MA 106-2023-2 and MA110-2023-2 (1st half): Linear Algebra*

Rekha Santhanam**

Department of Mathematics, I.I.T. Bombay

February 20, 2024

This set of slides contains the material presented in my classes (Divisions 2 & 3) of MA106, and the first half of MA110 in Spring 2024 at IIT Bombay. The primary content was developed by me and my co-instructor**, Prof. Ananthnarayan Hariharan using the reference: *Linear Algebra and its Applications* by G. Strang, 4th Ed., Thomson.

The topics covered are:

1. LINEAR EQUATIONS & MATRICES
   (a) Linear Equations & Pivots
   (b) Matrices
   (c) Gaussian Elimination
   (d) Null Space & Column Space: Introduction
2. VECTOR SPACES
   (a) Vector Spaces & Subspaces
   (b) Linear Span & Independence
   (c) Basis & Dimension
   (d) Null Space, Column Space & Row Space
   (e) Linear Transformations
3. EIGENVALUE DECOMPOSITION
   (a) Eigenvalues & Eigenvectors
   (b) Diagonalization
4. ORTHOGONALITY & PROJECTIONS

*There maybe mild differences between the class slides and these. Please use with care

(a) Orthogonality

(b) Projection and Least Squares method

(c) Gram-Schmidt Process and Applications.

APPENDIX I -DETERMINANTS

NOTE: (i) The notation in these slides is the same as that discussed in class.
(ii) Work out as many examples as you can.

Chapter 1. Linear Equations & Matrices

1.1 Linear Equations & Pivots

What is Linear Algebra?

Is $(d, c) = (950, 0)$ the only solution of

$$d = -25c + 950?$$

This equation has several solutions; $(d, c) = (-300, 50)$, $(700, 10)$, $(945, 0.2)$, $(-3450, -100)$, etc.

Are all these solutions **permissible?**

Definitely not $(50, -300)$, $(945, 0.2)$ or $(3450, -100)$. Further assume delivery costs force the following linear relation on the number of deliveries

Then, $d = 10c + 250.$

Solve $d = 10c + 250$, $d = -25c + 950$ simultaneously to get $(450, 20)$.

Key note: In general, we want all possible solutions to the given system, i.e., without any co unlike the introductory example.

Solving equations, Example

Solve the system: $(1)\ \ 2x + y = 5,$ $(2)\ \ x + 2y = 4.$

Elimination of variables: Eliminate x by $(2) - 1/2 \times (1)$ to get $y = 1$, or

Cramer's Rule (determinant): $y = \dfrac{\begin{vmatrix} 2 & 5 \\ 1 & 4 \end{vmatrix}}{\begin{vmatrix} 2 & 1 \\ 1 & 2 \end{vmatrix}} = \dfrac{8-5}{4-1} = 1$

In either case, back substitution gives $x = 2$

We could also solve for x first and use back substitution for y. **Why ?**

Key Note: For a large system, say 100 equations in 100 variables, elimination method is preferred, since computing 101 determinants of size 100×100 is time-consuming.

Geometry of linear equations

Row method:

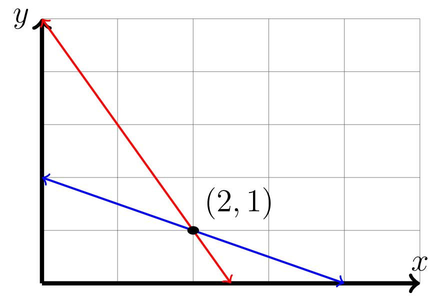
$\boxed{\text{2x+y=5}}$

and

$\boxed{\text{x+2y=4}}$

represent lines in $\mathbb{R}^2$ passing through $\boxed{(0,5)}$ and $\boxed{(5/2,0)}$ and through $\boxed{(0,2)}$ and $\boxed{(4,0)}$ respectively.

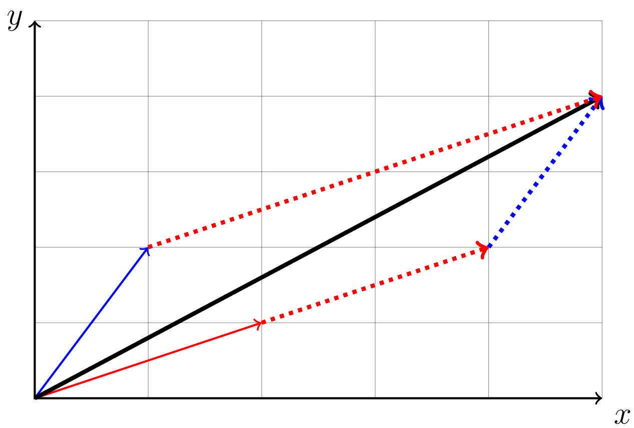
The intersection of the two lines is the unique point $(2,1)$. Hence $x=2$ and $y=1$ is the solution of above system of linear equations.



Column method: The system is $x\begin{pmatrix}2\\1\end{pmatrix}+y\begin{pmatrix}1\\2\end{pmatrix}=\begin{pmatrix}5\\4\end{pmatrix}$.

We need to find a *linear combination* of the column vectors on LHS to produce the column vector on RHS.

Geometrically this is same as completing the parallelogram with given directions and diagonal.



What are our choices of x and y here?

Equations in 3 variables: Geometry

Row method

A linear equation in 3 variables represents a plane in a 3 dimensional space $\mathbb{R}^3$.

Example: (1)

$$\boxed{\text{x+2y+3z=6}}$$

represents a plane passing through: $(0,0,2)$, $(0,3,0)$, $(6,0,0)$.

Example: (2)

$$\boxed{\text{x+2y+3z=0}}$$

represents a plane passing through: $(-2,1,0)$, $(-1,-1,1)$, $(2,-1,0)$.

In Example (2) we are looking for (x,y,z) such that $(x,y,z)\cdot(1,2,3)=0$, i.e., plane (2) is the set of all vectors perpendicular to the vector (1,2,3).

Equations in 3 variables: Examples

Example 1: (1) $x + 2y + 3z = 6$ (2) $x + 2y + 3z = 0$.

The two equations represent planes with normal vector (1,2,3) and are parallel to each other. **Exercise :** Prove this.

How many solutions can we find? There are *no solutions*.

Example 2: (1) $x + 2y + 3z = 0$ (2) $-x + 2y + z = 0$

The two equations represent planes passing through (0,0,0).

The intersection is non-empty, i.e., the system has at least one solution.
In fact, the *solution set* is a line passing through the origin.

Exercise: Find all the solutions in the second example.

3 equations in 3 variables

- Solving 3 by 3 system by the **row method** means finding an intersection of three planes, say P_1, P_2, P_3.

  This is same as the intersection of a line L

  (intersection of P_1 and P_2, if they are non-parallel) with the plane P_3.

- If the line L does not intersect the plane P_3, then the linear system has no solution, i.e., the system is *inconsistent*. Same is true if P_1 and P_2 were parallel.

- If the line L is contained in the plane P_3, then the system has infinitely many solutions.

  In this case, every point of L is a solution.

- **Exercise:** Workout some examples.

Linear Combinations

Column method:

Consider the 3×3 system:
x+2y+3z=2, -2x+3y=-5, -x+5y+2z=-4. Equivalently,

$$x \begin{pmatrix} 1 \\ -2 \\ -1 \end{pmatrix} + y \begin{pmatrix} 2 \\ 3 \\ 5 \end{pmatrix} + z \begin{pmatrix} 3 \\ 0 \\ 2 \end{pmatrix} = \begin{pmatrix} 2 \\ -5 \\ -4 \end{pmatrix}$$

We want a *linear combination* of the column vectors on LHS which is equal to RHS.
Observe: • $x = 1, y = -1, z = 1$ is a solution. **Q:** Is it unique?

• Since each column represents a vector in $\mathbb{R}^3$ from origin, we can find the solution geometrically, as in the 2×2 case.

Q: Can we do the same when number of variables are > 3?

Use other solving techniques to answer such questions.

Gaussian Elimination

Example: $\mathbf{2}u + v + w = 5, \quad 4u - 6v = -2, \quad -2u + 7v + 2w = 9.$

Algorithm: Eliminate u from last 2 equations by $(2) - \frac{4}{\mathbf{2}} \times (1)$, and $(3) - \frac{-2}{\mathbf{2}} \times (1)$ to get the *equivalent system:*

$$\mathbf{2}u + v + w = 5, \qquad \mathbf{-8}v - 2w = -12, \qquad \mathbf{8}v + 3w = 14$$

The coefficient used for eliminating a variable is called a *pivot*. The first pivot is 2. The second pivot is -8. The third pivot is 1. Eliminate v from the last equation to get an equivalent *triangular system*:

$\mathbf{2}u + v + w = 5, \quad \mathbf{-8}v - 2w = -12, \quad \mathbf{1} \cdot w = 2$

Solve this triangular system by *back substitution*, to get the *unique solution*
$w = 2, \quad v = 1, \quad u = 1.$

Matrix notation ($A\vec{x} = \vec{b}$) for linear systems

Consider the system
$2u + v + w = 5, \quad 4u - 6v = -2, \quad -2u + 7v + 2w = 9.$

Let $\vec{x} = \begin{pmatrix} u \\ v \\ w \end{pmatrix}$ be the unknown vector, and $\vec{b} = \begin{pmatrix} 5 \\ -2 \\ 9 \end{pmatrix}$.

The coefficient matrix is $A = \begin{pmatrix} 2 & 1 & 1 \\ 4 & -6 & 0 \\ -2 & 7 & 2 \end{pmatrix}$.

If we have m equations in n variables, then A has m rows and n columns, the column vector $\vec{b}$ has size m, and the unknown vector $\vec{x}$ has size n.

Notation: From now on, we will write $\vec{x}$ as x and $\vec{b}$ as b.

Elimination: Matrix form

Example: $2u + v + w = 5, \quad 4u - 6v = -2, \quad -2u + 7v + 2w = 9.$

Forward elimination in the *augmented* matrix form $[A|b]$:

(NOTE: The last column is the constant vector b).

$$\left(\begin{array}{ccc|c} \mathbf{2} & 1 & 1 & 5 \\ 4 & -6 & 0 & -2 \\ -2 & 7 & 2 & 9 \end{array}\right) \to \left(\begin{array}{ccc|c} \mathbf{2} & 1 & 1 & 5 \\ 0 & \mathbf{-8} & -2 & -12 \\ 0 & 8 & 3 & 14 \end{array}\right)$$

$$\to \left(\begin{array}{ccc|c} \mathbf{2} & 1 & 1 & 5 \\ 0 & \mathbf{-8} & -2 & -12 \\ 0 & 0 & \mathbf{1} & 2 \end{array}\right). \text{ Solution is: } x = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}.$$

Q: Is there a relation between 'pivots' and 'unique solution'?

Singular case: No solution

Example: $\mathbf{2}u + v + w = 5, \quad 4u - 6v = -2, \quad -2u + 7v + w = 9.$

$\boxed{\text{Step 1}}$ Eliminate u (using the 1st pivot 2) to get:

$\mathbf{2}u + v + w = 5, \quad \mathbf{-8}v - 2w = -12, \quad 8v + 2w = 14$

$\boxed{\text{Step 2}}$: Eliminate v (using the 2nd pivot -8) to get:

$\mathbf{2}u + v + w = 5, \quad -\mathbf{8}v - 2w = -12, \quad 0 = 2.$

The last equation shows that there is no solution,
i.e., the system is *inconsistent*.

Geometric reasoning: In Step 1, notice we get two distinct parallel planes $8v+2w = 12$ and $8v + 2w = 14$.

They have no point in common.

Note: The planes in the original system were not parallel, but in an equivalent system, we get two distinct parallel planes!

Singular Case: Infinitely many solutions

Example: $\mathbf{2}u + v + w = 5, \quad 4u - 6v = -2, \quad -2u + 7v + w = 7.$

$\boxed{\text{Step 1}}$ Eliminate u (using the 1st pivot 2) to get:

$\mathbf{2}u + v + w = 5, \quad \mathbf{-8}v - 2w = -12, \quad 8v + 2w = 12$

$\boxed{\text{Step 2}}$: Eliminate y (using the 2nd pivot -8) to get:

$\mathbf{2}u + v + w = 5, \quad -\mathbf{8}v - 2w = -12, \quad 0 = 0.$

There are only two equations. For every value of w, values for u and v are obtained by back-substitution, e.g, $(1, 1, 2)$ or $\left(\frac{7}{4}, \frac{3}{2}, 0\right)$. Hence the system has infinitely many solutions.

Geometric reasoning: In Step 1, notice we get two parallel planes $-8v - 2w = 12$ and $8v + 2w = 12$.

They give the same plane. Hence we are looking at the intersection of the two planes, $2u + v + w = 5$ and $8u + 2v = 12$, which is a line.

Some things to think about

- What are all the ways two different lines can intersect? What are all possible ways three different lines can intersect?
- What are all the ways two different planes can intersect? What are all possible ways three different plane can intersect?
- What is (if any) the geometric significance of the equation $x + y + z + w = 0$?
- Does the elimination method change the system of equations?
- Why does the solution set remain same all through the elimination method?

Singular Cases: Matrix Form

Eg. 1 $2u + v + w = 5, \quad 4u - 6v = -2, \quad -2u + 7v + w = 9.$

$$\begin{pmatrix} \mathbf{2} & 1 & 1 & | & 5 \\ 4 & -6 & 0 & | & -2 \\ -2 & 7 & 1 & | & 9 \end{pmatrix} \to \begin{pmatrix} \mathbf{2} & 1 & 1 & | & 5 \\ 0 & \mathbf{-8} & -2 & | & -12 \\ 0 & 0 & 0 & | & 2 \end{pmatrix}.$$

No Solution! Why?

Eg 2. $2u + v + w = 5, \quad 4u - 6v = -2, \quad -2u + 7v + w = 7.$

$$\begin{pmatrix} \mathbf{2} & 1 & 1 & | & 5 \\ 4 & -6 & 0 & | & -2 \\ -2 & 7 & 1 & | & 7 \end{pmatrix} \to \begin{pmatrix} \mathbf{2} & 1 & 1 & | & 5 \\ 0 & \mathbf{-8} & -2 & | & -12 \\ 0 & 0 & 0 & | & 0 \end{pmatrix}.$$

Infinitely many solutions! Why?

Q: Is there a relation between pivots and number of solutions? THINK!

Choosing pivots: Two examples

Example 1:

$-6v + 4w = -2, \quad u + v + 2w = 5, \quad 2u + 7v - 2w = 9.$

Forward elimination in the *augmented* matrix form $[A|b]$:

$$\begin{pmatrix} \mathbf{0} & -6 & 4 & | & -2 \\ 1 & 1 & 2 & | & 5 \\ 2 & 7 & -2 & | & 9 \end{pmatrix}$$

Coefficient of u in the first equation is 0. To get a non-zero coefficient we exchange the first two equations , i.e, interchange the first two rows of the matrix and get

$$\begin{pmatrix} 1 & 1 & 2 & | & 5 \\ 0 & -6 & 4 & | & -2 \\ 2 & 7 & -2 & | & 9 \end{pmatrix}$$

Exercise: Continue using elimination method; find all solutions.

Choosing pivots: Two examples

Example 2: 3 equations in 3 unknowns (u, v, w)

$0u + v + 2w = 1, \quad 0u + 6v + 4w = -2 \ , \quad 0u + 7v - 2w = -9.$

$$[A|b] = \begin{pmatrix} \mathbf{0} & \mathbf{1} & 2 & | & 1 \\ \mathbf{0} & 6 & 4 & | & -2 \\ \mathbf{0} & 7 & -2 & | & -9 \end{pmatrix} \to \begin{pmatrix} 0 & \mathbf{1} & 2 & | & 1 \\ 0 & 0 & \mathbf{-8} & | & -8 \\ 0 & 0 & -16 & | & -16 \end{pmatrix}$$

Coefficient of u is 0 in every equation.The first pivot is 1 and we eliminate v from the second and third equations. Solve for w and v to get $w = 1$, and $v = -1$.

Note: $(0, -1, 1)$ is a solution of the system. So is $(1, -1, 1)$.

In general, $(*, -1, 1)$ is a solution, for any real number $*$.

Observe: Unique solution is not an option. Why? This system has infinitely many solutions.

Q: Does such a system always have infinitely many solutions? **A:** Depends on the constant vector b.

Exercise: Find 3 vectors b for which the above system has (i) no solutions (ii) infinitely many solutions.

Summary: Pivots

- Can a pivot be zero? No (since we need to divide by it).
- If the first pivot (coefficient of 1st variable in 1st equation) is zero, then interchange it with next equation so that you get a non-zero first pivot. Do the same for other pivots.
- If the coefficient of the 1st variable is zero in every equation, consider the 2nd variable as 1st and repeat the previous step.
- Consider system of n equations in n variables.

  The non-singular case, i.e. the system has **exactly** n pivots:

  The system has a unique solution.

  The singular case, i.e., the system has **atmost** $n-1$ pivots: The system has no solutions, i.e., it is inconsistent, or it will have infinitely many solutions, provided it is consistent.

1.2 Matrices

What is a matrix?

A *matrix* is a collection of numbers arranged into a fixed number of rows and columns. If a matrix A has m rows and n columns, the size of A is $m \times n$.

The rows of A are denoted $A_{1*}, A_{2*}, \ldots, A_{m*}$, i.e., $A = \begin{pmatrix} A_{1*} \\ A_{2*} \\ \vdots \\ A_{m*} \end{pmatrix}$,

the columns are denoted $A_{*1}, A_{*2} \ldots, A_{*n}$, i.e.,

$A = \begin{pmatrix} A_{*1} & A_{*2} & \cdots & A_{*n} \end{pmatrix}$, and the (i,j)th entry is A_{ij} (or a_{ij}).

Operations on Matrices: Matrix Addition

Example 1. We know how to add two row or column vectors.

$$\begin{pmatrix}1 & 2 & 3\end{pmatrix} + \begin{pmatrix}-3 & -2 & -1\end{pmatrix} = \begin{pmatrix}-2 & 0 & 2\end{pmatrix} \quad \text{(component-wise)}$$

$\boxed{\text{We can add matrices if and only if they have the same size}}$,

and the addition is component-wise.

Example 2.

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

Thus

$$(A+B)_{i*} = A_{i*} + B_{i*} \text{ and } (A+B)_{*j} = A_{*j} + B_{*j}$$

Linear Systems: Multiplying a Matrix and a Vector

One row at a time (dot product): The system

$2u + v + w = 5, \quad 4u - 6v = -2, \quad -2u + 7v + 2w = 9$

can be rewritten using dot product as follows:

$$\begin{pmatrix}2 & 1 & 1\end{pmatrix}\begin{pmatrix}u\\ v\\ w\end{pmatrix} = 5, \quad \begin{pmatrix}4 & -6 & 0\end{pmatrix}\begin{pmatrix}u\\ v\\ w\end{pmatrix} = -2 \quad \text{and } \begin{pmatrix}-2 & 7 & 2\end{pmatrix}\begin{pmatrix}u\\ v\\ w\end{pmatrix} = 9.$$

Write the system in the $Ax = b$ form: $\begin{pmatrix}2 & 1 & 1\\ 4 & -6 & 0\\ -2 & 7 & 2\end{pmatrix}\begin{pmatrix}u\\ v\\ w\end{pmatrix} = \begin{pmatrix}2u+v+w\\ 4u-6v\\ -2u+7v+2w\end{pmatrix} = \begin{pmatrix}5\\ -2\\ 9\end{pmatrix}$

Note: No. of columns of A = length of the vector x.

Multiplication of a Matrix and a Vector

Dot Product (row method): Ax is obtained by taking dot product of each row of A with x.

If $A = \begin{pmatrix}A_{1*}\\ A_{2*}\\ A_{3*}\end{pmatrix}$, then $\boxed{Ax = \begin{pmatrix}A_{1*}\cdot x\\ A_{2*}\cdot x\\ A_{3*}\cdot x\end{pmatrix}}$

Linear Combinations (column method):

The column form of the system

$2u + v + w = 5, \quad 4u - 6v = -2, \quad -2u + 7v + 2w = 9$ is:

$$u\begin{pmatrix}2\\ 4\\ -2\end{pmatrix} + v\begin{pmatrix}1\\ -6\\ 7\end{pmatrix} + w\begin{pmatrix}1\\ 0\\ 2\end{pmatrix} = \begin{pmatrix}2 & 1 & 1\\ 4 & -6 & 0\\ -2 & 7 & 2\end{pmatrix}\begin{pmatrix}u\\ v\\ w\end{pmatrix}$$

Thus Ax is a linear combination of columns of A, with the coordinates of x as weights, i.e., $\boxed{Ax = uA_{*1} + vA_{*2} + wA_{*3}}$.

An Example

Let $A = \begin{pmatrix} 1 & 3 & -3 & -1 \\ 1 & 2 & 0 & -2 \\ 1 & 0 & -2 & 0 \end{pmatrix}$, $x = \begin{pmatrix} 2 \\ 1 \\ 1 \\ 2 \end{pmatrix}$, and $e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$.

$A_{1*} = \begin{pmatrix} 1 & 3 & -3 & -1 \end{pmatrix}$, $A_{2*} = \begin{pmatrix} 1 & 2 & 0 & -2 \end{pmatrix}$ $\quad A_{3*} = ?$.

Then $A_{1*} \cdot x = ?$, $A_{2*} \cdot x = 0$, $A_{3*} \cdot x = 0$, hence $Ax = \begin{pmatrix} ? \\ ? \\ ? \end{pmatrix}$.

Q: What is Ae_1? **A:** The first column A_{*1} of A.
Exercise:
What should x be so that $Ax = A_{*j}$, the jth column of A?

Observe: No. of rows of Ax = No. of rows of A,
and No. of columns of Ax = No. of columns of x.

Question: What can you say about the solutions of $Ax = 0$?

Operations on Matrices: Matrix Multiplication

Two matrices A and B can be multiplied if and only if

no. of columns of A = no. of rows of B.

If A is $m \times \underline{n}$ and B is $\underline{n} \times r$, then AB is $m \times r$.

Key Idea: We know how to multiply a matrix and a vector.

Column wise: Write B column-wise, i.e., let $B = \begin{pmatrix} B_{*1} & B_{*2} & \cdots & B_{*r} \end{pmatrix}$. Then

$$AB = \begin{pmatrix} AB_{*1} & AB_{*2} & \cdots & AB_{*r} \end{pmatrix}$$

Note: Each B_{*j} is a column vector of length n. Hence, AB_{*j} is a column vector of length m. So, the size of AB is $m \times r$.

Operations on Matrices: Matrix Multiplication

Row wise: Write A row-wise, i.e., let $A_{1*}, \ldots, A_{m*}$ be the rows of A. Then

$$AB = \begin{pmatrix} A_{1*} \\ \vdots \\ A_{m*} \end{pmatrix} B = \begin{pmatrix} A_{1*}B \\ \vdots \\ A_{m*}B \end{pmatrix}$$

Note: Each A_{i*} is a row vector of size $1 \times n$. Hence, $A_{i*}B$ is a row vector of size $1 \times r$. So, the size of AB is $m \times r$.

WORKING RULE:

The entry in the ith row and jth column of AB is the dot product of the ith row of A with the jth column of B, i.e., $(AB)_{ij} = A_{i*} \cdot B_{*j}$.

Properties of Matrix Multiplication

If A is $m \times n$, B is $n \times r$, C is $r \times l$.

- $(AB)_{ij} = A_{i*} \cdot B_{*j} =$ (ith row of A) $\cdot$ (jth column of B)
- jth column of $AB = A \cdot$ (jth column of B) , i.e., $(AB)_{*j} = AB_{*j}$.
- ith row of $AB =$(ith row of A) $\cdot$ B , i.e., $(AB)_{i*} = A_{i*}B$.
- (associativity) $(AB)C = A(BC)$. Why?
- (distributivity) $A(B+C) = AB + AC$. How to verify?

  $(B+C)D = BD + CD$. Why?
- (non-commutativity) $AB \neq BA$, in general. Why?

  Find examples.

Matrix Multiplication: Examples

Examples:

$$A = \begin{pmatrix} 2 & 1 & 1 \\ 4 & -6 & 0 \\ -2 & 7 & 2 \end{pmatrix}, B = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 2 & 4 \end{pmatrix}, I = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$ **(Identity)**

- $AB =$ **??**
- **size of** BA **is** $\_\_ \times \_\_$
- $BA = \begin{pmatrix} 4 & 10 & 7 \\ 4 & 18 & 10 \end{pmatrix}$,
- **and** $IA = A = AI$.

Questions to think about

- What does having a column of zeros in the augmented system signify for the solution of the corresponding system of linear equations? How are the pivots and solution set related?
- Recall Ae_j picks out the j^{th} column. What matrix multiplication will pick out the i^{th} row of A.
- The system $Ax = 0$ always has a solution. What does $Ax = 0$ having unique or infinitely many solutions signify geometrically for A?

Matrix Multiplication: Examples

Examples:

$$A = \begin{pmatrix} 2 & 1 & 1 \\ 4 & -6 & 0 \\ -2 & 7 & 2 \end{pmatrix}, \quad E = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$(Permutation) \quad P = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} e_2 & e_1 & e_3 \end{pmatrix}$$

Then $AP = \begin{pmatrix} Ae_2 & Ae_1 & Ae_3 \end{pmatrix} = \begin{pmatrix} A_{*2} & A_{*1} & A_{*3} \end{pmatrix}$

Exercise: Find EA and PA.

Question: Can you obtain EA and PA directly from A? How?

Transpose A^T of a Matrix A

Defn. The i-th row of A is the i-th column of A^T, the transpose of A and vice-versa. Hence if $A_{ij} = a$, then $(A^T)_{ji} = a$.

Example: If $A = \begin{pmatrix} 1 & 2 & 3 \\ 0 & -2 & 1 \end{pmatrix}$, then $A^T = \begin{pmatrix} 1 & 0 \\ 2 & -2 \\ 3 & 1 \end{pmatrix}$.

- If A is $m \times n$, then A^T is $n \times m$.
- If A is upper triangular, then A^T is lower triangular.
- $(A^T)^T = A$, $\boxed{(A+B)^T = A^T + B^T}$.
- $\boxed{(AB)^T = B^T A^T}$. *Proof.* Exercise.

Symmetric Matrix

Defn. If $A^T = A$, then A is called a *symmetric* matrix.

Note: A symmetric matrix is always $n \times n$.

Examples: $A = \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}$, $B = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ are symmetric.

- If A, B are symmetric, then AB may **NOT be symmetric.**

In the above case, $AB = \begin{pmatrix} 2 & 1 \\ 3 & 2 \end{pmatrix}$.

- If A and B are symmetric, then $A + B$is symmetric. Why?
- If A is a $n \times n$ matrix , $A + A^T$ is symmetric.Why?
- For any $m \times n$ matrix B, BB^T and $B^T B$ are symmetric.Why?

Exercise: If $A^T = -A$, we say that A is *skew-symmetric*.
Verify if similar observations are true for skew-symmetric matrices.

Inverse of a Matrix

Defn. Given A of size $n \times n$, we say B is an inverse of A if $AB = I = BA$. If this happens, we say A is *invertible*.

- What would be the inverse of $\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$?

- An inverse may not exist. Find an example. *Hint:* $n = 1$.

- An inverse of A, if it exists, has size $n \times n$.

- If the inverse of A exists, it is unique, and is denoted A^{-1}.Why unique?

Proof. Let B and C be inverses of A.

$$\begin{array}{ll} \Rightarrow BA = I & \text{by definition of inverse.} \\ \Rightarrow (BA)C = IC & \text{multiply both sides on the right by } C. \\ \Rightarrow B(AC) = IC & \text{by associativity.} \\ \Rightarrow BI = IC & \text{since } C \text{ is an inverse of } A. \\ \Rightarrow B = C & \text{by property of the identity matrix } I. \end{array}$$

- If A and B are invertible, what about AB? AB is invertible, with inverse $(AB)^{-1} = B^{-1}A^{-1}$.

Proof. Exercise.

- If A, B are invertible, what about $A + B$? $A + B$ may not be invertible.

Example: $I + (-I) = (0)$.

- If A is invertible, what about A^T? A^T is invertible with inverse $(A^T)^{-1} = (A^{-1})^T$.

Proof. Use $AA^{-1} = I$. Take transpose.

- If A is symmetric and invertible then, is A^{-1} symmetric?

Yes. *Proof.* Exercise!

- (Identity) $I^{-1} = I$.

Inverses and Linear Systems

- If A is invertible then the system $Ax = b$ has a solution,
for every constant vector b, namely $x = A^{-1}b$. Is this unique?

- Since $x = 0$ is always a solution of $Ax = 0$, if $Ax = 0$ has a non-zero solution, then A is not invertible by the last remark.

- If A is invertible, then the Gaussian elimination of A produces n pivots.

EXERCISE:

1. A diagonal matrix A is invertible if and only if ____.
(Hint: When are the diagonal entries pivots?)
2. When is an upper triangular matrix invertible?

• Since $AB = \begin{pmatrix} AB_{*1} \ AB_{*2} \cdots AB_{*n} \end{pmatrix}$ and $I = \begin{pmatrix} e_1 \ e_2 \cdots e_n \end{pmatrix}$, if $B = A^{-1}$, then B_{*j} is a solution of $Ax = e_j$ for all j.

• Strategy to find A^{-1}: Let A be an $n \times n$ invertible matrix. Solve $Ax = e_1$, $Ax = e_2$, . . . , $Ax = e_n$.

Solutions to Multiple Systems

Q: Let $A = \begin{pmatrix} 0 & 1 & -1 \\ 1 & 0 & 2 \\ 1 & 2 & 0 \end{pmatrix}$, $b_1 = \begin{pmatrix} -1 \\ 2 \\ 0 \end{pmatrix}$ $b_2 = \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}$. Solve for $Ax = b_1$ and $Ax = b_2$.

Do we apply Gaussian Elimination on two augmented matrices?

Rephrased question: Let $B = \begin{pmatrix} b_1 & b_2 \end{pmatrix}$. Is there a matrix C such that $AC = B$, i.e., such that $AC_{*1} = b_1$, $AC_{*2} = b_2$?

$$[A|B] = \left(\begin{array}{ccc|cc} 0 & 1 & -1 & -1 & 1 \\ 1 & 0 & 2 & 2 & 0 \\ 1 & 2 & 0 & 0 & 2 \end{array}\right) \xrightarrow{R_1 \leftrightarrow R_2} \left(\begin{array}{ccc|cc} \mathbf{1} & 0 & 2 & 2 & 0 \\ 0 & 1 & -1 & -1 & 1 \\ 1 & 2 & 0 & 0 & 2 \end{array}\right)$$

$$\xrightarrow{R_3 - R_1} \left(\begin{array}{ccc|cc} 1 & 0 & 2 & 2 & 0 \\ 0 & \mathbf{1} & -1 & -1 & 1 \\ 0 & 2 & -2 & -2 & 2 \end{array}\right) \xrightarrow{R_3 - 2R_2} \left(\begin{array}{ccc|cc} 1 & 0 & 2 & 2 & 0 \\ 0 & 1 & -1 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

Q: Are $Ax = b_1$ and $Ax = b_2$ both consistent?

Q: Given matrices A, $B = \begin{pmatrix} b_1 & b_2 \end{pmatrix}$, is there a matrix C such that $AC = B$?

$$[A|B] = \left(\begin{array}{ccc|cc} 0 & 1 & -1 & -1 & 1 \\ 1 & 0 & 2 & 2 & 0 \\ 1 & 2 & 0 & 0 & 2 \end{array}\right) \longrightarrow \left(\begin{array}{ccc|cc} 1 & 0 & 2 & 2 & 0 \\ 0 & 1 & -1 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

A solution to $Ax = b_1$ is $e_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$, and to $Ax = b_2$ is $e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$.

(Verify)! So $C = \begin{pmatrix} e_3 & e_2 \end{pmatrix}$ works! Is it unique?

Revisit the question about matrix inverses. Can you find inverse of a matrix this way?

Finding inverse of matrix

STRATEGY: . Let A be an $n \times n$ matrix. If $v_1, v_2, \ldots, v_n$ are solutions of $Ax = e_1$, $Ax = e_2$, . . . , $Ax = e_n$ respectively, then if it exists, $\boxed{A^{-1} = \begin{pmatrix} v_1 & v_2 & \cdots & v_n \end{pmatrix}.}$

If $Ax = e_j$ is not solvable for some j, then A is not invertible.

THUS, finding A^{-1} reduces to solving multiple systems of linear equations with the same coefficient matrix.

Consider the previous example, A. Is it invertible?

$$A = \begin{pmatrix} 0 & 1 & -1 \\ 1 & 0 & 2 \\ 1 & 2 & 0 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix}$$

Observe: In the above process, we used a *row exchange*: $R_1 \leftrightarrow R_2$ and *elimination using pivots*: $R_3 = R_3 - R_1$, $R_3 = R_3 - 2R_2$. Row operations can be achieved by left multiplication by special matrices.

1.3 Gaussian Elimination

Row Operations: Elementary Matrices

Example: $E\mathrm{x} = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} u \\ v \\ w \end{pmatrix} = \begin{pmatrix} u \\ v - 2u \\ w \end{pmatrix}$.

If $A = \begin{pmatrix} A_{*1} & A_{*2} & A_{*3} \end{pmatrix}$, then $EA = \begin{pmatrix} EA_{*1} & EA_{*2} & EA_{*3} \end{pmatrix}$.

Thus, EA has the same effect on A as the row operation $R_2 \mapsto R_2 + (-2)R_1$ on the matrix A.

Note: E is obtained from the identity matrix I by the row operation $R_2 \mapsto R_2 + (-2)R_1$.

Such a matrix (diagonal entries 1 and atmost one off-diagonal entry non-zero) is called an *elementary* matrix.

Notation: $\boxed{E := E_{21}(-2).}$ Similarly define $E_{ij}(\lambda)$.

Row Operations: Permutation Matrices

Example: $P\ x = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} u \\ v \\ w \end{pmatrix} = \begin{pmatrix} v \\ u \\ w \end{pmatrix}$

If $A = \begin{pmatrix} A_{*1} & A_{*2} & A_{*3} \end{pmatrix}$, then $PA = \begin{pmatrix} PA_{*1} & PA_{*2} & PA_{*3} \end{pmatrix}$.

Thus PA has the same effect on A as the row interchange $R_1 \leftrightarrow R_2$.

Note: We get P from the I by interchanging first and second rows. A matrix is called a *permutation* matrix if it is obtained from identity by row exchanges (possibly more than one).

Notation: $\boxed{P = P_{12}.}$ Similarly define P_{ij}.

Remark: Row operations correspond to multiplication by elementary matrices $E_{ij}(\lambda)$ or permutation matrices P_{ij} on the left.

Things to think about

• Complete the proofs left as exercise.

• Currently we can are unable to show that if $AB = I$ then $BA = I$ for square matrices A and B. Why so?

• Can you rephrase what we proved about transposes as a property of the transpose function from the set of $m \times n$ matrices to $n \times m$ matrices?

• Show that both Elementary matrices and Permutation matrices are invertible.

• Can you write down the precise inverse for a given elementary matrix or a permutation matrix.

Elementary Matrices: Inverses

For any $n \times n$ matrix A, observe that the row operations $R_2 \mapsto R_2 - 2R_1, R_2 \mapsto R_2 + 2R_1$ leave the matrix unchanged.

In matrix terms, $E_{21}(2)E_{21}(-2)A = IA = A$ since

$$E_{21}(-2)\ E_{21}(2) = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

• If $E_{21}(\lambda) = \begin{pmatrix} 1 & 0 & 0 \\ \lambda & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, what is your guess for $E_{21}(\lambda)^{-1}$? Verify.

• Let $P_{12} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} e_2^T \\ e_1^T \\ e_3^T \end{pmatrix}$. What is P_{12}^T? $P_{12}^T P_{12}$? P_{12}^{-1}?

Permutation Matrices: Inverses

Notice that the row interchange $R_1 \leftrightarrow R_2$ followed by $R_1 \leftrightarrow R_2$ leaves a matrix unchanged.

In matrix terms, $P_{12}P_{12}A = IA = A$, since

$$P_{12}P_{12} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

• Let P_{ij} be obtained by interchanging the ith and jth rows of I.
Show that $P_{ij}^T = P_{ij} = P_{ij}^{-1}$.

• Let $P = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} e_3^T \\ e_1^T \\ e_2^T \end{pmatrix}$. Show that $P = P_{12}P_{23}$.
Hence, $P^{-1} = (P_{12}P_{23})^{-1} = P_{23}^{-1}P_{12}^{-1} = P_{23}^T P_{12}^T = P^T$.

Elimination using Elementary Matrices

$$\text{Consider} \quad \begin{pmatrix} 2 & 1 & 1 \\ 4 & -6 & 0 \\ -2 & 7 & 2 \end{pmatrix} \begin{pmatrix} u \\ v \\ w \end{pmatrix} = \begin{pmatrix} 5 \\ -2 \\ 9 \end{pmatrix} \quad (Ax = b)$$

$\boxed{\text{Step 1}}$ Eliminate u by $R_2 \mapsto R_2 + (-2)R_1, \quad R_3 \mapsto R_3 + R_1$.

This corresponds to multiplying both sides on the left first by $E_{21}(-2)$ and then by $E_{31}(1)$. The equivalent system is:

$$E_{31}(1)E_{21}(-2)Ax = E_{31}(1)E_{21}(-2)b, \text{ i.e.,}$$
$$\begin{pmatrix} 2 & 1 & 1 \\ 0 & -8 & -2 \\ 0 & 8 & 3 \end{pmatrix} \begin{pmatrix} u \\ v \\ w \end{pmatrix} = \begin{pmatrix} 5 \\ -12 \\ 14 \end{pmatrix}.$$

Elimination using Elementary Matrices

$\boxed{\text{Step 2}}$ Eliminate v by $R_3 \mapsto R_3 + R_2$,

i.e., multiply both sides by $E_{32}(1)$ to get $Ux = c$,

where $U = E_{32}(1)E_{31}(1)E_{21}(-2)A = \begin{pmatrix} 2 & 1 & 1 \\ 0 & -8 & -2 \\ 0 & 0 & 1 \end{pmatrix}$ and $c = E_{32}(1)E_{31}(1)E_{21}(-2)b = \begin{pmatrix} 5 \\ -12 \\ 2 \end{pmatrix}$.

Elimination changed A to an upper triangular matrix and reduced the problem to solving $Ux = c$.

Observe: The pivots of the system $Ax = b$ are *the diagonal entries of* U.

Triangular Factorization

Thus $\boxed{Ax = b}$ is equivalent to $\boxed{Ux = c}$.

where

$$E_{32}(1)\ E_{31}(1)\ E_{21}(-2)\ A = U$$

Multiply both sides by $E_{32}(-1)$ on the left:

$$E_{31}(1)\ E_{21}(-2)\ A = E_{32}(-1)U$$

Multiply first by $E_{31}(-1)$ and then $E_{21}(2)$ on the left:

$$A = E_{21}(2)\ E_{31}(-1)\ E_{32}(-1)\ U = LU$$

where U is upper triangular, which is obtained by *forward elimination*, with diagonal entries as pivots and $L = E_{21}(2)\ E_{31}(-1)\ E_{32}(-1)$.

Note that each $E_{ij}(a)$ is a *lower triangular*. Product of lower triangular matrices is lower triangular. In particular L is lower triangular, where

$L = E_{21}(2)\ E_{31}(-1)\ E_{32}(-1) =$

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

Observe: L is lower triangular with diagonal entries 1 and *below the diagonals are* the multipliers.
($2, -1, -1$ in the earlier example).

LU Decomposition

If A is an $n \times n$ matrix, *with no row interchanges needed* in the Gaussian elimination of A, then $\boxed{A = L\,U}$, where

- U is an upper triangular matrix, which is obtained by forward elimination, with non-zero diagonal entries as pivots.

- L is a lower triangular with diagonal entries 1 and with the multipliers needed in the elimination algorithm below the diagonals.

Q: What happens if row exchanges are required?

LU Decomposition: with Row Exchanges

Example: $A = \begin{pmatrix}0&2\\3&4\end{pmatrix}$. A can not be factored as LU. (Why?) How to verify?

The 1st step in the Gaussian elimination of A is a row exchange.

$$P_{12}\ A = \begin{pmatrix}0&1\\1&0\end{pmatrix}\begin{pmatrix}0&2\\3&4\end{pmatrix} = \begin{pmatrix}3&4\\0&2\end{pmatrix}$$

Now elimination can be carried out without row exchanges.

- If A is an $n \times n$ non-singular matrix, then there is a matrix P which is a permutation matrix (needed to take care of row exchanges in the elimination process) such that $\boxed{PA = LU}$, where L and U are as defined earlier. Why?

Q: What happens when A is an $m \times n$ matrix? **A:** Coming Soon!

Application 1: Solving systems of equations

Let $A = \begin{pmatrix}1&2&3\\-2&-12&-5\\1&-6&3\end{pmatrix} = \begin{pmatrix}1&0&0\\-2&1&0\\1&1&1\end{pmatrix}\begin{pmatrix}1&2&3\\0&-8&1\\0&0&-1\end{pmatrix}$

To solve $Ax = b$, we can solve two triangular systems $Lc = b$ and $Ux = c$. Then $Ax = LUx = Lc = b$.

Take $b = \begin{pmatrix} 1 \\ 2 \\ 5 \end{pmatrix}$. First solve $\begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 5 \end{pmatrix}$.

We get $c_1 = 1$, $-2c_1 + c_2 = 2 \Rightarrow c_2 = 4$, and similarly $c_3 = 0$.

Now solve $\begin{pmatrix} 1 & 2 & 3 \\ 0 & -8 & 1 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} u \\ v \\ w \end{pmatrix} = \begin{pmatrix} 1 \\ 4 \\ 0 \end{pmatrix}$.

We get $w = 0$, $v = -1/2$, $u = 2$.

Applications: 2. Invertibility of a Matrix

Let A be $n \times n$, P, L and U as before be such that $PA = LU$.

- P is invertible and $P^{-1} = P^T \Rightarrow A = P^{-1}LU$.
- L is lower triangular, with diagonal entries 1 $\Rightarrow$ L is invertible.

Q: What is L^{-1}? e.g., Try $L = E_{21}(2)E_{31}(-1)E_{32}(-1)$ first.

- The non-zero diagonal entries of U are the pivots of A.

Thus, A invertible $\Rightarrow$ A has n pivots
$\Rightarrow$ all diagonal entries of U are non-zero $\Rightarrow$ U is invertible.

Why? Hint: U^T is invertible.

Conversely, suppose U is invertible. Then A is invertible and has n pivots. Why? Moreover, $A^{-1} =$ ______.

We have proved:

$\boxed{A \text{ is invertible} \Leftrightarrow U \text{ is invertible} \Leftrightarrow A \text{ has } n \text{ pivots.}}$

Computing the Inverse

Observe: $A = L\,U \Rightarrow A^{-1} = U^{-1}\,L^{-1}$.

Example: $A = \begin{pmatrix} 2 & 1 & 1 \\ 4 & -6 & 0 \\ -2 & 7 & 2 \end{pmatrix}$ is invertible. Find A^{-1}.

If $A^{-1} = \begin{pmatrix} x_1 & x_2 & x_3 \end{pmatrix}$, where x_i is the i-th column of A^{-1}, then $AA^{-1} = I$ gives three systems of linear equations

$$Ax_1 = e_1, \quad Ax_2 = e_2, \quad Ax_3 = e_3$$

where e_i is the i-th column of I. Since the coefficient matrix A is same in three systems, we can solve them simultaneously as follows:

Calculation of A^{-1} : Gauss-Jordan Method

Steps: $(A|I) \longrightarrow (U|L^{-1}) \longrightarrow (I|U^{-1}L^{-1})$.

$$(A \mid e_1 \quad e_2 \quad e_3) = \begin{pmatrix} 2 & 1 & 1 & | & 1 & 0 & 0 \\ 4 & -6 & 0 & | & 0 & 1 & 0 \\ -2 & 7 & 2 & | & 0 & 0 & 1 \end{pmatrix}$$

$$\xrightarrow[R_3+R_1]{R_2-2R_1} \begin{pmatrix} \mathbf{2} & 1 & 1 & | & 1 & 0 & 0 \\ 0 & -8 & -2 & | & -2 & 1 & 0 \\ 0 & 8 & 3 & | & 1 & 0 & 1 \end{pmatrix}$$

$$\xrightarrow{R_3+R_2} \begin{pmatrix} \mathbf{2} & 1 & 1 & | & 1 & 0 & 0 \\ 0 & \mathbf{-8} & -2 & | & -2 & 1 & 0 \\ 0 & 0 & \mathbf{1} & | & -1 & 1 & 1 \end{pmatrix}$$

$$\xrightarrow[R_1-R_3]{R_2+2R_3} \begin{pmatrix} \mathbf{2} & 1 & 0 & | & 2 & -1 & -1 \\ 0 & \mathbf{-8} & 0 & | & -4 & 3 & 2 \\ 0 & 0 & \mathbf{1} & | & -1 & 1 & 1 \end{pmatrix}$$

$$\xrightarrow{R_1+\frac{1}{8}R_2} \begin{pmatrix} \mathbf{2} & 0 & 0 & | & 12/8 & -5/8 & -6/8 \\ 0 & \mathbf{-8} & 0 & | & -4 & 3 & 2 \\ 0 & 0 & \mathbf{1} & | & -1 & 1 & 1 \end{pmatrix}$$

$$\text{Divide by pivots} \longrightarrow \begin{pmatrix} 1 & 0 & 0 & | & \mathbf{12/16} & \mathbf{-5/16} & \mathbf{-6/16} \\ 0 & 1 & 0 & | & \mathbf{4/8} & \mathbf{-3/8} & \mathbf{-2/8} \\ 0 & 0 & 1 & | & \mathbf{-1} & \mathbf{1} & \mathbf{1} \end{pmatrix}$$

$$= (I \mid U^{-1}L^{-1}) \quad = \quad (I \mid \mathbf{A^{-1}})$$

Echelon Form

Recall: If A is $n \times n$, then $PA = LU$, where P is a product of permutation matrices, L is lower triangular, U is upper triangular, and all of size $n \times n$.

Q: What happens when A is not a square matrix?

Let $A = \begin{pmatrix} 1 & 2 & 3 & 5 \\ 2 & 4 & 8 & 12 \\ 3 & 6 & 7 & 13 \end{pmatrix}$. By elimination, we see: $A \to \begin{pmatrix} 1 & 2 & 3 & 5 \\ 0 & 0 & 2 & 2 \\ 0 & 0 & -2 & -2 \end{pmatrix} \to \begin{pmatrix} 1 & 2 & 3 & 5 \\ 0 & 0 & 2 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix} = U$.

Thus $A = LU$, where $L = E_{21}(2)E_{31}(3)E_{32}(-1) = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & -1 & 1 \end{pmatrix}$.

If A is $m \times n$, we can find P, L and U as before. In this case, L and P will be $m \times m$ and U will be $m \times n$.

U has the following properties:

1. Pivots are the 1st nonzero entries in their rows.
2. Entries below pivots are zero, by elimination.
3. Each pivot lies to the right of the pivot in the row above.
4. Zero rows are at the bottom of the matrix.

U is called an *echelon form* of A.

What are all possible 2×2 echelon forms: Let $\bullet$ = pivot entry.

$\begin{pmatrix} \bullet & * \\ 0 & \bullet \end{pmatrix}$, $\begin{pmatrix} \bullet & * \\ 0 & 0 \end{pmatrix}$, $\begin{pmatrix} 0 & \bullet \\ 0 & 0 \end{pmatrix}$, and $\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$.

Row Reduced Form

To obtain the row reduced form R of a matrix A:

1) Get the echelon form U. 2) Make the pivots 1.

3) Make the entries above the pivots 0.

Ex: Find all possible 2×2 row reduced forms.

Eg. Let $A = \begin{pmatrix} 1 & 2 & 3 & 5 \\ 2 & 4 & 8 & 12 \\ 3 & 6 & 7 & 13 \end{pmatrix}$. Then $U = \begin{pmatrix} \mathbf{1} & 2 & 3 & 5 \\ 0 & 0 & \mathbf{2} & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix}$.

Divide by pivots: $R_2/2$ gives $\begin{pmatrix} \mathbf{1} & 2 & 3 & 5 \\ 0 & 0 & \mathbf{1} & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$.

By $R_1 = R_1 - 3R_2$, Row reduced form of A: $R = \begin{pmatrix} \mathbf{1} & 2 & 0 & 2 \\ 0 & 0 & \mathbf{1} & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$

U and R are used to solve $Ax = 0$ and $Ax = b$.

1.4 Null Space and Column Space: Introduction

Null Space: Solution of $Ax = 0$

Let A be $m \times n$. **Q:** For which $x \in \mathbb{R}^n$, is $Ax = 0$?

The Null Space of A, denoted by $N(A)$,
is the set of all vectors x in $\mathbb{R}^n$ such that $Ax = 0$.

Example 1: $A = \begin{pmatrix} 1 & 2 & 3 & 5 \\ 2 & 4 & 8 & 12 \\ 3 & 6 & 7 & 13 \end{pmatrix}$. Are the following in $N(A)$?

$x = \begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \end{pmatrix}$? $y = \begin{pmatrix} -5 \\ 0 \\ 0 \\ 1 \end{pmatrix}$? $z = \begin{pmatrix} -2 \\ 0 \\ -1 \\ 1 \end{pmatrix}$?

Note: x is in $N(A) \Leftrightarrow A_{1*} \cdot x = 0$, $A_{2*} \cdot x = 0$, and $A_{3*} \cdot x = 0$,
i.e., x is perpendicular to every row of A.

Linear Combinations in $N(A)$

EXAMPLE 1 (contd.): If $A = \begin{pmatrix} 1 & 2 & 3 & 5 \\ 2 & 4 & 8 & 12 \\ 3 & 6 & 7 & 13 \end{pmatrix}$, then $x = \begin{pmatrix} -2 & 1 & 0 & 0 \end{pmatrix}^T$ and $y = \begin{pmatrix} -2 & 0 & -1 & 1 \end{pmatrix}^T$ are in $N(A)$.

Q: What about $x + y = \begin{pmatrix} -4 & 1 & -1 & 1 \end{pmatrix}^T$, $-3 \cdot x = \begin{pmatrix} 6 & -3 & 0 & 0 \end{pmatrix}^T$?

REMARK: Let A be an $m \times n$ matrix, u, v be real numbers.

- The null space of A, $\boxed{N(A) \text{ contains vectors from } \mathbb{R}^n}$,
- If x, y are in $N(A)$, i.e., $Ax = 0$ and $Ay = 0$, then $A(ux + vy) = u(Ax) + v(Ay) = 0$, i.e., $ux + vy$ is in $N(A)$.

i.e., a linear combination of vectors in $N(A)$ is also in $N(A)$.

$\boxed{\text{Thus } N(A) \text{ is \textit{closed under} linear combinations}}$.

Finding N(A)

Key Point: $Ax = 0$ has the same solutions as $Ux = 0$, which has the same solutions as $Rx = 0$, i.e.,

$$\boxed{N(A) = N(U) = N(R)}.$$

Reason: If A is $m \times n$, and Q is an invertible $m \times m$ matrix, then $N(A) = N(QA)$. (Verify this)!

Example 2:

For $A = \begin{pmatrix} 1 & 2 & 3 & 5 \\ 2 & 4 & 8 & 12 \\ 3 & 6 & 7 & 13 \end{pmatrix}$, we have $Rx = \begin{pmatrix} \mathbf{1} & 2 & 0 & 2 \\ 0 & 0 & \mathbf{1} & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} t \\ u \\ v \\ w \end{pmatrix}$.

$Rx = 0$ gives $t + 2u + 2w = 0$ and $v + w = 0$.

i.e., $t = -2u - 2w$ and $v = -w$.

Null Space: Solution of $Ax = 0$

$Rx = 0$ gives $t = -2u - 2w$ and $v = -w$,

t and v are *dependent* on the values of u and w.

u and w are *free* and *independent*, i.e., we can choose any value for these two variables.

Special solutions:

$u = 1$ and $w = 0$, gives $x = \begin{pmatrix} -2 & 1 & 0 & 0 \end{pmatrix}^T$.

$u = 0$ and $w = 1$, gives $x = \begin{pmatrix} -2 & 0 & -1 & 1 \end{pmatrix}^T$.

The null space contains:

$$x = \begin{pmatrix} t \\ u \\ v \\ w \end{pmatrix} = \begin{pmatrix} -2u - 2w \\ u \\ -w \\ w \end{pmatrix} = u \begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \end{pmatrix} + w \begin{pmatrix} -2 \\ 0 \\ -1 \\ 1 \end{pmatrix},$$

i.e., all possible linear combinations of the special solutions.

Rank of A

$Ax = 0$ always has a solution: the trivial one, i.e., $x = 0$.

Main Q1: When does $Ax = 0$ have a non-zero solution?

A: When there is at least one free variable,
i.e., not every column of R contains a pivot.

To keep track of this, we define:

$\boxed{\text{rank}(A) = \text{number of columns containing pivots in } R}$.

If A is $m \times n$ and rank(A) = r, then

- rank(A) $\leq$ min$\{m, n\}$.
- no. of dependent variables = r.
- no. of free variables = $n - r$.
- $Ax = 0$ has only the 0 solution $\Leftrightarrow r = n$.
- $m < n \Rightarrow Ax = 0$ has non-zero solutions.

True/False: If $m \geq n$, then $Ax = 0$ has only the 0 solution.

$\boxed{\text{rank}(A) = \text{number of dependent variables in the system } Ax = 0.}$

Example: $R = \begin{pmatrix} 1 & 2 & 0 & 2 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$ when $A = \begin{pmatrix} 1 & 2 & 3 & 5 \\ 2 & 4 & 8 & 12 \\ 3 & 6 & 7 & 13 \end{pmatrix}$.

The no.of columns containing pivots in R is 2, $\Rightarrow$ rank(A) = 2. R contains a 2×2 identity matrix, namely the rows and columns corresponding to the pivots.

This is the row reduced form of the corresponding submatrix $\begin{pmatrix} 1 & 3 \\ 2 & 8 \end{pmatrix}$ of A, which is invertible, since it has 2 pivots.

Thus, $\boxed{\text{rank}(A) = r \Rightarrow A \text{ has an } r \times r \text{ invertible submatrix.}}$

State the converse. The converse is also true. Why?

Summary: Finding $N(A)$ **=** $N(U) = N(R)$

Let A be $m \times n$. To solve $Ax = 0$, find R and solve $Rx = 0$.

1. Find free (independent) and pivot (dependent) variables:
   pivot variables: columns in R with pivots ($\leftrightarrow t$ and v).
   free variables: columns in R without pivots ($\leftrightarrow u$ and w).

2. No free variables, i.e., rank(A) = $n \Rightarrow N(A) = 0$.

3. (a) If rank(A) $< n$, obtain a special solution:
   Set one free variable = 1, the other free variables = 0.
   Solve $Rx = 0$ to obtain values of pivot variables.

   (b) Find special solutions for each free variable.
   $N(A)$ = space of linear combinations of special solutions.

- This information is stored in a compact form in:
  $\boxed{\text{Null Space Matrix: Special solutions as columns.}}$

Solving $Ax = b$

Caution: If $b \neq 0$, solving $Ax = b$ may not be the same as solving $Ux = b$ or $Rx = b$.

Example: $Ax = \begin{pmatrix} 1 & 2 & 3 & 5 \\ 2 & 4 & 8 & 12 \\ 3 & 6 & 7 & 13 \end{pmatrix} \begin{pmatrix} t \\ u \\ v \\ w \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = b.$

Convert to $Ux = c$ and then $Rx = d$.

$$\left(\begin{array}{cccc|c} 1 & 2 & 3 & 5 & b_1 \\ 2 & 4 & 8 & 12 & b_2 \\ 3 & 6 & 7 & 13 & b_3 \end{array}\right) \to \left(\begin{array}{cccc|c} \mathbf{1} & 2 & 3 & 5 & b_1 \\ 0 & 0 & 2 & 2 & b_2 - 2b_1 \\ 0 & 0 & -2 & -2 & b_3 - 3b_1 \end{array}\right)$$

$$\to \left(\begin{array}{cccc|c} \mathbf{1} & 2 & 3 & 5 & b_1 \\ 0 & 0 & \mathbf{2} & 2 & b_2 - 2b_1 \\ 0 & 0 & 0 & 0 & b_3 + b_2 - 5b_1 \end{array}\right)$$

System is consistent $\Leftrightarrow b_3 + b_2 - 5b_1 = 0$, i.e., $b_3 = 5b_1 - b_2$

Solving $Ax = b$ **or** $Ux = c$ **or** $Rx = d$

$Ax = b$ has a solution $\Leftrightarrow b_3 = 5b_1 - b_2$.

for example, there is no solution when $b = \begin{pmatrix} 1 & 0 & 4 \end{pmatrix}^T$.

Suppose $b = \begin{pmatrix} 1 & 0 & 5 \end{pmatrix}^T$. Then $[A|b] \to$

$$\left(\begin{array}{cccc|c} \mathbf{1} & 2 & 3 & 5 & b_1 \\ 0 & 0 & \mathbf{2} & 2 & b_2 - 2b_1 \\ 0 & 0 & 0 & 0 & b_3 + b_2 - 5b_1 \end{array}\right) = \left(\begin{array}{cccc|c} \mathbf{1} & 2 & 3 & 5 & 1 \\ 0 & 0 & \mathbf{2} & 2 & -2 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

$$\to \left(\begin{array}{cccc|c} \mathbf{1} & 2 & 3 & 5 & 1 \\ 0 & 0 & \mathbf{1} & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right) \to \left(\begin{array}{cccc|c} \mathbf{1} & 2 & 0 & 2 & 4 \\ 0 & 0 & \mathbf{1} & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

$Ax = b$ is reduced to solving $Ux = c = \begin{pmatrix} 1 & -2 & 0 \end{pmatrix}^T$,

which is further reduced to solving $Rx = d = \begin{pmatrix} 4 & -1 & 0 \end{pmatrix}^T$.

that is, we want to solve

$$\begin{pmatrix} \mathbf{1} & 2 & 0 & 2 \\ 0 & 0 & \mathbf{1} & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} t \\ u \\ v \\ w \end{pmatrix} = \begin{pmatrix} 4 \\ -1 \\ 0 \end{pmatrix}$$

that is., $t = 4 - 2u - 2w$ and $v = -1 - w$

Set the free variables u and $w = 0$ to get $t = 4$ and $v = -1$
A particular solution: $\mathbf{x} = \begin{pmatrix} 4 & 0 & -1 & 0 \end{pmatrix}^T$.

Exercise: Check it is a solution i.e., check $Ax = b$.

Observe: In $Rx = d$, the vector d gives values for the pivot variables, when the free variables are 0.

General Solution of $Ax = b$

From $Rx = d$, we get $t = 4 - 2u - 2w$ and $v = -1 - w$, where u and w are free. Complete set of solutions to $Ax = b$:

$$\begin{pmatrix} t \\ u \\ v \\ w \end{pmatrix} = \begin{pmatrix} 4 - 2u - 2w \\ u \\ -1 - w \\ w \end{pmatrix} = \begin{pmatrix} 4 \\ 0 \\ -1 \\ 0 \end{pmatrix} + u \begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \end{pmatrix} + w \begin{pmatrix} -2 \\ 0 \\ -1 \\ 1 \end{pmatrix}.$$

To solve $Ax = b$ completely, reduce to $Rx = d$. Then:

1. Find $x_{\text{NullSpace}}$, i.e., $N(A)$, by solving $Rx = 0$.
2. Set free variables = 0, solve $Rx = d$ for pivot variables.

   This is a particular solution:$x_{\text{particular}}$.
3. Complete solutions: $x_{\text{complete}} = x_{\text{particular}} + x_{\text{NullSpace}}$

Exercise: Verify geometrically for a 1×2 matrix, say $A = \begin{pmatrix} 1 & 2 \end{pmatrix}$.

Exercise: Prove statement 3 for solutions of any $Ax = b$.

The Column Space of A

Q: Does $Ax = b$ have a solution? **A:** Not always.

Main Q2: When does $Ax = b$ have a solution?

If $Ax = b$ has a solution, then we can find numbers $x_1, \ldots, x_n$

such that $\begin{pmatrix} A_{*1} & \cdots & A_{*n} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = x_1 A_{*1} + \cdots + x_n A_{*n} = b,$

that is, b can be written as a linear combination of columns of A.

The *column space* of A, denoted $C(A)$;

is the set of all linear combinations of the columns of A

= $\{b$ in $\mathbb{R}^m$ such that $Ax = b$ is **consistent**$\}$.

Finding $C(A)$: Consistency of $Ax = b$

Example: Let $A = \begin{pmatrix} 1 & 2 & 3 & 5 \\ 2 & 4 & 8 & 12 \\ 3 & 6 & 7 & 13 \end{pmatrix}$. Then $Ax = b$, where $b = \begin{pmatrix} b_1 & b_2 & b_3 \end{pmatrix}^T$, has a solution whenever $-5b_1 + b_2 + b_3 = 0$.

- $C(A)$ is a plane in $\mathbb{R}^3$ passing through the origin with normal vector $\begin{pmatrix} -5 & 1 & 1 \end{pmatrix}^T$.
- $c = \begin{pmatrix} 1 & 0 & 4 \end{pmatrix}^T$ is not in $C(A)$ as $Ax = c$ is inconsistent.
- $d = \begin{pmatrix} 1 & 0 & 5 \end{pmatrix}^T$ is in $C(A)$ as $Ax = d$ is consistent.

Exercise: Write b as a linear combination of the columns of A.
(A different way of saying: Solve $Ax = b$).
$x = \begin{pmatrix} 4 & 0 & -1 & 0 \end{pmatrix}^T$ is a solution of $Ax = b$, and

$$\boxed{\begin{pmatrix} 1 & 0 & 5 \end{pmatrix}^T = 4A_{*1} + (-1)A_{*3}.}$$

Q: Can you write b as a different combination of $A_{*1}, \ldots, A_{*4}$?

Linear Combinations in $C(A)$

Let A be an $m \times n$ matrix, u and v be real numbers.

- The column space of A, $\boxed{C(A) \text{ contains vectors from } \mathbb{R}^m}$.
- If a, b are in $C(A)$, i.e., $Ax = a$ and $Ay = b$ for some x, y in $\mathbb{R}^n$, then $ua + vb = u(Ax) + v(Ay) = A(ux + vy) = Aw$, where $w = ux + vy$. Hence, if $w = \begin{pmatrix} w_1 & \cdots & w_n \end{pmatrix}^T$, then
$$ua + vb = w_1 A_{*1} + \cdots w_n A_{*n},$$
i.e., a linear combination of vectors in $C(A)$ is also in $C(A)$.
$$\boxed{\text{Thus, } C(A) \text{ is } \textit{closed under} \text{ linear combinations.}}$$
- If b is in $C(A)$, then b can be written as a linear combination of the columns of A in *as many ways* as the solutions of $Ax = b$.

Summary: $N(A)$ and $C(A)$

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

- The null space of A, $N(A)$ contains vectors from $\mathbb{R}^n$.
- $Ax = 0 \Leftrightarrow x$ is in $N(A)$.

- The column space of A, $C(A)$ contains vectors from $\mathbb{R}^m$.

- If B is the nullspace matrix of A, then $C(B) = N(A)$.

- $Ax = b$ is consistent $\Leftrightarrow$ b is in $C(A)$ $\Leftrightarrow$

  b can be written as a linear combination of the columns of A. This can be done in as many ways as the solutions of $Ax = b$.

- Let A be $n \times n$.

  A is *invertible* $\Leftrightarrow N(A) = \{0\} \Leftrightarrow C(A) = \mathbb{R}^n$. Why?

- $N(A)$ and $C(A)$ are closed under linear combinations.

Chapter 2. Vector Spaces

2.1 Vector Spaces and Subspaces

Vector Spaces: $\mathbb{R}^n$

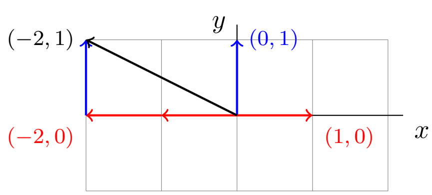
We begin with $\mathbb{R}^1, \mathbb{R}^2, \ldots, \mathbb{R}^n$, etc., where $\mathbb{R}^n$ consists of all column vectors of length n, i.e., $\mathbb{R}^n = \{x = \begin{pmatrix} x_1 & \cdots & x_n \end{pmatrix}^T$, where $x_1, \ldots, x_n$ are in $\mathbb{R}\}$.

We can add two vectors, and we can multiply vectors by scalars, (i.e., real numbers). Thus, we can take linear combinations in $\mathbb{R}^n$.

Examples:

$\mathbb{R}^1$ is the real line, $\mathbb{R}^3$ is the usual 3-dimensional space, and

$\mathbb{R}^2$ is represented by the x-y plane; the x and y co-ordinates are given by the two components of the vector.



Vector Spaces: Definition

Defn. A non-empty set V is a vector space if it is *closed under* vector addition (i.e., if x, y are in V, then $x + y$ must be in V) and scalar multiplication, (i.e., if x is in V, a is in $\mathbb{R}$, then $a * x$ must be in V) satisfying a few axioms.

Equivalently, $\boxed{x, y \text{ in } V,\ a, b \text{ in } \mathbb{R} \Rightarrow a * x + b * y \text{ must be in } V.}$

• A vector space is a triple $(V, +, *)$ with vector addition $+$ and scalar multiplication $*$(see next reading slide).

• The elements of V are called vectors and the scalars are chosen to be real numbers (for now).

- If the scalars are allowed to be complex numbers, then V is a *complex* vector space.
- **Primary Example:** $\mathbb{R}^n$. Under which operations.

Reading: Vector Spaces definition continued

Let x, y and z be vectors, a and b be scalars The vector addition and scalar multiplication are required to satisfy the following axioms:

- $x + y = y + x$ $\boxed{\text{Commutativity of addition}}$
- $(x + y) + z = x + (y + z)$ $\boxed{\text{Associativity of addition}}$
- There is a unique vector 0, such that $x + 0 = x$ $\boxed{\text{Existence of additive identity}}$
- For each x, there is a unique $-x$ such that $x+(-x) = 0$ $\boxed{\text{Existence of additive inverse}}$
- $1 * x = x$ $\boxed{\text{Unit property}}$
- $(a + b) * x = a * x + b * x, \quad a * (x + y) = a * x + a * y \ (ab) * x = a * (b * x)$ $\boxed{\text{Compatibility}}$

Notation: For a scalar a, and a vector x, we denote $a * x$ by ax.

Vector Spaces: Examples

1. $V = \{0\}$, the space consisting of only the zero vector.
2. $V = \mathbb{R}^n$, the n-dimensional space.
3. $V = \mathbb{R}^\infty =$ sequences of real numbers, e.g., $x = (0, 1, 0, 2, 0, 3, 0, 4, \ldots)$, with component-wise addition and scalar multiplication.
4. $V = \mathcal{M}_{m\times n}$, the set of $m \times n$ matrices, with entry-wise $+$ and $*$.
5. $V = \mathcal{P}$, the set of polynomials, e.g. $1 + 2x + 3x^2 + \cdots + 2023x^{2022}$, with term-wise $+$ and $*$.
6. $V = \mathcal{C}[0, 1]$, the set of continuous real-valued functions on the closed interval $[0, 1]$.e.g., x^2, e^x are vectors in V. How about $1/x$ and $1/(x - 5)$? Are they vectors in V?

   Vector addition and scalar multiplication are pointwise:

   $(f + g)(x) = f(x) + g(x)$ and $(a * f)(x) = af(x)$.

Subspaces: Definition and Examples

If V is a vector space, and W is a non-empty subset, then W is a *subspace* of V if:

$$\boxed{x,\ y \text{ in } W, \quad a,\ b \text{ in } \mathbb{R} \Rightarrow a * x + b * y \text{ are in } W.}$$

$$\boxed{\text{i.e., linear combinations stay in the subspace.}}$$

Examples:

1. $\{0\}$: The zero subspace and $\mathbb{R}^n$ itself.
2. $\{(x_1, x_2) : x_1 \geq 0, x_2 \geq 0\}$ is not a subspace of $\mathbb{R}^2$. Why?
3. The line $x - y = 1$ is not a subspace of $\mathbb{R}^2$. Why?

   Exercise: A line not passing through the origin is not a subspace of $\mathbb{R}^2$.
4. The line $x - y = 0$ is a subspace of $\mathbb{R}^2$. Why?

   Exercise: Any line passing through the origin is a subspace of $\mathbb{R}^2$.
5. Let A be an $m \times n$ matrix.

   The null space of A, $N(A)$, is a subspace of $\mathbb{R}^n$.

   The column space of A, $C(A)$, is a subspace of $\mathbb{R}^m$.

   Recall: They are both closed under linear combinations.
6. The set of 2×2 symmetric matrices is a subspace of $\mathcal{M}$. The set of 2×2 lower triangular matrices is also a subspace of $\mathcal{M}$.

   Q. Is the set of invertible 2×2 matrices a subspace of $\mathcal{M}$?
7. The set of convergent sequences is a subspace of $\mathbb{R}^\infty$. What about the set of sequences convergent to 1?
8. The set of differentiable functions is a subspace of $\mathcal{C}[0, 1]$. Is the same true for the set of functions integrable on $[0, 1]$? Create your own examples.
9. See the tutorial sheet for many more examples!

Exercise:(i) A subspace must contain the 0 vector!

(ii) Show that a subspace of a vector space is a vector space.

Examples: Subspaces of $\mathbb{R}^2$

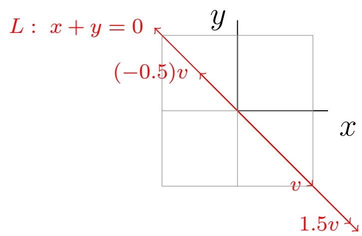
What are the subspaces of $\mathbb{R}^2$?

- $V = \{\begin{pmatrix}0 & 0\end{pmatrix}^T\}$.
- $V = \mathbb{R}^2$.

- What if V is neither of the above?

  Example:

  Suppose V contains a non-zero vector, say $v = \begin{pmatrix} -1 & 1 \end{pmatrix}^T$.

  V must contain the entire line $L: \ x + y = 0$, i.e., all multiples of v.

Let V be a subspace of $\mathbb{R}^2$ containing $v_1 = \begin{pmatrix} -1 & 1 \end{pmatrix}^T$. Then V must contain the entire line $L: \ x + y = 0$.

If $V \neq L$, it contains a vector v_2, which is not a multiple of v_1, say $v_2 = \begin{pmatrix} 0 & 1 \end{pmatrix}^T$.

Observe: $A = \begin{pmatrix} v_1 & v_2 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 1 & 1 \end{pmatrix}$ has two pivots,
$\Leftrightarrow$ A is invertible.
$\Leftrightarrow$ for any v in $\mathbb{R}^2$, $Ax = v$ is solvable,
$\Leftrightarrow$ v is in $C(A)$,
$\Leftrightarrow$ v can be written as a linear combination of v_1 and v_2.
$\Rightarrow$ v is in V, i.e., $V = \mathbb{R}^2$

To summarise: A subspace of $\mathbb{R}^2$, which is non-zero, and not $\mathbb{R}^2$, is a line passing through the origin. What happens in $\mathbb{R}^3$?

2.2 Linear Span and Linear Independence

Linear Span: Definition

Given a collection $S = \{v_1, v_2, \ldots, v_n\}$ in a vector space V, the *linear span* of S, denoted Span(S) or Span$\{v_1, \ldots, v_n\}$,
is the set of all linear combinations of v_1, $v_2, \ldots, v_n$, i.e.,

$$\boxed{\text{Span}(S) = \{v = a_1v_1 + \cdots + a_nv_n, \text{ for scalars } a_1, \ldots, a_n\}.}$$

Let $\{v_1, \ldots, v_n\}$ be n vectors in $\mathbb{R}^n$, $A = \begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix}$.
Note:

1. If $v_1, \ldots, v_n$ are in $\mathbb{R}^m$, Span$\{v_1, \ldots, v_n\} = C(A)$ for $A = \begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix}$, an $m \times n$ matrix. Thus v is in Span$\{v_1, \ldots, v_n\} \Leftrightarrow Ax = v$ is consistent.

2. Span$\{v_1, \ldots, v_n\} = \mathbb{R}^m \Leftrightarrow Ax = v$ is consistent for all $v \in \mathbb{R}^m \Leftrightarrow A$ has m pivots. This implies, $m \leq n$.

3. Let $m = n$. Then A is invertible $\Leftrightarrow$ A has n pivots $\Leftrightarrow$ $Ax = v$ is consistent for every v in $\mathbb{R}^n$ $\Leftrightarrow$ Span$\{v_1, \ldots, v_n\}$ = $\mathbb{R}^n$.

Example: Span$\{e_1, \ldots, e_n\} = \mathbb{R}^n$.

Linear Span: Examples

Examples:

1. Span$\{0\} = \{0\}$.

2. If $v \neq 0$ is a vector, Span$\{v\} = \{av, \text{ for scalars } a\}$.

   Geometrically (in $\mathbb{R}^m$): Span$\{v\}$ = the line in the direction of v passing through the origin.

3. Span$\left\{\begin{pmatrix} -1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}\right\} = \mathbb{R}^2$.

4. If A is $m \times n$, then Span$\{A_1, \ldots, A_n\} = C(A)$.

5. If $v_1, \ldots, v_k$ are the special solutions of A,

   then Span $\{v_1, \ldots, v_k\}$ = $N(A)$.

Remark: All of the above are subspaces.

Exercise: Span(S) is a subspace of V. Why?

6. Let $v_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$, $v_2 = \begin{pmatrix} 2 \\ 4 \\ 6 \end{pmatrix}$, $v_3 = \begin{pmatrix} 3 \\ 8 \\ 7 \end{pmatrix}$ and $v_4 = \begin{pmatrix} 5 \\ 12 \\ 13 \end{pmatrix}$. Is $v = \begin{pmatrix} 1 & 0 & 4 \end{pmatrix}^T$ in Span$\{v_1, v_2, v_3, v_4\}$?

Let $A = \begin{pmatrix} v_1 & \cdots & v_4 \end{pmatrix}$, and $b = \begin{pmatrix} b_1 & b_2 & b_3 \end{pmatrix}$.
Recall $Ax = b$ is solvable $\Leftrightarrow$ $5b_1 - b_2 - b_3 = 0$.
$\Rightarrow$ v is not in Span$\{v_1, v_2, v_3, v_4\}$,
and $w = \begin{pmatrix} 1 & 0 & 5 \end{pmatrix}^T = 4v_1 + (-1)v_3$ is in it.

Observe: $v_2 = 2v_1$ and $v_4 = 2v_1 + v_3$. Hence v_2, v_4 are in Span$\{v_1, v_3\}$ $\Rightarrow$ Span$\{v_1, v_2, v_3, v_4\}$ = Span$\{v_1, v_3,\}$.

Thus, $C(A) =$ the plane $P : (5x - y - z = 0)$ = Span$\{v_1, v_3\}$.

Question:
Is the **span** of two vectors in $\mathbb{R}^3$ always a plane?

7. Let $v_1 = \begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix}$, $v_2 = \begin{pmatrix} 4 \\ 5 \\ 3 \end{pmatrix}$, $v_3 = \begin{pmatrix} 6 \\ 7 \\ 5 \end{pmatrix}$ and $v_4 = \begin{pmatrix} 4 \\ 6 \\ 2 \end{pmatrix}$? .

   Is $v = \begin{pmatrix} 4 & 3 & 5 \end{pmatrix}^T$ in Span$\{v_1, v_2, v_3, v_4\}$? If yes, write v as a linear combination of v_1, v_2, v_3 and v_4.

Let $A = \begin{pmatrix} v_1 & \cdots & v_4 \end{pmatrix}$. The question can be rephrased as:
Question: Is v in $C(A)$, i.e., is $Ax = v$ solvable? If yes, find a solution.

Exercise: $Ax = \begin{pmatrix} a & b & c \end{pmatrix}^T$ is consistent $\Leftrightarrow$ $2a - b - c = 0$.

Observe and prove:
(i) that Span$\{v_1, v_2, v_3, v_4\}$ = $C(A)$ is a plane! **(ii)** that v is in Span$\{v_1, v_2, v_3, v_4\}$ (and $w = \begin{pmatrix} 4 & 3 & 4 \end{pmatrix}^T$ is not).

Solve $Ax = v$ using the row reduced form of A to get **particular** solution: $\begin{pmatrix} 4 & -1 & 0 & 0 \end{pmatrix}^T$ and $v = 4v_1 + (-1)v_2$.

Linear Independence: Example

With $v_1 = \begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix}$, $v_2 = \begin{pmatrix} 4 \\ 5 \\ 3 \end{pmatrix}$, $v_3 = \begin{pmatrix} 6 \\ 7 \\ 5 \end{pmatrix}$ and $v_4 = \begin{pmatrix} 4 \\ 6 \\ 2 \end{pmatrix}$

Observe: $v_3 = v_1 + v_2$ and $v_4 = -2v_1 + 2v_2$.
Hence v_3 and v_4 are in Span$\{v_1, v_2\}$.

Therefore, Span$\{v_1, v_2\}$ = Span$\{v_1, v_2, v_3, v_4\}$
$= C(A) =$ the plane $P : (2x - y - z = 0)$.

Question: Is the span of two vectors in $\mathbb{R}^3$ always a plane?
A: Not always. If v is a multiple of w, then Span$\{v, w\}$ = Span$\{w\}$, which is a line through the origin or zero.

Question: If v and w are not on the same line through the origin? **A:** Yes. v, w are examples of *linearly independent vectors.*

Linear Independence: Definition

The vectors $v_1, v_2, \ldots, v_n$ in a vector space V, are *linearly independent* if $\boxed{a_1v_1 + \cdots + a_nv_n = 0 \Rightarrow$

Equivalently, for every nonezero $(a_1, \ldots, a_n)^T$ in $\mathbb{R}^n$,
we have $a_1v_1 + \cdots + a_nv_n \neq 0$ in V.

The vectors $v_1, \ldots, v_n$ are *linearly dependent* if they are not linearly independent. i.e., we can find $(a_1, \ldots, a_n)^T \neq 0$ in $\mathbb{R}^n$, such that $a_1v_1 + \cdots + a_nv_n = 0$ in V.

Observe: When $V = \mathbb{R}^m$, if $A = \begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix}$, then

$\boxed{Ax = x_1v_1 + \cdots + x_nv_n = 0 \text{ has a } \textbf{non-trivial} \text{ solution,}}$

$\Leftrightarrow N(A) \neq 0 \Leftrightarrow v_1, \ldots, v_n$ are linearly **dependent** and

$$\boxed{Ax = x_1 v_1 + \cdots + x_n v_n = 0 \text{ has only the } \textbf{trivial} \text{ solution}}$$

$\Leftrightarrow N(A) = 0 \Leftrightarrow v_1, \ldots, v_n$ are linearly **independent**.

Linear Independence: Remarks

Remarks/Examples:

1. The zero vector 0 is not linearly independent. Why?
2. If $v \neq 0$, then it is linearly independent. Why?
3. v, w are not linearly independent $\Leftrightarrow$ one is a multiple of the other $\Leftrightarrow$ (for $V = \mathbb{R}^m$) they lie on the same line through the origin.
4. More generally, $v_1, \ldots, v_n$ are not linearly independent $\Leftrightarrow$ one of the v_i's can be written as a linear combination of the others, i.e., v_i is in Span$\{v_j : j = 1, \ldots n, j \neq i\}$.
5. Let A be $m \times n$. Then rank$(A) = n \Leftrightarrow N(A) = 0 \Leftrightarrow A_{*1}, \cdots, A_{*n}$ are linearly independent.

   In particular, if A is $n \times n$, A is invertible $\Leftrightarrow A_{*1}, \cdots, A_{*n}$ are linearly independent.

Example: $e_1, \ldots, e_n$ are linearly independent vectors in $\mathbb{R}^n$.

Linear Independence: Example

Example: Are the vectors $v_1 = \begin{pmatrix} 2 & 2 & 2 \end{pmatrix}^T$, $v_2 = \begin{pmatrix} 4 & 5 & 3 \end{pmatrix}^T$, $v_3 = \begin{pmatrix} 6 & 7 & 5 \end{pmatrix}^T$ and $v_4 = \begin{pmatrix} 4 & 6 & 2 \end{pmatrix}^T$ linearly independent?

For $A = \begin{pmatrix} v_1 & \cdots & v_4 \end{pmatrix}$, reduced form $R = \begin{pmatrix} 1 & 0 & 1 & -2 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix}$.

A has only 2 pivots $\Rightarrow N(A) \neq 0$, so v_1, v_2, v_3, v_4 are not independent. A non-trivial linear combination which is zero is $(1)v_1 + (1)v_2 + (-1)v_3 + (0)v_4$, or $(2)v_1 + (-2)v_2 + (0)v_3 + (1)v_4$.

- More generally, if $v_1, \ldots, v_n$ are vectors in $\mathbb{R}^m$, then

  $A = \begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix}$ is $m \times n$.

If $m < n$, then rank$(A) < n \Rightarrow N(A) \neq 0$. Thus

$$\boxed{\textbf{In } \mathbb{R}^m\textbf{, any set with more than } m \textbf{ vectors is linearly dependent.}}$$

Summary: Vector Spaces, Span and Independence

• **Vector space**: A triple $(V, +, *)$ which is closed under $+$ and $*$ with some additional properties satisfied by $+$ and $*$.

• **Subspace:** A non-empty subset W of V closed under linear combinations.

Let $V = \mathbb{R}^m$, $v_1, \ldots, v_n$ be in V, and $A = \begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix}$.

• For v in V, v is in Span$\{v_1, \ldots, v_n\} \Leftrightarrow Ax = v$ is consistent

• $v_1, \ldots, v_n$ are linearly independent
$\Leftrightarrow N(A) = 0 \Leftrightarrow$ rank$(A) = n$.

• In particular, with $n = m$, A is invertible
$\Leftrightarrow Ax = v$ is consistent for every v
$\Leftrightarrow$ Span$\{v_1, \ldots, v_n\}$ = $\mathbb{R}^n \Leftrightarrow$ rank$(A) = n$
$\Leftrightarrow N(A) = 0 \Leftrightarrow v_1, \ldots, v_n$ are linearly independent.

• If Span$\{v_1, \ldots, v_n\}$ = $\mathbb{R}^m$, then $m \le n$, and
any subset of $\mathbb{R}^m$ with more than m vectors is dependent.

2.3 Basis and Dimension

Basis: Introduction

Let $v_1 = \begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix}$, $v_2 = \begin{pmatrix} 4 \\ 5 \\ 3 \end{pmatrix}$, $v_3 = \begin{pmatrix} 6 \\ 7 \\ 5 \end{pmatrix}$, $v_4 = \begin{pmatrix} 4 \\ 6 \\ 2 \end{pmatrix}$, and $A = (v_1\ v_2\ v_3\ v_4)$. Can $C(A)$ =Span$\{v_1, v_2, v_3, v$ be spanned by less than 4 vectors?

Note: $v_3 = v_1 + v_2$ and $v_4 = -2v_1 + 2v_2 \Rightarrow C(A) =$ Span$\{v_1, v_2\}$.

Observe:

• The span of only v_1 or only v_2 is a line. Clearly v_1 is not on the line spanned by v_2 and vice versa.

Thus, $\{v_1, v_2\}$ is a *minimal spanning set* for $C(A)$.

• v_1 and v_2 are linearly independent and span $C(A)$.

• If v is in $C(A)$ = Span$\{v_1, v_2\}$, then v_1, v_2, v are linearly dependent. Why?

Thus, $\{v_1, v_2\}$ is a *maximal linearly independent set* in $C(A)$.

Any such set of vectors gives a *basis* of $C(A)$.

Basis: Definition

Defn. A subset $\mathcal{B}$ of a vector space V, is said to be a *basis* of V, if it is linearly independent and Span($\mathcal{B}$) = V.

Theorem: For any subset S of a vector space V, the following are equivalent:

• S is a maximal linearly independent set in V

• S is linearly independent and Span(S) = V.

• S is a minimal spanning set of V.

Note: Every vector space V has a basis.

Examples:

- By convention, the empty set is a basis for $V = \{0\}$.
- $\left\{ \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}$ is a basis for $\mathbb{R}^2$.
- $\{e_1, \ldots, e_n\}$ is a basis for $\mathbb{R}^n$, called the standard basis.
- A basis of $\mathbb{R}$ is just $\{1\}$.

Basis: Remarks

- Let $\mathcal{B} = \{v_1, \ldots, v_n\}$ be a basis for V and v a vector in V.

Span$(\mathcal{B}) = V \Rightarrow v = a_1 v_1 + \cdots + a_n v_n$ for scalars $a_1, \ldots, a_n$.
Linear independence $\Rightarrow$ this expression for v is unique. Thus

$$\boxed{\text{Every } v \in V \text{ can be } \textit{uniquely} \text{ written}} \; \boxed{\text{as a linear combination of } \{v_1, \ldots, v_n\}.}$$

Exercise: Prove this.

Q: Is the basis of a vector space unique? **A:** No.

e.g. $\{e_1, e_2\}$ is a basis for $\mathbb{R}^2$, so is $\left\{ \begin{pmatrix} -1 & 1 \end{pmatrix}^T, \begin{pmatrix} 0 & 1 \end{pmatrix}^T \right\}$, and so are the columns of any 2×2 invertible matrix.

Exercise: Find two different basis of $\mathbb{R}^3$.
The number of vectors in each basis of $\mathbb{R}^3$ is 3. Why?
Recall: If $v_1, \ldots, v_n$ span $\mathbb{R}^m$, then $m \le n$, and
if they are linear independent, then $n \le m$.

Coordinate Vector: Definition

- Let $\mathcal{B} = \{v_1, \ldots, v_n\}$ be a basis for V and v a vector in V.

Span$(\mathcal{B}) = V \Rightarrow v = a_1 v_1 + \cdots + a_n v_n$ for scalars $a_1, \ldots, a_n$.
Linear independence $\Rightarrow$ this expression for v is unique. Thus

$$\boxed{\text{Every } v \in V \text{ can be } \textit{uniquely} \text{ written}} \; \boxed{\text{as a linear combination of } \{v_1, \ldots, v_n\}.}$$

Exercise: Prove this!

Definition: If $v = a_1 v_1 + \cdots + a_n v_n$, then $(a_1, \ldots, a_n)^T \in \mathbb{R}^n$ is called the *coordinate vector* of v w.r.t. $\mathcal{B}$, denoted $[v]_{\mathcal{B}}$.

Note: $[v]_{\mathcal{B}}$ depends not only on the basis $\mathcal{B}$, but also the order of the elements in $\mathcal{B}$.

Question:

How does $[v]_{\mathcal{B}}$ change, if $\mathcal{B}$ is rewritten as $\{v_2, v_1, v_3, \ldots, v_n\}$?

Dimension of a Vector Space

Question: The number of vectors in each basis of $\mathbb{R}^3$ is 3. Why?
Recall: If $v_1, \dots, v_n$ span $\mathbb{R}^m$, then $m \leq n$, and
if they are linear independent, then $n \leq m$.

Defn.: More generally, if $v_1, \dots v_m$ and $w_1, \dots, w_n$ are both basis of V, then $m = n$. This is called the *dimension* of V. Thus

$$\boxed{\text{dim}(V) = \text{number of elements in a basis of } V.}$$

Examples: • $\dim(\{0\}) = 0$. • $\dim(\mathbb{R}^n) = n$.

• A line through origin in $\mathbb{R}^3$ is of the form $\mathbf{L} = \{tu \mid t \in \mathbb{R}\}$ for some u in $\mathbb{R}^3 \setminus \{0\}$. A basis for $\mathbf{L}$ is $\{\_\_\_\}$, and $\dim(\mathbf{L}) = \_\_\_$.

• The dimension of a plane $(\mathbf{P})$ in $\mathbb{R}^3$ is 2. Why?

• A basis for $\mathbb{C}$ as a vector space over $\mathbb{R}$ is $\{1, i\}$.
A basis for $\mathbb{C}$ as a *complex* vector space is $\{1\}$.
i.e., dim($\mathbb{C}$) = 2 as a $\mathbb{R}$-vector space and 1 as a $\mathbb{C}$-vector space.
Thus, dimension depends on the choice of scalars!

Basis: Remarks

Let dim $(V) = n$, $S = \{v_1, \dots, v_k\} \subseteq V$.

Recall: A basis is a minimal spanning set.

In particular, if Span$(S) = V$, then $k \geq n$, and S contains a basis of V, i.e., there exist $\{v_{i_1}, \dots, v_{i_n}\} \subseteq S$ which is a basis of V.

Example: The columns of a 3×4 matrix A with 3 pivots span $\mathbb{R}^3$. Hence the columns contain a basis of $\mathbb{R}^3$.

RECALL: A basis is a maximal linearly independent set.

In particular, if S is linear independent, then $k \leq n$, and S can be extended to a basis of V, i.e., there exist $w_1, \dots, w_{n-k}$ in V such that $\{v_1, \dots, v_k, w_1, \dots, w_{n-k}\}$ is a basis of V.

Example: The columns of a 3×2 matrix A with 2 pivots has linearly independent columns, and hence can be extended to a basis of $\mathbb{R}^3$.

Summary: Basis and Dimension

• A basis of a vector space V is a linearly independent subset $\mathcal{B}$ which spans V.

• A basis is a maximal linearly independent subset of V
$\Rightarrow$ any linearly independent subset in V
can be extended to a basis of V.

• A basis is a minimal spanning set of V
$\Rightarrow$ every spanning set of V contains a basis.

• The number of elements in each basis is the same,
and the dimension of V,
dim(V) = number of elements in a basis of V.

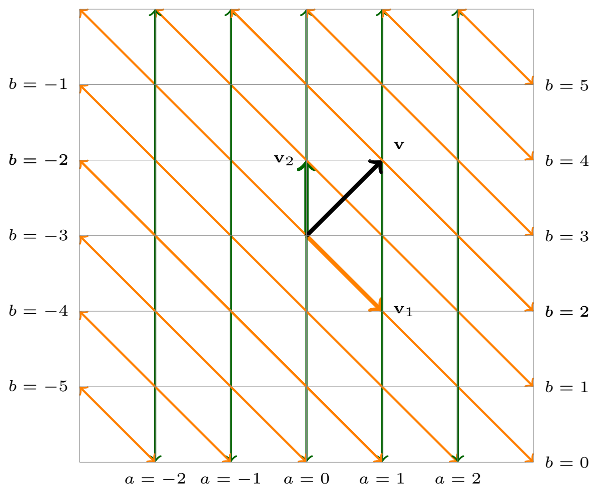
• $\mathcal{B} = \{v_1, \ldots, v_n\}$ is a basis for $V \Leftrightarrow$ every $v \in V$ can be uniquely written as a linear combination of $\{v_1, \ldots, v_n\}$.

• dim $(\mathbb{R}^n) = n$, and the set $\mathcal{B} = \{v_1, \ldots, v_n\}$ is a basis of $\mathbb{R}^n$
$\Leftrightarrow A = \begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix}$ is invertible.

Example: A basis for $\mathbb{R}^2$

Pick $\mathbf{v}_1 \neq 0$. Choose $\mathbf{v}_2$, not a multiple of $\mathbf{v}_1$. For any $\mathbf{v}$ in $\mathbb{R}^2$, there are unique scalars a and b such that $\mathbf{v} = a\mathbf{v}_1 + b\mathbf{v}_2$.

e.g., pick $\mathbf{v}_1 = (1, -1)^T$, $\mathbf{v}_2 = (0, 1)^T$, and let $\mathbf{v} = (1, 1)^T$.



Thus the lines $a = 0$ and $b = 0$ give a set of axes for $\mathbb{R}^2$,

and $\mathbf{v} = \mathbf{v}_1 + 2\mathbf{v}_2$.

With this basis $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2\}$, the coordinates of $\mathbf{v}$ will be $[\mathbf{v}]_{\mathcal{B}} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$.

Basis and Coordinates

A basis for $\mathcal{M}_{2\times 2}$, the vector space of 2×2 matrices , (called *standard the basis of $\mathcal{M}$*), is $\mathcal{B} = \{e_{11}, e_{12}, e_{21}, e_{22}\}$, where

$e_{11} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, e_{12} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, e_{21} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, e_{22} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$

(Verify this!) Hence $\dim(\mathcal{M}_{2\times 2}) = 4$.

Every 2×2 matrix $A = (a_{ij})$ can be written uniquely as

$$A = a_{11}e_{11} + a_{12}e_{12} + a_{21}e_{21} + a_{22}e_{22}.$$

Thus, the coordinate vector of A with respect to $\mathcal{B}$ is

$$[A]_{\mathcal{B}} = (a_{11}, a_{12}, a_{21}, a_{22})^T$$

Note: $[A]_{\mathcal{B}}$ completely determines A, once we fix $\mathcal{B}$, and order the elements in $\mathcal{B}$.

Since dim $(\mathcal{M}_{2\times 2}) = 4$, once we fix a basis, we will need 4 coordinates to describe each matrix.

Exercise: Find two bases (other than the standard one) and the dimension of $\mathcal{M}_{m\times n}$. Find $[e_{11}]_{\mathcal{B}}$ in both cases.

Coordinate Vectors: Examples

1. Consider the basis $\mathcal{B} = \{v_1 = (1, -1)^T, v_2 = (0, 1)^T\}$ of $\mathbb{R}^2$, and $v = (1, 1)^T$. Note that $v = 1v_1 + 2v_2$. Hence, the coordinate vector of v w.r.t. $\mathcal{B}$ is $[v]_{\mathcal{B}} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$.

2. **Exercise:** Show that $\mathcal{B} = \{1, x, x^2\}$ is a basis of $\mathcal{P}_2$

   (called the *standard basis* of $\mathcal{P}_2$).

   The coordinate vector of $v = 2x^2 - 3x + 1$ w.r.t. $\mathcal{B}$ is $[v]_{\mathcal{B}} = (1, -3, 2)^T$.

3. **Exercise:** Show that $\mathcal{B}' = \{1, (x-1), (x-1)^2, (x-1)^3\}$ is a basis of $\mathcal{P}_3$. HINT: Taylor expansion.

   Let $\mathcal{B}$ be the standard basis of $\mathcal{P}_3$. Then $[x^3]_{\mathcal{B}} = (\_\_, \_\_, \_\_, \_\_)^T$, and $[x^3]_{\mathcal{B}'} = (\_\_, \_\_, \_\_, \_\_)^T$.

Recall: To write the coordinates, we have a to fix a basis $\mathcal{B}$, and fix the order of elements in it!

2.4 Null Space, Column Space and Row Space

Subspaces Associated to a Matrix

Associated to an $m \times n$ matrix A, we have four subspaces:

- The **column space** of A: $C(A) = \text{Span}\{A_{*1}, \ldots A_{*n}\} = \{v : Ax = v \text{ is consistent}\} \subseteq \mathbb{R}^m$.

- The **null space** of A: $N(A) = \{x : Ax = 0\} \subseteq \mathbb{R}^n$.
- The **row space** of A = $\text{Span}\{A_{1*}, \ldots, A_{m*}\} = C(A^T) \subseteq \mathbb{R}^n$.
- The **left null space** of A = $\{x : x^T A = 0\} = N(A^T) \subseteq \mathbb{R}^m$.

Question: Why are the row space and the left null space subspaces?

Recall: Let U be the echelon form of A, and R its reduced form.

Then $\boxed{N(A) = N(U) = N(R).}$

Observe: The rows of U(and R) are linear combinations of the rows of A, and vice versa $\Rightarrow$ their row spaces are same, i.e.,

$$\boxed{C(A^T) = C(U^T) = C(R^T).}$$

We compute bases and dimensions of these special subspaces.

An Example

We illustrate how to find a basis and the dimension of the Null Space $N(A)$, the Column Space $C(A)$, and the Row Space $C(A^T)$ by using the following example.

Let $A = \begin{pmatrix} 2 & 4 & 6 & 4 \\ 2 & 5 & 7 & 6 \\ 2 & 3 & 5 & 2 \end{pmatrix}$.

Recall:

• The reduced form of A is $R = \begin{pmatrix} 1 & 0 & 1 & -2 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix}$.

• The 1st and 2nd are pivot columns ⇒ rank(A) = 2.

• $v = \begin{pmatrix} a & b & c \end{pmatrix}^T$ is in $C(A) \Leftrightarrow Ax = v$ is solvable $\Leftrightarrow 2a - b - c = 0$.

• We can compute special solutions to $Ax = 0$. The number of special solutions to $Ax = 0$ is the number of free variables.

The Null Space: $N(A)$

For $A = \begin{pmatrix} 2 & 4 & 6 & 4 \\ 2 & 5 & 7 & 6 \\ 2 & 3 & 5 & 2 \end{pmatrix}$, reduced form $R = \begin{pmatrix} 1 & 0 & 1 & -2 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix}$.

$$N(A) = \left\{ \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = \begin{pmatrix} -c + 2d \\ -c - 2d \\ c \\ d \end{pmatrix} = c \begin{pmatrix} -1 \\ -1 \\ 1 \\ 0 \end{pmatrix} + d \begin{pmatrix} 2 \\ -2 \\ 0 \\ 1 \end{pmatrix} \right\}.$$

= Span$\left\{ w_1 = \begin{pmatrix} -1 & -1 & 1 & 0 \end{pmatrix}^T, w_2 = \begin{pmatrix} 2 & -2 & 0 & 1 \end{pmatrix}^T \right\}$.

w_1, w_2 are linearly independent (Why?)

$\Rightarrow \mathcal{B} = \{w_1, w_2\}$ forms a basis for $N(A) \Rightarrow$ dim$(N(A)) = 2$.

$\boxed{\text{A basis for } N(A) \text{ is the set of special solutions.}}$

$\boxed{\dim(N(A)) = \text{no. of free variables = no. of variables - rank}(A)}$

dim$(N(A))$ is called nullity(A).

Show: $w = (-3, -7, 5, 1)^T$ is in $N(A)$. Find $[w]_{\mathcal{B}}$.

The Column Space: $C(A)$

For $A = \begin{pmatrix} 2 & 4 & 6 & 4 \\ 2 & 5 & 7 & 6 \\ 2 & 3 & 5 & 2 \end{pmatrix}$, reduced form $R = \begin{pmatrix} 1 & 0 & 1 & -2 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix}$.

Write $A = \begin{pmatrix} v_1 & v_2 & v_3 & v_4 \end{pmatrix}$ and $R = \begin{pmatrix} w_1 & w_2 & w_3 & w_4 \end{pmatrix}$.

Recall: Relations between the column vectors of A are the same as the relations between column vectors of R.

$\Rightarrow Ax = v_3$ has a solution has the same solution as $Rx = w_3$, and $Ax = v_4$ has a same solution as $Rx = w_4$.

Particular solutions are $(1,1,0,0)^T$ and $(-2,2,0,0)^T$ respectively $\Rightarrow v_3 = v_1 + v_2$, $v_4 = -2v_1 + 2v_2$.

Observe:

• v_1 and v_2 correspond to the pivot columns of A.
• $\{v_1, v_2\}$ are linearly independent. Why?
• $C(A) = \text{Span}\{v_1, \ldots, v_4\} = \text{Span}\{v_1, v_2\}$.

Thus $\mathcal{B} = \{v_1, v_2\}$ is a basis of $C(A)$. **Q:** What is $[v_i]_{\mathcal{B}}$?

The Rank-Nullity Theorem

More generally, for an $m \times n$ matrix A,

• Let $\text{rank}(A) = r$. The r pivot columns are linearly independent since their reduced form contains an $r \times r$ identity matrix.

• Each non-pivot column A_{*j} of A can be written as a linear combination of the pivots columns, by solving $Ax = A_{*j}$. Thus

$$\boxed{\text{A basis for } C(A) \text{ is given by the pivot columns of } A.}$$

$$\boxed{\dim(C(A)) = \text{no. of pivot variables} = \text{rank}(A).}$$

• A basis for $N(A)$ is given by the special solutions of A. Thus

$$\boxed{\dim(N(A)) = \text{no. of free variables} = \text{nullity}(A).}$$

RANK-NULLITY THEOREM: Let A be an $m \times n$ matrix. Then

$$\dim(C(A)) + \dim(N(A)) = \text{no. of variables} = n$$

The Row Space: $C(A^T)$

Recall: If $A = \begin{pmatrix} 2 & 4 & 6 & 4 \\ 2 & 5 & 7 & 6 \\ 2 & 3 & 5 & 2 \end{pmatrix}$, then $R = \begin{pmatrix} 1 & 0 & 1 & -2 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix}$.

Recall: R is obtained from A by taking non-zero scalar multiples of rows and their sums $\Rightarrow C(R^T) = C(A^T)$.

Observe: The non-zero rows of R will span $C(A^T)$, and they contain an identity submatrix $\Rightarrow$ they are linearly independent.

Thus, the non-zero rows of R form a basis for $C(R^T) = C(A^T)$.

Exercise: Give two different basis for $C(A^T)$.

Since the number of non-zero rows of R = number of pivots of A, we have:

$$\boxed{\dim C(A^T) = \text{no. of pivots of } A = \text{rank}(A).}$$

Recall: $\dim C(A^T) = \text{rank}(A^T)$. Thus,

$$\boxed{\text{rank}(A^T) = \dim\,(C(A^T)) = \text{rank}(A)}$$

Extra Reading: The Left Null Space - $N(A^T)$

The no. of columns of A^T is m.

By Rank-Nullity Theorem, rank$(A^T)+$ dim$(N(A^T)) = m$.

Hence:

$$\boxed{\text{dim}(N(A^T)) = m - \text{rank}(A).}$$

EXERCISE: Complete the example by finding a basis for $N(A^T)$. $A = \begin{pmatrix} 2 & 4 & 6 & 4 \\ 2 & 5 & 7 & 6 \\ 2 & 3 & 5 & 2 \end{pmatrix}$, reduced form $R = \begin{pmatrix} 1 & 0 & 1 & -2 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix}$.

Question. Can you use R to compute the basis for $N(A^T)$? Why not?

A. Need the reduced form of A^T which is $\begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$.

2.5 Linear Transformations

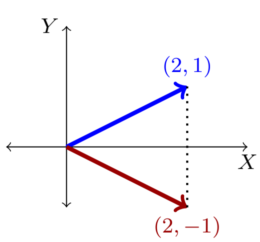
Matrices as Transformations: Examples

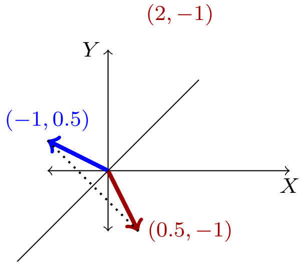
Let $A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. Then $A \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ -x_2 \end{pmatrix}$. Let $\mathbf{x} = (2,1)^T$. What is $\mathbf{A}x$? How does A transform x?

A reflects vectors across the X-axis.

Let $B = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. Then $B \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_2 \\ x_1 \end{pmatrix}$. If $\mathbf{x} = (-1, 0.5)^T$, then $\mathbf{B}x = (0.5, -1)^T$. How does B transform x?

B reflects vectors across the line $x_1 = x_2$.

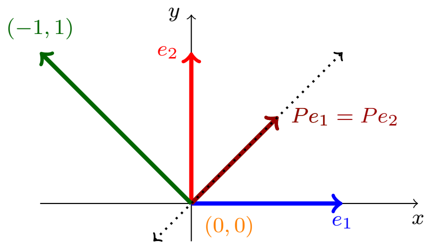




Q: Do reflections preserve scalar multiples? Sums of vectors?

Matrices as Transformations: Examples

- $P = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix}$ transforms $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ to $Px = \begin{pmatrix} \frac{x_1+x_2}{2} \\ \frac{x_1+x_2}{2} \end{pmatrix}$.



$Pe_1 = \begin{pmatrix} 1/2 \\ 1/2 \end{pmatrix} \begin{pmatrix} 1/2 \\ 1/2 \end{pmatrix} = Pe_2$.

P transforms the vector $\begin{pmatrix} -1 \\ 1 \end{pmatrix}$ to the origin.

Question: Geometrically, how is P transforming the vectors?
Answer: Projects onto the line $x_1 = x_2$.

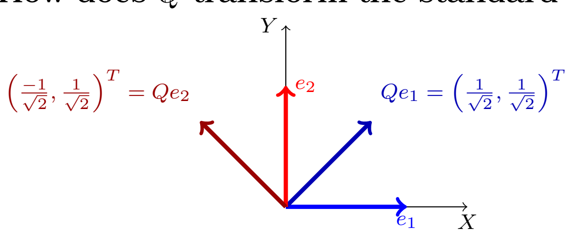
Question: What happens to sums of vectors when you project them? What about scalar multiples?

Question: Understand the effect of $P = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ on e_1 and e_2 and interpret what P represents geometrically!

Matrices as transformations: Examples

Let $Q = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} \cos(45^\circ) & -\sin(45^\circ) \\ \sin(45^\circ) & \cos(45^\circ) \end{pmatrix}$.

How does Q transform the standard basis vectors $\mathbf{e}_1$ and $\mathbf{e}_2$?



Q: What does the transformation $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto Qx$ represent geometrically?

Rotations also map sum of vectors to sum of their images and a scalar multiple of a vector to the scalar multiple of its image.

Matrices as Transformations

- An $m \times n$ matrix A transforms a vector x in $\mathbb{R}^n$ into the vector Ax in $\mathbb{R}^m$. Thus $T(x) = Ax$ defines a function $T : \mathbb{R}^n \to \mathbb{R}^m$.
- The domain of T is ___. The codomain of T is ___.

• Let $b \in \mathbb{R}^m$. Then b is in $C(A) \Leftrightarrow Ax = b$ is consistent $\Leftrightarrow T(x) = b$, i.e., b is in the image (or range) of T.

Hence, the range of T is ___.

Example: Let $A = \begin{pmatrix} 2 & 4 & 6 & 4 \\ 2 & 5 & 7 & 6 \\ 2 & 3 & 5 & 2 \end{pmatrix}$. Then $T(x) = Ax$ is a function with domain $\mathbb{R}^4$, codomain $\mathbb{R}^3$, and range equal to $C(A) = \{(a, b, c)^T \mid 2a - b - c = 0\} \subseteq \mathbb{R}^3$.

Question: How does T transform sums and scalar multiples of vectors?

Ans. Nicely! For scalars a and b, and vectors x and y,
$T(ax + by) = A(ax + by) = aAx + bAy = aT(x) + bT(y)$. Thus
$\boxed{T \text{ takes linear combinations to linear combinations.}}$

Linear Transformations

Defn. Let V and W be vector spaces.

• A *linear transformation* from V to W is a function $T : V \to W$ such that for $x, y \in V$, scalars a and b,

$$\boxed{T(ax + by) = aT(x) + bT(y)}$$

i.e., T takes linear combinations of vectors in V to the linear combinations of their images in W.

• If T is also a bijection, we say T is a *linear isomorphism*.

• The *image* (or *range*) of T is defined to be
$C(T) = \{y \in W \mid T(x) = y \text{ for some } x \in V\}$.

• The *kernel* (or *null space*) of T is defined as
$N(T) = \{x \in V \mid T(x) = 0\}$.

Main Example: Let A be an $m \times n$ matrix. Define $T(x) = Ax$.

• This defines a linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$.

• The image of T is the column space of A, i.e., $C(T) = C(A)$.

• The kernel of T is the null space of A, i.e., $N(T) = N(A)$.

Linear Transformations: Examples

Which of the following functions are linear transformations?

• $g : \mathbb{R}^3 \to \mathbb{R}^3$ defined as $g(x_1, x_2, x_3)^T = (x_1, x_2, 0)^T$

$ag(x) + bg(y) = ag \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} + bg \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} ax_1 \\ ax_2 \\ 0 \end{pmatrix} + \begin{pmatrix} by_1 \\ by_2 \\ 0 \end{pmatrix} = \begin{pmatrix} ax_1 + by_1 \\ ax_2 + by_2 \\ 0 \end{pmatrix} = g(ax + by)$ is a linear transformation.

Exercise: Find $N(g)$ and $C(g)$.

• $h : \mathbb{R}^3 \to \mathbb{R}^3$ defined as $h(x_1, x_2, x_3)^T = (x_1, x_2, 5)^T$.

Note: $h(0 + 0) \neq h(0) + h(0)$.

Observe: A linear transformation must map $0 \in V$ to $0 \in W$.

• $f : \mathbb{R}^2 \to \mathbb{R}^4$ defined by $f(x_1, x_2)^T = (x_1, 0, x_2, x_2^2)^T$.

Note: f transforms the Y-axis in $\mathbb{R}^2$ to $\{(0,0,y,y^2)^T \mid y \in \mathbb{R}\}$.

Observe: A linear transformation must transform a subspace of V into a subspace of W.

- $S : \mathcal{M}_{2\times 2} \to \mathbb{R}^4$ defined by $S\left(\begin{pmatrix} a & b \\ c & d \end{pmatrix}\right) = (a,b,c,d)^T$is a linear transformation.

Observe: S is also a bijection, and hence an isomorphism!
S is onto $\Rightarrow C(S) = \mathbb{R}^4$, and $S(A) = S(B) \Rightarrow A = B$,
i.e., S is one-one. In particular, $N(S) = \{0\}$.

Show that the following functions are linear transformations.

$T : \mathbb{R}^\infty \to \mathbb{R}^\infty$ defined by $T(x_1, x_2, \ldots) = (x_1 + x_2, x_2 + x_3, \ldots,)$.
Exercise: What is $N(T)$?

$S : \mathbb{R}^\infty \to \mathbb{R}^\infty$ defined by $S(x_1, x_2, \ldots) = (x_2, x_3, \ldots)$.
Exercise: Find $C(S)$, and a basis of $N(S)$.

Let $T : \mathcal{P}_2 \to \mathcal{P}_1$ be $S(a_0 + a_1x + a_2x^2) = a_1 + 4a_2x$.

Exercise: Show that dim $(N(T)) = 1$, and find $C(T)$.

Let $D : \mathcal{C}^\infty([0,1]) \to \mathcal{C}^\infty([0,1])$ defined as $Df = \frac{df}{dx}$.

Exercise: Is $D^2 = D \circ D$ linear? What about D^3?

Exercise: What is $N(D)$? $N(D^2)$? $N(D^k)$?

Question: Is integration linear?

Observe: Images and null spaces are subspaces!
Of which vector space?

Properties of Linear transformations

Let $\mathcal{B} = \{v_1, \ldots, v_n\} \subseteq V$, $T : V \to W$ be linear, and $T(\mathcal{B})$ =$\{T(v_1), \ldots, T(v_n)\}$. Then:

- $T(au + bv) = aT(u) + bT(v)$. In particular, $T(0) = 0$.
- $N(T)$ is a subspace of V. Why? $C(T)$ is a subspace of W. Why?
- If Span$(\mathcal{B}) = V$, is Span$\{T(\mathcal{B})\} = W$? **Note:** It is $C(T)$.

Conclusion: (i) If dim $(V) = n$, then dim $(C(T)) \leq n$.
(ii) T is onto $\Leftrightarrow$ Span$\{T(\mathcal{B}\}) = C(T) = W$.

- $T(u) = T(v) \Leftrightarrow u - v \in N(T)$.

Conclusion: T is one-one $\Leftrightarrow N(T) = 0$.• If $\mathcal{B} \subseteq V$ is linearly independent, is $\{T(\mathcal{B})\} \subseteq W$ linearly independent?

Hint: $a_1T(v_1) + \cdots a_nT(v_n) = 0 \Rightarrow a_1v_1 + \cdots a_nv_n \in N(T)$.

- $S : U \to V$, $T : V \to W$ are linear $\Rightarrow T \circ S : U \to W$ is linear. **Exercise:** Show that $N(S) \subseteq N(T \circ S)$.

How are $C(T \circ S)$ and $C(T)$ related?

Isomorphism of vector spaces

Recall: A linear map $T : V \to W$ is an *isomorphism* if T is also a bijection. **Notation:** $V \simeq W$.

Ques: If $T : V \to W$ is an isomorphism, is $T^{-1} : W \to V$ linear?

Recall: T is one-one $\Leftrightarrow N(T) = 0$ & T is onto $\Leftrightarrow C(T) = W$.

Thus $\boxed{T \text{ is an isomorphism} \Leftrightarrow N(T) = 0 \text{ and } C(T) = W.}$

Example: If V is the subspace of convergent sequences in $\mathbb{R}^\infty$, then $L : V \to \mathbb{R}$ given by $L(x_1, x_2, \ldots) = \lim_{n\to\infty} (x_n)$ is linear.

What is $N(L)$? $C(L)$? Is L one-one or onto?

Exercise: Given $A \in \mathcal{M}_{m\times n}$, let $T(x) = Ax$ for $x \in \mathbb{R}^n$.
Then T is an isomorphism $\Leftrightarrow m = n$ and A is invertible.

Exercise: In the previous examples, identify linear maps which are one-one, and those which are onto.

Example: $S\left(\begin{pmatrix} a & b \\ c & d \end{pmatrix}\right) = (a, b, c, d)^T$ is an isomorphism since $N(S) = 0$ and $C(S) = \mathbb{R}^4$. Thus $\mathcal{M}_{2\times 2} \simeq \mathbb{R}^4$. What is S^{-1}?

Linear Maps and Basis

- Consider $S : \mathcal{M}_{2\times 2} \to \mathbb{R}^4$ given by $S\left(\begin{pmatrix} a & b \\ c & d \end{pmatrix}\right) = (a, b, c, d)^T$.

Recall that $\{e_{11}, e_{12}, e_{21}, e_{22}\}$ is a basis of $\mathcal{M}_{2\times 2}$
such that $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ae_{11} + be_{12} + ce_{21} + de_{22}$.
Observe that $S(e_{11}) = e_1, S(e_{12}) = e_2, S(e_{21}) = e_3$, $S(e_{22}) = e_4$.
Thus, $S(A) = aS(e_{11}) + bS(e_{12}) + cS(e_{21}) + dS(e_{22})= ae_1 + be_2 + ce_3 + de_4 = (a, b, c, d)^T$.

General case:

If $\{v_1, \ldots, v_n\}$ is a basis of V, $T : V \to W$ is linear, $v \in V$,then $v = a_1v_1 + \cdots a_nv_n \Rightarrow T(v) = a_1T(v_1) + \cdots a_nT(v_n)$. Why? Thus,

$$\boxed{T \text{ is determined by its action on a basis,}}$$

i.e., for any n vectors $w_1, \ldots, w_n$ in W (not necessarily distinct), there is unique linear transformation $T : V \to W$ such that $T(v_1) = w_1, \ldots, T(v_n) = w_n$.

Finite-dimensional Vector Spaces

Important Observation: Let dim $(V) = n$, and $\mathcal{B} = \{v_1, \ldots, v_n\}$ be a basis of V. Define $T : V \to \mathbb{R}^n$ by $T(v_i) = e_i$.

e.g., If $v = v_1 + v_n$, $T(v) =$? If $v = 3v_2 - 5v_3$, $T(v) =$?
If $v = a_1v_1 + \cdots + a_nv_n$, $T(v) =$?

Thus $T(v) = [v]_{\mathcal{B}}$.
Is T a linear transformation? What is $N(T)$? What is $C(T)$?

Conclusion: $\boxed{\text{If dim (V) = n, then } V \simeq \mathbb{R}^n.}$

Question: Is $\mathcal{P}_3 \simeq \mathcal{M}_{2\times 2}$?

Key point: Composition of isomorphisms is an isomorphism, and inverse of an isomorphism is an isomorphism.

Exercise: Find 3 isomorphisms each from $\mathcal{P}_3$ to $\mathbb{R}^4$, and $\mathcal{M}_{2\times 2}$ to $\mathbb{R}^4$.

Linear maps from $\mathbb{R}^n$ to $\mathbb{R}^m$

Example: $T(e_1) = \begin{pmatrix} 3 \\ 1 \end{pmatrix}$, $T(e_2) = \begin{pmatrix} 2 \\ -1 \end{pmatrix}$, $T(e_3) = \begin{pmatrix} -5 \\ 0 \end{pmatrix}$
defines a linear map $T : \mathbb{R}^3 \to \mathbb{R}^2$.
If $x = (x_1, x_2, x_3)^T$, then $T(x) = T(x_1e_1 + x_2e_2 + x_3e_3) =$
$x_1T(e_1) + x_2T(e_2) + x_3T(e_3) = x_1 \begin{pmatrix} 3 \\ 1 \end{pmatrix} + x_2 \begin{pmatrix} 2 \\ -1 \end{pmatrix} + x_3 \begin{pmatrix} -5 \\ 0 \end{pmatrix}$, i.e., $T(x) = Ax$, where $A = \begin{pmatrix} 3 & 2 & -5 \\ 1 & -1 & 0 \end{pmatrix}$. **Q:** $A_{*j} =?$

If $x = (x_1, x_2, x_3)^T$, then $T(x) = Ax$, where $A = \begin{pmatrix} 3 & 2 & -5 \\ 1 & -1 & 0 \end{pmatrix}$, i.e., $A_{*j} = T(e_j)$.

General case: If $T : \mathbb{R}^n \to \mathbb{R}^m$ is linear, then
for $x = (x_1, \dots, x_n)^T$ in $\mathbb{R}^n$,

$$T(x) = x_1T(e_1) + \cdots x_nT(e_n) = Ax,$$

where $A = \begin{pmatrix} T(e_1) & \cdots & T(e_n) \end{pmatrix} \in \mathcal{M}_{m\times n}$, i.e., $A_{*j} = T(e_j)$.

Defn. A is called the *standard matrix* of T. Thus

$$\boxed{\text{Linear transformations from } \mathbb{R}^n \text{ to } \mathbb{R}^m}$$
$$\boxed{\text{are in one-one correspondence with } m \times n \text{ matrices.}}$$

Question : Can you imitate this if V and W are not $\mathbb{R}^n$ and $\mathbb{R}^m$?

Matrix Associated to a Linear Map: Example

$S : \mathcal{P}_2 \to \mathcal{P}_1$ given by $S(a_0 + a_1x + a_2x^2) = a_1 + 4a_2x$ is linear.
Question: Is there a matrix associated to S?
Expected size: 2×3. Why?
Idea: Construct an associated linear map $\mathbb{R}^3 \to \mathbb{R}^2$.
Use coordinate vectors! Fix bases $\mathcal{B} = \{1, x, x^2\}$ of $\mathcal{P}_2$, and $\mathcal{C} = \{1, x\}$ of $\mathcal{P}_1$ to do this.
Identify $f = a_0 + a_1x + a_2x^2 \in \mathcal{P}_2$ with $[f]_\mathcal{B} = (a_0, a_1, a_2)^T \in \mathbb{R}^3$,
and $S(f) \in \mathcal{P}_1$ with $[S(f)]_\mathcal{C} = (a_1, 4a_2)^T \in \mathbb{R}^2$.

The associated linear map $S' : \mathbb{R}^3 \to \mathbb{R}^2$ is defined by $S'(a_0, a_1, a_2)^T = (a_1, 4a_2)^T$, i.e., $S'([f]_\mathcal{B}) = [S(f)]_\mathcal{C}$, i.e.,

S' is defined by $S'(e_1) = (0, 0)^T$, $S'(e_2) = (1, 0)^T$, $S'(e_3) = (0, 4)^T \Rightarrow$ the standard matrix of S' is $A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 4 \end{pmatrix}$.

Q: How is A related to S?

Observe: $A_{*1} = [S(1)]_{\mathcal{C}}$, $A_{*2} = [S(x)]_{\mathcal{C}}$, $A_{*3} = [S(x^2)]_{\mathcal{C}}$. **Example:** The matrix of $S(a_0 + a_1x + a_2x^2) = a_1 + 4a_2x$, w.r.t. the bases $\mathcal{B} = \{1, x, x^2\}$ of $\mathcal{P}_2$ and $\mathcal{C} = \{1, x\}$ of $\mathcal{P}_1$ is $A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 4 \end{pmatrix}$ and $\boxed{A_{*1} = [S(1)]_{\mathcal{C}}, A_{*2} = [S(x)]_{\mathcal{C}}, A_{*3} = [S(x^2)]_{\mathcal{C}}.}$

Matrix Associated to a Linear Map

General Case: If $T : V \to W$ is linear, then the matrix of T w.r.t. the ordered bases $\mathcal{B} = \{v_1, \dots, v_n\}$ of V, and $\mathcal{C} = \{w_1, \dots, w_m\}$ of W, denoted $[T]_{\mathcal{C}}^{\mathcal{B}}$, is

$A = \big([T(v_1)]_{\mathcal{C}} \ \cdots \ [T(v_n)]_{\mathcal{C}}\big) \in \mathcal{M}_{m\times n}$.

Example: Projection onto the line $x_1 = x_2$

$P\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \frac{x_1+x_2}{2} \\ \frac{x_1+x_2}{2} \end{pmatrix}$ has standard matrix $\begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix}$.

This is the matrix of P w.r.t. the standard basis.

Question: What is $[P]_{\mathcal{B}}^{\mathcal{B}}$ where $\mathcal{B} = \{(1,1)^T, (-1,1)^T\}$?

Conclusion: The matrix of a transformation depends on the chosen basis. Some are better than others!

Chapter 3. Eigenvalue Decomposition.

3.1 Eigenvalues& Eigenvectors

Eigenvalues and Eigenvectors: Motivation

- Solve the differential equation for u: $\quad du/dt = 3u.$

  The solution is $u(t) = c\,e^{3t}$, $c \in \mathbb{R}$. With initial condition $u(0) = 2$, the solution is $u(t) = 2e^{3t}$.

- Consider the system of linear 1^{st} order differential equations (ODE) with constant coefficients:

$$du_1/dt = 4u_1 - 5u_2, \qquad du_2/dt = 2u_1 - 3u_2,$$

  How does one find the solution?

- Write the system in matrix form $\boxed{du/dt = Au,}$

  where $u(t) = \begin{pmatrix} u_1(t) \\ u_2(t) \end{pmatrix}$, $A = \begin{pmatrix} 4 & -5 \\ 2 & -3 \end{pmatrix}$.

- Assuming the solution is $u(t) = e^{\lambda t}\, v$, where $v = \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2$, we need to find λ and v.

Eigenvalues and Eigenvectors: Definition

We have $u_1' = 4u_1 - 5u_2$, $u_2' = 2u_1 - 3u_2$, where $u_1(t) = e^{\lambda t}\, x, u_2(t) = e^{\lambda t}\, y$

$$\lambda\, e^{\lambda t} x = 4e^{\lambda t} x - 5e^{\lambda t} y, \qquad \lambda\, e^{\lambda t} y = 2e^{\lambda t} x - 3e^{\lambda t} y.$$

Cancelling $e^{\lambda t}$, we get

Eigenvalue problem: Find λ and $v = (x, y)^T$ satisfying

$$4x - 5y = \lambda x, \qquad 2x - 3y = \lambda y.$$

In the matrix form, it is $\boxed{Av = \lambda v}$. This equation has two unknowns, λ and v.

If there exists a λ such that $Av = \lambda v$ has a non-zero solution v, then λ is called an eigenvalue of A and all *nonzero* v satisfying $Av = \lambda v$ are called eigenvectors of A associated to λ.

Question: How many eignevalues can A have? How do we find them & the associated eigenvectors? Reduce the number of unknowns!

Eigenvalues and Eigenvectors: Solving $Ax = \lambda x$

- Rewrite $Av = \lambda v$ as $\boxed{(A - \lambda I)v = 0.}$
- λ is an eigenvalue of A

  $\Leftrightarrow$ there is a nonzero v in the nullspace of $A - \lambda I$

  $\Leftrightarrow N(A - \lambda I) \neq 0$, i.e., dim $(N(A - \lambda I)) \geq 1$,

  $\Leftrightarrow A - \lambda I$ is not invertible

  $\Leftrightarrow \det(A - \lambda I) = 0$.
- $\det(A - \lambda I)$ is a polynomial in the variable λ of degree n. Hence it has *at most* n roots $\Rightarrow A$ has atmost n eigenvalues.
- $\det(A - \lambda I)$ is called the **characteristic polynomial** of A.
- If λ is an eigenvalue of A, then the nullspace of $A - \lambda I$ is called the **eigenspace** of A associated to eigenvalue λ.

Question: When is 0 an eigenvalue of A? What are the corresponding eigenvectors?

Eigenvalues and Eigenvectors: Example

TO SUMMARISE: An eigenvalue of A is a root (in $\mathbb{R}$) of its characteristic polynomial. Any non-zero vector in the corresponding eigenspace is an associated eigenvector.

Recall: The ODE system we want to solve is

$$u_1' = 4u_1 - 5u_2, \qquad u_2' = 2u_1 - 3u_2,$$

The solutions are $u_1(t) = e^{\lambda t}x, u_2(t) = e^{\lambda t}y$, where $(x, y)^T$ is a solution of:

$$\begin{pmatrix} 4 & -5 \\ 2 & -3 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \lambda \begin{pmatrix} x \\ y \end{pmatrix} \qquad (Av = \lambda v)$$

The characteristic polynomial of A is $\det(A - \lambda I)$

$$= \det \begin{pmatrix} 4 - \lambda & -5 \\ 2 & -3 - \lambda \end{pmatrix} = (4 - \lambda)(-3 - \lambda) + 10 = \lambda^2 - \lambda - 2 = (\lambda + 1)(\lambda - 2)$$

The eigenvalues of A are $\boxed{\lambda_1 = -1,\ \lambda_2 = 2.}$

Eigenvalues and Eigenvectors: Example

Eigenvectors v_1 and v_2 associated to $\lambda_1 = -1$ and $\lambda_2 = 2$ respectively are in:

$N(A - \lambda_1 I) = N(A + I)$, and $N(A - \lambda_2 I) = N(A - 2I)$.

Solving $(A + I)v = 0$, i.e., $\begin{pmatrix} 5 & -5 \\ 2 & -2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 0$, we get $N(A + I) = \left\{ \begin{pmatrix} y \\ y \end{pmatrix} \mid y \in \mathbb{R} \right\}$ and hence $v_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ is an eigenvector associated to $\lambda_1 = -1$.

Similarly, solving $(A - 2I)v = 0$ gives $N(A - 2I) = \left\{ \begin{pmatrix} \frac{5y}{2} \\ y \end{pmatrix} \mid y \in \mathbb{R} \right\}$. In particular, $v_2 = \begin{pmatrix} 5 \\ 2 \end{pmatrix}$ is an eigenvector associated to $\lambda_2 = 2$.

Thus, the system $du/dt = Au$ has two special solutions $e^{-t}v_1$ and $e^{2t}v_2$.

Reading Slide - Complete Solution to ODE

Note: When two functions satisfy $du/dt = Au$, then so do their linear combinations.

Complete solution: $u(t) = c_1 e^{-t} v_1 + c_2 e^{2t} v_2$, i.e., $\begin{pmatrix} u_1(t) \\ u_2(t) \end{pmatrix} = c_1 e^{-t} \begin{pmatrix} 1 \\ 1 \end{pmatrix} + c_2 e^{2t} \begin{pmatrix} 5 \\ 2 \end{pmatrix}$.

i.e. $u_1(t) = c_1 e^{-t} + 5c_2 e^{2t}, \quad u_2(t) = c_1 e^{-t} + 2c_2 e^{2t}$.

If we put initial conditions (IC) $u_1(0) = 8$ and $u_2(0) = 5$, then

$$c_1 + 5c_2 = 8, \quad c_1 + 2c_2 = 5 \Rightarrow c_1 = 3, \ c_2 = 1.$$

Hence the solution of the original ODE system with the given IC is

$$u_1(t) = 3e^{-t} + 5e^{2t}, \quad u_2(t) = 3e^{-t} + 2e^{2t}.$$

Finding Eigenvalues: Examples

In some cases it is easy to find the eigenvalues.

Example: $A = \begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix}$ is diagonal. Characteristic polynomial $(3 - \lambda)(2 - \lambda)$.

Eigenvalues: $\lambda_1 = 3, \ \lambda_2 = 2$.
Eigenvectors: $(A - 3I)v_1 = 0 \Rightarrow Av_1 = 3v_1$.

Can take $v_1 = e_1$

Similarly, an eigenvector associated to λ_2 is $v_2 = e_2$

Further, $\mathbb{R}^2$ has a basis consisting of eigenvectors of A: $\{e_1, e_2\}$.

Special case: If A is a diagonal matrix with diagonal entries $\lambda_1, \cdots, \lambda_n$, then

Eigenvalues: $\lambda_1, \cdots, \lambda_n$
Eigenvectors: $e_1, \cdots, e_n$, which form a basis for $\mathbb{R}^n$.

Finding Eigenvalues: Examples

Example: Projection onto the line $x = y$: $P = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix}$. $v_1 = \begin{pmatrix} 1 & 1 \end{pmatrix}^T$ projects onto itself $\Rightarrow \lambda_1 = 1$ with eigenvector v_1. $v_2 = \begin{pmatrix} 1 & -1 \end{pmatrix}^T \mapsto 0 \Rightarrow \lambda_2 = 0$ with eigenvector v_2. Further, $\{v_1, v_2\}$ is a basis of $\mathbb{R}^2$.

Question: Do a collection of eigenvectors always form a basis of $\mathbb{R}^n$?

A: No! **Example:** For $c \in \mathbb{R}$, let $A = \begin{pmatrix} c & 1 \\ 0 & c \end{pmatrix}$.

Characteristic Polynomial: $\det(A - \lambda I) = (c - \lambda)^2$.

Eigenvalues: $\lambda = c$.
Eigenvectors: $(A - I)v = 0 \Rightarrow v = (1\ 0)^T$

Question: Is it unique? Eigenspace of A is 1 dimensional $\Rightarrow$ $\mathbb{R}^2$ has no basis of eigenvectors of A.

Think: What is the advantage of a basis of eigenvectors?

Similarity and Eigenvalues

Defn. The $n \times n$ matrices A and B are *similar*, if there exists an invertible matrix P such that $P^{-1}AP = B$.

Observe: If $B = P^{-1}AP$, then (i) $\det(A) = \det(B)$, and (ii) $B^n = P^{-1}A^nP$ for each n.

Theorem: If A and B are similar, then they have the same characteristic polynomial. In particular, they have the same eigenvalues, $\det(A) = \det(B)$ and Trace(A) = Trace(B).

Proof. Given: $B = P^{-1}AP$. prove: $\det(A - \lambda I) = \det(B - \lambda I)$.

Note: It is enough to prove that $A - \lambda I$ and $B - \lambda I$ are similar!

Indeed, $B - \lambda I = P^{-1}AP - \lambda P^{-1}P$

$= P^{-1}(A - \lambda I)P.$ □

Ques: Why care?

Write $\det(A - \lambda I) = (\lambda_1 - \lambda) \cdots (\lambda_n - \lambda)$. Compare constant coeff.: $\det(A) = \lambda_1 \cdots \lambda_n = \det(B)$; Compare coeff. of λ^{n-1}: Sum of diagonal entries $= a_{11} + \cdots + a_{nn} =$ Trace of $A = \lambda_1 + \ldots + \lambda_n =$ Trace of B.

Ques: How are eigenvalues of A and B related?

Summary: Eigenvalues and Characteristic Polynomial

Let A be $n \times n$.

1. The *characteristic polynomial* of A is $\det(A - \lambda I)$ (of degree n) and its roots are the *eigenvalues* of A.
2. For each eigenvalue λ, the associated *eigenspace* is $N(A - \lambda I)$. To find it, solve $(A - \lambda I)v = 0$. Any non-zero vector in $N(A - \lambda I)$ is an *eigenvector* associated to λ.

3. If A is a diagonal matrix with diagonal entries $\lambda_1, \cdots, \lambda_n$, then its eigenvalues are $\lambda_1, \cdots, \lambda_n$ with associated eigenvectors $e_1, \cdots, e_n$ respectively.

4. Write $\det(A - \lambda I) = (\lambda_1 - \lambda) \cdots (\lambda_n - \lambda)$ and expand.

   Trace of A = $a_{11} + \cdots + a_{nn}$ (sum of diagonal entries)

   = $\lambda_1 + \ldots + \lambda_n$

   $\det(A) = \lambda_1 \cdots \lambda_n$

   THUS: If λ_1, ..., λ_n are real numbers, then $\mathrm{Tr}(A)$ = sum of eigenvalues, and $\det(A)$ = product of eigenvalues.

3.2 DIAGONALIZABILITY

Diagonizability: Introduction

Note: Finding roots of characteristic polynomials (and hence eigenvalues) is difficult in general.

For $n \geq 5$, no formula exists for roots. (Abel, Galois)

For $n = 3, 4$, formulae for root exist, but not easy to use.

Defn. An $n \times n$ matrix A is diagonalizable if A is similar to a diagonal matrix Λ, i.e., there is an invertible matrix P and a diagonal matrix Λ such that $P^{-1}AP = \Lambda$.

Importance of Diagonalizability:

Let the $n \times n$ matrix A be diagonalizable, i.e., $P^{-1}AP = \Lambda$, where P is invertible and Λ is diagonal. If this happens,

- The eigenvalues of A are the diagonal entries of Λ,
- $\det(A)$ is the product of the diagonal entries of Λ, and
- $\mathrm{Trace}(A)$ = sum of the diagonal entries of Λ.
- **Other Information:** e.g., what is $\mathrm{Trace}(A^n)$?

Diagonalization: Example

Example: $A = \begin{pmatrix} 1 & 5 & 6 \\ 0 & 2 & -4 \\ 0 & 0 & 3 \end{pmatrix}$ is triangular.

$\det(A - \lambda I) = (1 - \lambda)(2 - \lambda)(3 - \lambda)$.
Eigenvalues: $\lambda_1 = 1$, $\lambda_2 = 2$, $\lambda_3 = 3$.

$\boxed{\textbf{Note: } \text{If } A \text{ is triangular, its eigenvalues are on the diagonal}}$

Eigenvectors: $v_1 = e_1$, $v_2 = \begin{pmatrix} 5 & 1 & 0 \end{pmatrix}^T$, $v_3 = \begin{pmatrix} -7 & -4 & 1 \end{pmatrix}^T$. (**How?**)Further, $\{v_1, v_2, v_3\}$ is a basis of $\mathbb{R}^3$. Hence $P = \begin{pmatrix} v_1 & v_2 & v_3 \end{pmatrix}$ is invertible , and $AP = \begin{pmatrix} Av_1 & Av_2 & Av_3 \end{pmatrix} = \begin{pmatrix} v_1 & 2v_2 & 3v_3 \end{pmatrix} = P\Lambda$, where $\Lambda = \begin{pmatrix} 1 & & \\ & 2 & \\ & & 3 \end{pmatrix}$. Thus $P^{-1}AP = \Lambda$, i.e., A is diagonalizable.

Example: If $\mathcal{B} = \{v_1, v_2, v_3\}$, and $T(v) = Av$, then $[T]_{\mathcal{B}}^{\mathcal{B}} =$ ___.

Eigenvalue Decomposition (EVD)

Question: What is the advantage of a basis of $\mathbb{R}^n$ consisting of eigenvectors?

Let A be an $n \times n$ matrix with n eigenvectors $v_1, \dots, v_n$, associated to eigenvalues $\lambda_1, \dots, \lambda_n$. If $\mathcal{B} = \{v_1, \dots, v_n\}$ is a basis of $\mathbb{R}^n$, then the matrix $P = \begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix}$ is invertible.

Moreover, $AP = A\begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix} = \begin{pmatrix} Av_1 & \cdots & Av_n \end{pmatrix} = \begin{pmatrix} \lambda_1 v_1 & \cdots & \lambda_n v_n \end{pmatrix} = P\Lambda$, where $\Lambda = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$. Therefore $P^{-1}AP = \Lambda$, i.e., A is similar to a diagonal matrix.

Thus: $\boxed{\text{Eigenvectors diagonalize a matrix}}$

Eigenvalue Decomposition (EVD): Let A be diagonalizable.
With notation as above, we have $A = P\Lambda P^{-1}$.
This is called as the eigenvalue decompostion (EVD) of A.

Diagonizability and Eigenvectors

Theorem A is diagonalizable $\Leftrightarrow$ A has n linearly independent eigenvectors. In particular, $\mathbb{R}^n$ has a basis consisting of eigenvectors of A.

Proof. ($\Leftarrow$): Done! To prove ($\Rightarrow$), assume $P = \begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix}$ is an invertible matrix such that $P^{-1}AP = \Lambda = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$. Then $AP = P\Lambda$, i.e. $\begin{pmatrix} Av_1 & \dots & Av_n \end{pmatrix} = \begin{pmatrix} \lambda_1 v_1 & \dots & \lambda_n v_n \end{pmatrix}$. Therefore $v_1, \dots, v_n$ are eigenvectors of A. They are linearly independent since P is invertible. □

Question: Is every matrix is diagonalizable? **A:** No.

Examples: $Q = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ no eigenvalues (over $\mathbb{R}$)!

$P = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ not enough eigenvectors!

Diagonalizability: Summary

Thus: If an $n \times n$ matrix A has n linearly independent eigenvectors $v_1, \dots, v_n$, then A is diagonalizable. Moreover, if $\lambda_1, \dots, \lambda_n$ are the corresponding eigenvalues, then $P^{-1}AP = \Lambda$, where the diagonalizing matrix is $P = \begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix}$, and the diagonal matrix is $\Lambda = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$, i.e., $\boxed{P^{-1}AP = \Lambda}$, where

$\boxed{\text{The diagonal entries of } \Lambda \text{ are eigenvalues of } A \text{ and}}$

$$\boxed{\text{The columns of } P \text{ are corresponding eigenvectors of } A.}$$
$$\boxed{\text{The EVD of } A \text{ is } A = P\Lambda P^{-1}.}$$

Note: P need not be unique, e.g., replace v_1 by $2v_1$, etc.

When is A Diagonalizable?

Ans: When A has n linearly independent eigenvectors. **Ques:** When does A have n linearly independent eigenvectors?

- If $v_1, \ldots, v_r$ are eigenvectors of A associated to distinct eigenvalues $\lambda_1, \ldots, \lambda_r$, then $v_1, \ldots, v_r$ are linearly independent.

Proof. Suppose $v_1, \ldots, v_r$ are linearly dependent. Choose a linear relation involving minimum number of v_i's, say

(1) $a_1v_1 + \cdots + a_tv_t = 0$. ($1 < t \leq r$, t is minimal, $a_i \neq 0$)

Apply A to get $a_1\lambda_1v_1 + \cdots + a_t\lambda_tv_t = 0$ (2)

$\lambda_1(1) - (2)$ gives $a_2(\lambda_1 - \lambda_2)v_2 + \cdots + a_t(\lambda_1 - \lambda_t)v_t = 0$,

which contradicts the minimality of t. □

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

Proof. If $v_1, \ldots, v_n$ are eigenvectors associated to distinct eigenvalues $\lambda_1, \ldots, \lambda_n$, then $\{v_1, \ldots, v_n\}$ is linearly independent.

Then $P = \begin{pmatrix} v_1 & \ldots & v_n \end{pmatrix}$ is invertible, and $P^{-1}AP = \Lambda$ as seen earlier. Hence A is diagonalizable. □

Reading Slide - Eigenvalues of AB and $A + B$

- If λ is an eigenvalue of A, μ is an eigenvalue of B, is $\lambda\mu$ an eigenvalue of AB?

False Proof. $ABx = A(\mu x) = \mu(Ax) = \lambda\mu x$.

This is false since A and B may not have same eigenvector x.

- **Example:** $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, $B = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$, $AB = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$.

The eigenvalues of A and B are $0, 0$ and that of AB are $1, 0$.

- Eigenvalues of $A + B$ are NOT $\lambda + \mu$.

In above example, $A + B = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ has eigenvalues $1, -1$.

- If A and B have same eigenvectors associated to λ and μ, then $\lambda\mu$ and $\lambda + \mu$ are eigenvalues of AB and $A + B$ respectively.

Question: When do A and B have the same eigenvectors?

Extra Reading: Simultaneous Diagonalizability
Assume $\boxed{A \text{ and } B \text{ are diagonalizable}}$. Then A and B have same eigenvector matrix S if and only if $AB = BA$.

Proof. ($\Rightarrow$) Assume $S^{-1}AS = \Lambda_1$ and $S^{-1}BS = \Lambda_2$, where Λ_1 and Λ_2 are diagonal matrices. Then $AB = (S\Lambda_1 S^{-1})(S\Lambda_2 S^{-1}) = S(\Lambda_1\Lambda_2)S^{-1}$ and $BA = S(\Lambda_2\Lambda_1)S^{-1}$. Since $\Lambda_1\Lambda_2 = \Lambda_2\Lambda_1$, we

get $AB = BA$.

(*Part of* $\Leftarrow$) Assume $AB = BA$. If $Ax = \lambda x$, then $ABx = B(Ax) = B(\lambda x) = \lambda Bx$. If $Bx = 0$, then x is an eigenvector of B, associated to $\mu = 0$.If $Bx \neq 0$, then x and Bx both are eigenvectors of A, associated to λ.

Special case: Assume all the eigenspaces of A are one dimensional. Then $Bx = \mu x$ for some scalar $\mu \Rightarrow x$ is an eigenvector of B. We will not prove the general case.

Eigenvalues of A^k
• If $Av = \lambda v$, then $A^2 v = A(Av) = A(\lambda v) = \lambda(Av) = \lambda^2 v$. Similarly $A^k v = \lambda^k v$ for any $k \geq 0$.

Thus if v is an eigenvector of A with associated eigenvalue λ, then v is also an eigenvector of A^k with associated eigenvalue λ^k for $k \geq 0$. If A is invertible, then $\lambda \neq 0$. Hence, the same also holds for $k < 0$ since $A^{-1}v = \lambda^{-1}v$.

• If A is diagonalizable , then $P^{-1}AP = \Lambda$ is diagonal where columns of P are eigenvectors of A.

Since $(P^{-1}A^kP) = \Lambda^k$, which is diagonal, we see that A^k is diagonalizable, and the eigenvectors of A^k are same as eigenvectors of A. Similarly, the same also holds for $k < 0$ if A is invertible.

Question: What is the EVD of A^k.

Reading Slide - Application: Fibonacci Numbers
Let $F_0 = 0$, $F_1 = 1$ and $F_k = F_{k-1} + F_{k-2}$ for $k \geq 2$ define the Fibonacci sequence. What is the kth term?

If $u_k = \begin{pmatrix} F_{k+1} \\ F_k \end{pmatrix}$, then $\begin{pmatrix} F_{k+1} \\ F_k \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} F_k \\ F_{k-1} \end{pmatrix}$, i.e., $u_k = Au_{k-1}$ for $n \geq 1$, where $A = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \Rightarrow u_k = A^k u_0$ for $k \geq 1$.

Characteristic polynomial of A: $\lambda^2 - \lambda - 1$; Eigenvalues: $\lambda_1 = \dfrac{1+\sqrt{5}}{2}$, $\lambda_2 = \dfrac{1-\sqrt{5}}{2}$.

There are 2 distinct eigenvalues $\Rightarrow$ the associated eigenvectors x_1 and x_2 are linearly independent $\Rightarrow \{x_1, x_2\}$ is a basis for $\mathbb{R}^2$.

Write $u_0 = c_1x_1 + c_2x_2$. Then $u_k = A^k u_0 = A^k(c_1x_1 + c_2x_2)$

$$= c_1A^kx_1 + c_2A^kx_2 = c_1\left(\frac{1+\sqrt{5}}{2}\right)^k x_1 + c_2\left(\frac{1-\sqrt{5}}{2}\right)^k x_2.$$

Q: Find x_1, x_2, c_1 and c_2 and get the exact formula for F_k.

An Application: Predator-Prey Model

Let the owl and rat populations at time k be O_k and R_k respectively. Owls prey on the rats, so if there are no rats, the population of owls will go down by 50%. If there are no owls to prey on the rats, then the rat population will increase by 10%.

In particular, let us assume the rat and owl populations dependence is given as follows:

$$\begin{aligned} O_{k+1} &= 0.5O_k + 0.4R_k \\ R_{k+1} &= -pO_k + 1.1R_k \end{aligned}$$

The term $-p$ calculates the rats preyed by the owls.

Thus, if $P_k = \begin{pmatrix} O_k \\ R_k \end{pmatrix}$ and $A = \begin{pmatrix} 0.5 & 0.4 \\ -p & 1.1 \end{pmatrix}$, then $P_{k+1} = AP_k$ for all k. In particular, $P_k = A^k P_0$.

Exercise: If we start with a certain initial population of owls and rats, how many will be there in, say, 50 years, i.e., given P_0, what is P_{50}? What is the steady state, i.e., what is $\lim_{k\to\infty} P_k$?

An Application: Steady State

Suppose we have a system where the current state u_k depends on the previous one u_{k-1} linearly, i.e., $u_k = Au_{k-1}$. Then observe that $u_k = A^k u_0$. The steady state of the system is $u_\infty = \lim_{k\to\infty}(u_k)$. How do we find this?

- If u_0 is an eigenvector of A associated to λ, then $u_k = \lambda^k u_0$.
- Let $v_1, \ldots, v_r$ be eigenvectors of A associated respectively to $\lambda_1, \ldots, \lambda_r$. If $u_0 \in \text{Span}\{v_1, \ldots, v_r\}$, i.e., $u_0 = c_1v_1 + \cdots + c_rv_r$ for scalars $c_1, \ldots, c_r$, then $u_k = A^k u_0 = c_1A^kv_1 + \cdots + c_rA^kv_r = c_1\lambda_1^k v_1 + \cdots + c_r\lambda_r^k v_r$. In particular, if A is diagonalizable, then there is a basis of $\mathbb{R}^n$ of eigenvectors of A. Hence, this is applicable to every $u_0 \in \mathbb{R}^n$.

Let A be diagonalizable, and u_k represent population.

- Under what conditions will there be a population explosion?
- What conditions will force the population to become extinct?
- When does it stabilise (to a non-zero value)?

Hint: Depends on $|\lambda_i|$.

Extra Reading:Complex Eigenvalues

Example: Rotation by 90^o in $\mathbb{R}^2$ is given by $Q = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$. It has no real eigenvectors since rotation by 90^o changes the direction.

Q has eigenvalues, but they are not real. $\det(Q - \lambda I) = \lambda^2 + 1 \Rightarrow \lambda_1 = i$ and $\lambda_2 = -i$, where $i^2 = -1$. Let us compute the eigenvectors.

$(Q - iI)x_1 = \begin{pmatrix} -i & -1 \\ 1 & -i \end{pmatrix} x_1 = 0 \Rightarrow x_1 = \begin{pmatrix} 1 \\ -i \end{pmatrix}$, $(Q + iI)x_2 = \begin{pmatrix} i & -1 \\ 1 & i \end{pmatrix} x_2 = 0 \Rightarrow x_2 = \begin{pmatrix} 1 \\ i \end{pmatrix}$.

The eigenvalues, though imaginary, are distinct, hence eigenvectors are linearly independent.

If $P = \begin{pmatrix} x_1 & x_2 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ -i & i \end{pmatrix}$, then $P^{-1}QP = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}$.

Conclusion: We need complex numbers $\mathbb{C}$ even if we are working with real matrices. Over $\mathbb{C}$, an $n \times n$ matrix A always has n eigenvalues.

Reason: Fundamental theorem of Algebra

Every polynomial over $\mathbb{C}$ of degree n has n roots in $\mathbb{C}$.

Chapter 4. Orthogonality and Projections

4.1 Orthogonality

Inner product on $\mathbb{R}^n$

Defn. Define the **inner product** (dot product) of two vectors $v, w \in \mathbb{R}^n$ as $\boxed{v \cdot w = v^T w}$

For v, w in $\mathbb{R}^n$ and c in $\mathbb{R}$

- $v \cdot w = v^T w = v_1 w_1 + \cdots + v_n w_n = w^T v = w \cdot v$.
- (Bilinearity)

$$\begin{aligned} &(v+w)\cdot z = (v+w)^T z = v^T z + w^T z = v\cdot z + w \cdot z \\ &cv \cdot w = (cv)^T w = c\,(v^T w) = v^T (c\,w) = v \cdot cw. \end{aligned}$$

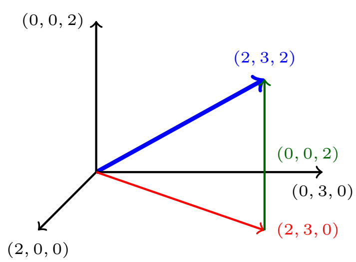
- $v \cdot v = v^T v \geq 0$ and $v^T v = 0$ if and only if $v = 0$.

Define **length** (or norm) of v in $\mathbb{R}^n$ to be $\boxed{\|v\| = \sqrt{v \cdot v}.}$

Henceforth we will use $v^T w$ directly to write the dot product.

Reading : Length of a vector in $\mathbb{R}^3$ and $\mathbb{R}^n$

Let $v = (2,3,2)$. By Pythagoras theorem, $||v|| = \sqrt{||(2,3,0)||^2 + ||(0,0,2)||^2}$



$= \sqrt{2^2 + 3^2 + 2^2} = \sqrt{17}$.

Generalize by induction: Let $v = (x_1, \cdots, x_n)^T \in \mathbb{R}^n$. Define

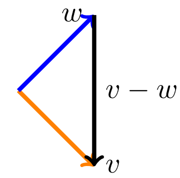
$\|v\| = \sqrt{\|(x_1, \cdots, x_{n-1}, 0)\|^2 + \|(0,0,\cdots,x_n)\|^2} = \sqrt{x_1^2 + \cdots + x_{n-1}^2 + x_n^2} = \sqrt{v^T v}$.

The length in $\mathbb{R}^n$ is compatible with the vector space structure. Let $v, w \in \mathbb{R}^n$ and $c \in \mathbb{R}$. Then,

- $\|v\| \geq 0$ and $\|v\| = 0$ if and only if $v = 0$.
- $\|cv\| = |c|\|v\|$ • $\|v + w\| \leq \|v\| + \|w\|$ (Triangle Inequality)

Orthogonal vectors in $\mathbb{R}^n$

We say vectors v and w in $\mathbb{R}^n$ are orthogonal (perpendicular) if they satisfy the Pythagoras theorem that is, $||v||^2 + ||w||^2 = ||v - w||^2$



$$\begin{aligned}
\|v\|^2 + \|w\|^2 &= (v - w)^T (v - w) \\
&= (v^T - w^T)(v - w) \\
&= v^T v - w^T v - v^T w + w^T w \\
&= \|v\|^2 - 2\, v^T w + \|w\|^2 \quad \text{(since } w^T v = v^T w \text{)}
\end{aligned}$$

Therefore, v and w are defined to be *orthogonal* if $v^T w =$

Think! What can be said about Span$\{v\}$ and Span$\{w\}$ when v and w are orthogonal to each other in $\mathbb{R}^3$?

Orthogonal and Orthonormal Sets

Defn. A set of *non-zero* vectors $\{v_1, \cdots, v_k\} \subset \mathbb{R}^n$, is said to be an **orthogonal set** if $v_i^T v_j = 0$ **for all** $i, j = 1, \cdots n, \; i \neq j$.

Examples: $\{(1, 3, 1), (-1, 0, 1)\} \subset \mathbb{R}^3$, $\{(2, 1, 0, -1), (0, 1, 0, 1), (-1, 1, 0, -1)\} \subseteq \mathbb{R}^4$, $\left\{(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}), (\frac{1}{\sqrt{2}}, 0, \frac{-1}{\sqrt{2}})\right\} \subseteq \mathbb{R}^3$, $\{e_1, \cdots, e_n\} \subseteq \mathbb{R}^n$.

Of these, the last two examples have all unit vectors (vectors of length one).

Defn. An orthogonal set $\{v_1, \cdots, v_k\} \subseteq \mathbb{R}^n$ with all unit vectors, i.e., $\|v_i\| = 1$ for all i, is called an **orthonormal set**.

Note: If $\{v_1, \cdots, v_k\}$ is an orthogonal set, then $\{u_1, \cdots, u_k\}$ is orthonormal, for $u_i = v_i / ||v_i||$.

Exercise: If $S = \{v_1, \ldots, v_k\}$ is an orthogonal set, then v_k is orthogonal to each $v \in$ Span$\{v_1, \ldots, v_{k-1}\}$.

Orthogonality and Linear Independence

Theorem: An orthogonal set in $\mathbb{R}^n$ is linearly independent. *Proof.* Let $\{v_1, \cdots, v_k\}$ be an orthogonal set in $\mathbb{R}^n$, i.e. $v_i \neq 0$ and $v_i^T v_j = 0$ for $i \neq j$. Note that for $i = j$, $v_i^T v_i = ||v_i||^2 \neq 0$.

Assume for some $a_1, \cdots, a_k \in \mathbb{R}$,

$$\begin{aligned}
a_1 v_1 + a_2 v_2 + \cdots + a_k v_k &= 0 \\
\Rightarrow (a_1 v_1 + a_2 v_2 + \cdots + a_k v_k)^T v_1 &= 0 \cdot v_1 = 0 \\
\Rightarrow (a_1 v_1^T + a_2 v_2^T + \cdots + a_k v_k^T)\; v_1 &= 0 \\
\Rightarrow a_1 v_1^T v_1 + a_2 v_2^T v_1 + \cdots + a_k v_k^T v_1 &= 0 \\
\Rightarrow a_1 ||v_1||^2 &= 0 \\
\Rightarrow a_1 &= 0 \text{ since } v_1 \neq 0
\end{aligned}$$

Similarly, we get $a_2 = \cdots = a_n = 0$. Hence $\{v_1, \cdots, v_k\}$ is linearly independent. **True/False:** Any matrix whose columns form an orthogonal set is invertible. Give example

Matrices with Orthogonal Columns

Let $A = [v_1 \ \cdots \ v_n]$ be $m \times n$. If $\{v_1, \ldots, v_n\}$ form an *orthonormal* set in $\mathbb{R}^m$, then

$$A^T A = \begin{pmatrix} v_1^T \\ \vdots \\ v_n^T \end{pmatrix} \begin{pmatrix} v_1 & \ldots & v_n \end{pmatrix} = \begin{pmatrix} v_1^T v_1 & \ldots & v_1^T v_n \\ \vdots & & \vdots \\ v_n^T v_1 & \ldots & v_n^T v_n \end{pmatrix} = I_n.$$

Defn. A square matrix A whose column vectors form an orthonormal set is called an orthogonal matrix.

If $Q = [u_1 \ \cdots \ u_n]$ is an orthogonal matrix, then
- $\{u_1, \ldots, u_n\}$ is an orthonormal set (by definition)
- $Q^T Q = I = QQ^T$ Why?
- $\|Qv\| = \sqrt{(Qv)^T(Qv)} = \sqrt{v^T Q^T Q v} = \sqrt{v^T x} = \|v\|$.

$\Rightarrow$ the only (real) eigenvalues of Q, if they exist, are ± 1.
- Row vectors of Q are orthonormal since $QQ^T = I$.

Orthogonal Matrices: Examples

Examples: 1. $\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$. 2. $\frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$.

3. $\frac{1}{3} \begin{pmatrix} 2 & 1 & -2 \\ 2 & -2 & 1 \\ 1 & 2 & 2 \end{pmatrix}$. 4. $\frac{1}{2} \begin{pmatrix} 1 & -1 & -1 & 1 \\ 1 & 1 & 1 & 1 \\ -1 & -1 & 1 & 1 \\ -1 & 1 & -1 & 1 \end{pmatrix}$.

Orthogonal Basis

Defn. A basis $\mathcal{B} = \{v_1, \ldots, v_k\}$ of a subspace V of $\mathbb{R}^n$ is an orthogonal basis if it is an orthogonal set, i.e., $v_i^T v_j = 0$ for $i \neq j$.
Furthermore, if $||v_i|| = 1$ for each i, then $\mathcal{B}$ is an orthonormal basis (or o.n.b.) of V.

Example: Consider the bases of $\mathbb{R}^2$: $\mathcal{B}_1 = \{w_1 = (8,0)^T, w_2 = (6,3)^T\}$,

$\mathcal{B}_2 = \{(8,0)^T, (0,3)^T\}$ and $\mathcal{B}_3 = \left\{ \left(\frac{8}{\sqrt{8^2+0^2}}, 0 \right)^T, \left(0, \frac{3}{\sqrt{0^2+3^2}} \right)^T \right\}$.

Then $\mathcal{B}_1$ is not orthogonal, $\mathcal{B}_2$ is an orthogonal basis, but not an orthonormal basis, and $\mathcal{B}_3$ is an orthonormal basis of $\mathbb{R}^2$.

Note: If $\{u_1, \ldots, u_k\} \subseteq \mathbb{R}^n$ is an orthonormal set, then it is an o.n.b. of $V = \text{Span}\{u_1, \ldots, u_k\}$.

Importance of Orthogonal Basis

Example : The set $\mathcal{B} = \{v_1 = (-1,1)^T, v_2 = (1,1)^T\}$ is a orthogonal basis of $\mathbb{R}^2$.

- Find $[v]_{\mathcal{B}} = (a,b)^T$: $\qquad v = av_1 + bv_2 = a(-1,1)^T + b(1,1)^T$

$v_1^T v = (-1,1)v = a(-1,1)(-1,1)^T = 2a = a||v_1||^2$

Then $a = \dfrac{v_1^T v}{2} = \dfrac{v_1^T v}{||v_1||^2}$ and $b = \dfrac{v_2^T v}{2} = \dfrac{v_2^T v}{||v_2||^2}$

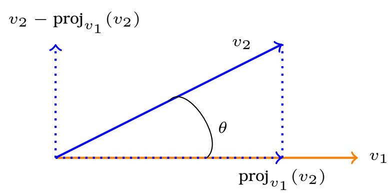
General Case: If $\mathcal{B} = \{v_1, \ldots, v_n\}$ is an o.n.b of V, then $[v]_{\mathcal{B}} = (c_1, \ldots, c_n)^T$, where $c_j = v_j^T v$.
Moreover, if $T : V \to V$ is linear, and $[T]_{\mathcal{B}}^{\mathcal{B}} = [a_{ij}]$, then
$[T]_{\mathcal{B}}^{\mathcal{B}} = \left([T(v_1)]_{\mathcal{B}} \quad \cdots \quad [T(v_n)]_{\mathcal{B}}\right) \Rightarrow a_{ij} = \_\_.$

Think!

When does T map orthogonal sets to orthogonal sets?

Orthogonal Basis and Projections

Every subspace of $\mathbb{R}^n$ has an orthogonal basis.
To construct one, we can start with any basis and modify it (Gram-Schmidt process).
First we see what happens in $\mathbb{R}^2$.



To construct an orthogonal basis in $\mathbb{R}^n$, we need to know how to find $\text{proj}_{v_1}(v_2)$ in $\mathbb{R}^n$.

Orthogonal Projections in $\mathbb{R}^n$

If $v(\neq 0), w \in \mathbb{R}^n$, then $\text{proj}_v(w)$, is a multiple of v and $w - \text{proj}_v(w)$ is orthogonal to v. Thus

$$\begin{aligned} \text{proj}_v w &= av \text{ for some } a \in \mathbb{R} \\ v^T(w - \text{proj}_v w) &= 0 \\ v^T w - v^T av = 0 &\Leftrightarrow a = \frac{v^T w}{v^T v} \end{aligned}$$

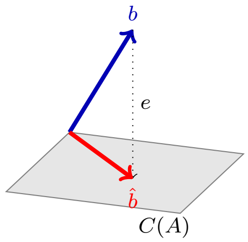
Therefore $\boxed{\text{proj}_v(w) = \left(\dfrac{v^T w}{v^T v}\right) v.}$

Example. If $w = \begin{pmatrix}1 & 1 & 1\end{pmatrix}^T$ and $v = \begin{pmatrix}1 & 2 & 3\end{pmatrix}^T$, then the orthogonal projection of w on Span$\{v\}$ is given by $\text{proj}_v(w) = \left(\dfrac{v^T w}{v^T v}\right) v = \dfrac{6}{14}\begin{pmatrix}1\\2\\3\end{pmatrix}$.

Application: This can be used to "solve" an inconsistent system of equations.

Linear Least Squares and Projections

Suppose system $Ax = b$ is inconsistent, i.e. $b \notin C(A)$. The error $E = \|Ax - b\|$ is the distance from b to $Ax \in C(A)$.



We want the least square solution $\hat{x}$ which minimizes E, i.e., we want to find $\hat{b}$ closest to b such that $A\hat{x} = \hat{b}$ is a consistent system.

Therefore, $\hat{b} = \text{proj}_{C(A)}(b)$ and $A\hat{x} = \hat{b}$. The error vector $e = b - A\hat{x}$ must be perpendicular to $C(A)$,

which is also the row space of A^T.

So, e must be in the left null space of A, $N(A^T)$, i.e.,

$$A^T(b - A\hat{x}) = 0 \text{ or } \boxed{A^T A\,\hat{x} = A^T b}$$

Therefore, to find $\hat{x}$, we need to solve $A^T A\,\hat{x} = A^T b$.

Linear Least Squares and Projections

Let A be $m \times n$. Then $A^T A$ is a symmetric $n \times n$ matrix.

- $\boxed{N(A^T A) = N(A)}$.

  Proof. $Ax = 0 \Rightarrow A^T Ax = 0$. So, $N(A) \subseteq N(A^T A)$.
  For the other inclusion, take $x \in N(A^T A)$.
  $A^T Ax = 0 \Rightarrow x^T(A^T Ax) = (Ax)^T(Ax) = \|Ax\|^2 = 0$
  $\Rightarrow Ax = 0$,i.e., $x \in N(A)$.

- Since $N(A) = N(A^T A)$, by rank-nullity theorem, rank$(A) = n - \dim\,(N(A)) =$ rank$(A^T A)$.

- A has linearly independent columns $\Leftrightarrow$ rank$(A) = n \Leftrightarrow$ rank$(A^T A) = n \Leftrightarrow A^T A$ is invertible.

- If rank$(A) = n$, then the least square solution of $Ax = b$ is given by $A^T A\,\hat{x} = A^T b \Rightarrow$ $\boxed{\hat{x} = (A^T A)^{-1} A^T b}$ and the orthogonal proj. of b on $C(A)$ is $\boxed{\hat{b} = A\hat{x} = Pb}$, where $\boxed{P = A(A^T A)^{-1} A^T}$ is the projection matrix. **Ques:** Is $P^2 = P$?

Linear Least Squares: Example

Example: Find the least square solution to the system

$$\begin{pmatrix} -1 & 2 \\ 2 & -3 \\ -1 & 3 \end{pmatrix} x = \begin{pmatrix} 4 \\ 1 \\ 2 \end{pmatrix} \quad (Ax = b)$$

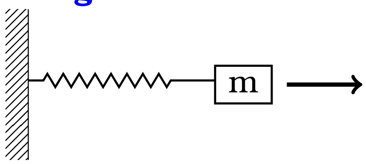
We need to solve $A^T A\,\hat{x} = A^T b$. Now $A^T b = \begin{pmatrix} -4 \\ 11 \end{pmatrix}$ and $A^T A = \begin{pmatrix} 6 & -11 \\ -11 & 22 \end{pmatrix}$.

$$[A^T A \mid A^T b] = \begin{pmatrix} 6 & -11 & | & -4 \\ -11 & 22 & | & 11 \end{pmatrix} \to \begin{pmatrix} 6 & -11 & | & -4 \\ 0 & 11/6 & | & 11/3 \end{pmatrix}.$$

Therefore $\hat{x_2} = 2$, and $\hat{x_1} = 3$.

Exercise: Find the projection matrix P, and check that $Pb = A\hat{x}$.

Reading Slide: Linear Least Squares: Application



Hooke's Law states that displacement x of the spring is directly proportional to the load (mass) applied, i.e., $m = kx$.

A student performs experiments to calculate spring constant k. The data collected says for loads $4, 7, 11$ kg applied, the displacement is $3, 5, 8$ inches respectively. Hence we have:

$$\begin{pmatrix} 3 \\ 5 \\ 8 \end{pmatrix} k = \begin{pmatrix} 4 \\ 7 \\ 11 \end{pmatrix} \qquad (ak = b).$$

Clearly the data is inconsistent.
Allowing for various errors, how do we find an estimate for k?
The method of least squares allows us to find a consistent system "close" to this one!
Exercise: Estimate k using the method of least squares.

Reading Slide: Line of Best Fit: Example

Question: We want to find the best line $y = C + Dx$ which fits the given data and gives least square error.

Data: (x, y) = $(-2, 4)$, $(-1, 3)$, $(0, 1)$, and $(2, 0)$.

The system $\begin{pmatrix} 1 & -2 \\ 1 & -1 \\ 1 & 0 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} C \\ D \end{pmatrix} = \begin{pmatrix} 4 \\ 3 \\ 1 \\ 0 \end{pmatrix}$ $(Ax = b)$
is inconsistent.

Find the least square solution by solving $A^T A\, \hat{x} = A^T b$.

Question: Find the best quadratic curve $y = C + Dx + Ex^2$ which fits the above data and gives least square error.

Hint. The first row of the matrix A in this case will be $[1 \quad -2 \quad 4]$.

Gram-Schmidt Process

If the set of vectors $v_1, \dots, v_r$ in $\mathbb{R}^n$ are linearly independent, then we can find an orthonormal set of vectors $q_1, \dots, q_r$ such that $\text{Span}\{v_1, \dots, v_r\} = \text{Span}\{q_1, \dots, q_r\}$.

First find an orthogonal set.

Let $w_1 = v_1$, $w_2 = v_2 - \text{proj}_{w_1}(v_2)$. Then $w_1 \perp w_2$ and $\text{Span}\{v_1, v_2\} = \text{Span}\{w_1, w_2\}$.

Let $c_1w_1 + c_2w_2$ be the projection of v_3 on $\text{Span}\{w_1, w_2\}$. Then $(v_3 - c_1w_1 - c_2w_2) \perp w_1$ and $(v_3 - c_1w_1 - c_2w_2) \perp w_2$. $\Rightarrow w_1^T(v_3 - c_1w_1 - c_2w_2) = 0 \Rightarrow c_1w_1 = \text{proj}_{w_1}(v_3)$ and similarly $c_2w_2 = \text{proj}_{w_2}(v_3)$. Therefore,

$$w_3 = v_3 - \text{proj}_{\text{Span}\{w_1,w_2\}}(v_3) = v_3 - \left(\frac{w_1^T v_3}{\|w_1\|^2}\right) w_1 - \left(\frac{w_2^T v_3}{\|w_2\|^2}\right) w_2.$$

$$\boxed{\text{Span}\{v_1, v_2, v_3\} = \text{Span}\{w_1, w_2, w_3\} \text{ and } w_1^T w_3 = 0, w_2^T w_3 = 0.}$$

By induction,

$$\begin{aligned} w_r := v_r - \text{proj}_{\text{Span}\{w_1,\dots w_{r-1}\}}(v_r) &= v_r - \text{proj}_{w_1}(v_r) - \text{proj}_{w_2}(v_r) - \cdots - \text{proj}_{w_{r-1}}(v_r) \\ &= v_r - \frac{w_1^T v_r}{\|w_1\|^2} w_1 - \frac{w_2^T v_r}{\|w_2\|^2} w_2 - \cdots - \frac{w_{r-1}^T v_r}{\|w_{r-1}\|^2} w_{r-1} \end{aligned}$$

Now take $q_1 = \dfrac{w_1}{\|w_1\|}$, $q_2 = \dfrac{w_2}{\|w_2\|}$, ..., $q_r = \dfrac{w_r}{\|w_r\|}$. Then $\{q_1, \dots, q_r\}$ is an orthonormal set and $W = \text{Span}\{v_1, \dots, v_r\} = \text{Span}\{w_1, \dots, w_r\} = \text{Span}\{q_1, \dots, q_r\}$.

In particular, $\{q_1, q_2, \dots, q_r\}$ is an *orthonormal basis* for W.

Exercise: Show that if $\{w_1, \dots, w_r\}$ is an orthogonal set, then

$$\boxed{\text{proj}_{\text{Span}\{w_1,\dots w_{i-1}\}}(v_i) = \text{proj}_{w_1}(v_i) + \text{proj}_{w_2}(v_i) + \cdots + \text{proj}_{w_{i-1}}(v_i).}$$

Gram-Schmidt Method: Example

Q: Let $S = \left\{ v_1 = \begin{pmatrix} 3 \\ 1 \\ -1 \\ 3 \end{pmatrix}, v_2 = \begin{pmatrix} -5 \\ 1 \\ 5 \\ -7 \end{pmatrix}, v_3 = \begin{pmatrix} 1 \\ 1 \\ -2 \\ 8 \end{pmatrix} \right\}$ and $W = \text{Span}(S)$. Find an orthonormal basis for W.

Exercise: First verify that $\{v_1, v_2, v_3\}$ are linearly independent. (Check that rank of $\begin{pmatrix} v_1 & v_2 & v_3 \end{pmatrix}$ is 3). Hence S is a basis of W. Use Gram-Schmidt method: $w_1 = v_1$, $w_2 = v_2 - \left(\dfrac{w_1^T v_2}{\|w_1\|^2}\right) w_1$

$$\Rightarrow \quad w_2 = v_2 - \left(\frac{-15+1-5-21}{9+1+1+9}\right) w_1 = v_2 - \left(\frac{-40}{20}\right) w_1 = v_2 + 2w_1 = \begin{pmatrix} 1 & 3 & 3 & -1 \end{pmatrix}^T.$$

Observe: v_1, $v_2 \in \text{Span}\{w_1, w_2\}$, w_1, $w_2 \in \text{Span}\{v_1, v_2\} \Rightarrow \text{Span}\{v_1, v_2\} = \text{Span}\{w_1, w_2\}$.

Recall $w_1 = \begin{pmatrix} 3 & 1 & -1 & 3 \end{pmatrix}^T$, $w_2 = \begin{pmatrix} 1 & 3 & 3 & -1 \end{pmatrix}^T$, and $v_3 = \begin{pmatrix} 1 & 1 & -2 & 8 \end{pmatrix}^T$. (Check $w_1^T w_2 = 0$).

Now $w_3 = v_3 - \left(\frac{w_1^T v_3}{\|w_1\|^2}\right) w_1 - \left(\frac{w_2^T v_3}{\|w_2\|^2}\right) w_2 = v_3 - \left(\frac{3+1+2+24}{20}\right) w_1 - \left(\frac{1+3-6-8}{20}\right) w_2$

$$\Rightarrow \quad w_3 = \begin{pmatrix} 1 \\ 1 \\ -2 \\ 8 \end{pmatrix} - \frac{3}{2}\begin{pmatrix} 3 \\ 1 \\ -1 \\ 3 \end{pmatrix} + \frac{1}{2}\begin{pmatrix} 1 \\ 3 \\ 3 \\ -1 \end{pmatrix} = \begin{pmatrix} -3 \\ 1 \\ 1 \\ 3 \end{pmatrix}.$$

Check $w_1^T w_3 = 0 = w_2^T w_3$ and $\text{Span}\{v_1, v_2, v_3\} = \text{Span}\{w_1, w_2, w_3\}$. Hence $\{w_1, w_2, w_3\}$ is an orthogonal basis of W. An orthonormal basis for W is $\left\{\frac{1}{\sqrt{20}} w_1, \frac{1}{\sqrt{20}} w_2, \frac{1}{\sqrt{20}} w_3\right\}$.

QR Factorization

Let $A = \begin{pmatrix} v_1 & \cdots & v_r \end{pmatrix}$ be an $n \times r$ matrix of rank r. Then $v_1, \ldots, v_r$ are linearly independent vectors in $\mathbb{R}^n$. By the Gram-Schmidt method, we get an orthonormal basis $\{q_1, \ldots, q_r\}$ of $C(A)$, where $q_i = \frac{w_i}{\|w_i\|}$ and $w_1 = v_1$, and for $k > 1$,

$$w_k = v_k - \left(\frac{w_1^T v_k}{\|w_1\|^2}\right) w_1 - \cdots - \left(\frac{w_{k-1}^T v_k}{\|w_{k-1}\|^2}\right) w_{k-1}.$$

Let $Q = \begin{pmatrix} q_1 & \ldots & q_r \end{pmatrix}$. $\boxed{\text{How are } A \text{ and } Q \text{ related?}}$

Note that $\text{Span}\{v_1, \ldots, v_k\} = \text{Span}\{w_1, \ldots, w_k\} = \text{Span}\{q_1, \ldots, q_k\}$ for all k. If $v_k = c_1 q_1 + \ldots + c_k q_k$, then $c_1 = q_1^T v_k, \quad c_2 = q_2^T v_k, \quad \ldots, \quad c_k = q_k^T v_k$. Thus
Hence $v_k = (q_1^T v_k) q_1 + \ldots + (q_k^T v_k) q_k$.

$$v_k = (q_1^T v_k) q_1 + \ldots + (q_k^T v_k) q_k \qquad \text{for each } k.$$

Therefore

$$\begin{pmatrix} v_1 & v_2 & \ldots & v_r \end{pmatrix} = \begin{pmatrix} q_1 & q_2 & \ldots & q_r \end{pmatrix} \begin{pmatrix} q_1^T v_1 & q_1^T v_2 & & q_1^T v_r \\ 0 & q_2^T v_2 & \ddots & q_2^T v_r \\ \vdots & \vdots & & \vdots \\ 0 & 0 & & q_r^T v_r \end{pmatrix}$$

i.e. $\boxed{A = QR}$, where the columns of Q form an orthonormal set and R is an invertible $r \times r$ matrix. **Question:** Why is R invertible?

This is called QR-factorization of A.

• If A is invertible $n \times n$, then $\boxed{A = QR}$, where Q is an orthogonal matrix and R is an invertible upper triangular matrix, both are $n \times n$ matrices.

Added remark: If A in least squares method has linearly independent columns, then QR factorization is useful in computing it.

Diagonalizing Symmetric Matrices: Example

Example: Let $A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$. Then $A - \lambda I = \begin{bmatrix} 1-\lambda & 1 & 1 \\ 1 & 1-\lambda & 1 \\ 1 & 1 & 1-\lambda \end{bmatrix}$ and

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

$(1-\lambda)[(1-\lambda)^2 - 1] - 1[1-\lambda-1] + 1[1-(1-\lambda)]$

$= (3-\lambda)\lambda^2$ Eigenvalues: $\lambda_1 = 3, \lambda_2 = 0, \lambda_3 = 0$.

To find $N(A - 3I)$, solve $A - 3I = \begin{bmatrix} -2 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$.

$N(A)$ is the plane $x + y + z = 0$. Hence, the associated eigenvectors are $v_1 = (1,1,1)^T$, $v_2 = (-1,0,1)^T$ and $v_3 = (0,-1,1)^T$.

Example: $A = Q\Lambda Q^T$

A has eigenvalues $\lambda_1 = 3, \lambda_2 = 0, \lambda_3 = 0$ with associated eigenvectors $v_1 = (1,1,1)^T$, $v_2 = (-1,0,1)^T$ and $v_3 = (0,-1,1)^T$. Note that v_2 and v_3 are linearly independent in $N(A)$. Observe $v_1^T v_2 = 0 = v_1^T v_3$.

How do we get an orthogonal Q such that $A = Q\Lambda Q^T$, where Λ is diagonal with entries 3, 0, 0 on the diagonal?

Steps: 1. Let $u_1 = v_1/\|v_1\|$.

2. Start with the basis $\{v_2, v_3\}$ of $N(A)$, and apply the Gram-Schimdt process to get an orthonormal basis $\{u_2, u_3\}$ for $N(A)$. Note that u_2 and u_3 are eigenvectors of A associated to $\lambda = 0$, and are linearly independent since they are non-zero orthogonal vectors.

3. Then $Q = [u_1 \;\; u_2 \;\; u_3]$ is orthogonal, and $Q^{-1}AQ = \Lambda$.

4. Since $Q^{-1} = Q^T$, $A = Q\Lambda Q^T$.

Diagonalizing Symmetric Matrices

Let A be a symmetric matrix, which is diagonalizable. Then there is an orthogonal matrix Q, and a diagonal matrix Λ such that $A = Q\Lambda Q^T$.

Observe: Eignevectors corresponding to distinct eigenvalues are orthogonal.

Proof. Let λ and μ be distinct eigenvalues of A with associated eigenvectors v and w repectively. Now, $\lambda(v^T w) = (\lambda v)^T w = (Av)^T w = v^T(A^T w) = v^T(Aw) = \mu(v^T w)$.

Since $\lambda \neq \mu$, this imples $v^T w = 0$, proving the result.

Step 1: Find the eigenvalues and the respective eigenvectors.

Step 2: Use Gram-Schmidt process to get an orthogonal basis for each eignespace.

Theorem: (Real Spectral Theorem)

Every symmetric matrix (with real entries) is diagonalizable, and hence decomposes as above.

Appendix 1: Determinants

Reading Slide - Determinants: Key Properties

Let A and B $n \times n$, and c a scalar.

- True/False: $\det(A+B) = \det(A) + \det(B)$.
- True/False: $\det(cA) = c\det(A)$.
- $\det(AB) = \det(A)\det(B)$.
- $\det(A) = \det(A^T)$.
- If A is orthogonal, i.e., $AA^T = I$, then $\det(A) =$
- If $A = [a_{ij}]$ is triangular, then $\det(A) =$ ___
- A is invertible $\Leftrightarrow$ $\det(A) \neq 0$.

  If this happens, then $\det(A^{-1}) =$ ___
- If $B = P^{-1}AP$ for an invertible matrix P,

  i.e., A and B are similar, then $\det(B) =$ ___
- If A is invertible, and $d_1, \ldots, d_n$ are the pivots of A,

  then $\det(A) =$ ___.

Reading Slide - Determinants: Defining Properties

Defn. The determinant function $\det : M_{n\times n}(\mathbb{R}) \to \mathbb{R}$ can be defined (uniquely) by its three basic properties.

- $\det(I) = 1$.
- The sign of determinant is reversed by a row exchange.

  Thus, if $B = P_{ij}A$, i.e., B is obtained from A by exchanging two rows, then $\det(B) = -\det(A)$. In particular, $\det(I) = 1 \Rightarrow \det(P_{ij}) = -1$.
- det is linear in each row separately, i.e. , we fix $n-1$ row vectors, say $v_2, \cdots, v_n$, then $\det\begin{pmatrix} - & v_2 & \cdots & v_n \end{pmatrix}^T : \mathbb{R}^n \to \mathbb{R}$ is a linear function.

I.e., for c, d in $\mathbb{R}$, and vectors u and v, if $A_{1*} = cu + dv$, we have

$$\det\begin{pmatrix} cu+dv & A_{2*} & \cdots & A_{n*} \end{pmatrix}^T = c\det\begin{pmatrix} u & A_{2*} & \cdots & A_{n*} \end{pmatrix}^T + d\det\begin{pmatrix} v & A_{2*} & \cdots & A_{n*} \end{pmatrix}^T.$$

There are n such equations (for n choices of rows).

Reading Slide - Determinants: Induced Properties

1. If two rows of A are equal, then $\det(A) = 0$.

   Proof. Suppose i-th and j-th rows of A are equal, i.e., $A_{i*} = A_{j*}$, then $A = P_{ij}A$.

   Hence $\det(A) = \det(P_{ij}A) = -\det(A) \Rightarrow \boxed{\det(A) = 0}$.

2. If B is obtained from A by $\boxed{R_i \mapsto R_i + aR_j,}$ then $\det(B) = \det(A)$.

3. If A is $n \times n$, and its row echelon form U is obtained without row exchanges, then $\det(U) = \det(A)$.

   Q: What happens if there are row exchanges? Exercise!

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

   Proof: Let the ith row of A be zero, i.e., $A_{i*} = 0$.

   Let B be obtained from A by $\boxed{R_i = R_i + R_j}$, i.e., $B = E_{ij}(1)A$. Then $B_{i*} = B_{j*}$.

   Exercise: Complete the proof.

Reading Slide - Determinants: Special Matrices

5. If $A = \begin{pmatrix} a_1 & & \\ & \ddots & \\ & & a_n \end{pmatrix}$ is diagonal, then $\det(A) = a_1 \cdots a_n$. (Use linearity).

6. If $A = (a_{ij})$ is triangular, then $\det(A) = a_{11} \dots a_{nn}$.

   Proof. If all a_{ii} are non-zero, then by elementary row operations, A reduces to the diagonal matrix $\begin{pmatrix} a_{11} & & \\ & \ddots & \\ & & a_{nn} \end{pmatrix}$ whose determinant is $a_{11} \cdots a_{nn}$.

   If at least one diagonal entry is zero, then elimination will produce a zero row $\Rightarrow \det(A) = 0$.

Reading Slide - Formula for Determinant: 2×2 case

Write $(a, b) = (a, 0) + (0, b)$, the sum of vectors in coordinate directions. Similarly write $(c, d) = (c, 0) + (0, d)$. By linearity,

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = \begin{vmatrix} a & 0 \\ c & d \end{vmatrix} + \begin{vmatrix} 0 & b \\ c & d \end{vmatrix} = \begin{vmatrix} a & 0 \\ c & 0 \end{vmatrix} + \begin{vmatrix} a & 0 \\ 0 & d \end{vmatrix} + \begin{vmatrix} 0 & b \\ c & 0 \end{vmatrix} + \begin{vmatrix} 0 & b \\ 0 & d \end{vmatrix}.$$

For an $n \times n$ matrix, each row splits into n coordinate directions, so the expansion of $\det(A)$ has n^n terms.

However, when two rows are in same coordinate direction, that term will be zero, e.g.,

$$\begin{vmatrix} a & 0 \\ c & 0 \end{vmatrix} = 0, \quad \begin{vmatrix} 0 & b \\ 0 & d \end{vmatrix} = 0, \quad \begin{vmatrix} 0 & b \\ c & 0 \end{vmatrix} = -\begin{vmatrix} c & 0 \\ 0 & b \end{vmatrix} = -bc \Rightarrow \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$$

The non-zero terms have to come in different columns. So, there will be $n!$ such terms in the $n \times n$ case.

Reading Slide - Formula for Determinant: $n \times n$ case
For $n \times n$ matrix $A = (a_{ij})$,

$$\det(A) = \sum_{\text{all permutations } P} (a_{1\alpha_1} \ldots a_{n\alpha_n}) \det(P).$$

The sum is over $n!$ permutations of numbers $(1, \ldots, n)$. Here a permutation $(i_1, i_2, \ldots, i_n)$ of $(1, 2, \ldots, n)$ corresponds to the product of permutation matrices $P = \begin{bmatrix} e_{i_1}^T \\ \vdots \\ e_{i_n}^T \end{bmatrix}$. Then $\det(P) = +1$ if the number of row exchanges in P needed to get I is even, and -1 if it is odd.

Reading Slide - Cofactors: 3×3 Case

$$\begin{aligned}
\det(A) &= \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \\
&= a_{11}\, a_{22}\, a_{33}\, (1) + a_{11}\, a_{23}\, a_{32}\, (-1) + a_{12}\, a_{21}\, a_{33}\, (-1) \\
&\quad + a_{12}\, a_{23}\, a_{31}\, (1) + a_{13}\, a_{21}\, a_{32}\, (1) + a_{13}\, a_{22}\, a_{31}\, (-1) \\
&= a_{11}(a_{22}a_{33} - a_{23}a_{32}) - a_{12}(a_{21}a_{33} - a_{23}a_{31}) \\
&\quad + a_{13}(a_{21}a_{32} - a_{22}a_{31}) \\
&= a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} + a_{12}(-1) \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} \\
&= a_{11}C_{11} + a_{12}C_{12} + a_{13}C_{13} \text{ where,}
\end{aligned}$$

$$C_{11} = \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix}, \quad C_{12} = (-1) \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix}, \quad C_{13} = \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}.$$

Reading Slide - Cofactors: $n \times n$ Case
Let C_{1j} be the coefficient of a_{1j} in the expansion

$$\det(A) = \sum_{\text{all permutations } P} (a_{1\alpha_1} \ldots a_{n\alpha_n}) \det(P)$$

Then $\boxed{\det(A) = a_{11}C_{11} + a_{12}C_{12} + \ldots + a_{1n}C_{1n}}$ where,

$$\begin{aligned}
C_{1j} &= \Sigma a_{2\beta_2} \ldots a_{n\beta_n} \det(P) \\
&= (-1)^{1+j} \det \begin{bmatrix} a_{21} & . & a_{2(j-1)} & a_{2(j+1)} & . & a_{2n} \\ \vdots & . & \vdots & . & \vdots & \\ a_{n1} & . & a_{n(j-1)} & a_{j+1} & . & a_{nn} \end{bmatrix} \\
&= (-1)^{1+j} \det(M_{1j}),
\end{aligned}$$

where M_{1j} is obtained from A by deleting the 1^{st} row and j^{th} column.

Extra Reading Slides - Determinants

The following set of slides contain some extra reading material on determinants for interested students of MA110 (Spring 2024)/ MA106 (Spring 2024).

$\mathbf{det}(AB) = \mathbf{det}(A)\,\mathbf{det}(B)$ (Proof)

7. $\boxed{\det(AB) = \det(A)\det(B)}$

   Proof. We may assume that B are invertible. Else, $\text{rank}(AB) \leq \text{rank}B \neq n \Rightarrow \text{rank}(AB) \neq n \Rightarrow AB$ is not invertible.

   $\boxed{\text{Hint:}}$ For fixed B, show that the function d defined by

   $$d(A) = \det(AB)/\det(B)$$

   satisfies the following properties

   (a) $d(I) = 1$.

   (b) If we interchange two rows of A, then d changes its sign.

   (c) d is a linear function in each row of A.

   Then d is the unique determinant function det and $\det(AB) = \det(A)\det(B)$. □

Determinants of Transposes (Proof)

8. $\boxed{\det(A) = \det(A^T)}$

   Proof. With U, L, and P, as usual write $PA = LU \Rightarrow A^T P^T = U^T L^T$ Since U and L are triangular, we get $\det(U) = \det(U^T)$ and $\det(L) = \det(L^T)$.

   Since $PP^T = I$ and $\det(P) = \pm 1$, we get $\det(P) = \det(P^T)$.

   Thus $\det(A) = \det(A^T)$. □

Determinants and Invertibility (Proof)

9. A is invertible if and only if $\det(A) \neq 0$.

   By elimination, we get an upper triangular matrix U, a lower triangular matrix L with diagonal entries 1, and a permutation matrix P, such that $PA = LU$.

   Observation 1: If A is singular, then $\det(A) = 0$.

This is because elimination produces a zero row in U and hence $\det(A) = \pm\det(U) = 0$.

Observation 2: If A is invertible, then $\det(A) \neq 0$.

This is because elimination produces n pivots, say $d_1, \dots, d_n$, which are non-zero. Then U is upper triangular, with diagonal entries $d_1, \dots, d_n \Rightarrow \det(A) = \pm\det(U) = \pm\, d_1 \cdots d_n \neq 0$.

Thus we have: A invertible $\Rightarrow \det(A) = \pm$(product of pivots).

Exercise: If AB is invertible, then so are A and B.

Exercise: A is invertible if and only if A^T is invertible.

Determinant: Geometric Interpretation (2×2)

INVERTIBILITY: Very often we are interested in knowing when a matrix is invertible. Consider $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. Then A is invertible if and only if A has full rank.

If a, c both are zero then clearly $\operatorname{rank}(A) < 2 \Rightarrow A$ is not invertible. Assume $a \neq 0$, else, interchange rows. The row operations $\begin{bmatrix} a & b \\ c & d \end{bmatrix} \xrightarrow{R_2 - c/aR_1} \begin{bmatrix} a & b \\ 0 & d - cb/a \end{bmatrix}$ show that A is invertible if and only if $d - cb/a \neq 0$, i.e., $ad - bc \neq 0$.

AREA: The area of the parallelogram with sides as vectors $v = (a, b)$ and $w = (c, d)$ is equal to $ad - bc$. Thus,

$\boxed{\text{A } 2 \times 2 \text{ matrix } A \text{ is singular } \Leftrightarrow}$

$\boxed{\text{its columns are on the same line}}$

$\boxed{\Leftrightarrow \text{ the area is zero.}}$

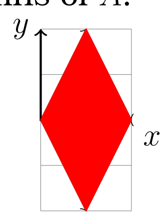
Determinant: Geometric Interpretation

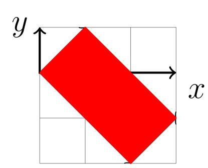
- Test for invertibility: An $n \times n$ matrix A is invertible $\Leftrightarrow \det(A) \neq 0$.
- n-dimensional volume: If A is $n \times n$, then $|\det(A)| =$ the volume of the box (in n-dimensional space $\mathbb{R}^n$) with edges as rows of A.

Examples: (1) The volume (area) of a line in $\mathbb{R}^2 = 0$.

(2) The determinant of the matrix $A = \begin{pmatrix} 1 & 2 \\ 1 & -2 \end{pmatrix}$ is $\boxed{-4}$.

(3) Let's compute the volume of the box (parallelogram) with edges as rows of A or columns of A.





$\boxed{= 4}$

Expansion along the i-th row (Proof)

If C_{ij} is the coefficient of a_{ij} in the formula of $\det(A)$, then

$\boxed{\det(A) = a_{i1}\,C_{i1} + \ldots + a_{in}\,C_{in}}$, where C_{ij} is determined as follows:

By $i-1$ row exchanges on A, get the matrix $B = \begin{pmatrix} A_{i*} & A_{1*} & .. & A_{(i-1)*} & A_{(i+1)*} & .. & A_{n*} \end{pmatrix}^T$

Since $\det(A) = (-1)^{i-1}\det(B)$, we get

$$C_{ij}(A) = (-1)^{i-1}C_{1j}(B) = (-1)^{i-1}(-1)^{j-1}\det(M)$$

where M is obtained from B by deleting 1^{st} row and j^{th} column. Here M is obtained from B by deleting its first row, and j-th column, and hence from A by deleting i-th row and j-th column. Write M as M_{ij}. Then $\boxed{C_{ij} = (-1)^{i+j}\det(M_{ij})}$

Expansion along the j-th column (Proof)

Note that $C_{ij}(A^T) = C_{ji}(A)$.

Hence, if we write $A^T = (b_{ij})$, then

$$\begin{aligned}\det(A) &= \det(A^T)\\ &= b_{j1}\,C_{j1}(A^T) + \ldots + b_{jn}\,C_{jn}(A^T)\\ &= a_{1j}\,C_{1j}(A) + \ldots + a_{nj}\,C_{nj}(A)\end{aligned}$$

This is the expansion of $\det(A)$ along j-th column of A.

Applications: 1. Computing A^{-1}

If $C = (C_{ij})$: cofactor matrix of A, then $\boxed{A\,C^T = \det(A)\,I}$

i.e., $\begin{bmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \ldots & a_{nn} \end{bmatrix} \begin{bmatrix} C_{11} & \ldots & C_{n1} \\ \vdots & \ddots & \vdots \\ C_{1n} & \ldots & C_{nn} \end{bmatrix} = \begin{bmatrix} \det(A) & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & \det(A) \end{bmatrix}$

Proof. We have seen that $a_{i1}C_{i1} + \ldots + a_{in}C_{in} = \det(A)$. Now $a_{11}C_{21} + a_{12}C_{22} + \ldots + a_{1n}C_{2n} =$ $\det \begin{bmatrix} a_{11} & \ldots & a_{1n} \\ a_{11} & \ldots & a_{1n} \\ a_{31} & \ldots & a_{3n} \\ . & & . \\ a_{n1} & \ldots & a_{nn} \end{bmatrix} = 0$. Similarly, if $i \neq j$, then $a_{i1}C_{j1} + a_{i2}C_{j2} + \ldots + a_{in}C_{jn} = 0$. $\square$

Remark. If A is invertible, then $A^{-1} = \dfrac{1}{\det(A)}C^T$.

For $n \geq 4$, this is *not* a good formula to find A^{-1}.
Use elimination to find A^{-1} for $n \geq 4$.
This formula is of theoretical importance.

Applications: 2. Solving $Ax = b$

Cramer's rule: If A is invertible, the $Ax = b$ has a unique solution.

$$x = A^{-1}b = \frac{1}{\det(A)} C^T b = \frac{1}{\det(A)} \begin{bmatrix} C_{11} & \cdots & C_{n1} \\ \vdots & & \vdots \\ C_{1n} & \cdots & C_{nn} \end{bmatrix} \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}$$

Hence $x_j = \dfrac{1}{\det(A)}(b_1 C_{1j} + b_2 C_{2j} + \cdots + b_n C_{nj}) = \dfrac{1}{\det(A)}\det(B_j)$,

where B_j is obtained by replacing j^{th} column of A by b, and $\det(B_j)$ is computed along the j^{th} column.

Remark: For $n \geq 4$, use elimination to solve $Ax = b$.

Cramer's rule is of theoretical importance.

Applications: 3. Volume of a box

$\boxed{\text{Assume the rows of } A \text{ are mutually orthogonal}}$. Then

$$AA^T = \begin{pmatrix} A_{1*} \\ \vdots \\ A_{n*} \end{pmatrix} \begin{pmatrix} (A_{1*})^T & \ldots & (A_{n*})^T \end{pmatrix} = \begin{pmatrix} l_1^2 & & 0 \\ & \ddots & \\ 0 & & l_n^2 \end{pmatrix}$$

where $l_i = \sqrt{(A^i)^T \cdot A^i}$ is the length of A^i. Since $\det(A) = \det(A^T)$,

we get $\boxed{|\det(A)| = l_1 \cdots l_n}$.

Since the edges of the box spanned by rows of A are at right angles, the volume of the box

$=$ the product of lengths of edges

$= |\det(A)|$.

Applications: 4. A Formula for Pivots

Observation: If row exchanges are not required, then the first k pivots are determined by the top-left $k \times k$ submatrices $\widetilde{A}_k$ of A.

Example. If $A = [a_{ij}]_{3\times 3}$, then $\widetilde{A}_1 = (a_{11})$, $\widetilde{A}_2 = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ and $\widetilde{A}_3 = A$.

Assume the pivots are $d_1, \ldots, d_n$, obtained without row exchange. Then

- $\det(\widetilde{A}_1) = a_{11} = d_1$
- $\det(\widetilde{A}_2) = d_1\, d_2 = \det(A_1)\, d_2$
- $\det(\widetilde{A}_3) = d_1\, d_2\, d_3 = \det(A_2)\, d_3$ *etc.*,
- If $\det(\widetilde{A}_k) = 0$, then we need a row exchange in elimination.
- Otherwise the k-th pivot is $\boxed{d_k = \det(\widetilde{A}_k)/\det(\widetilde{A}_{k-1})}$