# 7.4 Linear Independence. Rank of a Matrix. Vector Space

## Linear Independence/Dependence of Vectors

- Given any set of $m$ vectors $\mathbf{a}_{(1)}, \dots, \mathbf{a}_{(m)}$ (of same size), a **linear combination** of these vectors is of the form

$$c_1\mathbf{a}_{(1)} + c_2\,\mathbf{a}_{(2)} + \dots + c_m\mathbf{a}_{(m)}$$

where $c_1, c_2, \dots, c_m$ are any scalars.

- Now consider the equation

Zero vector of same size as $\mathbf{a}_{(j)}$

**(1)** $$c_1\mathbf{a}_{(1)} + c_2\,\mathbf{a}_{(2)} + \dots + c_m\mathbf{a}_{(m)} = \mathbf{0}$$

Can we find scalars $c_1, c_2, \dots, c_m$, to make (1) holds?

  - Trivial solution: $c_j$=0 for $j$=1,…,$m$, NOT interesting
  - **Interested in non-trivial solution:** some of the scalars $c_1, c_2, \dots, c_m$, nonzero

- **Question**: can we find non-trivial solution for scalars $c_1, c_2, \dots, c_m$, to satisfy (1)?

**Question**: can we find non-trivial solution for scalars $c_1, c_2, \ldots, c_m$, to satisfy (1)?

**No**

- Vectors $\mathbf{a}_{(1)}, \ldots, \mathbf{a}_{(m)}$ are **linearly independent** (or they form a *linear independent set*).
- Only solution to (1) is the trivial solution, i.e., all $c_j$'s are zero

**Yes**

- Vectors $\mathbf{a}_{(1)}, \ldots, \mathbf{a}_{(m)}$ are **linearly dependent** (or they form a *linearly dependent set*).
- A set of vectors that includes the **zero vector** must be **linearly dependent**. (Why?)

- For **linearly dependent** case, i.e., a non-trivial solution of $c_j$'s exists for (1), we have,
  - At least one of the vectors $\mathbf{a}_{(1)}, \ldots, \mathbf{a}_{(m)}$ can be expressed as a linear combination of other vectors, e.g., if (1) holds for some nonzero $c_j$'s, say, $c_1$, then $\mathbf{a}_{(1)} = k_2\,\mathbf{a}_{(2)} + \ldots + k_m\mathbf{a}_{(m)}$ where $k_j = -c_j / c_1$. (Hence the term linearly dependent)
  - We can remove $\mathbf{a}_{(1)}$ from the set without losing "content" or "information".
- Hence, if a set of vectors is **linearly dependent**, then we can reduce it (get rid of at least one, or perhaps more, if possible) until we get a **linearly independent** set.
- A **linearly independent** set is the smallest "truly essential" set – it cannot be reduced further without losing "content" or "information".

- **Example 1.5**: Given a set of vectors $\mathbf{a}_{(1)}, \ldots, \mathbf{a}_{(4)}$

$$\mathbf{a}_{(1)} = [1 \quad 1 \quad 0 \quad 1];$$
$$\mathbf{a}_{(2)} = [0 \quad 1 \quad 1 \quad 1];$$
$$\mathbf{a}_{(3)} = [1 \quad 0 \quad 0 \quad 2];$$
$$\mathbf{a}_{(4)} = [1 \quad 3 \quad 2 \quad 3]$$

  - The set of vectors is **linearly dependent** because we can find scalars $c_1=1$, $c_2=2$, $c_3=0$ , $c_4=-1$, such that

(1′) $$\mathbf{a}_{(1)} + 2\,\mathbf{a}_{(2)} - \mathbf{a}_{(4)} = \mathbf{0}$$

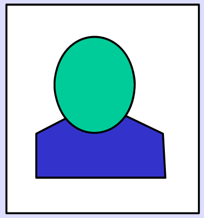
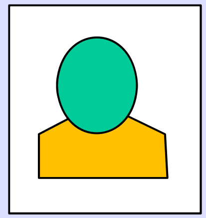
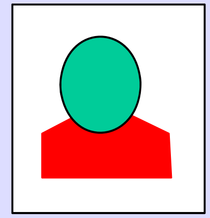
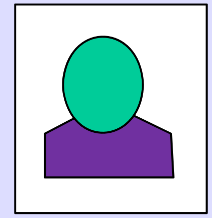
  - From (1′), we can express of one of the vectors of $\mathbf{a}_{(1)}$, $\mathbf{a}_{(2)}$, or $\mathbf{a}_{(4)}$ in terms of the other two.
  - If we choose to take out $\mathbf{a}_{(4)}$, then the reduced set $\{\mathbf{a}_{(1)}, \mathbf{a}_{(2)}, \mathbf{a}_{(3)}\}$ is **linearly independent** because now the equation $c_1\mathbf{a}_{(1)} + c_2\,\mathbf{a}_{(2)} + c_3\mathbf{a}_{(3)} = \mathbf{0}$ will only have trivial solution of all $c_j=0$!

- We can also choose to take out $\mathbf{a}_{(1)}$ and keep $\mathbf{a}_{(4)}$. So the reduced set $\{\mathbf{a}_{(2)}, \mathbf{a}_{(3)}, \mathbf{a}_{(4)}\}$ is **linearly independent**.
- Or, keep $\mathbf{a}_{(1)}$ and take out $\mathbf{a}_{(2)}$. So the reduced set $\{\mathbf{a}_{(1)}, \mathbf{a}_{(3)}, \mathbf{a}_{(4)}\}$ is **linearly independent**.
- In any case, $\mathbf{a}_{(3)}$ must be kept in the linearly independent set of 3 vectors.
- In conclusion, (1′) implies that one of the $\mathbf{a}_{(1)}$, $\mathbf{a}_{(2)}$, $\mathbf{a}_{(4)}$ can be taken out and keeping the remaining two with $\mathbf{a}_{(3)}$ will yield a **linearly independent** set.
- Note that in this example, 3 is the smallest number of vectors (properly selected) to possibly retains all "content" or "information" of the original set of 4 vectors.

- From another perspective, the 4 vectors $\{\mathbf{a}_{(1)}, \mathbf{a}_{(2)}, \mathbf{a}_{(3)}, \mathbf{a}_{(4)}\}$ forms a **linearly dependent set**.
- By proper selection, we can pick 3 vectors out of the 4 to form a **linearly independent set.**
- Going further, we can pick 2 vectors out of the 4 to form **linearly independent set**: $\{\mathbf{a}_{(1)}, \mathbf{a}_{(2)}\}$, $\{\mathbf{a}_{(1)}, \mathbf{a}_{(3)}\}$, $\{\mathbf{a}_{(1)}, \mathbf{a}_{(4)}\}$, $\{\mathbf{a}_{(2)}, \mathbf{a}_{(3)}\}$, $\{\mathbf{a}_{(2)}, \mathbf{a}_{(4)}\}$, and $\{\mathbf{a}_{(3)}, \mathbf{a}_{(4)}\}$.
- We can pick 1 vector out of the original 4 to form **linearly independent set**: $\{\mathbf{a}_{(1)}\}$, $\{\mathbf{a}_{(2)}\}$, $\{\mathbf{a}_{(3)}\}$, $\{\mathbf{a}_{(4)}\}$.
- Hence for this example, 3 is the maximum number of **linearly independent** vectors possibly extracted from the original 4 vectors.
- Recall that 3 is also the smallest number of **linearly independent** vectors to possibly contain the same "content" and "information" of the original set.

- Idea similar to picking agents of MI team:

Top Secret: Faces not to be shown

| | Agent #1 | Agent #2 | Agent #3 | Agent #4 |
|---|---|---|---|---|
| |  |  |  |  |
| Expertise: | Speed car<br>Explosive | Explosive<br>Computer | Speak Russian<br>Fast with knife | Computer<br>Speed car |

- **Linearly independent** set is the team with smallest number of members that keeps all the "skill" set (to achieve the mission)
- In this case, we do not need all 4 agents. We can just pick any two out of Agent #1, #2, and #4.
- However, Agent #3 must be kept.

- **Example 2**: Given a set of vectors $\mathbf{a}_{(1)}$, $\mathbf{a}_{(2)}$, $\mathbf{a}_{(3)}$

$$\mathbf{a}_{(1)} = [3 \quad 0 \quad 2 \quad 2];$$
$$\mathbf{a}_{(2)} = [-6 \quad 42 \quad 24 \quad 54];$$
$$\mathbf{a}_{(3)} = [21 \quad -21 \quad 0 \quad -15]$$

  - The set of vectors is **linearly dependent** because we can find non-trivial solution of scalars $c_1$=6, $c_2$=-0.5, $c_3$=-1, such that

$$6\mathbf{a}_{(1)} - ½\,\mathbf{a}_{(2)} - \mathbf{a}_{(4)} = \mathbf{0}$$

  - We can express of one of the vectors of $\mathbf{a}_{(1)}$, $\mathbf{a}_{(2)}$, or $\mathbf{a}_{(3)}$ in terms of the other two.
  - Maximum number of **linearly independent** vector is 2.

# Definition: Rank of a Matrix

The **rank** of a matrix **A** is the *maximum* number of linearly independent row vectors of **A**. It is denoted by rank **A**.

- Example: Let the vectors $\mathbf{a}_{(1)}, \ldots, \mathbf{a}_{(4)}$ of Example 1.5 form the rows of matrix **A**, i.e.,

$$\mathbf{A} = \begin{bmatrix} \mathbf{a}_{(1)} \\ \mathbf{a}_{(2)} \\ \mathbf{a}_{(3)} \\ \mathbf{a}_{(4)} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 2 \\ 1 & 3 & 2 & 3 \end{bmatrix}$$

  – Example 1.5 → matrix **A** can have 1, 2, or 3 linearly independent row, but 3 is the *maximum* number of linearly independent row vectors it can have
  – Rank **A** = 3

- Recall: a matrix $\mathbf{A}_1$ is **row-equivalent** to another matrix $\mathbf{A}_2$ if $\mathbf{A}_1$ can be obtained from $\mathbf{A}_2$ by (finitely many!) elementary row operations.
- Elementary row operations do **not** change the maximum number of linearly independent row vectors of a matrix.

**Theorem 1:**

**Row-equivalent matrices have the same rank.**

- **Idea**: we can (easier!) determine the rank of a matrix by converting it to row echelon form and determine the rank of the resulting matrix (which is the number of nonzero rows).

- **Example 1.5 revisited**: Matrix **A** formed by row vectors $\mathbf{a}_{(1)}, \ldots, \mathbf{a}_{(4)}$

$$\mathbf{A} = \begin{bmatrix} \mathbf{a}_{(1)} \\ \mathbf{a}_{(2)} \\ \mathbf{a}_{(3)} \\ \mathbf{a}_{(4)} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 2 \\ 1 & 3 & 2 & 3 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & -1 & 0 & 1 \\ 0 & 2 & 2 & 2 \end{bmatrix} \begin{array}{l} \\ \\ \text{Row 3 – Row 1} \\ \text{Row 4 – Row 1} \end{array}$$

$$\Rightarrow \begin{bmatrix} 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix} = \mathbf{R} \begin{array}{l} \\ \\ \text{Row 3 + Row 2} \\ \text{Row 4 – 2 Row 2} \end{array}$$

  - **R** is in row echelon form, and rank **A** = rank **R** =3. (Maximum number of linearly independent row)

## EXAMPLE 3 Determination of Rank

Given the matrix (recall previous Example 2):

$$\mathbf{A} = \begin{bmatrix} 3 & 0 & 2 & 2 \\ -6 & 42 & 24 & 54 \\ 21 & -21 & 0 & -15 \end{bmatrix} \qquad \text{(given)}$$

$$\Rightarrow \begin{bmatrix} 3 & 0 & 2 & 2 \\ 0 & 42 & 28 & 58 \\ 0 & -21 & -14 & -29 \end{bmatrix} \qquad \begin{matrix} \text{Row 2} + 2\ \text{Row 1} \\ \text{Row 3} - 7\ \text{Row 1} \end{matrix}$$

$$\Rightarrow \begin{bmatrix} 3 & 0 & 2 & 2 \\ 0 & 42 & 28 & 58 \\ 0 & 0 & 0 & 0 \end{bmatrix} = \mathbf{R} \qquad \text{Row 3} + \frac{1}{2}\ \text{Row 2.}$$

- **R** is in row echelon form, and rank **A** = rank **R** =2.

- The rank of a matrix in row echelon form is the number of nonzero rows.
- Why? Illustrate with **EXAMPLE 3**: The maximum number of linearly independent row vector in **R**?

$$\mathbf{R} = \begin{bmatrix} 3 & 0 & 2 & 2 \\ 0 & 42 & 28 & 58 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

  - Discard the zero row as including it will make any set of vectors linearly dependent
  - Is the remaining two nonzero row vectors **linearly independent**? Try to find constant $c_1$ and $c_2$ such that

$$\begin{aligned} & c_1[3 \quad 0 \quad 2 \quad 2] \\ + & c_2[0 \quad 42 \quad 28 \quad 58] \\ = & [0 \quad 0 \quad 0 \quad 0] \end{aligned}$$

$\Rightarrow$ **R** in REF makes it easy to see $c_1$ =0 and $c_2$ =0

  - Hence rank **R** = 2.

## Theorem 2: Determining the Linearly Independency or Dependency using matrix rank

- Given a set of vectors, we can put them as rows of a matrix and determine the rank of the resulting matrix.
- Specifically, consider $p$ vectors that each have $n$ components.
  - If the matrix formed with them as row vectors has rank $p$ → These vectors are **linearly independent**.
  - If the matrix formed has rank $< p$ → These vectors are **linearly dependent**

**Theorem 3** Matrices **A** and $\mathbf{A}^{\mathsf{T}}$ have same rank.

- The maximum number of linearly independent rows and columns in a matrix are the same.
- Hence, rank **A** also equals the maximum number of linearly independent column vectors of **A**.

- Why? Illustrate with an example

**Example:**

– Matrix $\mathbf{A} = \begin{bmatrix} 3 & 0 & 2 & 2 \\ -6 & 42 & 24 & 54 \\ 21 & -21 & 0 & -15 \end{bmatrix}$ has rank 2

– This means that the maximum number of linearly independent row vectors of **A** is 2, in fact,

$$6\ \text{Row1} - 0.5\ \text{Row2} - \text{Row3} = 0$$

## **Example** *(continued)* **:**

- Choose Row1 and Row2 as linearly independent set.
- All rows of **A** can be expressed in terms of Row1 and Row2:

$$\begin{bmatrix}3 & 0 & 2 & 2\end{bmatrix} = 1\times\begin{bmatrix}3 & 0 & 2 & 2\end{bmatrix} + 0\times\begin{bmatrix}-6 & 42 & 24 & 54\end{bmatrix}$$

$$\begin{bmatrix}-6 & 42 & 24 & 54\end{bmatrix} = 0\times\begin{bmatrix}3 & 0 & 2 & 2\end{bmatrix} + 1\times\begin{bmatrix}-6 & 42 & 24 & 54\end{bmatrix}$$

$$\begin{bmatrix}21 & -21 & 0 & -15\end{bmatrix} = 6\times\begin{bmatrix}3 & 0 & 2 & 2\end{bmatrix} - \tfrac{1}{2}\times\begin{bmatrix}-6 & 42 & 24 & 54\end{bmatrix}$$

- Now consider the column of matrix **A**:

$$\begin{bmatrix}3\\-6\\21\end{bmatrix} = 3\times\begin{bmatrix}1\\0\\6\end{bmatrix} - 6\times\begin{bmatrix}0\\1\\-\tfrac{1}{2}\end{bmatrix} \qquad \begin{bmatrix}2\\24\\0\end{bmatrix} = 2\times\begin{bmatrix}1\\0\\6\end{bmatrix} + 24\times\begin{bmatrix}0\\1\\-\tfrac{1}{2}\end{bmatrix}$$

$$\begin{bmatrix}0\\42\\-21\end{bmatrix} = 0\times\begin{bmatrix}1\\0\\6\end{bmatrix} + 42\times\begin{bmatrix}0\\1\\-\tfrac{1}{2}\end{bmatrix} \qquad \begin{bmatrix}2\\54\\-15\end{bmatrix} = 2\times\begin{bmatrix}1\\0\\6\end{bmatrix} + 54\times\begin{bmatrix}0\\1\\-\tfrac{1}{2}\end{bmatrix}$$

- Maximum number of linearly independent columns is also 2.

If a given matrix **A** of size *m* by *n:*

$$\operatorname{rank} \mathbf{A} \le m$$
$$\operatorname{rank} \mathbf{A}^{\mathrm{T}} \le n$$
$$\operatorname{rank} \mathbf{A} = \operatorname{rank} \mathbf{A}^{\mathrm{T}} \le \min(m, n)$$

Example: *m*=3, *n*=5

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 1 & 2 & 7 \\ 1.2 & 3 & 1 & 8 & 6 \\ 2 & 3 & 4 & 3 & 5 \end{bmatrix} \rightarrow \operatorname{rank} \mathbf{A} \le 3$$

Example: *m*=4, *n*=2

$$\mathbf{A} = \begin{bmatrix} 1 & 3 \\ 1 & 4 \\ 2.2 & 4 \\ 6 & 3 \end{bmatrix} \rightarrow \operatorname{rank} \mathbf{A} \le 2$$

## Theorem 4: Linear Dependence of Vectors

Consider $p$ vectors each having $n$ components. If $n < p$, then these vectors must be **linearly dependent**.

Reasoning:

- Form matrix **A** with the $p$ vectors.
- **A** is either $p$ by $n$ or $n$ by $p$, depending on whether the given vectors are row or column vectors
- Matrix **A** has rank(**A**)<$min(n,p)$=$n$ (given that $n < p$)
- **A** has at most $n$ linearly independent vectors
- The given $p>n$ vectors must be linearly dependent

**Example:** Given $p$ =5 vectors each with $n$=3 components.

$$\mathbf{a}_{(1)} = [1 \quad 1 \quad 2]; \; \mathbf{a}_{(2)} = [3 \quad 2 \quad 1]; \; \mathbf{a}_{(3)} = [3 \quad 4 \quad 9];$$

$$\mathbf{a}_{(4)} = [100 \quad 23 \quad 4]; \; \mathbf{a}_{(5)} = [2 \quad 8 \quad 2];$$

$n < p$ → Given set of vectors linearly dependent!

# Vector Space

Given a Vector Set
$(\mathbf{a}_{(1)}, \ldots, \mathbf{a}_{(p)})$

**V**

The set of vectors generated by taking linear combination: $\mathbf{a} = c_1\mathbf{a}_{(1)} + c_2\,\mathbf{a}_{(2)} + \ldots + c_p\mathbf{a}_{(p)}$ for any scalar values of $c_1, c_2, \ldots, c_p$

**The Span of the Vector Set** $(\mathbf{a}_{(1)}, \ldots, \mathbf{a}_{(p)})$

- We say that the given set of vectors *spans V*.
- *V*, the set of all linear combinations of given vectors $\mathbf{a}_{(1)}, \ldots, \mathbf{a}_{(p)}$, is called the **span** of these vectors.
- A span is a vector space.

# Definition: Vector Space

Consider a nonempty set $V$ of vectors (with same number of components) satisfying the following conditions:

- Given any two vectors **a** and **b** in $V$, then all their linear combinations $\alpha\mathbf{a} + \beta\mathbf{b}$ ($\alpha$, $\beta$ any real numbers) are also elements of $V$,
- Addition between vectors **a**, **b**, **c** in $V$ satisfying (see the laws (3) in Sec. 7.1):
  - (a) $\mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a}$
  - (b) $(\mathbf{a} + \mathbf{b}) + \mathbf{c} = \mathbf{a} + (\mathbf{b} + \mathbf{c})$
  - **(c)** $\mathbf{a} + \mathbf{0} = \mathbf{a}$ (Existence of **0** vector)
  - (d) $\mathbf{a} + (-\mathbf{a}) = \mathbf{0}$ (Existence of (**-a**) vector)
- Scalar multiplication of vectors **a**, **b** in $V$ satisfying (see the laws (4) in Sec. 7.1):
  - (a) $c(\mathbf{a} + \mathbf{b}) = c\mathbf{a} + c\mathbf{b}$ $c$, $k$ any finite real number
  - (b) $(c+k)\mathbf{a} = c\mathbf{a} + k\mathbf{a}$
  - (c) $c(k\mathbf{a}) = (ck)\mathbf{a}$
  - (d) $1\mathbf{a} = \mathbf{a}$

then $V$ is a vector space.

## Example:

Given **Vector Set** ($\mathbf{a}_{(1)}$=[1 0 1], $\mathbf{a}_{(2)}$=[2 0 2], $\mathbf{a}_{(3)}$=[3 0 1])

**V**

$$c_1 = 1, c_2 = 2, c_3 = 0.5 \rightarrow \mathbf{a} = \begin{bmatrix} 6.5 & 0 & 5.5 \end{bmatrix}$$

$$c_1 = 1, c_2 = 0.2, c_3 = 1 \rightarrow \mathbf{a} = \begin{bmatrix} 4.4 & 0 & 2.4 \end{bmatrix}$$

……………….. ……….

……………….. ……….

**The Span of the Vector Set** ($\mathbf{a}_{(1)}$, …, $\mathbf{a}_{(3)}$)

- $V$ contains vectors with special structure $[\ \alpha \quad 0 \quad \beta\ ]$, where scalars $\alpha$ and $\beta$ can be any value
- $V$ satisfies all conditions in the definition

→$V$ is a vector space

- **Note:** Not all set of vectors are vector space. A set must satisfy the conditions in the definition to qualify as Vector Space.

- **Example:** The space of row vector of 3 components:

$$V = \{ [a_1 \quad a_2 \quad a_3] \}$$

Check: $V$ satisfies all the conditions → Vector space!
Note: this means $R^3$ is a Vector space

- **Example:** The space of row vector of 3 components:

$$V_C = \{ [a_1 \quad a_2 \quad a_3], \text{with } a_1 + a_2 + a_3 = 0 \}$$

Check: $V_C$ satisfies all the conditions → Vector space!
In this case, $V_C$ is a subset of $R^3$. We called $V_C$ a subspace of of $R^3$. (More later)

- **Example:** The space of row vector of 3 components:

$$V = \{[a_1 \quad a_2 \quad a_3], \text{with } a_1, a_2, a_3 \text{ all integers } \}$$

Check: $V$ does not satisfy the first condition (and maybe others as well):

Given any two vectors **a** and **b** in $V$, then all their linear combinations $\alpha$**a** + β**b** ($\alpha$, β any real numbers) are also elements of $V$ →NOT a vector space.

- Example: $V = \{[a_1 \quad a_2 \quad a_3], \text{with } a_1 + a_2 = 9\}$

- Example: $V = \{[a_1 \quad a_2 \quad a_3], \text{with } a_1^2 = a_2^2 \}$

- Example:

$$V = \{[a_1 \quad a_2 \quad a_3], \text{with } a_1 + 3a_2 + 3a_3 = 0\}$$

- **Example:** Vector sets $\boldsymbol{S}$=($\mathbf{a}_{(1)}$=[2 5 3], $\mathbf{a}_{(2)}$=[1 0 2]) and $\boldsymbol{T}$=($\mathbf{b}_{(1)}$=[2 0 5], $\mathbf{b}_{(2)}$=[3 5 5]). Find intersection of Span($\boldsymbol{S}$) and Span($\boldsymbol{T}$).
  - Span($\boldsymbol{S}$): $[x\ y\ z] = \alpha[2\ 5\ 3] + \beta[1\ 0\ 2]$
    $\rightarrow -10\,x + y + 5z = 0$
  - Span($\boldsymbol{T}$): $[x\ y\ z] = \alpha[2\ 0\ 5] + \beta[3\ 5\ 5]$
    $\rightarrow\ 5\,x + y - 2z = 0$
  - Intersection of Span($\boldsymbol{S}$) and Span($\boldsymbol{T}$):

$$\left.\begin{aligned} -10\,x + y + 5z &= 0 \\ 5\,x + y - 2z &= 0 \end{aligned}\right\} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = c \begin{bmatrix} 0.6 \\ 1 \\ 1 \end{bmatrix}$$

Any real
finite number

**Reverse question**: given a vector space **V, **find the vector set that spans **V?**

Vector Set ($\mathbf{a}_{(1)}$, … , $\mathbf{a}_{(p)}$) to generate **V** by linear combination

?

Given Vector space **V**

**Answers:** Vector set to generate **V** is not unique (both in values and number)

## Previous Example Revisited:

Given $V$ containing vectors with special structure $[\ \alpha \quad 0 \quad \beta\ ]$, where scalars $\alpha$ and $\beta$ can be any value

**Vector Set to span V?**

**V**

$[6.5 \quad 0 \quad 5.5]$

$[4.4 \quad 0 \quad 2.4]$

…………. ……….

………….. ……….

-- One possible vector set to span $V$ is
S1= ($\mathbf{a}_{(1)=}[1\ 0\ 1]$, $\mathbf{a}_{(2)=}[2\ 0\ 2]$, $\mathbf{a}_{(3)=}[3\ \ 0\ \ 1]$)

-- Another set is S2= ($\mathbf{a}_{(1)=}[1\ 0\ 1]$, $\mathbf{a}_{(2)=}[2\ 0\ 1]$)

-- Or we have S3= ($\mathbf{a}_{(1)=}[1\ 0\ 0]$, $\mathbf{a}_{(2)=}[0\ 0\ 1]$)

-- There are actually infinite # of choices for this vector set!

**Observations from example:**

-- The set S1= ($\mathbf{a}_{(1)=}$[1 0 1], $\mathbf{a}_{(2)=}$[2 0 2], $\mathbf{a}_{(3)=}$[3 0 1]) is not a linear independent set. There is some redundant "information" contained among the three vectors.

-- On the other hand, both S2= ($\mathbf{a}_{(1)=}$[1 0 1], $\mathbf{a}_{(2)=}$[2 0 1]) and S3= ($\mathbf{a}_{(1)=}$[1 0 0], $\mathbf{a}_{(2)=}$[0 0 1]) are linearly independent set. S2 and S3 are called **basis** of **V**.

-- Both S2 and S3 need only two vectors to span **V** compared to three in S1.

-- We call the dimension of **V**: dim **V**=2. This means

* Vector space **V** has two "degrees" of freedom
* Maximum number of vectors in **V** that can be linearly independent to each other is 2. Pick any 3 vectors in **V** must be linearly dependent.
* Pick any two linearly independent vectors from **V** produces a **basis** that spans **V**.

## The Basis of Vector Space

- The maximum number of **linearly independent** vectors in $V$ is called the **dimension** of $V$ and is denoted by dim $V$. (Assume dim $V$ finite; infinite dim $V$ case in Sec. 7.9.)
- A **linearly independent** set in $V$ consisting of a maximum possible number of vectors in $V$ is called a **basis** for $V$. In other words, any largest possible set of linearly independent vectors in $V$ forms basis for $V$.
- If we add one or more vector to the basis, it will become linearly dependent.
- The number of vectors of a **basis** for $V$ equals dim $V$.

# Vector Space (continued)

- Definition **subspace** of a vector space **V**:

  An nonempty subset of **V** (including **V** itself) that forms a vector space with respect to the two algebraic operations (addition and scalar multiplication) defined for the vectors of **V**.

- Example:

  The space of row vector of 3 components:

$$\mathrm{V} = \{[a_1 \quad a_2 \quad a_3], \text{with } a_1 + a_2 + a_3 = 0\}$$

  is a vector space. The vector space

$$\mathrm{V}_1 = \{[a_1 \quad a_2 \quad a_3], \text{with } a_1 + a_2 = 0, a_3 = 0\}$$

  is a vector space (conditions checked) and a <u>subspace</u> of **V** . Both are subspace of $R^3$.

## Previous Example Revisited:

Given **V** containing vectors with special structure $[\ \alpha \quad 0 \quad \beta\ ]$, where scalars $\alpha$ and $\beta$ can be any value

-- Subspaces of **V**:

$V1 = [\ \alpha \quad 0 \quad 0\ ]$,

$V2 = [\ 0 \quad 0 \quad \beta\ ]$,

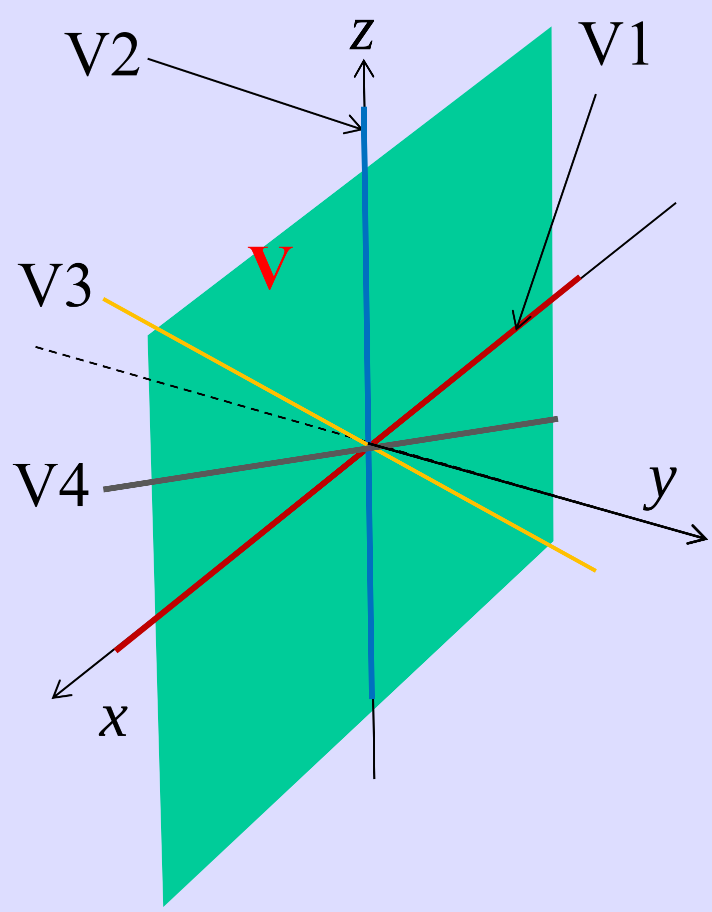
$V3 = \alpha\,[\ 1 \quad 0 \quad 1\ ]$,

where $a_1 = a_3$,

$V4 = \alpha\,[\ 2 \quad 0 \quad 1\ ]$,

where $a_1 = 2a_3$,

…………………..



# Theorem 5

**Vector Space $R^n$**

The space consisting of all vectors with $n$ components ($n$ real numbers) has dimension $n$. is a vector space and denoted by space $\boldsymbol{R}^n$.

$$R^n = \{\mathbf{a} = [a_1 \quad a_2 \quad \cdots \quad a_n], a_j,\ j = 1, \cdots, n \text{ can be any real numbers}\}$$

or

$$R^n = \{\mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}, a_j,\ j = 1, \cdots, n \text{ can be any real numbers}\}$$

# Theorem 6

## Row Space and Column Space

Given a matrix $A$,

- The row space of **A** is the vector space spanned by the rows of $A$.
- The column space of **A** is the vector space spanned by the columns of $A$.
- Because rank $A$= rank $A^T$, the maximum numbers of linearly independent row and column vectors are the same.
- The row and the column spaces of $A$ thus have the same dimension and both equal to rank $A$.

Given a $m$ x $n$ matrix **A**,

$$\mathbf{A} = \left.\begin{bmatrix} \leftarrow \ \mathbf{a}^{r}_{(1)} \ \rightarrow \\ \leftarrow \ \mathbf{a}^{r}_{(2)} \ \rightarrow \\ \vdots \\ \leftarrow \ \mathbf{a}^{r}_{(m)} \ \rightarrow \end{bmatrix}\right. = \underbrace{\begin{bmatrix} \uparrow & \uparrow & & \uparrow \\ \mathbf{a}^{c}_{(1)} & \mathbf{a}^{c}_{(2)} & \cdots & \mathbf{a}^{c}_{(n)} \\ \downarrow & \downarrow & & \downarrow \end{bmatrix}}_{\textbf{span}}$$

**span**

Row Space of **A**
(of row vectors
with $n$ element)

Column Space of **A**
(of column vectors
with $m$ element)

Dim (Row Space) = rank **A** = rank $\mathbf{A}^{T}$ = Dim (Column Space)

## Null Space

Given a matrix **A**, the space of column vectors **x** such that $\mathbf{Ax} = \mathbf{0}$ is called the **null space** of **A** and has the following properties,

- It is a vector space.
- It is the solution space of the homogeneous equation

  $\mathbf{Ax} = \mathbf{0}$.
- Its dimension is called the **nullity** of **A**.
- It satisfies the basic relation (later)

$$\text{rank } \mathbf{A} + \text{nullity } \mathbf{A} = \text{Number of columns of } \mathbf{A}. \tag{6}$$

# 7.5 Solutions of Linear Systems: Existence, Uniqueness

# Theorem 1

**Fundamental Theorem for Linear Systems**

**Recall: Existence of solution**

A linear system of $m$ equations in $n$ unknowns $x_1, \ldots, x_n$

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

can be expressed

$$\mathbf{Ax=b},$$

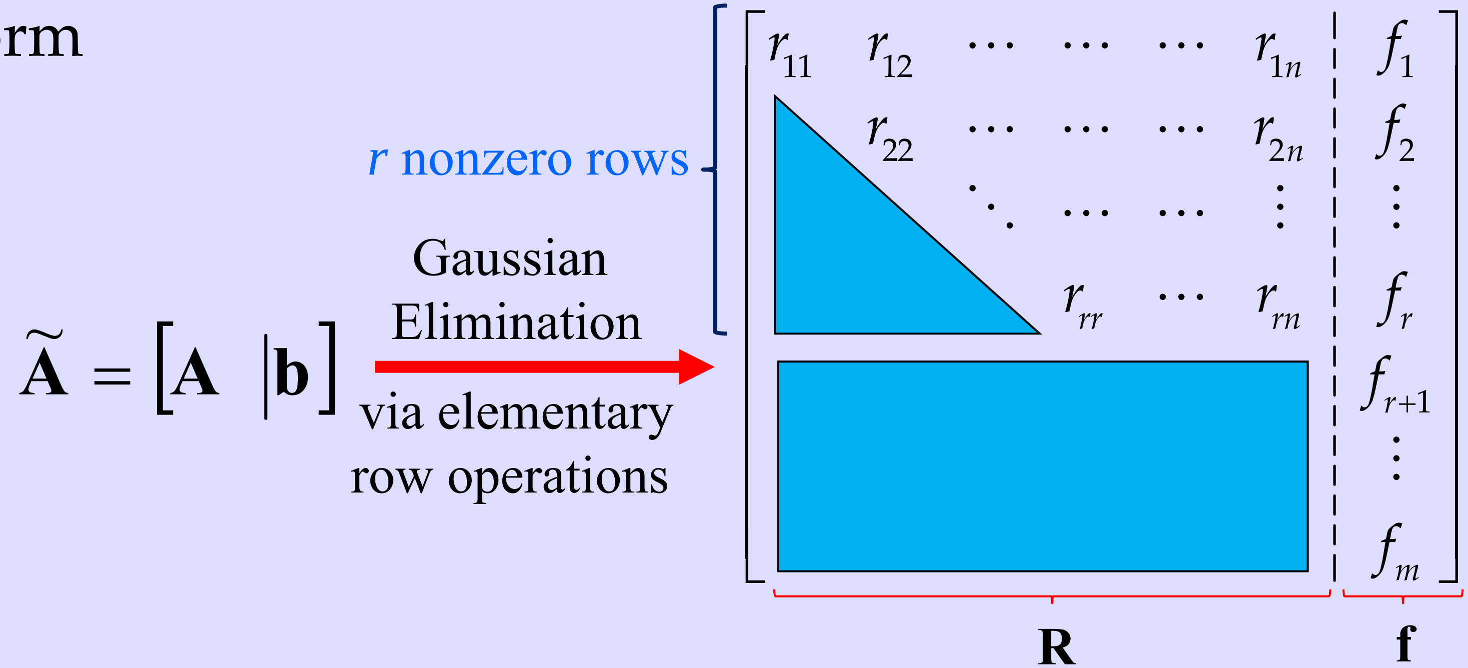
where the coefficient matrix **A** and augmented matrix $\tilde{\mathbf{A}}$ can be defined as

$$\mathbf{A} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \cdot & \cdots & \cdot \\ \cdot & \cdots & \cdot \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \quad and \quad \tilde{\mathbf{A}} = \left[\begin{array}{ccc|c} a_{11} & \cdots & a_{1n} & b_1 \\ \cdot & \cdots & \cdot & \cdot \\ \cdot & \cdots & \cdot & \cdot \\ a_{m1} & \cdots & a_{mn} & b_m \end{array}\right]$$

Existence of solution depend on whether (1) is **consistent** or **inconsistent**.

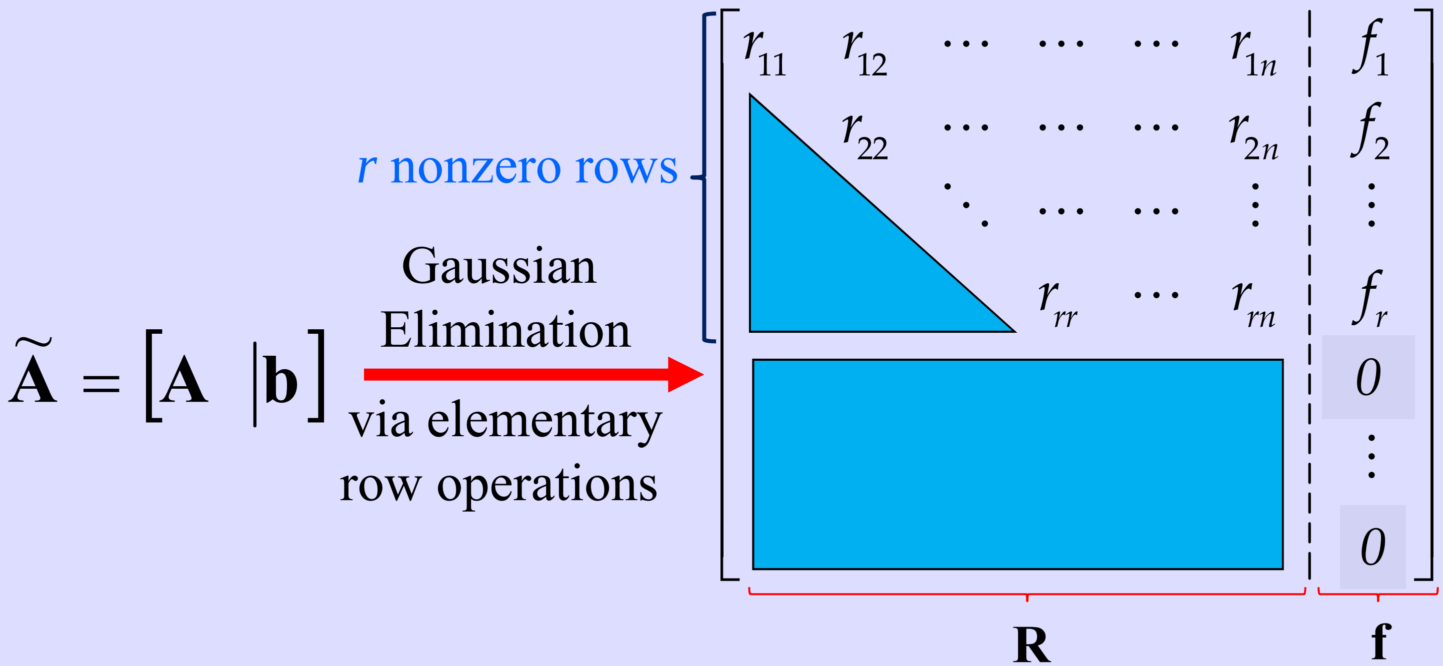
- Previously, we determine **consistency** or **inconsistency** via Gaussian elimination conducted through elementary row operations
- Here, look at the problem using concept of matrix rank.

Upon converting augmented matrix $\tilde{\mathbf{A}}$ into row echelon form




- Gaussian elimination does not change the rank of a matrix, so rank $\tilde{\mathbf{A}}$ = rank [**R** | **f**] and rank **A**=rank **R.** In this case, **R** have $r$ nonzero row, rank **A**=rank **R** = $r$.
- Recall, previously, equations (1) is **inconsistent**, i.e., no solution exist, if any of the $f_{r+1}, f_{r+2}, \ldots, f_m$ is nonzero. This hence implies, (1) is **inconsistent** if $\text{rank } \mathbf{A} \neq \text{rank } \tilde{\mathbf{A}}$.

- On the other hand, system of equations (1) is **consistent**, i.e., solution exist, if $f_{r+1}, f_{r+2}, \ldots, f_m$ are all zero. This hence implies (1) is **consistent** if rank $\mathbf{A}$ = rank $\tilde{\mathbf{A}}$ (= $r$).




- Furthermore, previously we have for consistent case,

(i) Unique solution if number of nonzero equation $r = n$.

(ii) Multiple solutions if number of nonzero equation $r < n$.

- Expressing the above results in terms of matrix rank, we have:
  (a) System (1) is **inconsistent** (i.e., no solution exist) iff rank $\mathbf{A} \neq$ rank $\widetilde{\mathbf{A}}$.

  (b) System (1) is **consistent** (i.e., solution exist) iff rank $\mathbf{A}$ = rank $\tilde{\mathbf{A}}$ **.** Furthermore,

  **(i) Unique Solution:** System (1) has precisely one solution iff this rank $\mathbf{A} = n$.
  **(ii) Multiple solutions:** System (1) has infinitely many solutions iff rank $\mathbf{A} < n$.

- Note for **consistent** case, because rank $\tilde{\mathbf{A}}$ = rank $\mathbf{A}$ we can just use rank $\mathbf{A}$ to describe case **(i)** and **(ii).**

- Also, understanding that rank $\mathbf{A}=r$, the result here is the same as we have previously.

## Further Characterization of solutions to Nonhomogeneous Linear System using Matrix Rank

If a nonhomogeneous linear system (1) **Ax=b** is **consistent**, then all of its solutions can be divided into two components:

$$\mathbf{x} = \mathbf{x}_0 + \mathbf{x}_h$$

where

- $\mathbf{x}_0$ *is any* (fixed) solution of nonhomogeneous system

  **Ax=b**, and

- $\mathbf{x}_h$ correspond to the solution(s) of the corresponding homogeneous system **Ax=0**.

The above applies to both the unique solution case and multiple solution case.

# Recall EXAMPLE 3 Consistent ( Multiple Solution) Case of Section 7.3

$n = 4$

$m = 3$

Free variables of $x_3$ and $x_4$

Solution is:

$$\mathbf{x} = \underbrace{\begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \end{bmatrix}}_{\mathbf{x}_0} + \underbrace{\begin{bmatrix} 0 \\ -1 \\ 1 \\ 0 \end{bmatrix} \alpha + \begin{bmatrix} -1 \\ 4 \\ 0 \\ 1 \end{bmatrix} \beta}_{\mathbf{x}_h}$$

Solution to homogeneous equation $\mathbf{A}\mathbf{x}_h = \mathbf{0}$

Any solution to given inhomogeneous equation $\mathbf{A}\mathbf{x}_0 = \mathbf{b}$

Solve the following linear system of three equations in four unknowns whose augmented matrix is

$$(5)\quad \left[\begin{array}{cccc|c} 3.0 & 2.0 & 2.0 & -5.0 & 8.0 \\ 0.6 & 1.5 & 1.5 & -5.4 & 2.7 \\ 1.2 & -0.3 & -0.3 & 2.4 & 2.1 \end{array}\right]. \qquad \text{Thus,} \qquad \begin{aligned} 3.0x_1 + 2.0x_2 + 2.0x_3 - 5.0x_4 &= 8.0 \\ \boxed{0.6x_1} + 1.5x_2 + 1.5x_3 - 5.4x_4 &= 2.7 \\ \boxed{1.2x_1} - 0.3x_2 - 0.3x_3 + 2.4x_4 &= 2.1. \end{aligned}$$

***Solution.*** As in the previous example, we circle pivots and box terms of equations and corresponding entries to be eliminated. We indicate the operations in terms of equations and operate on both equations and matrices.

***Step 1. Elimination of*** $x_1$ from the second and third equations by adding

$-0.6/3.0 = -0.2$ times the first equation to the second equation,

$-1.2/3.0 = -0.4$ times the first equation to the third equation.

This gives the following, in which the pivot of the next step is circled.

$$(6)\quad \left[\begin{array}{cccc|c} 3.0 & 2.0 & 2.0 & -5.0 & 8.0 \\ 0 & 1.1 & 1.1 & -4.4 & 1.1 \\ 0 & -1.1 & -1.1 & 4.4 & -1.1 \end{array}\right] \begin{array}{l} \\ \text{Row 2} - 0.2 \text{ Row 1} \\ \text{Row 3} - 0.4 \text{ Row 1} \end{array} \qquad \begin{aligned} 3.0x_1 + 2.0x_2 + 2.0x_3 - 5.0x_4 &= 8.0 \\ 1.1x_2 + 1.1x_3 - 4.4x_4 &= 1.1 \\ \boxed{-1.1x_2} - 1.1x_3 + 4.4x_4 &= -1.1. \end{aligned}$$

***Step 2. Elimination of*** $x_2$ from the third equation of (6) by adding

$1.1/1.1 = 1$ times the second equation to the third equation.

This gives

$$(7)\quad \left[\begin{array}{cccc|c} 3.0 & 2.0 & 2.0 & -5.0 & 8.0 \\ 0 & 1.1 & 1.1 & -4.4 & 1.1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right] \begin{array}{l} \\ \\ \text{Row 3} + \text{Row 2} \end{array} \qquad \begin{aligned} 3.0x_1 + 2.0x_2 + 2.0x_3 - 5.0x_4 &= 8.0 \\ 1.1x_2 + 1.1x_3 - 4.4x_4 &= 1.1 \\ 0 &= 0. \end{aligned}$$

***Back Substitution.*** From the second equation, $x_2 = 1 - x_3 + 4x_4$. From this and the first equation, $x_1 = 2 - x_4$. Since $x_3$ and $x_4$ remain arbitrary, we have infinitely many solutions. If we choose a value of $x_3$ and a value of $x_4$, then the corresponding values of $x_1$ and $x_2$ are uniquely determined.

***On Notation.*** If unknowns remain arbitrary, it is also customary to denote them by other letters $t_1, t_2, \cdots$. In this example we may thus write $x_1 = 2 - x_4 = 2 - t_2$, $x_2 = 1 - x_3 + 4x_4 = 1 - t_1 + 4t_2$, $x_3 = t_1$ (first arbitrary unknown), $x_4 = t_2$ (second arbitrary unknown). ■

# Homogeneous Linear System

A homogeneous linear system

$$
\textbf{(4)} \qquad
\begin{aligned}
a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= 0 \\
a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= 0 \\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= 0
\end{aligned}
$$

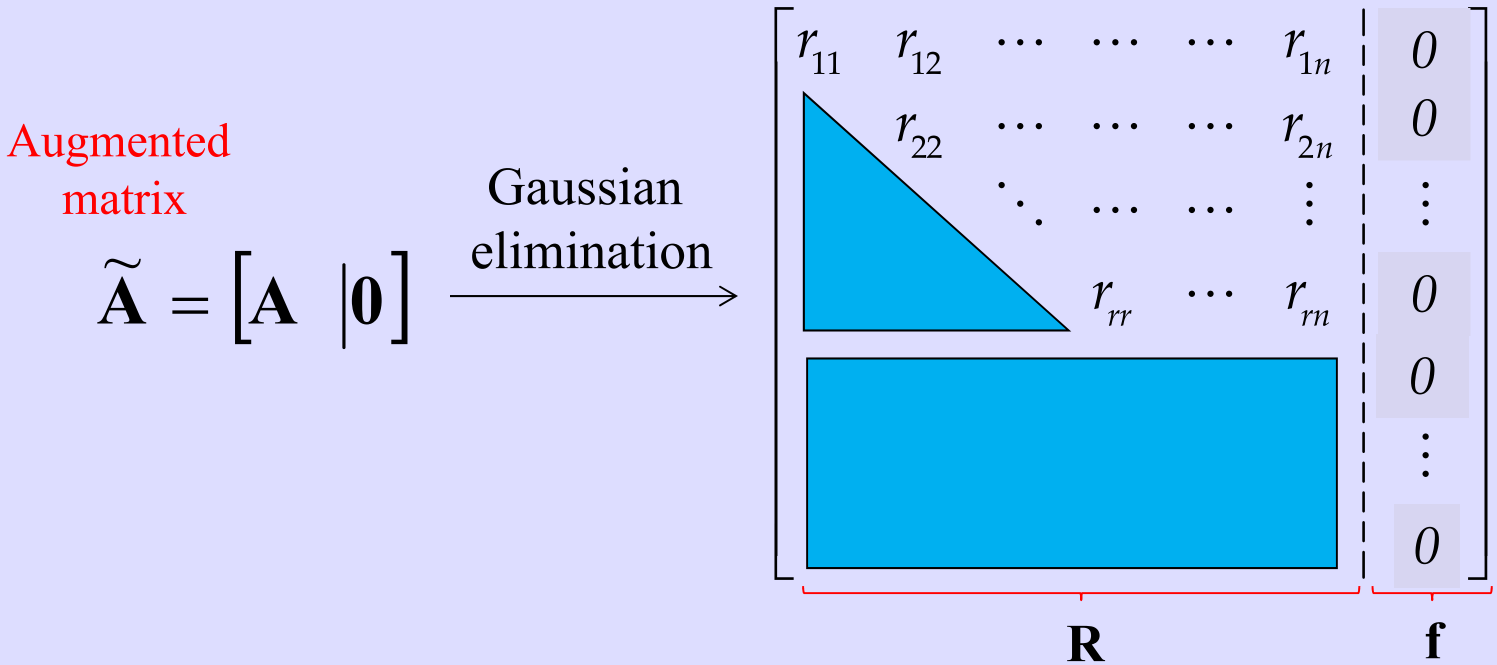
or in matrix form,

$$\mathbf{Ax=0},$$

always has the **trivial solution** $x_1 = 0, \ldots, x_n = 0$, i.e., $\mathbf{x=0}$.

Question is: when will **(4)** have **nontrivial solutions**, i.e., where all or some of the $x_1, \ldots, x_n$, are nonzero?

Using Gaussian elimination, i.e., elementary row operations to solve homogeneous equation **Ax=0**:




Hence, applying the same reasoning as before, we have for the homogeneous linear equation:

- if rank **A**=$n$, **Ax=0** has unique solution, i.e., **x=0**!
- if rank **A** < $n$, **Ax=0** have multiple solutions

In more details:

- If rank **A**=$n$, **Ax=0** has unique solution

  → **trivial solution**, **x=0**, is the only solution for the homogeneous linear system

- If rank **A**=$r < n$, **Ax=0** have multiple solutions

  → **nontrivial solution**, obtained by solving $r$ of the variables $x_1, \ldots, x_n$, in terms of the remaining ($n$-$r$) ***free*** variables (which can be assigned arbitrary values).

  → In this case, the space of the **nontrivial solution** is the linear combination of ($n$-$r$) linearly independent vectors, each of which satisfy **Ax=0.**

# More on Solution to Homogeneous Linear System

$m$ x $n$ matrix

column vector with $n$ entries

column vector with $m$ entries

For a homogeneous linear system $\mathbf{Ax=0}$,

- **Trivial solution** $\mathbf{x_h=0}$ exists iff rank $\mathbf{A} = n$.
- **Nontrivial solutions** exist iff rank $\mathbf{A} < n$. In this case,
  - Solutions of $\mathbf{Ax=0}$ form a vector space: If $\mathbf{x}_{(1)}$ and $\mathbf{x}_{(2)}$ are solution vectors of $\mathbf{Ax=0}$, then $\mathbf{x} = c_1\, \mathbf{x}_{(1)} + c_2\, \mathbf{x}_{(2)}$ with any scalars $c_1$ and $c_2$ is a solution vector of $\mathbf{Ax=0}$ as well and all other conditions satisfied.
  - The vector space as formed is of dimension $(n - r)$. (Why? Because there are $(n$-$r)$ ***free*** variables)

- The solution space of **Ax=0** is also the **null space** of **A** because **Ax** = **0** for every **x** in this solution space.
- The dimension of the **null space** of **A,** called **nullity** of **A** as defined previously, is $(n - r)$. Hence,

$$\text{nullity } \mathbf{A} = n - \text{rank } \mathbf{A}$$

or

**(5)** $$\text{rank } \mathbf{A} + \text{nullity } \mathbf{A} = n$$

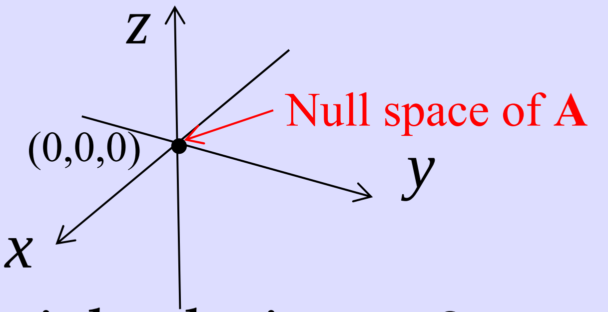
where $n$ is the number of unknowns (number of columns of **A**).

- Note: solutions of nonhomogeneous systems do not form a vector space.)

# Null Space Examples

Example N1:

$m = 3, n = 3$ $\quad \mathbf{A} = \begin{bmatrix} 1 & -1 & 1 \\ 2 & -1 & 3 \\ 1 & 1 & 4 \end{bmatrix}$



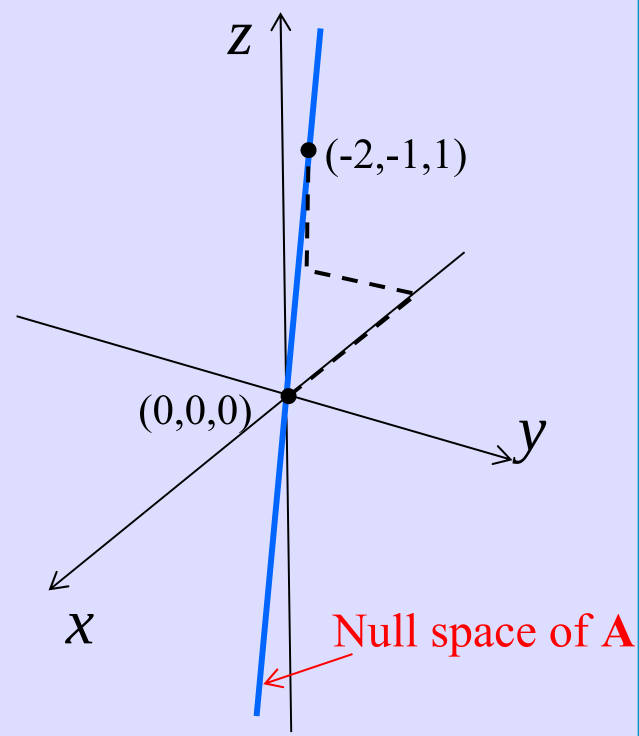

rank A=3=$n$, solution of **Ax=0** is the trivial solution **x=0**.

Example N2:

$m = 2$ $n = 3$ $\quad \mathbf{A} = \begin{bmatrix} 1 & -1 & 1 \\ 2 & -1 & 3 \end{bmatrix}$

rank **A**=2 < $n$, nontrivial solutions of **Ax=0** is the Null Space of dimension ($n$-$r$)=1 given by:

$$\mathbf{x}_h = \beta \begin{bmatrix} -2 \\ -1 \\ 1 \end{bmatrix}$$

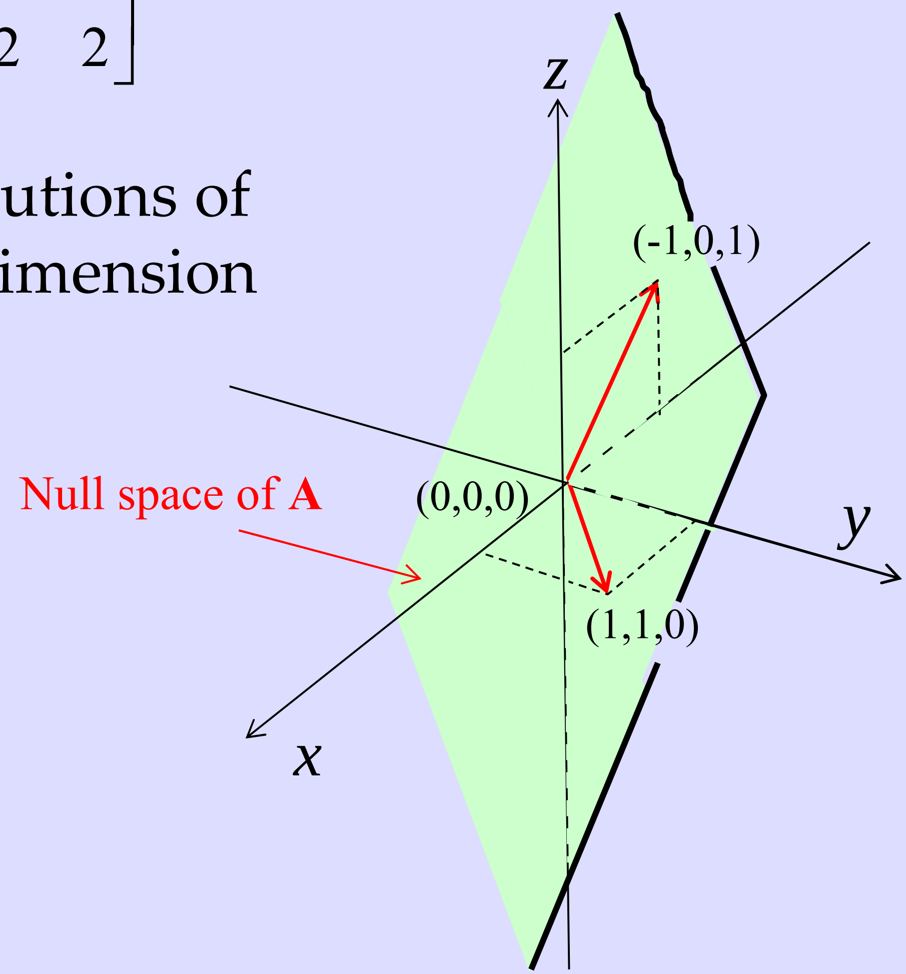


Example N3:

$$m = 2 \; n = 3 \qquad \mathbf{A} = \begin{bmatrix} 1 & -1 & 1 \\ 2 & -2 & 2 \end{bmatrix}$$

rank **A**=1 < $n$, nontrivial solutions of **Ax=0** is the Null Space of dimension ($n$-$r$)=2 given by:

$$\mathbf{x}_h = \alpha \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + \beta \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$



# Theorem 3

**Homogeneous Linear System**
**with Fewer Equations Than Unknowns**

A homogeneous linear system $\mathbf{Ax=0}$ with fewer equations than unknowns always has nontrivial solutions.

Reasoning:
Let $\mathbf{A}$ be $m$ by $n$ with $m < n$ (fewer equations than unknowns). Then, we have rank $\mathbf{A} \le m < n$, which yields the consistent (multiple solution) case.

Combining the previous results, we have the final characterization of the solution of the nonhomogeneous linear equations (1) in the **consistent** case: $\mathbf{x} = \mathbf{x}_0 + \mathbf{x}_h$

- If rank $\mathbf{A}=n$, $\mathbf{Ax=b}$ has unique solution; $\mathbf{Ax=0}$ has **trivial solution**, $\mathbf{x}_h = \mathbf{0}$.
- If rank $\mathbf{A}=r<n$, $\mathbf{Ax=b}$ have multiple solutions, with
  - $\mathbf{Ax=0}$ has **nontrivial solution** $\mathbf{x}_h$ obtained by solving $r$ of the variables $x_1, \ldots, x_n$, in terms of the remaining $(n\text{-}r)$ ***free*** variables in the row echelon form $\mathbf{Rx=f}$.
  - $\mathbf{x}_0$ is any solution of $\mathbf{Ax=b}$ or $\mathbf{Rx=f}$. Note that one can conveniently obtain $\mathbf{x}_0$ by assigning zero values to all free parameter.

# Redo EXAMPLE 3 Consistent ( Multiple Solution) Case of Section 7.3

$n = 4\,,\, m = 3$

**Solving for $\mathbf{x}_0$:**

set free variables of $x_3$=0 and $x_4$=0

$$\left[\begin{array}{cccc|c} 3 & 2 & 2 & -5 & 8 \\ 0 & 1.1 & 1.1 & -4.4 & 1.1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

$$\Rightarrow x_2 = 1 \Rightarrow x_1 = 2 \Rightarrow \mathbf{x}_0 = \begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

**Solving for $\mathbf{x}_h$:**

Let **f**=**0** and express $x_1$ and $x_2$ in terms of free variables $x_3=\alpha$ and $x_4=\beta$

$$\left[\begin{array}{cccc|c} 3 & 2 & 2 & -5 & 0 \\ 0 & 1.1 & 1.1 & -4.4 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

$$\Rightarrow \mathbf{x}_h = \begin{bmatrix} 0 \\ -1 \\ 1 \\ 0 \end{bmatrix}\alpha + \begin{bmatrix} -1 \\ 4 \\ 0 \\ 1 \end{bmatrix}\beta$$

In this case, Rank **A**=2 <$n$=4

→ nontrivial solution for $\mathbf{x}_h$

Solve the following linear system of three equations in four unknowns whose augmented matrix is

$$(5)\quad \left[\begin{array}{cccc|c} 3.0 & 2.0 & 2.0 & -5.0 & 8.0 \\ 0.6 & 1.5 & 1.5 & -5.4 & 2.7 \\ 1.2 & -0.3 & -0.3 & 2.4 & 2.1 \end{array}\right]. \qquad \text{Thus,} \qquad \begin{aligned} 3.0x_1 + 2.0x_2 + 2.0x_3 - 5.0x_4 &= 8.0 \\ \boxed{0.6x_1} + 1.5x_2 + 1.5x_3 - 5.4x_4 &= 2.7 \\ \boxed{1.2x_1} - 0.3x_2 - 0.3x_3 + 2.4x_4 &= 2.1. \end{aligned}$$

***Solution.*** As in the previous example, we circle pivots and box terms of equations and corresponding entries to be eliminated. We indicate the operations in terms of equations and operate on both equations and matrices.

***Step 1. Elimination of $x_1$*** from the second and third equations by adding

$$-0.6/3.0 = -0.2 \text{ times the first equation to the second equation,}$$

$$-1.2/3.0 = -0.4 \text{ times the first equation to the third equation.}$$

This gives the following, in which the pivot of the next step is circled.

$$(6)\quad \left[\begin{array}{cccc|c} 3.0 & 2.0 & 2.0 & -5.0 & 8.0 \\ 0 & 1.1 & 1.1 & -4.4 & 1.1 \\ 0 & -1.1 & -1.1 & 4.4 & -1.1 \end{array}\right] \begin{array}{l} \\ \text{Row 2} - 0.2 \text{ Row 1} \\ \text{Row 3} - 0.4 \text{ Row 1} \end{array} \qquad \begin{aligned} 3.0x_1 + 2.0x_2 + 2.0x_3 - 5.0x_4 &= 8.0 \\ 1.1x_2 + 1.1x_3 - 4.4x_4 &= 1.1 \\ \boxed{-1.1x_2} - 1.1x_3 + 4.4x_4 &= -1.1. \end{aligned}$$

***Step 2. Elimination of $x_2$*** from the third equation of (6) by adding

$$1.1/1.1 = 1 \text{ times the second equation to the third equation.}$$

This gives

$$(7)\quad \left[\begin{array}{cccc|c} 3.0 & 2.0 & 2.0 & -5.0 & 8.0 \\ 0 & 1.1 & 1.1 & -4.4 & 1.1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right] \begin{array}{l} \\ \\ \text{Row 3} + \text{Row 2} \end{array} \qquad \begin{aligned} 3.0x_1 + 2.0x_2 + 2.0x_3 - 5.0x_4 &= 8.0 \\ 1.1x_2 + 1.1x_3 - 4.4x_4 &= 1.1 \\ 0 &= 0. \end{aligned}$$

***Back Substitution.*** From the second equation, $x_2 = 1 - x_3 + 4x_4$. From this and the first equation, $x_1 = 2 - x_4$. Since $x_3$ and $x_4$ remain arbitrary, we have infinitely many solutions. If we choose a value of $x_3$ and a value of $x_4$, then the corresponding values of $x_1$ and $x_2$ are uniquely determined.

***On Notation.*** If unknowns remain arbitrary, it is also customary to denote them by other letters $t_1, t_2, \cdots$. In this example we may thus write $x_1 = 2 - x_4 = 2 - t_2$, $x_2 = 1 - x_3 + 4x_4 = 1 - t_1 + 4t_2$, $x_3 = t_1$ (first arbitrary unknown), $x_4 = t_2$ (second arbitrary unknown). ■

*Advanced Engineering Mathematics, 10/e* by Edwin Kreyszig

## MORE EXAMPLES

### Example M1: Consistent case: unique solution

Solve the linear system

$$x_1 - x_2 + x_3 = 2$$
$$2x_1 - x_2 + 3x_3 = 5$$
$$x_1 + x_2 + 4x_3 = 6$$

$m = 3\ ,\ n = 3$

Augmented matrix $\tilde{\mathbf{A}}$

$$\tilde{\mathbf{A}} = \left[\begin{array}{ccc:c} 1 & -1 & 1 & 2 \\ 2 & -1 & 3 & 5 \\ 1 & 1 & 4 & 6 \end{array}\right]$$

with the first three columns labelled $\mathbf{A}$ and the last column labelled $\mathbf{b}$.

Elementary row operations

Row echelon form

$$r = 3 \quad \left[\begin{array}{ccc:c} 1 & -1 & 1 & 2 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 2 \end{array}\right]$$

with the first three columns labelled $\mathbf{R}$ and the last column labelled $\mathbf{f}$.

rank $\mathbf{A}$ = rank $\tilde{\mathbf{A}}$

→ **Consistent**

rank $\mathbf{A}$ = $r$ =3= $n$

→ case (b) (i): Unique solution

Refer to Example N1 for $\mathbf{x}_h$

Solution:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \underbrace{\begin{bmatrix} -1 \\ -1 \\ 2 \end{bmatrix}}_{\mathbf{x}_o} + \underbrace{\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}}_{\mathbf{x}_h = \mathbf{0}}$$

# Example M2: Consistent case: multiple solutions

Solve the linear system

$$x_1 - x_2 + x_3 = 2$$
$$2x_1 - x_2 + 3x_3 = 5$$
$$x_1 + x_2 + 3x_3 = 4$$

$m = 3\ ,\ n = 3$

Augmented matrix $\tilde{\mathbf{A}}$

$$\tilde{\mathbf{A}} = \left[\begin{array}{ccc|c} 1 & -1 & 1 & 2 \\ 2 & -1 & 3 & 5 \\ 1 & 1 & 3 & 4 \end{array}\right] \quad \text{with } \mathbf{A} \text{ the first three columns and } \mathbf{b} \text{ the last}$$

Elementary row operations

Row echelon form

$$\left[\begin{array}{ccc|c} 1 & -1 & 1 & 2 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right] \quad \text{with } \mathbf{R} \text{ the first three columns and } \mathbf{f} \text{ the last}; \quad r = 2 \text{ (first two rows)}$$

rank $\mathbf{A}$ = rank $\tilde{\mathbf{A}}$

→ **Consistent**

rank $\mathbf{A}$ = $r$ =2 < $n$=3

→ case (b) (ii): multiple solutions

Free variables:

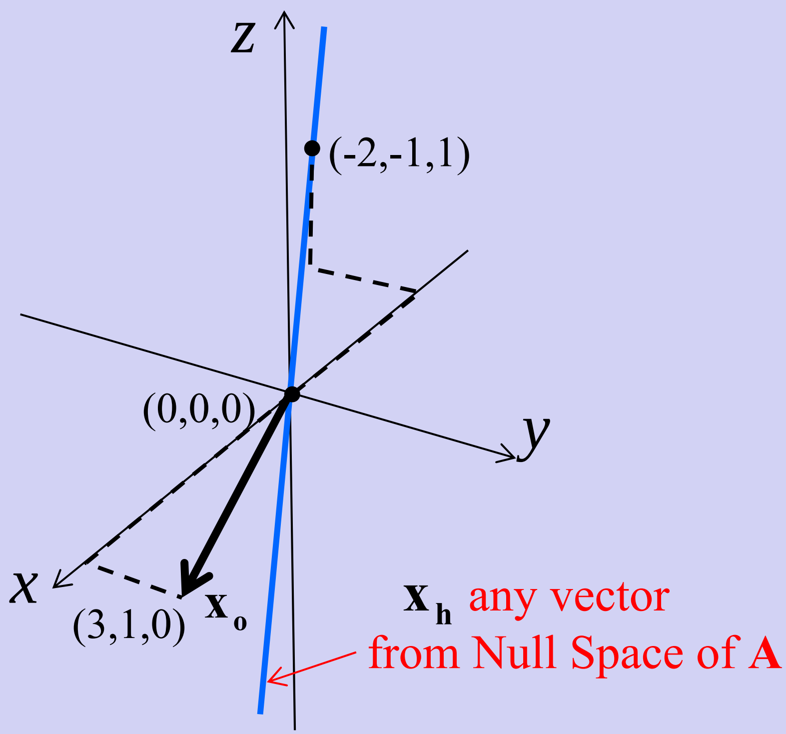
$x_3 = \beta$

Refer to Example N2 for $\mathbf{x}_h$

Solution:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \underbrace{\begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix}}_{\mathbf{x}_o} + \beta \underbrace{\begin{bmatrix} -2 \\ -1 \\ 1 \end{bmatrix}}_{\mathbf{x}_h}$$

# Solution space of Example M2: $\mathbf{x} = \mathbf{x}_0 + \mathbf{x}_h$



## Example M3: Consistent case: multiple solutions

Solve the linear system

$$x_1 - x_2 + x_3 = 2$$
$$2x_1 - 2x_2 + 2x_3 = 4$$
$$-x_1 + x_2 - x_3 = -2$$

$m = 3,\ n = 3$

Augmented matrix $\tilde{\mathbf{A}}$

$$\tilde{\mathbf{A}} = \left[\begin{array}{ccc:c} 1 & -1 & 1 & 2 \\ 2 & -2 & 2 & 4 \\ -1 & 1 & -1 & -2 \end{array}\right]$$

(with $\mathbf{A}$ the first three columns and $\mathbf{b}$ the last column)

Elementary row operations

Row echelon form

$$r = 1\ \left\{\left[\begin{array}{ccc:c} 1 & -1 & 1 & 2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]\right.$$

(with $\mathbf{R}$ the first three columns and $\mathbf{f}$ the last column)

rank $\mathbf{A}$ = rank $\tilde{\mathbf{A}}$

→ **Consistent**

rank $\mathbf{A}$ = $r$ =1 < $n$=3

→ case (b) (i): multiple solution

Free variables:

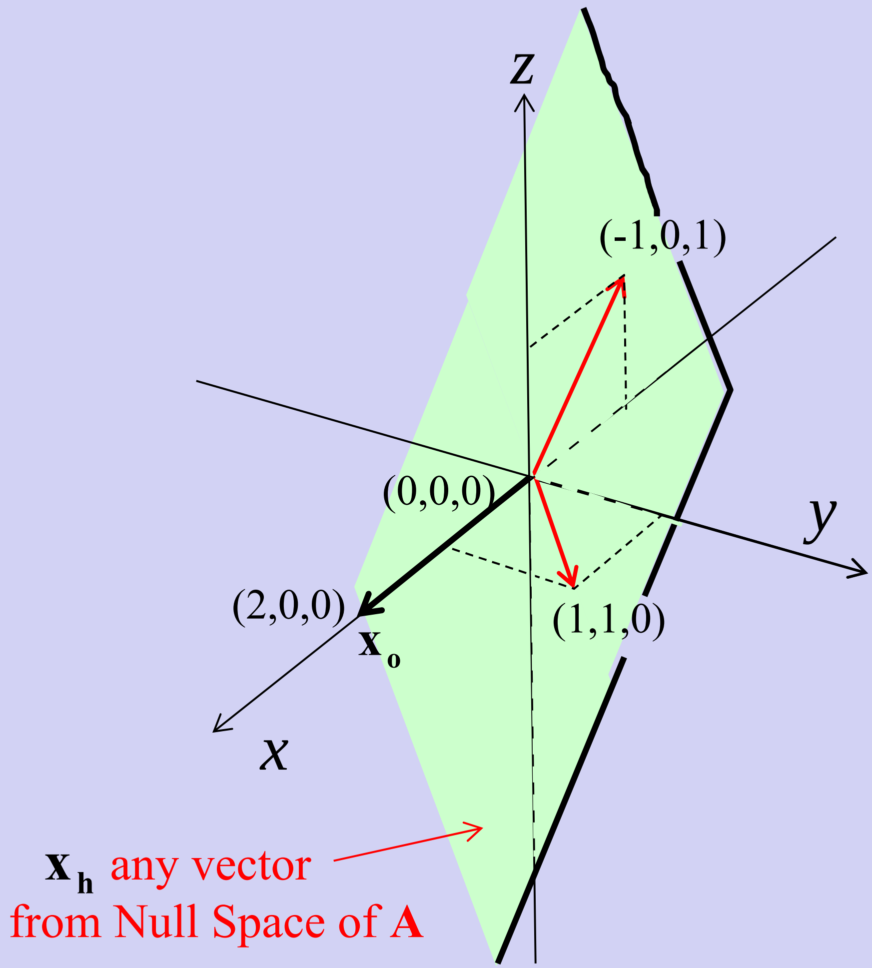
$x_2 = \alpha,\ x_3 = \beta$

Refer to Example N3 for $x_h$

Solution:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \underbrace{\begin{bmatrix} 2 \\ 0 \\ 0 \end{bmatrix}}_{\mathbf{x}_o} + \underbrace{\alpha \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + \beta \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}}_{\mathbf{x}_h}$$

# Solution space of Example M3: $\mathbf{x} = \mathbf{x}_0 + \mathbf{x}_h$



# 20.2 Linear Systems: LU-Factorization, Matrix Inversion

Note: **A** need to be square

An **LU-factorization** of a given square matrix **A** is of the form

$$\mathbf{A} = \mathbf{LU} \tag{2}$$

where **L** is *lower triangular* and **U** is *upper triangular*. For example,

$$\mathbf{A} = \begin{bmatrix} 2 & 3 \\ 8 & 5 \end{bmatrix} = \mathbf{LU} = \begin{bmatrix} 1 & 0 \\ 4 & 1 \end{bmatrix} \begin{bmatrix} 2 & 3 \\ 0 & -7 \end{bmatrix}.$$

Crucial idea: **L** and **U** can be efficiently obtained!

# Solution of Linear System Equations by LU decomposition (by Alan Turing in 1948)

Given the linear system **Ax=b,** with square matrix **A**:

- Conduct LU factorization to obtain lower triangular matrix **L** and upper triangular matrix **U**, such that **A=LU**
- Then **Ax=LUx=b** can be viewed as **y=Ux,** and **Ly=b**
- Specifically, we can solve **Ax=b** in two steps:
  **(i)** Solve **Ly=b** for **y** ← By forward substitution
  **(ii)** Solve **Ux=y** for **x** ← By back substitution
- **Doolittle's method:** LU factorization requiring **L** have main diagonal element of 1
- **Crout's method:** LU factorization requiring **U** have main diagonal element of 1

# Doolittle's method

- LU factorization: **A=LU** that requires **L** to have main diagonal element of 1
- Given square matrix **A**, convert **A** to become

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

$$= \underbrace{\begin{bmatrix} 1 & 0 & \cdots & 0 \\ m_{21} & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ m_{n1} & m_{n2} & \cdots & 1 \end{bmatrix}}_{\mathbf{L}} \underbrace{\begin{bmatrix} u_{11} & u_{12} & \cdots & u_{1n} \\ 0 & u_{22} & \cdots & u_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & u_{nn} \end{bmatrix}}_{\mathbf{U}}$$

- **To determine $m_{ij}$ and $u_{ij}$**

$$
\begin{aligned}
&u_{1k} = a_{1k} && k = 1, \cdots, n \\
&m_{j1} = \frac{a_{j1}}{u_{11}} && j = 2, \cdots, n \\
(4) \quad &u_{jk} = a_{jk} - \sum_{s=1}^{j-1} m_{js} u_{sk} && k = j, \cdots, n; \quad j \geqq 2 \\
&m_{jk} = \frac{1}{u_{kk}} \left( a_{jk} - \sum_{s=1}^{k-1} m_{js} u_{sk} \right) && j = k+1, \cdots, n; \quad k \geqq 2.
\end{aligned}
$$

9 parameters to determine

## Doolittle's Method

Solve the system in Example 1 of Sec. 20.1 by Doolittle's method.

***Solution.*** The decomposition (2) is obtained from

$$\mathbf{A} = [a_{jk}] = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \begin{bmatrix} 3 & 5 & 2 \\ 0 & 8 & 2 \\ 6 & 2 & 8 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ m_{21} & 1 & 0 \\ m_{31} & m_{32} & 1 \end{bmatrix} \begin{bmatrix} u_{11} & u_{12} & u_{13} \\ 0 & u_{22} & u_{23} \\ 0 & 0 & u_{33} \end{bmatrix}$$

by determining the $m_{jk}$ and $u_{jk}$, using matrix multiplication. By going through $\mathbf{A}$ row by row we get successively

| | | |
|---|---|---|
| $a_{11} = 3 = 1 \cdot u_{11} = u_{11}$ | $a_{12} = 5 = 1 \cdot u_{12} = u_{12}$ | $a_{13} = 2 = 1 \cdot u_{13} = u_{13}$ |
| $a_{21} = 0 = m_{21}u_{11}$<br>$m_{21} = 0$ | $a_{22} = 8 = m_{21}u_{12} + u_{22}$<br>$u_{22} = 8$ | $a_{23} = 2 = m_{21}u_{13} + u_{23}$<br>$u_{23} = 2$ |
| $a_{31} = 6 = m_{31}u_{11}$<br>$= m_{31} \cdot 3$<br>$m_{31} = 2$ | $a_{32} = 2 = m_{31}u_{12} + m_{32}u_{22}$<br>$= 2 \cdot 5 + m_{32} \cdot 8$<br>$m_{32} = -1$ | $a_{33} = 8 = m_{31}u_{13} + m_{32}u_{23} + u_{33}$<br>$= 2 \cdot 2 - 1 \cdot 2 + u_{33}$<br>$u_{33} = 6$ |

Thus the factorization (2) is

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

We first solve $\mathbf{Ly} = \mathbf{b}$, determining $y_1 = 8$, then $y_2 = -7$, then $y_3$ from $2y_1 - y_2 + y_3 = 16 + 7 + y_3 = 26$; thus (note the interchange in $\mathbf{b}$ because of the interchange in $\mathbf{A}$!)

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & -1 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 8 \\ -7 \\ 26 \end{bmatrix}. \qquad \text{Solution} \qquad \mathbf{y} = \begin{bmatrix} 8 \\ -7 \\ 3 \end{bmatrix}.$$

Then we solve $\mathbf{Ux} = \mathbf{y}$, determining $x_3 = \frac{3}{6}$ then $x_2$, then $x_1$, that is,

$$\begin{bmatrix} 3 & 5 & 2 \\ 0 & 8 & 2 \\ 0 & 0 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 8 \\ -7 \\ 3 \end{bmatrix}. \qquad \text{Solution} \qquad \mathbf{x} = \begin{bmatrix} 4 \\ -1 \\ \frac{1}{2} \end{bmatrix}.$$

This agrees with the solution in Example 1 of Sec. 20.1.

## EXAMPLE 2 of Section 7.3

Solve the linear system (after eliminating the first equation to make **A** square)

$$
\begin{aligned}
-x_1 + \quad x_2 - \quad x_3 &= 0 \\
10x_2 + 25x_3 &= 90 \\
20x_1 + 10x_2 \qquad &= 80.
\end{aligned}
$$

which is equivalent to $\begin{bmatrix} -1 & 1 & -1 \\ 0 & 10 & 25 \\ 20 & 10 & 0 \end{bmatrix} \mathbf{x} = \mathbf{b} = \begin{bmatrix} 0 \\ 90 \\ 80 \end{bmatrix}$

Write in LU form $\begin{bmatrix} -1 & 1 & -1 \\ 0 & 10 & 25 \\ 20 & 10 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ m_{21} & 1 & 0 \\ m_{31} & m_{32} & 1 \end{bmatrix} \begin{bmatrix} u_{11} & u_{12} & u_{13} \\ 0 & u_{22} & u_{23} \\ 0 & 0 & u_{33} \end{bmatrix}$

Solve for **L** and **U**

$$u_{11} = -1,$$
$$m_{21}u_{11} = 0 \Rightarrow m_{21} = 0,$$
$$m_{31}u_{11} = 20 \Rightarrow m_{31} = -20,$$

1st column of **A**

$$u_{12} = 1,$$
$$m_{21}u_{12} + u_{22} = 10 \Rightarrow u_{22} = 10,$$
$$m_{31}u_{12} + m_{32}u_{22} = 10 \Rightarrow m_{32} = 3,$$

2nd column of **A**

$$u_{13} = -1$$
$$m_{21}u_{13} + u_{23} = 25 \Rightarrow u_{23} = 25$$
$$m_{31}u_{13} + m_{32}u_{23} + u_{33} = 0 \Rightarrow u_{33} = 95$$

3rd column of **A**

Results:

$$\mathbf{A} = \begin{bmatrix} -1 & 1 & -1 \\ 0 & 10 & 25 \\ 20 & 10 & 0 \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -20 & 3 & 1 \end{bmatrix}}_{\mathbf{L}} \underbrace{\begin{bmatrix} -1 & 1 & -1 \\ 0 & 10 & 25 \\ 0 & 0 & -95 \end{bmatrix}}_{\mathbf{U}}$$

(i) Solve **Ly=b:**

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -20 & 3 & 1 \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 90 \\ 80 \end{bmatrix} \Rightarrow \begin{cases} y_1 = 0 \\ y_2 = 90 \\ y_3 = -190 \end{cases}$$

(ii) Solve **Ux=y**

$$\begin{bmatrix} -1 & 1 & -1 \\ 0 & 10 & 25 \\ 0 & 0 & -95 \end{bmatrix}\mathbf{x} = \begin{bmatrix} 0 \\ 90 \\ -190 \end{bmatrix} \Rightarrow \begin{cases} x_3 = 2 \\ x_2 = 4 \\ x_1 = 2 \end{cases}$$

(same results as before)

# Cholesky's Method

- **Cholesky's Method:** For *symmetric*, *positive definite* matrix $\mathbf{A}$ (i.e, $\mathbf{A}=\mathbf{A}^T$ and $\mathbf{x}^T\,\mathbf{A}\,\mathbf{x} > 0$ for all $\mathbf{x}=0$), we have $\mathbf{U}=\mathbf{L}^T$:

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ a_{21} & a_{22} & \cdots & a_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

$$= \underbrace{\begin{bmatrix} l_{11} & 0 & \cdots & 0 \\ l_{21} & l_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ l_{n1} & l_{n2} & \cdots & l_{nn} \end{bmatrix}}_{\mathbf{L}} \underbrace{\begin{bmatrix} l_{11} & l_{21} & \cdots & l_{n1} \\ 0 & l_{22} & \cdots & l_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & l_{nn} \end{bmatrix}}_{\mathbf{L}^T}$$

Diagonal elements not 1 in this case.

- Example:

$$\mathbf{A} = \begin{bmatrix} 4 & 2 & 14 \\ 2 & 17 & -5 \\ 14 & -5 & 83 \end{bmatrix} = \begin{bmatrix} 2 & 0 & 0 \\ 1 & 4 & 0 \\ 7 & -3 & 5 \end{bmatrix} \begin{bmatrix} 2 & 1 & 7 \\ 0 & 4 & -3 \\ 0 & 0 & 5 \end{bmatrix}$$

- General method to determine $l_{ij}$ from given symmetric, positive definite matrix **A**:

$$\begin{aligned} l_{11} &= \sqrt{a_{11}} && \\ l_{j1} &= \frac{a_{j1}}{l_{11}} && j = 2, \cdots, n \\ l_{jj} &= \sqrt{a_{jj} - \sum_{s=1}^{j-1} l_{js}^2} && j = 2, \cdots, n \\ l_{pj} &= \frac{1}{l_{jj}}\left(a_{pj} - \sum_{s=1}^{j-1} l_{js} l_{ps}\right) && p = j+1, \cdots, n; \quad j \geqq 2. \end{aligned} \tag{6}$$

- If matrix **A** symmetric, but not positive definite, Cholesky's method leads to complex matrix **L** and hence becomes impractical.

## Cholesky's Method

Solve by Cholesky's method:

$$
\begin{aligned}
4x_1 + 2x_2 + 14x_3 &= 14 \\
2x_1 + 17x_2 - 5x_3 &= -101 \\
14x_1 - 5x_2 + 83x_3 &= 155.
\end{aligned}
$$

6 parameters to determine

***Solution.*** From (6) or from the form of the factorization

$$
\begin{bmatrix} 4 & 2 & 14 \\ 2 & 17 & -5 \\ 14 & -5 & 83 \end{bmatrix} = \begin{bmatrix} l_{11} & 0 & 0 \\ l_{21} & l_{22} & 0 \\ l_{31} & l_{32} & l_{33} \end{bmatrix} \begin{bmatrix} l_{11} & l_{21} & l_{31} \\ 0 & l_{22} & l_{32} \\ 0 & 0 & l_{33} \end{bmatrix}
$$

we compute, in the given order,

$$
l_{11} = \sqrt{a_{11}} = 2 \qquad l_{21} = \frac{a_{21}}{l_{11}} = \frac{2}{2} = 1 \qquad l_{31} = \frac{a_{31}}{l_{11}} = \frac{14}{2} = 7
$$

$$
l_{22} = \sqrt{a_{22} - l_{21}^2} = \sqrt{17 - 1} = 4
$$

$$
l_{32} = \frac{1}{\cancel{l_{23}}}(a_{32} - l_{31}l_{21}) = \frac{1}{4}(-5 - 7 \cdot 1) = -3
$$

$l_{22}$

$$
l_{33} = \sqrt{a_{33} - l_{31}^2 - l_{32}^2} = \sqrt{83 - 7^2 - (-3)^2} = 5.
$$

EXAMPLE 2

This agrees with (5). We now have to solve $\mathbf{Ly} = \mathbf{b}$, that is,

$$\begin{bmatrix} 2 & 0 & 0 \\ 1 & 4 & 0 \\ 7 & -3 & 5 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 14 \\ -101 \\ 155 \end{bmatrix}. \qquad \text{Solution} \qquad \mathbf{y} = \begin{bmatrix} 7 \\ -27 \\ 5 \end{bmatrix}.$$

As the second step, we have to solve $\mathbf{Ux} = \mathbf{L}^{\mathsf{T}}\mathbf{x} = \mathbf{y}$, that is,

$$\begin{bmatrix} 2 & 1 & 7 \\ 0 & 4 & -3 \\ 0 & 0 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 7 \\ -27 \\ 5 \end{bmatrix}. \qquad \text{Solution} \qquad \mathbf{x} = \begin{bmatrix} 3 \\ -6 \\ 1 \end{bmatrix}.$$

# 7.6 For Reference: Second- and Third-Order Determinants

- A **determinant of second order** is denoted and defined for 2 by 2 matrix by

$$(1) \qquad D = \det \mathbf{A} = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}.$$

- Determinant denoted with ***bars*** (whereas a matrix uses ***brackets***).

- Given a linear systems of **2** equations in **2** unknowns

(2)
$$\begin{aligned} a_{11}x_1 + a_{12}x_2 &= b_1 \\ a_{21}x_1 + a_{22}x_2 &= b_2 \end{aligned}$$

Let $D$ be the determinant as given in (1).

- If $D \neq 0$, then system (2) is **consistent** with unique solution (later!). This unique solution is given by the **Cramer's rule:**

(3)
$$x_1 = \frac{\begin{vmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{vmatrix}}{D} = \frac{b_1a_{22} - a_{12}b_2}{D},$$

$$x_2 = \frac{\begin{vmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{vmatrix}}{D} = \frac{a_{11}b_2 - b_1a_{21}}{D}.$$

# Third-Order Determinants

A **determinant of third order** can be defined for 3 by 3 matrix by

$$(4)\quad D = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}\underbrace{\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix}} - a_{21}\underbrace{\begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix}} + a_{31}\underbrace{\begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix}}.$$

Minor corresponding to entry $a_{11}$, $a_{21}$, and $a_{31}$

Note that

- The signs on the right are $+ - +$.
- Each of the three terms on the right is an entry in the first column of $D$ times its corresponding **minor**, which is the second-order determinant obtained from $D$ by deleting the row and column of that entry;
- If we write out the minors in (4), we obtain

$$(4^*)\quad D = a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} + a_{21}a_{13}a_{32} - a_{21}a_{12}a_{33} + a_{31}a_{12}a_{23} - a_{31}a_{13}a_{22}.$$

- Given a linear systems of **3** equations in **3** unknowns

$$
\text{(5)} \qquad \begin{aligned}
a_{11}x_1 + a_{12}x_2 + a_{13}x_3 &= b_1 \\
a_{21}x_1 + a_{22}x_2 + a_{23}x_3 &= b_2 \\
a_{31}x_1 + a_{32}x_2 + a_{33}x_3 &= b_3
\end{aligned}
$$

  Let $D$ be the determinant as given in (4).

- If $D \neq 0$, system (5) is **consistent** with unique solution (later!). This unique solution is given by the **Cramer's rule:**

$$x_1 = \frac{D_1}{D}, \qquad x_2 = \frac{D_2}{D}, \qquad x_3 = \frac{D_3}{D}$$ , where

$$
\text{(6)} \qquad D_1 = \begin{vmatrix} b_1 & a_{12} & a_{13} \\ b_2 & a_{22} & a_{23} \\ b_3 & a_{32} & a_{33} \end{vmatrix}, \qquad D_2 = \begin{vmatrix} a_{11} & b_1 & a_{13} \\ a_{21} & b_2 & a_{23} \\ a_{31} & b_3 & a_{33} \end{vmatrix}, \qquad D_3 = \begin{vmatrix} a_{11} & a_{12} & b_1 \\ a_{21} & a_{22} & b_2 \\ a_{31} & a_{32} & b_3 \end{vmatrix}.
$$

- Note: $D_1$, $D_2$, $D_3$ are obtained by replacing Columns 1, 2, 3, respectively, by the column of the right sides of (5) in (4).

# 7.7 Determinants. Cramer's Rule

# Generalization to *n* by *n* Matrix

A **determinant of order** $n$ is a scalar associated with an $n \times n$ (***square***) matrix $\mathbf{A} = [a_{jk}]$, and is denoted by

(1)
$$D = \det \mathbf{A} = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{vmatrix}.$$

For $n = 1$, this determinant is defined by

(2)
$$D = a_{11}.$$

- **Def:** $M_{jk}$ , called the Minor of entry $a_{jk}$ , is a determinant of order $n - 1$, namely, the determinant of the submatrix of **A** obtained from **A** by omitting the row and column of the entry $a_{jk}$, that is, the $j$th row and the $k$th column.

$$\mathbf{A} = \begin{array}{cc} & \begin{array}{ccccccc} 1^{\text{st}} & \cdots\cdots & (k\text{-}1)^{\text{th}} & k^{\text{th}} & (k+1)^{\text{th}} & \cdots & n^{\text{th}}\text{ column} \end{array} \\ \begin{array}{c} 1^{\text{st}} \\ \vdots \\ (j\text{-}1)^{\text{th}} \\ j^{\text{th}} \\ (j+1)^{\text{th}} \\ \vdots \\ n^{\text{th}} \\ \text{row} \end{array} & \begin{bmatrix} a_{1,1} & \cdots & a_{1,k-1} & a_{1,k} & a_{1,k+1} & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{j-1,1} & \cdots & a_{j-1,k-1} & a_{j-1,k} & a_{j-1,k+1} & \cdots & a_{j-1,n} \\ a_{j,1} & \cdots & a_{j,k-1} & a_{j,k} & a_{j,k+1} & \cdots & a_{j,n} \\ a_{j+1,1} & \cdots & a_{j+1,k-1} & a_{j+1,k} & a_{j+1,k+1} & \cdots & a_{j+1,n} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & \cdots & a_{n,k-1} & a_{n,k} & a_{n,k+1} & \cdots & a_{n,n} \end{bmatrix} \end{array}$$

$$
\begin{array}{cc}
 & \begin{array}{cccccc} 1^{\text{st}} & \cdots & (k\text{-}1)^{\text{th}} & (k+1)^{\text{th}} & \cdots & n^{\text{th}} \text{ column} \end{array} \\
\begin{array}{c} 1^{\text{st}} \\ \vdots \\ (j\text{-}1)^{\text{th}} \\ (j+1)^{\text{th}} \\ \vdots \\ n^{\text{th}} \\ \text{row} \end{array} &
\underbrace{\begin{vmatrix}
a_{1,1} & \cdots & a_{1,k-1} & a_{1,k+1} & \cdots & a_{1,n} \\
\vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
a_{j-1,1} & \cdots & a_{j-1,k-1} & a_{j-1,k+1} & \cdots & a_{j-1,n} \\
a_{j+1,1} & \cdots & a_{j+1,k-1} & a_{j+1,k+1} & \cdots & a_{j+1,n} \\
\vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
a_{n,1} & \cdots & a_{n,k-1} & a_{n,k+1} & \cdots & a_{n,n}
\end{vmatrix}}
\end{array}
$$

- Take submatrix of **A** by omitting its *j*th row and the *k*th column
- $M_{jk}$ is the determinant of the resulting $(n-1) \times (n-1)$ submatrix

- For $n \geq 2$, the determinant of the $n \times n$ matrix A can be put in terms of the Minors $M_{jk}$ about the entries of the 1st row as follows:

$$D = a_{1,1}M_{1,1} - a_{1,2}M_{1,2} + a_{1,3}M_{1,3} + \cdots + (-1)^{n+1} a_{1,n}M_{1,n}$$

or about the entries of the 1st column,

$$D = a_{1,1}M_{1,1} - a_{2,1}M_{2,1} + a_{3,1}M_{3,1} + \cdots + (-1)^{n+1} a_{n,1}M_{n,1}$$

or about the entries of the $j$th row, $j=1,2,\ldots, n$

$$D = (-1)^{j+1} a_{j1}M_{j1} + (-1)^{j+2} a_{j2}M_{j2} + (-1)^{j+3} a_{j3}M_{j3} + \cdots + (-1)^{j+n} a_{jn}M_{jn}$$

$$= \sum_{k=1}^{n} (-1)^{j+k} a_{jk}M_{jk}$$

or about the entries of the $k$th column, $k=1,2,\ldots, n$

$$D = (-1)^{k+1} a_{1k}M_{1k} + (-1)^{k+2} a_{2k}M_{2k} + (-1)^{k+3} a_{3k}M_{3k} + \cdots + (-1)^{k+n} a_{nk}M_{nk}$$

$$= \sum_{j=1}^{n} (-1)^{j+k} a_{jk}M_{jk}$$

- Can also define Cofactor $C_{jk}$ of the entry $a_{jk}$ as

$$C_{jk} = (-1)^{j+k} M_{jk}$$

i.e., $C_{jk}$ absorbs the factor $(-1)^{j+k}$ in its definition so that the determinant can be expressed as:

(3a) $$D = a_{j1}C_{j1} + a_{j2}C_{j2} + \dots + a_{jn}C_{jn}$$

(about $j$th row, $j = 1, 2, \dots, \text{ or } n$)

or

(3b) $$D = a_{1k}C_{1k} + a_{2k}C_{2k} + \dots + a_{nk}C_{nk}$$

(about $k$th column $k = 1, 2, \dots, \text{ or } n$).

- As such, $D$ is defined in terms of $n$ determinants of order $n - 1$, each of which is, in turn, defined in terms of $n - 1$ determinants of order $n - 2$ and so on—until we finally arrive at first-order determinants.
- We may **expand** $D$ by any row or column, that is, one can choose to use in (3) the entries in any row or column.
- Result is same for $D$ no matter which columns or rows we choose in expanding the determinant.
- Definition is consistent with that of second and third order determinants before in (1) and (4) of Section 7.6.

# EXAMPLE 1

**Minors and Cofactors of a Third-Order Determinant**

In (4) of the previous section the minors and cofactors of the entries in the first column can be seen directly. For the entries in the second row the minors are

$$M_{21} = \begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix}, \quad M_{22} = \begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix}, \quad M_{23} = \begin{vmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{vmatrix}$$

and the cofactors are $C_{21} = -M_{21}$, $C_{22} = +M_{22}$, and $C_{23} = -M_{23}$.

Note also that the signs in $C_{jk}$ form a **checkerboard pattern**

$$\begin{matrix} + & - & + \\ - & + & - \\ + & - & + \end{matrix}$$

# EXAMPLE 2

## Expansions of a Third-Order Determinant

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

$$= 1(12 - 0) - 3(4 + 4) + 0(0 + 6) = -12.$$

This is the expansion by the first row. The expansion by the third column is

$$D = 0\begin{vmatrix} 2 & 6 \\ -1 & 0 \end{vmatrix} - 4\begin{vmatrix} 1 & 3 \\ -1 & 0 \end{vmatrix} + 2\begin{vmatrix} 1 & 3 \\ 2 & 6 \end{vmatrix} = 0 - 12 + 0 = -12.$$

Verify that the other four expansions also give the value $-12$.

# Application of Row elementary operations to Determinants

- Determinants as defined above is not computational efficient.
- Seek to find determinant by applying elementary row operations to convert (1) to an row echelon form.
- Given that we are dealing with square matrices, row echelon form in this case means "upper triangular" matrix.
- "Upper triangular" matrix determinant value very easily computed as product of its diagonal entries.
- Approach is **similar** (**but not the same!**) to what we did to matrices in Sec. 7.3. because determinant values **can be changed** by elementary row operations.

**THEOREM 1 Behavior of an *n*th-Order Determinant under Elementary Row Operations**

(a) Interchange of two rows multiplies the value of the determinant by −1.

$$\det\begin{bmatrix} \text{---}\ \mathbf{a}_{(1)}\ \text{---} \\ \vdots \\ \text{---}\ \mathbf{a}_{(j)}\ \text{---} \\ \vdots \\ \text{---}\ \mathbf{a}_{(k)}\ \text{---} \\ \vdots \\ \text{---}\ \mathbf{a}_{(n)}\ \text{---} \end{bmatrix} = \begin{vmatrix} \text{---}\ \mathbf{a}_{(1)}\ \text{---} \\ \vdots \\ \text{---}\ \mathbf{a}_{(j)}\ \text{---} \\ \vdots \\ \text{---}\ \mathbf{a}_{(k)}\ \text{---} \\ \vdots \\ \text{---}\ \mathbf{a}_{(n)}\ \text{---} \end{vmatrix} = (-1)\begin{vmatrix} \text{---}\ \mathbf{a}_{(1)}\ \text{---} \\ \vdots \\ \text{---}\ \mathbf{a}_{(k)}\ \text{---} \\ \vdots \\ \text{---}\ \mathbf{a}_{(j)}\ \text{---} \\ \vdots \\ \text{---}\ \mathbf{a}_{(n)}\ \text{---} \end{vmatrix} \text{ swap}$$

Matrix Operator for interchanging Row 2 and Row 4 of a 5 by 5 matrix:

$$\mathbf{RO}_{swap\,(R2,R4)} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

- Proof in 2 steps

1) Interchanging adjacent rows changes determinant value by (-1). Example: given matrix **M**, and **M′** obtained by interchanging , say, the 1st and 2nd row of **M**

$$\mathbf{M} = \begin{vmatrix} a_{1,1} & a_{1,2} & a_{1,3} & \cdots & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & a_{2,3} & \cdots & \cdots & a_{2,n} \\ a_{3,1} & a_{3,2} & a_{3,3} & \cdots & \cdots & a_{3,n} \\ \vdots & \vdots & \vdots & \ddots & \cdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & a_{n,3} & \cdots & \cdots & a_{n,n} \end{vmatrix}$$

Determinant $D$ (expanding about 1st row) of **M**:

$$D = a_{1,1}M_{1,1} - a_{1,2}M_{1,2} + \cdots + (-1)^{n+1} a_{1,n}M_{1,n}$$

$$\mathbf{M}' = \begin{vmatrix} a_{2,1} & a_{2,2} & a_{2,3} & \cdots & \cdots & a_{2,n} \\ a_{1,1} & a_{1,2} & a_{1,3} & \cdots & \cdots & a_{1,n} \\ a_{3,1} & a_{3,2} & a_{3,3} & \cdots & \cdots & a_{3,n} \\ \vdots & \vdots & \vdots & \ddots & \cdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & a_{n,3} & \cdots & \cdots & a_{n,n} \end{vmatrix}$$

Determinant $D'$ (expanding about 2nd row) of **M′**:

$$D' = -a_{1,1}M'_{2,1} + a_{1,2}M'_{2,2} - \cdots + (-1)^{n+2} a_{1,n}M'_{2,n}$$

- Proof in 2 steps

2) Noting that $M_{1,1}=M'_{2,1}$ , $M_{1,2}=M'_{2,2}$ , or $M_{1,j}=M_{2,j}$ in general, we have $D'=(-1)D$.

$$\mathbf{M} = \begin{vmatrix} a_{1,1} & a_{1,2} & a_{1,3} & \cdots & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & a_{2,3} & \cdots & \cdots & a_{2,n} \\ a_{3,1} & a_{3,2} & a_{3,3} & \cdots & \cdots & a_{3,n} \\ \vdots & \vdots & \vdots & \ddots & \cdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & a_{n,3} & \cdots & \cdots & a_{n,n} \end{vmatrix}$$

Determinant $D$ (expanding about 1st row) of M:

$$D = a_{1,1}M_{1,1} - a_{1,2}M_{1,2} + \cdots + (-1)^{n+1}a_{1,n}M_{1,n}$$

$$\mathbf{M}' = \begin{vmatrix} a_{2,1} & a_{2,2} & a_{2,3} & \cdots & \cdots & a_{2,n} \\ a_{1,1} & a_{1,2} & a_{1,3} & \cdots & \cdots & a_{1,n} \\ a_{3,1} & a_{3,2} & a_{3,3} & \cdots & \cdots & a_{3,n} \\ \vdots & \vdots & \vdots & \ddots & \cdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & a_{n,3} & \cdots & \cdots & a_{n,n} \end{vmatrix}$$

Determinant $D'$ (expanding about 2nd row) of M′:

$$D' = -a_{1,1}M'_{2,1} + a_{1,2}M'_{2,2} - \cdots + (-1)^{n+2}a_{1,n}M'_{2,n}$$

(b) Addition of a multiple of a row to another row does not alter the value of the determinant.

$$\begin{vmatrix} \text{——} & \mathbf{a}_{(1)} & \text{——} \\ & \vdots & \\ \text{——} & \mathbf{a}_{(j)} & \text{——} \\ & \vdots & \\ \text{——} & \mathbf{a}_{(k)} & \text{——} \\ & \vdots & \\ \text{——} & \mathbf{a}_{(n)} & \text{——} \end{vmatrix} = \begin{vmatrix} \text{——} & \mathbf{a}_{(1)} & \text{——} \\ & \vdots & \\ - & \mathbf{a}_{(j)} + c\mathbf{a}_{(k)} & - \\ & \vdots & \\ \text{——} & \mathbf{a}_{(k)} & \text{——} \\ & \vdots & \\ \text{——} & \mathbf{a}_{(n)} & \text{——} \end{vmatrix}$$

Matrix Operator for adding $c$ times Row 4 to Row 2 of a 5 by 5 matrix:

$$\mathbf{RO}_{Add\,(R2,cR4)} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & c & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

(c) Multiplication of a row by a nonzero constant $c$ multiplies the value of the determinant by $c$.

$$\begin{vmatrix} \text{---} & \mathbf{a}_{(1)} & \text{---} \\ & \vdots & \\ \text{---} & c\mathbf{a}_{(j)} & \text{---} \\ & \vdots & \\ \text{---} & \mathbf{a}_{(k)} & \text{---} \\ & \vdots & \\ \text{---} & \mathbf{a}_{(n)} & \text{---} \end{vmatrix} = c \begin{vmatrix} \text{---} & \mathbf{a}_{(1)} & \text{---} \\ & \vdots & \\ \text{---} & \mathbf{a}_{(k)} & \text{---} \\ & \vdots & \\ \text{---} & \mathbf{a}_{(j)} & \text{---} \\ & \vdots & \\ \text{---} & \mathbf{a}_{(n)} & \text{---} \end{vmatrix}$$

Matrix Operator for multiplying Row 2 of a 5 by 5 matrix by $c$:

$$\mathbf{RO}_{Mult\,(c,R2)} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & c & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

## EXAMPLE 4

**Evaluation of Determinants by converting to Upper Triangular Form via elementary row operations**

$$D = \begin{vmatrix} 2 & 0 & -4 & 6 \\ 4 & 5 & 1 & 0 \\ 0 & 2 & 6 & -1 \\ -3 & 8 & 9 & 1 \end{vmatrix}$$

$$= \begin{vmatrix} 2 & 0 & -4 & 6 \\ 0 & 5 & 9 & -12 \\ 0 & 2 & 6 & -1 \\ 0 & 8 & 3 & 10 \end{vmatrix}$$

Row 2 − 2 Row 1
(No change in *D* value)

Row 4 + 1.5 Row 1
(No change in *D* value)

## EXAMPLE 4 (continued)

$$= \begin{vmatrix} 2 & 0 & -4 & 6 \\ 0 & 5 & 9 & -12 \\ 0 & 0 & 2.4 & 3.8 \\ 0 & 0 & -11.4 & 29.2 \end{vmatrix}$$

Row 3 – 0.4 Row 2
(No change in *D* value)
Row 4 – 1.6 Row 2
(No change in *D* value)

$$= \begin{vmatrix} 2 & 0 & -4 & 6 \\ 0 & 5 & 9 & -12 \\ 0 & 0 & 2.4 & 3.8 \\ 0 & 0 & -0 & 47.25 \end{vmatrix}$$

Row 4 + 4.75 Row 3
(No change in *D* value)

$$= 2 \cdot 5 \cdot 2.4 \cdot 47.25 = 1134.$$

Note: Blue explanations always referring to the ***preceding determinant***

## EXAMPLE

**Evaluation of Determinants by converting to Upper Triangular Form via elementary row operations**

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

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

Row 2 - 0.5 Row 1
(No change in *D* value)
Row 3 - 0.5 Row 1
(No change in *D* value)

## EXAMPLE

$$= (-1)\begin{vmatrix} 2 & 4 & -2 & 6 \\ 0 & -1 & 3 & 1 \\ 0 & 0 & 6 & 1 \\ 0 & 2 & -6 & 3 \end{vmatrix}$$

Row 2 swap with Row 3
(Change in $D$ value by -1)

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

Row 4 +2 Row 2
(No change in $D$ value)

$$= 60$$

# THEOREM 2

**Further Properties of *n*th-Order Determinants**

**(a)–(c)** in Theorem 1 hold also for columns.

**(d) Transposition** leaves the value of a determinant unaltered, i.e., $\det \mathbf{A} = \det(\mathbf{A}^T)$

**(e) A zero row or column** renders the value of a determinant zero.

**(f ) Proportional rows or columns** render the value of a determinant zero. In particular, a determinant with two identical rows or columns has the value zero.

## Determinants and Rank

An $n \times n$ square matrix $\mathbf{A}$ has rank $n$ iff det $\mathbf{A} \neq 0$.

The above is readily concluded from the elementary row operations on a matrix.

$$
\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & \cdots & a_{nn} \end{bmatrix} \xrightarrow{\text{Elementary row operations}} \mathbf{R} = \begin{bmatrix} r_{11} & a_{12} & \cdots & \cdots & r_{1n} \\ & r_{22} & \cdots & \cdots & r_{2n} \\ & & \ddots & & \vdots \\ & & & \ddots & \vdots \\ & & & & r_{nn} \end{bmatrix}
$$

Zero diagonal element of **R,** if exists $\longleftrightarrow$ Rank $\mathbf{A} < n$

Det$\mathbf{A}$= $D$ =0

# THEOREM 3

## Rank in Terms of Determinants

- A **submatrix** is a matrix formed by selecting certain rows and columns from a bigger matrix.
- Consider an $m \times n$ matrix $\mathbf{A} = [a_{jk}]$:

1) $\mathbf{A}$ has rank $r$ iff $\mathbf{A}$ has an $r \times r$ submatrix with a nonzero determinant. Here, $r \geq 1$.
2) In the above case, any square submatrix of $\mathbf{A}$ that has more than $r$ rows, if exists, has a zero determinant value.

# THEOREM 4 Cramer's Rule

**Cramer's Theorem**
**(Solution of Linear Systems by Determinants)**

- If a linear system of $n$ equations in the same number of unknowns $x_1, \dots, x_n$

$$
\begin{aligned}
a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\
a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots&\cdots\cdots \\
a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n &= b_n
\end{aligned}
\tag{6}
$$

  has a nonzero coefficient determinant $D = \det \mathbf{A}$, the system is **consistent** with **unique solution**. (Why?)

- One can obtain this unique solution using Gaussian elimination or using the Cramer's rule

# THEOREM 4 (continued)

- The Cramer's rule solution is given by the formulas

$$\textbf{(7)} \quad x_1 = \frac{D_1}{D}, \quad x_2 = \frac{D_2}{D}, \quad \cdots, \quad x_n = \frac{D_n}{D} \qquad (\textbf{Cramer's rule})$$

where $D_k$ is the determinant obtained from $D$ by replacing the $k^{th}$ column in $D$ with a column having entries $b_1, \dots, b_n$.

- For the case that system (6) *is* ***homogeneous***, i.e., **b=0**, then:

-- *If* $D \neq 0$, Cramer's rule yield the trivial solution $x_1 = 0$, $x_2 = 0, \dots, x_n = 0$ for (6).

-- *If* $D = 0$, the ***homogeneous*** system has nontrivial solutions.

## Proof of Cramer's Solution

- Given the linear system of $n$ equations with unknowns $x_1, \ldots, x_n$

$$
\text{(6)} \qquad
\begin{aligned}
a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\
a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n &= b_n
\end{aligned}
$$

and that determinant $D$ = det $\mathbf{A}$, is nonzero, i.e., the system is **consistent** with **unique solution**.

- The Cramer's rule solution is

$$x_1 = \frac{D_1}{D}, \quad x_2 = \frac{D_2}{D}, \quad \cdots, \quad x_n = \frac{D_n}{D} \qquad \textbf{(Cramer's rule)}$$

- Just show solution for $x_1$, take linear combination of the equations as:

Cofactor

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

- With the result:

$$
D = \begin{vmatrix} a_{11} & a_{12} & \cdots & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & \cdots & a_{nn} \end{vmatrix}
\qquad
0 = \begin{vmatrix} a_{12} & a_{12} & \cdots & \cdots & a_{1n} \\ a_{22} & a_{22} & \cdots & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & \vdots \\ a_{n2} & a_{n2} & \cdots & \cdots & a_{nn} \end{vmatrix}
$$

$$
x_1(a_{11}C_{11} + a_{21}C_{21} + \cdots + a_{n1}C_{n1}) + x_2(a_{12}C_{11} + a_{22}C_{21} + \cdots + a_{n2}C_{n1})
$$
$$
+ \cdots + x_n(a_{1n}C_{11} + a_{2n}C_{21} + \cdots + a_{nn}C_{n1}) = (b_1C_{11} + b_2C_{21} + \cdots + b_nC_{n1})
$$

$$
0 = \begin{vmatrix} a_{1n} & a_{12} & \cdots & \cdots & a_{1n} \\ a_{2n} & a_{22} & \cdots & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & \vdots \\ a_{nn} & a_{n2} & \cdots & \cdots & a_{nn} \end{vmatrix}
\qquad
D_1 = \begin{vmatrix} b_1 & a_{12} & \cdots & \cdots & a_{1n} \\ b_2 & a_{22} & \cdots & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & \vdots \\ b_n & a_{n2} & \cdots & \cdots & a_{nn} \end{vmatrix}
$$

- Hence, $x_1 = \dfrac{D_1}{D}$