

Linear Systems

Matrices

Sivakumar Balasubramanian

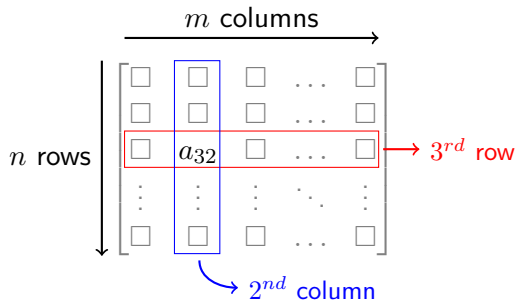
Department of Bioengineering
Christian Medical College, Bagayam
Vellore 632002

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. $\begin{cases} \text{ Tall/Skinny} & n > m \\ \text{ Square} & n = m \\ \text{ Wide/Fat} & n < m \end{cases}$

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. $\mathbf{x} = [1.45 \quad -3.1 \quad 12.4]$
- ▶ **Block matrices & Submatrices:** $\mathbf{A} = \begin{bmatrix} \mathbf{B} & \mathbf{C} \\ \mathbf{D} & \mathbf{E} \end{bmatrix}$. What are the dimensions of the different matrices?
- ▶ Matrices are also compact way to give a set of indexed column n -vectors, $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3 \dots \mathbf{x}_m$.

$$\mathbf{X} = [\mathbf{x}_1 \quad \mathbf{x}_2 \quad \mathbf{x}_3 \quad \dots \quad \mathbf{x}_m]$$

▶ **Zero matrix** $= \mathbf{0}_{n \times m} =$

$$\begin{bmatrix} 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{bmatrix}$$

Matrices

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

$$\mathbf{I}_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases} \quad \mathbf{I}_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = [\mathbf{e}_1 \quad \mathbf{e}_2 \quad \mathbf{e}_3]$$

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

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

- ▶ **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. \mathbf{A} is a $n \times m$ matrix, then its transpose is represented by \mathbf{A}^T , which is a $m \times n$ matrix.

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

Transpose converts between column and row vectors.

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

- ▶ **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*: $\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$
- *Associative*: $(\mathbf{A} + \mathbf{B}) + \mathbf{C} = \mathbf{A} + (\mathbf{B} + \mathbf{C})$
- *Addition with zero matrix*: $\mathbf{A} + \mathbf{0} = \mathbf{0} + \mathbf{A} = \mathbf{A}$
- *Transpose of sum*: $(\mathbf{A} + \mathbf{B})^T = \mathbf{A}^T + \mathbf{B}^T$

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

$$c_{ij} := \sum_{k=1}^p a_{ik} b_{kj} \quad \forall i \in \{1, \dots, n\} \quad \& \quad j \in \{1 \dots m\}$$

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

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

Matrix multiplication

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

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

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

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

Matrix multiplication

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

$$\mathbf{y} = \mathbf{x}^T \mathbf{A} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}^T \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{bmatrix} = \mathbf{x}^T [\mathbf{a}_1 \quad \mathbf{a}_2 \quad \dots \quad \mathbf{a}_m]$$

$$\mathbf{y} = [\mathbf{x}^T \mathbf{a}_1 \quad \mathbf{x}^T \mathbf{a}_2 \quad \dots \quad \mathbf{x}^T \mathbf{a}_m] = \sum_{i=1}^n x_i [a_{i1} \quad a_{i2} \quad \dots \quad a_{im}]$$

- ▶ Multiplying a row vector \mathbf{x} by a matrix \mathbf{A} produces a linear combination of the row of matrix \mathbf{A} . The row mixture is provided by \mathbf{x} .

Matrix multiplication

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

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

- **Inner product interpretation:** $c_{ij} = \tilde{\mathbf{a}}_i^T \mathbf{b}_j$, $i \in \{1 \dots n\}, j \in \{1 \dots m\}$
- **Column interpretation:** $\mathbf{C} = \mathbf{A} [\mathbf{b}_1 \quad \mathbf{b}_2 \quad \dots \quad \mathbf{b}_m] = [\mathbf{Ab}_1 \quad \mathbf{Ab}_2 \quad \dots \quad \mathbf{Ab}_m]$

- **Row interpretation :** $\mathbf{C} = \begin{bmatrix} \tilde{\mathbf{a}}_1^T \\ \tilde{\mathbf{a}}_2^T \\ \vdots \\ \tilde{\mathbf{a}}_n^T \end{bmatrix} \mathbf{B} = \begin{bmatrix} \tilde{\mathbf{a}}_1^T \mathbf{B} \\ \tilde{\mathbf{a}}_2^T \mathbf{B} \\ \vdots \\ \tilde{\mathbf{a}}_n^T \mathbf{B} \end{bmatrix}$

Matrix multiplication

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

$$\mathbf{xy}^T = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix} \begin{bmatrix} y_1 & y_2 & y_3 & \dots & y_n \end{bmatrix} = \begin{bmatrix} x_1 y_1 & x_1 y_2 & x_1 y_3 & \dots & x_1 y_n \\ x_2 y_1 & x_2 y_2 & x_2 y_3 & \dots & x_2 y_n \\ x_3 y_1 & x_3 y_2 & x_3 y_3 & \dots & x_3 y_n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_n y_1 & x_n y_2 & x_n y_3 & \dots & x_n y_n \end{bmatrix}$$

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

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

Properties of matrix multiplication

- ▶ **Not commutative:** $\mathbf{AB} \neq \mathbf{BA}$

The product of two matrices might not always be defined. When it is defined, \mathbf{AB} and \mathbf{BA} need not match.

- ▶ **Distributive:** $\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{BC}$ and $(\mathbf{A} + \mathbf{B})\mathbf{C} = \mathbf{AC} + \mathbf{BC}$

- ▶ **Associative:** $\mathbf{A}(\mathbf{BC}) = (\mathbf{AB})\mathbf{C}$

- ▶ **Transpose:** $(\mathbf{AB})^T = \mathbf{B}^T \mathbf{A}^T$

- ▶ **Scalar product:** $\alpha(\mathbf{AB}) = (\alpha\mathbf{A})\mathbf{B} = \mathbf{A}(\alpha\mathbf{B})$

Linear equations

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

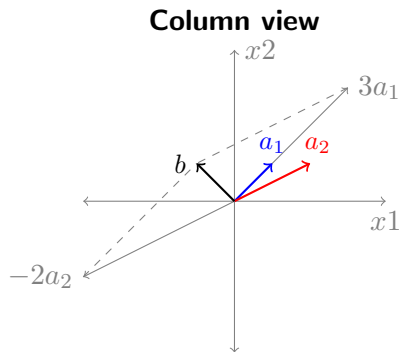
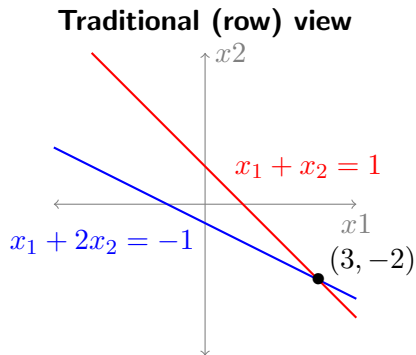
$$\left. \begin{array}{l} a_{11}x_1 + a_{12}x_2 \dots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 \dots + a_{2n}x_n = b_2 \\ a_{31}x_1 + a_{32}x_2 \dots + a_{3n}x_n = b_3 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 \dots + a_{mn}x_n = b_m \end{array} \right\} \longrightarrow \mathbf{Ax} = \mathbf{b}, \quad \mathbf{A}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

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

Geometry of linear equations

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

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

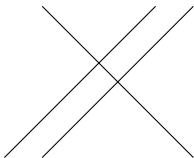


Solving linear equations

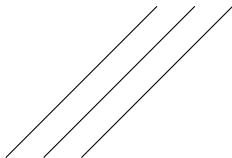
$$\mathbf{Ax} = \mathbf{b}, \quad \mathbf{A} \in \mathbb{R}^{m \times n}, \quad \mathbf{x} \in \mathbb{R}^n, \quad \mathbf{b} \in \mathbb{R}^m$$

- ▶ **Three possible situations:** NO SOLUTION, INFINITELY MANY SOLUTIONS, or UNIQUE SOLUTION.
- ▶ When do we have infinitely many or no solutions? In \mathbb{R}^3 , we can visualize the different situations.

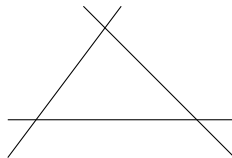
Two parallel planes



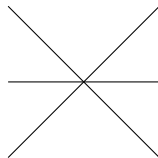
Three parallel planes



No intersection



Line intersection



Solving linear equations: Gaussian Elimination

$$a_{11}x_1 + a_{12}x_2 \dots + a_{1n}x_n = b_1 : E_1$$

$$a_{21}x_1 + a_{22}x_2 \dots + a_{2n}x_n = b_2 : E_2$$

$$a_{31}x_1 + a_{32}x_2 \dots + a_{3n}x_n = b_3 : E_3$$

$$\vdots$$

$$a_{m1}x_1 + a_{m2}x_2 \dots + a_{mn}x_n = b_m : E_m$$

- ▶ 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 αE_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

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

Gaussian Elimination

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

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

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

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.

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

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

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

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

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

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

Things to notice about the echelon form:

- ▶ If a particular 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 \mathbf{A} is defined as the number of basic columns in the row echelon form of the matrix \mathbf{A} .

Solving linear equations: Reduced Row Echelon Form

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

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

- ▶ 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 columns 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

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

Consider the following case,

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

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

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 $[\mathbf{A} \mid \mathbf{0}]$ with $\text{rank}(\mathbf{A}) = r$ and $r < n$ has the general solution of the form,

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

where, the variables $x_{f_1}, x_{f_2}, \dots, x_{f_{n-r}}$ 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? $\rightarrow \text{rank}(\mathbf{A}) = n$.

Solving linear equations: Non-homogenous Systems

$$a_{11}x_1 + a_{12}x_2 \dots + a_{1n}x_n = b_1$$

$$a_{21}x_1 + a_{22}x_2 \dots + a_{2n}x_n = b_2$$

$$a_{31}x_1 + a_{32}x_2 \dots + a_{3n}x_n = b_3 \longrightarrow [\mathbf{A} \mid \mathbf{b}]$$

$$\vdots$$

$$a_{m1}x_1 + a_{m2}x_2 \dots + a_{mn}x_n = b_m$$

Consider the following case,

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

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

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 $\text{rank}(A) = r$,

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

where, \mathbf{p} is the particular solution and $x_{f_1}, x_{f_2}, \dots, x_{f_{n-r}}$ are the free variables.

- ▶ When do we have a unique solution to this system? $\rightarrow \text{rank}(\mathbf{A}) = n$.
- ▶ What about the case when there are no solutions? When does that happen? \rightarrow *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 $[A \mid \mathbf{b}]$ has the same number of basic columns as \mathbf{A} .
- ▶ $[A \mid \mathbf{b}] \rightarrow [E \mid \mathbf{c}]$: \mathbf{c} is a non-basic column.
- ▶ $\text{rank}(\mathbf{A}) = \text{rank}([A \mid \mathbf{b}])$

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 $\mathbf{A} \in \mathbb{R}^{n \times n}$ can be decomposed 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**

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

**Scaling
row 2**

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

**Adding a multiple of
row 2 to row 3**

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

LU Factorization of a Matrix

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

PA = LU Factorization of a Matrix

- ▶ Consider the case: $\mathbf{A} = \begin{bmatrix} 1 & 1 & 0 \\ 2 & 2 & 1 \\ 3 & 4 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \neq \mathbf{LU}$
- ▶ 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,

$$\mathbf{PA} = \mathbf{LU}$$

where, **P** is the permuation matrix, which is the elementary matrix for row exchanges.

- ▶ In the current example, the following allows matrix factorization.

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

Linear transformations

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

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

- ▶ A general version is when the range of the function is not in \mathbb{R} but in \mathbb{R}^m :

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

- ▶ Such a function has a natural representation of the form $\mathbf{y} = \mathbf{A}\mathbf{x}$, $\mathbf{A} \in \mathbb{R}^{m \times n}$.
- ▶ Any linear transformation can be expressed as $\mathbf{y} = \mathbf{A}\mathbf{x}$.
- ▶ Matrices can be thought of as representing a particular linear transformation.

Another look at matrix multiplication

Why does matrix multiplication have this strange definition?

Consider the following two functions,

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

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

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

$$\mathbf{z} = \mathbf{A}(\mathbf{B}\mathbf{u}) = (\mathbf{A}\mathbf{B})\mathbf{u} \implies \mathbf{A}\mathbf{B} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix} = \begin{bmatrix} a\alpha + b\gamma & a\beta + b\delta \\ c\alpha + d\gamma & c\beta + d\delta \end{bmatrix}$$

This definition of matrix multiplication is the most natural for dealing with composition of linear transformations.

Four Fundamental Subspaces

- ▶ $C(\mathbf{A})$: **Column Space of \mathbf{A}** – the span of the columns of \mathbf{A} .

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

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

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

- ▶ $C(\mathbf{A}^T)$: **Row Space of \mathbf{A}** – the span of the rows of \mathbf{A} .

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

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

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

This is also called the **left nullspace** of \mathbf{A} .

Linear Independence

- ▶ Given a set of vectors $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$, $\mathbf{v}_i \in \mathbb{R}^m$, how can we determine if this set is linear independent?
- ▶ We need to verify, $a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \dots + a_n\mathbf{v}_n = \mathbf{0}$

$$\left[\mathbf{v}_1 \quad \mathbf{v}_2 \quad \dots \quad \mathbf{v}_n \right] \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \mathbf{V}\mathbf{a} = \mathbf{0} \left\} N(\mathbf{V}) = \{\mathbf{0}\}, \text{rank}(\mathbf{V}) = n$$

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

Examples

$$\underbrace{\begin{bmatrix} 2 & -1 & 0 & 1 \\ 4 & 3 & 1 & -1 \\ 6 & -2 & 5 & 1 \\ 2 & 3 & 0 & -1 \end{bmatrix}}_{\mathbf{A}} \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}}_{\mathbf{x}} = \underbrace{\begin{bmatrix} 1 \\ -2 \\ 0 \\ 6 \end{bmatrix}}_{\mathbf{b}}$$

- Now solve the above for a different $\mathbf{b} = [1 \ 1 \ 1 \ 1]^T$.

$$\begin{bmatrix} 2 & -1 & 0 \\ 4 & 3 & 1 \\ 6 & -2 & 5 \\ 2 & 3 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -2 \\ 13 \\ -17 \\ 14 \end{bmatrix}$$

- Now solve the above for a different $\mathbf{b} = [1 \ 1 \ 1 \ 1]^T$.

$$\begin{bmatrix} 2 & -1 & -2 \\ 4 & 3 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$$

- Now solve the above for a different $\mathbf{b} = [1 \ 1]^T$.

- Reduce this to the row echelon form:

$$\mathbf{A} = \begin{bmatrix} 1 & -2 & 1 \\ 2 & -4 & 1 \\ -1 & 2 & 2 \end{bmatrix}$$

Dimension and basis of the four fundamental subspaces

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

$$\underbrace{\begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ -5 & 3 & 1 \end{bmatrix}}_{\mathbf{E}} \mathbf{A} = \underbrace{\begin{bmatrix} 1 & -2 & 1 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \end{bmatrix}}_{\mathbf{R}}$$

$$\text{Pivot cols of } \mathbf{A}: \left\{ \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} \right\} \quad N(\mathbf{A}): x_2 \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}$$

$$\text{We can restructure } \mathbf{E}\mathbf{A} = \mathbf{R} \rightarrow \begin{bmatrix} \mathbf{E}_1 \\ \mathbf{E}_2 \end{bmatrix} \mathbf{A} = \begin{bmatrix} \mathbf{R}_1 \\ \mathbf{0} \end{bmatrix}$$

► **Column space** $C(\mathbf{A})$

- $\dim C(\mathbf{A}) = \text{rank}(\mathbf{A}) = r$
- Basis of $C(\mathbf{A})$ = Pivot columns of \mathbf{A} .

► **Nullspace** $N(\mathbf{A})$

- $\dim N(\mathbf{A}) = n - r$
- Basis of $N(\mathbf{A}) = \{\mathbf{h}_1, \mathbf{h}_2 \dots \mathbf{h}_{n-r}\}$.

► **Row space** $C(\mathbf{A}^T)$

- $\dim C(\mathbf{A}^T) = \text{rank}(\mathbf{A}^T) = \text{rank}(\mathbf{A}) = r$
- Basis of $C(\mathbf{A}^T)$ = Columns of \mathbf{R}_1^T .

► **Left Nullspace** $N(\mathbf{A}^T)$

- $\dim N(\mathbf{A}^T) = m - r$
- Basis of $N(\mathbf{A}^T)$ = Columns of \mathbf{E}_2^T

Matrix Inverse

- ▶ Consider the square matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$. $\mathbf{B} \in \mathbb{R}^{n \times n}$ is the inverse of \mathbf{A} , if $\mathbf{AB} = \mathbf{BA} = \mathbf{I}_n$, and \mathbf{B} is represented as \mathbf{A}^{-1} .
- ▶ Not all matrices have inverses. A matrix with an inverse is called **non-singular**, otherwise it is called **singular**.
- ▶ For a non-singular matrix \mathbf{A} , \mathbf{A}^{-1} is unique. \mathbf{A}^{-1} is both the left and right inverse.
- ▶ A matrix \mathbf{A} has an inverse, if and only if \mathbf{A} is full rank, i.e. $\text{rank}(\mathbf{A}) = n$
- ▶ The inverse of a non-singular matrix can be determined through Gauss-Jordan method.

$$[\mathbf{A}|\mathbf{I}] \xrightarrow{\text{Gauss-Jordan}} [\mathbf{I}|\mathbf{A}^{-1}]. \text{ Lets try: } \begin{bmatrix} 1 & -2 & 2 \\ 2 & 1 & 0 \\ 1 & 1 & -1 \end{bmatrix}$$

- ▶ $\mathbf{Ax} = \mathbf{b}$ can be solved as follows, $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$. *It is never solved like this in practice.*
- ▶ Inverse of product of matrices, $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$.
- ▶ $(\mathbf{A}^{-1})^{-1} = \mathbf{A}$ and $(\mathbf{A}^{-1})^T = (\mathbf{A}^T)^{-1}$