# Linear Control and Estimation

## Matrices

Sivakumar Balasubramanian

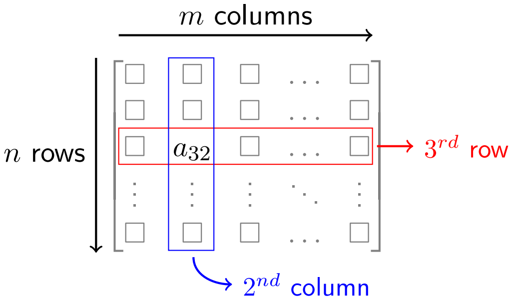
Department of Bioengineering
Christian Medical College, Bagayam
Vellore 632002

# References

- S Boyd, Applied Linear Algebra: Chapters 6, 7, 8, 10 and 11.
- G Strang, Linear Algebra: Chapters 1 and 2.

# Matrices

- **Matrices** are rectangular array of numbers. $\begin{bmatrix} 1.1 & -24 & \sqrt{2} \\ 0 & 1.12 & -5.24 \end{bmatrix}$




- Consider a matrix $A$ with $n$ rows and $m$ columns. $\begin{cases} \textbf{Tall/Skinny} & n > m \\ \textbf{Square} & n = m \\ \textbf{Wide/Fat} & n < m \end{cases}$

# Matrices

- $n$-vectors can be interpreted as $n \times 1$ matrices. These are called *column vectors*.
- A matrix with only one row is called a *row vector*, which can be referred to as $n$-row-vector. $x = \begin{bmatrix} 1.45 & -3.1 & 12.4 \end{bmatrix}$
- **Block matrices** & **Submatrices**: $A = \begin{bmatrix} B & C \\ D & E \end{bmatrix}$. What are the dimensions of the different matrices?
- Matrices are also compact way to give a set of indexed column $n$-vectors, $x_1, x_2, x_3 \ldots x_m$.

$$X = \begin{bmatrix} x_1 & x_2 & x_3 & \ldots & x_m \end{bmatrix}$$

- **Zero matrix**$= 0_{n \times m} = \begin{bmatrix} 0 & 0 & \ldots & 0 \\ 0 & 0 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 0 \end{bmatrix}$

# Matrices

- **Identity matrix** is a square $n \times n$ matrix with all zero elements, except the diagonals where all elements are 1.

$$I_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases} \qquad I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} e_1 & e_2 & e_3 \end{bmatrix}$$

- **Diagonal matrices** is a square matrix with non-zero elements on its diagonal.

$$\begin{bmatrix} 0.4 & 0 & 0 & 0 \\ 0 & -11 & 0 & 0 \\ 0 & 0 & 21 & 0 \\ 0 & 0 & 0 & 9.3 \end{bmatrix} = \textbf{diag}\left(0.4, -11, 21, 9.3\right)$$

- **Triangular matrices**: Are square matrices. *Upper triangular* $A_{ij} = 0, \forall i > j$; *Lower triangular* $A_{ij} = 0, \forall i < j$.

# Matrix operations

- **Transpose** switches the rows and columns of a matrix. $A$ is a $n \times m$ matrix, then its transpose is represented by $A^T$, which is a $m \times n$ matrix.

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} \implies A^T = \begin{bmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \\ a_{13} & a_{23} \end{bmatrix}$$

  Transpose converts between column and row vectors.

  What is the transpose of a block matrix? $A = \begin{bmatrix} B & C \\ D & E \end{bmatrix}$

- **Matrix addition** can only be carried out with matrices of same size. Like vectors we perform element wise addition.

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \begin{bmatrix} a_{11} + b_{11} & a_{12} + b_{12} \\ a_{21} + b_{21} & a_{22} + b_{22} \end{bmatrix}$$

# Matrix operations

- **Properties of matrix addition**:
  - *Commutative*: $A + B = B + A$
  - *Associative*: $(A + B) + C = A + (B + C)$
  - *Addition with zero matrix*: $A + 0 = 0 + A$
  - *Transpose of sum*: $(A + B)^T = A^T + B^T$
- **Scalar multiplication** Each element of the matrix gets multiplied by the scalar.

$$\alpha \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{bmatrix} \alpha a_{11} & \alpha a_{12} \\ \alpha a_{21} & \alpha a_{22} \end{bmatrix}$$

- We will mostly only deal with matrices with real entries. Such matrices are elements of the set $\mathbb{R}^{n \times m}$.
- Given the aforementioned matrix operations and their properties, is $\mathbb{R}^{n \times m}$ a vector space?

# Matrix multiplication

- It is possible to *multiply* two matrices $A \in \mathbb{R}^{n \times p}$ and $B \in \mathbb{R}^{p \times m}$ through *matrix multiplication* procedure.
- There is a product matrix $C := AB \in \mathbb{R}^{n \times m}$, if the number of columns of $A$ is equal to the number of rows of $B$.

$$C_{ij} := \sum_{k=1}^{p} A_{ik} B_{kj} \quad \forall i \in \{1, \ldots n\} \;\; \& \;\; j \in \{1 \ldots m\}$$

- *Inner product* is a special case of matrix multiplication between a *row vector* and a *column vector*.

$$x^T y = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}^T \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} x_1 & x_2 & \ldots & x_n \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \sum_{i=1}^{n} x_i y_i$$

# Matrix multiplication

- Consider a matrix $A \in \mathbb{R}^{n \times m}$ and a $m$-vector $x \in \mathbb{R}^m$. We can multiply $A$ and $x$ to obtain $y = Ax \in \mathbb{R}^n$.

$$y = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix} = \begin{bmatrix} \tilde{a}_1^T \\ \tilde{a}_2^T \\ \vdots \\ \tilde{a}_n^T \end{bmatrix} x = \begin{bmatrix} \tilde{a}_1^T x \\ \tilde{a}_2^T x \\ \vdots \\ \tilde{a}_n^T x \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^m a_{1i} x_i \\ \sum_{i=1}^m a_{2i} x_i \\ \vdots \\ \sum_{i=1}^m a_{ni} x_i \end{bmatrix}$$

$$y = \sum_{i=1}^{m} x_i \begin{bmatrix} a_{1i} \\ a_{2i} \\ \vdots \\ a_{ni} \end{bmatrix} = x_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{n2} \end{bmatrix} + \ldots + x_m \begin{bmatrix} a_{1m} \\ a_{2m} \\ \vdots \\ a_{nm} \end{bmatrix}$$

- Multiplying a matrix $A$ by a column vector $x$ produces a linear combination of the columns of matrix $A$. The column mixture is provided by $x$.

# Matrix multiplication

- We see a similar process in play when we multiply a row vector $x^T \in \mathbb{R}^n$ by a matrix $A \in \mathbb{R}^{n \times m}$.

$$y = x^T A = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}^T \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{bmatrix} = x^T \begin{bmatrix} a_1 & a_2 & \dots & a_m \end{bmatrix}$$

$$y = \begin{bmatrix} x^T a_1 & x^T a_2 & \dots & x^T a_m \end{bmatrix} = \sum_{i=1}^{n} x_i \begin{bmatrix} a_{i1} & a_{i2} & \dots & a_{im} \end{bmatrix}$$

- Multiplying a row vector $x$ by a matrix $A$ produces a linear combination of the row of matrix $A$. The row mixture is provided by $x$.

# Matrix multiplication

- Multiplying two matrices $A \in \mathbb{R}^{n \times p}$ and $B \in \mathbb{R}^{p \times m}$, we have $C \in \mathbb{R}^{n \times m}$,

$$C = AB = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1p} \\ a_{21} & a_{22} & \dots & a_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{np} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1m} \\ b_{21} & b_{22} & \dots & b_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ b_{p1} & b_{p2} & \dots & b_{pm} \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} & \dots & c_{1m} \\ c_{21} & c_{22} & \dots & c_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ c_{p1} & c_{n2} & \dots & c_{nm} \end{bmatrix}$$

- **Inner product interpretation**: $c_{ij} = \tilde{a}_i^T b_j$, $i \in \{1 \dots n\}, j \in \{1 \dots m\}$
- **Column interpretation**: $C = A \begin{bmatrix} b_1 & b_2 & \dots & b_m \end{bmatrix} = \begin{bmatrix} Ab_1 & Ab_2 & \dots & Ab_m \end{bmatrix}$
- **Row interpretation** : $C = \begin{bmatrix} \tilde{a}_1^T \\ \tilde{a}_2^T \\ \vdots \\ \tilde{a}_n^T \end{bmatrix} B = \begin{bmatrix} \tilde{a}_1^T B \\ \tilde{a}_2^T B \\ \vdots \\ \tilde{a}_n^T B \end{bmatrix}$

# Matrix multiplication

- **Outer product interpretation** Consider two $n$-vectors $x, y \in \mathbb{R}^n$. The *outer product* is defined as,

$$xy^T = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix} \begin{bmatrix} y_1 & y_2 & y_3 & \dots & y_n \end{bmatrix} = \begin{bmatrix} x_1y_1 & x_1y_2 & x_1y_3 & \dots & x_1y_n \\ x_2y_1 & x_2y_2 & x_2y_3 & \dots & x_2y_n \\ x_3y_1 & x_3y_2 & x_3y_3 & \dots & x_3y_n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_ny_1 & x_ny_2 & x_ny_3 & \dots & x_ny_n \end{bmatrix}$$

- We can represent the product between two matrices as the sum of outer products between the columns and $A$ and rows of $B$.

$$AB = \begin{bmatrix} a_1 & a_2 & a_3 & \dots & a_p \end{bmatrix} \begin{bmatrix} \tilde{b}_1^T \\ \tilde{b}_2^T \\ \tilde{b}_3^T \\ \vdots \\ \tilde{b}_p^T \end{bmatrix} = \sum_{i=1}^{p} a_i \tilde{b}_i^T$$

# Properties of matrix multiplication

- **Not commutative**: $AB \neq BA$
  The product of two matrices might not alwasys be defined. When it is defined, $AB$ and $BA$ need not match.
- **Distributive**: $A\left(B+C\right) = AB + BC$ and $\left(A+B\right)C = AC + BC$
- **Associative**: $A\left(BC\right) = \left(AB\right)C$
- **Transpose**: $\left(AB\right)^T = B^T A^T$
- **Scalar product**: $\alpha\left(AB\right) = \left(\alpha A\right)B = A\left(\alpha B\right)$

# Linear equations

- Matrices present a compact way to represent a set of linear equations. Consider the following,

$$\left.\begin{aligned} a_{11}x_1 + a_{12}x_2 \ldots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 \ldots + a_{2n}x_n &= b_2 \\ a_{31}x_1 + a_{32}x_2 \ldots + a_{3n}x_n &= b_3 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 \ldots + a_{mn}x_n &= b_m \end{aligned}\right\} \longrightarrow Ax = b, \;\; A \in \mathbb{R}^{m \times n}, \;\; x \in \mathbb{R}^n, \;\; b \in \mathbb{R}^m$$

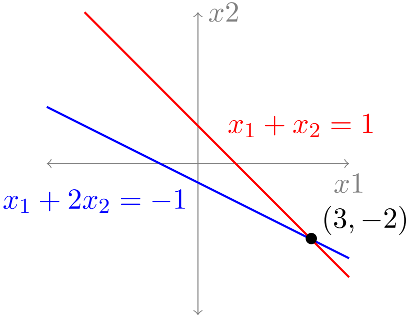
$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \ldots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \ldots & a_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \ldots & a_{mn} \end{bmatrix} \quad x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \quad b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

# Geometry of linear equations

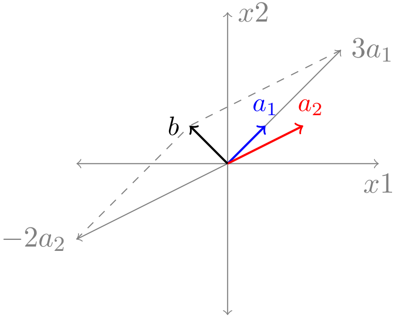
$$\left.\begin{aligned} x_1 + 2x_2 &= -1 \\ x_1 + x_2 &= 1 \end{aligned}\right\} \longrightarrow \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

Two ways to view this: **row view** and the **column view**.

**Traditional (row) view**




**Column view**

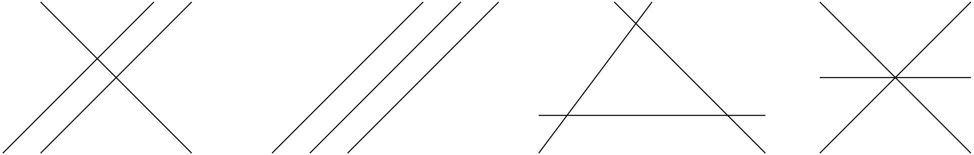


# Solving linear equations

$$Ax = b, \;\; A \in \mathbb{R}^{m \times n}, \;\; x \in \mathbb{R}^n, \;\; b \in \mathbb{R}^m$$

- **Three possible situations:** No solution, Infinitely many solutions, or Unique Solution.
- When do have infinitely many or no solutions? In $\mathbb{R}^3$, we can visualize the different situations.

Two parallel planes | Three parallel planes | No intersection | Line intersection



# Solving linear equations: Gaussian Elimination

$$
\begin{aligned}
a_{11}x_1 + a_{12}x_2 \ldots + a_{1n}x_n &= b_1 : E_1 \\
a_{21}x_1 + a_{22}x_2 \ldots + a_{2n}x_n &= b_2 : E_2 \\
a_{31}x_1 + a_{32}x_2 \ldots + a_{3n}x_n &= b_3 : E_3 \\
&\vdots \\
a_{m1}x_1 + a_{m2}x_2 \ldots + a_{mn}x_n &= b_m : E_m
\end{aligned}
$$

- Gaussian elimination is a systematic way of simplifying the above equations to an equivalent system that can be easily solved.
- Three simple operations are repeatedly performed:
    - Interchanging of equations $E_i$ and $E_j$.
    - Replacing equation $E_i$ by $\alpha E_i$, $\alpha \neq 0$.
    - Replacing equation $E_j$ by $E_j + \alpha E_i$, $\alpha \neq 0$.
- These three operations do not change the solution of the given linear system.

# Solving linear equations: Gaussian Elimination

**Augmented matrix**: $$\left[\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ a_{31} & a_{32} & \cdots & a_{3n} & b_3 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array}\right]$$

- We can work with the augmented matrix instead of the equations.
- Gaussian elimination is carried out on the entire matrix.
- The matrix is simplified to a point, from where one can easily:
  - find out the nature of the solutions for the system of equations; and
  - find the solution (with a bit of extra work), if they exist.

# Solving linear equations: Gaussian Elimination

$$\left.\begin{array}{rcrcrcr} x_1 & + & 2x_2 & - & x_3 & = & 1 \\ 2x_1 & + & 3x_2 & + & 4x_3 & = & 4 \\ -2x_1 & - & 4x_2 & + & x_3 & = & -3 \end{array}\right\} \longrightarrow \left[\begin{array}{rrr|r} 1 & 2 & -1 & 1 \\ 2 & 3 & 4 & 4 \\ -2 & -4 & 1 & -3 \end{array}\right]$$

## Gaussian Elimination

$$\left[\begin{array}{rrr|r} \underline{1} & 2 & -1 & 1 \\ 2 & 3 & 4 & 4 \\ -2 & -4 & 1 & -3 \end{array}\right] \longrightarrow \left[\begin{array}{rrr|r} \underline{1} & 2 & -1 & 1 \\ 0 & -1 & 6 & 2 \\ -2 & -4 & 1 & -3 \end{array}\right] \longrightarrow \left[\begin{array}{rrr|r} \underline{1} & 2 & -1 & 1 \\ 0 & \underline{-1} & 6 & 2 \\ 0 & 0 & \underline{-1} & -1 \end{array}\right]$$

Now, we can perform **back substitution** on this triangularized system of linear equations,

$$x_3 = 1; \;\; x_2 = 4; \;\; x_1 = -6$$

We can continue the simplification process through the **Gauss-Jordan** method.

# Solving linear equations: Gauss-Jordan Method

Continue the elimination upwards until all elements, except the ones in the main diagonal, are zero.

$$\left[\begin{array}{ccc|c} 1 & 2 & -1 & 1 \\ 0 & -1 & 6 & 2 \\ 0 & 0 & -1 & -1 \end{array}\right] \longrightarrow \left[\begin{array}{ccc|c} 1 & 2 & -1 & 1 \\ 0 & -1 & 0 & -4 \\ 0 & 0 & 1 & 1 \end{array}\right] \longrightarrow \left[\begin{array}{ccc|c} 1 & 2 & -1 & 1 \\ 0 & 1 & 0 & 4 \\ 0 & 0 & 1 & 1 \end{array}\right] \longrightarrow \left[\begin{array}{ccc|c} 1 & 2 & 0 & 2 \\ 0 & 1 & 0 & 4 \\ 0 & 0 & 1 & 1 \end{array}\right]$$

$$\left[\begin{array}{ccc|c} 1 & 2 & 0 & 2 \\ 0 & 1 & 0 & 4 \\ 0 & 0 & 1 & 1 \end{array}\right] \longrightarrow \left[\begin{array}{ccc|c} 1 & 0 & 0 & -6 \\ 0 & 1 & 0 & 4 \\ 0 & 0 & 1 & 1 \end{array}\right] \implies x_1 = -6; \;\; x_2 = 4; \;\; x_3 = 1;$$

Everything worked out well without any problems. What can go wrong here?

Try solving the these systems, $\left[\begin{array}{ccc|c} 1 & 2 & -1 & 1 \\ 2 & 3 & 4 & 4 \\ -2 & -4 & 2 & -3 \end{array}\right], \left[\begin{array}{ccc|c} 1 & 2 & -1 & 1 \\ 2 & 3 & 4 & 4 \\ -2 & -4 & 2 & -2 \end{array}\right]$

What is the difference between these two systems?

# Solving linear equations: Rectangular systems and Row Echelon Form

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 \ldots + a_{1n}x_n &= b_1 : E_1 \\ a_{21}x_1 + a_{22}x_2 \ldots + a_{2n}x_n &= b_2 : E_2 \\ a_{31}x_1 + a_{32}x_2 \ldots + a_{3n}x_n &= b_3 : E_3 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 \ldots + a_{mn}x_n &= b_m : E_m \end{aligned} \longrightarrow \left[\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ a_{31} & a_{32} & \cdots & a_{3n} & b_3 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array}\right]$$

Consider the following example,

$$\left[\begin{array}{ccccc|c} \underline{1} & -2 & 1 & 0 & 1 & 1 \\ 2 & -4 & 1 & -1 & -2 & 2 \\ -1 & 2 & 1 & 1 & 2 & -1 \end{array}\right] \longrightarrow \left[\begin{array}{ccccc|c} \underline{1} & -2 & 1 & 0 & 1 & 1 \\ 0 & 0 & \underline{-1} & -1 & -4 & 0 \\ 0 & 0 & 2 & 1 & 3 & 0 \end{array}\right] \longrightarrow \left[\begin{array}{ccccc|c} \underline{1} & -2 & 1 & 0 & 1 & 1 \\ 0 & 0 & \underline{-1} & -1 & -4 & 0 \\ 0 & 0 & 0 & \underline{-1} & -5 & 0 \end{array}\right]$$

# Solving linear equations: Rectangular systems and Row Echelon Form

$$\left[\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ a_{31} & a_{32} & \cdots & a_{3n} & b_3 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array}\right] \longrightarrow \begin{bmatrix} \underline{*} & * & * & * & * & * & * \\ 0 & 0 & \underline{*} & * & * & * & * \\ 0 & 0 & 0 & \underline{*} & * & * & * \\ 0 & 0 & 0 & 0 & \underline{*} & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & \underline{*} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

**Things to notice about the echelon form:**

- If a particualr row consists entirely of zeros, then all rows below that row also contain entirely of zeros.
- If the first non-zero entry in the $i^{th}$ row occurs in the $j^{th}$ position, then all elements below the $i^{th}$ row are zero from columns $1$ to $j$.

Columns containing pivot are called the **basic columns**.

**Rank of a matrix** $A$ is defined at the number of basic columns in the row echelon form of the matrix A.

# Solving linear equations: Reduced Row Echelon Form

$$\begin{bmatrix} \underline{*} & * & * & * & * & * & * \\ 0 & 0 & \underline{*} & * & * & * & * \\ 0 & 0 & 0 & \underline{*} & * & * & * \\ 0 & 0 & 0 & 0 & \underline{*} & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & \underline{*} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \xrightarrow{\text{Gauss-Jordan}} \begin{bmatrix} \underline{1} & * & 0 & 0 & 0 & * & 0 \\ 0 & 0 & \underline{1} & 0 & 0 & * & 0 \\ 0 & 0 & 0 & \underline{1} & 0 & * & 0 \\ 0 & 0 & 0 & 0 & \underline{1} & * & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \underline{1} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\left[\begin{array}{ccccc|c} \underline{1} & -2 & 1 & 0 & 1 & 1 \\ 0 & 0 & \underline{-1} & -1 & -4 & 0 \\ 0 & 0 & 0 & \underline{-1} & -5 & 0 \end{array}\right] \longrightarrow \left[\begin{array}{ccccc|c} \underline{1} & -2 & 0 & -1 & -3 & 1 \\ 0 & 0 & \underline{1} & 1 & 4 & 0 \\ 0 & 0 & 0 & \underline{-1} & -5 & 0 \end{array}\right] \longrightarrow \left[\begin{array}{ccccc|c} \underline{1} & -2 & 0 & 0 & 2 & 1 \\ 0 & 0 & \underline{1} & 0 & -1 & 0 \\ 0 & 0 & 0 & \underline{1} & 5 & 0 \end{array}\right]$$

- All non-basic columns can be represented as a linear combination of the basic columns.
- A non-basic columns is a linear combination of only the columns before it.
- Scaling factors for each basic comlumns is determined by the corresponding elements of the non-basic columns.

**The reduced row echelon form reveals structure in the original matrix $A$.**

# Solving linear equations: Homogenous Systems

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 \ldots + a_{1n}x_n &= 0 \\ a_{21}x_1 + a_{22}x_2 \ldots + a_{2n}x_n &= 0 \\ a_{31}x_1 + a_{32}x_2 \ldots + a_{3n}x_n &= 0 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 \ldots + a_{mn}x_n &= 0 \end{aligned} \longrightarrow \left[\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & 0 \\ a_{21} & a_{22} & \cdots & a_{2n} & 0 \\ a_{31} & a_{32} & \cdots & a_{3n} & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & 0 \end{array}\right]$$

Consider the following case,

$$\left[\begin{array}{rrrrr|r} 1 & -2 & 1 & 0 & 1 & 0 \\ 2 & -4 & 1 & -1 & -2 & 0 \\ -1 & 2 & 1 & 1 & 2 & 0 \end{array}\right] \longrightarrow \left[\begin{array}{rrrrr|r} 1 & -2 & 0 & 0 & 2 & 0 \\ 0 & 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 & 5 & 0 \end{array}\right]$$

$$\begin{aligned} x_1 - 2x_2 + 2x_5 &= 0 \\ x_3 - x_5 &= 0 \\ x_4 + 5x_5 &= 0 \end{aligned} \longrightarrow \begin{aligned} x_1 &= 2x_2 - 2x_5 \\ x_3 &= x_5 \\ x_4 &= -5x_5 \end{aligned} \longrightarrow \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 2x_2 - 2x_5 \\ x_2 \\ x_5 \\ 5x_5 \\ x_5 \end{bmatrix} = x_2 \begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + x_5 \begin{bmatrix} -2 \\ 0 \\ 1 \\ -5 \\ 1 \end{bmatrix}$$

## Solving linear equations: Homogenous Systems

- $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = x_2 \begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + x_5 \begin{bmatrix} -2 \\ 0 \\ 1 \\ -5 \\ 1 \end{bmatrix}$ represents the general solution of the system of equations.
- In general, any system $[A \mid 0]$ with $rank\,(A) = r$ and $m < n$ has the general solution of the form,

  $$x = x_{f_1}\mathbf{h}_1 + x_{f_2}\mathbf{h}_2 + \ldots + x_{f_{n-r}}\mathbf{h}_{n-r}$$

  where, the variables $x_{f_1}, x_{f_2}, \ldots, x_{f_{n-r}}$ are called the **free variables**.
- Free variables are the one corresponding to the non-basic columns; the variable variables corresponding to the basic columns are the **basic variables**.
- When does a homogenous system have a unique solution solution? $\longrightarrow rank\,(A) = m$.

# Solving linear equations: Non-homogenous Systems

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 \ldots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 \ldots + a_{2n}x_n &= b_2 \\ a_{31}x_1 + a_{32}x_2 \ldots + a_{3n}x_n &= b_3 \quad \longrightarrow [A \mid b] \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 \ldots + a_{mn}x_n &= b_m \end{aligned}$$

Consider the following case,

$$\left[\begin{array}{rrrrr|r} 1 & -2 & 1 & 0 & 1 & 1 \\ 2 & -4 & 1 & -1 & -2 & 2 \\ -1 & 2 & 1 & 1 & 2 & -1 \end{array}\right] \longrightarrow \left[\begin{array}{rrrrr|r} 1 & -2 & 0 & 0 & 2 & 1 \\ 0 & 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 & 5 & 0 \end{array}\right]$$

$$\begin{aligned} x_1 - 2x_2 + 2x_5 &= 1 \\ x_3 - x_5 &= 0 \\ x_4 + 5x_5 &= 0 \end{aligned} \longrightarrow \begin{aligned} x_1 &= 1 + 2x_2 - 2x_5 \\ x_3 &= x_5 \\ x_4 &= -5x_5 \end{aligned} \longrightarrow \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 1 + 2x_2 - 2x_5 \\ x_2 \\ x_5 \\ 5x_5 \\ x_5 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + x_5 \begin{bmatrix} -2 \\ 0 \\ 1 \\ -5 \\ 1 \end{bmatrix}$$

The general solution of a non-homogenous system is sum of the particular solution and the general solution of the associated homogenous system.

# Solving linear equations: Non-homogenous Systems

- The general solution for $[A \mid 0]$ with $rank\,(A) = r$,

$$x = \mathbf{p} + x_{f_1}\mathbf{h}_1 + x_{f_2}\mathbf{h}_2 + \ldots + x_{f_{n-r}}\mathbf{h}_{n-r}$$

  where, $\mathbf{p}$ is the particular solution and $x_{f_1}, x_{f_2}, \ldots, x_{f_{n-r}}$ are the free variables.
- When do we have a unique solution to this system? $\longrightarrow rank\,(A) = m$.
- What about the case when there are no solutions? When does that happen? $\longrightarrow$ *When the system is not consistent.*

$$\left[\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ a_{31} & a_{32} & \cdots & a_{3n} & b_3 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array}\right] \longrightarrow \left[\begin{array}{cccccc|c} 1 & * & 0 & 0 & 0 & * & c_1 \\ 0 & 0 & 1 & 0 & 0 & * & c_2 \\ 0 & 0 & 0 & 1 & 0 & * & c_3 \\ 0 & 0 & 0 & 0 & 1 & * & c_4 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & c_m \end{array}\right]$$

  There is a problem when $c_m \neq 0$
- The augmented matrix $[A \mid b]$ has the same number of basic columns as $A$.
- $[A \mid b] \rightarrow [E \mid c]$: $c$ is a non-basic column.
- $rank\,(A) = rank\,([A \mid b])$

# $LU$ Factorization of a Matrix

- A major theme of matrix algebra is to decompose matrices into simpler components that provide insights into the nature of the matrix.
- A full rank square matrix $A \in \mathbb{R}^{n \times n}$ can be decomosed into the product of a lower triangular and an upper triangular matrix.
- Matrices associated with the three elementary operations:

| **Inter-changing rows 2 and 4** | **Scaling row 2** | **Adding a multiple of row 2 to row 3** |
|---|---|---|
| $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$ | $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \alpha & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$ | $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & \alpha & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$ |

# $LU$ Factorization of a Matrix

- Consider the case: $A = \begin{bmatrix} 1 & -1 & 2 \\ 2 & 0 & 2 \\ 4 & 2 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 4 & 3 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 & 2 \\ 0 & 2 & -2 \\ 0 & 0 & -1 \end{bmatrix} = LU$
- LU factorization can be done only when no zero pivot is encountered during the Guassian elimination process.
- $Ax = b$ becomes $LUx = b$: This is decomposed into two triangular systems, $Ux = y, \;\; Ly = b$. First solve $Ly = b$ and then solve $Ux = y$
- Properties:
  - Diagonal elements of $L$ are $1$, and $U$ are not equal to zero.
  - $U$ is the final result of Guassian elimination, and $L$ is the matrix that reverses this process.
  - Element $l_{ij}$ of $L$ is the multiple of row $j$ used to eliminate the $a_{ij}$ element of $A$.
- Uses of the LU factorization:
  - Solving $Ax = b_i$ for several $b_i$s. $LU$ need to be calculated only once.
  - Factorization requires not extra space.

# $PA = LU$ Factorization of a Matrix

- Consider the case: $A = \begin{bmatrix} 1 & 1 & 0 \\ 2 & 2 & 1 \\ 3 & 4 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \neq LU$
- It turns out the second pivot become zero after the first elimination step, so $LU$ factorization cannot be done on $A$.
- The following however fixes this issue,

$$PA = LU$$

  where, $P$ is the permunation matrix, which is the elementary matrix for row exchanges.
- In the current example, the following allows matrix factorization.

$$PA = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 2 & 2 & 1 \\ 3 & 4 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = LU$$

# Linear functions

- We had earlier seen linear functions of the form $f : \mathbb{R}^n \mapsto \mathbb{R}$, which had the form,

$$y = f(x) = w^T x; \;\; w, x \in \mathbb{R}^n, \;\; y \in \mathbb{R}$$

- A generalization of this is when the range of the function is not in $\mathbb{R}$ but in $\mathbb{R}^n$:

$$y = f(x); \;\; x \in \mathbb{R}^n, \;\; y \in \mathbb{R}^m$$

- Such a function has a natural representation of the form $y = Ax, \;\; A \in \mathbb{R}^{m \times n}$. Can you prove that $y = Ax$ is linear?
- Any linear function can be represented in the form $y = Ax$. So, matrices can be viewed as representing a linear transformation.

# Another look at matrix multiplication

Why does matrix multiplication have this strange definition?

Consider the following two functions,

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = f\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} ax_1 + bx_2 \\ cx_1 + dx_2 \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = g\left(\begin{bmatrix} u_1 \\ u_2 \end{bmatrix}\right) = \begin{bmatrix} \alpha u_1 + \beta u_2 \\ \gamma u_1 + \delta u_2 \end{bmatrix} = \begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

$$h(u) = f(g(u)) = f\left(\begin{bmatrix} \alpha u_1 + \beta u_2 \\ \gamma u_1 + \delta u_2 \end{bmatrix}\right) = \begin{bmatrix} a\alpha u_1 + a\beta u_2 + b\gamma u_1 + b\delta u_2 \\ c\alpha u_1 + c\beta u_2 + d\gamma u_1 + d\delta u_2 \end{bmatrix}$$

$$= \begin{bmatrix} (a\alpha + b\gamma) u_1 + (a\beta + b\delta) u_2 \\ (c\alpha + d\gamma) u_1 + (c\beta + d\delta) u_2 \end{bmatrix} = \begin{bmatrix} a\alpha + b\gamma & a\beta + b\delta \\ c\alpha + d\gamma & c\beta + d\delta \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

$$\implies \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix} = \begin{bmatrix} a\alpha + b\gamma & a\beta + b\delta \\ c\alpha + d\gamma & c\beta + d\delta \end{bmatrix}$$

This definition of matrix multiplication is the most natural for dealing with composition of linear functions. It also turns to out to be the most useful.

# Four Fundamental Subspaces

- $C(A)$: **Column Space of** $A$ – the span of the columns of $A$.

$$C(A) = \{Ax \mid x \in \mathbb{R}^n\} \subseteq \mathbb{R}^m$$

- $N(A)$: **Nullspace of** $A$ – the set of all $x \in \mathbb{R}^n$ that are mapped to zero.

$$N(A) = \{x \mid Ax = 0\} \subseteq \mathbb{R}^n$$

- $C(A^T)$: **Row Space of** $A$ – the span of the rows of $A$.

$$C(A^T) = \{A^T y \mid y \in \mathbb{R}^m\} \subseteq \mathbb{R}^n$$

- $N(A^T)$: **Nullspace of** $A^T$ – the set of all $y \in \mathbb{R}^m$ that are mapped to zero by $A^T$.

$$N(A^T) = \{y \mid A^T y = 0\} \subseteq \mathbb{R}^m$$

This is also called the **left nullspace** of $A$.

# Linear Independence

- Given a set of vectors $\{v_1, v_2, \dots v_n\}, \;\; v_i \in \mathbb{R}^m$, how can we determine if this set is linear independent? Remember the Gram-Schmit algorithm?
- We need to verify, $\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n = 0$

$$\left.\begin{bmatrix} v_1 & v_2 & \cdots & v_n \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = V\alpha = 0 \right\} N(A) = \{0\}, \;\; rank(V) = n$$

- This is also equivalent to saying that when the $rank(A) = n \implies$ the columns of $A$ form an independent set of vectors.
- When do the rows of $A$ form an independent set?
- What about both rows and columns? When does that happen?

# Dimension and basis of the four fundamental subspaces

$$A = \begin{bmatrix} 1 & -2 & 1 & 0 & 1 \\ 2 & -4 & 1 & -1 & -2 \\ -1 & 2 & 1 & 1 & 2 \\ 2 & -4 & -2 & -2 & -4 \end{bmatrix}; \quad EA = R$$

$$\underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ -2 & 1 & 0 & 0 \\ -1 & 2 & 1 & 1 \\ 2 & -4 & -2 & -2 \end{bmatrix}}_{E} A = \underbrace{\begin{bmatrix} 1 & -2 & 1 & 0 & 1 \\ 0 & 0 & -1 & -1 & -4 \\ 0 & 0 & 0 & -1 & -5 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}}_{R}$$

**Pivot columns of** $A$: $\left\{ \begin{bmatrix} 1 \\ 2 \\ -1 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \\ -2 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \\ 1 \\ -2 \end{bmatrix} \right\}$

**Nullspace of** $A$: $x_2 h_1 + x_5 h_2; \;\; h_1 \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \;\; h_2 = \begin{bmatrix} 2 \\ 0 \\ -1 \\ -5 \\ 1 \end{bmatrix}$

We can restructure $EA = R \rightarrow \begin{bmatrix} E_1 \\ E_2 \end{bmatrix} A = \begin{bmatrix} R_1 \\ 0 \end{bmatrix}$

Consider the matrix $A \in \mathbb{R}^{m \times n}$

- **Column space** $N(A)$
  - dim $C(A) = rank\,(A) = r$
  - Basis of $C(A)$ = Pivot colums of $A$.
- **Nullspace** $N(A)$
  - dim $N(A) = n - r$
  - Basis of $N(A) = \{h_1, h_2 \dots h_{n-r}\}$.
- **Row space** $C(A^T)$
  - dim $C(A) = rank\left(A^T\right) = rank\,(A) = r$
  - Basis of $C(A^T)$ = Colums of $R_1^T$.
- **Nullspace** $N(A)$
  - dim $N(A) = n - r$
  - Basis of $N(A)$ = Colums of $E_2^T$

# Matrix Inverse

- Consider the square matrix $A \in \mathbb{R}^{n \times n}$. $B \in \mathbb{R}^{n \times n}$ is the inverse of $A$, if $AB = BA = I_n$, and $B$ is represented as $A{-}1$.
- Not all matrices have inverses. A matrix with an inverse is called **non-singular**, and a matrix that does not have an inverse is called **singular**.
- For a non-singular matrix $A$, $A^{-1}$ is unique. $A^{-1}$ is both the left and right inverse.
- A matrix $A$ has an inverse, if and only if $A$ is full rank, i.e. $rank\,(A) = n$
- The inverse of a non-signular matrix can be determined through Gauss-Jordan method.
  $[A|I] \xrightarrow{\text{Gauss-Jordan}} [I|A^{-1}]$
- $Ax = b$ can be solved as follows, $x = A^{-1}b$. *It is never solved like this in practice.*
- Inverse of product of matrices, $(AB)^{-1} = B^{-1}A^{-1}$
- $\left(A^{-1}\right)^{-1} = A$ and $\left(A^{-1}\right)^{T} = \left(A^{T}\right)^{-1}$