Applied Linear Algebra in Data Analysis

Linear Systems and Matrix Operations: Part 3

Sivakumar Balasubramanian

Department of Bioengineering
Christian Medical College, Bagayam
Vellore 632002

Simultaneous Linear Equations

We now move on to a more general case of linear equations – simultaneous linear equations.

We have p equations and q unknown variables.

$$\begin{aligned}
a_{11}x_1 + a_{12}x_2 + \cdots + a_{1q}x_q &= b_1 \\
a_{21}x_1 + a_{22}x_2 + \cdots + a_{2q}x_q &= b_2 \\
&\vdots \\
a_{p1}x_1 + a_{p2}x_2 + \cdots + a_{pq}x_q &= b_p
\end{aligned}$$

where, $a_{ij},\ b_i \in \mathbb{R},\ 1 \le i \le p,\ 1 \le j \le q$ are fixed and x_j are the unknown variables.

The goal is to find values for the unknown variables x_j, $1 \le j \le q$ that satisfy all the equations simultaneously.

Matrix Representation of Simultaneous Linear Equations

We can represent the above set of equations in a more compact matrix form.

$$\underbrace{\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1q} \\ a_{21} & a_{22} & \cdots & a_{2q} \\ \vdots & \vdots & \ddots & \vdots \\ a_{p1} & a_{p2} & \cdots & a_{pq} \end{bmatrix}}_{\mathbf{A} \in \mathbb{R}^{p \times q}} \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_q \end{bmatrix}}_{\mathbf{x} \in \mathbb{R}^{q}} = \underbrace{\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_p \end{bmatrix}}_{\mathbf{b} \in \mathbb{R}^{p}} \longrightarrow \mathbf{Ax} = \mathbf{b}$$

$\mathbf{A}$ is the coefficient matrix, $\mathbf{x}$ is the vector of unknowns, & $\mathbf{b}$ is the constant vector.

The number of rows of $\mathbf{A}$ equals the number of equations p, and the number of columns equals the number of unknown variables q.

$p \times q$ will be referred to as the size or shape of the matrix.

Matrices

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

$$\mathbf{A} = \begin{array}{ccc} & q \text{ columns} & \\ p \text{ rows} & \begin{bmatrix} \square & \square & \square & \cdots & \square \\ \square & \square & \square & \cdots & \square \\ \square & a_{32} & \square & \cdots & \square \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \square & \square & \square & \cdots & \square \end{bmatrix} & 3^{rd} \text{ row} \\ & 2^{nd} \text{ column} & \end{array}$$

Matrices

Consider a matrix $\mathbf{A}$ with n rows and m columns.

$$\mathbf{A} \longrightarrow \begin{cases} \textbf{Tall/Skinny} & p > q \\ \textbf{Square} & p = q \\ \textbf{Wide/Fat} & p < q \end{cases}$$

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.
$\mathbf{x}^\top = \begin{bmatrix} 1.45 & -3.1 & 12.4 \end{bmatrix}$

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

Matrices are also compact way to give a set of q indexed columns (p-vectors: $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3 \dots \mathbf{a}_q$), or a set of p indexed rows (q-vectors: $\tilde{\mathbf{a}}_1^\top, \tilde{\mathbf{a}}_2^\top, \dots \tilde{\mathbf{a}}_p^\top$)

$$\mathbf{A} = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3 & \cdots & \mathbf{a}_q \end{bmatrix} = \begin{bmatrix} \tilde{\mathbf{a}}_1^\top \\ \tilde{\mathbf{a}}_2^\top \\ \tilde{\mathbf{a}}_3^\top \\ \vdots \\ \tilde{\mathbf{a}}_p^\top \end{bmatrix}$$

Note that in this representation, the matrix looks like a block row or a block column.

The $p \times q$ matrix $\mathbf{A}$ can be thought of as a block row of q p-vectors, or a block column of p q-vectors.

Some Special Matrices

Zero matrix$= \mathbf{0}_{p\times q} = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}$

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

$$i_{kl} = \begin{cases} 1 & k = l \\ 0 & k \neq l \end{cases} \qquad \mathbf{I}_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \end{bmatrix} = \begin{bmatrix} \mathbf{e}_1^\top \\ \mathbf{e}_2^\top \\ \mathbf{e}_3^\top \end{bmatrix}$$

Some Special Matrices

Diagonal matrices is a square matrix with non-zero elements on its diagonal. $a_{ij} = 0, \forall i \neq j$

$$\mathbf{A} = \begin{bmatrix} 0.4 & 0 & 0 & 0 \\ 0 & -11 & 0 & 0 \\ 0 & 0 & 21 & 0 \\ 0 & 0 & 0 & 9.3 \end{bmatrix}$$

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

Useful Matrix Operations: Transpose

Transpose switches the rows and columns of a matrix. $\mathbf{A}$ is a $n \times m$ matrix, then its transpose is represented by $\mathbf{A}^\top$, which is a $m \times n$ matrix.

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

Transpose converts between column and row vectors.

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

Useful Matrix Operations: Transpose

Transpose switches the rows and columns of a matrix. $\mathbf{A}$ is a $n \times m$ matrix, then its transpose is represented by $\mathbf{A}^\top$, which is a $m \times n$ matrix.

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

Transpose converts between column and row vectors.

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

$$\mathbf{A}^\top = \begin{bmatrix} \mathbf{B}^\top & \mathbf{D}^\top \\ \mathbf{C}^\top & \mathbf{E}^\top \end{bmatrix}$$

Useful Matrix Operations: Matrix Addition

Matrix addition: Element-wise addition. Only matrices of the same size can be added together. Let $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{p \times q}$, then,

$$(\mathbf{A} + \mathbf{B})_{ij} = a_{ij} + b_{ij}, \;\; 1 \leq i \leq p, 1 \leq j \leq q$$

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*: $(\mathbf{A} + \mathbf{B})^\top = \mathbf{A}^\top + \mathbf{B}^\top$

Useful Matrix Operations: Scalar multiplication

Scalar multiplication Each element of the matrix gets multiplied by the scalar. Let $\mathbf{A} \in \mathbb{R}^{p \times q}$ and $\alpha \in \mathbb{R}$, then,

$$(\alpha \mathbf{A})_{ij} = \alpha a_{ij}$$

We will mostly only deal with matrices with real entries. Such matrices are elements of the set $\mathbb{R}^{p \times q}$.

Given the aforementioned matrix operations and their properties, is $\mathbb{R}^{p \times q}$ a vector space?

Useful Matrix Operations: Scalar multiplication

Scalar multiplication Each element of the matrix gets multiplied by the scalar. Let $\mathbf{A} \in \mathbb{R}^{p \times q}$ and $\alpha \in \mathbb{R}$, then,

$$(\alpha \mathbf{A})_{ij} = \alpha a_{ij}$$

We will mostly only deal with matrices with real entries. Such matrices are elements of the set $\mathbb{R}^{p \times q}$.

Given the aforementioned matrix operations and their properties, is $\mathbb{R}^{p \times q}$ a vector space?
Yes, because it satisfies all the properties of a vector space. Closed under addition and scalar multiplication..

Useful Matrix Operations: Matrix Multiplication

Matrix Multiplication: Let $\mathbf{A} \in \mathbb{R}^{p \times q}$ and $\mathbf{B} \in \mathbb{R}^{q \times r}$, then the product matrix $\mathbf{C} = \mathbf{AB} \in \mathbb{R}^{p \times r}$.

Multiplication operation between two matrices $\mathbf{A}$ and $\mathbf{B}$, $\mathbf{AB}$ can be done if and only if the number of columns of $\mathbf{A}$ equals the number of rows of $\mathbf{B}$.

The resulting matrix will have the number of rows of $\mathbf{A}$ and number of columns of $\mathbf{B}$.

The ij^{th} element of the resulting matrix $\mathbf{C} = \mathbf{AB}$ is given by,

$$c_{ij} := \sum_{k=1}^{q} a_{ik} b_{kj}, \ 1 \leq i \leq p, \ 1 \leq j \leq r$$

Can we compute $\mathbf{BA}$?

Useful Matrix Operations: Matrix Multiplication

We will start with some simple matrix multiplication operations: (a) inner product and (b) outer product.

Standard Inner Product: The standard inner product is a special case of matrix multiplication. Its a multiplication between a row vector and a column vector.

Consider two vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. We define the **standard inner product** as the following,

$$\mathbf{x}^\top \mathbf{y} = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \sum_{i=1}^{n} x_i y_i \in \mathbb{R}$$

Note, $\mathbf{x}^\top \in \mathbb{R}^{1\times n}$, $\mathbf{y} \in \mathbb{R}^{n\times 1}$, so the resulting product is a scalar in $\mathbf{x}^\top \mathbf{y} \in \mathbb{R}\,(= \mathbb{R}^{1\times 1})$.

Useful Matrix Operations: Matrix Multiplication

Outer Product: The outer product is another special case of matrix multiplication. Its a multiplication between a row vector and a column vector.

Consider two vectors $\mathbf{x} \in \mathbb{R}^p$ and $\mathbf{y} \in \mathbb{R}^q$. We define the **outer product** as the following,

$$\mathbf{x}\mathbf{y}^\top = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{bmatrix} \begin{bmatrix} y_1 & y_2 & \cdots & y_q \end{bmatrix} = \begin{bmatrix} x_1 y_1 & x_1 y_2 & \cdots & x_1 y_q \\ x_2 y_1 & x_2 y_2 & \cdots & x_2 y_q \\ \vdots & \vdots & \ddots & \vdots \\ x_p y_1 & x_p y_2 & \cdots & x_p y_q \end{bmatrix} \in \mathbb{R}^{p \times q}$$

Note, $\mathbf{x} \in \mathbb{R}^{p\times 1}$, $\mathbf{y}^\top \in \mathbb{R}^{1\times q}$, so the resulting product is a scalar in $\mathbf{x}\mathbf{y}^\top \in \mathbb{R}^{p\times q}$.

Useful Matrix Operations: Matrix Multiplication

We now move to matrix multiplication with column and row vectors.

Let's first find out what is possible and what the result will be. Consider the matrix $\mathbf{A} \in \mathbb{R}^{p \times q}$ and two vectors $\mathbf{x} \in \mathbb{R}^p$ and $\mathbf{y} \in \mathbb{R}^q$.

Which of the following is allowed?

1. $\mathbf{Ax}$?

Useful Matrix Operations: Matrix Multiplication

We now move to matrix multiplication with column and row vectors.

Let's first find out what is possible and what the result will be. Consider the matrix $\mathbf{A} \in \mathbb{R}^{p \times q}$ and two vectors $\mathbf{x} \in \mathbb{R}^p$ and $\mathbf{y} \in \mathbb{R}^q$.

Which of the following is allowed?

1. $\mathbf{Ax}$? **No**. Number of columns of $\mathbf{A} \neq$ number of rows of $\mathbf{x}$.
2. $\mathbf{Ay}$?

Useful Matrix Operations: Matrix Multiplication

We now move to matrix multiplication with column and row vectors.

Let's first find out what is possible and what the result will be. Consider the matrix $\mathbf{A} \in \mathbb{R}^{p \times q}$ and two vectors $\mathbf{x} \in \mathbb{R}^p$ and $\mathbf{y} \in \mathbb{R}^q$.

Which of the following is allowed?

1. $\mathbf{Ax}$? **No**. Number of columns of $\mathbf{A} \neq$ number of rows of $\mathbf{x}$.
2. $\mathbf{Ay}$? **Yes**. Result is in $\mathbb{R}^{p \times 1}$.
3. $\mathbf{xA}$?

Useful Matrix Operations: Matrix Multiplication

We now move to matrix multiplication with column and row vectors.

Let's first find out what is possible and what the result will be. Consider the matrix $\mathbf{A} \in \mathbb{R}^{p \times q}$ and two vectors $\mathbf{x} \in \mathbb{R}^p$ and $\mathbf{y} \in \mathbb{R}^q$.

Which of the following is allowed?

1. $\mathbf{Ax}$? **No**. Number of columns of $\mathbf{A} \neq$ number of rows of $\mathbf{x}$.
2. $\mathbf{Ay}$? **Yes**. Result is in $\mathbb{R}^{p \times 1}$.
3. $\mathbf{xA}$? **No**.
4. $\mathbf{yA}$?

Useful Matrix Operations: Matrix Multiplication

We now move to matrix multiplication with column and row vectors.

Let's first find out what is possible and what the result will be. Consider the matrix $\mathbf{A} \in \mathbb{R}^{p \times q}$ and two vectors $\mathbf{x} \in \mathbb{R}^p$ and $\mathbf{y} \in \mathbb{R}^q$.

Which of the following is allowed?

1. $\mathbf{Ax}$? **No**. Number of columns of $\mathbf{A} \neq$ number of rows of $\mathbf{x}$.
2. $\mathbf{Ay}$? **Yes**. Result is in $\mathbb{R}^{p \times 1}$.
3. $\mathbf{xA}$? **No**.
4. $\mathbf{yA}$? **No**.
5. $\mathbf{x}^\top \mathbf{A}$?

Useful Matrix Operations: Matrix Multiplication

We now move to matrix multiplication with column and row vectors.

Let's first find out what is possible and what the result will be. Consider the matrix $\mathbf{A} \in \mathbb{R}^{p \times q}$ and two vectors $\mathbf{x} \in \mathbb{R}^p$ and $\mathbf{y} \in \mathbb{R}^q$.

Which of the following is allowed?

1. $\mathbf{Ax}$? **No**. Number of columns of $\mathbf{A} \neq$ number of rows of $\mathbf{x}$.
2. $\mathbf{Ay}$? **Yes**. Result is in $\mathbb{R}^{p \times 1}$.
3. $\mathbf{xA}$? **No**.
4. $\mathbf{yA}$? **No**.
5. $\mathbf{x}^\top \mathbf{A}$? **Yes**. Result is in $\mathbb{R}^{1 \times q}$.
6. $\mathbf{y}^\top \mathbf{A}$?

Useful Matrix Operations: Matrix Multiplication

We now move to matrix multiplication with column and row vectors.

Let's first find out what is possible and what the result will be. Consider the matrix $\mathbf{A} \in \mathbb{R}^{p \times q}$ and two vectors $\mathbf{x} \in \mathbb{R}^p$ and $\mathbf{y} \in \mathbb{R}^q$.

Which of the following is allowed?

1. $\mathbf{Ax}$? **No**. Number of columns of $\mathbf{A} \neq$ number of rows of $\mathbf{x}$.
2. $\mathbf{Ay}$? **Yes**. Result is in $\mathbb{R}^{p \times 1}$.
3. $\mathbf{xA}$? **No**.
4. $\mathbf{yA}$? **No**.
5. $\mathbf{x}^\top \mathbf{A}$? **Yes**. Result is in $\mathbb{R}^{1 \times q}$.
6. $\mathbf{y}^\top \mathbf{A}$? **No**

Matrix-Column Vector Multiplication: A matrix $\mathbf{A} \in \mathbb{R}^{p\times q}$ can be post-multiplied by a column vector $\mathbf{x} \in \mathbb{R}^{q}$, resulting in another column vector $\mathbf{y} \in \mathbb{R}^{p}$.

$$\mathbf{y} = \mathbf{A}\mathbf{x} \longrightarrow \begin{cases} \mathbf{y} = \sum_{j=1}^{q} x_j \cdot \mathbf{a}_j, \\ y_i = \sum_{j=1}^{q} a_{ij}x_j, \;\; 1 \leq i \leq p \end{cases}$$

$\mathbf{Ax}$ can be viewed as the inner product of a block row and a column vector.

$$\mathbf{Ax} = \overbrace{\begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_q \end{bmatrix}}^{1\times q} \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_q \end{bmatrix}}_{q\times 1} = x_1 \cdot \mathbf{a}_1 + x_2 \cdot \mathbf{a}_2 + \cdots + x_q \cdot \mathbf{a}_q = \sum_{j=1}^{q} x_j \cdot \mathbf{a}_j$$

Useful Matrix Operations: Matrix Multiplication

Matrix-Column Vector Multiplication: A matrix $\mathbf{A} \in \mathbb{R}^{p\times q}$ can be post-multiplied by a column vector $\mathbf{x} \in \mathbb{R}^q$, resulting in another column vector $\mathbf{y} \in \mathbb{R}^p$.

$\mathbf{Ax}$ can be viewed as the outer product of the block column and a block row with one element.

$$\mathbf{Ax} = \overbrace{\begin{bmatrix} \tilde{\mathbf{a}}_1^\top \\ \tilde{\mathbf{a}}_2^\top \\ \cdots \\ \tilde{\mathbf{a}}_p^\top \end{bmatrix}}^{p\times 1} \underbrace{\begin{bmatrix} \mathbf{x} \end{bmatrix}}_{1\times 1} = \begin{bmatrix} \tilde{\mathbf{a}}_1^\top \mathbf{x} \\ \tilde{\mathbf{a}}_2^\top \mathbf{x} \\ \vdots \\ \tilde{\mathbf{a}}_p^\top \mathbf{x} \end{bmatrix} \implies y_i = \tilde{\mathbf{a}}_i^\top \mathbf{x} = \sum_{j=1}^{q} a_{ij} x_j$$

Useful Matrix Operations: Matrix Multiplication

Matrix-Column Vector Multiplication: A matrix $\mathbf{A} \in \mathbb{R}^{p \times q}$ can be post-multiplied by a column vector $\mathbf{x} \in \mathbb{R}^{q}$, resulting in another column vector $\mathbf{y} \in \mathbb{R}^{p}$.

$\mathbf{Ax}$ can be viewed as the outer product of the block column and a block row with one element.

$$\mathbf{Ax} = \overbrace{\begin{bmatrix} \tilde{\mathbf{a}}_1^\top \\ \tilde{\mathbf{a}}_2^\top \\ \cdots \\ \tilde{\mathbf{a}}_p^\top \end{bmatrix}}^{p \times 1} \underbrace{\begin{bmatrix} \mathbf{x} \end{bmatrix}}_{1 \times 1} = \begin{bmatrix} \tilde{\mathbf{a}}_1^\top \mathbf{x} \\ \tilde{\mathbf{a}}_2^\top \mathbf{x} \\ \vdots \\ \tilde{\mathbf{a}}_p^\top \mathbf{x} \end{bmatrix} \implies y_i = \tilde{\mathbf{a}}_i^\top \mathbf{x} = \sum_{j=1}^{q} a_{ij} x_j$$

Let's do an example: $\mathbf{A} = \begin{bmatrix} 2 & 3 & -1 & 2 \\ 1 & 0 & 2 & -1 \\ 5 & -2 & 1 & -1 \end{bmatrix}$ and $\mathbf{x} = \begin{bmatrix} 1 \\ -1 \\ 2 \\ 3 \end{bmatrix}$.

Useful Matrix Operations: Matrix Multiplication

Matrix-Column Vector Multiplication: A matrix $\mathbf{A} \in \mathbb{R}^{p \times q}$ can be post-multiplied by a column vector $\mathbf{x} \in \mathbb{R}^q$, resulting in another column vector $\mathbf{y} \in \mathbb{R}^p$.

$\mathbf{Ax}$ can be viewed as the outer product of the block column and a block row with one element.

$$\mathbf{Ax} = \overbrace{\begin{bmatrix} \tilde{\mathbf{a}}_1^\top \\ \tilde{\mathbf{a}}_2^\top \\ \cdots \\ \tilde{\mathbf{a}}_p^\top \end{bmatrix}}^{p \times 1} \underbrace{\begin{bmatrix} \mathbf{x} \end{bmatrix}}_{1 \times 1} = \begin{bmatrix} \tilde{\mathbf{a}}_1^\top \mathbf{x} \\ \tilde{\mathbf{a}}_2^\top \mathbf{x} \\ \vdots \\ \tilde{\mathbf{a}}_p^\top \mathbf{x} \end{bmatrix} \implies y_i = \tilde{\mathbf{a}}_i^\top \mathbf{x} = \sum_{j=1}^{q} a_{ij} x_j$$

Let's do an example: $\mathbf{A} = \begin{bmatrix} 2 & 3 & -1 & 2 \\ 1 & 0 & 2 & -1 \\ 5 & -2 & 1 & -1 \end{bmatrix}$ and $\mathbf{x} = \begin{bmatrix} 1 \\ -1 \\ 2 \\ 3 \end{bmatrix}$. $\mathbf{y} = \mathbf{Ax} = \begin{bmatrix} 3 \\ 2 \\ 8 \end{bmatrix}$

Row Vector-Matrix Multiplication: A matrix $\mathbf{A} \in \mathbb{R}^{p \times q}$ can be pre-multiplied by a row vector $\mathbf{x}^\top \in \mathbb{R}^{1 \times p}$, resulting in another row vector $\mathbf{y}^\top \in \mathbb{R}^{1 \times q}$.

$$\mathbf{y}^\top = \mathbf{x}^\top \mathbf{A} \longrightarrow \begin{cases} \mathbf{y}^\top = \sum_{j=1}^{p} x_j \cdot \tilde{\mathbf{a}}_j^\top, \\ y_i = \sum_{j=1}^{p} a_{ji} x_j, \;\; 1 \leq i \leq q \end{cases}$$

$\mathbf{x}^\top \mathbf{A}$ can be viewed as the inner product of a row and and a block column.

$$\mathbf{A}\mathbf{x} = \overbrace{\begin{bmatrix} x_1 & x_2 & \cdots & x_p \end{bmatrix}}^{1 \times p} \underbrace{\begin{bmatrix} \tilde{\mathbf{a}}_1^\top \\ \tilde{\mathbf{a}}_2^\top \\ \vdots \\ \tilde{\mathbf{a}}_p^\top \end{bmatrix}}_{p \times 1} = x_1 \cdot \tilde{\mathbf{a}}_1^\top + x_2 \cdot \tilde{\mathbf{a}}_2^\top + \cdots + x_p \cdot \tilde{\mathbf{a}}_p^\top = \sum_{i=1}^{p} x_j \cdot \tilde{\mathbf{a}}_j^\top$$

Useful Matrix Operations: Matrix Multiplication

Row Vector-Matrix Multiplication: A matrix $\mathbf{A} \in \mathbb{R}^{p\times q}$ can be pre-multiplied by a row vector $\mathbf{x}^\top \in \mathbb{R}^{1\times p}$, resulting in another row vector $\mathbf{y}^\top \in \mathbb{R}^{1\times q}$. $\mathbf{x}^\top\mathbf{A}$ can be viewed as the outer product of the block column and a block row with one element.

$$\mathbf{x}^\top\mathbf{A} = \overbrace{\begin{bmatrix}\mathbf{x}^\top\end{bmatrix}}^{1\times 1}\underbrace{\begin{bmatrix}\mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_q\end{bmatrix}}_{1\times q} = \begin{bmatrix}\mathbf{x}^\top\mathbf{a}_1 & \mathbf{x}_2^\top\mathbf{a}_2 & \cdots & \mathbf{x}^\top\mathbf{a}_q\end{bmatrix} \implies y_i = \mathbf{x}^\top\mathbf{a}_i = \sum_{j=1}^{q} a_{ij}x_j$$

Useful Matrix Operations: Matrix Multiplication

Row Vector-Matrix Multiplication: A matrix $\mathbf{A} \in \mathbb{R}^{p \times q}$ can be pre-multiplied by a row vector $\mathbf{x}^\top \in \mathbb{R}^{1 \times p}$, resulting in another row vector $\mathbf{y}^\top \in \mathbb{R}^{1 \times q}$. $\mathbf{x}^\top \mathbf{A}$ can be viewed as the outer product of the block column and a block row with one element.

$$\mathbf{x}^\top \mathbf{A} = \overbrace{\begin{bmatrix} \mathbf{x}^\top \end{bmatrix}}^{1 \times 1} \underbrace{\begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_q \end{bmatrix}}_{1 \times q} = \begin{bmatrix} \mathbf{x}^\top \mathbf{a}_1 & \mathbf{x}_2^\top \mathbf{a}_2 & \cdots & \mathbf{x}^\top \mathbf{a}_q \end{bmatrix} \implies y_i = \mathbf{x}^\top \mathbf{a}_i = \sum_{j=1}^{q} a_{ij} x_j$$

Let's do an example: $\mathbf{A} = \begin{bmatrix} 2 & 3 & -1 & 2 \\ 1 & 0 & 2 & -1 \\ 5 & -2 & 1 & -1 \end{bmatrix}$ and $\mathbf{x}^\top = \begin{bmatrix} -1 & 2 & 1 \end{bmatrix}$

Useful Matrix Operations: Matrix Multiplication

Row Vector-Matrix Multiplication: A matrix $\mathbf{A} \in \mathbb{R}^{p\times q}$ can be pre-multiplied by a row vector $\mathbf{x}^\top \in \mathbb{R}^{1\times p}$, resulting in another row vector $\mathbf{y}^\top \in \mathbb{R}^{1\times q}$. $\mathbf{x}^\top\mathbf{A}$ can be viewed as the outer product of the block column and a block row with one element.

$$\mathbf{x}^\top\mathbf{A} = \overbrace{\begin{bmatrix}\mathbf{x}^\top\end{bmatrix}}^{1\times 1} \underbrace{\begin{bmatrix}\mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_q\end{bmatrix}}_{1\times q} = \begin{bmatrix}\mathbf{x}^\top\mathbf{a}_1 & \mathbf{x}_2^\top\mathbf{a}_2 & \cdots & \mathbf{x}^\top\mathbf{a}_q\end{bmatrix} \implies y_i = \mathbf{x}^\top\mathbf{a}_i = \sum_{j=1}^{q} a_{ij}x_j$$

Let's do an example: $\mathbf{A} = \begin{bmatrix}2 & 3 & -1 & 2\\ 1 & 0 & 2 & -1\\ 5 & -2 & 1 & -1\end{bmatrix}$ and $\mathbf{x}^\top = \begin{bmatrix}-1 & 2 & 1\end{bmatrix}$

$$\mathbf{y}^\top = \mathbf{x}^\top\mathbf{A} = \begin{bmatrix}5 & -5 & 6 & -5\end{bmatrix}$$

Useful Matrix Operations: Matrix Multiplication

Now, we are ready to look at matrix-matrix multiplication. Consider two matrices $\mathbf{A} \in \mathbb{R}^{p \times q}$ and $\mathbf{B} \in \mathbb{R}^{q \times r}$, then the product matrix $\mathbf{C} = \mathbf{AB} \in \mathbb{R}^{p \times r}$.

There are four ways to view the multiplication operation between two matrices $\mathbf{A}$ and $\mathbf{B}$,

- **Inner Product View**: Computes the individual elements of $\mathbf{AB}$.
- **Outer Product View**: Expresses $\mathbf{AB}$ as a sum of outer products.
- **Column View**: Columns of $\mathbf{AB}$ as linear combinations of columns of $\mathbf{A}$.
- **Row View**: Rows of $\mathbf{AB}$ as linear combinations of rows of $\mathbf{B}$.

Useful Matrix Operations: Matrix Multiplication

Inner Product View: Computes the individual elements of $\mathbf{AB}$. We get this view by treating $\mathbf{A}$ as a block column and $\mathbf{B}$ as a block row.

$$\mathbf{C} = \overbrace{\begin{bmatrix} \tilde{\mathbf{a}}_1^\top \\ \tilde{\mathbf{a}}_2^\top \\ \vdots \\ \tilde{\mathbf{a}}_p^\top \end{bmatrix}}^{p\times 1} \underbrace{\begin{bmatrix} \mathbf{b}_1 & \mathbf{b}_2 & \cdots & \mathbf{b}_r \end{bmatrix}}_{1\times q} = \begin{bmatrix} \tilde{\mathbf{a}}_1^\top \mathbf{b}_1 & \tilde{\mathbf{a}}_1^\top \mathbf{b}_2 & \cdots & \tilde{\mathbf{a}}_1^\top \mathbf{b}_r \\ \tilde{\mathbf{a}}_2^\top \mathbf{b}_1 & \tilde{\mathbf{a}}_2^\top \mathbf{b}_2 & \cdots & \tilde{\mathbf{a}}_2^\top \mathbf{b}_r \\ \vdots & \vdots & \ddots & \vdots \\ \tilde{\mathbf{a}}_p^\top \mathbf{b}_1 & \tilde{\mathbf{a}}_p^\top \mathbf{b}_2 & \cdots & \tilde{\mathbf{a}}_p^\top \mathbf{b}_r \end{bmatrix} \implies c_{ij} = \tilde{\mathbf{a}}_i^\top \mathbf{b}_j$$

Let's do an example: $\mathbf{A} = \begin{bmatrix} 2 & 3 & -1 & 2 \\ 1 & 0 & 2 & -1 \\ 5 & -2 & 1 & -1 \end{bmatrix}$ and $\mathbf{B} = \begin{bmatrix} 1 & 2 \\ -1 & 1 \\ 2 & 1 \\ 0 & 1 \end{bmatrix}$.

Useful Matrix Operations: Matrix Multiplication

Outer Product View: Computes $\mathbf{AB}$ as a sum of outer product matrices. We get this view by treating $\mathbf{A}$ as a block row and $\mathbf{B}$ as a block column.

$$\mathbf{C} = \overbrace{\begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_q \end{bmatrix}}^{1\times q} \underbrace{\begin{bmatrix} \tilde{\mathbf{b}}_1^\top \\ \tilde{\mathbf{b}}_2^\top \\ \vdots \\ \tilde{\mathbf{b}}_q^\top \end{bmatrix}}_{q\times 1} = \mathbf{a}_1\tilde{\mathbf{b}}_1^\top + \mathbf{a}_2\tilde{\mathbf{b}}_2^\top + \cdots + \mathbf{a}_q\tilde{\mathbf{b}}_q^\top = \sum_{i=1}^{q} \mathbf{a}_i\tilde{\mathbf{b}}_i^\top$$

Let's do an example: $\mathbf{A} = \begin{bmatrix} 2 & 3 & -1 & 2 \\ 1 & 0 & 2 & -1 \\ 5 & -2 & 1 & -1 \end{bmatrix}$ and $\mathbf{B} = \begin{bmatrix} 1 & 2 \\ -1 & 1 \\ 2 & 1 \\ 0 & 1 \end{bmatrix}$.

Useful Matrix Operations: Matrix Multiplication

Column View: The columns of $\mathbf{AB}$ are linear combinations of the columns of $\mathbf{A}$, with the mixture for each column coming from the columns of $\mathbf{B}$. We get this view by treating $\mathbf{B}$ as a block row and $\mathbf{A}$ as a matrix.

$$\mathbf{C} = \mathbf{A}\begin{bmatrix}\mathbf{b}_1 & \mathbf{b}_2 & \cdots & \mathbf{b}_r\end{bmatrix} = \begin{bmatrix}\mathbf{Ab}_1 & \mathbf{Ab}_2 & \cdots & \mathbf{Ab}_r\end{bmatrix} \implies \mathbf{c}_i = \mathbf{Ab}_i$$

Let's do an example: $\mathbf{A} = \begin{bmatrix} 2 & 3 & -1 & 2 \\ 1 & 0 & 2 & -1 \\ 5 & -2 & 1 & -1 \end{bmatrix}$ and $\mathbf{B} = \begin{bmatrix} 1 & 2 \\ -1 & 1 \\ 2 & 1 \\ 0 & 1 \end{bmatrix}$.

Useful Matrix Operations: Matrix Multiplication

Row View: The rows of $\mathbf{AB}$ are linear combinations of the rows of $\mathbf{B}$, with the mixture for each row coming from the rows of $\mathbf{A}$. We get this view by treating $\mathbf{A}$ as a block column and $\mathbf{B}$ as a matrix.

$$\mathbf{C} = \begin{bmatrix} \tilde{\mathbf{a}}_1^\top \\ \tilde{\mathbf{a}}_2^\top \\ \vdots \\ \tilde{\mathbf{a}}_p^\top \end{bmatrix} \mathbf{B} = \begin{bmatrix} \tilde{\mathbf{a}}_1^\top \mathbf{B} \\ \tilde{\mathbf{a}}_2^\top \mathbf{B} \\ \vdots \\ \tilde{\mathbf{a}}_p^\top \mathbf{B} \end{bmatrix} \implies \tilde{\mathbf{c}}_i^\top = \tilde{\mathbf{a}}_i^\top \mathbf{B}$$

Let's do an example: $\mathbf{A} = \begin{bmatrix} 2 & 3 & -1 & 2 \\ 1 & 0 & 2 & -1 \\ 5 & -2 & 1 & -1 \end{bmatrix}$ and $\mathbf{B} = \begin{bmatrix} 1 & 2 \\ -1 & 1 \\ 2 & 1 \\ 0 & 1 \end{bmatrix}$.

Properties of Matrix Multiplication

Not commutative: $\mathbf{AB} \neq \mathbf{BA}$
The product of two matrices might not alwasys be defined. When it is defined, $\mathbf{AB}$ and $\mathbf{BA}$ need not match.

Distributive: $\mathbf{A}\left(\mathbf{B}+\mathbf{C}\right) = \mathbf{AB} + \mathbf{BC}$ and $\left(\mathbf{A}+\mathbf{B}\right)\mathbf{C} = \mathbf{AC} + \mathbf{BC}$

Associative: $\mathbf{A}\left(\mathbf{BC}\right) = \left(\mathbf{AB}\right)\mathbf{C}$

Transpose: $\left(\mathbf{AB}\right)^\top = \mathbf{B}^\top\mathbf{A}^\top$

Scalar product: $\alpha\left(\mathbf{AB}\right) = \left(\alpha\mathbf{A}\right)\mathbf{B} = \mathbf{A}\left(\alpha\mathbf{B}\right)$

Simultaneous Linear Equations

With that detour into matrices and some useful matrix operations, we come back to a systems of linear equations.

$$\left.\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1q}x_q &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2q}x_q &= b_2 \\ &\vdots \\ a_{p1}x_1 + a_{p2}x_2 + \cdots + a_{pq}x_q &= b_p \end{aligned}\right. \longrightarrow \mathbf{Ax} = \mathbf{b}$$

where, $a_{ij},\ b_i \in \mathbb{R},\ 1 \leq i \leq p,\ 1 \leq j \leq q$ are fixed and x_j are the unknown variables. $\mathbf{A} \in \mathbb{R}^{p \times q}$ is the coefficient matrix, $\mathbf{x} \in \mathbb{R}^q$ is the vector of unknowns, and $\mathbf{b} \in \mathbb{R}^p$ is the right hand side vector.

From the matrix multiplication rules, the left hand side is the linear column combinations of the columns of $\mathbf{A}$, which must equal the vector $\mathbf{b}$ on the right hand side, for some choice of $\mathbf{x}$.

We are interested in finding such an $\mathbf{x}$, if it exists.

Simultaneous Linear Equations

We will now relate the nature of the solution to $\mathbf{A}\mathbf{x} = \mathbf{b}$ to the properties of the matrix $\mathbf{A}$ and $\mathbf{b}$.

$$x_1 \cdot \mathbf{a}_1 + x_2 \cdot \mathbf{a}_2 + \cdots + x_q \cdot \mathbf{a}_q = \mathbf{b}$$

We would like to know:

- $\mathbf{b} \in \operatorname{span}\left(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_q\right)$?
- Are the columns of $\mathbf{A}$ linearly independent?

Guassian eliminiation or the Gauss-Jordan method perform row operations on the "augmented matrix" $[\mathbf{A} \mid \mathbf{b}]$ and reveal answers to the above questions and the solution $\mathbf{x}$, if it exists.

$$\text{Augmented Matrix} = \left[\mathbf{A} \mid \mathbf{b}\right] = \left[\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1q} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2q} & b_2 \\ a_{31} & a_{32} & \cdots & a_{3q} & b_3 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{p1} & a_{p2} & \cdots & a_{pq} & b_p \end{array}\right]$$

Simultaneous Linear Equations

Before we look at the row operations, let's looks solutions to some simple systems of linear equations. System where $\mathbf{A}$ has some special properties.

$$\mathbf{A}\mathbf{x} = \mathbf{b}$$

Write down the system of linear equations for each case and solve for $\mathbf{x}$.

1. $\mathbf{A} = \mathbf{I}_p$.
2. $\mathbf{A} \in \mathbb{R}^{p \times p}$ is a diagonal matrix with non-zero diagonal elements.
3. $\mathbf{A} \in \mathbb{R}^{p \times p}$ is upper triangular with non-zero diagonal elements.
4. $\mathbf{A} \in \mathbb{R}^{p \times p}$ is lower triangular with non-zero diagonal elements.

Guassian elimination attempts to convert a square $\mathbf{A}$ into an upper triangular matrix using row operations.

While Gauss-Jordan elimination attempts to convert an arbitrary $\mathbf{A}$ into its *simplest possible form* (identity matrix if possible) using row operations.

Row Operations on the Augmented Matrix

Three row operations are allowed on the augmented matrix $[\mathbf{A} \mid \mathbf{b}]$ without changing the solution to the system of equations $\mathbf{Ax} = \mathbf{b}$. Let E_i and E_j represent the i^{th} and j^{th} equations in the system respectively.

- Interchanging of equations E_i and E_j.
- Replacing equation E_i by αE_i, $\alpha \neq 0$.
- Replacing equation E_j by $E_j + \alpha E_i$, $\alpha \neq 0$.

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

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

The blue-colored elements along the main diagonal are the *pivots* of the matrix. Note that all pivots are non-zero.

Back substitution reveals the solution: $x_3 = 1;\ x_2 = 4;\ x_1 = -6$.

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

Solving linear equations: Gauss-Jordan Method

Continue the elimination upwards until all elements above and below the pivots are zero, and normalize the pivots.

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

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

Solving linear equations: Gauss-Jordan Method

Continue the elimination upwards until all elements above and below the pivots are zero, and normalize the pivots.

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

$$\left[\begin{array}{rrr|r} \underline{1} & 2 & 0 & 2 \\ 0 & \underline{1} & 0 & 4 \\ 0 & 0 & \underline{1} & 1 \end{array}\right] \longrightarrow \left[\begin{array}{rrr|r} \underline{1} & 0 & 0 & -6 \\ 0 & \underline{1} & 0 & 4 \\ 0 & 0 & \underline{1} & 1 \end{array}\right] \Longrightarrow x_1 = -6; \quad x_2 = 4; \quad x_3 = 1;$$

Things can go wrong! Lets solve these: $\left[\begin{array}{rrr|r} 1 & 2 & -1 & 1 \\ 2 & 3 & 4 & 4 \\ -2 & -4 & 2 & -3 \end{array}\right], \left[\begin{array}{rrr|r} 1 & 2 & -1 & 1 \\ 2 & 3 & 4 & 4 \\ -2 & -4 & 2 & -2 \end{array}\right]$

What is the difference between these two systems?

Solving linear equations: Rectangular systems and Row Echelon Form

For rectangular systems of equations, the Gauss-Jordan elimination results in a matrix in **row echelon form**.

If the diagnonal element is zero and if row exchanges are not possible, we skip that column and move to the next column to find a non-zero pivot element.

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

$$\left[\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1q} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2q} & b_2 \\ a_{31} & a_{32} & \cdots & a_{3q} & b_3 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{p1} & a_{p2} & \cdots & a_{pq} & b_p \end{array}\right] \longrightarrow \begin{bmatrix} \underline{*} & * & * & * & * & * & * \\ 0 & 0 & \underline{*} & * & * & * & * \\ 0 & 0 & 0 & \underline{*} & * & * & * \\ 0 & 0 & 0 & 0 & \underline{*} & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & \underline{*} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Things to notice about the echelon form:

- If a 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.

Solving linear equations: Rectangular systems and Row Echelon Form

$$\left[\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1q} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2q} & b_2 \\ a_{31} & a_{32} & \cdots & a_{3q} & b_3 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{p1} & a_{p2} & \cdots & a_{pq} & b_p \end{array}\right] \longrightarrow \begin{bmatrix} \underline{*} & * & * & * & * & * & * \\ 0 & 0 & \underline{*} & * & * & * & * \\ 0 & 0 & 0 & \underline{*} & * & * & * \\ 0 & 0 & 0 & 0 & \underline{*} & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & \underline{*} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Columns containing pivot are called the ***basic columns***, and the others are called the ***non-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.

Solving linear equations: Reduced Row Echelon Form

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

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

The reduced row echelon form reveals structure in the original matrix $\mathbf{A}$.

Solving linear equations: Homogenous Systems

A homogenous system of linear equations,

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

Consider the following case,

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

We choose the free variables as x_2 and x_5, variables corresponding to the non-basic columns.

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

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

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

  where, the variables $x_{f_1}, x_{f_2}, \ldots, x_{f_{q-r}}$ 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? $\longrightarrow rank\,(\mathbf{A}) = q$.

Solving linear equations: Non-homogenous Systems

A non-homogenous system of linear equations: $\mathbf{Ax} = \mathbf{b} \neq \mathbf{0}$.

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

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

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

Solving linear equations: Non-homogenous Systems

What about the case when there are no solutions? When does that happen? $\longrightarrow$ *When the system is not consistent.*

$$\left[\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1q} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2q} & b_2 \\ a_{31} & a_{32} & \cdots & a_{3q} & b_3 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{p1} & a_{p2} & \cdots & a_{pq} & b_p \end{array}\right] \longrightarrow i \left[\begin{array}{ccccc|c} 1 & * & 0 & 0 & * & c_1 \\ 0 & 0 & 1 & 0 & * & c_2 \\ 0 & 0 & 0 & 1 & * & c_3 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & c_p \end{array}\right]$$

There is a problem when $c_p \neq 0$.

The augmented matrix $[\mathbf{A} \mid \mathbf{b}]$ has the same number of basic columns as $\mathbf{A}$.

$[\mathbf{A} \mid \mathbf{b}] \rightarrow [\mathbf{E} \mid \mathbf{c}] \longrightarrow \mathbf{c}$ is a non-basic column.

$$rank\,(\mathbf{A}) = rank\,([\mathbf{A} \mid \mathbf{b}])$$

Solving linear equations: Non-homogenous Systems

A non-homogenous system of linear equations: $\mathbf{Ax} = \mathbf{b} \neq \mathbf{0}$.

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

This system of equations has no solutions since the last row implies $0 = 3$ which is a contradiction.

This system is **inconsistent**.