Linear Algebra

Engineering Mathematics In Action

Course Designed by Prof. Amrik Sen
Course Conducted by Prof. Amrik Sen and Prof. Sanjay K. Bose
Plaksha University

FIELD (Definition): A field is a set $\mathbb{F}$ of numbers with the property that if $a, b \in \mathbb{F}$, then $a + b$, $a - b$, ab and $\frac{a}{b}$ are also in $\mathbb{F}$ (assuming, of course, that $b \neq 0$ *in the expression* $\frac{a}{b}$.

e.g. $\mathbb{Q}, \mathbb{R}$ and $\mathbb{C}$ are fields of numbers

$\mathbb{N}$ and $\mathbb{Z}$ are **not** fields of numbers!

$\mathbb{Q}$ - Rational Numbers
$\mathbb{R}$ - Real Numbers
$\mathbb{C}$ - Complex Numbers

$\mathbb{N}$: Natural Numbers (positive integers)
$\mathbb{Z}$: Integers

VECTOR SPACES (Definition): A vector space, $\boldsymbol{\mathcal{V}}$ consists of a set $\mathbb{V}$ of vectors, a field $\mathbb{F}$ of scalars, and **two** operations:

i. **Vector Addition:** *if* $v, w \in \mathbb{V}$, *then* $v + w \in \mathbb{V}$

ii. **Scalar Multiplication:** $c \in \mathbb{F}$ *and* $v \in \mathbb{V}$ *produces a new vector* $cv \in \mathbb{V}$

These scalars and vectors also satisfy the following **axioms**

i. ***Associativity** of addition:* $(v + u) + w = v + (u + w) \;\; \forall v, u, w \in \mathbb{V}$
ii. ***Associativity** of multiplication:* $(ab)u = a(bu)$, *for any* $a, b \in \mathbb{F}, u \in \mathbb{V}$
iii. ***Distributivity:*** $(a + b)u = au + bu \;\; and \;\; a(u + v) = au + av$
$$\forall\, a, b \in \mathbb{F},\; u \in \mathbb{V},\; v \in \mathbb{V}$$
iv. ***Unitarity**:* $1u = u \;\; \forall u \in \mathbb{V}$
v. ***Existence of zero**:* $\exists 0 \in \mathbb{V} \;\; s.t. \;\; u + 0 = u \quad \forall u \in \mathbb{V}$
vi. ***Negation**:* *For every* $u \in \mathbb{V}, \exists(-u) \in \mathbb{V} \;\; s.t. \;\; u + (-u) = 0 \in \mathbb{V}$

VECTOR SPACES (Examples):

1) Let $\mathbb{V}$ be the set of $n \times 1$ column matrices (vectors), $\mathbb{F}$ be the field of *reals* $\mathbb{R}$, and the laws of vector addition and scalar multiplication are defined as:

$$\begin{pmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \\ \cdot \\ \cdot \\ y_n \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \cdot \\ \cdot \\ x_n + y_n \end{pmatrix} \quad \text{and} \quad c \begin{pmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ x_n \end{pmatrix} = \begin{pmatrix} cx_1 \\ cx_2 \\ \cdot \\ \cdot \\ cx_n \end{pmatrix}.$$

HW: Verify that the above indeed constitutes a vector space!
(Check that the axioms are satisfied.)

VECTOR SPACES (Examples):

2) Let $\mathbb{V}$ be the set of all continuous functions $f: \mathbb{R} \to \mathbb{R}$, let the field of scalars be $\mathbb{R}$, and let the operations be as usually defined.

HW: Verify that the above indeed constitutes a vector space!

3) Let $\mathbb{V}$ be the set of all continuous functions $f: \mathbb{R} \to \mathbb{R}$ that satisfy the equation $f'' = -f$. (Can you think of any function that satisfies this property? Cosine, Sine?)

Let the field of scalars be $\mathbb{R}$. The operations are defined in the usual manner. *Hint: Suppose $f_1, f_2 \in \mathbb{V}, c \in \mathbb{R}$; then $(f_1 + f_2)'' = f_1'' + f_2'' = -f_1 - f_2 = -(f_1 + f_2)$; and $(cf_1)'' = cf_1'' = c(-f_1) = -(cf_1)$.*
Are these results consistent with the definition of the vector space?
Also check whether all axioms are compliant?

LINEAR INDEPENDENCE OF VECTORS

Definition (Linearly_dependent_vectors):

Let $\mathcal{V}$ be a vector space and $\mathcal{X} \subset \mathcal{V}$ be a non-empty subset. Then $\mathcal{X}$ is ***linearly dependent*** if there are distinct vectors $v_1, v_2, \ldots, v_k \in \mathcal{X}, and$ $scalars$ $c_1, c_2, \ldots, c_k$ ***(not all of them zero)***, *s.t.* $c_1 v_1 + c_2 v_2 + \ldots + c_k v_k = 0$.

This is equivalent to saying that *at least one of the vectors* v_i *can be expressed as a linear combination of the others, i.e.* $v_i = \sum_{j \neq i} -(\frac{c_j}{c_i}) v_j$

Definition (Linearly independent vectors):

A subset which is not linearly dependent is said to be ***linearly independent***. Thus a set of distinct vectors $\{v_1, v_2, \ldots, v_k\}$ is linearly independent if and only if an equation of the form $c_1 v_1 + c_2 v_2 + \ldots + c_k v_k = 0$ always implies that $c_1 = c_2 = \ldots = c_k = 0$.

Geometrical Interpretation of Linear Dependence

Let V_1, V_2, V_3 be the vectors in 3D-Euclidean space $\mathbb{R}^3$ with a common origin. If these vectors form a *linearly dependent* set, then one of them, say V_1, can be expressed as a linear combination of the other two: $V_1 = aV_2 + bV_3$. This implies, by the parallelogram law, that the three vectors are **co-planar.**

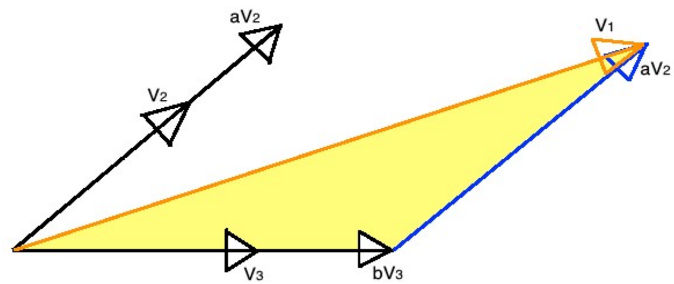


Fig. 1: Linear dependence of vectors is equivalent to coplanar geometry

In fact, **linearly dependent set of vectors with common origin ⟺ co-planar.**

Can you think of a similar interpretation of vectors in $\mathbb{R}^2$?

Consider what happens when we have three vectors **A**, **B** and **C**, from a common origin, in a 2-dimensional vector space, where –

A = $\begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$ **B** = $\begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$ and **C** = $\begin{pmatrix} c_1 \\ c_2 \end{pmatrix}$

Can I represent the vector **C** as a linear combination of the vectors **A** and **B** such as **C** = α**A** + β**B** ?

Yes, if we choose α and β as $\alpha = \dfrac{c_1 b_2 - c_2 b_1}{a_1 b_2 - a_2 b_1}$ and $\beta = \dfrac{c_2 a_1 - c_1 a_2}{a_1 b_2 - a_2 b_1}$

Looking at these, we can immediately conclude that this cannot be done if A and B are collinear, because then $\dfrac{a_1}{a_2} = \dfrac{b_1}{b_2}$

What can you conclude when either α or β or both become zero?

Example

Show that these vectors are linearly dependent in $\mathbb{R}^2$

$$\mathbf{A} = \begin{pmatrix} -1 \\ 2 \end{pmatrix} \quad \mathbf{B} = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \quad \text{and} \quad \mathbf{C} = \begin{pmatrix} 2 \\ -4 \end{pmatrix}$$

We choose scalars c_1, c_2, c_3 such that -

$$c_1 \begin{pmatrix} -1 \\ 2 \end{pmatrix} + c_2 \begin{pmatrix} 1 \\ 2 \end{pmatrix} + c_3 \begin{pmatrix} 2 \\ -4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

This gives

$$\begin{aligned} -c_1 + c_2 + 2c_3 &= 0 \\ 2c_1 + 2c_2 - 4c_3 &= 0 \end{aligned}$$

Since the number of unknowns is more than the number of equations, there will be a non-trivial solution

Therefore, the vectors are Linearly Dependent

Example

Are the polynomials $x+1$, $x+2$, x^2-1 linearly independent in the vector space $P_3(\mathbb{R})$?

Notation: $P_3(\mathbb{R})$ is the set of polynomials of less than degree 3 with real coefficients

We choose scalars c_1, c_2, c_3 such that

$$c_1(x+1)+c_2(x+2)+c_3(x^2-1)=0$$

This gives

$$\begin{aligned} c_1 + 2c_2 - c_3 &= 0 \\ c_1 + c_2 &= 0 \\ c_3 &= 0 \end{aligned}$$

Clearly, this can only have the trivial solution $c_1 = c_2 = c_3 = 0$

Therefore, the polynomials are Linearly Independent

Example

In the vector space $V= P(\mathbb{R})$, consider the subset $S = \{x-1, x^2+1, x^3-x^2-x+3\}$. Is S linearly dependent or linearly independent?

Consider $$a_1(x-1)+a_2(x^2+1)+a_3(x^3-x^2-x+3)=0$$

Equating the coefficients of the powers of x to zero for each term in the LHS, we get -

$$\begin{aligned} -a_1 + a_2 + 3a_3 &= 0 \\ a_1 \qquad\quad - a_3 &= 0 \\ a_2 - a_3 &= 0 \\ a_3 &= 0 \end{aligned}$$

The only solution to this linear homogenous system is the trivial solution, so the vectors in the subset S are **linearly independent**

BASIS OF A VECTOR SPACE (Definition)

Let $\mathbb{X}$ be a non-empty subset of a vector space $\mathcal{V}$. Then $\mathbb{X}$ is called a *basis* of $\mathcal{V}$ if **both** the following are true:

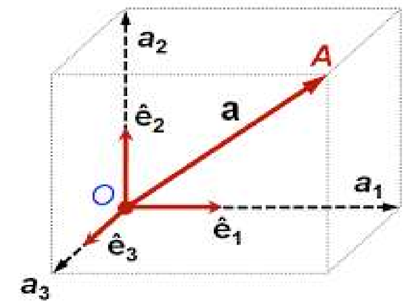
i. $\mathbb{X}$ is linearly independent cannot generate an element of $\mathbb{X}$ as linear combination of the other elements of $\mathbb{X}$

ii. $\mathbb{X}$ generates $\mathcal{V}$ (i.e. $\mathbb{X}$ *spans* $\mathcal{V}$) any element of $\mathcal{V}$ can be generated as a linear combination of the elements of $\mathbb{X}$

What is the meaning of "*spans*"?

Technically, it means that every element (vector) in the space $\mathcal{V}$ can be expressed as a linear combination of the elements of the set $\mathbb{X}$.

Examples of Bases



1. Basis of $\mathbb{R}^n$: $\boldsymbol{e_1} = \begin{pmatrix} 1 \\ 0 \\ . \\ . \\ 0 \end{pmatrix}$, $\boldsymbol{e_2} = \begin{pmatrix} 0 \\ 1 \\ . \\ . \\ 0 \end{pmatrix}$, $\ldots$, $\boldsymbol{e_n} = \begin{pmatrix} 0 \\ 0 \\ . \\ . \\ 1 \end{pmatrix}$ form a basis of $\mathbb{R}^n$ because (i) they are

linearly independent (by inspection), and (ii) they *span* $\mathbb{R}^n$ because $c_1\boldsymbol{e_1} + c_2\boldsymbol{e_2} + \cdots c_n\boldsymbol{e_n} = \begin{pmatrix} c_1 \\ c_2 \\ . \\ . \\ c_n \end{pmatrix}$

generates any vector in $\mathbb{R}^n$ depending on the values of $c_i \forall i = 1,2,\ldots,n$

Examples of Bases (continued):

2. Let $\mathbb{P}_n$ be a vector space of all polynomial functions of degree n or less. The basis of $\mathbb{P}_n$ is $\{1, x, x^2, \ldots, x^n\}$, the set of monomials.
(This is not a unique basis set because $\{p_0(x), p_1(x), \ldots, p_n(x)\}$ also forms a basis where $p_i(x)$ is a polynomial in $\mathbb{P}_n$ of degree i.)

3. Let $\mathbb{M}_{m\times n}(\mathbb{F})$ denote the set of $m \times n$ matrices with entries in $\mathbb{F}$. Then $\mathbb{M}_{m\times n}(\mathbb{F})$ is a vector space over $\mathbb{F}$. Vector addition is just matrix addition and scalar multiplication is defined in the obvious way (by multiplying each entry of the matrix by the same scalar). The zero vector is just the zero matrix. One possible choice of basis is the matrices with a single entry equal to 1 and all other entries 0.
(We will study the vector space of matrices in more detail in subsequent lectures!)

PROPERTIES OF BASES:

1. Must every vector space have a basis?

*Ans: Every **non-zero, finitely generated** vector space has a basis!*

2. Does a vector space have a unique basis?

Ans: Usually a vector space will have many bases. e.g., the vector space $\mathbb{R}^2$ has the basis $\{\begin{pmatrix} 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 2 \\ -1 \end{pmatrix}\}$ as well as the standard basis $\{\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}\}$.

3. What is the dimension of a vector space?

Ans: $dim(\mathcal{V}) = no.\ of\ elements\ (vectors)\ in\ the\ basis\ (basis\ set)$.

Can you think of a vector space whose dimension is infinite?

A Few Other Things –

Finitely Generated Vector Space: One where you only need a finite number of elements to generate the vector space using linear combinations, e.g. $\mathbb{R}^2$ needs only (0,1) & (1,0) to generate all vectors in $\mathbb{R}^2$

Infinite Dimensional Vector Space (example): Let *P* be the vector space of all polynomials in *X* with rational coefficients. *P* is infinite dimensional. To see this – If *P* is given by the span of *k* polynomials in *P*, $p_1 \dots p_k$ where *m* is the maximum of the degrees of $p_1 \dots p_k$. Then x^{m+1} is a vector which cannot be written as a combination of $p_1 \dots p_k$. This is a contradiction so *P* cannot be finite dimensional.

VECTOR SPACES: *We will study more about the* ***vector space of*** $\boldsymbol{m \times n}$ ***matrices*** *over the reals and their mathematical utilities!*
In fact much of this course is a study about matrices and their applications in engineering.

Vector space of $m \times n$ matrices over the reals, $\mathbb{M}_{m \times n}(\mathbb{R})$ and associated vector spaces

1. So what are matrices?

Ans: *Matrices are convenient arrangement of numbers in rows and columns lending a compact structure that are amenable to mathematical laws (laws or rules of matrix algebra) that are a consequence of matrix operations like addition, multiplication, transpose, inverse, etc:*

A+B = B+A Commutative Law of Addition (A+B)+C = A+(B+C) Associative Law of Addition
A+0 = A (AB)C = A(BC) Associative Law of Multiplication
AI = A = IA I: Identity Matrix A(B+C) = AB+AC (A+B)C = AC+BC Distributive Laws
A-B = A+(-1)B (cd)A = c(dA) c(AB) = (cA)B = A(cB)
c(A+B) = cA+cB (c+d)A = cA+dA
$(A+B)^T = A^T + B^T$
$(AB)^T = B^TA^T$ *note the change in order of A and B*

Here, $A, B, C \in \mathbb{M}_{m \times n}(\mathbb{R}), a_{ij} \in \mathbb{R}, c, d \in \mathbb{F}$ *where* $\mathbb{F} \equiv \mathbb{R}$ *in our discussion in this chapter.*

We can generalize a matrix as consisting of appropriately defined sub-matrices (does not have to be just numbers)!

How do you take the transpose of a matrix?

Taking the transpose of a matrix $A = \begin{pmatrix} a_{11} & a_{12} & \cdot & \cdot & \cdot & a_{1n} \\ a_{21} & a_{22} & \cdot & \cdot & \cdot & a_{2n} \\ \cdot & \cdot & \cdot & & & \cdot \\ \cdot & \cdot & & \cdot & & \cdot \\ \cdot & \cdot & & & \cdot & \cdot \\ a_{m1} & a_{m2} & \cdot & \cdot & \cdot & a_{mn} \end{pmatrix}$

m rows and n columns

gives $A^T = \begin{pmatrix} a_{11} & a_{21} & \cdot & \cdot & \cdot & a_{m1} \\ a_{12} & a_{22} & \cdot & \cdot & \cdot & a_{m2} \\ \cdot & \cdot & \cdot & & & \cdot \\ \cdot & \cdot & & \cdot & & \cdot \\ \cdot & \cdot & & & \cdot & \cdot \\ a_{1n} & a_{2n} & \cdot & \cdot & \cdot & a_{mn} \end{pmatrix}$

n rows and m columns

2. The rules of matrix algebra guarantee that $\mathbb{M}_{m\times n}(\mathbb{R})$ is a vector space!

3. Special cases: i) When $n = 1$ in $\mathbb{M}_{m\times n}(\mathbb{R})$, we recover the familiar Euclidean space $\mathbb{R}^m$,
and ii) When $m = 1$, we recover the vector space $\mathbb{R}_n$ of all real n –row vectors.

4. A practical application of matrices: A system of linear equations can be expressed in matrix form and the entire mathematical machinery of matrices can be unleashed to find and analyze the solution(s) of the said system of linear equations.

There is a seamless hierarchy of what are known as ***TENSORS*** *in mathematical parlance, the most simplest tensor being scalars (tensors of rank 0), the next in the hierarchy are vectors (tensors of rank 1), followed by matrices (tensors of rank 2), etc. It must be noted that not all matrices are tensors, but all tensors of rank 2 are definitely matrices!*

Example 4.1: Consider the two set of linear equations –

$$\begin{aligned} 2x - y &= 0 \\ -x + 2y &= 3 \end{aligned} \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{(i)}$$

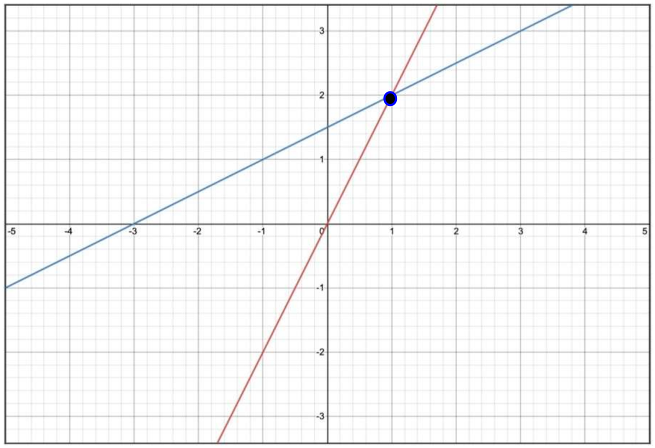
We begin by sketching out the respective straight lines

$$\boldsymbol{2x - y = 0} \text{ and } \boldsymbol{-x + 2y = 3}.$$

The solution of this system is the *intersection point* of these two straight lines, which is $x = 1, y = 2$.

If we express this system of linear equations in matrix-vector notation, then we have $A\boldsymbol{x} = \boldsymbol{b}$, where $A = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$ is the *coefficient matrix*, $\boldsymbol{x} = \begin{pmatrix} x \\ y \end{pmatrix}$, and $\boldsymbol{b} = \begin{pmatrix} 0 \\ 3 \end{pmatrix}$.

The system (i) corresponds to the *row picture*, and the solution strategy presented above lends a geometrical interpretation of this *row picture*.



Points which follow 2x-y=0 lie on the red line. Points which follow –x+2y=3 lie on the blue line.
Our solution is the point which lies on both the lines, i.e. their intersection point.
Now think of what happens when there are -
(a) No intersection between the two lines (except at infinity) or
(b) Infinitely many such points.

There is an alternative (*and sometimes more useful*) geometrical picture, the *column picture*, which lends a different interpretation of the situation at hand.

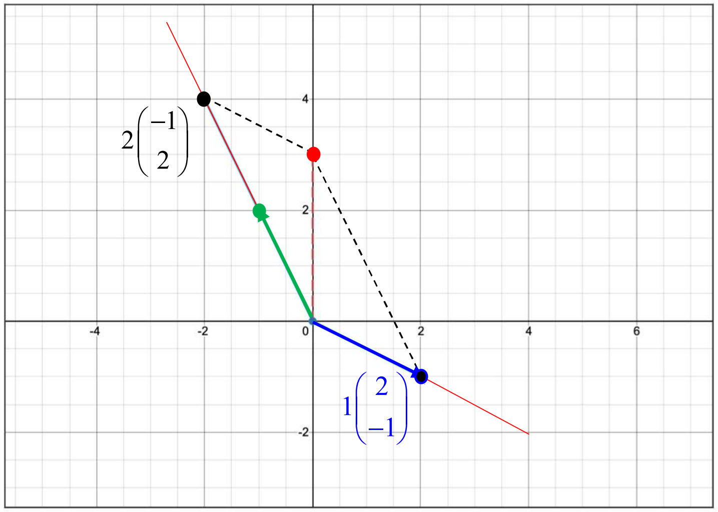
By careful inspection, we notice that the earlier system of equations can be re-written as follows:

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

Here the system of equations is expressed in terms of a **linear combination** of the column vectors of A.

If we treat the columns as vectors in 2D Euclidean space, and consider the correct solutions (say we somehow know that $x = 1 \text{ and } y = 2$), then we have the following –

*"The resultant of adding the two vectors (properly scaled, i.e. weighted) on the left-hand side is identically equal to the vector **b** on the right-hand side."*



Suggest doing this by taking the projection on one vector, along the direction of the other

Throughout our study about linear algebra, the idea of *linear combination* and *column vectors of a matrix* will play a very important role in terms of the mathematical machinery as well as interpretation of the physical picture.

Of course, the question arises whether we can always solve this system as follows:

$$\boldsymbol{x} = A^{-1}\boldsymbol{b} \qquad \text{for any given 2D vector } \boldsymbol{b}?$$

This will be possible only if A is invertible, i.e., A^{-1} exists! In that case, the columns of A will span the entire 2D Euclidean plane (and $\boldsymbol{x} = A^{-1}\boldsymbol{b}$ will be the solution for any 2D vector $\boldsymbol{b}$), i.e. there will be one combination of the columns which will give $\boldsymbol{b}$.

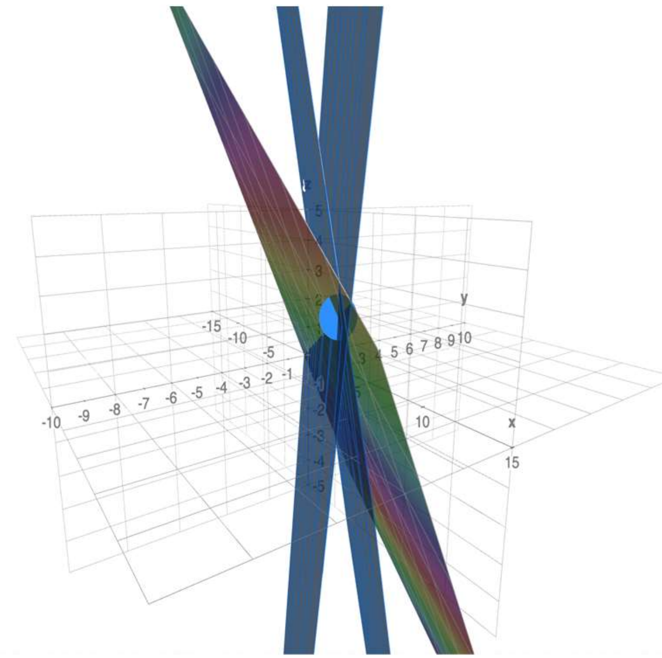
This should the lead us to ask when is A invertible? i.e., ***when will the columns of A span the entire 2D Euclidean plane?***
The answer should be obvious by inspecting the above 2D graph: ***whenever the columns of A are linearly independent.***

Note that *spanning a plane* here is analogous to *obtaining any vector in the plane by a linear combination of vectors.*

Example 4.2: Consider the following system of linear equations-

$$\begin{aligned} 2x + 8y + 4z &= 2 \\ 2x + 5y + z &= 5 \\ 4x + 10y - z &= 1 \end{aligned}$$

(ii)



This system can be expressed in matrix form as: $A\boldsymbol{x} = \boldsymbol{b}$ where

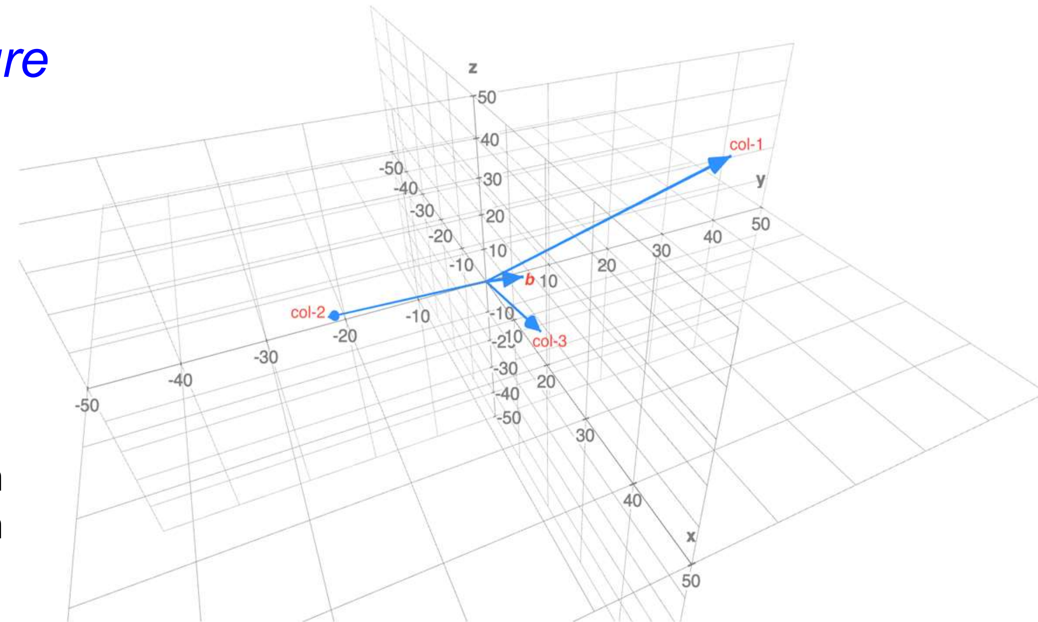
$$A = \begin{pmatrix} 2 & 8 & 4 \\ 2 & 5 & 1 \\ 4 & 10 & -1 \end{pmatrix}, \boldsymbol{x} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}, \text{ and } \boldsymbol{b} = \begin{pmatrix} 2 \\ 5 \\ 1 \end{pmatrix}.$$

Here A is called the *coefficient matrix.*

The three planes defined by the system of equations (ii) intersect at a point $x = 11, y = -4, z = 3$ which is the solution.

This is the *Row Picture.*

Now let us examine the *Column Picture*



$$x\begin{bmatrix}2\\2\\4\end{bmatrix}+y\begin{bmatrix}8\\5\\10\end{bmatrix}+z\begin{bmatrix}4\\1\\-1\end{bmatrix}=\begin{bmatrix}2\\5\\1\end{bmatrix}$$

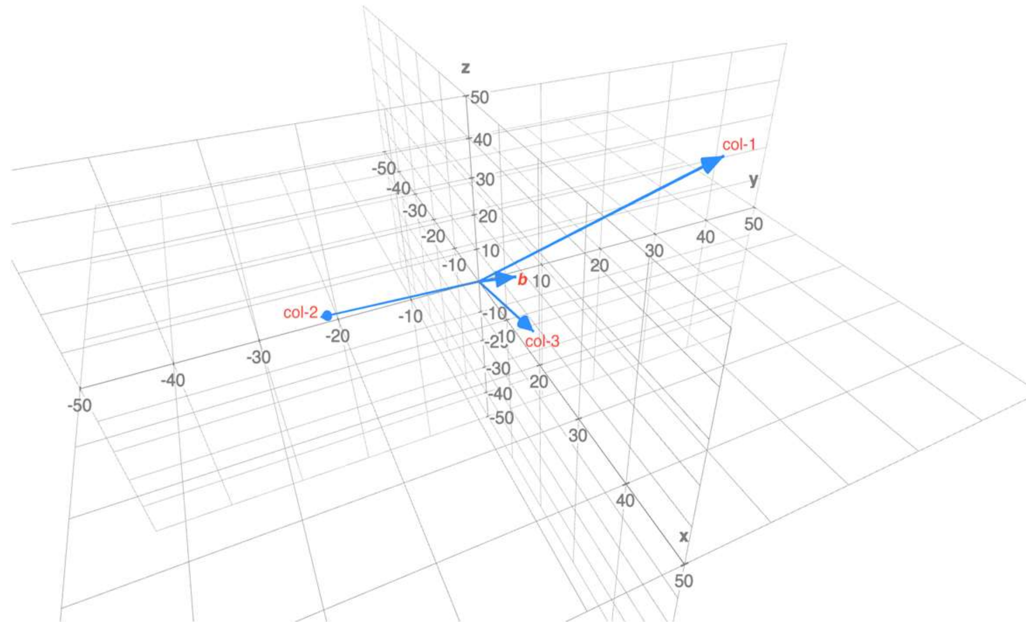
Like in the previous example, a given 3D-vector $\boldsymbol{b}$ can be obtained by the linear combination of the column vectors of A with the appropriate coefficients (x, y, z).

The appropriate coefficients on the l.h.s. form the solution set.

Augmented Matrix $\tilde{A}$, *where* $\tilde{A}=\begin{pmatrix}2 & 8 & 4 & 2\\2 & 5 & 1 & 5\\4 & 10 & -1 & 1\end{pmatrix}$ We will find this to be useful later

This *column picture* and the *linear combination of the columns* allow us to determine, geometrically, the conditions when a unique solution can be attained for any 3D-vector $\boldsymbol{b}$.

The answer is, once again, when A is invertible, i.e. when the columns of A are linearly independent (or equivalently when the column vectors of A are **not** co-planar), *i.e. in that case, the columns vectors of A span the 3-D volume and a proper combination of them will give* ***b***



If the column vectors of A were co-planar, then a linear combination of them will not span the entire 3D plane and a unique solution will not be possible.

A Peek into the Future - "Motivating the **Reduced Row Echelon Form**"

Note that the route to obtain the solution involves finding the inverse of A. This may be difficult especially if the number of unknown variables (and thereby the number of equations) are large.

Computing the inverse of a large matrix becomes simpler by transforming the original coefficient matrix into what is called a **Reduced Row-Echelon Form**
This form lies at the heart of several numerical techniques to solve systems of linear equations and also characterizes several important features of the coefficient matrix.

In any case, as a prelude to what is to come soon, the solution to this system of linear equations can be obtained by transforming the **Augmented Matrix** $\tilde{A}$, *where* $\tilde{A}=\left(\begin{array}{ccc|c} 2 & 8 & 4 & 2 \\ 2 & 5 & 1 & 5 \\ 4 & 10 & -1 & 1 \end{array}\right)$, into its **Reduced Row-Echelon Form** $rref(\tilde{A}) = \begin{pmatrix} 1 & 0 & 0 & 11 \\ 0 & 1 & 0 & -4 \\ 0 & 0 & 1 & 3 \end{pmatrix}$.

From this, the solutions can be directly obtained as $x = 11, y = -4, z = 3$. We will study this next.
We will study subsequently why this turns out to be the case. The point here is that the matrix structure and its rules (laws of matrix algebra) are useful to solve such systems of linear equations.
We will later do a lab project to further understand the power of this technique to solve engineering problems.

5. Reduced Row-Echelon Form or rref (Definition):

A matrix is said to be in rref if it satisfies **all** the following conditions

i. If a row has non-zero entries, then the first non-zero entry is a 1, known as the *leading 1* (or *pivot*) in this row.
ii. If a column has a leading 1, then all the other entries in that columns are 0.
iii. If a row contains a leading 1, then each row above it contains a leading 1 further to the left.

The third condition implies that rows of 0's, if any, appear at the bottom of the matrix.

6. Types of elementary row operations (in order to obtain the rref):

i. Divide a row by a non-zero scalar.
ii. Subtract a multiple of a row from another row.
iii. Swap two rows.

Question: Why can we do these operations?

Answer: Since the rows are the rows of the corresponding system of linear equations, these operations are such that doing them will not affect the solution to the equations.

We will later see that points 5 and 6 above form the core of the powerful GAUSS-JORDAN ELIMINATION approach to solve systems of linear equations.

Example:

System of Equations

Ax=b

$$\left|\begin{array}{c} x+y+z=3 \\ 2x-3y-z=-8 \\ -x+2y+2z=3 \end{array}\right| \quad A=\begin{pmatrix} 1 & 1 & 1 \\ 2 & -3 & -1 \\ -1 & 2 & 2 \end{pmatrix} \quad \widetilde{A}=\begin{pmatrix} 1 & 1 & 1 & 3 \\ 2 & -3 & -1 & -8 \\ -1 & 2 & 2 & 3 \end{pmatrix}$$

Augmented Matrix

Augmented Matrix

$$\begin{pmatrix} 1 & 1 & 1 & 3 \\ 2 & -3 & -1 & -8 \\ -1 & 2 & 2 & 3 \end{pmatrix} \to \begin{pmatrix} 1 & 1 & 1 & 3 \\ 0 & 1 & 3 & -2 \\ 0 & 3 & 3 & 6 \end{pmatrix} \to \begin{pmatrix} 1 & 1 & 1 & 3 \\ 0 & 1 & 3 & -2 \\ 0 & 0 & -6 & 12 \end{pmatrix} \to \begin{pmatrix} 1 & 1 & 1 & 3 \\ 0 & 1 & 3 & -2 \\ 0 & 0 & 1 & -2 \end{pmatrix}$$

Row Echelon Form

Row Echelon Form

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

Reduced Row Echelon Form

Solution $\quad x=1, \quad y=4, \quad z=-2$

7. Linear Transformations (Definition):

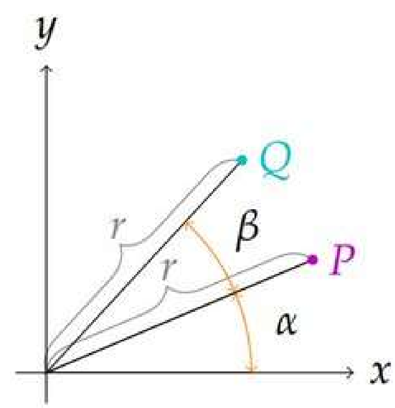
A function $T: \mathbb{R}^n \to \mathbb{R}^m$ is called a *linear transformation* if $\exists A \in \mathbb{M}_{m\times n}(\mathbb{R})$ such that

$T(\boldsymbol{x}) = \boldsymbol{A}\boldsymbol{x}, \forall \boldsymbol{x} \in \mathbb{R}^{\boldsymbol{n}}$ *m*X1 $\leftarrow$ *m*X*n* *n*X1

This is a mapping from *n*-dimensional space to *m*-dimensional space

Example: The rotation matrix $\begin{pmatrix} cos\theta & -sin\theta \\ sin\theta & cos\theta \end{pmatrix}$ is a linear transformation which rotates a vector in $\mathbb{R}^2$by θ.

We shall derive the rotation matrix that counter-clockwise rotates a 2D point $P=(x,y)$, β radians around the origin, to yield the rotated point Q. Let r and α be the polar coordinates of P. This means



$$P=(x,y)=r(\cos(\alpha),\sin(\alpha))$$

In order to rotate this point β radians, we simply add β to α. That is,

$$Q=r(\cos(\alpha+\beta),\sin(\alpha+\beta))$$

Applying the trigonometric addition formulas gives

$$\begin{aligned}
Q&=r(\cos(\alpha)\cos(\beta)-\sin(\alpha)\sin(\beta),\sin(\alpha)\cos(\beta)+\cos(\alpha)\sin(\beta))\\
&=(\cos(\beta)r\cos(\alpha)-\sin(\beta)r\sin(\alpha),\sin(\beta)r\cos(\alpha)+\cos(\beta)r\sin(\alpha))\\
&=(\cos(\beta)x-\sin(\beta)y,\sin(\beta)x+\cos(\beta)y)\\
&=\begin{pmatrix}\cos(\beta) & -\sin(\beta)\\ \sin(\beta) & \cos(\beta)\end{pmatrix}\begin{pmatrix}x\\ y\end{pmatrix}=\begin{pmatrix}\cos(\beta) & -\sin(\beta)\\ \sin(\beta) & \cos(\beta)\end{pmatrix}P
\end{aligned}$$

To conclude, applying the blue matrix to P gives us the rotated point Q.

Ques: Given $T: \mathbb{R}^n \to \mathbb{R}^m$, how do we find A?

………. n columns ………….

m rows

Ans: $$A = \begin{pmatrix} | & | & & & | \\ T(\boldsymbol{e_1}) & T(\boldsymbol{e_2}) & \cdot & \cdot & T(\boldsymbol{e_n}) \\ | & | & & & | \end{pmatrix}$$

where $\boldsymbol{e_i}$ is the i^{th} standard basis element of $\mathbb{R}^n$.

Finding *A*, given the transformation *T* – Another Example

The transformation *T* is given as - $T\begin{pmatrix}1\\2\end{pmatrix}=\begin{pmatrix}2\\3\end{pmatrix} \quad T\begin{pmatrix}3\\2\end{pmatrix}=\begin{pmatrix}1\\2\end{pmatrix}$

Method 1: Manipulate directly to find

$$T\begin{pmatrix}1\\0\end{pmatrix}=\begin{pmatrix}-1/2\\-1/2\end{pmatrix} \quad T\begin{pmatrix}0\\1\end{pmatrix}=\begin{pmatrix}5/4\\7/4\end{pmatrix} \quad \Rightarrow \quad A=\begin{pmatrix}-1/2 & 5/4\\-1/2 & 7/4\end{pmatrix}$$

Method 2: Assume and solve for a_1, a_2, b_1, b_2

$$A=\begin{pmatrix}a_1 & b_1\\a_2 & b_2\end{pmatrix} \quad then \quad \begin{pmatrix}a_1 & b_1\\a_2 & b_2\end{pmatrix}\begin{pmatrix}1\\2\end{pmatrix}=\begin{pmatrix}2\\3\end{pmatrix} \quad \begin{pmatrix}a_1 & b_1\\a_2 & b_2\end{pmatrix}\begin{pmatrix}3\\2\end{pmatrix}=\begin{pmatrix}1\\2\end{pmatrix}$$

Method 3: Using Matrix Inversion

$$A\begin{pmatrix}1 & 3\\2 & 2\end{pmatrix}=\begin{pmatrix}2 & 1\\3 & 2\end{pmatrix} \quad \Rightarrow \quad A=\begin{pmatrix}2 & 1\\3 & 2\end{pmatrix}\begin{pmatrix}1 & 3\\2 & 2\end{pmatrix}^{-1}=\begin{pmatrix}2 & 1\\3 & 2\end{pmatrix}\begin{pmatrix}-1/2 & 3/4\\1/2 & -1/4\end{pmatrix}=\begin{pmatrix}-1/2 & 5/4\\1/2 & 7/4\end{pmatrix}$$

Continuing with this example -

$$T\,\vec{x} = A\,\vec{x} = \begin{pmatrix} -\frac{1}{2} & \frac{5}{4} \\ -\frac{1}{2} & \frac{7}{4} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

transforms a 2-dimensional vector in ***V**, i.e.* $\begin{pmatrix} x \\ y \end{pmatrix}$

to another two dimensional vector in ***W*** $\begin{pmatrix} -x\frac{1}{2} + y\frac{5}{4} \\ -x\frac{1}{2} + y\frac{7}{4} \end{pmatrix}$

This gives the **image** of the transformation *T* in the target space *W* spanned by the basis

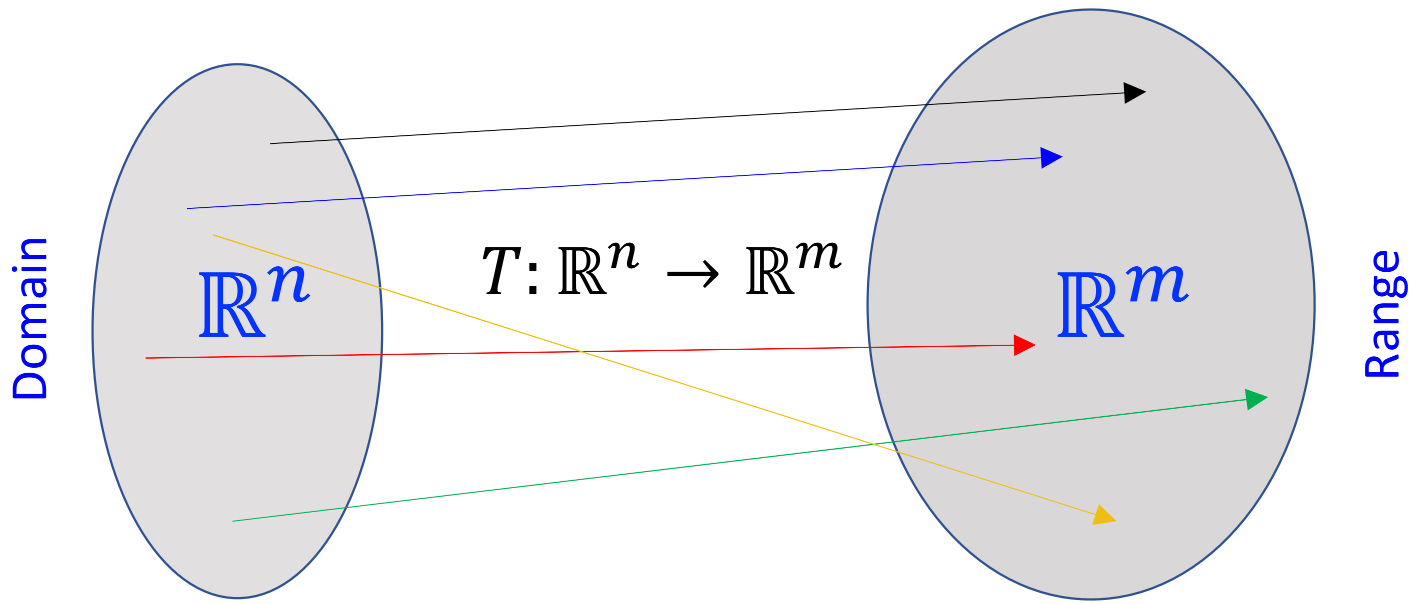
$$\begin{pmatrix} -\frac{1}{2} \\ -\frac{1}{2} \end{pmatrix}, \begin{pmatrix} -\frac{5}{4} \\ -\frac{7}{4} \end{pmatrix}$$

When does this transformation create a {0} in the target space?

When *x*=0 and *y*=0

Ker(*T*), the kernel of the transformation

Transformation from $\mathbb{R}^n \to \mathbb{R}^m$



An invertible linear transformation is **a map between vector spaces and with an inverse map** which is also a linear transformation.

Theorem: A $n \times n$ matrix A is invertible $\Leftrightarrow rref(A) = I_n \equiv rank(A) = n$.

In that case, the matrix will have *n* independent rows (or columns)

Finding the inverse of a matrix: $A \in \mathbb{M}_{n \times n}(\mathbb{R})$

In order to find A^{-1}, form the augmented matrix $\tilde{A} = (A|\, I_n)$ and compute $rref(\tilde{A})$

- If $rref(\tilde{A})$ is of the form $(I_n|\, B)$, then $A^{-1} = B$
- If $rref(\tilde{A})$ is of any other form, then A is **not** invertible.

$(AB)^{-1} = B^{-1}A^{-1}$ note the change in order of *A* and *B*

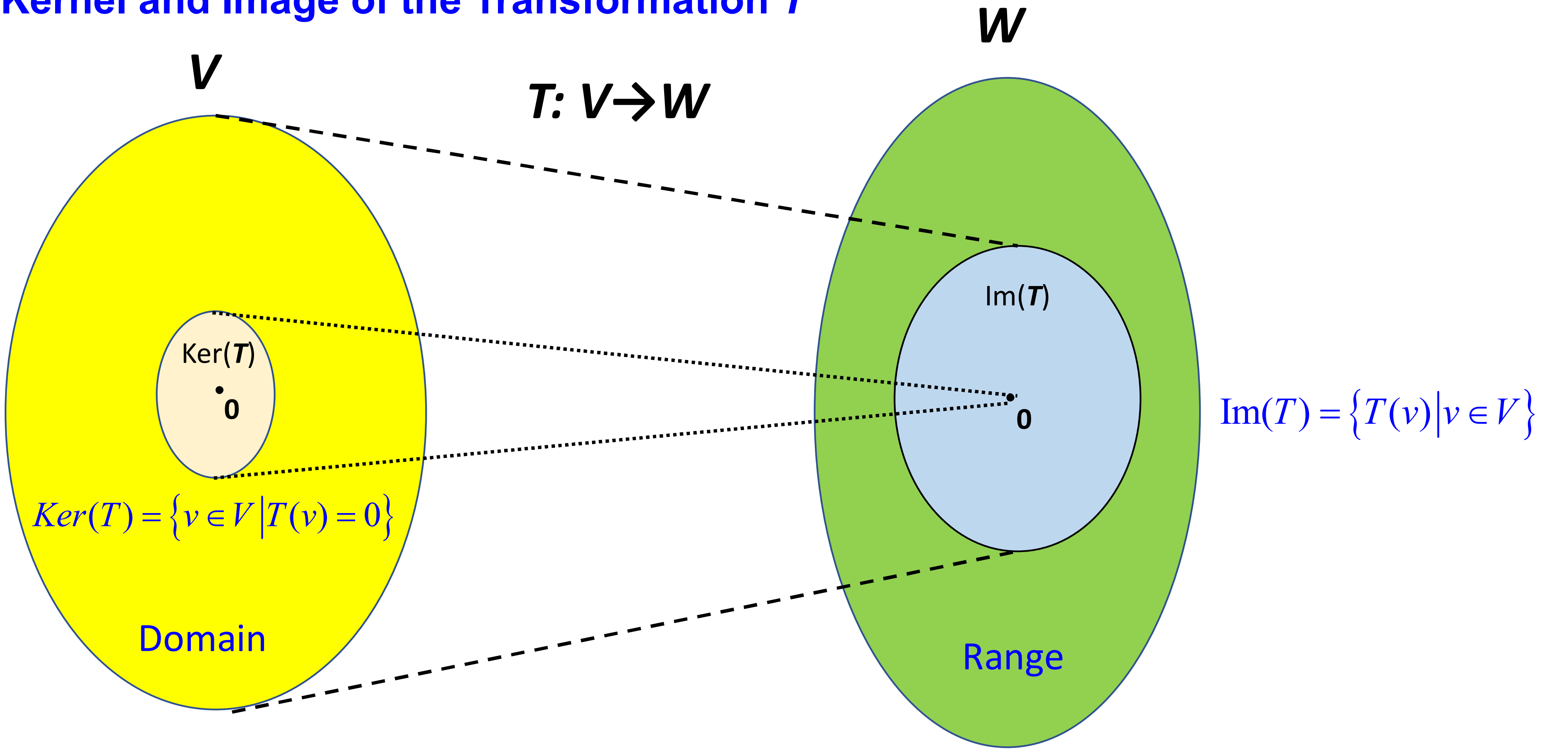
Find Inverse of $\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$

Find rref $\begin{pmatrix} 2 & 1 & 1 & 0 \\ 1 & 2 & 0 & 1 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 1/2 & 1/2 & 0 \\ 1 & 2 & 0 & 1 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 1/2 & 1/2 & 0 \\ 0 & 3/2 & -1/2 & 1 \end{pmatrix}$

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

$$\left(I \;\middle|\; \begin{matrix} 2/3 & -1/3 \\ -1/3 & 2/3 \end{matrix} \right)$$

Kernel and Image of the Transformation *T*



8. Image or Range of a Matrix/Linear Transformation (Definition)

$Im(A) = Im(T)$ is the ***span*** of the column vectors of A.

Que: Find a basis of the image of $A = \begin{pmatrix} 1 & 2 & 2 & -5 & 6 \\ -1 & -2 & -1 & 1 & -1 \\ 4 & 8 & 5 & -8 & 9 \\ 3 & 6 & 1 & 5 & -7 \end{pmatrix} = \begin{pmatrix} | & | & | & | & | \\ \boldsymbol{a_1} & \boldsymbol{a_2} & \boldsymbol{a_3} & \boldsymbol{a_4} & \boldsymbol{a_5} \\ | & | & | & | & | \end{pmatrix}$

and determine $dim(Im(A))$

Ans: A basis of the image of A can be found as $\begin{pmatrix} 1 \\ -1 \\ 4 \\ 3 \end{pmatrix}, \begin{pmatrix} 2 \\ -1 \\ 5 \\ 1 \end{pmatrix}$ which gives $dim(Im(A)) = 2$

See next slide

To find the basis of Im(*A*), we need to identify the redundant columns of *A* from amongst all the column vectors of *A*.

By mere inspection of *A*, it will be hard to tell which of the columns of *A* are redundant (i.e. linearly dependent on the others).

So we will transform *A* to $B = rref(A) = \begin{pmatrix} 1 & 2 & 0 & 3 & -4 \\ 0 & 0 & 1 & -4 & 5 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} | & | & | & | & | \\ \boldsymbol{b_1} & \boldsymbol{b_2} & \boldsymbol{b_3} & \boldsymbol{b_4} & \boldsymbol{b_5} \\ | & | & | & | & | \end{pmatrix}$.

The redundant columns of *B* correspond to the redundant columns of *A* and are easy to spot from *B*

These are the columns that do not contain a leading 1, i.e., $\boldsymbol{b_2} = 2\boldsymbol{b_1}, \boldsymbol{b_4} = 3\boldsymbol{b_1} - 4\boldsymbol{b_3},\ \boldsymbol{b_5} = -4\boldsymbol{b_1} + 5\boldsymbol{b_3}$

Therefore, the redundant columns of A are $\boldsymbol{a_2} = 2\boldsymbol{a_1},\ \boldsymbol{a_4} = 3\boldsymbol{a_1} - 4\boldsymbol{a_3}$, and $\boldsymbol{a_5} = -4\boldsymbol{a_1} + 5\boldsymbol{a_3}$

and the non-redundant columns of A are $\boldsymbol{a_1}$ and $\boldsymbol{a_3}$ which form a basis of image of A as given in the previous slide

9. Kernel of T (Definition):

Kernel of T (or equivalently the **null space of $\boldsymbol{A}, \boldsymbol{Null(A)}$**):

The set of all $x \in \mathbb{R}^n$ s.t. $T(\boldsymbol{x}) = A\boldsymbol{x} = \boldsymbol{0}$

Q) Find a basis of the kernel of A (equivalently, $Null(A)$) and determine $dim(Ker(A)) = dim(null(A))$.

Ans) Most importantly $Ker(A) = Ker(rref(A)) = Ker(B)$.

So we might as well solve for $\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} s.t.\ B\boldsymbol{x} = \boldsymbol{0}$.

This is done by considering the augmented matrix $\tilde{B} = (B|\mathbf{0})$ from which we have the following:

$$\begin{aligned} x_1 + 2x_2 + 0x_3 + 3x_4 - 4x_5 &= 0 \\ 0x_1 + 0x_2 + x_3 - 4x_4 + 5x_5 &= 0 \end{aligned}$$

How we can obtain the rref(A) is given in the next slide.

or equivalently,

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

where $x_2 = \alpha, x_4 = \beta, x_5 = \gamma$ are set arbitrarily.

Transforming a Matrix to a Reduced Row Echelon Form

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

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

$$\begin{bmatrix} 1 & 2 & 2 & -5 & 6 \\ 0 & 0 & 1 & -4 & 5 \\ 0 & 0 & -3 & 12 & -15 \\ 3 & 6 & 1 & 5 & -7 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 2 & 2 & -5 & 6 \\ 0 & 0 & 1 & -4 & 5 \\ 0 & 0 & -3 & 12 & -15 \\ 0 & 0 & -5 & 20 & -25 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 2 & 2 & -5 & 6 \\ 0 & 0 & 1 & -4 & 5 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -5 & 20 & -25 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 2 & 2 & -5 & 6 \\ 0 & 0 & 1 & -4 & 5 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

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

Leave it to you to work out the details of what is done in each step!

Therefore,

$$\boldsymbol{x} = \begin{pmatrix} -2\alpha - 3\beta + 4\gamma \\ \alpha \\ 4\beta - 5\gamma \\ \beta \\ \gamma \end{pmatrix} = \begin{pmatrix} -2\alpha & -3\beta & +4\gamma \\ \alpha & & \\ & 4\beta & -5\gamma \\ & \beta & \\ & & \gamma \end{pmatrix} = \alpha \begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \beta \begin{pmatrix} -3 \\ 0 \\ 4 \\ 1 \\ 0 \end{pmatrix} + \gamma \begin{pmatrix} 4 \\ 0 \\ -5 \\ 0 \\ 1 \end{pmatrix}.$$

The $Null(A)$ is spanned by these basis vectors $\begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} -3 \\ 0 \\ 4 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 4 \\ 0 \\ -5 \\ 0 \\ 1 \end{pmatrix}$

and $dim(Null(A)) = 3$

Something for you to try out –

Question: Find the basis for the null space of the matrix

$$A = \begin{pmatrix} 1 & -1 & 1 & 2 \\ 2 & 1 & 3 & 2 \\ 1 & 5 & 3 & -2 \end{pmatrix}$$ and determine its dimension

Answer:

$$\begin{pmatrix} -4/3 \\ -1/3 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -4/3 \\ 2/3 \\ 0 \\ 1 \end{pmatrix}$$

and the dimension of null space of A is 2 rref (A)=

1	0	4/3	4/3
0	1	1/3	-2/3
0	0	0	0

Basis of Image of A = $\begin{pmatrix} 1 & -1 \\ 2 & 1 \\ 1 & 5 \end{pmatrix}$

dim Im(A) = 2

$$x_1 + 0 + \frac{4}{3}x_3 + \frac{4}{3}x_4 = 0 \quad \Rightarrow \quad x_1 = -\frac{4}{3}x_3 - \frac{4}{3}x_4$$

$$0 + x_2 + \frac{1}{3}x_3 - \frac{2}{3}x_4 = 0 \quad \Rightarrow \quad x_2 = -\frac{1}{3}x_3 + \frac{2}{3}x_4$$

$$x_3 = \alpha \quad x_4 = \beta \quad \Rightarrow \quad \vec{X} = \begin{pmatrix} -\frac{4}{3}\alpha - \frac{4}{3}\beta \\ -\frac{1}{3}\alpha + \frac{2}{3}\beta \\ \alpha \\ \beta \end{pmatrix} = \alpha \begin{pmatrix} -\frac{4}{3} \\ -\frac{1}{3} \\ 1 \\ 0 \end{pmatrix} + \beta \begin{pmatrix} -\frac{4}{3} \\ \frac{2}{3} \\ 0 \\ 1 \end{pmatrix}$$

Basis for the Null Space of A $\begin{pmatrix} -\frac{4}{3} \\ -\frac{1}{3} \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -\frac{4}{3} \\ \frac{2}{3} \\ 0 \\ 1 \end{pmatrix}$

Something for you to try out –

Que: Find the basis for the null space of the matrix
and determine its dimension

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

Answer: $\begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad \begin{pmatrix} -2 \\ 0 \\ -1 \\ 1 \\ 0 \end{pmatrix}$ and the dimension of null space of *A* is 2 rref (A)= $\begin{pmatrix} 1 & -1 & 0 & 2 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$

Basis of Image of A = $\begin{pmatrix} 1 \\ -1 \\ 1 \\ 2 \end{pmatrix} \begin{pmatrix} -1 \\ 0 \\ -2 \\ -1 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix}$

& *dim Im*(A) = 3

$$rref(A)\,\vec{x} = 0 \quad \begin{aligned} x_1 - x_2 \quad + 2x_4 &= 0 \\ x_3 + x_4 &= 0 \\ x_5 &= 0 \end{aligned} \qquad \begin{aligned} x_1 &= x_2 - 2x_4 \\ x_3 &= -x_4 \\ x_5 &= 0 \end{aligned}$$

$$\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = x_2 \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + x_4 \begin{pmatrix} -2 \\ 0 \\ -1 \\ 1 \\ 0 \end{pmatrix}$$

Basis for the Null Space of A $\begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \begin{pmatrix} -2 \\ 0 \\ -1 \\ 1 \\ 0 \end{pmatrix}$ *dim Ker*(A) = 2

10. Theorem: $A \in \mathbb{M}_{m\times n}(\mathbb{R})$. Then $Ker(A) = \{\mathbf{0}\} \Leftrightarrow rank(A) = n$.

For a square matrix the statement is true when A is invertible

When *A* is invertible, $rref(A) = I_n \Longrightarrow$ no. of pivots = *n* = rank(*A*) by definition.

Further, $A\boldsymbol{x} = 0$ *can be solved by considering the augmented matrix* $rref(A|\mathbf{0}) = (I_n|\mathbf{0})$) which gives us $x_1 = 0, x_2 = 0, x_3 = 0$ which gives $Ker(A) = \{\mathbf{0}\}$. The converse is obvious.

11. Theorem (*Rank-nullity theorem*): For any $m \times n$ matrix A, the following is known as the *fundamental theorem of linear algebra*:

$$dim(Null(A)) + dim(Im(A)) \quad = n$$

$$or \;\; equivalently,$$

$$(nullity\ of A) + (rank\ of\ A) \qquad = n$$

Proof of the Rank Nullity Theorem

$\boldsymbol{A}$ m x n matrix with **rank** $\boldsymbol{r}$ and **nullity** $\boldsymbol{\ell}$

Rank Nullity Theorem claims that $\boldsymbol{r} + \boldsymbol{\ell} = \boldsymbol{n}$

Consider the matrix equation $\boldsymbol{Ax} = \mathbf{0}$ and assume that $\boldsymbol{A_{rref}}$ = rref ($\boldsymbol{A}$)

- The elementary row operations which reduce $\boldsymbol{A}$ to $\boldsymbol{A_{rref}}$ do not change the row space of $\boldsymbol{A}$ or the rank of $\boldsymbol{A}$
- The number of components in $\boldsymbol{x}$ is n, which is also the number of columns of $\boldsymbol{A}$ and of $\boldsymbol{A_{rref}}$
- Since $\boldsymbol{A_{rref}}$ has only r nonzero rows (because its rank is r), $n-r$ of the variables $x_1, x_2, \ldots, x_n$ in $\boldsymbol{x}$ are free for us to choose as parameters in the general solution of $\boldsymbol{Ax} = \mathbf{0}$
- But the number of free variables, i.e. the number of parameters in the general solution of $\boldsymbol{Ax} = \mathbf{0}$, is the nullity of $\boldsymbol{A}$.
- The nullity of A is then $n - r$, and the statement of the theorem, $\boxed{r + \ell = r + (n - r) = n}$ follows

Example

$$A = \begin{pmatrix} 1 & -2 & 0 & 4 \\ 3 & 1 & 1 & 0 \\ -1 & -5 & -1 & 8 \end{pmatrix} \qquad A_{rref} = rref(A) = \begin{pmatrix} 1 & 0 & \frac{2}{7} & \frac{4}{7} \\ 0 & 1 & \frac{1}{7} & -\frac{12}{7} \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

To solve the system ***Ax*=*0***, choose *x*3 and *x*4 and solve for *x*1 and *x*2 in terms of *x*3 and *x*4

$$x_1 = -\frac{2}{7}x_3 - \frac{4}{7}x_4$$

$$x_2 = -\frac{1}{7}x_3 + \frac{12}{7}x_4$$

$$\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = x_3 \begin{pmatrix} -\frac{2}{7} \\ -\frac{1}{7} \\ 1 \\ 0 \end{pmatrix} + x_4 \begin{pmatrix} -\frac{4}{7} \\ \frac{12}{7} \\ 0 \\ 1 \end{pmatrix}$$

Im(***A***) is spanned by

$$\begin{pmatrix} 1 \\ 3 \\ -1 \end{pmatrix} \begin{pmatrix} -2 \\ 1 \\ -5 \end{pmatrix}$$

and *dim*(*Im*(***A***))=2

Null(***A***) or *Ker*(***A***)is spanned by

$$\begin{pmatrix} -\frac{2}{7} \\ -\frac{1}{7} \\ 1 \\ 0 \end{pmatrix} \begin{pmatrix} -\frac{4}{7} \\ -\frac{12}{7} \\ 0 \\ 1 \end{pmatrix}$$

and *dim*(*Im*(***A***))=2

Rank-Nullity Theorem is satisfied

Mini-project: Transactions on an accounting system

Objective: This simple mini-project will demonstrate an application from economics/accounting whereby we will be required to compute the basis of the null space of a certain matrix. This basis will represent the most fundamental unit of transaction in a closed accounting system.

Description: Consider a *closed* accounting system with n accounts, say $\alpha_1, \alpha_2, \alpha_3, \ldots \alpha_n$. At any instant, each account has a balance which can be a *credit* (positive), *debit* (negative), or *zero*. Since the accounting system must at all times be in balance, the sum of the balances of all the accounts will always be zero. Now suppose that a *transaction* is applied to the system. By this we mean that there is a flow of funds between accounts of this system. If as a result of the transaction the balance of account α_i changes by an amount t_i, then the transaction can be represented by an n-column vector with entries $t_1, t_2, t_3, \ldots, t_n$. Since the accounting system must still be in balance after the transaction has been applied, the sum of the t_is will be zero.

The transactions correspond to column vectors of the form $\boldsymbol{t} = \begin{pmatrix} t_1 \\ t_2 \\ \cdot \\ \cdot \\ \cdot \\ t_n \end{pmatrix}$. Vectors of this form are easily seen to constitute a subspace $\mathbb{T}$ of the vector space $\mathbb{R}^n$; $\mathbb{T}$ is called the *transaction space.*

Questions:

Q1) Construct a matrix A such that $\mathbb{T}$ is just the null space of A.

Q2) Deduce the reduced row-echelon form, $\tilde{A}$ of A.

Q3) Consider the solution of the equation $\tilde{A}\boldsymbol{x} = 0$. Express $\boldsymbol{x}$ as a linear combination of the most canonical column vectors.

Q4) Deduce the basis of the *transaction space* $\mathbb{T}$.

Q5) What is the dimension of $\mathbb{T}$? Given that $\mathbb{T} \subset \mathbb{R}^n$, does the dimension of $\mathbb{T}$, you have just computed, make sense? Why?

Q6) Justify why your answer to Q4) above represents the most fundamental activity in this accounting system?

Computational solutions to systems of linear equations

Example 1: Let us consider the following system of equations

$$x_1 - x_2 + x_3 + x_4 = 2 \ldots\ldots\ldots\ldots (i)$$
$$x_1 + x_2 + x_3 - x_4 = 3 \ldots\ldots\ldots\ldots (ii)$$
$$x_1 + 3x_2 + x_3 - 3x_4 = 1 \ldots\ldots\ldots\ldots (iii)$$

Does this set of equations have a solution?

Let's perform the following operations to eliminate x_1:

$eq.(v): (ii) - (i)$ and $eq.(vi): (iii) - (i)$

$$x_1 - x_2 + x_3 + x_4 = 2 \ldots\ldots\ldots\ldots (iv)$$
$$x_2 - x_4 = \frac{1}{2} \ldots\ldots\ldots\ldots (v)$$
$$4x_2 - 4x_4 = -1 \ldots\ldots\ldots\ldots (vi)$$

Next we will attempt to eliminate x_2 from $eq.(vi)$: $eq.(ix): (vi) - 4(v)$,

$$\begin{aligned} x_1 - x_2 + x_3 + x_4 &= 2 \ldots\ldots\ldots\ldots (vii) \\ x_2 - x_4 &= \tfrac{1}{2} \ldots\ldots\ldots\ldots (viii) \\ 0 &= -3 \ldots\ldots\ldots\ldots (ix)\ OOPS! \end{aligned}$$

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

This is a **contradiction** and hence the above system of equations is **inconsistent** (no solutions)

This conclusion should have been an obvious one as here we have only THREE equations to solve for FOUR unknowns!

Example 2: Now consider the following system of equations

$$
\begin{aligned}
x_1 + 4x_2 + 2x_3 &= -2 \ldots\ldots\ldots\ldots\ldots\ldots (i)\\
-2x_1 - 8x_2 + 3x_3 &= 32 \ldots\ldots\ldots\ldots\ldots\ldots (ii)\\
x_2 + x_3 &= 1 \ldots\ldots\ldots\ldots\ldots\ldots (iii)
\end{aligned}
$$

Does the above system of equations have a solution?

$Eq.(v)$: $(ii) + 2(i)$ gives

$$
\begin{aligned}
x_1 + 4x_2 + 2x_3 &= -2 \ldots\ldots\ldots\ldots\ldots\ldots (iv)\\
7x_3 &= 28 \ldots\ldots\ldots\ldots\ldots\ldots (v)\\
x_2 + x_3 &= 1 \ldots\ldots\ldots\ldots\ldots\ldots (vi)
\end{aligned}
$$

Swap $eq.(v) and (vi)$:

$$
\begin{aligned}
x_1 + 4x_2 + 2x_3 &= -2 \ldots\ldots\ldots\ldots\ldots\ldots (vii)\\
x_2 + x_3 &= 1 \ldots\ldots\ldots\ldots\ldots\ldots (viii)\\
7x_3 &= 28 \ldots\ldots\ldots\ldots\ldots\ldots (ix)
\end{aligned}
$$

Eq. (xii): $\frac{1}{7}(ix)$

$$\begin{aligned} x_1 + 4x_2 + 2x_3 &= -2 \ldots\ldots\ldots\ldots\ldots\ldots (x) \\ x_2 + x_3 &= 1 \ldots\ldots\ldots\ldots\ldots\ldots (xi) \\ x_3 &= 4 \ldots\ldots\ldots\ldots\ldots\ldots (xii) \end{aligned}$$

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

Now that we know $x_3 = 4$, we can *back substitute* the knowns and compute the remaining unknowns: $x_2 = 1 - x_3 = -3$ and $x_1 = -2 - 4x_2 - 2x_3 = 2$.
Therefore, this system of equation has a unique solution.

This example will be very similar to the numerical technique we will learn in this section known as *Gauss elimination.*

Before we study this new method (Gauss elimination), let us look at one more example next

Example 3: Consider the following system of equations

$$\begin{aligned} x_1 + 3x_2 + 3x_3 + 2x_4 &= 1 \ldots\ldots\ldots\ldots\ldots (i) \\ 2x_1 + 6x_2 + 9x_3 + 5x_4 &= 5 \ldots\ldots\ldots\ldots\ldots (ii) \\ -x_1 - 3x_2 + 3x_3 &= 5 \ldots\ldots\ldots\ldots\ldots (iii) \end{aligned}$$

$Eq. (ii): (ii) - 2(i) and (iii): (iii) + (i)$ gives

$$\begin{aligned} x_1 + 3x_2 + 3x_3 + 2x_4 &= 1 \ldots\ldots\ldots\ldots\ldots (i) \\ 3x_3 + x_4 &= 3 \ldots\ldots\ldots\ldots\ldots (ii) \\ 6x_3 + 2x_4 &= 6 \ldots\ldots\ldots\ldots\ldots (iii) \end{aligned}$$

Note x_2 has disappeared from eqs. $(ii) and (iii)$; so we proceed to the next unknown x_3!

Example 3 (continued) **:**

Eq. (ii): $\frac{1}{3}(ii)$ followed by $eq.(iii)$: $(iii) - 2(ii)$:

$$rref = \begin{pmatrix} 1 & 3 & 0 & 1 & -2 \\ 0 & 0 & 1 & \frac{1}{3} & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\begin{aligned} x_1 + 3x_2 + 3x_3 + 2x_4 &= 1 \ldots\ldots\ldots\ldots\ldots (i) \\ x_3 + \tfrac{1}{3}x_4 &= 1 \ldots\ldots\ldots\ldots\ldots (ii) \\ 0 &= 0 \ldots\ldots\ldots\ldots\ldots (iii) \end{aligned}$$

Here the third equation tells us nothing and can be ignored.
x_4 and x_2 can be assigned arbitrary values: $x_4 = c, x_2 = d$; to recover (by back substitution) $x_1 = -2 - c - 3d; x_3 = 1 - c/3$.

Stated more completely, one of x_3 and x_4 and one of x_1 and x_2 must be assigned arbitrary values and the other found by back substitution

This system of linear equations has infinitely many solutions!

Example Summary

Example 1

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

- No solution
- System of Equations Not Consistent

Example 2

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

Gauss Elimination

Gauss Jordan Elimination

- Unique Solution
- Number of Independent Equations = Number of Unknowns

Example 3

$$rref = \begin{pmatrix} 1 & 3 & 0 & 1 & -2 \\ 0 & 0 & 1 & \frac{1}{3} & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

- Infinitely Many Solutions
- Number of Independent Equations < Number of Unknowns

Gauss Elimination method (direct computational method!)**:**

Here we are solving system of linear equations of the form:

$$
\begin{aligned}
a_{11}x_1 + a_{12}x_2 + \cdots a_{1n}x_n &= b_1 \ldots\ldots\ldots\ldots (i) \\
a_{21}x_1 + a_{22}x_2 + \cdots a_{2n}x_n &= b_2 \ldots\ldots\ldots\ldots (ii) \\
\cdot\cdot \quad &= \cdot \\
\cdot\cdot \quad &= \cdot \\
\cdot\cdot \quad &= \cdot \\
a_{m1}x_1 + a_{m2}x_2 + \cdots a_{mn}x_n &= b_m \ldots\ldots\ldots\ldots (m^{th}\, eq.)
\end{aligned}
$$

Note that this is the same as the equation ***Ax* = *b***

(***A*** is a *mxn* matrix, ***x*** is a *nx1* column vector and ***b*** is a *mx1* column vector.)

Procedure: (Gauss Elimination Method)

i) Find an equation in which x_1 appears and, if necessary, interchange this equation with the first equation. Thus we can assume that x_1 appears in the first equation.

ii) Multiply eq. (i) by a suitable non-zero scalar in such a way as to make the coefficient of x_1 equal to 1.

iii) Subtract suitable multiples of eq. (i) from eqs. (ii) through (m) in order to eliminate x_1 from each of these equations.

iv) Inspect equations (ii) through (m) and find the first equation which involves one of the unknowns $x_2, \ldots, x_n$, say x_{i_2}. By interchanging equations once again, we can suppose that x_{i_2} appears in eq. (ii).

v) Multiply eq. (ii) by a suitable non-zero scalar to make the coefficient of x_{i_2} equal to 1.

Procedure: (Gauss Elimination Method) …….. continued….

vi) Subtract multiples of eq. (ii) from eq. (iii) through (m) to eliminate x_{i_2} from each these equations.

$$
\begin{array}{ccl}
x_{i_1} + * \, x_{i_2} + \cdots * \, x_n & = * \ldots\ldots\ldots\ldots & (i) \\
x_2 + \cdots * \, x_n & = * \ldots\ldots\ldots\ldots & (ii) \\
\cdot\cdot & = \cdot & \\
\cdot\cdot & = \cdot & \\
\cdot\cdot & = \cdot & \\
x_{i_r} + \cdots * \, x_n & = * \ldots\ldots\ldots\ldots & (r^{th}\, eq.) \\
0 & = * & \\
\cdot & = * & \\
\cdot & = * & \\
0 & = * \ldots\ldots\ldots\ldots & (m^{th}\, eq.)
\end{array}
$$

vii) Examine eqs. (iii) through (m) and find the first one that involves an unknown other than x_1 and x_{i_2}, say x_{i_3}. Interchange equations so that x_{i_3} appears in eq. (iii).

Procedure: (Gauss Elimination Method) …….. continued….

This elimination procedure continues in this manner producing the so called pivotal unknowns $x_1 = x_{i_1}, x_{i_2}, \cdots, x_{i_r}$ until we reach a linear system in which no further unknowns occur in the equations beyond the r^{th} equation. A linear system of this sort is said to be in echelon form.

The i_j are integers which satisfy $1 = i_1 < i_2 < \cdots < i_r \leq n$.

After arriving at the echelon form, we use back substitution to solve for the unknowns $x_1, x_2, \ldots, x_n$.

Q) What can be said about the solution(s) of the linear system by inspecting the echelon form?

Theorem:

(i) A linear system is consistent if and only if all the entries on the right hand sides of those equations in echelon form which contain no unknowns are zero.

(ii) If the system is consistent, the non-pivotal unknowns can be given arbitrary values; the general solution is then obtained by using back-substitution to solve for the pivotal unknowns.

(iii) The system has a unique solution if and only if all the unknowns are pivotal.

Matrix Form of Gauss Elimination

Do (i) $R_i \leftrightarrow R_j$

(ii) $R_i : R_i + cR_j$

(iii) $R_i : cR_i$

to put the *Augmented Matrix* in the *row echelon form*

The matrix in row-echelon form will have a descending staircase structure:

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

For this, consider our earlier example once again –

$$\begin{pmatrix} 1 & 3 & 3 & 2 \\ 2 & 6 & 9 & 5 \\ -1 & -3 & 3 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 1 \\ 5 \\ 5 \end{pmatrix}$$

$$\begin{aligned} x_1 + 3x_2 + 3x_3 + 2x_4 &= 1 \ldots\ldots\ldots\ldots\ldots (i) \\ 2x_1 + 6x_2 + 9x_3 + 5x_4 &= 5 \ldots\ldots\ldots\ldots\ldots (ii) \\ -x_1 - 3x_2 + 3x_3 &= 5 \ldots\ldots\ldots\ldots\ldots (iii) \end{aligned}$$

In Matrix Vector Form $A\boldsymbol{x} = \boldsymbol{b}$

with augmented matrix $\tilde{A} = (A \,|\, b) = \begin{pmatrix} 1 & 3 & 3 & 2 & | & 1 \\ 2 & 6 & 9 & 5 & | & 5 \\ -1 & -3 & 3 & 0 & | & 5 \end{pmatrix}$

Having *row echelon form* $rref(\tilde{A}) = \begin{pmatrix} 1 & 3 & 3 & 2 & | & 1 \\ 0 & 0 & 1 & 1/3 & | & 1 \\ 0 & 0 & 0 & 0 & | & 0 \end{pmatrix}$

For $$rref(\widetilde{A}) = \begin{pmatrix} 1 & 3 & 3 & 2 & | & 1 \\ 0 & 0 & 1 & 1/3 & | & 1 \\ 0 & 0 & 0 & 0 & | & 0 \end{pmatrix}$$

The non-pivotal entries are given arbitrary values

$$x_2 = d \text{ and } x_4 = c$$

The pivotal entries are then calculated as

$$x_1 = -2 - c - 3d \text{ and } x_3 = 1 - c/3$$

Consider the system

Augmented Matrix

$$\left|\begin{array}{c} 2x_1+8x_2+4x_3=2 \\ 2x_1+5x_2+x_3=5 \\ 4x_1+10x_2-x_3=1 \end{array}\right| \quad Ax=b \qquad A=\begin{pmatrix} 2 & 8 & 4 \\ 2 & 5 & 1 \\ 4 & 10 & -1 \end{pmatrix} \quad x=\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \quad b=\begin{pmatrix} 2 \\ 5 \\ 1 \end{pmatrix} \quad \widetilde{A}=\begin{pmatrix} 2 & 8 & 4 & 2 \\ 2 & 5 & 1 & 5 \\ 4 & 10 & -1 & 1 \end{pmatrix}$$

$$\begin{pmatrix} 2 & 8 & 4 & 2 \\ 2 & 5 & 1 & 5 \\ 4 & 10 & -1 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & 4 & 2 & 1 \\ 2 & 5 & 1 & 5 \\ 4 & 10 & -1 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & 4 & 2 & 1 \\ 0 & -3 & -3 & 3 \\ 0 & -6 & -9 & -3 \end{pmatrix} \to \begin{pmatrix} 1 & 4 & 2 & 1 \\ 0 & 1 & 1 & -1 \\ 0 & 2 & 3 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & 4 & 2 & 1 \\ 0 & 1 & 1 & -1 \\ 0 & 0 & 1 & 3 \end{pmatrix}$$

Row-Echelon Form

$$x_3=3 \quad x_2=-4 \quad x_3=11 \quad \Rightarrow \quad x=\begin{pmatrix} 11 \\ -4 \\ 3 \end{pmatrix}$$

Directly from last row

By Back Substitution

Solution

Consider the system

Augmented Matrix

$$\left|\begin{matrix} 2x_1 + 8x_2 + 4x_3 = 2 \\ 2x_1 + 5x_2 + x_3 = 5 \\ 4x_1 + 10x_2 - x_3 = 1 \end{matrix}\right| \quad Ax = b \quad A = \begin{pmatrix} 2 & 8 & 4 \\ 2 & 5 & 1 \\ 4 & 10 & -1 \end{pmatrix} \quad x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \quad b = \begin{pmatrix} 2 \\ 5 \\ 1 \end{pmatrix} \quad \tilde{A} = \begin{pmatrix} 2 & 8 & 4 & 2 \\ 2 & 5 & 1 & 5 \\ 4 & 10 & -1 & 1 \end{pmatrix}$$

$$\begin{pmatrix} 2 & 8 & 4 & 2 \\ 2 & 5 & 1 & 5 \\ 4 & 10 & -1 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & 4 & 2 & 1 \\ 2 & 5 & 1 & 5 \\ 4 & 10 & -1 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & 4 & 2 & 1 \\ 0 & -3 & -3 & 3 \\ 0 & -6 & -9 & -3 \end{pmatrix} \to \begin{pmatrix} 1 & 4 & 2 & 1 \\ 0 & 1 & 1 & -1 \\ 0 & -6 & -9 & -3 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & -2 & 5 \\ 0 & 1 & 1 & -1 \\ 0 & 0 & -3 & -9 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 & -2 & 5 \\ 0 & 1 & 1 & -1 \\ 0 & 0 & -3 & -9 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & -2 & 5 \\ 0 & 1 & 1 & -1 \\ 0 & 0 & 1 & 3 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 0 & 11 \\ 0 & 1 & 0 & -4 \\ 0 & 0 & 1 & 3 \end{pmatrix} \qquad x = \begin{pmatrix} 11 \\ -4 \\ 3 \end{pmatrix}$$

Reduced Row Echelon Form

Solution

Iterative Methods for Linear Systems

Jacobi Method

Gauss-Siedel Method

Apparently, Gauss had very little to do with this but did mention it to a student in a private letter!

Applies only to matrix equations with **non-zero elements on the diagonals**.
Convergence of the iteration is guaranteed if

(a) the matrix is either **strictly diagonally dominant** (*diagonal term is larger than the sum of the non-diagonal terms*) or

(b) the matrix is **symmetric and positive definite** (*symmetric and all its eigenvalues are positive, or equivalently, all its pivots are positive*)

System of Equations

$$a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1$$
$$a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2$$
$$\vdots$$
$$a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = b_n$$

Initial Guess

$$x^{(0)} = \begin{bmatrix} x_1^{(0)} \\ x_2^{(0)} \\ \vdots \\ x_n^{(0)} \end{bmatrix}$$

Starting with the initial guess $x^{(0)}$, use the equations to iteratively calculate better and better estimates – $x^{(0)}$, $x^{(1)}$, $x^{(2)}$, ...until the results are sufficiently accurate!

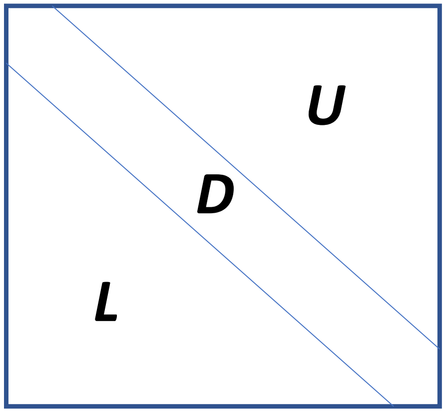
The difference between the two methods is basically that –

- Jacobi's method updates all the values of $x^{(i)}$ before starting on the values of $x^{(i+1)}$
- Gauss-Siedel starts using an updated component of $x^{(i)}$ immediately (i.e. for updating later components of $x^{(i)}$)

Jacobi's method is generally a little slower but has one major advantage in our modern day and age!
What would you imagine that to be?

System of Equations ***Ax=b***

$$
\begin{aligned}
a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\
a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\
&\vdots \\
a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n &= b_n
\end{aligned}
$$



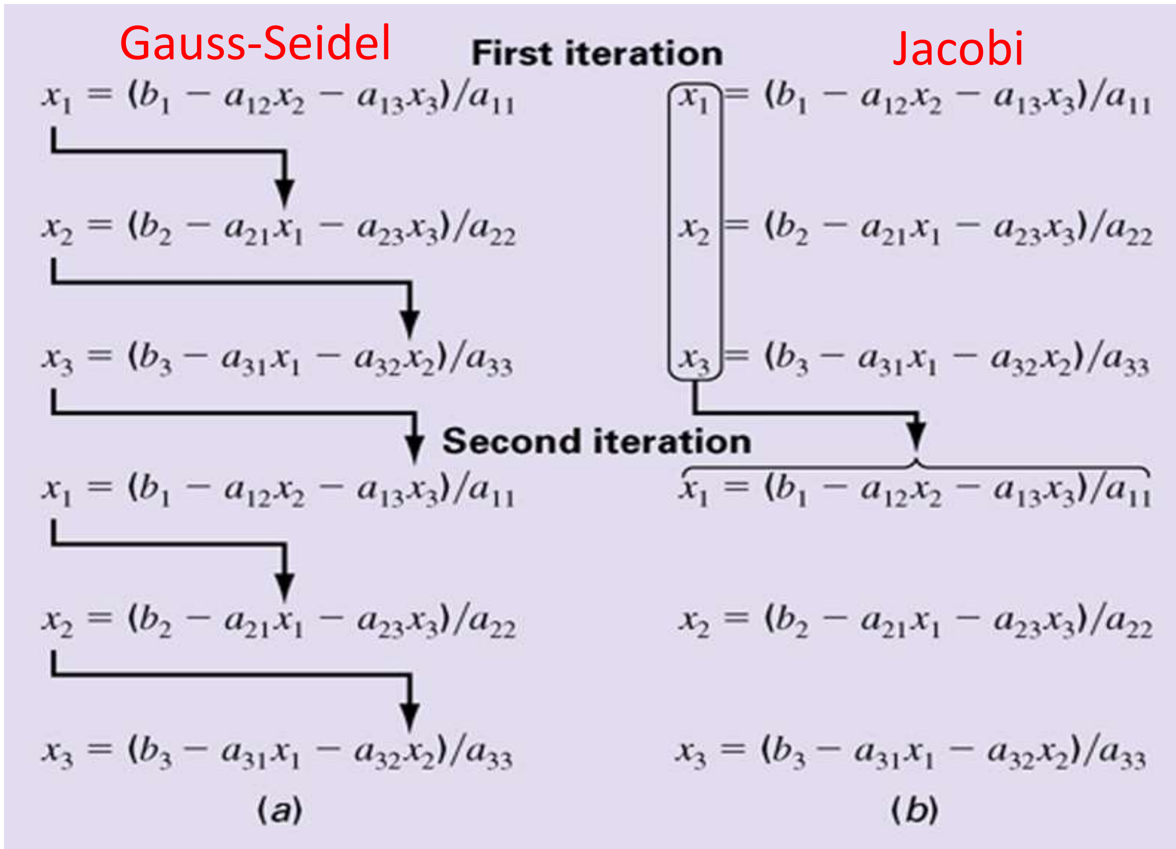
* Organize the equations, possibly by rearranging them, to ensure that the diagonal elements are **non-zero**.

Write ***A*** as the sum of three matrices ***A = D + U + L*** *where* ***D*** is a Diagonal Matrix with non-zero elements, ***L*** is a Lower Triangular Matrix, ***U*** is an Upper Triangular Matrix

* We can then also scale each equation by the corresponding diagonal element to make that 1.

We can then write ***A*** in the form of ***I+U+L*** where ***I*** is an Identity Matrix. Note that ***U, L*** here are different from the ***U, L*** above

Graphical depiction of the difference between (*a*) the Gauss-Seidel and (*b*) the Jacobi iterative methods for solving simultaneous linear algebraic equations.



Example for 3 equations with 3 unknowns

Jacobi Method

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n &= b_n \end{aligned}$$

Start with an Initial Guess

$$x^{(0)} = \begin{bmatrix} x_1^{(0)} \\ x_2^{(0)} \\ \vdots \\ x_n^{(0)} \end{bmatrix}$$

$$x_1^{(1)} = \frac{1}{a_{11}}\left(b_1 - a_{12}x_2^{(0)} - \ldots\ldots - a_{1n}x_n^{(0)}\right)$$

$$x_2^{(1)} = \frac{1}{a_{22}}\left(b_2 - a_{21}x_1^{(0)} - a_{23}x_3^{(0)} - \ldots\ldots - a_{2n}x_n^{(0)}\right)$$

....

$$x_n^{(1)} = \frac{1}{a_{nn}}\left(b_n - a_{n1}x_1^{(0)} - a_{n2}x_2^{(0)} - \ldots\ldots - a_{n,n-1}x_{n-1}^{(0)}\right)$$

In iteration round $(k+1)$

$$x_i^{(k+1)} = \frac{1}{a_{ii}}\left[b_i - \sum_{j=1}^{i-1} a_{ij}x_j^{(k)} - \sum_{j=i+1}^{n} a_{ij}x_j^{(k)}\right]$$

The solution is iteratively refined numerically

Jacobi Method (An Example)

Consider using this to solve the following system of linear equations

$$
\begin{aligned}
E_1 &: 10x_1 - x_2 + 2x_3 &&= 6 \\
E_2 &: -x_1 + 11x_2 - x_3 + 3x_4 &&= 25 \\
E_3 &: 2x_1 - x_2 + 10x_3 - x_4 &&= -11 \\
E_2 &: 3x_2 - x_3 + 8x_4 &&= 15
\end{aligned}
$$

Note that the actual solution of this system is $x = \begin{pmatrix} 1 \\ 2 \\ -1 \\ 1 \end{pmatrix}$

Jacobi Method (Example continued)

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 0 & \frac{1}{10} & -\frac{1}{5} & 0 \\ \frac{1}{11} & 0 & \frac{1}{11} & -\frac{3}{11} \\ -\frac{1}{5} & \frac{1}{10} & 0 & \frac{1}{10} \\ 0 & -\frac{3}{8} & \frac{1}{8} & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} + \begin{pmatrix} \frac{3}{5} \\ \frac{25}{11} \\ -\frac{11}{10} \\ \frac{15}{8} \end{pmatrix}$$

$$x^{(k)} = Tx^{(k-1)} + c \quad k = 1, 2, 3, \ldots..$$

Stopping Criteria as $\dfrac{\left\| x^{(k)} - x^{(k-1)} \right\|}{\left\| x^{(k)} \right\|} <$ *tolerance*

Start with $x^{(0)} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$

Then -

$$x^{(1)} = \begin{pmatrix} 0.6 \\ 2.2727 \\ -1.10 \\ 1.875 \end{pmatrix} \quad x^{(2)} = \begin{pmatrix} 1.0473 \\ 1.7159 \\ -0.8052 \\ 0.8852 \end{pmatrix}$$

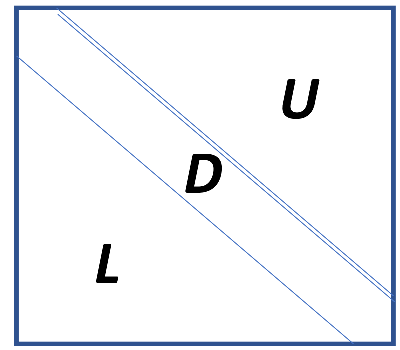
$$x^{(3)} = \begin{pmatrix} 0.9326 \\ 2.0530 \\ -1.0493 \\ 1.1309 \end{pmatrix} \ldots\ldots\ldots \; x^{(10)} = \begin{pmatrix} 1.0001 \\ 1.9998 \\ -0.9998 \\ 0.9998 \end{pmatrix}$$

Matrix Iterative Form of the Jacobi Method

Ax=b

Write ***A*** as the sum of three matrices ***A = D + L + U***

where ***D*** is a Diagonal Matrix
L is a Lower Triangular Matrix
U is an Upper Triangular Matrix



$$Ax = b \quad \Rightarrow \quad (D+L+U)x = b$$

$$x^{(k)} = D^{-1}\left(b-(L+U)x^{(k-1)}\right) = -\left(D^{-1}(L+U)\right)x^{(k-1)} + D^{-1}b$$

or, equivalently

$$x_i^{(k+1)} = \frac{1}{a_{ii}}\left(b_{ii} - \sum_{\substack{j=1 \\ j\neq i}}^{n} a_{ij}x_j^{(k)}\right), \quad i = 1, 2, \ldots\ldots, n$$

Same as before but written differently

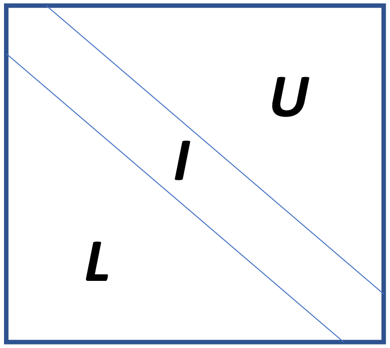
How would these change if we write ***A = I + L + U*** ?

Matrix Iterative Form of the Jacobi Method

Ax=b

If ***A*** written as the sum of three matrices ***A = I + L + U***

where ***I*** is a IdentityMatrix
L is a Lower Triangular Matrix
U is an Upper Triangular Matrix



$$Ax = b \quad \Rightarrow \quad (I + L + U)x = b$$

$$x^{(k)} = b - (L+U)x^{(k-1)} = -(L+U)x^{(k-1)} + b$$

or, equivalently
$$x_i^{(k+1)} = \left(b_{ii} - \sum_{\substack{j=1 \\ j \neq i}}^{n} a_{ij} x_j^{(k)} \right), \quad i = 1, 2, \ldots\ldots, n$$

Gauss-Seidel Iteration Method *Ax=b*

Example

$$
\begin{aligned}
x_1 - 0.25x_2 - 0.25x_3 \qquad &= 50 \\
-0.25x_1 + x_2 - \qquad 0.25x_4 &= 50 \\
-0.25x_1 + \qquad x_3 - 0.25x_4 &= 25 \\
-0.25x_2 - 0.25x_3 + x_4 &= 25
\end{aligned}
\quad \Rightarrow \quad
\begin{aligned}
x_1 &= \qquad 0.25x_2 + 0.25x_3 \qquad + 50 \\
x_2 &= 0.25x_1 + \qquad\qquad 0.25x_4 + 50 \\
x_3 &= 0.25x_1 + \qquad\qquad 0.25x_4 + 25 \\
x_4 &= \qquad 0.25x_2 + 0.25x_3 \qquad + 25
\end{aligned}
$$

Start with an Initial Guess

$$x_1^{(0)} = 100 \quad x_2^{(0)} = 100 \quad x_3^{(0)} = 100 \quad x_4^{(0)} = 100$$

Then

$$
\begin{aligned}
x_1^{(1)} &= 0.25x_2^{(0)} + 0.25x_3^{(0)} + 50 = 100 \\
x_2^{(1)} &= 0.25x_1^{(1)} + 0.25x_4^{(0)} + 50 = 100 \\
x_3^{(1)} &= 0.25x_1^{(1)} + 0.25x_4^{(0)} + 25 = 75 \\
x_4^{(1)} &= 0.25x_2^{(1)} + 0.25x_3^{(1)} + 25 = 68.75
\end{aligned}
\qquad
\begin{aligned}
x_1^{(2)} &= 0.25x_2^{(1)} + 0.25x_3^{(1)} + 50 = 93.75 \\
x_2^{(2)} &= 0.25x_1^{(2)} + 0.25x_4^{(1)} + 50 = 90.62 \\
x_3^{(2)} &= 0.25x_1^{(2)} + 0.25x_4^{(1)} + 25 = 65.62 \\
x_4^{(1)} &= 0.25x_2^{(2)} + 0.25x_3^{(2)} + 25 = 64.06
\end{aligned}
$$

Gauss-Seidel Iteration Method (Matrix Form)

We assume that a_{jj}=1 for j=1,..., n by rearranging the equations so that no diagonal coefficient is zero and then divide each equation by its corresponding diagonal coefficient.

Write ***A*** as the sum of three matrices ***A = I + L + U***

Ax=b ***(I+L+U)x=b*** ***x=b-Lx-Ux***

If we ensure that below the main diagonal we take the NEW approximations and above the main diagonal the OLD approximations, then we can write the iterations as –

$$x^{(m+1)} = b - Lx^{(m+1)} - Ux^{(m)} \qquad a_{jj} = 1$$

where $x^{(m)} = \left[x_j^{(m)} \right]$ is the m^{th} approximation
and $x^{(m+1)} = \left[x_j^{(m+1)} \right]$ is the $(m+1)^{th}$ approximation

See algorithm in the next slide

Algorithm for Gauss-Seidel Iteration (A, b, $x^{(m)}$, ε, *N*) A is a *nxn* matrix with $a_{jj}\neq 0$

INPUT: **A, b,** initial approximation $\mathbf{x}^{(0)}$, tolerance ε>0, maximum number of iterations *N*

OUTPUT: Approximate solution $\mathbf{x}^{(m)}=\left[x_j^{(m)}\right]$ or failure message that $\mathbf{x}^{(N)}$ does not satisfy the required tolerance

For *m* = 0,….. *N*-1 do:

For *j* = 1, …….., *n,* do:

$$x_j^{(m+1)} = \frac{1}{a_{jj}}\left(b_j - \sum_{k=1}^{j-1} a_{jk}x_k^{(m=1)} - \sum_{k=j+1}^{n} a_{jk}x_k^{(m)}\right)$$

End

If $\max_j \left|x_j^{(m+1)} - x_j^{(m)}\right| < \varepsilon,$ then OUTPUT $\boldsymbol{x}^{(m+1)}$, STOP; *Procedure completed successfully*

END

OUTPUT: No solution obtained satisfying the given tolerance condition, *Procedure completed unsuccessfully*

Question: For how long should an iteration be continued?

Answer: Typically, a **Tolerance Bound** and a **Maximun Number of Iterations** will be given. The iterations should be continued until the Tolerance Bound is met (Success!) or the iteration count reaches the maximum number of iterations (Failure!)

As an example, consider the previous slide where the Tolerance Bound was ε and *N* was the maximum number of iterations. The iterations continue until either the norm of the error in the solution became less than ε, i.e.

$$\max_j \left| x_j^{(m+1)} - x_j^{(m)} \right| < \varepsilon$$

Absolute Value of the maximum change in any one component of the ***x*** vector from one iteration to the next is less than tolerance

or the iteration count reached its limit *N*

Normalizing this may be a better option

Convergence of the Iterative Procedure

An iteration method for solving **Ax=b** is said to converge for an initial $\mathbf{x}^{(0)}$ if the corresponding iterative sequence $\mathbf{x}^{(0)}$, $\mathbf{x}^{(1)}$, $\mathbf{x}^{(2)}$ .... converges to a solution of the given system.

For an iteration of the type $\mathbf{x}^{(m+1)}=\mathbf{C}\mathbf{x}^{(m)}+ (\ldots)$ converges is guaranteed for every $\mathbf{x}^{(0)}$ if and only if the eigenvalues of the matrix $\mathbf{C} = [c_{jk}]$ have absolute values less than 1.

If the **spectral radius** of **C** = maximum of the absolute values of the eigenvalues is small then the convergence is rapid

Sufficient Convergence Condition is that the Matrix Norm of **C** is less than 1, $\|C\| < 1$

Frobenius Norm $\quad \|C\| = \sqrt{\sum_{j=1}^{n}\sum_{k=1}^{n} c_{jk}^2}$

Column "Sum" Norm $\quad \|C\| = \max_k \sum_{j=1}^{n} |c_{jk}|$

Row "Sum" Norm $\quad \|C\| = \max_j \sum_{k=1}^{n} |c_{jk}|$

A **sufficient condition** for the Jacobi or Gauss-Seidel iterations to converge is that the absolute value of the diagonal element of each row of ***A*** is larger than the sum of the absolute value of the other elements in the row. Matrix ***A*** is then said to be **Diagonally Dominant.**

In some cases, the iterations may still converge even though this condition is not satisfied.

Useful Properties of Norms

$$\|\vec{x}\| \geq 0 \qquad \forall \vec{x}$$

$$\|\vec{x}\| = 0 \qquad \Leftrightarrow \vec{x} = \vec{0}$$

$$\|\alpha\vec{x}\| = |\alpha|\|\vec{x}\| \quad \forall\ \alpha \in R, \forall\ \vec{x} \in R^n$$

$$\|\vec{x} + \vec{y}\| \leq \|\vec{x}\| + \|\vec{y}\| \quad \forall\ \vec{x}, \vec{y} \in R^n$$

Reading Assignments:

(i) Practical challenges in Gauss elimination:
Pivoting: https://235d9ee8-8e8c-4d7b-a842-264ad94cf102.filesusr.com/ugd/334434_d627cffea18f4f81b195c15b37ec990e.pdf
Partial pivoting: https://235d9ee8-8e8c-4d7b-a842-264ad94cf102.filesusr.com/ugd/334434_e58e6bc8907f486b9de1253d60c52e3a.pdf

(ii) Arithmetic complexity of Gauss elimination: https://235d9ee8-8e8c-4d7b-a842-264ad94cf102.filesusr.com/ugd/334434_5a3eab64a8b0442cabd729aa5defab45.pdf

(iii) **Gauss Jordan elimination**: (similar but relies on *reduced row echelon form* instead of simply the row echelon form that is used in Gauss Elimination)

(iv) **LU decomposition**: https://235d9ee8-8e8c-4d7b-a842-264ad94cf102.filesusr.com/ugd/334434_5a3eab64a8b0442cabd729aa5defab45.pdf