Computational Linear Algebra
EK103

Giacomo Cappelletto

21/1/2

Contents

Chapter 1

Basics

1.1 Vectors, Norms and Products

Note:-

Let us consider two vectors in $\mathbb{R}^3$:

$$u = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \quad \text{and} \quad v = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}.$$

We wish to compute their magnitudes (norms and norm-squared), the angle between them, and the plane that they span. These methods are directly applicable to computational tools such as MATLAB.

Definition 1.1.1: Norm of a Vector

For a vector $x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$, its norm is

$$\|x\| = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}.$$

In many programming languages (including MATLAB), this is computed via `norm(x)`, while the square of the norm is $\|x\|^2 = x \cdot x = x_1^2 + \cdots + x_n^2$.
Norm squared is the result of the dot product of a vector with itself. For example, the norm squared of x is

$$\|x\|^2 = x \cdot x = x_1^2 + \cdots + x_n^2 = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \cdot \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix}.$$

Example 1.1.1 (Norms and Norm-Squared of u and v)

$$\|u\| = \sqrt{1^2 + 1^2 + 1^2} = \sqrt{3}, \quad \|v\| = \sqrt{1^2 + (-1)^2 + 1^2} = \sqrt{3}.$$

Thus, both vectors have the same magnitude $\sqrt{3}$. Their squared norms are

$$\|u\|^2 = 3, \quad \|v\|^2 = 3.$$

In MATLAB notation, one could write:

- `norm(u)` or `norm(u,2)` for the norm of u.
- `dot(u,u)` or `norm(u)^2` for $\|u\|^2$.

Definition 1.1.2: Angle Between Two Vectors

The angle θ between two nonzero vectors u and v in $\mathbb{R}^n$ is given by

$$\theta = \arccos\left(\frac{u \cdot v}{\|u\|\|v\|}\right).$$

Example 1.1.2 (Angle Between u and v)

First, compute the dot product:

$$u \cdot v = (1)(1) + (1)(-1) + (1)(1) = 1 - 1 + 1 = 1.$$

Hence,

$$\theta = \arccos\left(\frac{u \cdot v}{\|u\|\|v\|}\right) = \arccos\left(\frac{1}{\sqrt{3}\,\sqrt{3}}\right) = \arccos\left(\tfrac{1}{3}\right).$$

In MATLAB, one could write:

```
theta = acos(dot(u,v)/(norm(u)*norm(v)));
```

Definition 1.1.3: Plane Spanned by Two Vectors

The plane containing vectors u and v and passing through the origin is given by

$$\{\, \alpha\, u + \beta\, v \mid \alpha, \beta \in \mathbb{R}\}.$$

An equivalent description is all points $x \in \mathbb{R}^3$ such that $x \cdot (u \times v) = 0$.

Example 1.1.3 (Plane Containing u and v)

- *Span form:*

$$\text{Plane} = \left\{ \alpha \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + \beta \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} \;\middle|\; \alpha, \beta \in \mathbb{R} \right\}.$$

- *Normal form:* The cross product

$$u \times v = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 1 & 1 & 1 \\ 1 & -1 & 1 \end{vmatrix} = (2,\, 0,\, -2).$$

  Hence, the plane also can be described by the set of points $x = (x_1, x_2, x_3)$ for which

$$(2,\, 0,\, -2) \cdot (x_1, x_2, x_3) = 0 \quad \Longrightarrow \quad 2\,x_1 - 2\,x_3 = 0 \quad \Longrightarrow \quad x_1 = x_3.$$

In many computational environments, one simply keeps the span form or uses a symbolic package to compute the cross product and normal equation.

Definition 1.1.4: Cross Product

Construct a system of linear equations where the dot product of the vector is orthogonal to both the vectors in the matrix.

$$\begin{bmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Example 1.1.4 (Finding the Plane Spanned by Two Vectors)

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = t \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$$

Then, we must formulate an equation for any vector perpendicular to the normal of the plane, i.e. the cross product of the two original vectors.

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} = 0$$

Hence,

$$y_1 - y_3 = 0$$

Definition 1.1.5: Dot product

We can take a vector $\vec{v}$ in $\mathbb{R}^n$ and a vector $\vec{w}$ in $\mathbb{R}^n$. Then, the dot product of $\vec{v}$ and $\vec{w}$ is defined as

$$\vec{v} \cdot \vec{w} = \vec{v}^T \vec{w} = \vec{w}^T \vec{v}$$

Therefore

$$\|\vec{v}\|^2 = \vec{v}^T \vec{v}$$

Definition 1.1.6: Scalar Multiplication

Scalar multiplication is the operation of multiplying a vector by a scalar. The result is a new vector with the same direction as the original vector, but with a magnitude that is the product of the original magnitude and the scalar.

$$t \cdot \vec{v} = t \cdot \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} t\, a_1 \\ t\, a_2 \\ \vdots \\ t\, a_n \end{bmatrix}$$

Where

$$\|t \cdot \vec{v}\| = \|\vec{v}\| \cdot t$$

Definition 1.1.7: Vector Addition

Vector addition is the operation of adding two vectors together. The result is a new vector that is the sum of the two original vectors.

$$\vec{v} + \vec{w} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} = \begin{bmatrix} a_1 + b_1 \\ a_2 + b_2 \\ \vdots \\ a_n + b_n \end{bmatrix}$$

Definition 1.1.8: Matrix to Vector Multiplication

Matrix to vector multiplication is the operation of multiplying a matrix by a vector. The result is a new vector that is the result of the matrix-vector multiplication. Matrices are represented as $n \times m$, where n is the number of rows and m is the number of columns. The number of columns of the matrix must be equal to the number of rows of the vector.

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n \end{bmatrix}$$

Definition 1.1.9: Matrix to Matrix Multiplication

Matrix to matrix multiplication is the non-commutative operation of multiplying two matrices together. The result is a new matrix that is the product of the two original matrices.
For two matrices A and B to be multiplied, the number of columns of A must be equal to the number of rows of B. If A is an $m \times n$ matrix and B is an $n \times p$ matrix, then their product $C = AB$ is an $m \times p$ matrix. The element c_{ij} of the resulting matrix C is computed as:

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$$

where a_{ik} is the element from the i-th row and k-th column of matrix A, and b_{kj} is the element from the k-th row and j-th column of matrix B.

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1p} \\ b_{21} & b_{22} & \cdots & b_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{np} \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1p} \\ c_{21} & c_{22} & \cdots & c_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ c_{m1} & c_{m2} & \cdots & c_{mp} \end{bmatrix}$$

Example 1.1.5 (Matrix to Matrix Multiplication)

Consider the matrices

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 7 & 8 & 9 \\ 10 & 11 & 12 \end{bmatrix}.$$

Their product $C = AB$ is computed as follows:

$$C = \begin{bmatrix} 1\cdot 7 + 2\cdot 10 & 1\cdot 8 + 2\cdot 11 & 1\cdot 9 + 2\cdot 12 \\ 3\cdot 7 + 4\cdot 10 & 3\cdot 8 + 4\cdot 11 & 3\cdot 9 + 4\cdot 12 \\ 5\cdot 7 + 6\cdot 10 & 5\cdot 8 + 6\cdot 11 & 5\cdot 9 + 6\cdot 12 \end{bmatrix} = \begin{bmatrix} 27 & 30 & 33 \\ 61 & 68 & 75 \\ 95 & 106 & 117 \end{bmatrix}.$$

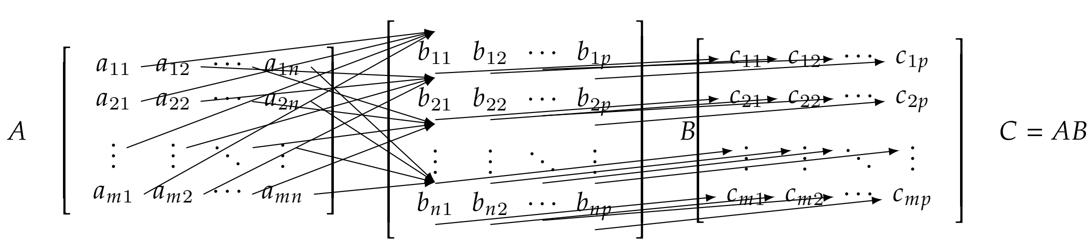


Figure 1.1: Matrix-Matrix Multiplication Diagram

Definition 1.1.10: Matrix Multiplication of a Matrix with Itself

When a matrix A is multiplied by its transpose A^T, the resulting matrix is a symmetric matrix. The element c_{ij} of the resulting matrix $C = AA^T$ is computed as:

$$c_{ij} = \sum_{k=1}^{n} a_{ik} a_{jk}$$

where a_{ik} is the element from the i-th row and k-th column of matrix A, and a_{jk} is the element from the j-th row and k-th column of matrix A.

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1m} \\ c_{21} & c_{22} & \cdots & c_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ c_{m1} & c_{m2} & \cdots & c_{mm} \end{bmatrix}$$

The zeros in this matrix represent orthogonality between the corresponding rows of the original matrix A. Specifically, if $c_{ij} = 0$, it means that the i-th row and the j-th row of matrix A are orthogonal to each other.

Example 1.1.6 (Orthogonality in Matrix Multiplication)

Consider the matrix

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Its transpose is

$$A^T = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

The product $C = AA^T$ is computed as follows:

$$C = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Notice that the resulting matrix is the identity matrix, which is symmetric and has zeros in all off-diagonal elements, indicating that the rows (and columns) of the original matrix A are orthogonal to each other.

1.1.1 Interpretation of vectors in $\mathbb{R}^{2,3}$

Definition 1.1.11: Position Vector

A position vector is a vector that describes the position of an object in space with reference to an origin.

Definition 1.1.12: Translational Vector

A translational vector is a vector that describes the displacement of an object in space with reference to an origin.

1.2 Rotation Matrices

Note:-

We are considering vectors in $\mathbb{R}^2$, denoted by

$$\vec{v} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

We also have a 2×2 matrix (linear operator or transformation) given by

$$A = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}.$$

Note:-

Applying A to an input vector $\vec{v}_i$ produces the output vector $\vec{v}_o$:

$$A\,\vec{v}_i = \vec{v}_o.$$

Explicitly,

$$\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}.$$

Question 1: Transforming a region

Suppose we restrict the input vectors $\vec{v}_i$ to the square region in the (x_1, x_2)-plane with coordinates

$$0 \leqslant x_1 \leqslant 1, \quad 0 \leqslant x_2 \leqslant 1.$$

- How does A map this square region in the input space to a region in the (y_1, y_2)-plane?
- Geometrically, what does that image region look like?

Solution: We can answer this question *mechanically* by observing what happens to the corners of the square, or by spanning the space with two basis vectors:

$$\vec{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \vec{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Each point in the square can be written as

$$x_1\,\vec{e}_1 + x_2\,\vec{e}_2 \quad \text{with } 0 \leq x_1, x_2 \leq 1.$$

Applying A to these basis vectors, we get

$$A\,\vec{e}_1 = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad A\,\vec{e}_2 = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}.$$

Hence, the image of the unit square (spanned by $\vec{e}_1$ and $\vec{e}_2$) is the parallelogram spanned by

$$A\,\vec{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \text{and} \quad A\,\vec{e}_2 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}.$$

In other words, every point (x_1, x_2) in the original square is mapped to

$$x_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} x_1 + 2\,x_2 \\ x_2 \end{bmatrix},$$

with $0 \leq x_1, x_2 \leq 1$. Geometrically, this results in a parallelogram in the (y_1, y_2)-plane whose vertices are $(0, 0)$, $(1, 0)$, $(1 + 2, 1)$, and $(2, 1)$.

Note:-

Thus, restricting $\vec{v}_i$ to a square region in the input space restricts $\vec{v}_o$ to a parallelogram in the output space.

Algorithm 1: Mapping a unit square under the linear transformation A

Input: Input vector $\vec{v}_i = [x_1\ x_2]^T$ with $0 \le x_1, x_2 \le 1$
Output: Output vector $\vec{v}_o = [y_1\ y_2]^T$ in the parallelogram

```
/* Matrix A: */
A ← [[1, 2], [0, 1]];
/* Apply A to input: */
y1 ← x1 + 2 x2;
y2 ← x2;
return v_o;
```

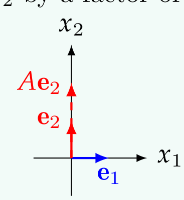
$$A \leftarrow \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix};$$

$$y_1 \leftarrow x_1 + 2\,x_2; \quad y_2 \leftarrow x_2; \quad \textbf{return } \vec{v}_o;$$

1.2.1 Examples of 2D Matrix Transformations

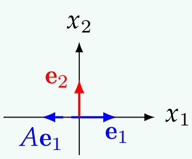
Example 1.2.1 ($A = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$)

Transformation: Vertical stretch scaling x_2 by a factor of 2.



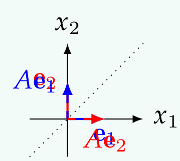
Example 1.2.2 ($A = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$)

Transformation: Reflection about the vertical axis.

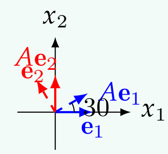


Example 1.2.3 ($A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$)

Transformation: Reflection about the line $x_1 = x_2$.

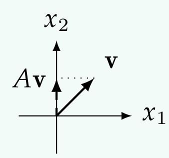


Example 1.2.4 ($A = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$)



Transformation: Counterclockwise rotation by θ about the origin.

Example 1.2.5 ($A = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$)



Transformation: Projection onto the vertical axis.

1.3 Mapping

$$[x_1,\ x_2] \longmapsto \text{point } (x_1, x_2).$$

Under the linear transformation given by

$$[x_1,\ x_2] \mapsto \begin{bmatrix} x_1 + 2x_2 \\ x_2 \end{bmatrix},$$

we see that points are "stretched to the right" in the x_1-direction, while the x_2 coordinate remains unchanged.

Method 1 – Vector Method

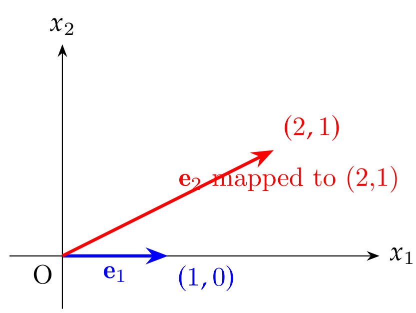
The matrix under consideration is

$$A = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}.$$

The columns of this matrix are:

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 2 \\ 1 \end{bmatrix}.$$

These are the images of the standard basis vectors $\mathbf{e}_1$ and $\mathbf{e}_2$, respectively.



Method 2 – Vertex Method

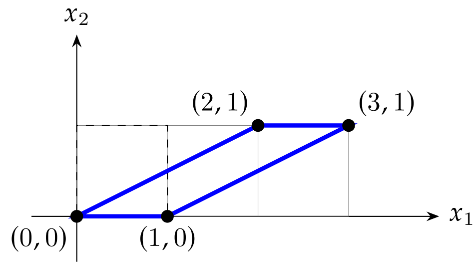
Consider the unit square with vertices

$$(0,0), \quad (1,0), \quad (1,1), \quad (0,1).$$

We apply A to each vertex:

$$A\begin{bmatrix}0\\0\end{bmatrix}=\begin{bmatrix}0\\0\end{bmatrix},\quad A\begin{bmatrix}1\\0\end{bmatrix}=\begin{bmatrix}1\\0\end{bmatrix},\quad A\begin{bmatrix}0\\1\end{bmatrix}=\begin{bmatrix}2\\1\end{bmatrix},\quad A\begin{bmatrix}1\\1\end{bmatrix}=\begin{bmatrix}3\\1\end{bmatrix}.$$

Hence the new vertices are $(0,0)$, $(1,0)$, $(2,1)$, and $(3,1)$. Plotting these points yields the transformed parallelogram (the image of the original unit square).



Definition 1.3.1: Rules for Linear Transformations

1. **Straight lines remain straight.**
2. **Parallel lines remain parallel.**
3. **Distances along lines scale in a consistent, proportional way.**

Justification

1. Straight lines stay straight. A line in parametric form is:

$$r(t)=t\,\mathbf{v}+\mathbf{w}.$$

Applying A gives:

$$A(r(t))=A(t\,\mathbf{v}+\mathbf{w})=t\,A(\mathbf{v})+A(\mathbf{w}),$$

which is again a parametric line.

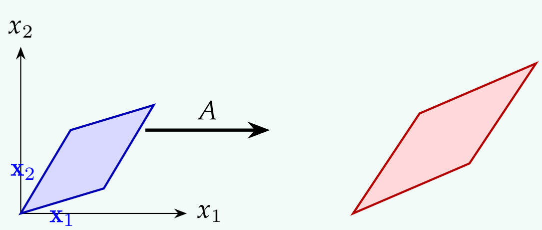
2. Parallel lines stay parallel. If two lines are parallel, their direction vectors are scalar multiples of each other. After applying A, the resulting direction vectors are $A(\mathbf{v})$ for each original direction $\mathbf{v}$. Since A is linear, any scalar multiples remain so, preserving parallelism.

3. Distances scale proportionally. For two points $\mathbf{v}_1$ and $\mathbf{v}_2$, the difference is $\mathbf{d}=\mathbf{v}_2-\mathbf{v}_1$. Under A:

$$A\mathbf{v}_2-A\mathbf{v}_1=A(\mathbf{v}_2-\mathbf{v}_1)=A(\mathbf{d}).$$

Thus the new distance $\|\mathbf{d}'\|=\|A(\mathbf{d})\|$ is a consistent transform of $\|\mathbf{d}\|$, depending on the nature of A.

Example 1.3.1 (A matrix A that acts on a parallelogram (spanned by two vectors $\mathbf{x}_1,\mathbf{x}_2$) will produce another parallelogram in the output plane.)



Here, lines remain lines, parallels remain parallel.

1.4 A System of Linear Equations

We have a matrix A of size $m \times n$, multiplying an unknown vector $\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ (which belongs to $\mathbb{R}^n$), producing a result vector $\begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_m \end{bmatrix}$ in $\mathbb{R}^m$:

$$A_{m\times n} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_m \end{bmatrix}.$$

This can be viewed as a *system of m linear equations in n unknowns.* We are often interested in two main questions:

1. **Does a solution exist?** That is, can we find $x_1, x_2, \ldots, x_n$ so that $A\mathbf{x} = \mathbf{v}$? Geometrically, this asks if $\mathbf{v}$ lies in the *column space* (or image) of A.
2. **If at least one solution exists, is it unique or are there infinitely many?** Uniqueness is typically tied to whether the columns of A are linearly independent (and whether m, n are related in a way that gives a single solution). If there are *fewer* pivots than unknowns, or if the system is underdetermined, infinitely many solutions can occur.

Key intuition:

- The question of *existence* boils down to whether $\mathbf{v}$ is in the span of the columns of A.
- The question of *uniqueness* depends on whether those columns form a set of independent vectors and on the relationship between m and n.

1.5 How to Determine Consistency, Uniqueness, and the Number of Solutions for a Linear System

Note:-

We consider a linear system

$$\begin{cases} a_{11}x_1 \ + \ a_{12}x_2 \ + \ \cdots + a_{1n}x_n \ = \ b_1 \\ a_{21}x_1 \ + \ a_{22}x_2 \ + \ \cdots + a_{2n}x_n \ = \ b_2 \\ \vdots \\ a_{m1}x_1 \ + \ a_{m2}x_2 \ + \ \cdots + a_{mn}x_n \ = \ b_m \end{cases}$$

and write it in *augmented matrix* form:

$$\begin{bmatrix} A \mid \mathbf{b} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{bmatrix}.$$

- We perform *elementary row operations* (EROs) on $\begin{bmatrix} A \mid \mathbf{b} \end{bmatrix}$ to attempt to solve the system. The EROs are:
  1. **Scaling (multiplying a row by a nonzero constant).**

2. **Row replacement** ($R_i = R_i + c\,R_j$ for $i \neq j$).
3. **Interchanging two rows** ($R_i \leftrightarrow R_j$).

These operations do not change the solution set of the system.

Row Echelon Form (REF) We say $\left[A \mid \mathbf{b}\right]$ is in *row echelon form* if:

- All rows of all zeros (if any) are at the bottom of the matrix.
- Each *pivot* (leftmost nonzero entry in a nonzero row) is strictly to the right of the pivot in the row above.

A pivot column in A corresponds to a *leading variable* (or *basic variable*), and any other column (apart from the augmented column) is called a *free column* (its corresponding variable is a *free variable*).
Consistency & Number of Solutions

- If, in the augmented matrix, there is a pivot in the last column (meaning a row of the form

$$[\,0 \quad 0 \quad \cdots \quad 0 \mid c\,], \quad c \neq 0,$$

) then the system is *inconsistent* (no solutions). This corresponds to an equation $0 = c$ where $c \neq 0$, which is impossible.

- If no such contradiction is found, then at least one solution exists (the system is *consistent*).
- The number of pivot columns in A (i.e. the number of leading variables) tells us whether solutions are unique or infinite:
  1. If the number of pivots equals the number of unknowns n, then there is exactly one solution (assuming no inconsistency).
  2. If the number of pivots is less than n, then there are free variables, implying infinitely many solutions (again, assuming no inconsistency).

Example 1.5.1 (Example: Augmented Matrix and REF)

$$\begin{bmatrix} 1 & 2 & 3 & 5 \\ 2 & 4 & 6 & 9 \\ 3 & 6 & 9 & 15 \end{bmatrix} \longrightarrow \text{(REF)}.$$

One performs row operations to get an upper-triangular or echelon form. If the last row becomes something like

$$[\,0 \;\; 0 \;\; 0 \mid c\,], \quad c \neq 0,$$

then there is no solution. Otherwise, we identify pivot columns, read off the relationships among variables, and find a general solution (unique or infinite).

Definition 1.5.1: Free Columns and Basic Columns

Once an augmented matrix is in REF, we label each pivot column (in A) as a *basic column*, and any other column (except the last augmented column) as a *free column*.

- If x_j corresponds to a free column, we may label x_j as a parameter (e.g. $t_1, t_2, \ldots$).
- The variables in pivot columns can then be written in terms of these parameters.

In this manner, we get a general solution describing the entire solution set to $A\mathbf{x} = \mathbf{b}$.

Algorithm to Convert $\left[A \mid \mathbf{b}\right]$ to REF(A)

1. *Select a candidate row*: Choose the topmost row *among those not yet having a pivot* in which a pivot might appear.

2. *Pivot search and possible row swap*: Among this candidate row and those below it, find a row having a leftmost nonzero entry in the desired pivot column. Interchange (swap) that row with the candidate row if needed, placing a nonzero entry where your pivot should be.

3. *Declare the pivot and eliminate below*: Scale the pivot row (if desired) so that the pivot becomes 1. Then use row replacement to produce zeros below that pivot in the same column.

4. Move to the next row down and next column to the right, and repeat until you have a row echelon form.

One can then further use row replacement operations to clear the entries *above* each pivot, yielding the *reduced* row echelon form (RREF). However, for most solution purposes, REF is already sufficient to read off whether solutions exist, how many, and so on.

1.5.1 FRR

Algorithm 2: Forward Row Reduction (Forward Elimination)

Input: Matrix A of size $m \times n$ and vector $\mathbf{b}$ of size m

Output: Upper triangular matrix A and modified vector $\mathbf{b}$

```
/* Forward elimination process */
for k ← 1 to min(m, n) do
    for i ← k + 1 to m do
        if A_kk ≠ 0 then
            f ← A_ik / A_kk;
            for j ← k to n do
                A_ij ← A_ij − f · A_kj;
            end
            b_i ← b_i − f · b_k;
        end
    end
end
return A, b;
```

1.6 Back Row Reduction

Note:-

Back row reduction, also known as back substitution, is a method used to solve a system of linear equations that has been transformed into an upper triangular form through Gaussian elimination. This method involves solving the equations starting from the last row and moving upwards.

Definition 1.6.1: Back Row Reduction

Consider a system of linear equations represented in matrix form as $A\mathbf{x} = \mathbf{b}$, where A is an upper triangular matrix:

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}.$$

The solution is obtained by solving the last equation first and then substituting the obtained values into the preceding equations.

Example 1.6.1 (Example of Back Row Reduction)

Consider the following upper triangular system:

$$\begin{cases} 2x_1 + 3x_2 + 4x_3 = 5, \\ 0x_1 + 6x_2 + 7x_3 = 8, \\ 0x_1 + 0x_2 + 9x_3 = 10. \end{cases}$$

We start with the last equation:

$$9x_3 = 10 \quad \Rightarrow \quad x_3 = \frac{10}{9}.$$

Next, we substitute x_3 into the second equation:

$$6x_2 + 7\left(\frac{10}{9}\right) = 8 \quad \Rightarrow \quad 6x_2 + \frac{70}{9} = 8 \quad \Rightarrow \quad 6x_2 = 8 - \frac{70}{9} \quad \Rightarrow \quad x_2 = \frac{2}{9}.$$

Finally, we substitute x_2 and x_3 into the first equation:

$$2x_1 + 3\left(\frac{2}{9}\right) + 4\left(\frac{10}{9}\right) = 5 \quad \Rightarrow \quad 2x_1 + \frac{6}{9} + \frac{40}{9} = 5 \quad \Rightarrow \quad 2x_1 = 5 - \frac{46}{9} \quad \Rightarrow \quad x_1 = \frac{1}{9}.$$

Thus, the solution is:

$$x_1 = \frac{1}{9}, \quad x_2 = \frac{2}{9}, \quad x_3 = \frac{10}{9}.$$

Algorithm 3: Back Row Reduction (Back Substitution)

Input: Upper triangular matrix A of size $n \times n$ and vector $\mathbf{b}$ of size n
Output: Solution vector $\mathbf{x}$ of size n

```
/* Initialize solution vector */
for i ← n to 1 do
    x_i ← b_i;
    for j ← i + 1 to n do
        x_i ← x_i − A_ij · x_j;
    end
    x_i ← x_i / A_ii;
end
return x;
```

Example 1.6.2 (Full SLE solving)

$$\begin{bmatrix} 1 & 1 & 2 & 2 \\ 1 & -1 & 0 & -2 \\ 2 & 0 & m & 0 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 1 & 2 & 2 \\ 0 & -2 & -2 & -4 \\ 0 & -2 & m-4 & -4 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 1 & 2 & 2 \\ 0 & -2 & -2 & -4 \\ 0 & 0 & m-2 & 0 \end{bmatrix}$$

given m=0

$$\begin{bmatrix} 1 & 1 & 2 & 2 \\ 0 & -2 & -2 & -4 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 1 & 2 & 2 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Chapter 2

Linear Combinations

2.1 Matrix–Vector Product as Linear Combination

$$Ax = \begin{bmatrix} 1 & -2 \\ 3 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ 3 \end{bmatrix} + x_2 \begin{bmatrix} -2 \\ 2 \end{bmatrix} = b.$$

Intuition: "Scaling each column of A by its corresponding entry in x."

In other words, to compute Ax, you take x_1 times the first column of A plus x_2 times the second column of A.

Example 2.1.1 (A Specific Choice of x)

If

$$A = \begin{bmatrix} 1 & -2 \\ 3 & 2 \end{bmatrix},$$

then the first column is

$$a_1 = \begin{bmatrix} 1 \\ 3 \end{bmatrix} \quad \text{and} \quad a_2 = \begin{bmatrix} -2 \\ 2 \end{bmatrix}.$$

A linear combination of these columns is

$$x_1 \begin{bmatrix} 1 \\ 3 \end{bmatrix} + x_2 \begin{bmatrix} -2 \\ 2 \end{bmatrix} = b.$$

As an example, if $x_1 = -\frac{1}{3}$ and $x_2 = 1$,

$$-\frac{1}{3} \begin{bmatrix} 1 \\ 3 \end{bmatrix} + \begin{bmatrix} -2 \\ 2 \end{bmatrix} = \begin{bmatrix} -\frac{1}{3} - 2 \\ -1 + 2 \end{bmatrix} = \begin{bmatrix} -\frac{7}{3} \\ 1 \end{bmatrix}.$$

In the notes, a similar combination yields $\begin{bmatrix} -2 \\ -2 \end{bmatrix}$. The key point is that any pair (x_1, x_2) gives a vector in $\mathbb{R}^2$.

Definition 2.1.1: Span

Span of a set of vectors is the collection of all linear combinations of those vectors. Concretely, for

$$a_1 = \begin{bmatrix} 1 \\ 3 \end{bmatrix}, \quad a_2 = \begin{bmatrix} -2 \\ 2 \end{bmatrix},$$

we write

$$\mathrm{Span}\{\, a_1, a_2\} = \Big\{ x_1 a_1 + x_2 a_2 \;\Big|\; x_1, x_2 \in \mathbb{R} \Big\}.$$

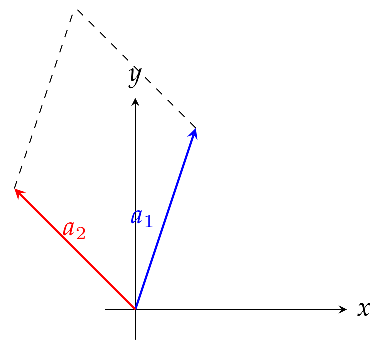
Any vector b not in this span means $Ax = b$ has no solution.

2.2 Linear Combinations and the Span

Vectors "outside" the span of $\{a_1, a_2\}$ are precisely those b for which the system $Ax = b$ is not consistent. Equivalently, they are not expressible as a linear combination of a_1 and a_2.

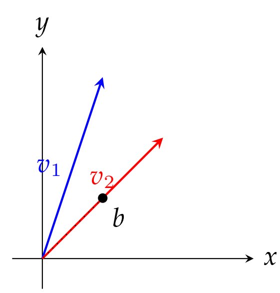
2.3 Geometric Plots in the Notes

1. *First Plot (showing the columns a_1 and a_2):*



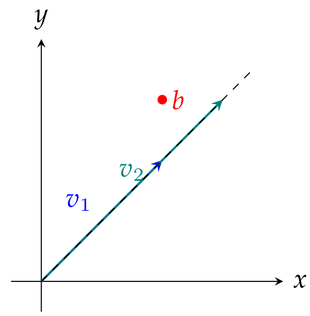
   Any linear combination $x_1 a_1 + x_2 a_2$ lands in the parallelogram structure (and its extensions) formed by these two vectors.

2. *Second Plot (showing a vector not in the span)*:



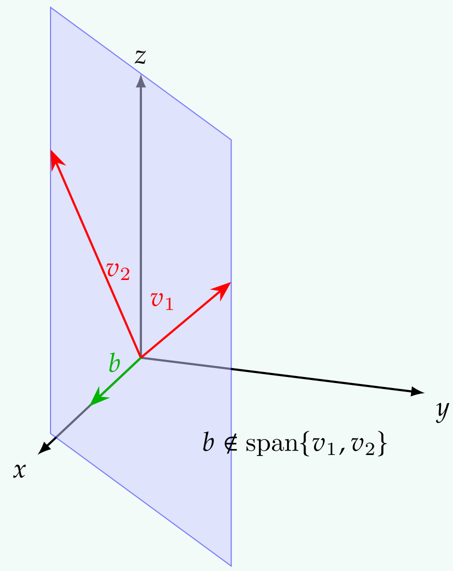
   If v_1 and v_2 do not cover $(1, 1)$ by any linear combination, then $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ is not in their span, hence no solution exists for $Ax = b$.

3. *Plot illustrating parallel vectors*:



Here, v_1 and v_2 are multiples of each other, so their span is just the single dashed line. The vector b off that line cannot be written as any combination of v_1 and v_2.

Example 2.3.1 ($\vec{b}$ not in span of $\vec{v_1}, \vec{v_2}$)



2.3.1 Key Conclusions

- Span$\{a_1, a_2\}$ is the set of all vectors b for which $Ax = b$ has a solution.
- If b is not in that span, there is no solution.
- Geometrically, if a_1 and a_2 are not multiples of each other, their span is a 2D plane through the origin in $\mathbb{R}^2$. If they are multiples, the span is just a single line, and most vectors in $\mathbb{R}^2$ lie outside that line (no solution).
- $span a_1$ is a line or point. $span a_1, a_2$ is a plane, line or point. and so on

2.3.2 Dropping a vector to retain the same span

Question 2: Drop vector a_3 to retain the same span as a_1, a_2 and a_3.

$$a_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

$$a_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$a_3 = \begin{bmatrix} 3 \\ 3 \end{bmatrix}$$

Solution: yes, $span a_1, a_2, a_3 = span a_1, a_2$ because a_3 is a linear combination of a_1 and a_2 $(a_1 + a_2 a_3)$

2.4 Evaluating Linear Dependence Procedurally

Example 2.4.1 (Finding linear dependence of a set of vectors.)

$$\vec{v}_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \vec{v}_2 = \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix}, \vec{v}_3 = \begin{bmatrix} 5 \\ 2 \\ -1 \end{bmatrix}$$

Since v_3 is a linear combination of v_1 and v_2, we can drop v_3 to retain the same span as v_1 and v_2.

$$span v_1, v_2, v_3 = span v_1, v_2$$

As

$$v_3 = 2v_1 - v_2$$

Note:-

Working towards a procedure for evaluating linear dependence of a set of vectors.

$$-v_1 + 2v_2 - v_3 = 0$$

$$\begin{bmatrix} \vec{v}_1 & \vec{v}_2 & \vec{v}_3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

We know that one solution is

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

Since this SLE has one particular solution as well as the trivial one, then it must have infinitely many solutions.

Then if $A\vec{x} = 0$ has infinitely many solutions, then the set of vectors (columns) in A is linearly dependent.

Conversely, if $A\vec{x} = 0$ has only one solution ($\vec{x} = \vec{0}$), then the set of vectors (columns) in A is linearly independent.

Definition 2.4.1: Linear Independence

We say a set of vectors $\{v_1, v_2, \cdots, v_n\}$ is *linearly independent* if the only solution to the SLE

$$\begin{bmatrix} \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

is the trivial solution, i.e. $\vec{x} = \vec{0}$.

$$A' = \begin{bmatrix} 1 & 3 & 5 & 0 \\ 2 & 2 & 2 & 0 \\ 3 & 1 & -1 & 0 \end{bmatrix} \xrightarrow{R_2=R_2-2R_1, \quad R_3=R_3-3R_1} \begin{bmatrix} 1 & 3 & 5 & 0 \\ 0 & -4 & -8 & 0 \\ 0 & -8 & -16 & 0 \end{bmatrix} \xrightarrow{R_3=R_3-2R_2} \begin{bmatrix} 1 & 3 & 5 & 0 \\ 0 & -4 & -8 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} = \text{REF}(A')$$

1 Free variable $\implies$ Linearly Independent

$$\text{REF}(A') \xrightarrow{R_2=R_2/(-4)} \begin{bmatrix} 1 & 3 & 5 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \xrightarrow{R_1=R_1-3R_2} \begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} = \text{RREF}(A') \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

Then

$$\begin{bmatrix} \vec{v}_1 & \vec{v}_2 & \vec{v}_3 \end{bmatrix} t_3 \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

And therefore the linear combination in its homogeneous form implies

$$t_3(\vec{v}_1 - 2\vec{v}_2 + \vec{v}_3) = \vec{0}$$

Definition 2.4.2: Identity Matrix

special Square matrix where the diagonal elements are 1 and the off-diagonal elements are 0.

$$I_n = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

Note that $I_n\vec{x} = \vec{x}$ and $I_n' = I_n$ and therefore is called a **symmetric matrix**