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 $x = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$, 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

$$\theta \;=\; \arccos\Bigl(\frac{u \cdot v}{\|u\|\|v\|}\Bigr).$$

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,

$$\theta \;=\; \arccos\Bigl(\frac{u \cdot v}{\|u\|\|v\|}\Bigr) \;=\; \arccos\Bigl(\frac{1}{\sqrt{3}\,\sqrt{3}}\Bigr) \;=\; \arccos\bigl(\tfrac{1}{3}\bigr).$$

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:*

$$\text{Plane} = \Bigl\{\, \alpha \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \;+\; \beta \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} \;\Big|\; \alpha, \beta \in \mathbb{R} \Bigr\}.$$

- *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:

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

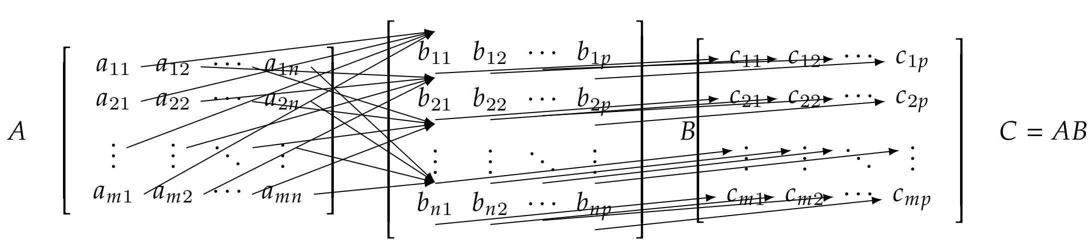


Figure 1.1: Matrix-Matrix Multiplication Diagram

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;
```

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

$y_1 \leftarrow x_1 + 2\,x_2$;

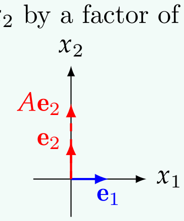
$y_2 \leftarrow x_2$;

return $\vec{v}_o$;

1.2.1 Examples of 2D Matrix Transformations

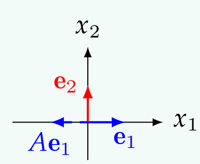
Example 1.2.1 ($A = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$)

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



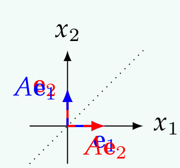
Example 1.2.2 ($A = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$)

Transformation: Reflection about the vertical axis.

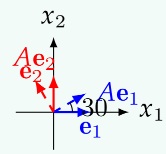


Example 1.2.3 ($A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$)

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

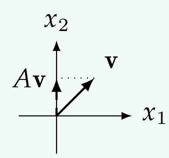


Example 1.2.4 ($A = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$)



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

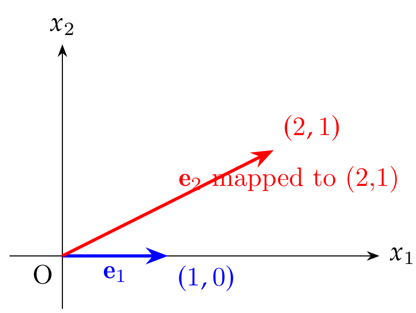
The matrix under consideration is

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

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

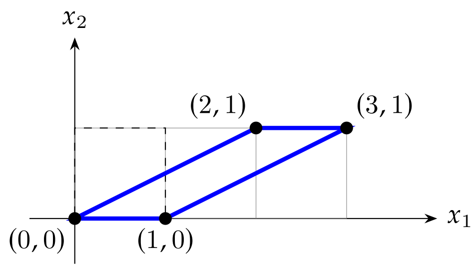
Consider the unit square with vertices

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

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.

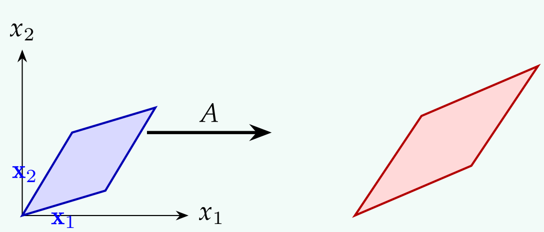
2. Parallel lines stay parallel. If two lines are parallel, their direction vectors are scalar multiples of each other. After applying A, the resulting direction vectors are $A(\mathbf{v})$ for each original direction $\mathbf{v}$. Since A is linear, any scalar multiples remain so, preserving parallelism.

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.