Applied Linear Algebra in Data Analysis

Linear Systems and Matrix Operations: Part 4

Sivakumar Balasubramanian

Department of Bioengineering
Christian Medical College, Bagayam
Vellore 632002

Matrix Representation of Simultaneous Linear Equations

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

The general solution to the system of equations is given by,

$$\mathbf{x} = \mathbf{x}_p + \mathbf{x}_h \in \mathbb{R}^q$$

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

Does the set of all homogeneous solutions form a subspace space of $\mathbb{R}^q$?

Matrix Representation of Simultaneous Linear Equations

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

The general solution to the system of equations is given by,

$$\mathbf{x} = \mathbf{x}_p + \mathbf{x}_h \in \mathbb{R}^q$$

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

Does the set of all homogeneous solutions form a subspace space of $\mathbb{R}^q$? **No**, because the set is not closed under addition or scaling.

Linear Functions

Definition: Linear Function

A function f that maps vectors from $\mathbb{R}^n$ to $\mathbb{R}$, i.e. $f : \mathbb{R}^n \longrightarrow \mathbb{R}$, which satisfies the superposition principle is called a *linear function*.

$$\textbf{Superposition} : f\left(\alpha \mathbf{x} + \beta \mathbf{y}\right) = \alpha f\left(\mathbf{x}\right) + \beta f\left(\mathbf{y}\right)$$

where $\alpha, \beta \in \mathbb{R}$ and $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$.

Any linear fucntion $f : \mathbb{R}^n \to \mathbb{R}$can be represented as the inner product with a fixed vector $\mathbf{w} \in \mathbb{R}^n$.

$$\mathbb{R} \ni y = f\left(\mathbf{x}\right) = \mathbf{w}^\top \mathbf{x} = \sum_{i=1}^{n} w_i x_i$$

The vector $\mathbf{w}$ is the representation of the linear function f.

Linear Functions

How do we find out the vector $\mathbf{w} \in \mathbb{R}^n$ of an arbitrary function $f : \mathbb{R}^n \to \mathbb{R}$?

Linear Functions

How do we find out the vector $\mathbf{w} \in \mathbb{R}^n$ of an arbitrary function $f : \mathbb{R}^n \to \mathbb{R}$?

Choose n vectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$, and let the corresponding function values be $y_1 = f(\mathbf{x}_1), y_2 = f(\mathbf{x}_2), \ldots, y_n = f(\mathbf{x}_n)$. We know that, $\mathbf{w}^\top \mathbf{x}_i = y_i$ or $\mathbf{x}_i^\top \mathbf{w} = y_i$ (why?).

Then we have the following set of linear equations,

$$\begin{bmatrix} \mathbf{x}_1^\top \mathbf{w} \\ \mathbf{x}_2^\top \mathbf{w} \\ \vdots \\ \mathbf{x}_n^\top \mathbf{w} \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \implies \begin{bmatrix} \mathbf{x}_1^\top \\ \mathbf{x}_2^\top \\ \vdots \\ \mathbf{x}_n^\top \end{bmatrix} \mathbf{w} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \implies \mathbf{X}\mathbf{w} = \mathbf{y}$$

When can we solve for $\mathbf{w}$ uniquely?

Linear Functions

How do we find out the vector $\mathbf{w} \in \mathbb{R}^n$ of an arbitrary function $f : \mathbb{R}^n \to \mathbb{R}$?

Choose n vectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$, and let the corresponding function values be $y_1 = f(\mathbf{x}_1), y_2 = f(\mathbf{x}_2), \ldots, y_n = f(\mathbf{x}_n)$. We know that, $\mathbf{w}^\top \mathbf{x}_i = y_i$ or $\mathbf{x}_i^\top \mathbf{w} = y_i$ (why?).

Then we have the following set of linear equations,

$$\begin{bmatrix} \mathbf{x}_1^\top \mathbf{w} \\ \mathbf{x}_2^\top \mathbf{w} \\ \vdots \\ \mathbf{x}_n^\top \mathbf{w} \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \implies \begin{bmatrix} \mathbf{x}_1^\top \\ \mathbf{x}_2^\top \\ \vdots \\ \mathbf{x}_n^\top \end{bmatrix} \mathbf{w} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \implies \mathbf{X}\mathbf{w} = \mathbf{y}$$

When can we solve for $\mathbf{w}$ uniquely? When $\mathbf{X}$ is a square matrix and invertible (full rank).

What is the simplest possible $\mathbf{X}$?

How do we find out the vector $\mathbf{w} \in \mathbb{R}^n$ of an arbitrary function $f : \mathbb{R}^n \to \mathbb{R}$?

Choose n vectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$, and let the corresponding function values be $y_1 = f(\mathbf{x}_1), y_2 = f(\mathbf{x}_2), \ldots, y_n = f(\mathbf{x}_n)$. We know that, $\mathbf{w}^\top \mathbf{x}_i = y_i$ or $\mathbf{x}_i^\top \mathbf{w} = y_i$ (why?).

Then we have the following set of linear equations,

$$\begin{bmatrix} \mathbf{x}_1^\top \mathbf{w} \\ \mathbf{x}_2^\top \mathbf{w} \\ \vdots \\ \mathbf{x}_n^\top \mathbf{w} \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \implies \begin{bmatrix} \mathbf{x}_1^\top \\ \mathbf{x}_2^\top \\ \vdots \\ \mathbf{x}_n^\top \end{bmatrix} \mathbf{w} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \implies \mathbf{X}\mathbf{w} = \mathbf{y}$$

When can we solve for $\mathbf{w}$ uniquely? When $\mathbf{X}$ is a square matrix and invertible (full rank).

What is the simplest possible $\mathbf{X}$? When $\mathbf{X} = \mathbf{I}_n$.

Linear Transformations/Maps

Definition: Linear Transformation/Map

A *linear transformation* or *linear map* is a function $\mathbf{f} : \mathbb{R}^q \mapsto \mathbb{R}^p$ that satisfies the following properties:

$$\textbf{Superposition} : \mathbf{f}\left(\alpha \mathbf{x} + \beta \mathbf{y}\right) = \alpha \mathbf{f}\left(\mathbf{x}\right) + \beta \mathbf{f}\left(\mathbf{y}\right)$$

where $\alpha, \beta \in \mathbb{R}$ and $\mathbf{x}, \mathbf{y} \in \mathbb{R}^q$ and $\mathbf{f}\left(\mathbf{x}\right), \mathbf{f}\left(\mathbf{y}\right) \in \mathbb{R}^p$.

Each component of the vector $\mathbf{f}\left(\mathbf{x}\right)$ is a linear function of $\mathbf{x}$.

$$\mathbf{y} = \mathbf{f}\left(\mathbf{x}\right) = \begin{bmatrix} f_1\left(\mathbf{x}\right) \\ f_2\left(\mathbf{x}\right) \\ \vdots \\ f_n\left(\mathbf{x}\right) \end{bmatrix} = \begin{bmatrix} \tilde{\mathbf{a}}_1^\top \mathbf{x} \\ \tilde{\mathbf{a}}_2^\top \mathbf{x} \\ \vdots \\ \mathbf{a}_n^\top \mathbf{x} \end{bmatrix} = \mathbf{A}\mathbf{x}, \quad \mathbf{A} \in \mathbb{R}^{p \times q}$$

$\tilde{\mathbf{a}}_i$ vectors represents the individual linear function f_i. The matrix $\mathbf{A}$ is representation of the linear transformation $\mathbf{f}$.

Linear Transformations/Maps

How do we find out the $\mathbf{A} \in \mathbb{R}^{p \times q}$ of an arbitrary linear transformation $\mathbf{f} : \mathbb{R}^q \to \mathbb{R}^p$?

Linear Transformations/Maps

How do we find out the $\mathbf{A} \in \mathbb{R}^{p \times q}$ of an arbitrary linear transformation $\mathbf{f} : \mathbb{R}^q \to \mathbb{R}^p$?

Choose a linear independent set of q vectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_q \in \mathbb{R}^q$. This can be used to identify the individual vectors $\tilde{\mathbf{a}}$, and thus help identify the matrix $\mathbf{A}$.

What is the simplest possible set of vectoes $\mathbf{x}_1 \cdots \mathbf{x}_q$?

Linear Transformations/Maps

How do we find out the $\mathbf{A} \in \mathbb{R}^{p \times q}$ of an arbitrary linear transformation $\mathbf{f} : \mathbb{R}^q \rightarrow \mathbb{R}^p$?

Choose a linear independent set of q vectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_q \in \mathbb{R}^q$. This can be used to identify the individual vectors $\tilde{\mathbf{a}}$, and thus help identify the matrix $\mathbf{A}$.

What is the simplest possible set of vectoes $\mathbf{x}_1 \cdots \mathbf{x}_q$?

When $\mathbf{x}_i = \mathbf{e}_i$.

How do we find out the $\mathbf{A} \in \mathbb{R}^{p\times q}$ of an arbitrary linear transformation $\mathbf{f} : \mathbb{R}^q \rightarrow \mathbb{R}^p$?

Choose a linear independent set of q vectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_q \in \mathbb{R}^q$. This can be used to identify the individual vectors $\tilde{\mathbf{a}}$, and thus help identify the matrix $\mathbf{A}$.

What is the simplest possible set of vectoes $\mathbf{x}_1 \cdots \mathbf{x}_q$?

When $\mathbf{x}_i = \mathbf{e}_i$.

Each of these inputs will directly read out the columns of the matrix $\mathbf{A}$.

Matrix Multipliation represents Composition of Linear Transformations

Consider the following two linear transformations,

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

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

What will be $\mathbf{z} = \left(\mathbf{f} \circ \mathbf{g}\right)\left(\mathbf{u}\right) = \mathbf{f}\left(\mathbf{g}\left(\mathbf{u}\right)\right) = \mathbf{h}\left(\mathbf{u}\right)$?

Matrix Multipliation represents Composition of Linear Transformations

Consider the following two linear transformations,

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

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

What will be $\mathbf{z} = (\mathbf{f} \circ \mathbf{g})(\mathbf{u}) = \mathbf{f}(\mathbf{g}(\mathbf{u})) = \mathbf{h}(\mathbf{u})$?

$$\mathbf{z} = \mathbf{h}(\mathbf{u}) = \mathbf{f}(\mathbf{g}(\mathbf{u})) = \mathbf{f}(\mathbf{B}\mathbf{u}) = (\mathbf{A}\mathbf{B})\,\mathbf{u}$$

Through simple substitution, we can find that,

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

Geometry of Linear Equations

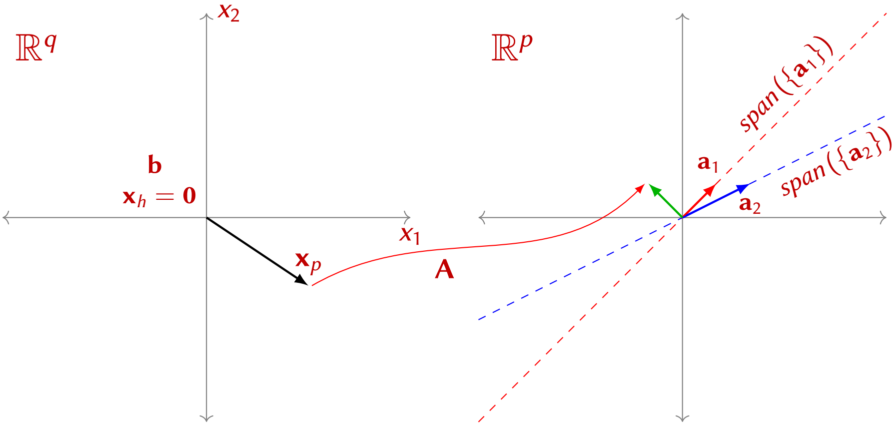
We can interpret linear equations geometrically as a transformation from one space to another.

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

Geometry of Linear Equations

We can interpret linear equations geometrically as a transformation from one space to another.

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

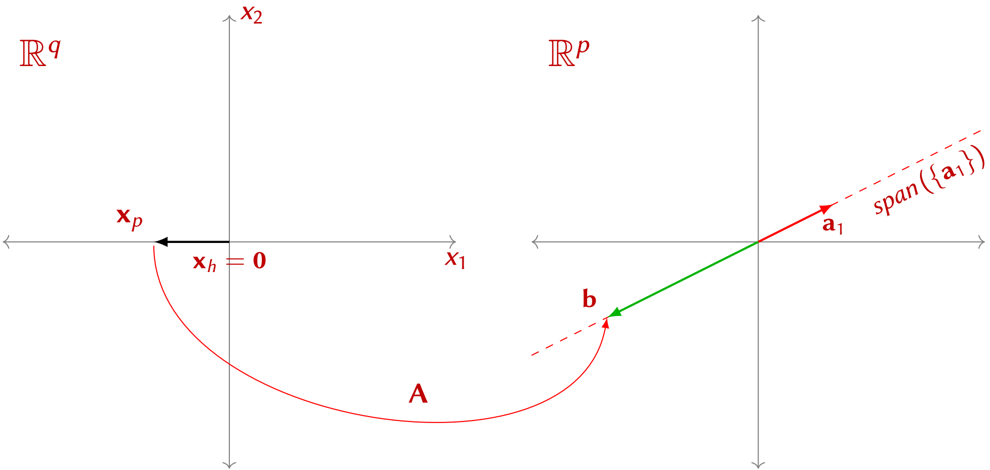


Geometry of Linear Equations

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

Geometry of Linear Equations

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

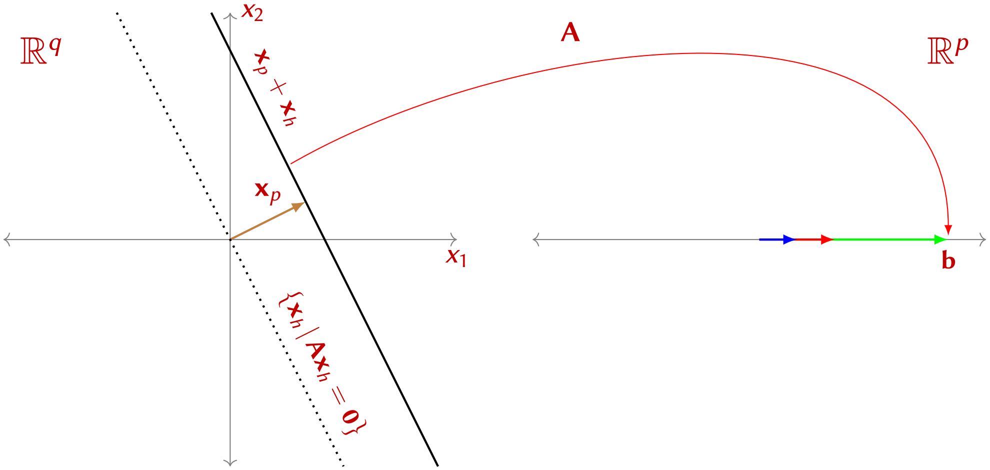


Geometry of Linear Equations

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

Geometry of Linear Equations

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

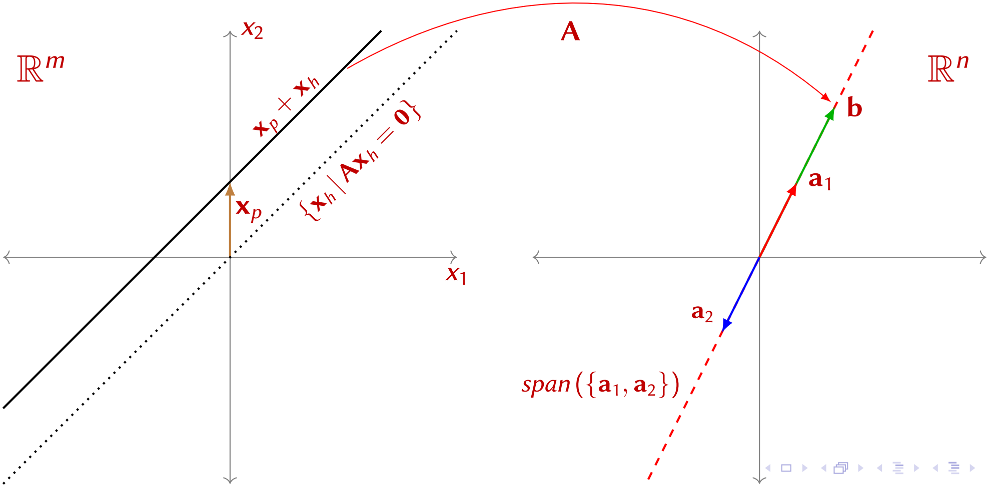


Geometry of Linear Equations

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

Geometry of Linear Equations

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



Four Fundamental Subspace of a matrix $\mathbf{A} \in \mathbb{R}^{p \times q}$ (or a linear transformation)

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

$$\mathcal{C}(\mathbf{A}) = \{\mathbf{A}\mathbf{x} \,|\, \mathbf{x} \in \mathbb{R}^q\} \subseteq \mathbb{R}^p$$

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

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

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

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

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

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

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

Examples

Let's find the four fundamental subspaces of the the following matrices.

Example 1: $\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix}$

Examples

Let's find the four fundamental subspaces of the the following matrices.

Example 1: $\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix}$

Example 2: $\mathbf{A} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$

Examples

Let's find the four fundamental subspaces of the the following matrices.

Example 1: $\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix}$

Example 2: $\mathbf{A} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$

Example 3: $\mathbf{A} = \begin{bmatrix} 2 & 1 \end{bmatrix}$

Examples

Let's find the four fundamental subspaces of the the following matrices.

Example 1: $\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix}$

Example 2: $\mathbf{A} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$

Example 3: $\mathbf{A} = \begin{bmatrix} 2 & 1 \end{bmatrix}$

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

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 decomosed into the product of a lower triangular and an upper triangular matrix.
- Matrices associated with the three elementary operations:

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

LU Factorization of a Matrix

- Consider the case: $\mathbf{A} = \begin{bmatrix} 1 & -1 & 2 \\ 2 & 0 & 2 \\ 4 & 2 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 4 & 3 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 & 2 \\ 0 & 2 & -2 \\ 0 & 0 & -1 \end{bmatrix} = \mathbf{LU}$
- **LU** factorization can be done only when no zero pivot is encountered during the Guassian elimination process.
- $\mathbf{Ax} = \mathbf{b}$ becomes $\mathbf{LUx} = \mathbf{b}$: This is decomposed into two triangular systems, $\mathbf{Ux} = \mathbf{y}, \quad \mathbf{Ly} = \mathbf{b}$. First solve $\mathbf{Ly} = \mathbf{b}$ and then solve $\mathbf{Ux} = \mathbf{y}$
- Properties:
  - Diagonal elements of **L** are 1, and **U** are not equal to zero.
  - **U** is the final result of Guassian elimination, and **L** is the matrix that reverses this process.
- 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.

$\mathbf{PA} = \mathbf{LU}$ 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 permunation matrix, which is the elementary matrix for row exchanges.
- In the current example, the following allows matrix factorization.

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