7 \end{bmatrix} = \begin{bmatrix} 28 & -45 \\ 2 & -4 \end{bmatrix}$$

LINEAR ALGEBRA

Properties of matrix multiplication

ILLINOIS

Department of Mathematics

Definition. The identity matrix I_n of size n is defined as

$$I_n = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix}.$$

Theorem 8. *Let A be an $m \times n$ matrix and let B and C be matrices for which the indicated sums and products are defined.*

(a) $A(BC) = (AB)C$ *(associative law of multiplication)*

(b) $A(B + C) = AB + AC$, $(B + C)A = BA + CA$ *(distributive laws)*

(c) $r(AB) = (rA)B = A(rB)$ *for every scalar* r,

(d) $A(rB + sC) = rAB + sAC$ *for every scalars* r, s *(linearity of matrix multiplication)*

(e) $I_mA = A = AI_n$ *(identity for matrix multiplication)*

Warning. Properties above are analogous to properties of real numbers. But **NOT ALL** properties of real number also hold for matrices.

Example. Let $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$. Determine AB and BA. Are these matrices the same?

Solution.

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

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

No. These matrices are not the same: $AB \neq BA$. Matrix multiplication is *not* commutative.

Example. Let $A = \begin{bmatrix} 1 & 2 & 0 \\ 3 & 0 & 1 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 2 \\ 0 & 1 \\ -2 & 4 \end{bmatrix}$. Compute $(AB)^T$, $A^T B^T$ and $B^T A^T$.

Solution.

$$AB = \begin{bmatrix} 1 & 2 & 0 \\ 3 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 0 & 1 \\ -2 & 4 \end{bmatrix} = \begin{bmatrix} 1 & 4 \\ 1 & 10 \end{bmatrix}$$

$$(AB)^T = \begin{bmatrix} 1 & 1 \\ 4 & 10 \end{bmatrix}$$

$$A^T B^T = \begin{bmatrix} 1 & 3 \\ 2 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & -2 \\ 2 & 1 & 4 \end{bmatrix} = \begin{bmatrix} 7 & 3 & 10 \\ 2 & 0 & -4 \\ 2 & 1 & 4 \end{bmatrix}$$

$$B^T A^T = \begin{bmatrix} 1 & 0 & -2 \\ 2 & 1 & 4 \end{bmatrix} \begin{bmatrix} 1 & 3 \\ 2 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 4 & 10 \end{bmatrix}$$

Theorem 9. *The transpose of a product is the product of transposes* **in opposite order**:

$$(AB)^T = B^T A^T$$

Definition. We write A^k for $A \cdots A$, k-times; that is A^k is obtained by multiplying A k-times with itself.

Question. For which matrices A does A^k make sense? If A is $m \times n$ what can m and n be?

Solution.

To be able to multiply A by any $m \times n$-matrix, we need that $m = n$.

Example. Determine $\begin{bmatrix} 1 & 0 \\ 3 & 2 \end{bmatrix}^3$.

Solution.

$$\begin{aligned} \begin{bmatrix} 1 & 0 \\ 3 & 2 \end{bmatrix}^3 &= \begin{bmatrix} 1 & 0 \\ 3 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 3 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 3 & 2 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 \\ 9 & 4 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 3 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 21 & 8 \end{bmatrix} \end{aligned}$$

Remark. In general, calculating high powers of large matrix in this way is very hard. We will later learn a clever way to do this efficiently.

LINEAR ALGEBRA

Elementary matrices

ILLINOIS

Department of Mathematics

Definition. An **elementary matrix** is one that is obtained by performing a single elementary row operation on an identity matrix.
A **permutation matrix** is one that is obtained by performing row exchanges on an identity matrix.

Example. Consider $E_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, $E_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$, $E_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \end{bmatrix}$. Are these matrices elementary matrices?

Solution.

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \underset{R_2 \to 3R_2}{\rightsquigarrow} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \underset{R_2 \leftrightarrow R_3}{\rightsquigarrow} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \underset{R_3 \to R_3 + 2R_1}{\rightsquigarrow} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \end{bmatrix}$$

Question. Let A be a 3×3-matrix. What happens to A if you multiply it by one of E_1, E_2 and E_3?

Solution.

$$E_1 A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 3a_{21} & 3a_{22} & 3a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

$$E_2 A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{31} & a_{32} & a_{33} \\ a_{21} & a_{22} & a_{23} \end{bmatrix}$$

$$E_3 A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} + 2a_{11} & a_{32} + 2a_{12} & a_{33} + 2a_{13} \end{bmatrix}$$

Theorem 10. *If an elementary row operation is performed on an $m \times n$-matrix A, the resulting matrix can be written as EA, where the $m \times m$-matrix E is created by performing the same row operations on I_m.*

Theorem 11. *Let A, B be two $m \times n$-matrices and row-equivalent. Then there is a sequence $m \times m$-elementary matrices $E_1, \ldots, E_\ell$ such that*

$$E_\ell \ldots E_1 A = B.$$

Example. Consider $A = \begin{bmatrix} 0 & 1 \\ 1 & 2 \\ 2 & 4 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 2 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$. Find two elementary matrices E_1, E_2 such that $E_2 E_1 A = B$.

Solution.

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 2 \\ 2 & 4 \end{bmatrix} \underset{R_1 \leftrightarrow R_2}{\rightsquigarrow} \begin{bmatrix} 1 & 2 \\ 0 & 1 \\ 2 & 4 \end{bmatrix} \underset{R_3 \to R_3 - 2R_1}{\rightsquigarrow} \begin{bmatrix} 1 & 2 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} = B$$

$$\text{Set } E_1 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \text{and } E_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix}.$$

$$E_2 E_1 A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 2 \\ 2 & 4 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} = B$$

Example. From the previous example, we know

$$\underbrace{\begin{bmatrix}1 & 2\\0 & 1\\0 & 0\end{bmatrix}}_{B} = \begin{bmatrix}1 & 0 & 0\\0 & 1 & 0\\-2 & 0 & 1\end{bmatrix}\begin{bmatrix}0 & 1 & 0\\1 & 0 & 0\\0 & 0 & 1\end{bmatrix}\underbrace{\begin{bmatrix}0 & 1\\1 & 2\\2 & 4\end{bmatrix}}_{A}.$$

Find elementary matrices E_1^{-1} and E_2^{-1} such that $A = E_1^{-1}E_2^{-1}B$.

Solution.

$$A = \begin{bmatrix}0 & 1\\1 & 2\\2 & 4\end{bmatrix} \underset{R_1\leftrightarrow R_2}{\leftsquigarrow} \begin{bmatrix}1 & 2\\0 & 1\\2 & 4\end{bmatrix} \underset{R_3\to R_3+2R_1}{\leftsquigarrow} \begin{bmatrix}1 & 2\\0 & 1\\0 & 0\end{bmatrix} = B$$

$$\text{Set } E_1^{-1} = \begin{bmatrix}0 & 1 & 0\\1 & 0 & 0\\0 & 0 & 1\end{bmatrix} \text{ and } E_2^{-1} = \begin{bmatrix}1 & 0 & 0\\0 & 1 & 0\\2 & 0 & 1\end{bmatrix}.$$

LINEAR ALGEBRA

Inverse of a matrix

ILLINOIS
Department of Mathematics

The inverse of a real number a is denoted by a^{-1}. For example, $7^{-1} = 1/7$ and

$$7 \cdot 7^{-1} = 7^{-1} \cdot 7 = 1.$$

Remark. Not all real numbers have inverse. 0^{-1} is not well defined, since there is no real number b such that $0 \cdot b = 1$.

Remember that the identity matrix I_n is the $n \times n$-matrix

$$I_n = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

Definition. An $n \times n$ matrix A is said to be **invertible** if there is an $n \times n$ matrix C satisfying

$$CA = AC = I_n$$

where I_n is the $n \times n$ identity matrix. We call C the **inverse** of A.

Example. What is the inverse of $\begin{bmatrix} 1 & 0 & 0 \\ 5 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$?

Solution.

Elementary matrices are *invertible*, because row operations are *reversible*. So the inverse matrix is the elementary matrix corresponding to $R_2 \to R_2 - 5R_1$: $\begin{bmatrix} 1 & 0 & 0 \\ -5 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

Theorem 12. *Let A be an invertible matrix, then its inverse C is unique.*

Proof. Assume B and C are both inverses of A. Then

$$BA = AB = I_n \tag{I}$$

$$CA = AC = I_n \tag{II}$$

Thus

$$B = BI_n \underset{\text{(II)}}{=} BAC \underset{\text{(I)}}{=} I_n C = C$$

Theorem 13. *Suppose A and B are invertible. Then:*

(a) A^{-1} is invertible and $\left(A^{-1}\right)^{-1} = A$ (i.e. A is the inverse of A^{-1}).

(b) AB is invertible and $(AB)^{-1} = B^{-1}A^{-1}$.

(c) A^T is invertible and $\left(A^T\right)^{-1} = \left(A^{-1}\right)^T$.

Proof.

(a)

$$AA^{-1} = I = A^{-1}A \quad \checkmark$$

(b)

$$(B^{-1}A^{-1})(AB) = B^{-1}IB = B^{-1}B = I \quad \checkmark$$
$$(AB)(B^{-1}A^{-1}) = AIA^{-1} = AA^{-1} = I \quad \checkmark$$

(c)

$$A^T(A^{-1})^T = (A^{-1}A)^T = I^T = I \quad \checkmark$$
$$(A^{-1})^T A^T = (AA^{-1})^T = I^T = I \quad \checkmark$$

- We will write A^{-1} for the inverse of A. Multiplying by A^{-1} is like "dividing by A."
- Do not write $\frac{A}{B}$. Why?

It is unclear whether this means AB^{-1} or $B^{-1}A$, and these two matrices are *different*.

- Fact: if $AB = I$, then $A^{-1} = B$ and so $BA = I$. (Not so easy to show at this stage.)
- Not all $n \times n$ matrices are invertible. For example, the 2×2 matrix

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

is not invertible. Try to find an inverse!

$$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} c & d \\ 0 & 0 \end{bmatrix} \neq I_2$$

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & a \\ 0 & c \end{bmatrix} \neq I_2$$

Theorem 14. *Let A be an invertible $n \times n$ matrix. Then for each $\mathbf{b}$ in $\mathbb{R}^n$, the equation $A\mathbf{x} = \mathbf{b}$ has the unique solution $\mathbf{x} = A^{-1}\mathbf{b}$.*

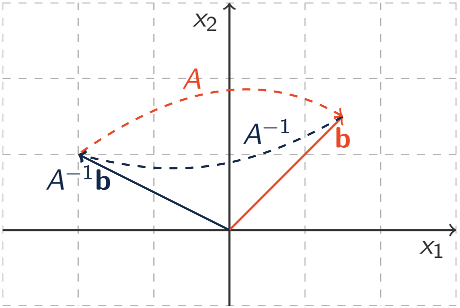
Proof.

The vector $A^{-1}\mathbf{b}$ is a solution, because

$$A(A^{-1}\mathbf{b}) = (AA^{-1})\mathbf{b} = I_n\mathbf{b} = \mathbf{b}$$

Suppose there is another solution $\mathbf{w}$, then $A\mathbf{w} = \mathbf{b}$. Thus

$$\mathbf{w} = I_n\mathbf{w} = A^{-1}A\mathbf{w} = A^{-1}\mathbf{b}.$$



Question. If A is an invertible $n \times n$ matrix, then A has how many pivot columns?

A must have n pivots, because otherwise $A\mathbf{x} = \mathbf{b}$ would not have a solution of each $\mathbf{b}$.

LINEAR ALGEBRA

Computing the inverse

ILLINOIS

Department of Mathematics

Question. When is the 1×1 matrix $\begin{bmatrix} a \end{bmatrix}$ invertible?

Solution.

When $a \neq 0$. Its inverse is $[\,\frac{1}{a}\,]$.

Theorem 15. *Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. If $ad - bc \neq 0$, then A is invertible and*

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

If $ad - bc = 0$, then A is not invertible.

Proof.

$$\frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \frac{1}{ad - bc} \begin{bmatrix} da - bc & db - bd \\ -ca + ac & -cb + ad \end{bmatrix}$$
$$= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Theorem 16. *Let A be an $n \times n$-matrix. The following are equivalent:*

- *A is invertible.*
- *the reduced echelon form of A is I_n.*

Proof.

Suppose A can be row-reduced to the identity matrix.

$$A = A_0 \rightsquigarrow A_1 \rightsquigarrow \cdots \rightsquigarrow A_m = I_n$$

Thus there are elementary matrices $E_1, \ldots, E_m$ such that

$$E_m E_{m-1} \ldots E_1 A = I_n.$$

Thus

$$A^{-1} = E_m E_{m-1} \ldots E_1 = E_m E_{m-1} \ldots E_1 I_n.$$

Theorem 17. *Suppose A is invertible. The every sequence of elementary row operations that reduces A to I_n will also transform I_n to A^{-1}.*

Algorithm.

- Place A and I side-by-side to form an augmented matrix $[\ A \mid I\ ]$.
  This is an $n \times 2n$ matrix (Big Augmented Matrix), instead of $n \times (n+1)$.
- Perform row operations on this matrix (which will produce identical operations on A and I).
- By Theorem:

$$[\ A \mid I\ ] \text{ will row reduce to } \left[\ I \mid A^{-1}\ \right]$$

or A is not invertible.

Example. Find the inverse of $A = \begin{bmatrix} 2 & 0 & 0 \\ -3 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$, if it exists.

Solution.

$$\left[A \mid I\ \right] = \left[\begin{array}{ccc|ccc} 2 & 0 & 0 & 1 & 0 & 0 \\ -3 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \end{array}\right] \sim \cdots \sim \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{1}{2} & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & \frac{3}{2} & 1 & 0 \end{array}\right]$$

Example. Let's do the previous example step by step.

Solution.

$$[\,A \mid I\,] = \left[\begin{array}{ccc|ccc} 2 & 0 & 0 & 1 & 0 & 0 \\ -3 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \end{array}\right]$$

$$\underset{R_1 \to \frac{1}{2}R_1}{\rightsquigarrow} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{1}{2} & 0 & 0 \\ -3 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \end{array}\right]$$

$$\underset{R_2 \to R_2 + 3R_1}{\rightsquigarrow} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 1 & \frac{3}{2} & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \end{array}\right]$$

$$\underset{R_2 \leftrightarrow R_3}{\rightsquigarrow} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{1}{2} & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & \frac{3}{2} & 1 & 0 \end{array}\right]$$

LINEAR ALGEBRA

LU decomposition

ILLINOIS

Department of Mathematics

Definition. An $n \times n$ matrix A is called

upper triangular if it is of the form $\begin{bmatrix} \star & \star & \star & \star & \star \\ 0 & \star & \star & \star & \star \\ 0 & 0 & \star & \star & \star \\ 0 & 0 & 0 & \ddots & \vdots \\ 0 & 0 & 0 & 0 & * \end{bmatrix}$,

lower triangular if it is of the form $\begin{bmatrix} \star & 0 & 0 & 0 & 0 \\ \star & \star & 0 & 0 & 0 \\ \star & \star & \star & 0 & 0 \\ \star & \star & \star & \ddots & \vdots \\ \star & \star & \star & \star & \star \end{bmatrix}$.

Example. Give a few examples of upper and lower triangular matrices!

Solution.

Upper triangular matrices:

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

Lower triangular matrices:

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

Neither upper nor lower triangular: $\begin{bmatrix} 0 & 2 \\ 2 & 1 \end{bmatrix}$

Theorem 18. *The product of two lower (upper) triangular matrices is lower (upper) triangular.*

Example. Consider the row operation $R_i \to R_i + cR_j$ where $j < i$. What can you say about the elementary matrix E corresponding this row operation? What about E^{-1}?

Solution.

Such a matrix is lower-triangular! Take $R_3 \to R_3 + cR_2$, then

$$E = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & c & 1 \end{bmatrix}.$$

Recall that $E^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -c & 1 \end{bmatrix}$. So E^{-1} is also lower-triangular.

Remark. The inverse of a lower (upper) triangular matrix (if it exists) is again lower (upper) triangular.

Definition. A matrix A has an **LU decomposition** if there is a lower triangular matrix L and a upper triangular matrix U such that $A = LU$.

Theorem 19. *Let A be an $n \times n$-matrix. If A can be brought to echelon form just using row operations of the form $R_i \to R_i + cR_j$ where $j < i$, then A has an LU-decomposition.*

Proof.

$$A \rightsquigarrow A_1 \rightsquigarrow A_2 \cdots \rightsquigarrow A_m = U$$

using row operations of the form $R_i \to R_i + cR_j$ where $j < i$. Then there are lower-triangular matrices $E_1, \dots, E_m$ such that

$$E_m E_{m-1} \dots E_1 A = U.$$

Multiply the equation by $E_1^{-1} \dots E_{m-1}^{-1} E_m^{-1}$, we get

$$A = \underbrace{E_1^{-1} \dots E_{m-1}^{-1} E_m^{-1}}_{=:L} U.$$

Remark. Not every matrix has an LU decomposition.
The LU decomposition of a matrix is not unique.

Example. Let $A = \begin{bmatrix} 2 & 1 & 1 \\ 4 & -6 & 0 \\ -2 & 7 & 2 \end{bmatrix}$. Determine its LU decomposition.

Solution.

$$\begin{bmatrix} 2 & 1 & 1 \\ 4 & -6 & 0 \\ -2 & 7 & 2 \end{bmatrix} \underset{R_2 \to R_2 - 2R_1}{\rightsquigarrow} \begin{bmatrix} 2 & 1 & 1 \\ 0 & -8 & -2 \\ -2 & 7 & 2 \end{bmatrix} \underset{R_3 \to R_3 + R_1}{\rightsquigarrow} \begin{bmatrix} 2 & 1 & 1 \\ 0 & -8 & -2 \\ 0 & 8 & 3 \end{bmatrix} \underset{R_3 \to R_3 + R_2}{\rightsquigarrow} \begin{bmatrix} 2 & 1 & 1 \\ 0 & -8 & -2 \\ 0 & 0 & 1 \end{bmatrix}$$

$$U = \begin{bmatrix} 2 & 1 & 1 \\ 0 & -8 & -2 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 & 1 \\ 4 & -6 & 0 \\ -2 & 7 & 2 \end{bmatrix} = E_3 E_2 E_1 A$$

$$E_1^{-1} E_2^{-1} E_3^{-1} U = \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & -1 & 1 \end{bmatrix}}_{=:L} \begin{bmatrix} 2 & 1 & 1 \\ 0 & -8 & -2 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 1 & 1 \\ 4 & -6 & 0 \\ -2 & 7 & 2 \end{bmatrix} = A$$

Remark. In previous example $L = \begin{bmatrix} 1 & 0 & 0 \\ \ell_{21} & 1 & 0 \\ \ell_{31} & \ell_{32} & 1 \end{bmatrix}$ where ℓ_{ij} is the factor between pivot and the entry you want to make zero in the elimination process. This works in general if you do the row operations **in the right order**.

Step 1:

$$A = \begin{bmatrix} 2 & 1 & 1 \\ 4 & -6 & 0 \\ -2 & 7 & 2 \end{bmatrix} \underset{R_2 \to R_2 \boxed{-2} R_1, R_3 \to R_3 + \boxed{1} R_1}{\rightsquigarrow} = \begin{bmatrix} 2 & 1 & 1 \\ 0 & -8 & -2 \\ 0 & 8 & 3 \end{bmatrix}$$

Then set $\ell_{21} = 2$ and $\ell_{31} = -1$.

Step 2:

$$\begin{bmatrix} 2 & 1 & 1 \\ 0 & -8 & -2 \\ 0 & 8 & 3 \end{bmatrix} \underset{R_3 \to R_3 + \boxed{1} \cdot R_2}{\rightsquigarrow} = \begin{bmatrix} 2 & 1 & 1 \\ 0 & -8 & -2 \\ 0 & 0 & 1 \end{bmatrix}.$$

Set $\ell_{32} = -1$,' $U := \begin{bmatrix} 2 & 1 & 1 \\ 0 & -8 & -2 \\ 0 & 0 & 1 \end{bmatrix}$, and $L := \begin{bmatrix} 1 & 0 & 0 \\ \ell_{21} & 1 & 0 \\ \ell_{31} & \ell_{32} & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & -1 & 1 \end{bmatrix}$.

Example. Determine the LU-decomposition of $\begin{bmatrix} 1 & 2 & 2 \\ 4 & 4 & 4 \\ 4 & 4 & 8 \end{bmatrix}$.

Solution.

$$\begin{bmatrix} 1 & 2 & 2 \\ 4 & 4 & 4 \\ 4 & 4 & 8 \end{bmatrix} \underset{R_2 \to R_2 \boxed{-4} R_1, R_3 \to R_3 \boxed{-4} R_1}{\rightsquigarrow} \begin{bmatrix} 1 & 2 & 2 \\ 0 & -4 & -4 \\ 0 & -4 & 0 \end{bmatrix} \underset{R_3 \to R_3 \boxed{-1} R_2}{\rightsquigarrow} \begin{bmatrix} 1 & 2 & 2 \\ 0 & -4 & -4 \\ 0 & 0 & 4 \end{bmatrix}.$$

Thus

$$U := \begin{bmatrix} 1 & 2 & 2 \\ 0 & -4 & -4 \\ 0 & 0 & 4 \end{bmatrix}, L := \begin{bmatrix} 1 & 0 & 0 \\ 4 & 1 & 0 \\ 4 & 1 & 1 \end{bmatrix}.$$

Check that $A = LU$.

Remark. It is important that you do the row operations in the right order!

LINEAR ALGEBRA

Solving linear systems using LU decomposition

ILLINOIS

Department of Mathematics

Theorem 20. *Let A be an $n \times n$-matrix such that $A = LU$, where L is a lower triangular matrix and U is a upper triangular matrix, and let $\mathbf{b} \in \mathbb{R}^n$. In order to find a solution of the linear system*

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

it is enough to find a solution of the linear system

$$U\mathbf{x} = \mathbf{c},$$

where $\mathbf{c}$ satisfies $L\mathbf{c} = \mathbf{b}$.

Proof.

If $L\mathbf{c} = \mathbf{b}$ and $U\mathbf{x} = \mathbf{c}$, then

$$A\mathbf{x} = (LU)\mathbf{x} = L(U\mathbf{x}) = L\mathbf{c} = \mathbf{b}$$

On the other hand, suppose $A\mathbf{x} = \mathbf{b}$. We take $\mathbf{c}$ to be the vector $U\mathbf{x}$. Then $L\mathbf{c}$ is equal to $L(U\mathbf{x}) = A\mathbf{x} = \mathbf{b}$, so in total we have both $U\mathbf{x} = \mathbf{c}$ and $L\mathbf{c} = \mathbf{b}$.

Example. Find a solution to the linear system $\begin{bmatrix} 2 & 1 & 1 \\ 4 & -6 & 0 \\ -2 & 7 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 5 \\ -2 \\ 9 \end{bmatrix}$.

Solution.

Recall the LU decomposition $\begin{bmatrix} 2 & 1 & 1 \\ 4 & -6 & 0 \\ -2 & 7 & 2 \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & -1 & 1 \end{bmatrix}}_{L} \underbrace{\begin{bmatrix} 2 & 1 & 1 \\ 0 & -8 & -2 \\ 0 & 0 & 1 \end{bmatrix}}_{U}$.

$$\underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & -1 & 1 \end{bmatrix}}_{L} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \underbrace{\begin{bmatrix} 5 \\ -2 \\ 9 \end{bmatrix}}_{\mathbf{b}} \qquad \underbrace{\begin{bmatrix} 2 & 1 & 1 \\ 0 & -8 & -2 \\ 0 & 0 & 1 \end{bmatrix}}_{U} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \underbrace{\begin{bmatrix} 5 \\ -12 \\ 2 \end{bmatrix}}_{\mathbf{c}}.$$

Forward substitution:

Row 1 $\to c_1 = 5$.

Row 2 $\to 10 + c_2 = -2 \rightsquigarrow c_2 = -12$.

Row 3 $\to -5 + 12 + c_3 = 9 \rightsquigarrow c_3 = 2$.

Backwards substitution:

Row 3 $\to x_3 = 2$.

Row 2 $\to -8x_2 - 4 = -12 \rightsquigarrow x_2 = 1$.

Row 1 $\to 2x_1 + 1 + 2 = 5 \rightsquigarrow x_1 = 1$.

Question. Why do we care about LU decomposition if we already have Gaussian elimination?

Solution.

Imagine you solve many linear systems $A\mathbf{x} = \mathbf{b}$ where A always stays the same, but b varies.

You have to do the LU-factorization of A only once, and not for every $\mathbf{b}$.

Finding solutions of the two triangular systems above takes $2n^2 - n$ arithmetic operations.

You need $(4n^3 + 9n^2 - 7n)/6$ arithmetic operations for Gaussian elimination.

Theorem 21. *Let A be $n \times n$ matrix. Then there is a permutation matrix P such that PA has an LU-decomposition.*

Proof.

If A can be brought to echelon form with the help of row exchanges, we can do those exchange first.

So there is a permutation matrix P such that PA can be brought to echelon form without row exchanges.

Example. Let $A = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 1 & 0 \\ 2 & 1 & 0 \end{bmatrix}$. Find a permutation matrix P such that PA has a LU decomposition.

Solution.

$$A = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 1 & 0 \\ 2 & 1 & 0 \end{bmatrix} \underset{R_1 \leftrightarrow R_3}{\rightsquigarrow} \begin{bmatrix} 2 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad \text{Set } P = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}.$$

$$PA = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \underset{R_2 \to R_2 - .5R_1}{\rightsquigarrow} \underbrace{\begin{bmatrix} 2 & 1 & 0 \\ 0 & .5 & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{=:U}$$

Thus

$$PA = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ .5 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{=:L} \begin{bmatrix} 2 & 1 & 0 \\ 0 & .5 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

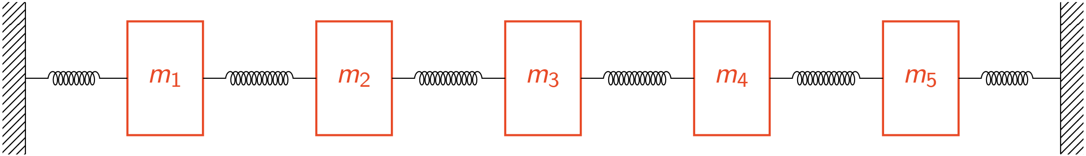
LINEAR ALGEBRA

Spring-mass systems

ILLINOIS

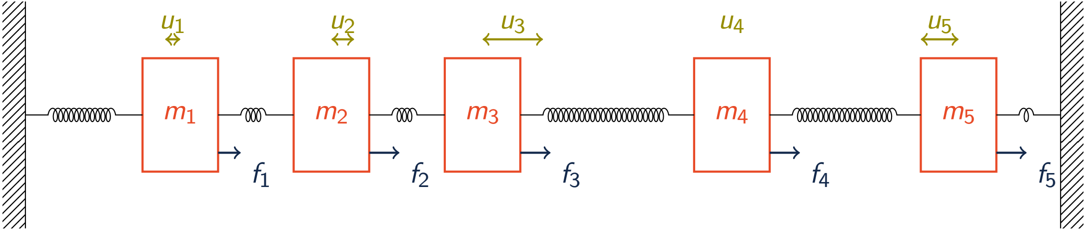
Department of Mathematics

Example. Consider the following spring-mass system, consisting of five masses and six springs fixed between two walls.



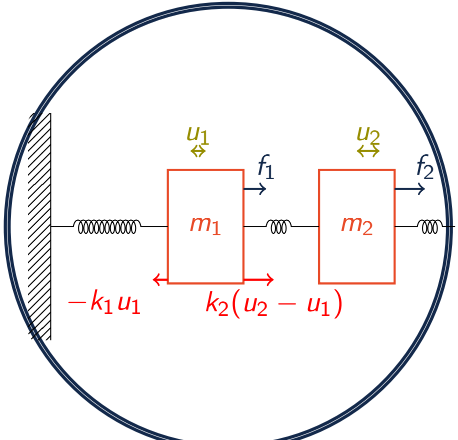
Goal.

- Add (steady) applied forces f_1, f_2, f_3, f_4, f_5 a (steady) applied force on mass i.
- Compute the displacements u_1, u_2, u_3, u_4, u_5 of the fives masses.



Equilibrium. In the equilibrium, the forces at each mass add up to 0.

Hooke's law. The force F needed to extend or compress a spring by some distance u is proportional to that distance; that is $F = -ku$, where k is a constant factor characteristic of the spring, called its **stiffness**. Let k_i be the stiffness of the i-th spring.



Springs:

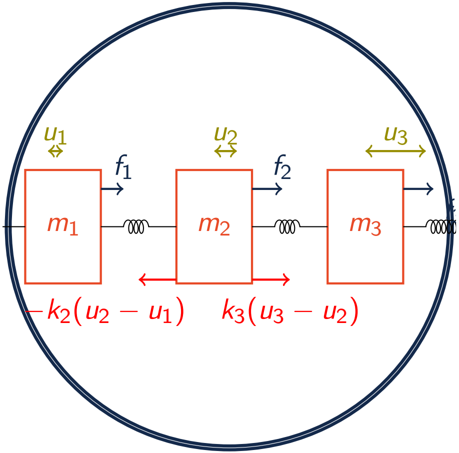
- $u_1 > 0 \leadsto$ spring 1 is extended.
- $u_2 - u_1 > 0 \leadsto$ spring 2 is extended.

Forces at m_1:

- applied forces f_1 (if positive, pushes m_1 to the right).
- spring 1: $-k_1 u_1$ (since $u_1 > 0$, pulls m_1 to the left).
- spring 2: $k_2(u_2 - u_1)$ (since $u_2 - u_1 > 0$, pulls m_1 to the right).

Equilibrium:

$$f_1 - k_1 u_1 + k_2(u_2 - u_1) = 0 \leadsto (k_1 + k_2)u_1 - k_2 u_2 = f_1$$



Springs:

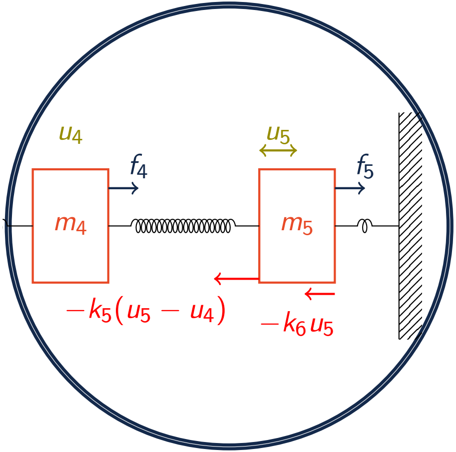
- $u_2 - u_1 > 0 \rightsquigarrow$ spring 2 is extended.
- $u_3 - u_2 > 0 \rightsquigarrow$ spring 3 is extended.

Forces at m_2:

- applied forces f_2 (if positive, pushes m_2 to the right).
- spring 2: $-k_2(u_2 - u_1)$ (since $(u_2 - u_1) > 0$, pulls m_2 to the left).
- spring 3: $k_3(u_3 - u_2)$ (since $u_3 - u_2 > 0$, pulls m_2 to the right).

Equilibrium at m_2:

$$f_2 - k_2(u_2 - u_1) + k_3(u_3 - u_2) = 0$$
$$\rightsquigarrow -k_2 u_1 + (k_2 + k_3)u_2 - k_3 u_3 = f_2$$



Springs:

- $u_5 - u_4 > 0 \rightsquigarrow$ spring 5 is extended.
- $u_5 > 0 \rightsquigarrow$ spring 6 is compressed.

Forces at m_5:

- applied forces f_5 (if positive, pushes m_5 to the right).
- spring 5: $-k_5(u_5 - u_4)$ (since $(u_5 - u_4) > 0$, pulls m_5 to the left).
- spring 6: $-k_6 u_5$ (since $u_5 > 0$, pushes m_5 to the left).

Equilibrium at m_5:

$$f_5 - k_5(u_5 - u_4) - k_6 u_5 = 0$$

$$\rightsquigarrow (k_5 + k_6) u_5 - k_5 u_4 = f_5$$

Equilibrium equations.

$$
\begin{aligned}
(k_1+k_2)u_1 - k_2 u_2 &= f_1\\
-k_2 u_1 + (k_2+k_3)u_2 - k_3 u_3 &= f_2\\
-k_3 u_2 + (k_3+k_4)u_3 - k_4 u_4 &= f_3\\
-k_4 u_3 + (k_4+k_5)u_4 - k_5 u_5 &= f_4\\
-k_5 u_4 + (k_5+k_6)u_5 &= f_5.
\end{aligned}
\qquad
\left[\begin{array}{ccccc|c}
k_1+k_2 & -k_2 & 0 & 0 & 0 & f_1\\
-k_2 & k_2+k_3 & -k_3 & 0 & 0 & f_2\\
0 & -k_3 & k_3+k_4 & -k_4 & 0 & f_3\\
0 & 0 & -k_4 & k_4+k_5 & -k_5 & f_4\\
0 & 0 & 0 & -k_5 & k_5+k_6 & f_5
\end{array}\right]
$$

Remark.

- The purpose of this example is not to find the precise solutions to these equations, but rather to show you that linear equations appear naturally in engineering.
- **Finite element method:** object is broken up into many small parts with connections between the different parts $\rightsquigarrow$ gigantic spring-mass system where the forces from spring correspond to the interaction between the different parts.
- In practice: millions of equations, not just five!
- Solve not for a single force vector $\mathbf{f}$, but for many different vectors. Thus the coefficient matrix stays the same $\rightsquigarrow$ LU-decomposition can make a difference in how quickly such systems can be solved.

LINEAR ALGEBRA

Inner Product and Orthogonality

ILLINOIS

Department of Mathematics

Definition. The **inner product** of $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$ is

$$\mathbf{v} \cdot \mathbf{w} = \mathbf{v}^T \mathbf{w}.$$

Example. If $\mathbf{v} = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}$ and $\mathbf{w} = \begin{bmatrix} w_1 \\ \vdots \\ w_n \end{bmatrix}$, then $\mathbf{v} \cdot \mathbf{w}$ is ...

Solution.

$$v_1 w_1 + v_2 w_2 + \cdots + v_n w_n$$

Question. Why is $\mathbf{v} \cdot \mathbf{w} = \mathbf{w} \cdot \mathbf{v}$?

Solution.

Note that $(\mathbf{v} \cdot \mathbf{w})^T = \mathbf{v} \cdot \mathbf{w}$, because it is a 1×1-matrices. Thus

$$\mathbf{v}^T \mathbf{w} = (\mathbf{v}^T \mathbf{w})^T = \mathbf{w}^T (\mathbf{v}^T)^T = \mathbf{w}^T \mathbf{v}$$

Question. Why is $\mathbf{v} \cdot \mathbf{v}$ always larger or equal to 0? For which $\mathbf{v}$ is $\mathbf{v} \cdot \mathbf{v} = 0$?

Solution.

$$\mathbf{v} \cdot \mathbf{v} = v_1 v_1 + \cdots + v_n v_n = \underbrace{v_1^2}_{\geq 0} + \cdots + \underbrace{v_n^2}_{\geq 0} \geq 0.$$

Thus $\mathbf{v} \cdot \mathbf{v} = 0$ if and only if $\mathbf{v} = \mathbf{0}$.

Theorem 22. *Let $\mathbf{u}, \mathbf{v}$ and $\mathbf{w}$ be vectors in $\mathbb{R}^n$, and let c be any scalar. Then*

(a) $\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$

(b) $(\mathbf{u} + \mathbf{v}) \cdot \mathbf{w} = \mathbf{u} \cdot \mathbf{w} + \mathbf{v} \cdot \mathbf{w}$

(c) $(c\mathbf{u}) \cdot \mathbf{v} = c(\mathbf{u} \cdot \mathbf{v}) = \mathbf{u} \cdot (c\mathbf{v})$

(d) $\mathbf{u} \cdot \mathbf{u} \geq \mathbf{0}$, *and* $\mathbf{u} \cdot \mathbf{u} = \mathbf{0}$ *if and only if* $\mathbf{u} = \mathbf{0}$.

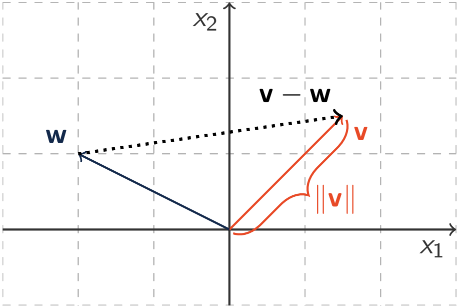
Definition. Let $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$.

The **norm** (or **length**) of $\mathbf{v}$ is

$$\|\mathbf{v}\| = \sqrt{\mathbf{v}\cdot\mathbf{v}} = \sqrt{v_1^2 + \cdots + v_n^2}.$$

The **distance** between $\mathbf{v}$ and $\mathbf{w}$ is

$$\operatorname{dist}(\mathbf{v}, \mathbf{w}) = \|\mathbf{v} - \mathbf{w}\|.$$



Example. Compute $\left\| \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix} \right\|$ and $\operatorname{dist}\left(\begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right)$.

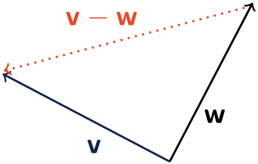
Solution.

$$\left\| \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix} \right\| = \sqrt{\begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix}} = \sqrt{2^2 + (-1)^2 + 1^2} = \sqrt{6}, \quad \operatorname{dist}\left(\begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right) = \left\| \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix} \right\| = \sqrt{6}.$$

Definition. Let $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$. We say $\mathbf{v}$ and $\mathbf{w}$ are **orthogonal** if $\mathbf{v} \cdot \mathbf{w} = 0$.

Theorem 23. *Let $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$ be non-zero. Then $\mathbf{v}$ and $\mathbf{w}$ are orthogonal if and only if they are perpendicular (ie. if they form a right angle).*

Proof.



$$\begin{aligned}
&\mathbf{v} \perp \mathbf{w} \\
&\iff \|\mathbf{v}\|^2 + \|\mathbf{w}\|^2 = \|\mathbf{v} - \mathbf{w}\|^2 \\
&\iff \mathbf{v} \cdot \mathbf{v} + \mathbf{w} \cdot \mathbf{w} = (\mathbf{v} - \mathbf{w}) \cdot (\mathbf{v} - \mathbf{w}) \\
&\iff \mathbf{v} \cdot \mathbf{v} + \mathbf{w} \cdot \mathbf{w} = \mathbf{v} \cdot \mathbf{v} - 2\mathbf{v} \cdot \mathbf{w} + \mathbf{w} \cdot \mathbf{w} \\
&\iff \mathbf{v} \cdot \mathbf{w} = 0
\end{aligned}$$

Example. Are $\begin{bmatrix}1\\1\end{bmatrix}, \begin{bmatrix}1\\-1\end{bmatrix}$ orthogonal? Are $\begin{bmatrix}1\\1\end{bmatrix}, \begin{bmatrix}1\\-2\end{bmatrix}$ orthogonal?

Solution.

$$\begin{bmatrix}1\\1\end{bmatrix} \cdot \begin{bmatrix}1\\-1\end{bmatrix} = 1 \cdot 1 + 1 \cdot (-1) = 0, \qquad \begin{bmatrix}1\\1\end{bmatrix} \cdot \begin{bmatrix}1\\-2\end{bmatrix} = 1 \cdot 1 + 1 \cdot (-2) = -1$$

Definition. A set of vectors in $\mathbb{R}^n$ is **pairwise orthogonal** if each pairing of them is orthogonal. We call such a set an **orthogonal set**.

Example. Find a non-zero $\mathbf{v} \in \mathbb{R}^3$ such that $\begin{bmatrix}1\\1\\0\end{bmatrix}, \begin{bmatrix}1\\-1\\0\end{bmatrix}$ and $\mathbf{v}$ form an orthogonal set?

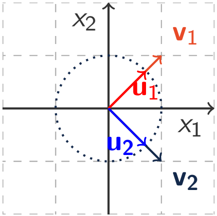
Solution.

Find $\mathbf{v} = \begin{bmatrix}v_1\\v_2\\v_3\end{bmatrix}$ such that $\begin{bmatrix}1\\1\\0\end{bmatrix} \cdot \mathbf{v} = \begin{bmatrix}1\\-1\\0\end{bmatrix} \cdot \mathbf{v} = 0$. So $v_1 + v_2 = 0$ and $v_1 - v_2 = 0$.

Thus $v_1 = v_2 = 0$. So we can take $v = \begin{bmatrix}0\\0\\1\end{bmatrix}$.

Definition. A **unit vector** in $\mathbb{R}^n$ is vector of length 1.

Example. Let $\mathbf{v} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$. Is $\mathbf{v}$ a unit vector?



Solution.

Since $\mathbf{v} \cdot \mathbf{v} = 5$ and $\|\mathbf{v}\| = \sqrt{5}$. However, $\mathbf{u} = \frac{\mathbf{v}}{\|\mathbf{v}\|} = \frac{\mathbf{v}}{\sqrt{5}}$ *is* a unit vector. The unit vector $\mathbf{u}$ is called the **normalization** of $\mathbf{v}$.

Definition. A set of vectors form an **orthonormal set** if this set is an orthogonal set and all vectors in the set are unit vectors.

Example. Let $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$. Do $\mathbf{v}_1$ and $\mathbf{v}_2$ form an orthonormal set?

Solution.

They are orthogonal but not orthonormal, since not unit vectors. Normalize to get a orthonormal set $\mathbf{u_1} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \mathbf{u_2} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$.

LINEAR ALGEBRA

Subspaces of $\mathbb{R}^n$

ILLINOIS

Department of Mathematics

Definition. A non-empty subset H of $\mathbb{R}^n$ is a **subspace** of $\mathbb{R}^n$ if it satisfies the following two conditions:

- If $\mathbf{u}, \mathbf{v} \in H$, then the sum $\mathbf{u} + \mathbf{v} \in H$. ($H$ is *closed under vector addition*).
- If $\mathbf{u} \in H$ and c is a scalar, then $c\mathbf{u} \in H$. (H is *closed under scalar multiplication.*)

Theorem 24. *Let $\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_m} \in \mathbb{R}^n$. Then Span$(\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_m})$ is a subspace of $\mathbb{R}^n$.*

Proof.

Let

$$\mathbf{u} = c_1\mathbf{v}_1 \cdots + c_m\mathbf{v_m}, \text{ and } \mathbf{w} = d_1\mathbf{v}_1 \cdots + d_m\mathbf{v_m}.$$

Then

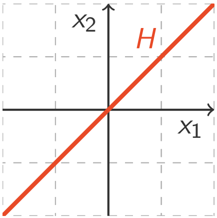
$$\mathbf{u} + \mathbf{w} = c_1\mathbf{v}_1 \cdots + c_m\mathbf{v_m} + d_1\mathbf{v}_1 \cdots + d_m\mathbf{v_m} = (c_1 + d_1)\mathbf{v}_1 \cdots + (c_m + d_m)\mathbf{v_m}.$$

and

$$c\mathbf{u} = c(c_1\mathbf{v}_1 \cdots + c_m\mathbf{v_m}) = cc_1\mathbf{v}_1 \cdots + cc_m\mathbf{v_m}.$$

Example. Is $H = \left\{ \begin{bmatrix} x \\ x \end{bmatrix} : x \in \mathbb{R} \right\}$ a subspace of $\mathbb{R}^2$?

Solution.



Let $\begin{bmatrix} a \\ a \end{bmatrix}, \begin{bmatrix} b \\ b \end{bmatrix}$ be in H.

- $\begin{bmatrix} a \\ a \end{bmatrix} + \begin{bmatrix} b \\ b \end{bmatrix} = \begin{bmatrix} a+b \\ a+b \end{bmatrix}$ is in H.
- $c \begin{bmatrix} a \\ a \end{bmatrix} = \begin{bmatrix} ca \\ ca \end{bmatrix}$ is in H.

Example. Let $Z = \left\{ \begin{bmatrix} 0 \\ 0 \end{bmatrix} \right\}$. Is Z a subspace of $\mathbb{R}^2$?

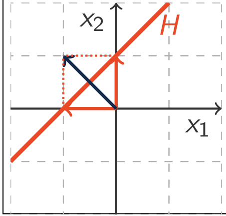
Solution.

Yes.

$$\begin{bmatrix} 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0+0 \\ 0+0 \end{bmatrix} \in Z, \quad c \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} c0 \\ c0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \in Z.$$

Example. Let $H = \left\{ \begin{bmatrix} x \\ x+1 \end{bmatrix} : x \in \mathbb{R} \right\}$. Is H a subspace of $\mathbb{R}^2$?

Solution.

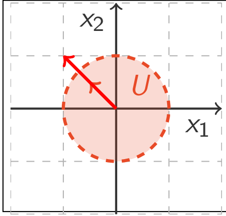


H is not a subspace of $\mathbb{R}^2$, because H is not closed under addition:

$$\begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \end{bmatrix} \in H, \text{ but } \begin{bmatrix} 0 \\ 1 \end{bmatrix} + \begin{bmatrix} -1 \\ 0 \end{bmatrix} \notin H.$$

Example. Is $U = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^2 \ : \ x^2 + y^2 < 1 \right\}$ a subspace of $\mathbb{R}^2$?

Solution.

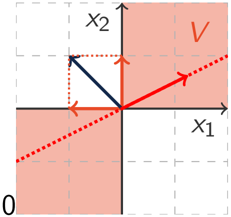


U is not a subspace, because U is not closed under scalar multiplication:

$$\begin{bmatrix} -0.5 \\ 0.5 \end{bmatrix} \in U, \text{ but } 2 \begin{bmatrix} -0.5 \\ 0.5 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \end{bmatrix} \notin U.$$

Example. Consider $V = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^2 \ : \ xy \geq 0 \right\}$.

Solution.



V is not a subspace, because it is not closed under addition:

$$\begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \end{bmatrix} \in V, \text{ but } \begin{bmatrix} 0 \\ 1 \end{bmatrix} + \begin{bmatrix} -1 \\ 0 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \end{bmatrix} \notin V.$$

But V is closed under scalar multiplication.

Question. Is $W = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^2 \ : \ xy = 0 \right\}$ a subspace?

Solution.

W is not a subspace, because it is not closed under addition:

$$\begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix} \in W, \text{ but } \begin{bmatrix} 0 \\ 1 \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \notin W.$$

LINEAR ALGEBRA

Column spaces and Nullspaces

ILLINOIS

Department of Mathematics

Definition. The **column space**, written as $\text{Col}(A)$, of an $m \times n$ matrix A is the set of all linear combinations of the columns of A. If $A = \begin{bmatrix} \mathbf{a_1} & \mathbf{a_2} & \cdots & \mathbf{a_n} \end{bmatrix}$, then
$\text{Col}(A) = \text{span}\,(\mathbf{a_1}, \mathbf{a_2}, \ldots, \mathbf{a_n})$.

Example. Describe the column space of $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$.

Solution.

$\text{Col}(A) = \text{span}(\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \end{bmatrix}) = \text{span}(\begin{bmatrix} 1 \\ 0 \end{bmatrix})$. This is the x_1 axis in $\mathbb{R}^2$.

Theorem 25. *Let A be an $m \times n$ matrix. Then $\text{Col}(A)$ is a subspace of $\mathbb{R}^m$.*

Proof.

$\text{Col}(A)$ is a span $\rightsquigarrow$ Spans are subspaces

It is a subspace of $\mathbb{R}^m$, because the columns of A are in $\mathbb{R}^m$.

Theorem 26. *Let A be an $m \times n$ matrix and $\mathbf{b} \in \mathbb{R}^m$. Then $\mathbf{b}$ is in $\mathrm{Col}(A)$ if and only if the linear system $A\mathbf{x} = \mathbf{b}$ has a solution.*

Proof.

Let $A = \begin{bmatrix} \mathbf{a}_1 & \dots \mathbf{a}_n \end{bmatrix}$. Suppose $A\mathbf{x} = \mathbf{b}$, where $\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$.

$$\mathbf{b} = A\mathbf{x} = \underbrace{x_1\mathbf{a_1} + x_2\mathbf{a_2} + \cdots + x_n\mathbf{a_n}}_{(\textit{lin. comb. of } \mathbf{a_1},\ldots,\mathbf{a_n})}$$

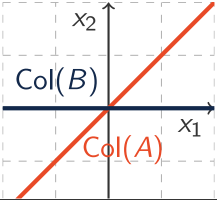
Thus:
$\mathbf{b}$ is a linear combination of the columns of A $\iff$ $A\mathbf{x} = \mathbf{b}$ is consistent.

Question. Let A and B be two row-equivalent matrices. Is $\mathrm{Col}(A) = \mathrm{Col}(B)$?

Solution.

No. Take $A = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$.

Then $\mathrm{Col}(A) = \mathrm{span}(\begin{bmatrix} 1 \\ 1 \end{bmatrix}) \neq \mathrm{span}(\begin{bmatrix} 1 \\ 0 \end{bmatrix}) = \mathrm{Col}(B)$.



Definition. The **nullspace** of an $m \times n$ matrix A, written as $\mathrm{Nul}(A)$, is the set of all solutions to the homogeneous equation $A\mathbf{x} = \mathbf{0}$; that is, $\mathrm{Nul}(A) = \{\mathbf{v} \in \mathbb{R}^n \ : \ A\mathbf{v} = \mathbf{0}\}$.

Theorem 27. *The nullspace of an $m \times n$ matrix A is a subspace of $\mathbb{R}^n$.*

Proof.

$\mathrm{Nul}(A)$ is non-empty since $\mathbf{0} \in \mathrm{Nul}(A)$. Suppose $A\mathbf{u} = \mathbf{0}$ and $A\mathbf{v} = \mathbf{0}$. Then

$$A(\mathbf{u} + \mathbf{v}) = A\mathbf{u} + A\mathbf{v} = \mathbf{0} + \mathbf{0} = \mathbf{0}.$$

$$A(c\mathbf{u}) = c\,(A\mathbf{u}) = c\,(\mathbf{0}) = \mathbf{0}.$$

Thus $\mathrm{Nul}(A)$ is closed under addition and scalar multiplication.

Example. Let $H := \left\{ \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} \ : \ v_1 + v_2 - v_3 = 0 \right\}$. Find a matrix A such that $H = \mathrm{Nul}(A)$.

Solution.

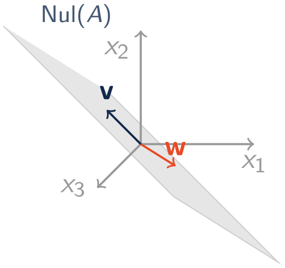
$v_1 + v_2 - v_3 = 0$ if and only if $\begin{bmatrix} 1 & 1 & -1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = 0$. Thus $\mathrm{Nul}\left(\begin{bmatrix} 1 & 1 & -1 \end{bmatrix}\right) = H$.

Example. Let $A = \begin{bmatrix} 1 & 1 & -1 \end{bmatrix}$. Find two vectors $\mathbf{v}, \mathbf{w}$ such that $\mathrm{Nul}(A) = \mathrm{span}(\mathbf{v}, \mathbf{w})$.

Solution.

Let $u \in \mathrm{Nul}(A)$. Then

$$\begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} -u_2 + u_3 \\ u_2 \\ u_3 \end{bmatrix} = u_2 \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + u_3 \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}.$$



Thus $\mathrm{Nul}(A) = \mathrm{span}\left(\overbrace{\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}}^{=:\mathbf{v}}, \overbrace{\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}}^{=:\mathbf{w}} \right)$

Question. Is there a matrix B such that $\mathrm{Nul}(A) = \mathrm{Col}(B)$?

Solution.

$$\text{Yes. Set } B = \begin{bmatrix} \mathbf{v} & \mathbf{w} \end{bmatrix} = \begin{bmatrix} -1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Theorem 28. *Let A be an $m \times n$ matrix, let $\mathbf{b} \in \mathbb{R}^m$, and let $\mathbf{w} \in \mathbb{R}^n$ such that $A\mathbf{w} = \mathbf{b}$. Then $\{\mathbf{v} \in \mathbb{R}^n \ : \ A\mathbf{v} = \mathbf{b}\} = \mathbf{w} + \mathrm{Nul}(A)$.*

Proof.

Let $\mathbf{v} \in \mathbb{R}^n$ be such that $A\mathbf{v} = \mathbf{b}$. Then

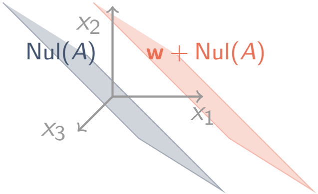
$$A(\mathbf{v} - \mathbf{w}) = A\mathbf{v} - A\mathbf{w} = \mathbf{b} - \mathbf{b} = \mathbf{0}.$$

Therefore, $\mathbf{v} - \mathbf{w} \in \mathrm{Nul}(A)$. Thus $\mathbf{v} \in \mathbf{w} + \mathrm{Nul}(A)$.

Example. Let $A = \begin{bmatrix} 1 & 1 & -1 \end{bmatrix}$ and let $\mathbf{b} = 1$. Observe that $A \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}^T = \mathbf{b}$. Use this to describe $\{\mathbf{v} \in \mathbb{R}^n \ : \ A\mathbf{v} = \mathbf{b}\}$.

Solution.

$$\{\mathbf{v} \in \mathbb{R}^n \ : \ A\mathbf{v} = \mathbf{b}\}$$
$$= \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + \mathrm{Nul}(A) = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + \mathrm{span}\left(\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \right).$$



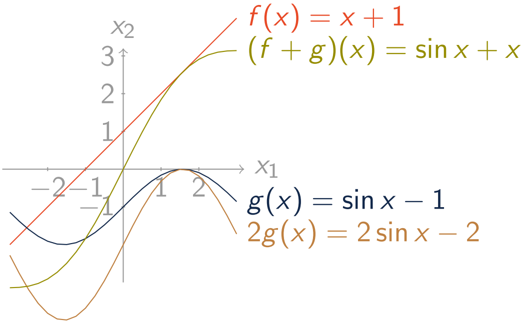
LINEAR ALGEBRA

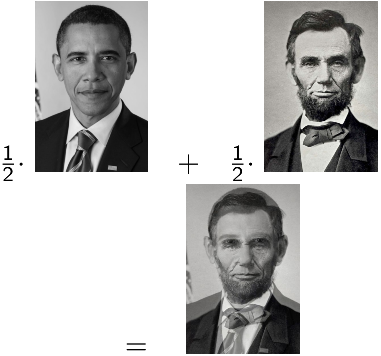
Abstract vector spaces

ILLINOIS

Department of Mathematics

- The most important property of column vectors in $\mathbb{R}^n$ is that you can take *linear combinations* of them.
- There are many mathematical objects $X, Y, \ldots$ for which a linear combination $cX + dY$ make sense, and have the usual properties of linear combination in $\mathbb{R}^n$.
- We are going to define a *vector space* in general as a collection of objects for which linear combinations make sense. The objects of such a set are called *vectors*.





Definition. A **vector space** is a non-empty set V of objects, called **vectors**, for which linear combinations make sense. More precisely: on V there are defined two operations, called *addition* and *multiplication* by scalars (real numbers),satisfying the following axioms for all $u, v, w \in V$ and for all scalars $c, d \in \mathbb{R}$:

- $\mathbf{u} + \mathbf{v}$ is in V. (V is "closed under addition".)
- $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$.
- $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$.
- There is a vector (called the zero vector) $\mathbf{0}_V$ in V such that $\mathbf{u} + \mathbf{0}_V = \mathbf{u}$.
- For each $\mathbf{u}$ in V, there is a vector $-\mathbf{u}$ in V satisfying $\mathbf{u} + (-\mathbf{u}) = \mathbf{0}_V$.
- $c\mathbf{u}$ is in V. (V is "closed under scalar multiplication".)
- $c(\mathbf{u} + \mathbf{v}) = c\mathbf{u} + c\mathbf{v}$.
- $(c + d)\mathbf{u} = c\mathbf{u} + d\mathbf{u}$.
- $(cd)\mathbf{u} = c(d\mathbf{u})$.
- $1\mathbf{u} = \mathbf{u}$.

In particular, we may talk about linear combinations and span within a vector space, e.g. $3\mathbf{u} + 2\mathbf{v}$ or $\text{span}(\mathbf{u}, \mathbf{v})$.

Example. Explain how the set of functions $\mathbb{R} \to \mathbb{R}$ is a vector space.

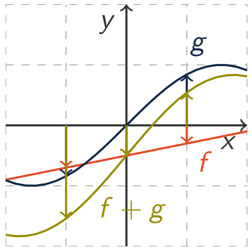
Solution.

Let f, g be two function from $\mathbb{R}$ to $\mathbb{R}$. We define $f + g$ by

$$(f+g)(x) = f(x) + g(x).$$

For a scalar c define cf by

$$(cf)(x) = cf(x).$$



Example. Explain how the set of all 2×2 matrices is a vector space.

Solution.

Addition:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} + \begin{bmatrix} e & f \\ g & h \end{bmatrix} = \begin{bmatrix} a+e & b+f \\ c+g & d+h \end{bmatrix}$$

Scalar Multiplication:

$$e \cdot \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} ea & eb \\ ec & ed \end{bmatrix}$$

Definition. Let V be a vector space. A non-empty subset $W \subseteq V$ is a **subspace** of V if

- $\mathbf{u} + \mathbf{v} \in W$ for all $\mathbf{u}, \mathbf{v} \in W$ (*closed under addition*)
- $c\mathbf{u} \in W$ for all $\mathbf{u} \in W$ and $c \in \mathbb{R}$ (*closed under scalar multiplication*)

Example. Explain why the set of all symmetric 2×2 matrices is a subspace of the vector space of 2×2 matrices?

Solution.

The set of symmetric matrices of a given size is non-empty since the zero matrix is symmetric. Let A, B be two symmetric 2×2 matrices. So $A^T = A$ and $B^T = B$.

$$(A+B)^T = A^T + B^T = A + B \qquad (cA)^T = cA^T = cA.$$

Thus the set is closed under addition and scalar multiplication.

Question. Is the set of all invertible 2×2 matrices a subspace of the vector space of 2×2 matrices?

Solution.

No. The set is not closed under addition:

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

Example. Let $\mathbb{P}_n$ be the set of all polynomials of degree at most n, that is

$$\mathbb{P}_n = \{a_0 + a_1 t + a_2 t^2 + \cdots + a_n t^n \ : \ a_0, \ldots, a_n \in \mathbb{R}\}.$$

Explain why it is a vector space. Can you think of it as a subspace of the vector space of all functions $\mathbb{R} \to \mathbb{R}$?

Solution.

Let $\mathbf{p}(t) = a_0 + a_1 t + \cdots + a_n t^n$ and $\mathbf{q}(t) = b_0 + b_1 t + \cdots + b_n t^n$ be two polynomials of degree at most n. Then

$$(p+q)(t) = (a_0 + b_0) + (a_1 + b_1)t + \cdots + (a_n + b_n)t^n$$

is also a polynomial of degree at most n. And

$$(c\mathbf{p})(t) = (ca_0) + (ca_1)t + \cdots + (ca_n)t^n$$

is also a polynomial of degree at most n.

LINEAR ALGEBRA

Linear Independence

ILLINOIS

Department of Mathematics

Definition. Vectors $\mathbf{v}_1, \ldots, \mathbf{v}_p$ are said to be **linearly independent** if the equation

$$x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_p\mathbf{v}_p = \mathbf{0}$$

has only the trivial solution (namely, $x_1 = x_2 = \cdots = x_p = 0$).
We say the vectors are **linearly dependent** if they are not linearly independent.

Example. Are $\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 2 \end{bmatrix}$ linearly independent?

Solution.

No, the equation

$$x_1 \begin{bmatrix} 1 \\ 1 \end{bmatrix} + x_2 \begin{bmatrix} 2 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

has the solution $(-2, 1)$.

The problem:
$\begin{bmatrix} 2 \\ 2 \end{bmatrix} \in \operatorname{span} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

Theorem 29. *Vectors $\mathbf{v}_1, \ldots, \mathbf{v}_p$ are linear dependent if and only if there is $i \in \{1, \ldots, p\}$ such that $\mathbf{v}_i \in \operatorname{span}(\mathbf{v}_1, \ldots, \mathbf{v}_{i-1}, \mathbf{v}_{i+1}, \ldots, \mathbf{v}_p)$.*

Question. A single non-zero vector $\mathbf{v}_1$ is always linearly independent. Why?

Solution.

Because $x_1\mathbf{v}_1 = \mathbf{0}$ only for $x_1 = 0$.

Question. Two vectors $\mathbf{v}_1, \mathbf{v}_2$ are linearly independent if and only if neither of the vectors is a multiple of the other. Why?

Solution.

Because if $x_1\mathbf{v_1} + x_2\mathbf{v_2} = \mathbf{0}$ with, say, $x_2 \neq 0$, then $\mathbf{v_2} = -\frac{x_1}{x_2}\mathbf{v_1}$.

Question. Vectors $\mathbf{v_1}, \ldots, \mathbf{v_p}$ containing the zero vector are linearly dependent.

Solution.

Because if, say, $\mathbf{v_1} = \mathbf{0}$, then $\mathbf{v_1} + 0\mathbf{v_2} + \cdots + 0\mathbf{v_p} = \mathbf{0}$.

Example. Consider $\begin{bmatrix}1\\1\\1\end{bmatrix}, \begin{bmatrix}1\\2\\3\end{bmatrix}, \begin{bmatrix}-1\\1\\3\end{bmatrix}$. Are these vectors linearly independent?

Solution.

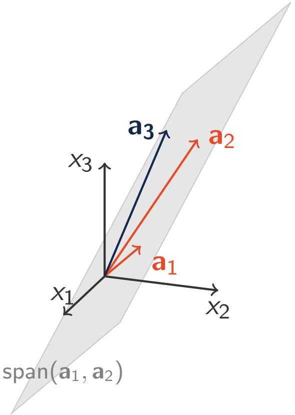
Solve: $$x_1\begin{bmatrix}1\\1\\1\end{bmatrix} + x_2\begin{bmatrix}1\\2\\3\end{bmatrix} + x_3\begin{bmatrix}-1\\1\\3\end{bmatrix} = \begin{bmatrix}0\\0\\0\end{bmatrix}$$

Solve corresponding linear system:

$$\left[\begin{array}{ccc|c}1 & 1 & -1 & 0\\ 1 & 2 & 1 & 0\\ 1 & 3 & 3 & 0\end{array}\right] \underset{RREF}{\rightsquigarrow} \left[\begin{array}{ccc|c}1 & 0 & -3 & 0\\ 0 & 1 & 2 & 0\\ 0 & 0 & 0 & 0\end{array}\right]$$

Solution: x_3 is free. $x_2 = -2x_3$, and $x_1 = 3x_3$. Take $x_3 = 1$:

$$3\begin{bmatrix}1\\1\\1\end{bmatrix} - 2\begin{bmatrix}1\\2\\3\end{bmatrix} + \begin{bmatrix}-1\\1\\3\end{bmatrix} = \begin{bmatrix}0\\0\\0\end{bmatrix}.$$



Question. In the previous example we determined the following linear dependence:

$$3\begin{bmatrix}1\\1\\1\end{bmatrix} - 2\begin{bmatrix}1\\2\\3\end{bmatrix} + \begin{bmatrix}-1\\1\\3\end{bmatrix} = \mathbf{0}.$$

Can you write this as a matrix equation?

Solution.

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

Idea. The Null space determines linear (in)dependence!

Theorem 30. *Let A be an $m \times n$ matrix. The following are equivalent:*

- *The columns of A are linearly independent.*
- *$A\mathbf{x} = \mathbf{0}$ has only the solution $\mathbf{x} = \mathbf{0}$.*
- *A has n pivots.*
- *there are no free variables for $A\mathbf{x} = \mathbf{0}$.*

Question. Let $\mathbf{v}_1, \ldots, \mathbf{v}_n \in \mathbb{R}^m$. If $n > m$, then $\mathbf{v}_1, \ldots \mathbf{v}_n$ are linearly dependent. Why?

Solution.

Let $A = \begin{bmatrix} \mathbf{v_1} & \ldots & \mathbf{v_n} \end{bmatrix}$. This is an $m \times n$ matrix.

$\rightsquigarrow$ The matrix can have at most m pivots.

$\rightsquigarrow$ Since $n > m$, A does not have n pivots.

$\rightsquigarrow$ By Theorem, the columns of A have to be linearly dependent.

Question. Consider an $m \times n$-matrix A in echelon form. The pivot columns of A are linearly independent. Why?

Solution.

For example, let $A = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3 \end{bmatrix}$ and suppose that $\mathbf{a}_1, \mathbf{a}_3$ are the pivot columns.

$\rightsquigarrow$ The matrix $\begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_3 \end{bmatrix}$ has 2 pivots.

$\rightsquigarrow$ The vectors $\mathbf{a}_1, \mathbf{a}_3$ are linearly independent.

LINEAR ALGEBRA

Basis and Dimension

ILLINOIS

Department of Mathematics

Definition. Let V be a vector space. A sequence of vectors $(\mathbf{v}_1, \ldots, \mathbf{v}_p)$ in V is a **basis** of V if

- $V = \operatorname{span}(\mathbf{v}_1, \ldots, \mathbf{v}_p)$, and
- $(\mathbf{v}_1, \ldots, \mathbf{v}_p)$ are linearly independent.

Example. Check that both $\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right)$ and $\left(\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right)$ are bases of $\mathbb{R}^2$.

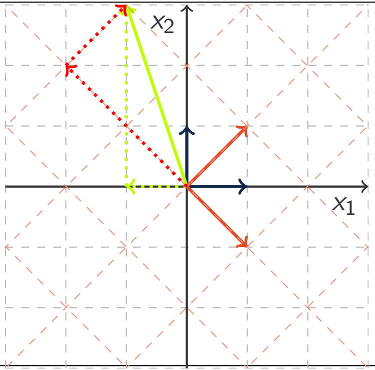
Solution.

Linear independence:

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \rightsquigarrow 2 \text{ pivots}, \qquad \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \rightsquigarrow \begin{bmatrix} 1 & 1 \\ 0 & -2 \end{bmatrix} \rightsquigarrow 2 \text{ pivots}.$$

Spanning set:

$$\begin{bmatrix} a \\ b \end{bmatrix} = a \begin{bmatrix} 1 \\ 0 \end{bmatrix} + b \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \frac{a+b}{2} \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \frac{a-b}{2} \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$



Theorem 31. *Every two bases in a vector space V contain the same numbers of vectors.*

Definition. The number of vectors in a basis of V is the **dimension** of V.

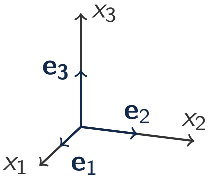
Example. What is the dimension of $\mathbb{R}^n$?

Solution.

Let

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \ldots, \quad \mathbf{e}_n = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$$

Note that $(\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$ is a basis of $\mathbb{R}^n$:



$$\begin{bmatrix} \mathbf{e}_1 & \ldots & \mathbf{e_n} \end{bmatrix} = I_n \rightsquigarrow \text{ n pivots}, \qquad \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} = v_1\mathbf{e_1} + \cdots + v_n\mathbf{e_n}.$$

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

Theorem 32. *Suppose that V has dimension d. Then*

- *A sequence of d vectors in V are a basis if they span V.*
- *A sequence of d vectors in V are a basis if they are linearly independent.*

Proof.

- Take d vectors that span V, say $\mathbf{v}_1, \dots \mathbf{v}_d$.
  So $V = \operatorname{span}(\mathbf{v}_1, \dots \mathbf{v}_{\mathbf{d}})$.
  Suppose $\mathbf{v}_1, \dots \mathbf{v}_d$ are not linearly independent.
  Then one of the them, say $\mathbf{v}_d$, is in the span of the other vectors.
  So $\mathbf{v}_d \in \operatorname{span}(\mathbf{v}_1, \dots, \mathbf{v}_{d-1})$.
  Drop $\mathbf{v}_d$!
  Remaining $d-1$ vectors still span V. Thus $V = \operatorname{span}(\mathbf{v}_1, \dots, \mathbf{v}_{d-1})$.
  Repeat until the remaining vectors are linearly independent.
  Eventually, we have $\mathbf{v}_1, \dots, \mathbf{v}_\ell$ linearly independent such that $\operatorname{span}(\mathbf{v}_1, \dots, \mathbf{v}_\ell) = V$.
  Contradicts $\dim V = d$, since $\ell < d$.

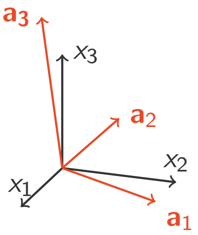
Example. Is $\left(\begin{bmatrix}1\\2\\0\end{bmatrix}, \begin{bmatrix}0\\1\\1\end{bmatrix}, \begin{bmatrix}1\\0\\3\end{bmatrix}\right)$ a basis of $\mathbb{R}^3$?

Solution.

The set has 3 elements.

Since $\dim \mathbb{R}^3 = 3$, it is a basis if and only if the vectors are linearly independent.

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



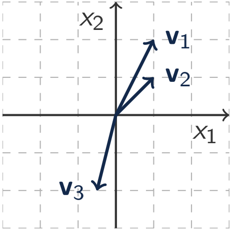
Since each column contains a pivot, the three vectors are linearly independent.

Hence, $\left(\begin{bmatrix}1\\2\\0\end{bmatrix}, \begin{bmatrix}0\\1\\1\end{bmatrix}, \begin{bmatrix}1\\0\\3\end{bmatrix}\right)$ a basis of $\mathbb{R}^3$.

Theorem 33. *A basis is a minimal spanning set of V; that is the elements of the basis span V but you cannot delete any of these elements and still get all of V.*

Example. Produce a basis of $\mathbb{R}^2$ from the vectors

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix} -.5 \\ -2 \end{bmatrix}.$$



Solution.

Three vectors in $\mathbb{R}^2$ have to be linearly dependent.
Here, we notice that $\mathbf{v}_2 = 1.5\mathbf{v}_1 + \mathbf{v}_3$.
The remaining vectors $\{\mathbf{v}_1, \mathbf{v}_3\}$ are a basis for $\mathbb{R}^2$, because the two vectors are clearly linearly independent.

Example. Produce a basis of $\mathbb{R}^2$ from the vector $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$.

Solution.

Add any vector not in span $\left(\begin{bmatrix} 2 \\ 1 \end{bmatrix}\right)$. For example, $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$. So $\left(\begin{bmatrix} 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix}\right)$ is a basis.

LINEAR ALGEBRA

Bases and dimensions of the four fundamental subspaces

ILLINOIS

Department of Mathematics

Algorithm. To find a basis for $\mathrm{Nul}(A)$:

- Find the parametric form of the solutions to $A\mathbf{x} = \mathbf{0}$.
- Express solutions $\mathbf{x}$ as a linear combination of vectors with the free variables as coefficients.
- Use these vectors as a basis of $\mathrm{Nul}(A)$.

Example. Find a basis for $\mathrm{Nul}(A)$ where $A = \begin{bmatrix} 3 & 6 & 6 & 3 & 9 \\ 6 & 12 & 15 & 0 & 3 \end{bmatrix}$.

Solution.

$$\left[\begin{array}{ccccc|c} 3 & 6 & 6 & 3 & 9 & 0 \\ 6 & 12 & 15 & 0 & 3 & 0 \end{array}\right] \underset{\text{RREF}}{\rightsquigarrow} \left[\begin{array}{ccccc|c} 1 & 2 & 0 & 5 & 13 & 0 \\ 0 & 0 & 1 & -2 & -5 & 0 \end{array}\right] \rightsquigarrow \begin{aligned} x_1 &= -2x_2 - 5x_4 - 13x_5 \\ x_3 &= 2x_4 + 5x_5 \end{aligned}$$

The solutions to $A\mathbf{x} = \mathbf{0}$ are:

$$\begin{bmatrix} -2x_2 - 5x_4 - 13x_5 \\ x_2 \\ 2x_4 + 5x_5 \\ x_4 \\ x_5 \end{bmatrix} = x_2 \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} -5 \\ 0 \\ 2 \\ 1 \\ 0 \end{bmatrix} + x_5 \begin{bmatrix} -13 \\ 0 \\ 5 \\ 0 \\ 1 \end{bmatrix} \rightsquigarrow \text{basis: } \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -5 \\ 0 \\ 2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -13 \\ 0 \\ 5 \\ 0 \\ 1 \end{bmatrix}$$

Definition. The **rank** of a matrix is the number of pivots it has.

Theorem 34. *Let A be an $m \times n$ matrix with rank r. Then* $\dim \operatorname{Nul}(A) = n - r$.

Remark. Let $A = \begin{bmatrix} \mathbf{a}_1 & \dots \mathbf{a}_n \end{bmatrix}$ and let $U = \begin{bmatrix} \mathbf{u}_1 & \dots \mathbf{u}_n \end{bmatrix}$ be an echelon form of A. Explain why

$$x_1\mathbf{u_1} + \cdots + x_n\mathbf{u_n} = \mathbf{0} \iff x_1\mathbf{a_1} + \cdots + x_n\mathbf{a_n} = \mathbf{0}.$$

Solution.

Because A and U are row-equivalent:

$$x_1\mathbf{u_1} + \cdots + x_n\mathbf{u_n} = \mathbf{0} \iff U\mathbf{x} = \mathbf{0} \iff A\mathbf{x} = \mathbf{0} \iff x_1\mathbf{a_1} + \cdots + x_n\mathbf{a_n} = \mathbf{0}$$

Theorem 35. *Let A be an $m \times n$ matrix with rank r. The pivot columns of A form a basis of* $\operatorname{Col}(A)$. *In particular,* $\dim \operatorname{Col}(A) = r$.

Proof.

Let U be RREF of A.

$\rightsquigarrow$ Then the pivot columns of U are linearly independent.

$\rightsquigarrow$ By Remark, pivot columns of A are linearly independent.

Example. Find a basis for $\mathrm{Col}(A)$ where $A = \begin{bmatrix} 1 & 2 & 0 & 4 \\ 2 & 4 & -1 & 3 \\ 3 & 6 & 2 & 22 \\ 4 & 8 & 0 & 16 \end{bmatrix}$.

Solution.

$$A = \begin{bmatrix} 1 & 2 & 0 & 4 \\ 2 & 4 & -1 & 3 \\ 3 & 6 & 2 & 22 \\ 4 & 8 & 0 & 16 \end{bmatrix} \underset{\text{RREF}}{\rightsquigarrow} U = \begin{bmatrix} 1 & 2 & 0 & 4 \\ 0 & 0 & 1 & 5 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

This are the pivot columns, the first and third.

Hence, a basis for $\mathrm{Col}(A)$ is $\left\{ \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \\ 2 \\ 0 \end{bmatrix} \right\}$. An aside: Note that for U we have column $\mathbf{u}_2 = 2\mathbf{u_1}$ and $\mathbf{u_4} = 4\mathbf{u_1} + 5\mathbf{u_3}$.
The same is true for the columns of A!

Example. Let $A = \begin{bmatrix} 1 & 3 \\ 2 & 6 \end{bmatrix}$. The RREF of A is $U = \begin{bmatrix} 1 & 3 \\ 0 & 0 \end{bmatrix}$. Is $\text{Col}(A) = \text{Col}(U)$? Is $\text{Col}(A^T) = \text{Col}(U^T)$?

Solution.

The second columns of A and U are redundant:

$$\text{Col}(A) = \text{span}\left(\begin{bmatrix} 1 \\ 2 \end{bmatrix}\right) \neq \text{span}\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) = \text{Col}(U)$$

Row operations do not preserve the column space.

The second rows of A and U are redundant:

$$\text{Col}(A^T) = \text{span}\left(\begin{bmatrix} 1 \\ 3 \end{bmatrix}\right) = \text{span}\left(\begin{bmatrix} 1 \\ 3 \end{bmatrix}\right) = \text{Col}(U^T)$$

Row operations preserve the row space.

Theorem 36. *Let A, B be two row-equivalent matrices. Then $\text{Col}(A^T) = \text{Col}(B^T)$.*

Theorem 37. *Let A be $m \times n$ matrix with rank r. Then the non-zero rows in an echelon form of A form a basis of $\text{Col}(A^T)$, and thus $\dim(A^T) = r$.*

Example. Find a basis for $\mathrm{Col}(A^T)$ where $A = \begin{bmatrix} 1 & 2 & 0 & 4 \\ 2 & 4 & -1 & 3 \\ 3 & 6 & 2 & 22 \\ 4 & 8 & 0 & 16 \end{bmatrix}$.

Solution.

$$A = \begin{bmatrix} 1 & 2 & 0 & 4 \\ 2 & 4 & -1 & 3 \\ 3 & 6 & 2 & 22 \\ 4 & 8 & 0 & 16 \end{bmatrix} \underset{\text{RREF}}{\rightsquigarrow} U = \begin{bmatrix} 1 & 2 & 0 & 4 \\ 0 & 0 & 1 & 5 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \rightsquigarrow \left(\begin{bmatrix} 1 \\ 2 \\ 0 \\ 4 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 5 \end{bmatrix} \right) \text{ basis of } \mathrm{Col}(A^T).$$

Theorem 38. *Let A be an $m \times n$ matrix with rank r. Then*

- $\dim \mathrm{Col}(A) = \dim \mathrm{Col}(A^T) = r$.
- $\dim \mathrm{Nul}(A) = n - r$.

Question. What is $\dim \mathrm{Nul}(A^T)$? What is $\dim \mathrm{Col}(A) + \dim \mathrm{Nul}(A)$?

Solution.

$$\dim \mathrm{Nul}(A^T) = m - \dim \mathrm{Col}(A^T) = m - r. \qquad \dim \mathrm{Col}(A) + \dim \mathrm{Nul}(A) = r + (n - r) = n.$$

LINEAR ALGEBRA

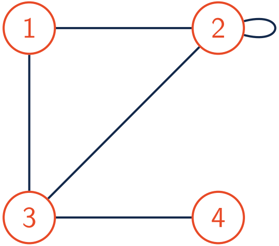
Graphs

ILLINOIS

Department of Mathematics

Definition. A **graph** is a set of nodes (or: vertices) that are connected through edges.

Example. A graph with 4 nodes and 5 edges:



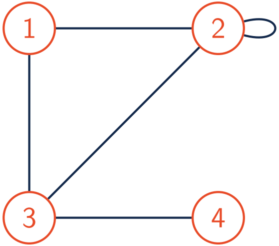
$$\begin{bmatrix} 0 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

Definition. Let $\mathcal{G}$ be a graph with n nodes. The **adjacency matrix** of $\mathcal{G}$ is the $n \times n$-matrix $A = (a_{ij})$ such that

$$a_{ij} = \begin{cases} 1 & \text{if there is an edge between node } i \text{ and node } j \\ 0 & \text{otherwise .} \end{cases}$$

Definition. A **walk of length** k on a graph of is a sequence of $k+1$ vertices and k edges between two nodes (including the start and end) that may repeat.A **path** is walk in which all vertices are distinct.

Example. Count the number of walks of length 2 from node 2 to node 3 and the number of walks of length 3 from node 3 back to node 3:



- Node 2 to Node 3: 2 walks of length 2
- Node 3 to Node 3: 3 walks of length 3

Definition. A graph is **connected** if for every pair of nodes i and j there is a walk from node i to node j. A graph is **disconnected** if it is not connected.

Theorem 39. *Let $\mathcal{G}$ be a graph and let A be its adjacency matrix. Then the entry in the i-th row and j-th column of A^ℓ is the number of walks of length ℓ from node j to node i on $\mathcal{G}$.*

Proof.

We just consider the case $\ell = 2$. Then the entry in the i-th row and j-th column of A^2 is:

$$(A^2)_{ij} = \begin{bmatrix} a_{i1} & \dots & a_{in} \end{bmatrix} \begin{bmatrix} a_{j1} \\ \vdots \\ a_{jn} \end{bmatrix} = a_{i1}a_{1j} + a_{i2}a_{2j} + \cdots + a_{in}a_{nj}$$

Each summand corresponds to a walk of length $\ell = 2$ from node j to i. For example:

$a_{i2}a_{2j} = 1 \iff$ there is an edge between node j and node 2 and an edge between node 2 and i.

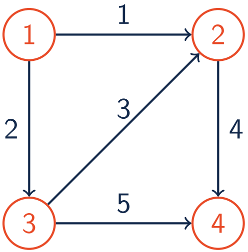
Since each walk from node j to node i has go through some node,the sum

$$a_{i1}a_{1j} + a_{i2}a_{2j} + \cdots + a_{in}a_{nj}$$

is equal to the number of walks of length $\ell = 2$ from node j to node i.

Definition. A **directed graph** is a set of vertices connected by edges, where the edges have a direction associated with them.

Example. A graph with 4 nodes and 5 edges:



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

We can also talk about adjacency matrices of directed graphs. We use the following convention:

Definition. Let G be a directed graph with m edges and n nodes. The **adjacency matrix** of G is the $n \times n$ matrix $A = (a_{i,j})_{i,j}$ with

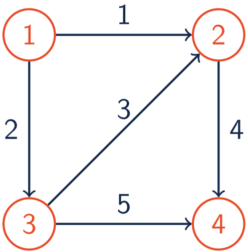
$$a_{i,j} = \begin{cases} 1, & \text{if there is a directed edge from node } j \text{ to node } i \\ 0, & \text{otherwise} \end{cases}$$

Directed graphs have another important associated matrix:

Definition. Let G be a directed graph with m edges and n nodes. The **edge-node incidence matrix** of G is the $m \times n$ matrix $A = (a_{i,j})_{i,j}$ with

$$a_{i,j} = \begin{cases} -1, & \text{if edge } i \text{ leaves node } j \\ +1, & \text{if edge } i \text{ enters node } j \\ 0, & \text{otherwise} \end{cases}$$

Example. A graph with 4 nodes and 5 edges:

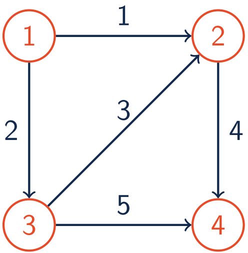


$$\begin{bmatrix} -1 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & -1 & 0 & 1 \\ 0 & 0 & -1 & 1 \end{bmatrix}$$

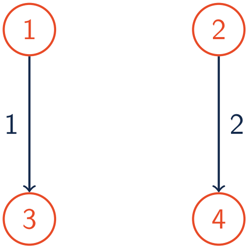
Definition. A **connected component** of an undirected graph is a part in which any two vertices are connected to each other by paths, and which is connected to no additional vertices in the rest of the graph. The connected components of a directed graph are those of its underlying undirected graph. A graph is **connected** if only has one connected component.

Example.

A graph with one connected component:

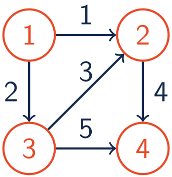


A graph with two connected components:



Theorem 40. *Let $\mathcal{G}$ be a directed graph and let A be its edge-node incidence matrix. Then* $\dim \operatorname{Nul}(A)$ *is equal to the number of connected components of $\mathcal{G}$.*

Example. We explain the idea in the case of the following graph $\mathcal{G}$. Let A be its edge-node incidence matrix, a 5×4-matrix. Note that $\operatorname{Nul}(A) \subseteq \mathbb{R}^4$.



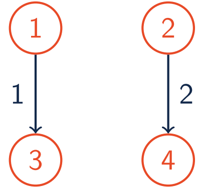
$$\begin{bmatrix} -1 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & -1 & 0 & 1 \\ 0 & 0 & -1 & 1 \end{bmatrix}$$

Think of $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$ as assigning to potential to each node. If $\mathbf{x} \in \operatorname{Nul}(A)$:

$$\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} = A\mathbf{x} = \begin{bmatrix} -1 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & -1 & 0 & 1 \\ 0 & 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} -x_1 + x_2 \\ -x_1 + x_3 \\ -x_2 + x_3 \\ -x_2 + x_4 \\ -x_3 + x_4 \end{bmatrix} \rightsquigarrow \mathbf{x} = \begin{bmatrix} c \\ c \\ c \\ c \end{bmatrix}$$

Each edge forces the potential of the nodes it connects to be equal. Thus if two nodes are connected by a path, they have the same assigned potential.

Example. Find a basis of the null space of the edge-node incidence matrix of the following graph:



Solution.

Let A be the edge-node incidence matrix. It is a 2×4-matrix. Let $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} \in \mathbb{R}^4$.

If $A\mathbf{x} = 0$, then

$$\underbrace{x_1 = x_3}_{\text{connected by an edge}} \quad \text{and} \quad \underbrace{x_2 = x_4}_{\text{connected by an edge}}$$

$\rightsquigarrow$ Nul(A) has the following basis: $\begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix}$

Just to make sure:

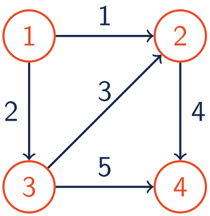
$$A = \begin{bmatrix} -1 & 0 & 1 & 0 \\ 0 & -1 & 0 & 1 \end{bmatrix}$$

$$\underset{RREF}{\rightsquigarrow} \begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \end{bmatrix}$$

Definition. A **cycle** in an undirected graph is a path in which all edges are distinct and the only repeated vertices are the first and last vertices. By cycles of a directed graph we mean those of its underlying undirected graph.

Example. Find cycles in the following graph.

Solution.



$$edge_2, edge_3, -edge_1 \rightsquigarrow \begin{bmatrix} -1 \\ 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

$$edge_5, -edge_4, -edge_3 \rightsquigarrow \begin{bmatrix} 0 \\ 0 \\ -1 \\ -1 \\ 1 \end{bmatrix}$$

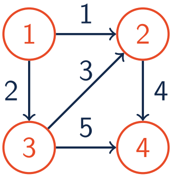
$$edge_2, edge_5, -edge_4, -edge_1 \rightsquigarrow \begin{bmatrix} -1 \\ 1 \\ 0 \\ -1 \\ 1 \end{bmatrix}$$

$$\rightsquigarrow \text{Cycle1} + \text{Cycle2} = \text{Cycle3}$$

Definition. The span of all cycle vectors of a graph $\mathcal{G}$ is called the **cycle space**.

Theorem 41. *Let $\mathcal{G}$ be a directed graph and let A be its edge-node incidence matrix. Then the cycle space of $\mathcal{G}$ is equal to* $\operatorname{Nul}(A^T)$.

Example. We explain the idea in the case of the following graph $\mathcal{G}$. Let A be its edge-node incidence matrix, a 5×4-matrix. Note that $\operatorname{Nul}(A^T) \subseteq \mathbb{R}^5$.



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

Think of $\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \end{bmatrix}$ as assigning a flow to each edge. If $\mathbf{y} \in \operatorname{Nul}(A^T)$:

$$\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} = A^T\mathbf{y} = \begin{bmatrix} -1 & -1 & 0 & 0 & 0 \\ 1 & 0 & 1 & -1 & 0 \\ 0 & 1 & -1 & 0 & -1 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \end{bmatrix} = \begin{bmatrix} -y_1 - y_2 \\ y_1 + y_3 - y_4 \\ y_2 - y_3 - y_5 \\ y_4 + y_5 \end{bmatrix}$$

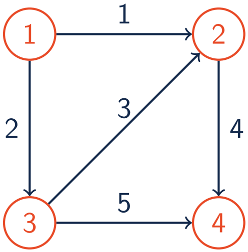
$\rightsquigarrow$ $\mathbf{y} \in \operatorname{Nul}(A^T)$ if and only if the inflow equals the outflow at each node.

What is the simplest way to balance flow? Assign flow around cycles.

Example (ctd.).

Let's solve $A^T\mathbf{y} = \mathbf{0}$.

$$A^T = \begin{bmatrix} -1 & -1 & 0 & 0 & 0 \\ 1 & 0 & 1 & -1 & 0 \\ 0 & 1 & -1 & 0 & -1 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix} \underset{RREF}{\leadsto} \begin{bmatrix} 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & -1 & 0 & -1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$



$$\mathbf{y} \in \mathrm{Nul}(A^T) \leadsto y_1 = -y_3 - y_5,\ y_2 = y_3 + y_5,\ y_4 = -y_5 \leadsto \mathbf{y} = \begin{bmatrix} -y_3 - y_5 \\ y_3 + y_5 \\ y_3 \\ -y_5 \\ y_5 \end{bmatrix} = y_3 \underset{\text{Cycle1}}{\begin{bmatrix} -1 \\ 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}} + y_5 \underset{\text{Cycle3}}{\begin{bmatrix} -1 \\ 1 \\ 0 \\ -1 \\ 1 \end{bmatrix}}$$

$\leadsto$ Cycle1 and Cycle3 form a basis of the cycle space of $\mathcal{G}$.

LINEAR ALGEBRA

Orthogonal complements

ILLINOIS

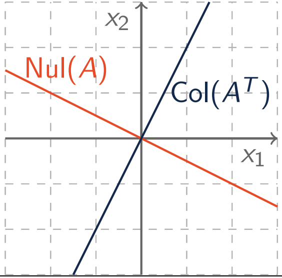
Department of Mathematics

Example. Let $A = \begin{bmatrix} 1 & 2 \\ 2 & 4 \\ 3 & 6 \end{bmatrix}$. Find bases of $\mathrm{Nul}(A)$ and $\mathrm{Col}(A^T)$.

Solution.

$$A \underset{RREF}{\rightsquigarrow} \begin{bmatrix} 1 & 2 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

$$\mathrm{Nul}(A) = \mathrm{span}\left\{ \begin{bmatrix} -2 \\ 1 \end{bmatrix} \right\}, \quad \mathrm{Col}(A^T) = \mathrm{span}\left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix} \right\}$$



Definition. Let W be a subspace of $\mathbb{R}^n$. The **orthogonal complement** of W is the subspace $W^\perp$ of all vectors that are orthogonal to W; that is

$$W^\perp := \{\mathbf{v} \in \mathbb{R}^n \; : \; \mathbf{v} \cdot \mathbf{w} = 0 \text{ for all } \mathbf{w} \in W\}.$$

Remark. Observe that $(W^\perp)^\perp = W$.

Theorem 42. *Let A be an $m \times n$ matrix. Then* $\mathrm{Nul}(A)$ *is the orthogonal complement of* $\mathrm{Col}(A^T)$*; that is* $\mathrm{Nul}(A) = \mathrm{Col}(A^T)^\perp$.

Proof.

Suppose $\mathbf{u} \in \mathrm{Nul}(A)$ and $\mathbf{v} \in \mathrm{Col}(A^T)$. Then $A\mathbf{u} = \mathbf{0}$ and there is $\mathbf{w} \in \mathbb{R}^m$ such that $\mathbf{v} = A^T\mathbf{w}$.

$$\rightsquigarrow \mathbf{v} \cdot \mathbf{u} = (A^T\mathbf{w})^T\mathbf{u} = \mathbf{w}^T A\mathbf{u} = \mathbf{w}^{\mathbf{T}}\mathbf{0} = 0.$$

Suppose that $\mathbf{v} \in \mathrm{Col}(A^T)^\perp$. Let $A^T = \begin{bmatrix} \mathbf{R}_1 & \dots & \mathbf{R}_m \end{bmatrix}$. Then

$$(A\mathbf{v})^T = \mathbf{v}^T \begin{bmatrix} \mathbf{R}_1 & \mathbf{R}_2 & \dots & \mathbf{R}_m \end{bmatrix} = \begin{bmatrix} \mathbf{v}^T\mathbf{R}_1 & \mathbf{v}^T\mathbf{R}_2 & \dots & \mathbf{v}^T\mathbf{R}_m \end{bmatrix} = \begin{bmatrix} 0 & 0 & \dots & 0 \end{bmatrix}.$$

So $A\mathbf{v} = \mathbf{0}$. Thus $\mathbf{v} \in \mathrm{Nul}(A)$.

Remark. It follows that

- $\mathrm{Nul}(A)^\perp = \mathrm{Col}(A^T)$.
- $\mathrm{Nul}(A^T) = \mathrm{Col}(A)^\perp$.

Theorem 43. *Let V be a subspace of $\mathbb{R}^n$. Then* $\dim V + \dim V^\perp = n$.

Proof.

Write $V = \mathrm{Col}(A)$ for some $n \times m$ matrix A.

$$\dim V + \dim V^\perp = \dim \mathrm{Col}(A) + \dim \mathrm{Nul}(A^T) = \dim \mathrm{Col}(A^T) + \dim \mathrm{Nul}(A^T) = n$$

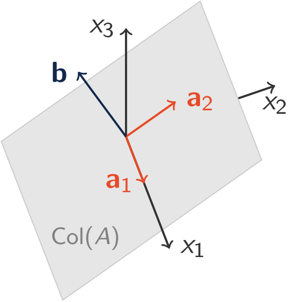
Example. Find a basis of the orthogonal complement of span $\left(\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \right)$.

Solution.

Strategy: Find A such that span is $\mathrm{Col}(A)$. Then find a basis of $\mathrm{Nul}(A^T)$.

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} \underset{\text{Transpose}}{\rightsquigarrow} A^T = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix}$$

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ -x_3 \\ x_3 \end{bmatrix} = x_3 \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} \rightsquigarrow \left(\begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} \right) \text{ basis of } \mathrm{Nul}(A^T).$$



LINEAR ALGEBRA

Coordinates

ILLINOIS

Department of Mathematics

Theorem 44. *Let $(\mathbf{v}_1, \ldots, \mathbf{v}_p)$ be a basis of V. Then every vector $\mathbf{w}$ in V can be expressed uniquely as*

$$\mathbf{w} = c_1\mathbf{v}_1 + \cdots + c_p\mathbf{v}_p.$$

Proof.

Let $c_1, \ldots, c_p$ and $d_1, \ldots, d_p$ such that $w = c_1\mathbf{v}_1 + \cdots + c_p\mathbf{v}_p = d_1\mathbf{v}_1 + \cdots + d_p\mathbf{v}_p$.

$$\begin{aligned} &\implies (c_1 - d_1)\mathbf{v}_1 + \cdots + (c_p - d_p)\mathbf{v}_p = \mathbf{0}. \\ &\implies c_1 - d_1 = \cdots = c_p - d_p = 0 \implies c_1 = d_1, \ldots, c_p = d_p. \end{aligned}$$

Definition. Let $\mathcal{B} = (\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p)$ be an (ordered) basis of V, and let $\mathbf{w} \in V$.The **coordinate vector $\mathbf{w}_{\mathcal{B}}$** of $\mathbf{w}$ with respect to the basis $\mathcal{B}$ is

$$\mathbf{w}_{\mathcal{B}} = \begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ \vdots \\ c_p \end{bmatrix}, \quad \text{if } \mathbf{w} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_p\mathbf{v}_p.$$

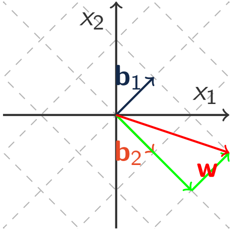
Example. Let $V = \mathbb{R}^2$, and consider the bases

$$\mathcal{B} := \left(\mathbf{b}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \mathbf{b}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right), \qquad \mathcal{E} := \left(\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right).$$

Let $\mathbf{w} = \begin{bmatrix} 3 \\ -1 \end{bmatrix}$. Determine $\mathbf{w}_{\mathcal{B}}$ and $\mathbf{w}_{\mathcal{E}}$.

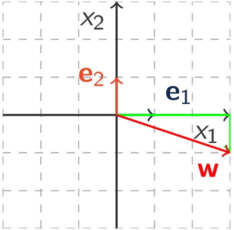
Solution.

$$\begin{bmatrix} 3 \\ -1 \end{bmatrix} = c_1 \begin{bmatrix} 1 \\ 1 \end{bmatrix} + c_2 \begin{bmatrix} 1 \\ -1 \end{bmatrix} = 1 \begin{bmatrix} 1 \\ 1 \end{bmatrix} + 2 \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$



$$\rightsquigarrow \mathbf{w}_{\mathcal{B}} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

$$\begin{bmatrix} 3 \\ -1 \end{bmatrix} = c_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + c_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix} = 3 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + (-1) \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$



$$\rightsquigarrow \mathbf{w}_{\mathcal{E}} = \begin{bmatrix} 3 \\ -1 \end{bmatrix}$$

Definition. In $\mathbb{R}^n$ let $\mathbf{e}_i$ denote the vector with a 1 in the i-th coordinate and 0's elsewhere. The **standard basis of** $\mathbb{R}^n$ is the ordered basis $\mathcal{E}_n := (\mathbf{e}_1, \ldots, \mathbf{e}_n)$.

Question. For all $\mathbf{v} \in \mathbb{R}^n$, we have $\mathbf{v} = \mathbf{v}_{\mathcal{E}_n}$. Why?

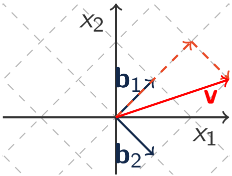
Solution.

$$\mathbf{v} = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} = v_1\mathbf{e}_1 + \cdots + v_n\mathbf{e}_n \rightsquigarrow \mathbf{v}_{\mathcal{E}_n} = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}$$

Example. Consider the basis $\mathcal{B} := \left(\mathbf{b}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \mathbf{b}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}\right)$ of $\mathbb{R}^2$. Let $\mathbf{v} \in \mathbb{R}^2$ be such that $\mathbf{v}_{\mathcal{B}} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$. Can you determine $\mathbf{v}$?

Solution.

$$\mathbf{v}_{\mathcal{B}} = \begin{bmatrix} 2 \\ 1 \end{bmatrix} \rightsquigarrow \mathbf{v} = 2\mathbf{b}_1 + 1\mathbf{b}_2 = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$$



Definition. Let $\mathcal{B}$ and $\mathcal{C}$ be two bases of $\mathbb{R}^n$. The **change of basis matrix** $I_{\mathcal{C},\mathcal{B}}$ is the matrix such that for all $\mathbf{v} \in \mathbb{R}^n$

$$I_{\mathcal{C},\mathcal{B}}\mathbf{v}_{\mathcal{B}} = \mathbf{v}_{\mathcal{C}}$$

Theorem 45. *Let $\mathcal{B} = (\mathbf{b}_1, \ldots, \mathbf{b_n})$ be a basis of $\mathbb{R}^n$. Then*

$$I_{\mathcal{E}_n,\mathcal{B}} = \begin{bmatrix} \mathbf{b}_1 & \ldots & \mathbf{b}_n \end{bmatrix}$$

That is, for all $\mathbf{v} \in \mathbb{R}^n$,

$$\mathbf{v} = \begin{bmatrix} \mathbf{b}_1 & \ldots & \mathbf{b}_n \end{bmatrix} \mathbf{v}_{\mathcal{B}}.$$

Proof.

Let $\mathbf{v}_{\mathcal{B}} = \begin{bmatrix} v_{B,1} \\ \vdots \\ v_{B,n} \end{bmatrix}$. Then

$$\begin{bmatrix} \mathbf{b}_1 & \ldots & \mathbf{b}_n \end{bmatrix} \mathbf{v}_{\mathcal{B}} = \begin{bmatrix} \mathbf{b}_1 & \ldots & \mathbf{b}_n \end{bmatrix} \begin{bmatrix} v_{B,1} \\ \vdots \\ v_{B,n} \end{bmatrix} = v_{B,1}\mathbf{b}_1 + \cdots + v_{B,n}\mathbf{b}_n = \mathbf{v}.$$

Question. Let $\mathcal{B} = (\mathbf{b_1}, \ldots, \mathbf{b_n})$ be a basis of $\mathbb{R}^n$. How do you compute $I_{\mathcal{B},\mathcal{E}_n}$?

Solution.

$$\mathbf{v} = I_{\mathcal{E}_n,\mathcal{B}}\mathbf{v}_{\mathcal{B}} \rightsquigarrow I_{\mathcal{E}_n,\mathcal{B}}^{-1}\mathbf{v} = \mathbf{v}_{\mathcal{B}} \rightsquigarrow I_{\mathcal{B},\mathcal{E}_n} = I_{\mathcal{E}_n,\mathcal{B}}^{-1}$$

Question. Let $\mathcal{B}$ and $\mathcal{C}$ be two bases of $\mathbb{R}^n$. How do you compute $I_{\mathcal{B},\mathcal{C}}$?

Solution.

$$I_{\mathcal{B},\mathcal{E}_n} I_{\mathcal{E}_n,\mathcal{C}}\mathbf{v}_{\mathcal{C}} = I_{\mathcal{B},\mathcal{E}_n}\mathbf{v} = \mathbf{v}_{\mathcal{B}} \rightsquigarrow I_{\mathcal{B},\mathcal{C}} = I_{\mathcal{B},\mathcal{E}_n} I_{\mathcal{E}_n,\mathcal{C}}$$

Remark. Another way to compute $I_{\mathcal{C},\mathcal{B}}$:

$$I_{\mathcal{C},\mathcal{B}} = \begin{bmatrix} (\mathbf{b}_1)_{\mathcal{C}} & \ldots & (\mathbf{b}_n)_{\mathcal{C}} \end{bmatrix}$$

LINEAR ALGEBRA

Orthogonal and Orthonormal bases

ILLINOIS

Department of Mathematics

Theorem 46. *Let $\mathbf{v}_1, \ldots, \mathbf{v}_m \in \mathbb{R}^n$ be non-zero and pairwise orthogonal. Then $\mathbf{v}_1, \ldots, \mathbf{v}_m$ are linearly independent.*

Proof.

Suppose that

$$c_1\mathbf{v_1} + \cdots + c_m\mathbf{v_m} = \mathbf{0}.$$

Take the inner product of $\mathbf{v_1}$ on both sides.

$$\begin{aligned} 0 &= \mathbf{v_1} \cdot (c_1\mathbf{v_1} + \cdots + c_m\mathbf{v_m}) \\ &= c_1\mathbf{v_1}\cdot\mathbf{v_1} + c_2\mathbf{v_1}\cdot\mathbf{v_2} + \cdots + c_m\mathbf{v_1}\cdot\mathbf{v_m} \\ &= c_1\mathbf{v_1}\cdot\mathbf{v_1} = c_1\|\mathbf{v_1}\|^2 \end{aligned}$$

But $\|\mathbf{v_1}\| \neq 0$ and so $c_1 = 0$. Similarly, we find that $c_2 = 0, \ldots, c_m = 0$. Therefore, the vectors are linearly independent.

Remark. The theorem implies that a set of n orthonormal vectors in $\mathbb{R}^n$ is a basis of $\mathbb{R}^n$.

Definition. An **orthogonal basis** (an **orthonormal basis**) is an orthogonal set of vectors (an orthonormal set of vectors) that forms a basis.

Theorem 47. *Let $\mathcal{B} := (\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n)$ be an orthogonal basis of $\mathbb{R}^n$, and let $\mathbf{v} \in \mathbb{R}^n$. Then*

$$\mathbf{v} = \frac{\mathbf{v} \cdot \mathbf{b}_1}{\mathbf{b}_1 \cdot \mathbf{b}_1}\mathbf{b}_1 + \ldots + \frac{\mathbf{v} \cdot \mathbf{b}_n}{\mathbf{b}_n \cdot \mathbf{b}_n}\mathbf{b}_n.$$

Solution.

Let $c_1, \ldots, c_n$ be such that

$$\mathbf{v} = c_1\mathbf{b}_1 + \cdots + c_n\mathbf{b}_n.$$

Just calculate

$$\begin{aligned} \mathbf{b_1} \cdot \mathbf{v} &= \mathbf{b_1} \cdot (c_1\mathbf{b_1} + c_2\mathbf{b_2} + \cdots + c_n\mathbf{b_n}) \\ &= c_1\mathbf{b_1} \cdot \mathbf{b_1} + c_2\mathbf{b_1} \cdot \mathbf{b_2} + \cdots + c_n\mathbf{b_1} \cdot \mathbf{b_n} \\ &= c_1\mathbf{b}_1 \cdot \mathbf{b}_1 \rightsquigarrow c_1 = \frac{\mathbf{v} \cdot \mathbf{b}_1}{\mathbf{b}_1 \cdot \mathbf{b}_1} \end{aligned}$$

Remark. When $\mathcal{B}$ is orthonormal, then $\mathbf{b}_i \cdot \mathbf{b}_i = 1$ for $i = 1, \ldots, n$.

Example. Let $\mathcal{U}$ be the orthonormal basis $\left(\mathbf{u}_1 = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\1\end{bmatrix}, \mathbf{u}_2 = \frac{1}{\sqrt{2}}\begin{bmatrix}-1\\1\end{bmatrix}\right)$ of $\mathbb{R}^2$. Let $\mathbf{v} = \begin{bmatrix}2\\3\end{bmatrix}$. Determine $\mathbf{v}_{\mathcal{U}}$ using the formula from the previous theorem!

Solution.

$$\mathbf{v}_{\mathcal{U}} = \begin{bmatrix}\mathbf{u_1}\cdot\mathbf{v}\\ \mathbf{u_2}\cdot\mathbf{v}\end{bmatrix} = \begin{bmatrix}\frac{5}{\sqrt{2}}\\ \frac{1}{\sqrt{2}}\end{bmatrix}$$

Example. With $\mathcal{U}$ as above, compute the change of basis matrix $I_{\mathcal{U},\mathcal{E}_2}$?

Solution.

$$I_{\mathcal{U},\mathcal{E}_2} = I_{\mathcal{E}_2,\mathcal{U}}^{-1} = \left(\frac{1}{\sqrt{2}}\begin{bmatrix}1 & -1\\1 & 1\end{bmatrix}\right)^{-1} = \frac{1}{\sqrt{2}}\begin{bmatrix}1 & 1\\-1 & 1\end{bmatrix} = I_{\mathcal{E}_2,\mathcal{U}}^T$$

Theorem 48. *Let $\mathcal{U} = (\mathbf{u_1}, \ldots, \mathbf{u_n})$ be an orthonormal basis of $\mathbb{R}^n$. Then*

$$I_{\mathcal{U},\mathcal{E}_n} = \begin{bmatrix} \mathbf{u}_1 & \ldots & \mathbf{u}_n \end{bmatrix}^T.$$

Proof.

$$\mathbf{v}_{\mathcal{U}} = \begin{bmatrix} \mathbf{u}_1 \cdot \mathbf{v} \\ \vdots \\ \mathbf{u}_n \cdot \mathbf{v} \end{bmatrix} = \begin{bmatrix} \mathbf{u}_1^T \mathbf{v} \\ \vdots \\ \mathbf{u}_n^T \mathbf{v} \end{bmatrix} = \begin{bmatrix} \mathbf{u}_1^T \\ \vdots \\ \mathbf{u}_n^T \end{bmatrix} \mathbf{v} \rightsquigarrow I_{\mathcal{U},\mathcal{E}_n} = \begin{bmatrix} \mathbf{u}_1 & \ldots & \mathbf{u}_n \end{bmatrix}^T$$

Definition. An $n \times n$-matrix Q is **orthogonal** if $Q^{-1} = Q^T$.

Remark. The columns of an orthogonal matrix form an orthonormal basis. Why?

Solution.

$$I_n = Q^T Q = \begin{bmatrix} \mathbf{q}_1^T \\ \vdots \\ \mathbf{q}_n^T \end{bmatrix} \begin{bmatrix} \mathbf{q}_1 & \ldots & \mathbf{q}_n \end{bmatrix} = \begin{bmatrix} \mathbf{q}_1^T \mathbf{q}_1 & \mathbf{q}_1^T \mathbf{q}_2 & \cdots \\ \mathbf{q}_2^T \mathbf{q}_1 & \mathbf{q}_2^T \mathbf{q}_2 & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix} \rightsquigarrow \mathbf{q}_i \cdot \mathbf{q}_j = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j \end{cases}$$

LINEAR ALGEBRA

Linear Transformation

ILLINOIS

Department of Mathematics

Definition. Let V and W be vector spaces. A map $T : V \to W$ is a **linear transformation** if

$$T(a\mathbf{v} + b\mathbf{w}) = aT(\mathbf{v}) + bT(\mathbf{w})$$

for all $\mathbf{v}, \mathbf{w} \in V$ and all $a, b \in \mathbb{R}$.

Example. Let $V = \mathbb{R}, W = \mathbb{R}$. Then the map $f(x) = 3x$ is linear, $g(x) = 2x - 2$ is not. Why?

Solution.

If $x, y \in \mathbb{R}$, then

$$f(ax + by) = 3(ax + by) = a \cdot 3x + b \cdot 3y = af(x) + bf(y).$$

What about the function $g(x) = 2x - 2$?

$$g(0) + g(0) = 2 \cdot 0 - 2 + 2 \cdot 0 - 2 = -4 \neq -2 = g(0) = g(0 + 0).$$

Remark. $T(\mathbf{0}_V) = T(0 \cdot \mathbf{0}_V) = 0 \cdot T(\mathbf{0}_V) = \mathbf{0}_W \rightsquigarrow T(\mathbf{0}_V) = \mathbf{0}_W$

Question. Let A be an $m \times n$ matrix, and consider the map $T : \mathbb{R}^n \to \mathbb{R}^m$ given by $T(\mathbf{v}) := A\mathbf{v}$. Is this a linear transformation?

Solution.

Because matrix multiplication is linear.

$$A(c\mathbf{v} + d\mathbf{w}) = cA\mathbf{v} + dA\mathbf{w}$$

The left-hand side is $T(c\mathbf{v} + d\mathbf{w})$ and the right-hand side is $cT(\mathbf{v}) + dT(\mathbf{w})$.

Example. Let $\mathbb{P}_n$ be the vector space of all polynomials of degree at most n. Consider the map $T : \mathbb{P}_n \to \mathbb{P}_{n-1}$ given by $T(p(t)) := \frac{d}{dt}p(t)$. This map is linear! Why?

Solution.

Because differentiation is linear:

$$\frac{d}{dt}\left[ap(t) + bq(t)\right] = a\frac{d}{dt}p(t) + b\frac{d}{dt}q(t).$$

The left-hand side is $T(ap(t) + bq(t))$ and the right-hand side is $aT(p(t)) + bT(q(t))$.

Theorem 49. *Let V, W be two vector spaces, let $T : V \to W$ be a linear transformation and let $(\mathbf{v}_1, \ldots, \mathbf{v}_n)$ be a basis of V. Then T is completely determined by the values $T(\mathbf{v}_1), \ldots, T(\mathbf{v}_n)$.*

Solution.

Take any $\mathbf{v} \in V$.

It can be written as $\mathbf{v} = c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n$ because $(\mathbf{v}_1, \ldots, \mathbf{v}_n)$ is a basis and hence spans V.

Hence by the linearity of T:

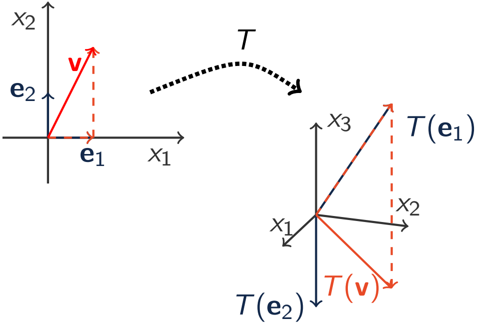
$$T(\mathbf{v}) = T(c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n) = c_1 T(\mathbf{v}_1) + \cdots + c_n T(\mathbf{v}_n).$$

So we know how to write $T(\mathbf{v})$ as long as we know $T(\mathbf{v}_1), \ldots, T(\mathbf{v}_n)$.

Example. Let $T : \mathbb{R}^2 \to \mathbb{R}^3$ be a linear transformation with $T\left(\begin{bmatrix}1\\0\end{bmatrix}\right) = \begin{bmatrix}1\\2\\3\end{bmatrix}$ and $T\left(\begin{bmatrix}0\\1\end{bmatrix}\right) = \begin{bmatrix}0\\0\\-2\end{bmatrix}$. What is $T\left(\begin{bmatrix}1\\2\end{bmatrix}\right)$?

Solution.

$$\begin{aligned}
T\left(\begin{bmatrix}1\\2\end{bmatrix}\right) &= T\left(1\begin{bmatrix}1\\0\end{bmatrix} + 2\begin{bmatrix}0\\1\end{bmatrix}\right)\\
&= T\left(\begin{bmatrix}1\\0\end{bmatrix}\right) + 2T\left(\begin{bmatrix}0\\1\end{bmatrix}\right)\\
&= \begin{bmatrix}1\\2\\3\end{bmatrix} + 2\begin{bmatrix}0\\0\\-2\end{bmatrix} = \begin{bmatrix}1\\2\\-1\end{bmatrix}
\end{aligned}$$



Theorem 50. *Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. Then there is a $m \times n$ matrix A such that*

- $T(\mathbf{v}) = A\mathbf{v}, \quad$ *for all* $\mathbf{v} \in \mathbb{R}^n$.
- $A = \begin{bmatrix} T(\mathbf{e}_1) & T(\mathbf{e}_2) & \dots & T(\mathbf{e}_n) \end{bmatrix}$, *where* $(\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n)$ *is the standard basis of* $\mathbb{R}^n$.

Proof.

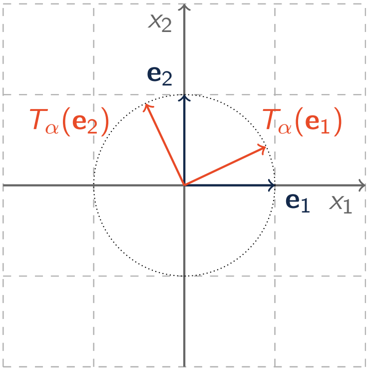
Let $\mathbf{v} = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}$. We can write $\mathbf{v} = v_1\mathbf{e}_1 + v_2\mathbf{e}_2 + \cdots + v_n\mathbf{e}_n$. Then

$$\begin{aligned} T(\mathbf{v}) &= T(v_1\mathbf{e}_1 + v_2\mathbf{e}_2 + \cdots + v_n\mathbf{e}_n) \\ &= v_1 T(\mathbf{e}_1) + v_2 T(\mathbf{e}_2) + \cdots + v_n T(\mathbf{e}_n) \\ &= \begin{bmatrix} T(\mathbf{e}_1) & T(\mathbf{e}_2) & \dots & T(\mathbf{e}_n) \end{bmatrix} \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} = A\mathbf{v}. \end{aligned}$$

Remark. We call this A the **coordinate matrix of** T with respect to the standard bases - we write $T_{\mathcal{E}_m, \mathcal{E}_n}$.

Example. Let $T_\alpha : \mathbb{R}^2 \to \mathbb{R}^2$ be the "rotation over α radians (counterclockwise)" map, that is $T_\alpha(\mathbf{v})$ is the vector obtained by rotating $\mathbf{v}$ over angle α. Find the 2×2 matrix A_α such that $T_\alpha(\mathbf{v}) = A_\alpha \mathbf{v}$ for all $\mathbf{v} \in \mathbb{R}^2$.

Solution.



We just need to find what happens under (counterclockwise) rotation to the standard basis vectors. By high school trigonometry,

$$T_\alpha\left(\begin{bmatrix}1\\0\end{bmatrix}\right) = \begin{bmatrix}\cos(\alpha)\\ \sin(\alpha)\end{bmatrix}, \quad T_\alpha\left(\begin{bmatrix}0\\1\end{bmatrix}\right) = \begin{bmatrix}-\sin(\alpha)\\ \cos(\alpha)\end{bmatrix},$$

So our matrix is $A_\alpha = \begin{bmatrix}\cos(\alpha) & -\sin(\alpha)\\ \sin(\alpha) & \cos(\alpha)\end{bmatrix}$.

This is called the rotation matrix for angle α.

LINEAR ALGEBRA

Coordinate matrices of linear transformations

ILLINOIS

Department of Mathematics

Last time: *Linear transformation is matrix multiplication*. For every linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$, there is an $m \times n$ matrix A such that $T(\mathbf{v}) = A\mathbf{v}$.

Today: The same in abstract vector spaces.

Theorem 51. *Let V, W be two vector space, let $\mathcal{B} = (\mathbf{b}_1, \ldots, \mathbf{b}_n)$ be a basis of V and $\mathcal{C} = (\mathbf{c}_1, \ldots, \mathbf{c}_m)$ be a basis of W, and let $T\colon V \to W$ be a linear transformation. Then there is a $m \times n$ matrix $T_{\mathcal{C},\mathcal{B}}$ such that*

- $T(\mathbf{v})_{\mathcal{C}} = T_{\mathcal{C},\mathcal{B}}\mathbf{v}_{\mathcal{B}}$, *for all* $\mathbf{v} \in V$.
- $T_{\mathcal{C},\mathcal{B}} = \begin{bmatrix} T(\mathbf{b}_1)_{\mathcal{C}} & T(\mathbf{b}_2)_{\mathcal{C}} & \ldots & T(\mathbf{b}_n)_{\mathcal{C}} \end{bmatrix}$.

$$\begin{tikzcd}
\mathbf{v} : \text{vector in } V \arrow[r, "\text{apply } T"] \arrow[d, "\text{write in coordinates wrt } \mathcal{B}"'] & \text{vector in } W \; : \; T(\mathbf{v}) \arrow[d, "\text{write in coordinates wrt } \mathcal{C}"] \\
\mathbf{v}_{\mathcal{B}} \; : \; \text{coordinate vector in } \mathbb{R}^n \arrow[r, "\text{multiply by } T_{\mathcal{C},\mathcal{B}}"] & \text{coordinate vector in } \mathbb{R}^m \; : \; T_{\mathcal{C},\mathcal{B}}\mathbf{v}
\end{tikzcd}$$

Example. Let $D : \mathbb{P}_2 \to \mathbb{P}_1$ be given by $D(p(t)) = \frac{d}{dt}p(t)$. Consider the bases $\mathcal{B} = (1, t, t^2)$ and $\mathcal{C} = (1, t)$ of $\mathbb{P}_2$ and $\mathbb{P}_1$. Determine $D_{\mathcal{C},\mathcal{B}}$.

Solution.

$$D(1) = 0 = 0 \cdot 1 + 0 \cdot t, D(t) = 1 = 1 \cdot 1 + 0 \cdot t, D(t^2) = 2t = 0 \cdot 1 + 2 \cdot t$$

$$D_{\mathcal{C},\mathcal{B}} = \begin{bmatrix} D(1)_{\mathcal{C}} & D(t)_{\mathcal{C}} & D(t^2)_{\mathcal{C}} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

Question. Consider $p(t) = 2 - t + 3t^2$ in $\mathbb{P}_2$. Compute $D(p(t))_{\mathcal{C}}$ and $D_{\mathcal{C},\mathcal{B}}p(t)_{\mathcal{B}}$.

Solution.

$$D(2 - t + 3t^2) = -1 + 6t \rightsquigarrow D(p(t))_{\mathcal{C}} = \begin{bmatrix} -1 \\ 6 \end{bmatrix}$$

$$p(t)_{\mathcal{B}} = \begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix} \rightsquigarrow D_{\mathcal{C},\mathcal{B}}p(t)_{\mathcal{B}} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix} = \begin{bmatrix} -1 \\ 6 \end{bmatrix}$$

Example. Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be such that $T(\mathbf{v}) = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix} \mathbf{v}$. Consider $\mathcal{B} := \left(\mathbf{b}_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \mathbf{b}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right)$. Compute $T_{\mathcal{B},\mathcal{B}}$. Let $\mathbf{v} = \mathbf{b}_1 + \mathbf{b}_2$. Use $T_{\mathcal{B},\mathcal{B}}$ to compute $T(\mathbf{v})$.

Solution.

$$T_{\mathcal{B},\mathcal{B}} = \begin{bmatrix} T(\mathbf{b_1})_{\mathcal{B}} & T(\mathbf{b_2})_{\mathcal{B}} \end{bmatrix}.$$

$$T(\mathbf{b_1}) = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = 2 \begin{bmatrix} 1 \\ -1 \end{bmatrix} = 2\mathbf{b}_1 + 0\mathbf{b}_2,$$

$$T(\mathbf{b_2}) = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = 4 \begin{bmatrix} 1 \\ 1 \end{bmatrix} = 0\mathbf{b}_1 + 4\mathbf{b}_2$$

This means that the coordinate matrix is simply

$$T_{\mathcal{B},\mathcal{B}} = \begin{bmatrix} 2 & 0 \\ 0 & 4 \end{bmatrix}$$

Note that $\mathbf{v}_{\mathcal{B}} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

$$\begin{aligned} T(\mathbf{v})_{\mathcal{B}} &= T_{\mathcal{B},\mathcal{B}} \mathbf{v}_{\mathcal{B}} \\ &= \begin{bmatrix} 2 & 0 \\ 0 & 4 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \end{bmatrix} \end{aligned}$$

$$\rightsquigarrow T(\mathbf{v}) = 2\mathbf{b}_1 + 4\mathbf{b}_2 = \begin{bmatrix} 6 \\ 2 \end{bmatrix}$$

Theorem 52. *Let $T : \mathbb{R}^m \to \mathbb{R}^n$ be a linear transformation and $\mathcal{A}$ and $\mathcal{B}$ be two bases of $\mathbb{R}^m$ and $\mathcal{C}, \mathcal{D}$ be two bases of $\mathbb{R}^n$. Then*

$$T_{\mathcal{C},\mathcal{A}} = I_{\mathcal{C},\mathcal{D}} T_{\mathcal{D},\mathcal{B}} I_{\mathcal{B},\mathcal{A}}.$$

Proof.

We compute

$$\begin{aligned} I_{\mathcal{C},\mathcal{D}} T_{\mathcal{D},\mathcal{B}} I_{\mathcal{B},\mathcal{A}} \mathbf{v}_{\mathcal{A}} &= I_{\mathcal{C},\mathcal{D}} T_{\mathcal{D},\mathcal{B}} \mathbf{v}_{\mathcal{B}} \\ &= I_{\mathcal{C},\mathcal{D}} T(\mathbf{v})_{\mathcal{D}} \\ &= T(\mathbf{v})_{\mathcal{C}} \end{aligned}$$

$$\begin{tikzcd} (\mathbb{R}^m, \mathcal{A}) \arrow[r, "\text{apply } T_{\mathcal{C},\mathcal{A}}"] \arrow[d, "I_{\mathcal{B},\mathcal{A}}"'] & (\mathbb{R}^n, \mathcal{C}) \\ (\mathbb{R}^m, \mathcal{B}) \arrow[r, "\text{apply } T_{\mathcal{D},\mathcal{B}}"] & (\mathbb{R}^n, \mathcal{D}) \arrow[u, "I_{\mathcal{C},\mathcal{D}}"'] \end{tikzcd}$$

Since $T_{\mathcal{C},\mathcal{A}}$ is the only matrix with $T_{\mathcal{C},\mathcal{A}} \mathbf{v}_{\mathcal{A}} = T(\mathbf{v})_{\mathcal{C}}$, we have

$$T_{\mathcal{C},\mathcal{A}} = I_{\mathcal{C},\mathcal{D}} T_{\mathcal{D},\mathcal{B}} I_{\mathcal{B},\mathcal{A}}.$$

Example. Consider $\mathcal{E} := \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}$ and $\mathcal{B} := \left\{ \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}$ as before. Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be again the linear transformation that $\mathbf{v} \mapsto \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix} \mathbf{v}$. Determine $T_{\mathcal{B},\mathcal{B}}$.

Solution.

$$T_{\mathcal{B},\mathcal{B}} = I_{\mathcal{B},\mathcal{E}} T_{\mathcal{E},\mathcal{E}} I_{\mathcal{E},\mathcal{B}}.$$

We already calculated that

$$I_{\mathcal{B},\mathcal{E}} = \frac{1}{2} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}, I_{\mathcal{E},\mathcal{B}} = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}.$$

Since $\mathcal{E}$ is the standard basis,

$$T_{\mathcal{E},\mathcal{E}} = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}.$$

Therefore

$$T_{\mathcal{B},\mathcal{B}} = \frac{1}{2} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 4 \end{bmatrix}$$

.

LINEAR ALGEBRA

Determinants

ILLINOIS

Department of Mathematics

Definition. The **determinant** of

- a 2×2 matrix is $\det\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = ad - bc$,
- a 1×1 matrix is $\det([a]) = a$.

Remark. Recall that

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

Goal. Define the determinant of an $n \times n$-matrix such that

$$A \text{ is invertible} \iff \det(A) \neq 0.$$

Notation. We will write both $\det\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right)$ and $\begin{vmatrix} a & b \\ c & d \end{vmatrix}$ for the determinant.

Definition. The **determinant** is the operation that assigns to each $n \times n$-matrix a number and satisfies the following conditions:

- **(Normalization)** $\det I_n = 1$,
- It is affected by elementary row operations as follows:
  - **(Replacement)** Adding a multiple of one row to another row *does not change* the determinant.
  - **(Interchange)** Interchanging two different rows *reverses the sign* of the determinant.
  - **(Scaling)** Multiplying all entries in a row by s, *multiplies* the determinant by s.

Example. Compute $\det\left(\begin{bmatrix} 2 & 3 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 6 \end{bmatrix}\right)$.

Solution.

$$\begin{vmatrix} 2 & 3 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 6 \end{vmatrix} \overset{\substack{R_2 \to R_2 - \frac{1}{3}R_3 \\ R_1 \to R_1 - \frac{1}{2}R_3}}{=} \begin{vmatrix} 2 & 3 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 6 \end{vmatrix} \overset{R_1 \to R_1 - 3R_2}{=} \begin{vmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 6 \end{vmatrix} \overset{\substack{R_1 \to \frac{1}{2}R_1 \\ R_3 \to \frac{1}{6}R_3}}{=} 2 \cdot (6) \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} = 12$$

Theorem 53. *The determinant of a triangular matrix is the product of the diagonal entries.*

Example. Compute $\begin{vmatrix} 1 & 2 & 0 \\ 3 & -1 & 2 \\ 2 & 0 & 1 \end{vmatrix}$.

Solution.

$$\begin{aligned}
\begin{vmatrix} 1 & 2 & 0 \\ 3 & -1 & 2 \\ 2 & 0 & 1 \end{vmatrix} &\overset{\substack{R2\to R2-3R1 \\ R3\to R3-2R1}}{=} \begin{vmatrix} 1 & 2 & 0 \\ 0 & -7 & 2 \\ 0 & -4 & 1 \end{vmatrix} \\
&\overset{R3\to R3-\frac{4}{7}R2}{=} \begin{vmatrix} 1 & 2 & 0 \\ 0 & -7 & 2 \\ 0 & 0 & -\frac{1}{7} \end{vmatrix} \\
&= 1 \cdot (-7) \cdot (-\frac{1}{7}) = 1
\end{aligned}$$

Example. (Re)Discover the formula for $\begin{vmatrix} a & b \\ c & d \end{vmatrix}$.

Solution.

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} \overset{R2 \to R2 - \frac{c}{a} R1}{=} \begin{vmatrix} a & b \\ 0 & d - \frac{c}{a} b \end{vmatrix} = a \left(d - \frac{c}{a} b \right) = ad - bc$$

NB: this only works if $a \neq 0$. What do you do if $a = 0$? Swap rows!

Example. Suppose A is a 3×3 matrix with $\det(A) = 5$. What is $\det(2A)$?

Solution.

A has three rows.

Multiplying each of them by 2 of them produces $2A$.

Hence, $\det(2A) = 2^3 \det(A) = 40$.

Theorem 54. *Let A be an $n \times n$-matrix. Then $\det(A) = 0$ if and only if A is not invertible.*

Proof.

Note that if A and B are row-equivalent,then

$$\det(A) = 0 \iff \det(B) = 0,$$

because elementary row operations do not change whether the determinant is 0 or not.

Thus A is invertible if and only if A is row-equivalent to I_n if and only if $\det(A) \neq 0$.

Theorem 55. *Let A, B be two $n \times n$-matrices. Then $\det(AB) = \det(A)\det(B)$.*

Example. If A is invertible, then $\det(A^{-1}) = \frac{1}{\det(A)}$. Why?

Solution.

$$1 = \det(I_n) = \det(AA^{-1}) = \det(A)\det(A^{-1}) \rightsquigarrow \det(A^{-1}) = \frac{1}{\det(A)}.$$

Theorem 56. *Let A be an $n \times n$-matrix. Then* $\det(A^T) = \det(A)$.

Proof.

Let P be a permutation matrix, L be a lower-triangular matrix and U be an upper triangular matrix such that $PA = LU$.

For simplicity, assume P is the permutation matrix of a single row swap. Then $P = P^T$.

$$\begin{aligned} \det(P)\det(A) &= \det(PA) = \det(LU) = \det(L)\det(U), \\ \det(A^T)\det(P^T) &= \det((PA)^T) = \det((LU)^T) = \det(U^T)\det(L^T). \end{aligned}$$

However, $\det(L^T) = \det(L)$ and $\det(U) = \det(U^T) \rightsquigarrow \det(A) = \det(A^T)$.

Remark. $\det(A^T) = \det(A)$ means that everything you know about determinants in terms of rows of A is also true for the *columns* of A. In particular:

- If you exchange two columns in a determinant, the determinant changes by a factor of -1.
- You can add a multiple of a column to another column without changing the determinant.
- Multipying each entry of a column by a scalar s, change the determinant by a factor of s.

LINEAR ALGEBRA

Cofactor expansion

ILLINOIS

Department of Mathematics

Notation. Let A be an $n \times n$-matrix. We denote by A_{ij} the matrix obtained from matrix A by deleting the i-th row and j-th column of A.

Example. Let $A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \\ 13 & 14 & 15 & 16 \end{bmatrix}$. Find A_{23} and A_{43}.

Solution.

$$A_{23} = \begin{bmatrix} 1 & 2 & 4 \\ 9 & 10 & 12 \\ 13 & 14 & 16 \end{bmatrix}, \qquad A_{43} = \begin{bmatrix} 1 & 2 & 4 \\ 5 & 6 & 8 \\ 9 & 10 & 12 \end{bmatrix}$$

Definition. Let A be an $n \times n$-matrix. The **$(\mathbf{i}, \mathbf{j})$-cofactor** of A is the scalar C_{ij} defined by

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

Theorem 57. *Let A be an $n \times n$-matrix. Then for every $i, j \in \{1, \ldots, n\}$*

$$\begin{aligned} \det A &= a_{i1} C_{i1} + a_{i2} C_{i2} + \cdots + a_{in} C_{in} \qquad (\textit{expansion across row } i) \\ &= a_{1j} C_{1j} + a_{2j} C_{2j} + \cdots + a_{nj} C_{nj} \qquad (\textit{expansion down column } j) \end{aligned}$$

Example. Compute $\begin{vmatrix} 1 & 2 & 0 \\ 3 & -1 & 2 \\ 2 & 0 & 1 \end{vmatrix}$ by cofactor expansion across row 1.

Solution.

$$\begin{vmatrix} 1 & 2 & 0 \\ 3 & -1 & 2 \\ 2 & 0 & 1 \end{vmatrix} = 1 \cdot (-1)^{1+1} \cdot \begin{vmatrix} + & & \\ & -1 & 2 \\ & 0 & 1 \end{vmatrix} + 2 \cdot (-1)^{1+2} \cdot \begin{vmatrix} & - & \\ 3 & & 2 \\ 2 & & 1 \end{vmatrix} + 0 \cdot (-1)^{1+3} \cdot \begin{vmatrix} & & + \\ 3 & -1 & \\ 2 & 0 & \end{vmatrix}$$

$$= 1 \cdot \begin{vmatrix} -1 & 2 \\ 0 & 1 \end{vmatrix} - 2 \cdot \begin{vmatrix} 3 & 2 \\ 2 & 1 \end{vmatrix} + 0 \cdot \begin{vmatrix} 3 & -1 \\ 2 & 0 \end{vmatrix} = 1 \cdot (-1) - 2 \cdot (-1) + 0 = 1$$

Example. Compute $\begin{vmatrix} 1 & 2 & 0 \\ 3 & -1 & 2 \\ 2 & 0 & 1 \end{vmatrix}$ by cofactor expansion down column 2 and by cofactor expansion down column 3.

Solution.

$$\begin{vmatrix} 1 & 2 & 0 \\ 3 & -1 & 2 \\ 2 & 0 & 1 \end{vmatrix} = 2\cdot(-1)^{1+2}\cdot\begin{vmatrix} 3 & 2 \\ 2 & 1 \end{vmatrix} + (-1)\cdot(-1)^{2+2}\cdot\begin{vmatrix} 1 & 0 \\ 2 & 1 \end{vmatrix} + 0\cdot(-1)^{3+2}\cdot\begin{vmatrix} 1 & 0 \\ 3 & 2 \end{vmatrix}$$

$$= -2\cdot(-1) + (-1)\cdot 1 - 0 = 1$$

$$\begin{vmatrix} 1 & 2 & 0 \\ 3 & -1 & 2 \\ 2 & 0 & 1 \end{vmatrix} = 0\cdot(-1)^{1+3}\cdot\begin{vmatrix} 3 & -1 \\ 2 & 0 \end{vmatrix} + 2\cdot(-1)^{2+3}\cdot\begin{vmatrix} 1 & 2 \\ 2 & 0 \end{vmatrix} + 1\cdot(-1)^{3+3}\cdot\begin{vmatrix} 1 & 2 \\ 3 & -1 \end{vmatrix}$$

$$= 0 - 2\cdot(-4) + 1\cdot(-7) = 1$$

Question. Why is the method of cofactor expansion not practical for large n?

Solution.

To compute the determinant of a large $n \times n$ matrix,

- one reduces to n determinants of size $(n-1) \times (n-1)$,
- then $n(n-1)$ determinants of size $(n-2) \times (n-2)$,
- and so on.

In the end, we have $n! = n(n-1)\cdots 3 \cdot 2 \cdot 1$ many numbers to add.
WAY TOO MUCH WORK! Already

$$25! = 15511210043330985984000000 \approx 1.55 \cdot 10^{25}.$$

P.-S. Laplace (1749–1827)

Remark. The cofactor expansion is also called **Laplace expansion**, because it is due Pierre-Simon Laplace.

LINEAR ALGEBRA

Eigenvectors and Eigenvalues

ILLINOIS

Department of Mathematics

Definition. Let A be an $n \times n$ matrix. An **eigenvector** of A is a nonzero $\mathbf{v} \in \mathbb{R}^n$ such that

$$A\mathbf{v} = \lambda \mathbf{v},$$

where λ is a scalar, known as the **eigenvalue** associated with $\mathbf{v}$.

Example. Verify that $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ is an eigenvector of $A = \begin{bmatrix} 0 & -2 \\ -4 & 2 \end{bmatrix}$. Is $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ an eigenvector?

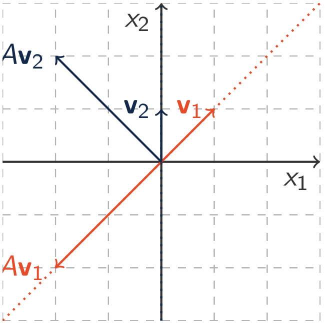
Solution.

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

Hence, $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ is an eigenvector of A with eigenvalue -2.

$$A \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 & -2 \\ -4 & 2 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -2 \\ 2 \end{bmatrix} \neq \lambda \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

for any scalar λ! Thus $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ is not an eigenvector of A.



Example. Find the eigenvectors and the eigenvalues of $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$.

Solution.

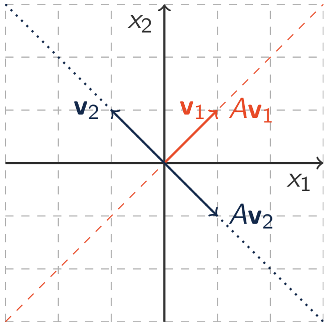
$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_2 \\ x_1 \end{bmatrix}$$

$$A \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} = 1 \cdot \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$\rightsquigarrow \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ is an eigenvector with eigenvalue 1.

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

$\rightsquigarrow \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ is an eigenvector with eigenvalue -1.



Example. Find the eigenvectors and the eigenvalues of $B = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$.

Solution.

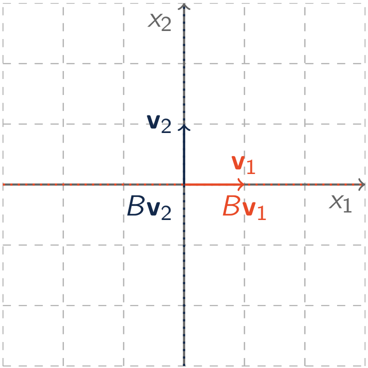
$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 \\ 0 \end{bmatrix}$$

$$B \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} = 1 \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

$\rightsquigarrow \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ is an eigenvector with eigenvalue 1.

$$B \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} = 0 \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$\rightsquigarrow \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ is an eigenvector with eigenvalue 0.



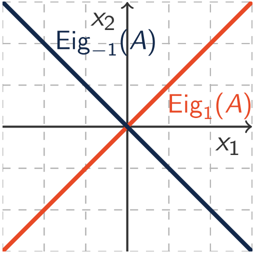
Definition. Let λ be an eigenvalue of A. The **eigenspace** of A associated with λ is the set

$$\mathrm{Eig}_\lambda(A) := \{\mathbf{v} \;:\; A\mathbf{v} = \lambda\mathbf{v}\}.$$

It consists of all the eigenvectors of A with eigenvalue λ and the zero vector.

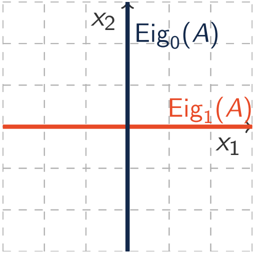
Example. Draw the eigenspaces of the two matrices $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$.

Solution.



$$A\begin{bmatrix} x \\ x \end{bmatrix} = 1 \cdot \begin{bmatrix} x \\ x \end{bmatrix}$$

$$A\begin{bmatrix} -x \\ x \end{bmatrix} = -1 \cdot \begin{bmatrix} -x \\ x \end{bmatrix}$$



$$B\begin{bmatrix} x \\ 0 \end{bmatrix} = 1 \cdot \begin{bmatrix} x \\ 0 \end{bmatrix}$$

$$B\begin{bmatrix} 0 \\ x \end{bmatrix} = 0 \cdot \begin{bmatrix} 0 \\ x \end{bmatrix}$$

LINEAR ALGEBRA

Computing Eigenvalues and Eigenvectors

ILLINOIS

Department of Mathematics

Theorem 58. *Let A be an $n \times n$ matrix and λ be a scalar. Then λ is an eigenvalue of A if and only if* $\det(A - \lambda I) = 0$.

Proof.

$$
\begin{aligned}
& \lambda \text{ is an eigenvalue of } A. \\
\Leftrightarrow\ & \text{there is a non-zero } \mathbf{v} \in \mathbb{R}^n \text{ such that } A\mathbf{v} = \lambda\mathbf{v}. \\
\Leftrightarrow\ & \text{there is a non-zero } \mathbf{v} \in \mathbb{R}^n \text{ such that } A\mathbf{v} - \lambda\mathbf{v} = \mathbf{0}. \\
\Leftrightarrow\ & \text{there is a non-zero } \mathbf{v} \in \mathbb{R}^n \text{ such that } (A - \lambda I)\mathbf{v} = \mathbf{0}. \\
\Leftrightarrow\ & \mathrm{Nul}(A - \lambda I) \neq \{\mathbf{0}\}. \\
\Leftrightarrow\ & A - \lambda I \text{ is not invertible.} \\
\Leftrightarrow\ & \det(A - \lambda I) = 0.
\end{aligned}
$$

Theorem 59. *Let A be an $n \times n$ matrix. Then $p_A(t) := \det(A - tI)$ is a polynomial of degree n. Thus A has at most n eigenvalues.*

Definition. We call $p_A(t)$ the **characteristic polynomial** of A.

Example. Find the eigenvalues of $A = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$.

Solution.

$$A - \lambda I = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 3-\lambda & 1 \\ 1 & 3-\lambda \end{bmatrix}$$

$$\det(A - \lambda I) = \begin{vmatrix} 3-\lambda & 1 \\ 1 & 3-\lambda \end{vmatrix} = (3-\lambda)^2 - 1$$

$$= \lambda^2 - 6\lambda + 8 = 0 \rightsquigarrow \lambda_1 = 2, \lambda_2 = 4$$

Theorem 60. *Let A be $n \times n$ matrix and let λ be eigenvalue of A. Then*

$$\mathrm{Eig}_\lambda(A) = \mathrm{Nul}(A - \lambda I).$$

Example. Determine the eigenspaces of $A = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$.

Solution.

Eigenvalue $\lambda_1 = 2$:

$$A - 2I = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \underset{RREF}{\rightsquigarrow} \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$$

$$\rightsquigarrow \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -x_2 \\ x_2 \end{bmatrix} = x_2 \begin{bmatrix} -1 \\ 1 \end{bmatrix} \rightsquigarrow \mathrm{Nul}(A - 2I) = \mathrm{span}\left(\begin{bmatrix} -1 \\ 1 \end{bmatrix} \right)$$

Eigenvalue $\lambda_2 = 4$:

$$A - 4I = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} \underset{RREF}{\rightsquigarrow} \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix}$$

$$\rightsquigarrow \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_2 \\ x_2 \end{bmatrix} = x_2 \begin{bmatrix} 1 \\ 1 \end{bmatrix} \rightsquigarrow \mathrm{Nul}(A - 4I) = \mathrm{span}\left(\begin{bmatrix} 1 \\ 1 \end{bmatrix} \right)$$

Example. Find the eigenvalues and eigenspaces of $A = \begin{bmatrix} 3 & 2 & 3 \\ 0 & 6 & 10 \\ 0 & 0 & 2 \end{bmatrix}$.

Solution.

$$\det(A - \lambda I) = \begin{vmatrix} 3-\lambda & 2 & 3 \\ 0 & 6-\lambda & 10 \\ 0 & 0 & 2-\lambda \end{vmatrix} = (3-\lambda)(6-\lambda)(2-\lambda)$$

$\rightsquigarrow$ A has eigenvalues $2, 3, 6$. The eigenvalues of a triangular matrix are its diagonal entries.

$$\lambda_1 = 2: \quad A - 2I = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 4 & 10 \\ 0 & 0 & 0 \end{bmatrix} \underset{RREF}{\rightsquigarrow} \begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & 2.5 \\ 0 & 0 & 0 \end{bmatrix} \rightsquigarrow \mathrm{Nul}(A - 2I) = \mathrm{span}\left(\begin{bmatrix} 2 \\ -5/2 \\ 1 \end{bmatrix} \right)$$

$$\lambda_2 = 3: \quad A - 3I = \begin{bmatrix} 0 & 2 & 3 \\ 0 & 3 & 10 \\ 0 & 0 & -1 \end{bmatrix} \underset{RREF}{\rightsquigarrow} \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \rightsquigarrow \mathrm{Nul}(A - 3I) = \mathrm{span}\left(\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \right)$$

$$\lambda_3 = 6: \quad A - 6I = \begin{bmatrix} -3 & 2 & 3 \\ 0 & 0 & 10 \\ 0 & 0 & -4 \end{bmatrix} \underset{RREF}{\rightsquigarrow} \begin{bmatrix} 1 & \frac{-2}{3} & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \rightsquigarrow \mathrm{Nul}(A - 6I) = \mathrm{span}\left(\begin{bmatrix} \frac{2}{3} \\ 1 \\ 0 \end{bmatrix} \right)$$

LINEAR ALGEBRA

Properties of Eigenvectors and Eigenvalues

ILLINOIS

Department of Mathematics

Definition. Let A be an $n \times n$ matrix and let λ be an eigenvalue of A.

- The **algebraic multiplicity** of λ is its multiplicity as a root of the characteristic polynomial, that is, the largest integer k such that $(t-\lambda)^k$ divides $p_A(t)$.
- The **geometric multiplicity** of λ is the dimension of the eigenspace $\mathrm{Eig}_\lambda(A)$ of λ.

Example. Find the eigenvalues of $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ and determine their algebraic and geometric multiplicities.

Solution.

- $\det(A - \lambda I) = \begin{vmatrix} 1-\lambda & 1 \\ 0 & 1-\lambda \end{vmatrix} = (1-\lambda)^2$

  So: $\lambda = 1$ is the only eigenvalue (it has algebraic multiplicity 2).
- $\lambda = 1: \;\; A - I = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \rightsquigarrow \mathrm{Nul}(A - I) = \mathrm{span}\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right)$

  Only dimension 1! $\rightsquigarrow$ Geometric multiplicity 1.

Theorem 61. *Let A be an $n \times n$ matrix and let $\mathbf{v}_1, \ldots, \mathbf{v}_m$ be eigenvectors of A corresponding to different eigenvalues. Then $\mathbf{v}_1, \ldots, \mathbf{v}_m$ are linearly independent.*

Proof.

For simplicity, assume $m = 2$. Let λ_1, λ_2 be such that $A\mathbf{v}_1 = \lambda_1 \mathbf{v}_1$ and $A\mathbf{v}_2 = \lambda_2 \mathbf{v}_2$. Suppose there are scalars c_1, c_2 such that

$$c_1 \mathbf{v}_1 + c_2 \mathbf{v}_2 = \mathbf{0}.$$

Now multiply this relation with A:

$$A(c_1 \mathbf{v}_1 + c_2 \mathbf{v_2}) = c_1 A\mathbf{v}_1 + c_2 A\mathbf{v_2} = c_1 \lambda_1 \mathbf{v}_1 + c_2 \lambda_2 \mathbf{v}_2 = \mathbf{0}$$

Subtracting λ_2 times the first equation from the second, we obtain

$$c_1(\lambda_1 - \lambda_2)\mathbf{v}_1 = \mathbf{0}.$$

Since $\lambda_1 \neq \lambda_2$, we get that $c_1 = 0$. Also $c_2 = 0$, since $\mathbf{v}_2 \neq \mathbf{0}$.

Definition. Let $A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}$. The **trace** of A is the sum of the diagonal entries of A; that is

$$\mathrm{Tr}(A) = a_{11} + a_{22} + \cdots + a_{nn}.$$

Theorem 62. *Let A be $n \times n$ matrix with eigenvalues $\lambda_1, \dots, \lambda_n$. Then*

- $\mathrm{Tr}(A) = \lambda_1 + \lambda_2 + \cdots + \lambda_n$.
- $\det(A) = \lambda_1 \cdot \lambda_2 \cdot \cdots \cdot \lambda_n$.

Theorem 63. *Let A be a 2×2-matrix. Then the characteristic polynomial of A is*

$$p(\lambda) = \lambda^2 - \mathrm{Tr}(A)\lambda + \det(A).$$

Solution.

$$p(\lambda) = \begin{vmatrix} a_{11} - \lambda & a_{12} \\ a_{21} & a_{22} - \lambda \end{vmatrix} = (a_{11} - \lambda)(a_{22} - \lambda) - a_{12}a_{21} = \lambda^2 - \underbrace{(a_{11} + a_{22})}_{\mathrm{Tr}(A)}\lambda + \underbrace{(a_{11}a_{22} - a_{12}a_{21})}_{\det(A)}$$

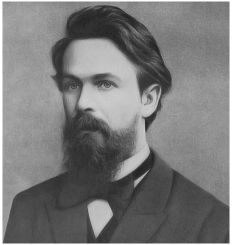
LINEAR ALGEBRA

Markov matrices

ILLINOIS

Department of Mathematics

Definition. An $n \times n$ matrix A is a **Markov matrix** (or: stochastic matrix) if it has only non-negative entries, and the entries in each column add up to 1.
A vector in $\mathbb{R}^n$ is a **probability vector** (or: stochastic vector) if has only non-negative entries, and the entries add up to 1.



A. Markov (1856–1922)

Example. Give examples of Markov matrices and probability vectors.

Solution.

$$\begin{bmatrix} .1 & .5 \\ .9 & .5 \end{bmatrix}, \begin{bmatrix} 0 & .25 & .4 \\ 1 & .25 & .2 \\ 0 & .5 & .4 \end{bmatrix}, \begin{bmatrix} .1 \\ .25 \\ .05 \\ .5 \\ .1 \end{bmatrix}.$$ Note that $\begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}$ is not a probability vector, but $\frac{1}{10}\begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}$ is.

Question. Let A be $n \times n$ Markov matrix and $\mathbf{v} \in \mathbb{R}^n$ be a probability vector. Is $A\mathbf{v}$ also a probability vector?

Solution.

$$A\mathbf{v} = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{bmatrix} \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} v_1 a_{11} + \cdots + v_n a_{1n} \\ \vdots \\ v_1 a_{n1} + \cdots + v_n a_{nn} \end{bmatrix}$$

$$\rightsquigarrow (v_1 a_{11} + \cdots + v_n a_{1n}) + \cdots + (v_1 a_{n1} + \cdots + v_n a_{nn})$$

$$= v_1 \underbrace{(a_{11} + \cdots + a_{n1})}_{=1} + \cdots + v_n \underbrace{(a_{1n} + \cdots + a_{nn})}_{=1} = v_1 + \cdots + v_n = 1.$$

Theorem 64. *Let A be a Markov matrix. Then*

- 1 *is an eigenvalue of A and every other eigenvalue λ of A satisfies $|\lambda| \leq 1$.*
- *If A has only positive entries, then any other eigenvalue satisfies $|\lambda| < 1$.*

Definition. A **stationary** probability vector of a Markov matrix is probability vector $\mathbf{v}$ that is an eigenvector of A corresponding to the eigenvaule 1.

Theorem 65. *Let A be an $n \times n$-Markov matrix with only positive entries and let $\mathbf{z} \in \mathbb{R}^n$ be a probability vector. Then*

$$\mathbf{z}_\infty := \lim_{k\to\infty} A^k \mathbf{z} \text{ exists,}$$

and $\mathbf{z}_\infty$ is a stationary probability vector of A (ie. $A\mathbf{z}_\infty = \mathbf{z}_\infty$).

Proof.

For simplicity, assume that A has n different eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$.

Since A is a Markov matrix and has only positive entries, we can assume that $\lambda_1 = 1$ and $|\lambda_i| < 1$ for all $i = 2, \dots, n$.

Let $\mathbf{v_1}, \mathbf{v_2}, \dots, \mathbf{v_n} \in \mathbb{R}^n$ such that $A\mathbf{v}_i = \lambda_i \mathbf{v_i}$.

Since eigenvectors to different eigenvalues are linear independent, the above eigenvectors form a basis of $\mathbb{R}^n$.

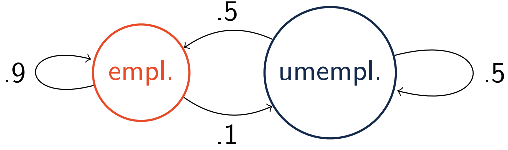
Thus there are scalar $c_1, \dots, c_n$ such that $\mathbf{z} = c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n$.

$$\rightsquigarrow A^k\mathbf{z} = c_1\lambda_1^k\mathbf{v_1} + \cdots + c_n\lambda_n^k\mathbf{v_n} \to c_1\mathbf{v_1},$$

because $|\lambda_i| < 1$ for each $i = 2, \dots, n$. Why is $c_1 \neq 0$? Because $A^k\mathbf{z}$ is a probability vector.

Example. Consider a fixed population of people with or without job. Suppose that each year, 50% of those unemployed find a job while 10% of those employed lose their job.
What is the unemployment rate in the long term equilibrium?

Solution.



x_t: % of population employed at time t
y_t: % of population unemployed at time t

$$\begin{bmatrix} x_{t+1} \\ y_{t+1} \end{bmatrix} = \begin{bmatrix} .9x_t + .5y_t \\ .1x_t + .5y_t \end{bmatrix} = \begin{bmatrix} .9 & .5 \\ .1 & .5 \end{bmatrix} \begin{bmatrix} x_t \\ y_t \end{bmatrix}$$

$$\begin{bmatrix} x_\infty \\ y_\infty \end{bmatrix} = \begin{bmatrix} 0.9 & 0.5 \\ 0.1 & 0.5 \end{bmatrix} \begin{bmatrix} x_\infty \\ y_\infty \end{bmatrix}$$

$\rightsquigarrow \begin{bmatrix} x_\infty \\ y_\infty \end{bmatrix}$ stationary probability vector

$$A - 1I = \begin{bmatrix} -0.1 & 0.5 \\ 0.1 & -0.5 \end{bmatrix} \underset{RREF}{\rightsquigarrow} \begin{bmatrix} 1 & -5 \\ 0 & 0 \end{bmatrix}$$

$$\rightsquigarrow \mathrm{Nul}\left(\begin{bmatrix} -0.1 & 0.5 \\ 0.1 & -0.5 \end{bmatrix} \right) = \mathrm{span}\left(\begin{bmatrix} 5 \\ 1 \end{bmatrix} \right)$$

Since $x_\infty + y_\infty = 1 \rightsquigarrow \begin{bmatrix} x_\infty \\ y_\infty \end{bmatrix} = \begin{bmatrix} 5/6 \\ 1/6 \end{bmatrix}$.

Unemployment rate in the long term equilibrium is $1/6$

LINEAR ALGEBRA

Diagonalization

ILLINOIS

Department of Mathematics

Definition. A square matrix A is said to be **diagonalizable** if there is a invertible matrix P and a diagonal matrix D such that $A = PDP^{-1}$.

Theorem 66. *Let A be an $n \times n$ matrix that has n linearly independent eigenvectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ with associated eigenvalues $\lambda_1, \ldots, \lambda_n$. Then A is diagonalizable as PDP^{-1}, where*

$$P = \begin{bmatrix} \mathbf{v_1} & \ldots & \mathbf{v_n} \end{bmatrix} \text{ and } D = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$$

Proof.

$$\begin{aligned} AP &= A \begin{bmatrix} \mathbf{v_1} & \ldots & \mathbf{v_n} \end{bmatrix} = \begin{bmatrix} A\mathbf{v_1} & \ldots & A\mathbf{v_n} \end{bmatrix} \\ &= \begin{bmatrix} \lambda_1 \mathbf{v_1} & \ldots & \lambda_n \mathbf{v_n} \end{bmatrix} = \begin{bmatrix} \mathbf{v_1} & \ldots & \mathbf{v_n} \end{bmatrix} \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} = PD \\ &\rightsquigarrow A = PDP^{-1} \end{aligned}$$

Example. Diagonalize $A = \begin{bmatrix} 6 & -1 \\ 2 & 3 \end{bmatrix}$.

Solution.

$$\det(A - \lambda I) = \begin{vmatrix} 6-\lambda & -1 \\ 2 & 3-\lambda \end{vmatrix} = (6-\lambda)(3-\lambda) + 2 = \lambda^2 - 9\lambda + 20 = (\lambda - 4)(\lambda - 5)$$

$$\lambda_1 = 4: \qquad A - 4I = \begin{bmatrix} 2 & -1 \\ 2 & -1 \end{bmatrix} \underset{RREF}{\rightsquigarrow} \begin{bmatrix} 1 & -\frac{1}{2} \\ 0 & 0 \end{bmatrix} \rightsquigarrow \mathrm{Nul}(A - 4I) = \mathrm{span}\left(\begin{bmatrix} \frac{1}{2} \\ 1 \end{bmatrix}\right)$$

$$\lambda_2 = 5: \qquad A - 5I = \begin{bmatrix} 1 & -1 \\ 2 & -2 \end{bmatrix} \underset{RREF}{\rightsquigarrow} \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix} \rightsquigarrow \mathrm{Nul}(A - 5I) = \mathrm{span}\left(\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right)$$

$$D = \begin{bmatrix} 4 & 0 \\ 0 & 5 \end{bmatrix}, P = \begin{bmatrix} \frac{1}{2} & 1 \\ 1 & 1 \end{bmatrix} \quad \rightsquigarrow \quad A = PDP^{-1}.$$

Definition. Vectors $\mathbf{v}_1, \dots, \mathbf{v}_n$ form an **eigenbasis** of $n \times n$ matrix A if $\mathbf{v}_1, \dots, \mathbf{v}_n$ form a basis of $\mathbb{R}^n$ and $\mathbf{v}_1, \dots, \mathbf{v}_n$ are all eigenvectors of A.

Theorem 67. *The following are equivalent for $n \times n$ matrix A:*

- *A has an eigenbasis.*
- *A is diagonalizable.*
- *The geometric multiplicities of all eigenvalues of A sums up to n.*

Theorem 68. *Let A be $n \times n$ matrix and let $\mathcal{B} = (\mathbf{v}_1, \dots, \mathbf{v}_n)$ be an eigenbasis of A. Then there is a diagonal matrix D such that*

$$A = I_{\mathcal{E}_n,\mathcal{B}} D I_{\mathcal{B},\mathcal{E}_n}.$$

Proof.

Set $P = \begin{bmatrix} \mathbf{v}_1 & \dots & \mathbf{v}_n \end{bmatrix}$. Recall that $I_{\mathcal{E}_n,\mathcal{B}} = P$ and $I_{\mathcal{B},\mathcal{E}_n} = I_{\mathcal{E}_n,\mathcal{B}}^{-1}$.

$$A = PDP^{-1} = I_{\mathcal{E}_n,\mathcal{B}} D I_{\mathcal{B},\mathcal{E}_n}.$$

$\rightsquigarrow$ *Diagonalization is base change to eigenbasis!*

Question. Let A be an $n \times n$ with n distinct eigenvalues. Is A diagonalizable?

Solution.

Yes. Eigenvectors to different eigenvalue are linearly independent. Thus A has n linearly independent eigenvectors.

Question. Is $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ diagonalizable?

Solution.

$$\det(A - \lambda I) = \begin{vmatrix} -\lambda & 1 \\ 0 & -\lambda \end{vmatrix} = \lambda^2$$

$$\lambda_1, \lambda_2 = 0 : \qquad A - 0I = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \rightsquigarrow \mathrm{Nul}(A) = \mathrm{span}\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right)$$

$\rightsquigarrow$ Only one linearly independent eigenvector. No eigenbasis

LINEAR ALGEBRA

Powers of Matrices

ILLINOIS

Department of Mathematics

Idea. If A has an eigenbasis, then we can raise A to large powers easily!

Theorem 69. *If $A = PDP^{-1}$, where D is a diagonal matrix, then for any m,*

$$A^m = PD^mP^{-1}$$

Proof.

$$\begin{aligned}
A^m &= (PDP^{-1})^m \\
&= \underbrace{(PDP^{-1}) \cdot (PDP^{-1}) \cdots (PDP^{-1})}_{m \text{ times}} \\
&= (PD)(P^{-1} \cdot P)(DP^{-1}) \cdots (PDP^{-1}) \\
&= PD \cdot DP^{-1} \cdots PDP^{-1} \\
&= PD \cdot D \cdots D \cdot P^{-1} \\
&= PD^mP^{-1}
\end{aligned}$$

Remark. Finding D^m is easy!

$$D^m = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}^m = \begin{bmatrix} (\lambda_1)^m & & \\ & \ddots & \\ & & (\lambda_n)^m \end{bmatrix}$$

Example. Let $A = \begin{bmatrix} \frac{1}{2} & 0 & 0 \\ -\frac{1}{2} & 1 & 6 \\ 0 & 0 & 2 \end{bmatrix}$. The matrix has the following eigenvectors:

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \text{ with } \lambda_1 = \frac{1}{2}, \mathbf{v}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \text{ with } \lambda_2 = 1, \mathbf{v}_3 = \begin{bmatrix} 0 \\ 6 \\ 1 \end{bmatrix} \text{ with } \lambda_3 = 2.$$

Find A^{100}.

Solution.

$$\overbrace{\begin{bmatrix} \frac{1}{2} & 0 & 0 \\ -\frac{1}{2} & 1 & 6 \\ 0 & 0 & 2 \end{bmatrix}}^{A} = \overbrace{\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 6 \\ 0 & 0 & 1 \end{bmatrix}}^{P} \overbrace{\begin{bmatrix} \frac{1}{2} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}}^{D} \overbrace{\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 6 \\ 0 & 0 & 1 \end{bmatrix}^{-1}}^{P^{-1}} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 6 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{2} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & -6 \\ 0 & 0 & 1 \end{bmatrix}$$

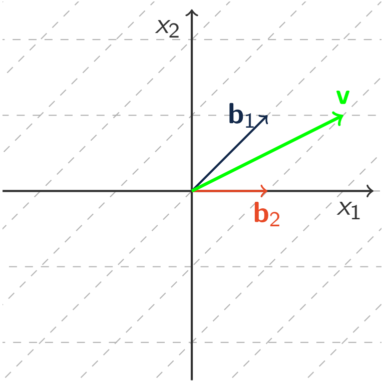
$$\mathrm{A}^{100} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 6 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{2}^{100} & 0 & 0 \\ 0 & 1^{100} & 0 \\ 0 & 0 & 2^{100} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & -6 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{2^{100}} & 0 & 0 \\ (\frac{1}{2^{100}} - 1) & 1 & (6 \cdot 2^{100} - 6) \\ 0 & 0 & 2^{100} \end{bmatrix}$$

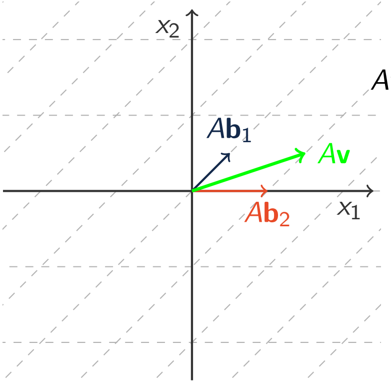
Example. Let A be a 2×2 matrix such that

$\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ is an eigenvector with eigenvalue $\frac{1}{2}$, $\quad \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ is an eigenvector with eigenvalue 1.

Let $\mathbf{v} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$. Graphically determine how $A^n\mathbf{v}$ behaves as $n \to \infty$.

Solution.





$$\begin{aligned} A^n\mathbf{v} &= A^n \left(\begin{bmatrix} 1 \\ 1 \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right) \\ &= A^n \begin{bmatrix} 1 \\ 1 \end{bmatrix} + A^n \begin{bmatrix} 1 \\ 0 \end{bmatrix} \\ &= (\frac{1}{2})^n \begin{bmatrix} 1 \\ 1 \end{bmatrix} + 1^n \begin{bmatrix} 1 \\ 0 \end{bmatrix} \\ &\to \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \text{as } n \to \infty \end{aligned}$$

LINEAR ALGEBRA

Matrix exponential

ILLINOIS

Department of Mathematics

Definition. Let A be an $n \times n$-matrix. We define the **matrix exponential** e^{At} as

$$e^{At} = I + At + \frac{(At)^2}{2!} + \frac{(At)^3}{3!} + \dots$$

Example. Compute e^{At} for $A = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$.

Solution.

$$At = \begin{bmatrix} 2t & 0 \\ 0 & t \end{bmatrix} \rightsquigarrow (At)^k = \begin{bmatrix} (2t)^k & 0 \\ 0 & t^k \end{bmatrix}$$

$$e^{At} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 2t & 0 \\ 0 & 1t \end{bmatrix} + \frac{1}{2!} \begin{bmatrix} (2t)^2 & 0 \\ 0 & t^2 \end{bmatrix} + \dots$$

$$= \begin{bmatrix} 1 + 2t + \frac{(2t)^2}{2!} + \dots & 0 \\ 0 & 1 + t + \frac{t^2}{2!} + \dots \end{bmatrix} = \begin{bmatrix} e^{2t} & 0 \\ 0 & e^t \end{bmatrix}$$

Theorem 70. *Let A be an $n \times n$ matrix. Then*

- *The series in the definition of e^{At} always converges,*
- $e^{At}e^{As} = e^{A(t+s)}$,
- $e^{At}e^{-At} = I_n$,
- $\frac{d}{dt}(e^{At}) = Ae^{At}$.

Theorem 71. *Let A be a $n \times n$ matrix such that $A = PDP^{-1}$ for some invertible matrix P and some diagonal matrix D. Then*

$$e^{At} = Pe^{Dt}P^{-1}.$$

Proof.

Recall that

$$A^k = PD^kP^{-1}.$$

$$\begin{aligned} e^{At} &= I + At + \frac{(At)^2}{2!} + \frac{(At)^3}{3!} + \dots \\ &= I + PDP^{-1}t + \frac{PD^2P^{-1}t^2}{2!} + \frac{PD^3P^{-1}t^3}{3!} + \dots \\ &= P\left(I + Dt + \frac{D^2t^2}{2!} + \frac{D^3t^3}{3!} + \dots\right)P^{-1} \\ &= Pe^{Dt}P^{-1} \end{aligned}$$

Example. Let $A = \begin{bmatrix} -2 & 1 \\ 1 & -2 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} -1 & 0 \\ 0 & -3 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}^{-1}$. Compute e^{At}.

Solution.

$$
\begin{aligned}
e^{At} &= Pe^{Dt}P^{-1} \\
&= \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} e^{-t} & 0 \\ 0 & e^{-3t} \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}^{-1} \\
&= \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} e^{-t} & 0 \\ 0 & e^{-3t} \end{bmatrix} \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{bmatrix} \\
&= \begin{bmatrix} e^{-t} & e^{-3t} \\ e^{-t} & -e^{-3t} \end{bmatrix} \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{bmatrix} \\
&= \begin{bmatrix} \frac{1}{2}e^{-t} + \frac{1}{2}e^{-3t} & \frac{1}{2}e^{-t} - \frac{1}{2}e^{-3t} \\ \frac{1}{2}e^{-t} - \frac{1}{2}e^{-3t} & \frac{1}{2}e^{-t} + \frac{1}{2}e^{-3t} \end{bmatrix}
\end{aligned}
$$

LINEAR ALGEBRA

Linear Differential Equations

ILLINOIS

Department of Mathematics

Definition. A **linear (first order) differential equation** is an equation of the form

$$\frac{d\mathbf{u}}{dt} = A\mathbf{u}$$

where $\mathbf{u}$ is function from $\mathbb{R}$ to $\mathbb{R}^n$. A further condition $\mathbf{u}(0) = \mathbf{v}$, for some $\mathbf{v}$ in $\mathbb{R}^n$ is called a **initial condition**.

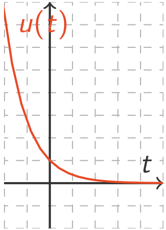
Theorem 72. *Let A be an $n \times n$ matrix and $\mathbf{v} \in \mathbb{R}^n$. The solution of the differential equation $\frac{d\mathbf{u}}{dt} = A\mathbf{u}$ with initial condition $\mathbf{u}(0) = \mathbf{v}$ is $\mathbf{u}(t) = e^{At}\mathbf{v}$.*

Example. Let $n = 1$. Find the solution of the differential equation $\frac{du}{dt} = -u$ and $u(0) = 1$.

Solution.

In this case, $A = \begin{bmatrix} -1 \end{bmatrix}$. Thus

$$u(t) = e^{-t} \cdot 1 = e^{-t}$$



Theorem 73. *Let A be an $n \times n$-matrix, and let $\mathbf{v} \in \mathbb{R}^n$ be an eigenvector of A with eigenvalue λ. Then $e^{At}\mathbf{v} = e^{\lambda t}\mathbf{v}$.*

Proof.

$$
\begin{aligned}
e^{At}\mathbf{v} &= \left(I + At + \frac{(At)^2}{2!} + \frac{(At)^3}{3!} + \dots\right)\mathbf{v} = \left(\mathbf{v} + tA\mathbf{v} + \frac{t^2A^2\mathbf{v}}{2!} + \frac{t^3A^3\mathbf{v}}{3!} + \dots\right) \\
&= \left(\mathbf{v} + \lambda t\mathbf{v} + \frac{(\lambda t)^2\mathbf{v}}{2!} + \frac{(\lambda t)^3\mathbf{v}}{3!} + \dots\right) = \left(1 + \lambda t + \frac{(\lambda t)^2}{2!} + \frac{(\lambda t)^3}{3!} + \dots\right)\mathbf{v} = e^{\lambda t}\mathbf{v}
\end{aligned}
$$

Theorem 74. *Let A be an $n \times n$-matrix and $(\mathbf{v}_1, \dots, \mathbf{v}_n)$ be an eigenbasis of A with eigenvalues $\lambda_1, \dots, \lambda_n$. If $\mathbf{v} = c_1\mathbf{v}_1 + \dots + c_n\mathbf{v}_n$,then the unique solution to the differential equation $\frac{d\mathbf{u}}{dt} = A\mathbf{u}$ with initial condition $\mathbf{u}(0) = \mathbf{v}$ is*

$$
e^{At}\mathbf{v} = c_1e^{\lambda_1 t}\mathbf{v}_1 + \dots + c_ne^{\lambda_n t}\mathbf{v}_n.
$$

Proof.

$$
\begin{aligned}
e^{At}\mathbf{v} &= e^{At}(c_1\mathbf{v}_1 + \dots + c_n\mathbf{v}_n) = c_1e^{At}\mathbf{v}_1 + \dots + c_ne^{At}\mathbf{v}_n \\
&= c_1e^{\lambda_1 t}\mathbf{v}_1 + \dots + c_ne^{\lambda_n t}\mathbf{v}_n.
\end{aligned}
$$

Example. Let $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. Solve the differential equation $\frac{d\mathbf{u}}{dt} = A\mathbf{u}$ with initial condition $\mathbf{u}(0) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$.

Solution.

- Eigenvalues: $\det(A - \lambda I) = \lambda^2 - 1 \rightsquigarrow \lambda_1 = 1, \lambda_2 = -1$
- Eigenvectors: $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ eigenvector with eigenvalue 1, $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$ eigenvector with eigenvalue -1
- Express $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ in terms of the eigenbasis: $\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 1 \\ 1 \end{bmatrix} + \frac{1}{2}\begin{bmatrix} 1 \\ -1 \end{bmatrix}$.
-

$$e^{At}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = e^{At}\left(\frac{1}{2}\begin{bmatrix} 1 \\ 1 \end{bmatrix} + \frac{1}{2}\begin{bmatrix} 1 \\ -1 \end{bmatrix}\right) = \frac{1}{2}e^{At}\begin{bmatrix} 1 \\ 1 \end{bmatrix} + \frac{1}{2}e^{At}\begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

$$= \frac{1}{2}e^{t}\begin{bmatrix} 1 \\ 1 \end{bmatrix} + \frac{1}{2}e^{-t}\begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} \frac{1}{2}e^{t} + \frac{1}{2}e^{-t} \\ \frac{1}{2}e^{t} - \frac{1}{2}e^{-t} \end{bmatrix} \to \infty \text{ (as } t \to \infty)$$

Example. Let $A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$. Solve the differential equation $\frac{d\mathbf{u}}{dt} = A\mathbf{u}$ with initial condition $\mathbf{u}(0) = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$.

Solution.

- Eigenvalues: $\det(A - \lambda I) = \lambda^2 - 2\lambda - 3 = (\lambda - 3)(\lambda + 1) \rightsquigarrow \lambda_1 = 3, \lambda_2 = -1$
- Eigenvectors: $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ eigenvector with eigenvalue 3, $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$ eigenvector with eigenvalue -1
- Express $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ in terms of the eigenbasis: $\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 \\ 1 \end{bmatrix} + -\frac{1}{2} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$.
-

$$e^{At} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = e^{At}(\frac{1}{2} \begin{bmatrix} 1 \\ 1 \end{bmatrix} + -\frac{1}{2} \begin{bmatrix} 1 \\ -1 \end{bmatrix}) = \frac{1}{2} e^{At} \begin{bmatrix} 1 \\ 1 \end{bmatrix} + -\frac{1}{2} e^{At} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

$$= \frac{1}{2} e^{3t} \begin{bmatrix} 1 \\ 1 \end{bmatrix} - \frac{1}{2} e^{-t} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} e^{3t} - \frac{1}{2} e^{-t} \\ \frac{1}{2} e^{3t} + \frac{1}{2} e^{-t} \end{bmatrix} \to \infty \text{ (as } t \to \infty)$$

LINEAR ALGEBRA

Orthogonal projection onto a line

ILLINOIS

Department of Mathematics

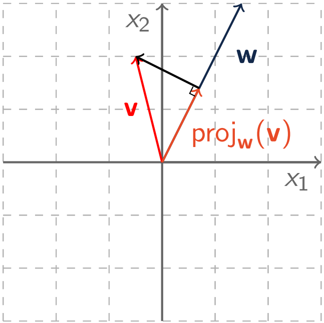
Definition. Let $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$. The **orthogonal projection** of $\mathbf{v}$ onto the line spanned by $\mathbf{w}$ is

$$\text{proj}_{\mathbf{w}}(\mathbf{v}) := \frac{\mathbf{w} \cdot \mathbf{v}}{\mathbf{w} \cdot \mathbf{w}} \mathbf{w}.$$

Theorem 75. *Let $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$. Then $\text{proj}_{\mathbf{w}}(\mathbf{v})$ is the point in $\text{span}(\mathbf{w})$ closest to $\mathbf{v}$; that is*

$$\text{dist}(\mathbf{v}, \text{proj}_{\mathbf{w}}(\mathbf{v})) = \min_{\mathbf{u} \in \text{span}(\mathbf{w})} \text{dist}(\mathbf{v}, \mathbf{u}).$$

Proof.



- We know $\text{proj}_{\mathbf{w}}(\mathbf{v}) \in \text{span}(\mathbf{w})$. So $\text{proj}_{\mathbf{w}}(\mathbf{v}) = c\mathbf{w}$ for some scalar c.
- Observe that $\mathbf{v} - \text{proj}_{\mathbf{w}}(\mathbf{v})$ orthogonal to $\mathbf{w}$.
- So

$$\begin{aligned} 0 &= \mathbf{w} \cdot (\mathbf{v} - \text{proj}_{\mathbf{w}}(\mathbf{v})) \\ &= \mathbf{w} \cdot (\mathbf{v} - c\mathbf{w}) = \mathbf{w} \cdot \mathbf{v} - c\mathbf{w} \cdot \mathbf{w} \\ \rightsquigarrow c &= \frac{\mathbf{w} \cdot \mathbf{v}}{\mathbf{w} \cdot \mathbf{w}} \end{aligned}$$

Remark. Note that $\mathbf{v} - \text{proj}_{\mathbf{w}}(\mathbf{v})$ (called the **error term**) is in $\text{span}(\mathbf{w})^{\perp}$.

$$\mathbf{v} = \underbrace{\text{proj}_{\mathbf{w}}(\mathbf{v})}_{\in \text{span}(\mathbf{w})} + \underbrace{\mathbf{v} - \text{proj}_{\mathbf{w}}(\mathbf{v})}_{\in \text{span}(\mathbf{w})^{\perp}}$$

Example. Find the orthogonal projection of $\mathbf{v} = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$ onto the line spanned by $\mathbf{w} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$.

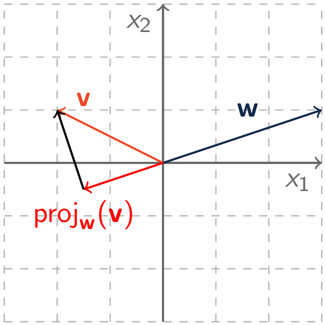
Solution.

$$\text{proj}_{\mathbf{w}}(\mathbf{v}) = \frac{\mathbf{w} \cdot \mathbf{v}}{\mathbf{w} \cdot \mathbf{w}} \mathbf{w} = \frac{-6 + 1}{3^2 + 1^2} \begin{bmatrix} 3 \\ 1 \end{bmatrix}$$

$$= -\frac{5}{10} \begin{bmatrix} 3 \\ 1 \end{bmatrix} = \begin{bmatrix} -1.5 \\ -.5 \end{bmatrix}$$

The error term is

$$\mathbf{v} - \text{proj}_{\mathbf{w}}(\mathbf{v}) = \begin{bmatrix} -2 \\ 1 \end{bmatrix} - \begin{bmatrix} -1.5 \\ -.5 \end{bmatrix} = \begin{bmatrix} -0.5 \\ 1.5 \end{bmatrix}.$$



Theorem 76. *Let* $\mathbf{w} \in \mathbb{R}^n$*. Then for all* $\mathbf{v} \in \mathbb{R}^n$

$$\text{proj}_{\mathbf{w}}(\mathbf{v}) = \left(\frac{1}{\mathbf{w} \cdot \mathbf{w}} \mathbf{w}\mathbf{w}^T\right) \mathbf{v}.$$

Proof.

Trick: $c\mathbf{u} = \mathbf{u}c$ for every scalar c and $\mathbf{u} \in \mathbb{R}^n$ (Think of c as 1×1 matrix).

Rewrite the formula for $\text{proj}_{\mathbf{w}}(\mathbf{v})$.

$$\begin{aligned} \text{proj}_{\mathbf{w}}(\mathbf{v}) &= \frac{\mathbf{w} \cdot \mathbf{v}}{\mathbf{w} \cdot \mathbf{w}} \mathbf{w} = \frac{1}{\mathbf{w} \cdot \mathbf{w}} (\mathbf{w}^T \mathbf{v}) \mathbf{w} = \frac{1}{\mathbf{w} \cdot \mathbf{w}} \mathbf{w} (\mathbf{w}^T \mathbf{v}) \\ &= \frac{1}{\mathbf{w} \cdot \mathbf{w}} (\mathbf{w}\mathbf{w}^T) \mathbf{v} = \left(\frac{1}{\mathbf{w} \cdot \mathbf{w}} \mathbf{w}\mathbf{w}^T\right) \mathbf{v}. \end{aligned}$$

Remark. Note that $\left(\frac{1}{\mathbf{w} \cdot \mathbf{w}} \mathbf{w}\mathbf{w}^T\right)$ is an $n \times n$ matrix, we call the **orthogonal projection matrix** onto $\text{span}(\mathbf{w})$.

Example. Let $\mathbf{w} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. Find the orthogonal projection matrix P onto span($\mathbf{w}$). Use it to calculate the projections of $\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ onto span($\mathbf{w}$).

Solution.

$$P = \frac{1}{\mathbf{w} \cdot \mathbf{w}} \mathbf{w}\mathbf{w}^T = \frac{1}{2} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}.$$

- If $\mathbf{v} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, then $\text{proj}_{\mathbf{w}}(\mathbf{v}) = P\mathbf{v} = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$
- If $\mathbf{v} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, then $\text{proj}_{\mathbf{w}}(\mathbf{v}) = P\mathbf{v} = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \mathbf{v}$
- If $\mathbf{v} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$, then $\text{proj}_{\mathbf{w}}(\mathbf{v}) = P\mathbf{v} = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

LINEAR ALGEBRA

Orthogonal projection onto a subspace

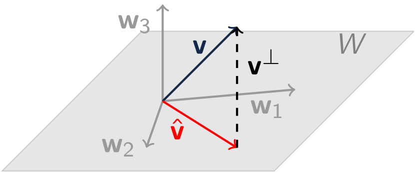
ILLINOIS

Department of Mathematics

Theorem 77. *Let W be a subspace of $\mathbb{R}^n$ and $\mathbf{v} \in \mathbb{R}^n$. Then $\mathbf{v}$ can be written uniquely as*

$$\mathbf{v} = \underbrace{\hat{\mathbf{v}}}_{\text{in } W} + \underbrace{\mathbf{v}^{\perp}}_{\text{in } W^{\perp}}$$

Proof.



Note that $\dim W + \dim W^{\perp} = n$. Let $(\mathbf{w}_1, \dots, \mathbf{w}_\ell)$ be a basis of W and let $(\mathbf{u}_1, \dots, \mathbf{u}_m)$ be a basis of $W^{\perp}$. Since $\ell + m = n$, $(\mathbf{w}_1, \dots, \mathbf{w}_\ell, \mathbf{u}_1, \dots, \mathbf{u}_m)$ is a basis of $\mathbb{R}^n$. Thus

$$\mathbf{v} = \underbrace{c_1\mathbf{w}_1 + \dots + c_\ell\mathbf{w}_\ell}_{=:\hat{\mathbf{v}}} + \underbrace{d_1\mathbf{u}_1 + \dots d_m\mathbf{u}_m}_{=:\mathbf{v}^{\perp}}$$

Definition. We say $\hat{\mathbf{v}}$ is the **orthogonal projection** of $\mathbf{v}$ onto W - written $\mathrm{proj}_W(\mathbf{v})$.

Remark. If $(\mathbf{w}_1, \dots, \mathbf{w}_m)$ is an orthogonal basis of W, then

$$\mathrm{proj}_W(\mathbf{v}) = \left(\frac{\mathbf{v} \cdot \mathbf{w}_1}{\mathbf{w}_1 \cdot \mathbf{w}_1}\right)\mathbf{w}_1 + \dots + \left(\frac{\mathbf{v} \cdot \mathbf{w}_m}{\mathbf{w}_m \cdot \mathbf{w}_m}\right)\mathbf{w}_m.$$

Example. Let $W = \operatorname{span}\left(\underbrace{\begin{bmatrix} 3 \\ 0 \\ 1 \end{bmatrix}}_{=:\mathbf{w}_1}, \underbrace{\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}}_{=:\mathbf{w}_2} \right)$ and $\mathbf{v} = \begin{bmatrix} 0 \\ 3 \\ 10 \end{bmatrix}$. Compute $\operatorname{proj}_W(\mathbf{v})$.

Solution.

$$\begin{aligned}
\operatorname{proj}_W(\mathbf{v}) &= \frac{\mathbf{v} \cdot \mathbf{w}_1}{\mathbf{w}_1 \cdot \mathbf{w}_1}\mathbf{w}_1 + \frac{\mathbf{v} \cdot \mathbf{w}_2}{\mathbf{w}_2 \cdot \mathbf{w}_2} w_2 \\
&= \frac{\begin{bmatrix} 0 \\ 3 \\ 10 \end{bmatrix} \cdot \begin{bmatrix} 3 \\ 0 \\ 1 \end{bmatrix}}{\begin{bmatrix} 3 \\ 0 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 3 \\ 0 \\ 1 \end{bmatrix}} \begin{bmatrix} 3 \\ 0 \\ 1 \end{bmatrix} + \frac{\begin{bmatrix} 0 \\ 3 \\ 10 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}}{\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \\
&= \frac{10}{10} \begin{bmatrix} 3 \\ 0 \\ 1 \end{bmatrix} + 3 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \\ 1 \end{bmatrix}
\end{aligned}$$

This calculation only works for orthogonal $\mathbf{w_1}, \mathbf{w_2}$!

$$\begin{aligned}
\mathbf{v}^{\perp} &= \begin{bmatrix} 0 \\ 3 \\ 10 \end{bmatrix} - \begin{bmatrix} 3 \\ 3 \\ 1 \end{bmatrix} \\
&= \begin{bmatrix} -3 \\ 0 \\ 9 \end{bmatrix}
\end{aligned}$$

Theorem 78. *The projection map* $\operatorname{proj}_W : \mathbb{R}^n \to \mathbb{R}^n$ *that sends* $\mathbf{v}$ *to* $\operatorname{proj}_W(\mathbf{v})$ *is linear.*

Definition. The matrix P_W is the matrix $(\operatorname{proj}_W)_{\mathcal{E}_n,\mathcal{E}_n}$ that represents proj_W with respect to the standard basis. We call P_W the **orthogonal projection matrix** onto W.

Example. Compute P_W for $W = \operatorname{span}\left(\begin{bmatrix}3\\0\\1\end{bmatrix}, \begin{bmatrix}0\\1\\0\end{bmatrix}\right)$.

Solution.

$$\operatorname{proj}_W\left(\begin{bmatrix}1\\0\\0\end{bmatrix}\right) = \frac{\begin{bmatrix}1\\0\\0\end{bmatrix}\cdot\begin{bmatrix}3\\0\\1\end{bmatrix}}{\begin{bmatrix}3\\0\\1\end{bmatrix}\cdot\begin{bmatrix}3\\0\\1\end{bmatrix}}\begin{bmatrix}3\\0\\1\end{bmatrix} + \frac{\begin{bmatrix}1\\0\\0\end{bmatrix}\cdot\begin{bmatrix}0\\1\\0\end{bmatrix}}{\begin{bmatrix}0\\1\\0\end{bmatrix}\cdot\begin{bmatrix}0\\1\\0\end{bmatrix}}\begin{bmatrix}0\\1\\0\end{bmatrix} = \frac{3}{10}\begin{bmatrix}3\\0\\1\end{bmatrix}$$

$$\operatorname{proj}_W\left(\begin{bmatrix}0\\1\\0\end{bmatrix}\right) = \begin{bmatrix}0\\1\\0\end{bmatrix}$$

$$\operatorname{proj}_W\left(\begin{bmatrix}0\\0\\1\end{bmatrix}\right) = \frac{\begin{bmatrix}0\\0\\1\end{bmatrix}\cdot\begin{bmatrix}3\\0\\1\end{bmatrix}}{10}\begin{bmatrix}3\\0\\1\end{bmatrix} + \frac{\begin{bmatrix}0\\0\\1\end{bmatrix}\cdot\begin{bmatrix}0\\1\\0\end{bmatrix}}{1}\begin{bmatrix}0\\1\\0\end{bmatrix} = \frac{1}{10}\begin{bmatrix}3\\0\\1\end{bmatrix}$$

$$\rightsquigarrow P_W = \begin{bmatrix}\frac{9}{10} & 0 & \frac{3}{10}\\ 0 & 1 & 0\\ \frac{3}{10} & 0 & \frac{1}{10}\end{bmatrix}$$

Question. Let P_W be the orthogonal projection matrix onto W in $\mathbb{R}^n$, and let $\mathbf{v} \in \mathbb{R}^n$.

- If $P_W\mathbf{v} = \mathbf{v}$, what can you say about $\mathbf{v}$?
- If $P_W\mathbf{v} = \mathbf{0}$, what can you say about $\mathbf{v}$?

Solution.

Let $\mathbf{v} = \mathbf{w} + \mathbf{u}$, where $\mathbf{w} \in W$ and $\mathbf{u} \in W^\perp$. Then $\mathbf{w} = P_W\mathbf{v}$.

If $P_W\mathbf{v} = \mathbf{v}$, we get that $\mathbf{w} = \mathbf{v}$ and thus $\mathbf{v}$ is in W.

If $P_W\mathbf{v} = \mathbf{0}$, then $\mathbf{u} = \mathbf{v}$ and thus $\mathbf{v}$ in $W^\perp$.

Question. What is the orthogonal projection matrix $P_{W^\perp}$ for projecting onto $W^\perp$?

Solution.

Set $Q = I - P_W$, where I is the identity. Then $P_{W^\perp} = Q$. Why?

Let $\mathbf{v}$ be in $\mathbb{R}^n$. In order to show that $Q\mathbf{v}$ is the projection of $\mathbf{v}$ onto $W^\perp$, we need to check that $Q\mathbf{v} \in W^\perp$ and $\mathbf{v} - Q\mathbf{v} \in (W^\perp)^\perp$.

Since $P_W\mathbf{v}$ is the projection of $\mathbf{v}$ onto W, $Q\mathbf{v} = \mathbf{v} - P_W\mathbf{v}$ is in $W^\perp$.

Now observe that $\mathbf{v} - Q\mathbf{v} = P_W\mathbf{v}$. Thus $\mathbf{v} - Q\mathbf{v}$ is in $W = (W^\perp)^\perp$.

LINEAR ALGEBRA

Least squares solutions

ILLINOIS

Department of Mathematics

Goal. Suppose $A\mathbf{x} = \mathbf{b}$ is inconsistent. Can we still find something like a best solution?

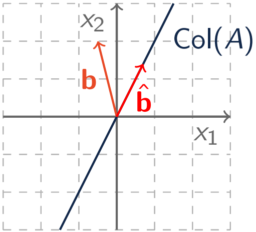
Definition. Let A be an $m \times n$ matrix and $\mathbf{b} \in \mathbb{R}^m$. A **least squares solution** (short: LSQ solution) of the system $A\mathbf{x} = \mathbf{b}$ is a vector $\hat{\mathbf{x}} \in \mathbb{R}^n$ such that

$$\operatorname{dist}(A\hat{\mathbf{x}}, \mathbf{b}) = \min_{\mathbf{x} \in \mathbb{R}^n} \operatorname{dist}(A\mathbf{x}, \mathbf{b}).$$

Theorem 79. *Let A be an $m \times n$ matrix and $\mathbf{b} \in \mathbb{R}^m$. Then $\hat{\mathbf{x}}$ is an LSQ solution to $A\mathbf{x} = \mathbf{b}$ if and only if $A\hat{\mathbf{x}} = \operatorname{proj}_{\operatorname{Col}(A)}(\mathbf{b})$.*

Proof.

Recall that $\operatorname{Col}(A) = \{A\mathbf{x} \ : \ \mathbf{x} \in \mathbb{R}^n\}$.



Let $\hat{\mathbf{x}} \in \mathbb{R}^n$. So $A\hat{\mathbf{x}} \in \operatorname{Col}(A)$.

$$\operatorname{dist}(A\hat{\mathbf{x}}, \mathbf{b}) = \min_{\mathbf{x} \in \mathbb{R}^n} \operatorname{dist}(A\mathbf{x}, \mathbf{b}) \iff \operatorname{dist}(A\hat{\mathbf{x}}, \mathbf{b}) = \min_{\mathbf{w} \in \operatorname{Col}(A)} \operatorname{dist}(\mathbf{w}, \mathbf{b})$$

$$\iff A\hat{\mathbf{x}} = \operatorname{proj}_{\operatorname{Col}(A)}(\mathbf{b})$$

Example. Let $A = \begin{bmatrix} 1 & 1 \\ -1 & 1 \\ 0 & 0 \end{bmatrix}$ and $\mathbf{b} = \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}$. Find an LSQ solution of $A\mathbf{x} = \mathbf{b}$.

Solution.

$$\hat{\mathbf{b}} = \mathrm{proj}_{\mathrm{Col}(A)}(\mathbf{b}) = \frac{\begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}}{\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}} \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} + \frac{\begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}}{\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} + \frac{3}{2} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}.$$

Solve $A\hat{\mathbf{x}} = \hat{\mathbf{b}}$:

$$\begin{bmatrix} 1 & 1 \\ -1 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}$$

We have already solved equation in the process when computing $\hat{\mathbf{b}}$: $\hat{\mathbf{x}} = \begin{bmatrix} 1/2 \\ 3/2 \end{bmatrix}$.

Theorem 80. *Let A be an $m \times n$ matrix and $\mathbf{b} \in \mathbb{R}^m$. Then $\hat{\mathbf{x}}$ is an LSQ solution to $A\mathbf{x} = \mathbf{b}$ if and only if $A^T A\hat{\mathbf{x}} = A^T\mathbf{b}$.*

Proof.

$$\begin{aligned}
&\hat{\mathbf{x}} \text{ is a least squares solution of } A\mathbf{x} = \mathbf{b} \\
\iff\ & A\hat{\mathbf{x}} - \mathbf{b} \text{ is as small as possible} \\
\iff\ & A\hat{\mathbf{x}} - \mathbf{b} \text{ is orthogonal to } \operatorname{Col}(A) \\
\overset{FTLA}{\iff}\ & A\hat{\mathbf{x}} - \mathbf{b} \text{ is in } \operatorname{Nul}(A^T) \\
\iff\ & A^T(A\hat{\mathbf{x}} - \mathbf{b}) = \mathbf{0} \\
\iff\ & A^T A\hat{\mathbf{x}} = A^T\mathbf{b}
\end{aligned}$$

Theorem 81. *Let A be an $m \times n$ matrix with linearly independent columns and $\mathbf{b} \in \mathbb{R}^m$. Then*

$$\operatorname{proj}_{\operatorname{Col}(A)}(\mathbf{b}) = A(A^T A)^{-1}A^T\mathbf{b}.$$

Proof.

Let $\hat{\mathbf{x}}$ be an LSQ solution of $A\mathbf{x} = \mathbf{b}$. The hypotheses implies that $(A^T A)$ is invertible. Then

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

$$\rightsquigarrow \operatorname{proj}_{\operatorname{Col}(A)}(\mathbf{b}) = A\hat{\mathbf{x}} = A(A^T A)^{-1}A^T\mathbf{b}$$

Example. Let $A = \begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{bmatrix}$ and $\mathbf{b} = \begin{bmatrix} 2 \\ 0 \\ 11 \end{bmatrix}$. Find a LSQ solution of $A\mathbf{x} = \mathbf{b}$.

Solution.

$$A^T A = \begin{bmatrix} 4 & 0 & 1 \\ 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 17 & 1 \\ 1 & 5 \end{bmatrix}$$

$$A^T \mathbf{b} = \begin{bmatrix} 4 & 0 & 1 \\ 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 0 \\ 11 \end{bmatrix} = \begin{bmatrix} 19 \\ 11 \end{bmatrix}$$

The normal equations $A^T A\hat{\mathbf{x}} = A^T\mathbf{b}$ are $\begin{bmatrix} 17 & 1 \\ 1 & 5 \end{bmatrix} \hat{\mathbf{x}} = \begin{bmatrix} 19 \\ 11 \end{bmatrix}$.

Solving, we find $\hat{\mathbf{x}} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$.

The projection of $\mathbf{b}$ onto $\mathrm{Col}(A)$ is $A\hat{\mathbf{x}} = \begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 4 \\ 4 \\ 3 \end{bmatrix}$.

LINEAR ALGEBRA

Linear Regression

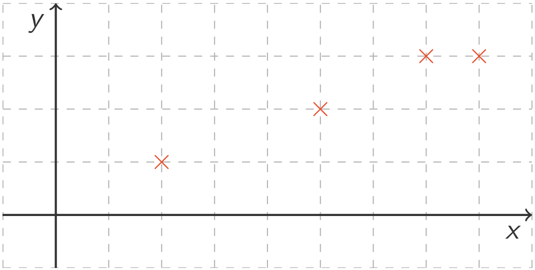
ILLINOIS

Department of Mathematics

Example. Are there $\beta_1, \beta_2 \in \mathbb{R}$ such that the data points

$$(x_1, y_1) = (2, 1), (x_2, y_2) = (5, 2),$$
$$(x_3, y_3) = (7, 3), (x_4, y_4) = (8, 3)$$

all lie on the line $y = \beta_1 + \beta_2 x$.



Solution.

Need that for $i = 1, 2, 3, 4 \quad y_i = \beta_1 + \beta_2 x_i$.

In matrix form:

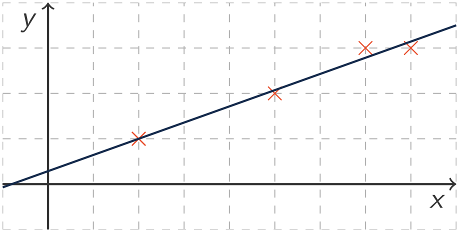
$$\underbrace{\begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ 1 & x_3 \\ 1 & x_4 \end{bmatrix}}_{\text{design matrix } X} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} = \underbrace{\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix}}_{\text{observation vector } \mathbf{y}}$$

$$\left[\begin{array}{cc|c} 1 & 2 & 1 \\ 1 & 5 & 2 \\ 1 & 7 & 3 \\ 1 & 8 & 3 \end{array}\right] \underset{\textit{reduce}}{\rightsquigarrow} \left[\begin{array}{cc|c} 1 & 2 & 1 \\ 0 & 3 & 1 \\ 0 & 0 & \frac{1}{3} \\ 0 & 0 & 0 \end{array}\right]$$

The system is inconsistent.

Example (ctd). Find β_1, β_2 such that the line $y = \beta_1 + \beta_2 x$ best fits the data points:

$$(x_1, y_1) = (2,1), (x_2, y_2) = (5,2),$$
$$(x_3, y_3) = (7,3), (x_4, y_4) = (8,3)$$



Solution.

Find a LSQ to

$$\underbrace{\begin{bmatrix} 1 & 2 \\ 1 & 5 \\ 1 & 7 \\ 1 & 8 \end{bmatrix}}_{\text{design matrix } X} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} = \underbrace{\begin{bmatrix} 1 \\ 2 \\ 3 \\ 3 \end{bmatrix}}_{\text{observation vector } \mathbf{y}}$$

Normal equations:

$$\overbrace{\begin{bmatrix} 4 & 22 \\ 22 & 142 \end{bmatrix}}^{X^T X} \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix} = \overbrace{\begin{bmatrix} 9 \\ 57 \end{bmatrix}}^{X^T \mathbf{y}}$$

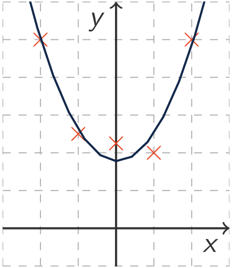
$$\rightsquigarrow \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix} = \begin{bmatrix} 2/7 \\ 5/14 \end{bmatrix}$$

$\rightsquigarrow$ LSQ-line: $y = \frac{2}{7} + \frac{5}{14} x$

Example. A scientist tries to find the relation between the mysterious quantities x and y. She measures the following values:

x	-2	-1	0	1	2
y	5	2.5	2.25	2	5

Find $\beta_1, \beta_2, \beta_3$ such that $y = \beta_1 + \beta_2 x + \beta_3 x^2$ best fits the data.



Solution.

For $i = 1, \ldots, 5:\ \ \beta_1 + x_i\beta_2 + x_i^2\beta_3 = y_i$.

$$\rightsquigarrow \underbrace{\begin{bmatrix} 1 & -2 & 4 \\ 1 & -1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \end{bmatrix}}_{=:X} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} = \underbrace{\begin{bmatrix} 5 \\ 2.5 \\ 2.25 \\ 2 \\ 5 \end{bmatrix}}_{=:\mathbf{y}}$$

$$\overbrace{\begin{bmatrix} 5 & 0 & 10 \\ 0 & 10 & 0 \\ 10 & 0 & 34 \end{bmatrix}}^{X^T X} \begin{bmatrix} \hat\beta_1 \\ \hat\beta_2 \\ \hat\beta_3 \end{bmatrix} = \overbrace{\begin{bmatrix} 16.75 \\ -0.5 \\ 44.5 \end{bmatrix}}^{X^T\mathbf{y}}$$

$$\rightsquigarrow \begin{bmatrix} \hat\beta_1 \\ \hat\beta_2 \\ \hat\beta_3 \end{bmatrix} = \begin{bmatrix} 1.78 \\ -0.05 \\ 0.79 \end{bmatrix}$$

$\rightsquigarrow$ LSQ-quadratic: $y = 1.78 - 0.05x + 0.79x^2$

Question. Suppose you are given a data set (u_i, v_i, y_i), $i = 1, \ldots, n$. You expect that y depends linearly on u and v. Which linear system do you have to solve to determine this dependency?

Solution.

Need: $y_i = \beta_1 + \beta_2 u_i + \beta_3 v_i$ for each $i = 1, \ldots, n$.

$$\rightsquigarrow \underbrace{\begin{bmatrix} 1 & u_1 & v_1 \\ 1 & u_2 & v_2 \\ \vdots & \vdots & \vdots \\ 1 & u_n & v_n \end{bmatrix}}_{\text{design matrix}} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} = \underbrace{\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}}_{\text{observation vector}}$$

Question. Suppose you are given a data set (x_i, y_i), $i = 1, \ldots, n$. You expect that y depends linearly on x and $\sin(x)$. Which linear system to you have to solve to determine this dependency?

Solution.

Need: $y_i = \beta_1 + \beta_2 x_i + \beta_3 \sin(x_i)$ for each $i = 1, \ldots, n$.

$$\rightsquigarrow \underbrace{\begin{bmatrix} 1 & x_1 & \sin(x_1) \\ 1 & x_2 & \sin(x_2) \\ \vdots & \vdots & \vdots \\ 1 & x_n & \sin(x_n) \end{bmatrix}}_{\text{design matrix}} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} = \underbrace{\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}}_{\text{observation vector}}$$

LINEAR ALGEBRA

Gram–Schmidt Method

ILLINOIS

Department of Mathematics

Theorem 82. *Every subspace of $\mathbb{R}^n$ has an orthonormal basis.*

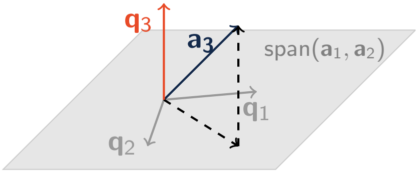
Algorithm. (Gram-Schmidt orthonormalization) Given a basis $\mathbf{a}_1, \ldots, \mathbf{a}_m$, produce an orthogonal basis $\mathbf{b}_1, \ldots, \mathbf{b}_m$ and an orthonormal basis $\mathbf{q}_1, \ldots, \mathbf{q}_m$.

$$\mathbf{b}_1 = \mathbf{a}_1, \qquad \mathbf{q}_1 = \frac{\mathbf{b}_1}{\|\mathbf{b}_1\|}$$

$$\mathbf{b}_2 = \mathbf{a}_2 - \underbrace{\operatorname{proj}_{\operatorname{span}(\mathbf{q}_1)}(\mathbf{a}_2)}_{=(\mathbf{a}_2 \cdot \mathbf{q}_1)\mathbf{q}_1}, \qquad \mathbf{q}_2 = \frac{\mathbf{b}_2}{\|\mathbf{b}_2\|}$$

$$\mathbf{b}_3 = \mathbf{a}_3 - \underbrace{\operatorname{proj}_{\operatorname{span}(\mathbf{q}_1, \mathbf{q}_2)}(\mathbf{a}_3)}_{(\mathbf{a}_3 \cdot \mathbf{q}_1)\mathbf{q}_1 + (\mathbf{a}_3 \cdot \mathbf{q}_2)\mathbf{q}_2} \qquad \mathbf{q}_3 = \frac{\mathbf{b}_3}{\|\mathbf{b}_3\|}$$

$$\cdots \qquad \cdots$$



Remark.

- $\operatorname{span}(\mathbf{q}_1, \ldots, \mathbf{q}_i) = \operatorname{span}(\mathbf{a}_1, \ldots, \mathbf{a}_i)$ for $i = 1, \ldots, m$.
- $\mathbf{q}_j \notin \operatorname{span}(\mathbf{a}_1, \ldots, \mathbf{a}_i)$ for all $j > i$.

Example. Let $V = \text{span}\left(\begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 3 \end{bmatrix} \right)$. Use the Gram-Schmidt method to find an orthonormal basis of V.

Solution.

Set $\mathbf{b}_1 = \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix}$ and $\mathbf{q}_1 = \frac{\mathbf{b}_1}{\|\mathbf{b}_1\|} = \frac{1}{3} \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix}$.

$$\mathbf{b}_2 = \begin{bmatrix} 0 \\ 0 \\ 3 \end{bmatrix} - \left(\begin{bmatrix} 0 \\ 0 \\ 3 \end{bmatrix} \cdot \frac{1}{3} \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix} \right) \frac{1}{3} \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 3 \end{bmatrix} - \frac{2}{3} \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} -\frac{4}{3} \\ -\frac{2}{3} \\ \frac{5}{3} \end{bmatrix}$$

$$\rightsquigarrow \mathbf{q}_2 = \frac{1}{\sqrt{45}} \begin{bmatrix} -4 \\ -2 \\ 5 \end{bmatrix}$$

Theorem 83. *(QR decomposition) Let A be an $m \times n$ matrix of rank n. There is is an $m \times n$-matrix Q with orthonormal columns and an upper triangular $n \times n$ invertible matrix R such that* $A = QR$.

Proof.

Let's do the 3×3 case: $A = \begin{bmatrix} \mathbf{a_1} & \mathbf{a_2} & \mathbf{a_3} \end{bmatrix}$. Using Gram-Schmidt we obtain orthonormal vectors $\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3$ such that

- $(\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3)$ is a basis of $\operatorname{Col}(A)$
- $\mathbf{a_1} \in \operatorname{span}(\mathbf{q}_1)$, $\mathbf{a_2} \in \operatorname{span}(\mathbf{q}_1, \mathbf{q}_2)$ and $\mathbf{a_3} \in \operatorname{span}(\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3)$

$$\rightsquigarrow \mathbf{a}_1 = (\mathbf{a}_1 \cdot \mathbf{q}_1)\,\mathbf{q_1}, \qquad \mathbf{a}_2 = (\mathbf{a}_2 \cdot \mathbf{q}_1)\,\mathbf{q_1} + (\mathbf{a}_2 \cdot \mathbf{q}_2)\,\mathbf{q_2},$$
$$\mathbf{a}_3 = (\mathbf{a}_3 \cdot \mathbf{q}_1)\,\mathbf{q_1} + (\mathbf{a}_3 \cdot \mathbf{q}_2)\,\mathbf{q_2} + (\mathbf{a}_3 \cdot \mathbf{q}_3)\,\mathbf{q_3}$$

$$\rightsquigarrow A = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3 \end{bmatrix} = \begin{bmatrix} \mathbf{q}_1 & \mathbf{q}_2 & \mathbf{q}_3 \end{bmatrix} \begin{bmatrix} \mathbf{a}_1 \cdot \mathbf{q}_1 & \mathbf{a}_2 \cdot \mathbf{q}_1 & \mathbf{a}_3 \cdot \mathbf{q}_1 \\ 0 & \mathbf{a}_2 \cdot \mathbf{q}_2 & \mathbf{a}_3 \cdot \mathbf{q}_2 \\ 0 & 0 & \mathbf{a}_3 \cdot \mathbf{q}_3 \end{bmatrix}$$

Example. Find the QR decomposition of $A = \begin{bmatrix} 1 & 2 & 4 \\ 0 & 0 & 5 \\ 0 & 3 & 6 \end{bmatrix}$.

Solution.

Apply Gram-Schmidt on columns of A:

$$\mathbf{q}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \quad \mathbf{b}_2 = \begin{bmatrix} 2 \\ 0 \\ 3 \end{bmatrix} - \left(\begin{bmatrix} 2 \\ 0 \\ 3 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \right) \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 3 \end{bmatrix} \rightsquigarrow \mathbf{q}_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

$$\mathbf{b}_3 = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} - \left(\begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \right) \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} - \left(\begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right) \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 5 \\ 0 \end{bmatrix} \rightsquigarrow \mathbf{q}_3 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

$$\rightsquigarrow Q = [\mathbf{q}_1 \ \mathbf{q}_2 \ \mathbf{q}_3] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \rightsquigarrow R = Q^T A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 & 4 \\ 0 & 0 & 5 \\ 0 & 3 & 6 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 4 \\ 0 & 3 & 6 \\ 0 & 0 & 5 \end{bmatrix}$$

LINEAR ALGEBRA

Spectral Theorem

ILLINOIS

Department of Mathematics

Theorem 84. *Let A be a symmetric $n \times n$ matrix. Then A has an orthonormal basis of eigenvectors.*

Proof.

We will just show that eigenvectors to different eigenvalues are orthogonal. Let $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$ be such that

$$A\mathbf{v} = \lambda_1 \mathbf{v} \text{ and } A\mathbf{w} = \lambda_2 \mathbf{w}, \qquad \text{for } \lambda_1 \neq \lambda_2.$$

$$\begin{aligned}
\lambda_1(\mathbf{v} \cdot \mathbf{w}) \quad &= (\lambda_1 \mathbf{v}) \cdot \mathbf{w} \\
&= (A\mathbf{v}) \cdot \mathbf{w} \\
&= (A\mathbf{v})^T \mathbf{w} \\
&= \mathbf{v}^T A^T \mathbf{w} \\
&= \mathbf{v}^T A \mathbf{w} \qquad \longleftarrow \text{ because } A^T = A \\
&= \mathbf{v} \cdot (A\mathbf{w}) \\
&= \lambda_2(\mathbf{v} \cdot \mathbf{w}) \\
&\rightsquigarrow \mathbf{v} \cdot \mathbf{w} = 0, \; \text{since } \lambda_1 \neq \lambda_2.
\end{aligned}$$

Theorem 85. *Let A be a symmetric $n \times n$ matrix. Then there is a diagonal matrix D and a matrix Q with orthonormal columns such that $A = QDQ^T$.*

Proof.

Let Q be the matrix whose columns are orthonormal eigenvectors.
Then Q is orthogonal, that is $Q^{-1} = Q^T$. Thus,

$$A = QDQ^{-1} = QDQ^T.$$

Question. If A is an $n \times n$ matrix with an orthonormal eigenbasis, is it symmetric?

Solution.

Since A has an orthonormal eigenbasis, there is a diagonal matrix D and a matrix Q with orthonormal columns such that $A = QDQ^T$.

$$A^T = (QDQ^T)^T = (Q^T)^T D^T Q^T = QDQ^T = A$$

Example. Let $A = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$. Write A as QDQ^T, where D is diagonal and Q has orthonormal columns.

Solution.

A has eigenvalues 2 and 4.

$\lambda_1 = 2$: Eigevector $\mathbf{v}_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$. Normalized, we get $\mathbf{q}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$.

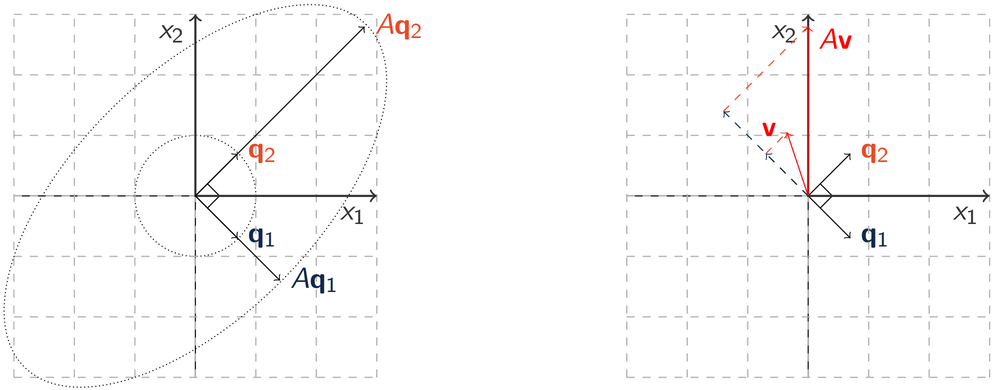
$\lambda_2 = 4$: Eigenvector $\mathbf{v}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. Normalized, we get $\mathbf{q}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

$D = \begin{bmatrix} 2 & 0 \\ 0 & 4 \end{bmatrix}$ and $Q = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$

$$A = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 4 \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$$

Example. Recall that $A = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 4 \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$, and $\mathbf{q}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ and $\mathbf{q}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. Suppose $\mathbf{v} = -\mathbf{q}_1 + \frac{1}{2}\mathbf{q}_2$. Graphically, determine $A\mathbf{v}$.

Solution.



LINEAR ALGEBRA

Singular Value Decomposition

ILLINOIS

Department of Mathematics

Definition. Let A be an $m \times n$ matrix. A **singular value decomposition** of A is a decomposition $A = U\Sigma V^T$ where

- U is an $m \times m$ matrix with orthonormal columns,
- Σ is an $m \times n$ rectangular diagonal matrix with non-negative numbers on the diagonal,
- V is an $n \times n$ matrix with orthonormal columns.

Remark. The diagonal entries $\sigma_i = \Sigma_{ii}$ which are positive are called the singular values of A. We usually arrange them in decreasing order, that is $\sigma_1 \geq \sigma_2 \geq \dots$

Question. Let A be an $m \times n$ matrix with rank r. Recall why

- $\mathrm{Nul}(A^T A) = \mathrm{Nul}(A)$ and $\mathrm{Nul}(AA^T) = \mathrm{Nul}(A^T)$.
- $A^T A$ is symmetric and has rank r.

Solution.

$$\mathbf{v} \in \mathrm{Nul}(A) \rightsquigarrow A\mathbf{v} = \mathbf{0} \rightsquigarrow A^T A\mathbf{v} = A^T\mathbf{0} = \mathbf{0} \rightsquigarrow \mathbf{v} \in \mathrm{Nul}(A^T A)$$

$$\mathbf{v} \in \mathrm{Nul}(A^T A) \rightsquigarrow A^T A\mathbf{v} = \mathbf{0} \rightsquigarrow (A\mathbf{v}) \cdot (A\mathbf{v}) = (A\mathbf{v})^T A\mathbf{v} = \mathbf{v}^T A^T A\mathbf{v} = \mathbf{v}^T\mathbf{0} = 0 \rightsquigarrow \mathbf{v} \in \mathrm{Nul}(A).$$

$$\begin{aligned} n &= \dim \mathrm{Col}(A^T A) + \dim \mathrm{Nul}(A^T A) \\ &= \dim \mathrm{Col}(A^T A) + \dim \mathrm{Nul}(A) = \dim \mathrm{Col}(A^T A) + n - r \rightsquigarrow \dim \mathrm{Col}(A^T A) = r. \end{aligned}$$

Algorithm. Let A be an $m \times n$ matrix with rank r.

- Find orthonormal eigenbasis $(\mathbf{v}_1, \dots, \mathbf{v}_n)$ of $A^T A$ with eigenvalues $\lambda_1 \geq \dots \geq \lambda_r > \lambda_{r+1} = 0 = \dots = \lambda_n$.
- Set $\sigma_i = \sqrt{\lambda_i}$ for $i = 1, \dots, n$.
- Set $\mathbf{u}_1 = \frac{1}{\sigma_1} A\mathbf{v}_1$, ..., $\mathbf{u}_r = \frac{1}{\sigma_r} A\mathbf{v}_r$. (Magic: orthonormal!)
- Find $\mathbf{u}_{r+1}, \dots, \mathbf{u}_m \in \mathbb{R}^m$ such that $(\mathbf{u}_1, \dots, \mathbf{u}_m)$ is an orthonormal basis of $\mathbb{R}^m$.
- Set

$$U = \begin{bmatrix} \mathbf{u}_1 & \dots & \mathbf{u}_m \end{bmatrix}, \qquad \Sigma = \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_{\min\{m,n\}} \end{bmatrix}, V = \begin{bmatrix} \mathbf{v}_1 & \dots & \mathbf{v}_n \end{bmatrix}$$

Proof that $A = U\Sigma V^T$.

Note that $\mathbf{v}_{r+1}, \dots, \mathbf{v}_n \in \mathrm{Nul}(A^T A)$. Thus $\mathbf{v}_{r+1}, \dots, \mathbf{v}_n \in \mathrm{Nul}(A)$.

$$\begin{aligned} \rightsquigarrow AV &= A \begin{bmatrix} \mathbf{v}_1 & \dots & \mathbf{v}_n \end{bmatrix} = \begin{bmatrix} A\mathbf{v}_1 & \dots & A\mathbf{v}_n \end{bmatrix} \\ &= \begin{bmatrix} \sigma_1 \mathbf{u}_1 & \dots & \sigma_r \mathbf{u}_r & \mathbf{0} & \dots & \mathbf{0} \end{bmatrix} = U\Sigma \rightsquigarrow A = U\Sigma V^T. \end{aligned}$$

Example. Compute the SVD of $A = \begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix}$.

Solution.

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

Eigenvalues: $\lambda_1 = 3, \lambda_2 = 1, \lambda_3 = 0$

$$\text{Eigenbasis: } \underbrace{\begin{bmatrix} \frac{1}{\sqrt{6}} \\ -\frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \end{bmatrix}}_{:=\mathbf{v}_1}, \underbrace{\begin{bmatrix} -\frac{1}{\sqrt{2}} \\ 0 \\ \frac{1}{\sqrt{2}} \end{bmatrix}}_{:=\mathbf{v}_2}, \underbrace{\begin{bmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{bmatrix}}_{:=\mathbf{v}_3}$$

Singular values: $\sigma_1 := \sqrt{3}, \sigma_2 = 1$

$$V = \begin{bmatrix} \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \\ -\frac{2}{\sqrt{6}} & 0 & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \end{bmatrix}, \Sigma = \begin{bmatrix} \sqrt{3} & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

$$\mathbf{u}_1 := \frac{1}{\sigma_1} A\mathbf{v_1} = \frac{1}{\sqrt{3}} \begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{6}} \\ -\frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \end{bmatrix}$$

$$= \frac{1}{\sqrt{3}} \begin{bmatrix} -\frac{3}{\sqrt{6}} \\ \frac{3}{\sqrt{6}} \end{bmatrix} = \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$$

$$\mathbf{u}_2 = \frac{1}{\sigma_2} A\mathbf{v_2} = \begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ 0 \\ \frac{1}{\sqrt{2}} \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$$

$$\rightsquigarrow U = \begin{bmatrix} -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$$

Theorem 86. *Let A be an $m \times n$ matrix with rank r, and let $U = \begin{bmatrix} \mathbf{u}_1 & \dots \mathbf{u}_m \end{bmatrix}$, $V = \begin{bmatrix} \mathbf{v}_1 & \dots \mathbf{v}_n \end{bmatrix}$, Σ be such that $A = U\Sigma V^T$ is a SVD of A. Then*

- $(\mathbf{u}_1, \dots, \mathbf{u}_r)$ *is a basis of* $\mathrm{Col}(A)$.
- $(\mathbf{v}_1, \dots, \mathbf{v}_r)$ *is a basis of* $\mathrm{Col}(A^T)$.
- $(\mathbf{u}_{r+1}, \dots, \mathbf{u}_m)$ *is a basis of* $\mathrm{Nul}(A^T)$.
- $(\mathbf{v}_{r+1}, \dots, \mathbf{v}_n)$ *is a basis of* $\mathrm{Nul}(A)$.

Proof.

For simplicity, assume U, V, Σ are constructed using our algorithm.

Note that

$$\mathbf{u}_1 = \frac{1}{\sigma_1} A\mathbf{v}_1, \dots, \mathbf{u}_r = \frac{1}{\sigma_r} A\mathbf{v}_r$$

$$\rightsquigarrow \mathbf{u}_1, \dots, \mathbf{u}_r \in \mathrm{Col}(A)$$

$$\underset{\dim \mathrm{Col}(A) = r}{\rightsquigarrow} (\mathbf{u}_1, \dots, \mathbf{u}_r) \text{ basis of } \mathrm{Col}(A)$$

$$\rightsquigarrow (\mathbf{u}_{r+1}, \dots, \mathbf{u}_m) \text{ basis of } \mathrm{Col}(A)^\perp = \mathrm{Nul}(A^T)$$

Recall that

$$(\mathbf{v}_{r+1}, \dots, \mathbf{v}_n) \in \mathrm{Nul}(A^T A) = \mathrm{Nul}(A).$$

Since $\dim(\mathrm{Nul}(A)) = n - r$,

$$\rightsquigarrow (\mathbf{v}_{r+1}, \dots, \mathbf{v}_n) \text{ basis of } \mathrm{Nul}(A)$$

$$\rightsquigarrow (\mathbf{v}_1, \dots, \mathbf{v}_r) \text{ basis of } \mathrm{Nul}(A)^\perp = \mathrm{Col}(A^T)$$

LINEAR ALGEBRA

Low rank approximation via SVD

ILLINOIS

Department of Mathematics

Theorem 87. *Let A be an $m \times n$ matrix with rank r, and let $U = \begin{bmatrix} \mathbf{u}_1 & \dots \mathbf{u}_m \end{bmatrix}$, $V = \begin{bmatrix} \mathbf{v}_1 & \dots \mathbf{v}_n \end{bmatrix}$ be matrices with with orthonormal columns and Σ be a rectangular diagonal $m \times n$ matrix such that $A = U\Sigma V^T$ is an SVD of A. Then*

$$
\begin{aligned}
A &= \sigma_1 \mathbf{u}_1 \mathbf{v}_1^T + \sigma_2 \mathbf{u}_2 \mathbf{v}_2^T + \dots + \sigma_r \mathbf{u}_r \mathbf{v}_r^T \\
&= \begin{bmatrix} \mathbf{u}_1 & \dots & \mathbf{u}_r \end{bmatrix} \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_r \end{bmatrix} \begin{bmatrix} \mathbf{v}_1 & \dots & \mathbf{v}_r \end{bmatrix}^T
\end{aligned}
$$

Proof.

Suppose $m \leq n$.

$$
\begin{aligned}
A &= \begin{bmatrix} \mathbf{u}_1 & \dots & \mathbf{u}_m \end{bmatrix} \begin{bmatrix} \sigma_1 & & & & \\ & \ddots & & & \\ & & \sigma_m & & \end{bmatrix} \begin{bmatrix} \mathbf{v}_1 & \dots \mathbf{v}_n \end{bmatrix}^T \\
&= \begin{bmatrix} \sigma_1 \mathbf{u}_1 & \dots & \sigma_m \mathbf{u}_m & \mathbf{0} & \dots & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{v}_1 & \dots & \mathbf{v}_n \end{bmatrix}^T \\
&= \sigma_1 \mathbf{u}_1 \mathbf{v}_1^T + \dots + \sigma_m \mathbf{u}_m \mathbf{v}_m^T \\
&= \sigma_1 \mathbf{u}_1 \mathbf{v}_1^T + \dots + \sigma_r \mathbf{u}_r \mathbf{v}_r^T \qquad \text{since } \sigma_{r+1} = \dots = \sigma_m = 0
\end{aligned}
$$

Example. Use

$$\begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} = \begin{bmatrix} -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} \sqrt{3} & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{6}} & -\frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \end{bmatrix}$$

to write $\begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix}$ as a sum of rank 1 matrices.

Solution.

$$\begin{aligned} \begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} &= \sqrt{3} \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{6}} & -\frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} \end{bmatrix} + 1 \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{bmatrix} \\ &= \begin{bmatrix} -\frac{1}{2} & 1 & -\frac{1}{2} \\ \frac{1}{2} & -1 & \frac{1}{2} \end{bmatrix} + \begin{bmatrix} -\frac{1}{2} & 0 & \frac{1}{2} \\ -\frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix} \end{aligned}$$

Definition. For $k \le r$, define

$$A_k = \sigma_1 \mathbf{u}_1 \mathbf{v}_1^T + \sigma_2 \mathbf{u}_2 \mathbf{v}_2^T + \cdots + \sigma_k \mathbf{u}_k \mathbf{v}_k^T.$$

Idea. If $\sigma_1 \gg \sigma_2 \gg \ldots ...$, then A_k is a good approximation of A.

Definition. Let A be an $m \times n$ matrix with rank r. A **compact singular value decomposition** of A is a decomposition $A = U_c \Sigma_c V_c^T$ where

- $U_c = \begin{bmatrix} \mathbf{u}_1 & \ldots \mathbf{u}_r \end{bmatrix}$ is an $m \times r$ matrix with orthonormal columns,
- Σ_c is an $r \times r$ diagonal matrix with positive diagonal elements,
- $V_c = \begin{bmatrix} \mathbf{v}_1 & \ldots \mathbf{v}_r \end{bmatrix}$ is an $n \times r$ matrix with orthonormal columns.

Question. What is $\mathrm{Col}(V_c)$? What is $\mathrm{Col}(U_c)$?

Solution.

If we use our usual algorithm to compute an SVD, then

$$\mathrm{Col}(V_c) = \mathrm{Col}(A^T) \text{ and } \mathrm{Col}(U_c) = \mathrm{Col}(A).$$

This is also true if you use a different algorithm.

Example. Use

$$\begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} = \begin{bmatrix} -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} \sqrt{3} & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{6}} & -\frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \end{bmatrix}$$

to find a compact SVD of $\begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix}$.

Solution.

Note that $\begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix}$ has rank 2. Set

$$U_c = \begin{bmatrix} -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}, \ \Sigma_c = \begin{bmatrix} \sqrt{3} & 0 \\ 0 & 1 \end{bmatrix}, V_c = \begin{bmatrix} \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} \\ -\frac{2}{\sqrt{6}} & 0 \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} \end{bmatrix}$$

LINEAR ALGEBRA

The Pseudo-Inverse

ILLINOIS

Department of Mathematics

Definition. Let A be an $m \times n$ matrix with rank r. Given the compact singular value decomposition $A = U_c \Sigma_c V_c^T$ where

- $U_c = \begin{bmatrix} \mathbf{u}_1 & \ldots \mathbf{u}_r \end{bmatrix}$ is an $m \times r$ matrix with orthonormal columns,
- Σ_c is an $r \times r$ diagonal matrix with positive diagonal elements,
- $V_c = \begin{bmatrix} \mathbf{v}_1 & \ldots \mathbf{v}_r \end{bmatrix}$ is an $n \times r$ matrix with orthonormal columns,

we define the **pseudoinverse** A^+ of A as $V_c \Sigma_c^{-1} U_c^T$.

Example. Recall that

$$A := \begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} = \begin{bmatrix} -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} \sqrt{3} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} \\ -\frac{2}{\sqrt{6}} & 0 \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} \end{bmatrix}^T .$$

Determine A^+.

Solution.

$$A^+ = \begin{bmatrix} \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} \\ -\frac{2}{\sqrt{6}} & 0 \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{3}} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} \\ -\frac{2}{\sqrt{6}} & 0 \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} = \begin{bmatrix} -\frac{2}{3} & -\frac{1}{3} \\ \frac{1}{3} & -\frac{1}{3} \\ \frac{1}{3} & \frac{2}{3} \end{bmatrix}$$

Theorem 88. *Let* $\mathbf{v} \in \mathrm{Col}(A^T)$ *and* $\mathbf{w} \in \mathrm{Col}(A)$. *Then* $A^+A\mathbf{v} = \mathbf{v}$ *and* $AA^+\mathbf{w} = \mathbf{w}$.

Proof.

Since $\mathrm{Col}(V_c) = \mathrm{Col}(A^T)$ and V_c has orthonormal columns, we have for every $\mathbf{x} \in \mathbb{R}^n$

$$\mathrm{proj}_{\mathrm{Col}\, A^T}(\mathbf{x}) = \mathrm{proj}_{\mathrm{Col}(V_c)}(\mathbf{x}) = V_c V_c^T \mathbf{x}.$$

Thus since $\mathbf{v} \in \mathrm{Col}(A^T)$, we get

$$A^+A\mathbf{v} = V_c \Sigma_c^{-1} U_c^T U_c \Sigma_c V_c^T \mathbf{v} = V_c V_c^T \mathbf{v} = \mathrm{proj}_{\mathrm{Col}\, A^T}(\mathbf{v}) = \mathbf{v}.$$

Since $\mathrm{Col}(U_c) = \mathrm{Col}(A)$ and U_c has orthonormal columns, we have for every $\mathbf{x} \in \mathbb{R}^m$

$$\mathrm{proj}_{\mathrm{Col}\, A}(\mathbf{x}) = \mathrm{proj}_{\mathrm{Col}(U_c)}(\mathbf{x}) = U_c U_c^T \mathbf{x}.$$

Thus

$$AA^+\mathbf{w} = U_c \Sigma_c V_c^T V_c \Sigma_c^{-1} U_c^T \mathbf{w} = U_c U_c^T \mathbf{w} = \mathrm{proj}_{\mathrm{Col}\, A}(\mathbf{w}) = \mathbf{w}.$$

Remark. If A is $n \times n$ and invertible, then $\mathrm{Col}(A) = \mathbb{R}^n$. Thus $A^{-1} = A^+$.

Question. Let $\mathbf{v} \in \mathbb{R}^n$ such that $\underbrace{\mathbf{v}_r}_{\in \mathrm{Col}(A^T)} + \underbrace{\mathbf{v}_n}_{\in \mathrm{Nul}(A)} = \mathbf{v}$. What is $A^+A\mathbf{v}$?

Solution.

$$A^+A\mathbf{v} = A^+A(\mathbf{v}_r + \mathbf{v}_n) = A^+A\mathbf{v}_r + A^+A\mathbf{v}_n = \mathbf{v}_r.$$

Theorem 89. *Let A be an $m \times n$ matrix and let $\mathbf{b} \in \mathbb{R}^m$. Then $A^+\mathbf{b}$ is the LSQ solution of $A\mathbf{x} = \mathbf{b}$ (with minimum length).*

Proof.

Observe that

$$AA^+\mathbf{b} = U_c\Sigma_c V_c^T V_c \Sigma_c^{-1} U_c^T \mathbf{b} = U_c U_c^T \mathbf{b} = \mathrm{proj}_{\mathrm{Col}\,A}(\mathbf{b})$$

Thus $A^+\mathbf{b}$ is a LSQ solution of $A\mathbf{x} = \mathbf{b}$.

Remark. This is particularly useful, when solving many different LSQ problems of the form $A\mathbf{x} = \mathbf{b}$, where A stays the same, but $\mathbf{b}$ varies.

LINEAR ALGEBRA

PCA

ILLINOIS

Department of Mathematics

Setup.

- Given m objects, we measure the same n variables.
- Thus m samples of n-dimensional data $\rightsquigarrow$ $m \times n$ matrix (each row is a sample)
- Analyse this matrix to understand what drives the variance in the data.

Definition. Let $X = \begin{bmatrix} \mathbf{a}_1 & \dots & \mathbf{a}_m \end{bmatrix}^T$ be an $m \times n$ matrix. We define the **column average** $\mu(X)$ of X as

$$\mu(X) := \frac{1}{m}(\mathbf{a}_1 + \dots + \mathbf{a}_m).$$

We say X is **centered** if $\mu(X) = \mathbf{0}$.
For X centered, we define the **covariance matrix** $\mathrm{cov}(X)$ of X as $\frac{1}{m-1}X^T X$.

Remark.

- Not centered, replace X by $\begin{bmatrix} \mathbf{a}_1 - \mu(X) & \dots & \mathbf{a}_m - \mu(X) \end{bmatrix}^T$.
- If the columns of X are orthogonal, then $\mathrm{cov}(X)$ is a diagonal matrix $\rightsquigarrow$ each variable is independent.
- What if not? Idea: Pick an eigenbasis of $X^T X$.

Principal component analysis

- Input: centered $m \times n$-matrix X.
- Compute $\mathrm{cov}(X)$.
- Since $\mathrm{cov}(X)$ is symmetric, we can find an orthonormal eigenbasis $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ of $\mathrm{cov}(X)$ with eigenvalues $\lambda_1 \geq \cdots \geq \lambda_n \geq 0$.
- Write $\mathrm{cov}(X)$ as a sum of rank 1 matrices:

$$\mathrm{cov}(X) = \lambda_1 \mathbf{v}_1 \mathbf{v}_1^T + \cdots + \lambda_n \mathbf{v}_n \mathbf{v}_n^T.$$

- Each principal component $\mathbf{v}_i$ explains part of the variance of the data. The larger λ_i, the more of the variances is explained by $\mathbf{v}_i$.

Example. Let X be a centered 30×2-matrix. We plot the data and see that

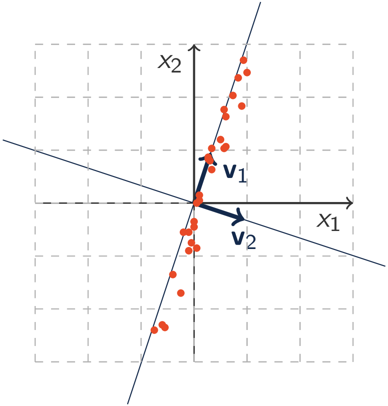
$$\mathrm{cov}(X) = \begin{bmatrix} 7.515 & 20.863 \\ 20.863 & 63.795 \end{bmatrix}.$$

An orthonormal eigenbasis is

$$\mathbf{v}_1 = \begin{bmatrix} 0.314 \\ 0.95 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 0.95 \\ -0.314 \end{bmatrix}$$

with $\lambda_1 = 70.685$ and $\lambda_2 = 0.625$. Thus

$$\mathrm{cov}(X) = \underbrace{\begin{bmatrix} 6.97 & 21.085 \\ 21.085 & 63.793 \end{bmatrix}}_{\lambda_1 \mathbf{v}_1 \mathbf{v}_1^T} + \underbrace{\begin{bmatrix} 0.564 & -0.186 \\ -0.186 & 0.0616 \end{bmatrix}}_{\lambda_2 \mathbf{v_2} \mathbf{v_2}^T}.$$



PCA using SVD.

- Let X be a centered data matrix. Observe that $X^TX = (m-1)\,\mathrm{cov}(X)$.
- To find orthonormal eigenbasis of $\mathrm{cov}(X)$, it is enough to find orthonormal eigenbasis of X^TX.
- Compute the SVD of $X \rightsquigarrow X = U\Sigma V^T$.
- The columns of $V = \begin{bmatrix} \mathbf{v}_1 & \dots & \mathbf{v}_n \end{bmatrix}$ are the desired orthonormal eigenbasis.
- If σ_i is the singular value for $\mathbf{v_i}$, then

$$\lambda_i = \frac{\sigma_i^2}{m-1}$$

LINEAR ALGEBRA

Review of Complex Numbers

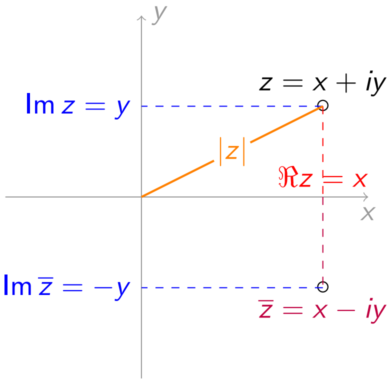
ILLINOIS

Department of Mathematics

Definition. $\mathbb{C} = \{x + iy \mid x, y \in \mathbb{R}\}$ where $i = \sqrt{-1}$, or $i^2 = -1$.

- The real part of z, denoted $\Re(z)$ is defined by $\Re(z) = x$.
- The imaginary part of z, denoted $\mathrm{Im}(z)$ is defined by $\mathrm{Im}(z) = y$.
- The complex conjugate of z, denoted $\overline{z}$, is defined by $\overline{z} = x - iy$.
- The absolute value, or magnitude of z, denoted $|z|$ or $\|z\|$, is given by

$$|z| = \sqrt{x^2 + y^2}.$$



Any point in $\mathbb{R}^2$ can be viewed as a complex number: $\binom{x}{y} \leftrightarrow x + iy$

Definition. Given $z = x + iy$, $w = u + iv$, we define

$$z + w = (x + u) + i(y + v)$$

$$\begin{aligned} zw &= (x + iy)(u + iv) \\ &= xu + x(iv) + (iy)u + (iy)(iv) \\ &= (xu - yv) + i(xv + yu) \end{aligned}$$

Example. Compute $(3 + 2i) + (4 - i)$ and $(3 + 2i)(4 - i)$.

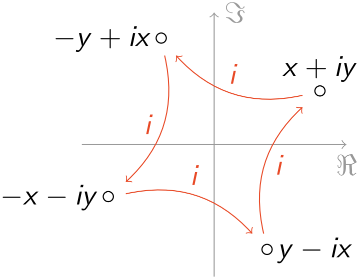
Solution.

$$\begin{aligned} (3 + 2i) + (4 - i) &= 7 + i \\ (3 + 2i)(4 - i) &= 14 + 5i \end{aligned}$$

Example. Compute $i(x + iy)$, $i^2(x + iy)$ and $i^3(x + iy)$ and $i^4(x + iy)$.

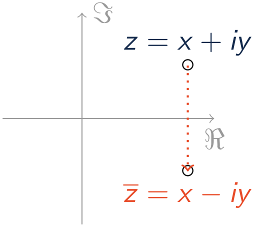
Solution.

$$\begin{aligned} i(x + iy) &= -y + ix \\ i^2(x + iy) &= i(-y + ix) = -x - iy \\ i^3(x + iy) &= i(-x - iy) = y - ix \\ i^4(x + iy) &= i(y - ix) = x + iy \end{aligned}$$



Theorem 90. *Let $z \in \mathbb{C}$.*

- $\overline{\overline{z}} = z$
- $|z|^2 = z\overline{z}$
- $|z| = |\overline{z}|$



Proof.

Given $z = x + iy$: $\overline{z} = x - iy$ and $|z| = \sqrt{x^2 + y^2}$.

$$\overline{\overline{z}} = \overline{\overline{x + iy}} = \overline{x - iy} = x + iy$$

$$\begin{aligned} z\overline{z} &= (x + iy)(x - iy) \\ &= x^2 - x(iy) + (iy)x - (iy)(iy) \\ &= x^2 + y^2 = |z|^2. \end{aligned}$$

$$|z| = |x + iy| = \sqrt{x^2 + y^2} = \sqrt{x^2 + (-y)^2} = |x - iy| = |\overline{z}|$$

LINEAR ALGEBRA

Complex Linear Algebra

ILLINOIS

Department of Mathematics

Goal. Use complex numbers (instead of real numbers) as scalars.

Definition. The (complex) vector space $\mathbb{C}^n$ is of all *complex* column vectors $\mathbf{z} = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{bmatrix}$,

where $z_1, z_2, \ldots, z_n$ are complex numbers.

- Now multiplication by a complex scalar makes sense.
- We can define subspaces, span, independence, basis, dimension for $\mathbb{C}^n$ in the usual way.
- We can multiply complex vectors by complex matrices. Column space and Null space still make sense.
- The only difference is the dot product, you need to use the complex conjugate to get a good notion of length:

$$\begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix} \cdot \begin{bmatrix} w_1 \\ \vdots \\ w_n \end{bmatrix} = z_1 \bar{w}_1 + z_2 \bar{w}_2 + \ldots z_n \bar{w}_n.$$

Example. Find the complex eigenvectors and eigenvalues of $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$.

Solution.

$$\det(A - \lambda I) = \begin{vmatrix} -\lambda & -1 \\ 1 & -\lambda \end{vmatrix} = \lambda^2 + 1$$

No real eigenvalues! The (complex) eigenvalues are $\lambda_1 = i$ and $\lambda_2 = -i$.

$$\lambda_1 = i: \ A - iI = \begin{bmatrix} -i & -1 \\ 1 & -i \end{bmatrix} \underset{R_1 \to R_1 + iR_2}{\rightsquigarrow} \begin{bmatrix} 0 & -1 + i(-i) \\ 1 & -i \end{bmatrix} \underset{R_1 \leftrightarrow R_2}{\rightsquigarrow} \begin{bmatrix} 1 & -i \\ 0 & 0 \end{bmatrix}$$

$$\rightsquigarrow \operatorname{Nul}(A - iI) = \operatorname{span}\left(\begin{bmatrix} i \\ 1 \end{bmatrix} \right)$$

$$\lambda_2 = -i: \ A + iI = \begin{bmatrix} i & -1 \\ 1 & i \end{bmatrix} \underset{R_1 \to R_1 - iR_2}{\rightsquigarrow} \begin{bmatrix} 0 & -1 - i(i) \\ 1 & i \end{bmatrix} \underset{R_1 \leftrightarrow R_2}{\rightsquigarrow} \begin{bmatrix} 1 & i \\ 0 & 0 \end{bmatrix}$$

$$\rightsquigarrow \operatorname{Nul}(A + iI) = \operatorname{span}\left(\begin{bmatrix} -i \\ 1 \end{bmatrix} \right)$$

Let's check: $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} i \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ i \end{bmatrix} = i \begin{bmatrix} i \\ 1 \end{bmatrix}$

Definition. Let A be an $m \times n$-matrix. The **conjugate matrix** $\overline{A}$ of A is obtained from A by taking the complex conjugate of each entry of A.

Theorem 91. *Let A be a matrix with real entries and λ is an eigenvalue of A. Then $\bar{\lambda}$ is also a eigenvalue. Furthermore, if $\mathbf{v}$ is an eigenvector with eigenvalue λ, then $\bar{\mathbf{v}}$ is an eigenvector with eigenvalue $\bar{\lambda}$.*

Proof.

Suppose $A\mathbf{v} = \lambda\mathbf{v}$. Observe that $\bar{A} = A$, because A only has real entries. Thus

$$A\overline{\mathbf{v}} = \overline{A}\overline{\mathbf{v}} = \overline{A\mathbf{v}} = \overline{\lambda\mathbf{v}} = \overline{\lambda}\overline{\mathbf{v}}.$$

Thus $\overline{\mathbf{v}}$ is an eigenvector of A to the eigenvalue $\overline{\lambda}$.

Definition. Let A be an $m \times n$-matrix. The **conjugate transpose** A^H of A is defined as $\overline{A}^T$. We say the matrix A is **Hermitian** if $A = A^H$.

Example. Find the eigenvectors and eigenvalues of $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$. Can complex numbers help you find an eigenbasis of A?

- $\det(A - \lambda I) = \begin{vmatrix} 1-\lambda & 1 \\ 0 & 1-\lambda \end{vmatrix} = (1-\lambda)^2$

  So: $\lambda = 1$ is the only eigenvalue (it has algebraic multiplicity 2).

- $\lambda = 1: \; A - I = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \implies \mathrm{Nul}(A) = \mathrm{span}\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right)$

  So the eigenspace still has only dimension 1. Complex numbers didn't help.

- We can not find an eigenbasis for this matrix.

  This kind of problem cannot really be fixed.

  We have to lower our expectations and look for generalized eigenvectors.

  These are solutions to $(A - \lambda I)^2 \mathbf{x} = \mathbf{0}, (A - \lambda I)^3 \mathbf{x} = \mathbf{0}, \ldots$