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. $\begin{bmatrix} 1.1 & -24 & \sqrt{2} \\ 0 & 1.12 & -5.24 \end{bmatrix}$

$$\begin{array}{ccc} & m \text{ columns} & \\ n \text{ rows} & \begin{bmatrix} \square & \square & \square & \cdots & \square \\ \square & \square & \square & \cdots & \square \\ \square & a_{32} & \square & \cdots & \square \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \square & \square & \square & \cdots & \square \end{bmatrix} & \begin{array}{c} \\ \\ \to 3^{rd} \text{ row} \\ \\ \\ \end{array} \\ & 2^{nd} \text{ column} & \end{array}$$

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

- **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 indexed column n-vectors, $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3 \ldots \mathbf{x}_m$.
$$\mathbf{X} = \begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \mathbf{x}_3 & \ldots & \mathbf{x}_m \end{bmatrix}$$
- **Zero matrix**$= \mathbf{0}_{n \times m} = \begin{bmatrix} 0 & 0 & \ldots & 0 \\ 0 & 0 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 0 \end{bmatrix}$
- **Identity matrix** is a square $n \times n$ matrix with all zero elements, except the diagonals where all elements are 1.
$$i_{mn} = \begin{cases} 1 & m = n \\ 0 & m \neq n \end{cases} \quad \mathbf{I}_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \end{bmatrix}$$

Matrices

- **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

- **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? $\mathbf{A} = \begin{bmatrix} \mathbf{B} & \mathbf{C} \\ \mathbf{D} & \mathbf{E} \end{bmatrix}$

Matrix operations: Matrix Addition

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

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

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

Matrix operations: Scalar multiplication

- **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 operations: Matrix multiplication

- A useful multiplication operation can be defined for matrices.
- 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 this *matrix multiplication* procedure.
- The 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 $\mathbf{B}$.

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

Matrix multiplication

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

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

Matrix multiplication: Post-multiplication by a column vector

- 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 $\mathbf{y} = \mathbf{A}\mathbf{x} \in \mathbb{R}^n$.

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

- Post-multiplying a matrix $\mathbf{A}$ by a column vector $\mathbf{x}$ results in a linear combination of the columns of matrix $\mathbf{A}$.

- $\mathbf{x}$ provides the column mixture.

Matrix multiplication: Pre-multiplication by a row vector

- Let $\mathbf{x}^\top \in \mathbb{R}^n$ and $\mathbf{A} \in \mathbb{R}^{n \times m}$, then $\mathbf{y} = \mathbf{x}^\top \mathbf{A}$.

$$\mathbf{y} = \begin{bmatrix} x_1 & \dots & x_n \end{bmatrix} \begin{bmatrix} a_{11} & \dots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nm} \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^n x_i a_{i1} & \dots & \sum_{i=1}^n x_i a_{im} \end{bmatrix} = \sum_{i=1}^n x_i \tilde{\mathbf{a}}_i^\top$$

where, $\tilde{\mathbf{a}}_i^\top = \begin{bmatrix} a_{i1} & \dots & a_{im} \end{bmatrix}$

- Pre-multiplying a matrix $\mathbf{A}$ by a row vector $\mathbf{x}$ results in a linear combination of the rows of $\mathbf{A}$.

- $\mathbf{x}^\top$ provides the row mixture.

Matrix multiplication

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

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

- **Four interpretations of matrix multiplication.**
  1. Inner-Product interpretation
  2. Column interpretation
  3. Row interpretation
  4. Outer product interpretation.

Matrix multiplication: Inner-product Interpreation

$$\mathbf{C} = \mathbf{A}\mathbf{B}, \;\; \mathbf{A} \in \mathbb{R}^{n \times p}, \, \mathbf{B} \in \mathbb{R}^{p \times m}, \, \mathbf{C} \in \mathbb{R}^{n \times m}$$

- ij^{th} element of $\mathbf{C}$ is the inner product of the i^{th} row of $\mathbf{A}$ and the j^{th} column of $\mathbf{B}$.

$$c_{ij} = \sum_{k=1}^{m} a_{ik} b_{kj} = \tilde{\mathbf{a}}_i^\top \mathbf{b}_j$$

where, $i \in \{1 \ldots n\}, j \in \{1 \ldots m\}$

Matrix multiplication: Column interpretation

$$\mathbf{C} = \mathbf{A}\mathbf{B}, \;\; \mathbf{A} \in \mathbb{R}^{n \times p}, \, \mathbf{B} \in \mathbb{R}^{p \times m}, \, \mathbf{C} \in \mathbb{R}^{n \times m}$$

- Columns of $\mathbf{C}$ are the linear combinations of the columns of $\mathbf{A}$.

$$\mathbf{C} = \mathbf{A}\begin{bmatrix}\mathbf{b}_1 & \mathbf{b}_2 & \ldots & \mathbf{b}_m\end{bmatrix} = \begin{bmatrix}\mathbf{A}\mathbf{b}_1 & \mathbf{A}\mathbf{b}_2 & \ldots & \mathbf{A}\mathbf{b}_m\end{bmatrix}$$

- j^{th} column of $\mathbf{C}$ is the linear combination of the columns of $\mathbf{A}$

$$\mathbf{c}_j = \sum_{k=1}^{p} b_{kj}\mathbf{a}_k$$

Matrix multiplication: Row interpretation

$$\mathbf{C} = \mathbf{AB}, \;\; \mathbf{A} \in \mathbb{R}^{n \times p}, \, \mathbf{B} \in \mathbb{R}^{p \times m}, \, \mathbf{C} \in \mathbb{R}^{n \times m}$$

- Rows of $\mathbf{C}$ are the linear combinations of the rows of $\mathbf{B}$.

$$\mathbf{C} = \begin{bmatrix} \tilde{\mathbf{a}}_1^\top \\ \tilde{\mathbf{a}}_2^\top \\ \dots \\ \tilde{\mathbf{a}}_n^\top \end{bmatrix} \mathbf{B} = \begin{bmatrix} \tilde{\mathbf{a}}_1^\top \mathbf{B} \\ \tilde{\mathbf{a}}_2^\top \mathbf{B} \\ \dots \\ \tilde{\mathbf{a}}_n^\top \mathbf{B} \end{bmatrix}$$

- i^{th} row of $\mathbf{C}$ is the linear combination of the rows of $\mathbf{B}$

$$\tilde{\mathbf{c}}_i^\top = \sum_{k=1}^{p} a_{ik} \tilde{\mathbf{b}}_k^\top$$

Matrix multiplication: Outer product interpretation

- **Outer product**: Product between a colum vector and a row vector. Let $\mathbf{x} \in \mathbb{R}^n$ and $\mathbf{y} \in \mathbb{R}^m$. The *outer product* is defined as,

$$\mathbf{x}\mathbf{y}^\top = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \begin{bmatrix} y_1 & y_2 & \ldots & y_m \end{bmatrix} = \begin{bmatrix} x_1 y_1 & x_1 y_2 & \ldots & x_1 y_m \\ x_2 y_1 & x_2 y_2 & \ldots & x_2 y_m \\ \vdots & \vdots & \ddots & \vdots \\ x_n y_1 & x_n y_2 & \ldots & x_n y_m \end{bmatrix} \in \mathbb{R}^{n \times m}$$

Matrix multiplication: Outer product interpretation

$$\mathbf{C} = \mathbf{A}\mathbf{B}, \ \ \mathbf{A} \in \mathbb{R}^{n \times p}, \mathbf{B} \in \mathbb{R}^{p \times m}, \mathbf{C} \in \mathbb{R}^{n \times m}$$

- $\mathbf{C}$ can be written as the sum of p outer products of columns of $\mathbf{A}$ and rows of $\mathbf{B}$.

$$\mathbf{C} = \mathbf{A}\mathbf{B} = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3 & \ldots & \mathbf{a}_p \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{b}}_1^\top \\ \tilde{\mathbf{b}}_2^\top \\ \tilde{\mathbf{b}}_3^\top \\ \vdots \\ \tilde{\mathbf{b}}_p^\top \end{bmatrix} = \sum_{i=1}^{p} \mathbf{a}_i \tilde{\mathbf{b}}_i^\top$$

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

Linear equations

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

$$\left.\begin{aligned} a_{11}x_1 + a_{12}x_2 \ldots + a_{1m}x_m &= b_1 \\ a_{21}x_1 + a_{22}x_2 \ldots + a_{2m}x_m &= b_2 \\ a_{31}x_1 + a_{32}x_2 \ldots + a_{3m}x_m &= b_3 \\ &\vdots \\ a_{n1}x_1 + a_{n2}x_2 \ldots + a_{nm}x_m &= b_n \end{aligned}\right\} \longrightarrow \mathbf{Ax} = \mathbf{b}, \;\; \mathbf{A} \in \mathbb{R}^{n \times m}, \;\; \mathbf{x} \in \mathbb{R}^{n}, \;\; \mathbf{b} \in \mathbb{R}^{m}$$

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \ldots & a_{1m} \\ a_{21} & a_{22} & a_{23} & \ldots & a_{2m} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \ldots & a_{nm} \end{bmatrix} \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix} \quad \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

Linear equations in control problems

$\mathbf{x}$: Input $\mathbf{b}$: Output $\mathbf{A}$: System dynamics

$$\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1m} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2m} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{nm} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix}$$

Linear equations in estimation problems

$\mathbf{x}$: Parameter $\mathbf{b}$: Measurements $\mathbf{A}$: System characteristics

$$\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1m} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2m} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{nm} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix}$$

Rank of a matrix $\mathbf{A}$

- **Rank of a matrix $\mathbf{A}$**: dimension of the subspace spanned by the columns of $\mathbf{A}$ or the rows of $\mathbf{A} \in \mathbb{R}^{n \times m}$.

$$\begin{aligned} rank\left(\mathbf{A}\right) &= \dim span\left(\left\{\mathbf{a}_1, \mathbf{a}_2, \ldots \mathbf{a}_m\right\}\right) \rightarrow \text{Column rank} \\ &= \dim span\left(\left\{\tilde{\mathbf{a}}_1^\top, \tilde{\mathbf{a}}_2^\top, \ldots \tilde{\mathbf{a}}_n^\top\right\}\right) \rightarrow \text{Row rank} \end{aligned}$$

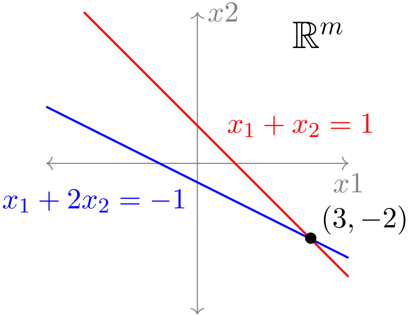
- Column Rank is always equal to the row rank.
- Rank tells us the number of independent columns/row in the matrix.
- **Full rank matrix** $\mathbf{A}$: $rank\left(\mathbf{A}\right) = \min\left(n, m\right)$
  Rank deficient matrix $\mathbf{A}$: $rank\left(\mathbf{A}\right) < \min\left(n, m\right)$

Geometry of linear equations

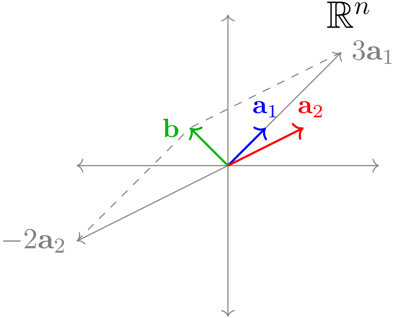
$$\left.\begin{aligned} x_1 + 2x_2 &= -1 \\ x_1 + x_2 &= 1 \end{aligned}\right\} \longrightarrow \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

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

Traditional (row) view



Column view

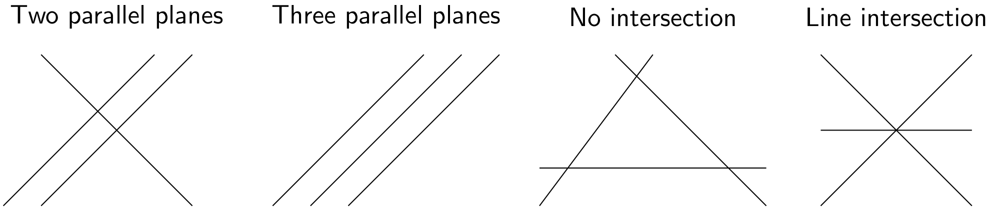


Solutions of linear equations

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

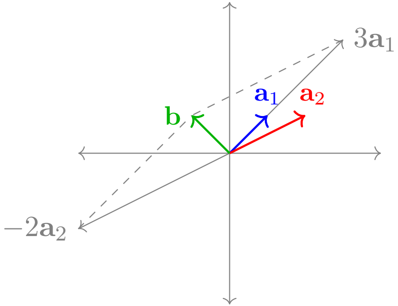
- **Three possible situations:** No solution, Infinitely many solutions, or Unique Solution.

- When do have infinitely many or no solutions? In $\mathbb{R}^3$, we can visualize the different situations.



Understanding $\mathbf{Ax} = \mathbf{b}$: Unique solution

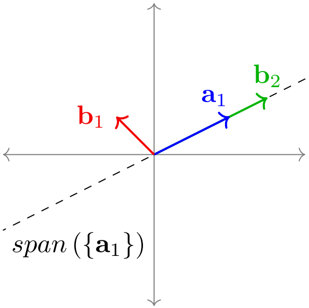
$$\begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$



- Square matrix
- Linearly independent set of columns $\{\mathbf{a}_1, \mathbf{a}_2\}$
- $\mathbf{b} \in span\left(\{\mathbf{a}_1, \mathbf{a}_2\}\right)$.
- Always solvable, and give an unique solution.

Understanding $\mathbf{Ax} = \mathbf{b}$: Unique solution or No solution

1. $\begin{bmatrix} 2 \\ 1 \end{bmatrix} \begin{bmatrix} x_1 \end{bmatrix} = \mathbf{b}_1 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$
2. $\begin{bmatrix} 2 \\ 1 \end{bmatrix} \begin{bmatrix} x_1 \end{bmatrix} = \mathbf{b}_2 = \begin{bmatrix} 3 \\ 1.5 \end{bmatrix}$



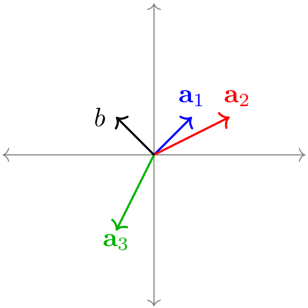
- Tall matrix
- Linearly independent set of columns $\{\mathbf{a}_1\}$

$\mathbf{b}_1 \notin span\left(\{\mathbf{a}_1\}\right) \implies$ Not solvable.

$\mathbf{b}_2 \in span\left(\{\mathbf{a}_1\}\right) \implies$ Solvable with unqiue solution.

Understanding $\mathbf{Ax} = \mathbf{b}$: Infinitely many solution

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



- Fat matrix
- Linearly dependent set of columns $\{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3\}$
- $\mathbf{b} \in span\left(\{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3\}\right)$.
- Always solvable, with infinitely many solutions.

Understanding $\mathbf{Ax} = \mathbf{b}$: Conditions for different types of solutions

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

Full rank $\mathbf{A}$:

- $rank\left(\mathbf{A}\right) = n \implies$ **Always solvable**

$$\begin{cases} n = m & \implies \text{Unique solution} \\ n < m & \implies \text{Infinitely many solutions} \end{cases}$$

- $rank\left(\mathbf{A}\right) = m \implies$ **No infinite solutions**

$$\begin{cases} m = n & \implies \text{Unique solution} \\ m < n & \rightarrow \begin{cases} \mathbf{b} \in span\left(\mathbf{a}_1, \ldots \mathbf{a}_m\right) \implies \text{Unique solution} \\ \mathbf{b} \notin span\left(\mathbf{a}_1, \ldots \mathbf{a}_m\right) \implies \text{No solution} \end{cases} \end{cases}$$

Understanding $\mathbf{Ax} = \mathbf{b}$: Conditions for different types of solutions

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

Rank deficient $\mathbf{A}$:

- $rank\left(\mathbf{A}\right) < \min\left(n, m\right) \implies$ **No unique solution**

$$\begin{cases} \mathbf{b} \in span\left(\mathbf{a}_1, \ldots \mathbf{a}_m\right) \implies \text{Infinitely many solutions} \\ \mathbf{b} \notin span\left(\mathbf{a}_1, \ldots \mathbf{a}_m\right) \implies \text{No solution} \end{cases}$$

Understanding $\mathbf{Ax} = \mathbf{b}$: Conditions for different types of solutions

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

- $\mathbf{b} \notin span\left(\mathbf{a}_1, \ldots \mathbf{a}_m\right) \implies$ No solution
- $\mathbf{b} \in span\left(\mathbf{a}_1, \ldots \mathbf{a}_m\right) \implies \begin{cases} rank\left(\mathbf{A}\right) = m \implies \text{Unique} \\ rank\left(\mathbf{A}\right) < m \implies \text{Infinitely many solutions} \end{cases}$

General solution of linear equations

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

- Assuming that this system can be solved, the most general form of the solution is,

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

  where, $\mathbf{x}_p$ is called the particular solution, and $\mathbf{x}_h$ is the homogenous solution.

- **Homogenous solution**: Solution of the equation $\mathbf{A}\mathbf{x} = \mathbf{0}$.

- The set of all homogenous solutions of $\mathbf{A}$ – $\{\mathbf{x}_h \,|\, \mathbf{A}\mathbf{x}_h = \mathbf{0}\}$ – form a subspace of $\mathbb{R}^m$.

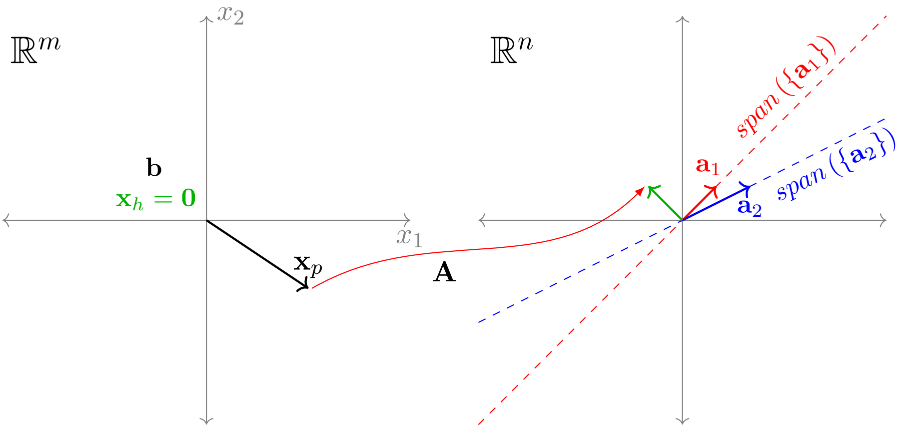
Geometry of the general solution

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

$$\mathbf{A} = \begin{bmatrix} 1 & -1 \\ 2 & -2 \end{bmatrix}, \; \mathbf{b} = \begin{bmatrix} 1.5 \\ 3 \end{bmatrix}$$

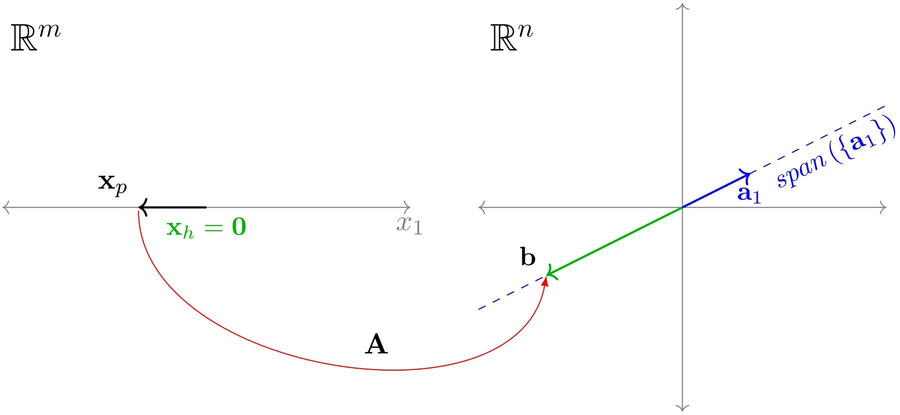
Geometry of the general solution

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



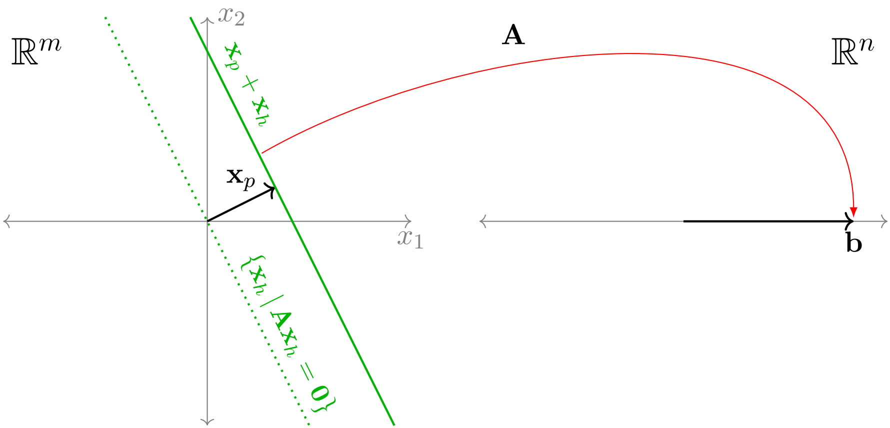
Geometry of the general solution

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



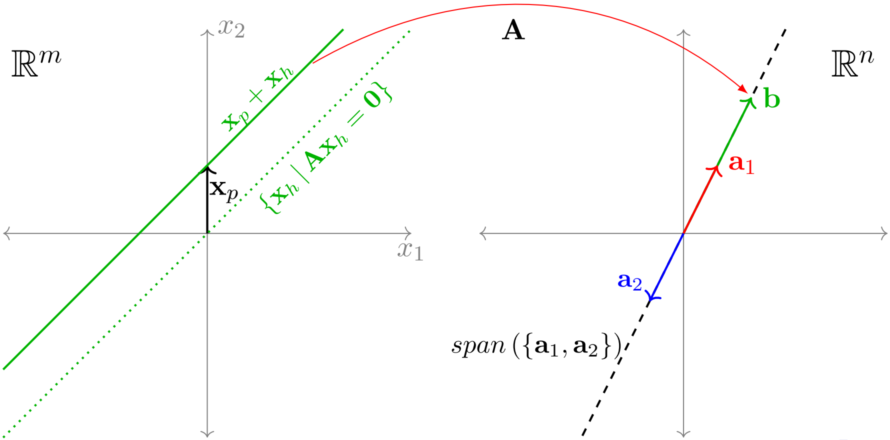
Geometry of the general solution

$$\mathbf{A} = \begin{bmatrix} 2 & 1 \end{bmatrix}, \; \mathbf{b} = \begin{bmatrix} 5 \end{bmatrix}$$



Geometry of the general solution

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



Linear transformations

- Linear functions $f : \mathbb{R}^m \mapsto \mathbb{R}$,

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

- Generalization of the linear function is when its range $\mathbb{R}^n$:

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

- These can be represented as, $\mathbf{y} = \mathbf{A}\mathbf{x}, \;\; \mathbf{A} \in \mathbb{R}^{n \times m}$.

- Matrices can be thought of as representing a particular linear transformation.

Why does matrix multiplication have this strange definition?

Consider the following two functions,

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

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

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

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

Matrix multiplication represents the composition of linear transformations.

Four Fundamental Subspaces of $\mathbf{A} \in \mathbb{R}^{n \times m}$

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

$$\mathcal{C}\left(\mathbf{A}\right) = \left\{\mathbf{A}\mathbf{x} \,|\, \mathbf{x} \in \mathbb{R}^m\right\} \subseteq \mathbb{R}^n$$

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

$$\mathcal{N}\left(\mathbf{A}\right) = \left\{\mathbf{x} \,|\, \mathbf{A}\mathbf{x} = \mathbf{0}\right\} \subseteq \mathbb{R}^m$$

- $\mathcal{C}\left(\mathbf{A}^\top\right)$: **Row Space of $\mathbf{A}$** – the span of the rows of $\mathbf{A}$.

$$\mathcal{C}\left(\mathbf{A}^\top\right) = \left\{\mathbf{A}^\top\mathbf{y} \,|\, \mathbf{y} \in \mathbb{R}^n\right\} \subseteq \mathbb{R}^m$$

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

$$\mathcal{N}\left(\mathbf{A}^\top\right) = \left\{\mathbf{y} \,|\, \mathbf{A}^\top\mathbf{y} = \mathbf{0}\right\} \subseteq \mathbb{R}^n$$

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

Linear Independence

- Given a set of vectors $\{\mathbf{v}_1, \mathbf{v}_2, \ldots \mathbf{v}_m\}, \;\; \mathbf{v}_i \in \mathbb{R}^n$, how can we determine if this set is linear independent?

- We need to verify, $a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_m\mathbf{v}_m = 0$

$$\left.\begin{bmatrix}\mathbf{v}_1 & \cdots & \mathbf{v}_m\end{bmatrix}\begin{bmatrix}a_1 \\ \vdots \\ a_m\end{bmatrix} = \begin{bmatrix}0 \\ \vdots \\ 0\end{bmatrix} = \mathbf{V}\mathbf{a} = \mathbf{0}\right\} \mathcal{N}\left(\mathbf{V}\right) = \left\{\mathbf{0}\right\}, \;\; rank\left(\mathbf{V}\right) = n$$

- This is also equivalent to saying that when the $rank\left(\mathbf{A}\right) = n \implies$ 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?

Dimension of the four fundamental subspaces

- **Column space** $C(\mathbf{A})$
  - $\dim C(\mathbf{A}) = rank\,(\mathbf{A}) = r$
- **Nullspace** $N(\mathbf{A})$
  - $\dim N(\mathbf{A}) = n - r$
- **Row space** $C(\mathbf{A}^\top)$
  - $\dim C(\mathbf{A}^\top) = rank\left(\mathbf{A}^\top\right) = rank\,(\mathbf{A}) = r$
- **Left Nullspace** $N(\mathbf{A}^\top)$
  - $\dim N(\mathbf{A}^\top) = m - r$

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. $rank\left(\mathbf{A}\right) = n$
- $\mathbf{Ax} = \mathbf{b}$ 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, $\left(\mathbf{AB}\right)^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$.
- $\left(\mathbf{A}^{-1}\right)^{-1} = \mathbf{A}$ and $\left(\mathbf{A}^{-1}\right)^{\top} = \left(\mathbf{A}^{T}\right)^{-1}$