Contents

1. Linear Algebra

1.1. Matrices

The **determinant** is a function that associates a number to a square matrix. Given a $n \times n$ matrix A, its determinant, denoted as $\det(A)$ or $\mid A \mid$, is defined recursively as follows:

$$\det(A) = \begin{cases} \sum_{j=1}^{n} (-1)^{i+j} a_{i,j} \det(M_{i,j}) & \text{if } n > 1 \\ a_{11} & \text{otherwise} \end{cases}$$

Where j is any column of the matrix A chosen at random and $M_{i,j}$ is the matrix obtained by removing the i-th row and j-th column from A. The formula can also be applied with respect to rows instead of columns.
When the matrix has dimension $n = 2$, the formula can actually be simplified as follows:

$$\det(A) = (a_{1,1} \cdot a_{2,2}) - (a_{2,1} \cdot a_{1,2})$$

Exercise 1.1.1: Given the matrix $A = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 1 & 1 & 0 \end{pmatrix}$, compute its determinant.

Solution: The fastest way to compute a determinant is to pick the row/column that has the most zeros, because the number of $\det(M_{i,j})$ to compute is the smallest. In the case of A, the best choices are: the second row, the first column, the third row and the third column. Suppose the first column is chosen:

$$\det(A) = \sum_{i=1}^{3} (-1)^{i+j} a_{i,j} \det(M_{i,j}) =$$

$$(-1)^{1+1} a_{1,1} \det(M_{1,1}) + (-1)^{2+1} a_{2,1} \det(M_{2,1}) + (-1)^{3+1} a_{3,1} \det(M_{3,1}) =$$

$$1 \cdot \begin{vmatrix} 1 & 2 \\ 1 & 0 \end{vmatrix} - \cancel{0 \cdot \begin{vmatrix} 2 & 3 \\ 1 & 0 \end{vmatrix}} + 1 \cdot \begin{vmatrix} 2 & 3 \\ 1 & 2 \end{vmatrix} = (1 \cdot 0 - 1 \cdot 2) + (2 \cdot 2 - 3 \cdot 1) = 0 - 2 + 4 - 3 = -1$$

□

Theorem 1.1.1: A matrix is invertible if and only if its determinant is not zero.

Theorem 1.1.2: The determinant of a triangular matrix is equal to the product of the elements on its diagonal.

Proof: Consider an upper triangular matrix A and pick the first column to apply the formula:

$$\det(A) = \begin{vmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ 0 & a_{2,2} & \dots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_{n,n} \end{vmatrix} = a_{1,1} \begin{vmatrix} a_{2,2} & \dots & a_{2,n} \\ \vdots & \ddots & \vdots \\ 0 & \dots & a_{n,n} \end{vmatrix} = \dots = \prod_{i=1}^{n} a_{i,i}$$

The same is achieved for a lower triangular matrix by picking the first row. □

Theorem 1.1.3 (Binet's Theorem): The determinant is a multiplicative function. That is to say, given two matrices A and B:

$$\det(AB) = \det(A)\det(B)$$

Theorem 1.1.4: The determinant is invariant with respect to transposition.

Theorem 1.1.5: Given an invertible matrix A, the determinant of its inverse is the reciprocal of the determinant of A:

$$\det(A^{-1}) = \frac{1}{\det(A)}$$

The **trace** of a square matrix is defined as the sum of the elements on its diagonal:

$$\mathrm{tr}(A) = \sum_{i=1}^{n} a_{i,i}$$

1.2. Vector Spaces

Let V be a set, whose elements are called **vectors**. A vector $\underline{v}$ is denoted as $\underline{v} = (v_1, v_2, ..., v_n)$, where each v_i with $1 \leq i \leq n$ is called the i-th **component** of $\underline{v}$.
Let $+$ be an operation on such set, a *sum* of vectors, that has two vectors as arguments and returns another vector. That is, foreach $\left(\underline{x}, \underline{y}\right) \in V \times V$ there exists a vector $\underline{v} \in V$ such that $\underline{x} + \underline{y} = \underline{v}$.
Let $\cdot$ be another operation, a *product* between a vector and a real number, that has a real number and a vector as argument and returns another vector. That is, foreach $\lambda \in \mathbb{R}$ and $\underline{v} \in V$ there exists a vector $\underline{w} \in V$ such that $\lambda \cdot \underline{v} = \underline{w}$.
Suppose those operations possess the following properties:

- $(V, +)$ is an Abelian group;
- The product has the distributive property, such that for every $\lambda \in \mathbb{R}$ and for every $\underline{x}, \underline{y} \in V$ it is true that $\lambda \cdot \left(\underline{x} + \underline{y}\right) = \lambda \cdot \underline{x} + \lambda \cdot \underline{y}$;
- The product has the associative property, such that for every $\lambda, \mu \in \mathbb{R}$ and for every $\underline{x} \in V$ it is true that $(\lambda + \mu) \cdot \underline{x} = \lambda \cdot \underline{x} + \mu \cdot \underline{x}$;
- For every vector $\underline{v} \in V$, it is true that $1 \cdot \underline{v} = \underline{v}$.

If that is the case, the set V is called **vector space**. It should be noted that it does not matter what the elements of a vector space actually are (be they numbers, functions, polynomals, etcetera); as long as the aforementioned properties hold for the two operations, such set shares all of the properties that a vector space possesses.
For the sake of readability, the product between a real number and a vector is often represented without the dot. That is to say, the expressions $\lambda \cdot \underline{x}$ and $\lambda\underline{x}$ have the same meaning.

Exercise 1.2.1: Denote as $\mathbb{R}^n$ the set containing all vectors of real components[1] in the n-dimensional plane. Prove that $\mathbb{R}^n$ constitutes a vector space.

Solution: It is possible to define both a sum between two vectors in the n-dimensional plane and a product between a vector in the n-dimensional space and a real number. To sum two vectors in the n-dimensional space, it suffices to sum each component with each component. To multiply a vector in the n-dimensional space with a real number it suffices to multiply each component by that number:

$$\begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} + \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix} = \begin{pmatrix} v_1 + w_1 \\ v_2 + w_2 \\ \vdots \\ v_n + w_n \end{pmatrix} \qquad \lambda \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} \lambda v_1 \\ \lambda v_2 \\ \vdots \\ \lambda v_n \end{pmatrix}$$

[1]This is a misnomer.

Both operations obey the properties stated:

- $(\mathbb{R}^n, +)$ constitutes an Abelian group. Infact:
  - The sum has the associative property:

$$\begin{pmatrix} v_1 + w_1 \\ v_2 + w_2 \\ \vdots \\ v_n + w_n \end{pmatrix} + \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} = \begin{pmatrix} v_1 + w_1 + u_1 \\ v_2 + w_2 + u_2 \\ \vdots \\ v_n + w_n + u_n \end{pmatrix} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} + \begin{pmatrix} w_1 + u_1 \\ w_2 + u_2 \\ \vdots \\ w_n + u_n \end{pmatrix}$$

  - There exists an identity element, in the form of the vector whose components are all zero:

$$\begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} v_1 + 0 \\ v_2 + 0 \\ \vdots \\ v_n + 0 \end{pmatrix} = \begin{pmatrix} 0 + v_1 \\ 0 + v_2 \\ \vdots \\ 0 + v_n \end{pmatrix} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}$$

  - Each vector in the n-dimensional space has an inverse element with respect to the sum, that is the same vector multiplied by -1:

$$\begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} + (-1)\begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} + \begin{pmatrix} -v_1 \\ -v_2 \\ \vdots \\ -v_n \end{pmatrix} = \begin{pmatrix} v_1 + (-v_1) \\ v_2 + (-v_2) \\ \vdots \\ v_n + (-v_n) \end{pmatrix} = \begin{pmatrix} -v_1 + v_1 \\ -v_2 + v_2 \\ \vdots \\ -v_n + v_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

  - The sum has the commutative property:

$$\begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} + \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix} = \begin{pmatrix} v_1 + w_1 \\ v_2 + w_2 \\ \vdots \\ v_n + w_n \end{pmatrix} = \begin{pmatrix} w_1 + v_1 \\ w_2 + v_2 \\ \vdots \\ w_n + v_n \end{pmatrix} = \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix} + \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}$$

- The product has the associative property:

$$(\lambda + \mu)\begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} (\lambda + \mu)v_1 \\ (\lambda + \mu)v_2 \\ \vdots \\ (\lambda + \mu)v_n \end{pmatrix} = \begin{pmatrix} \lambda v_1 + \mu v_1 \\ \lambda v_2 + \mu v_2 \\ \vdots \\ \lambda v_n + \mu v_n \end{pmatrix} = \begin{pmatrix} \lambda v_1 \\ \lambda v_2 \\ \vdots \\ \lambda v_n \end{pmatrix} + \begin{pmatrix} \mu v_1 \\ \mu v_2 \\ \vdots \\ \mu v_n \end{pmatrix} = \lambda\begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} + \mu\begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}$$

- The product has the distributive property:

$$\lambda\left(\begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} + \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix}\right) = \lambda\begin{pmatrix} v_1 + w_1 \\ v_2 + w_2 \\ \vdots \\ v_n + w_n \end{pmatrix} = \begin{pmatrix} \lambda(v_1 + w_1) \\ \lambda(v_2 + w_2) \\ \vdots \\ \lambda(v_n + w_n) \end{pmatrix} = \begin{pmatrix} \lambda v_1 + \lambda w_1 \\ \lambda v_2 + \lambda w_2 \\ \vdots \\ \lambda v_n + \lambda w_n \end{pmatrix} = \lambda\begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} + \lambda\begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix}$$

- Multiplying a vector by the number 1 leaves the vector unchanged:

$$1\begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} 1v_1 \\ 1v_2 \\ \vdots \\ 1v_n \end{pmatrix} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}$$

□

Exercise 1.2.2: Denote as $\mathbb{P}_n$ the set containing all polynomials with real coefficients and degree less than or equal to n. Prove that $\mathbb{P}_n$ constitutes a vector space.

Solution: It is possible to define both a sum between two polynomials with real coefficients and degree $\leq n$ and a product between a polynomial with real coefficients and degree $\leq n$ and a real number. To sum two such polynomials it suffices to sum the coefficients of their monomials having the same degree:

$$\left(a_n x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0\right) + \left(b_m x^m + b_{m-1}x^{m-1} + \ldots + b_1 x + b_0\right) =$$
$$a_n x^n + a_{n-1}x^{n-1} + \ldots + (a_m + b_m)x^m + (a_{m-1} + b_{m-1})x^{m-1} + \ldots + (a_1 + b_1)x + (a_0 + b_0)$$

To multiply a polynomial with real coefficients and degree $\leq n$ with a real number it suffices to multiply each coefficient of its monomials by such number:

$$\lambda\left(a_n x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0\right) = (\lambda a_n)x^n + (\lambda a_{n-1})x^{n-1} + \ldots + (\lambda a_1)x + (\lambda a_0)$$

Both operations satisfy the properties required.

□

Given a vector space V, a set W is said to be a **subspace** of V if it's a subset of V and it's itself a vector space with respect to the same operations defined for V.

Theorem 1.2.1: Let V be a vector space. To prove that a set W is a subspace of V it suffices to prove that it is a subset of V and is algebraically closed with respect to the same operations defined for V.

Exercise 1.2.3: Consider the vector space $\mathbb{R}^3$. Prove that the set W_1 is a subspace of $\mathbb{R}^3$ while W_2 is not.

$$W_1 = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} : \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbb{R}^3, x_1 + x_2 = 0 \right\} \qquad W_2 = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} : \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbb{R}^3, x_2 = 2x_3 + 1 \right\}$$

Solution: The first set is a subspace of $\mathbb{R}^3$ because it is a subset of V and is algebraically closed:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} x_1 \\ -x_1 \\ x_3 \end{pmatrix} + \begin{pmatrix} y_1 \\ -y_1 \\ y_3 \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ -x_1 - y_1 \\ x_3 + y_3 \end{pmatrix} \Rightarrow x_2 + y_2 = -x_1 - y_1 \Rightarrow x_2 + y_2 + (x_1 + y_1) = 0$$

$$\lambda \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \lambda \begin{pmatrix} x_1 \\ -x_1 \\ x_3 \end{pmatrix} = \begin{pmatrix} \lambda x_1 \\ -\lambda x_1 \\ \lambda x_3 \end{pmatrix} \Rightarrow \lambda x_2 = -\lambda x_1 \Rightarrow \lambda(x_1 + x_2) = 0$$

The second one, on the other hand, is not:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} x_1 \\ 2x_3 + 1 \\ x_3 \end{pmatrix} + \begin{pmatrix} y_1 \\ 2y_3 + 1 \\ y_3 \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ 2x_3 + 2y_3 + 2 \\ x_3 + y_3 \end{pmatrix} \Rightarrow 2x_3 + 2y_3 + 2 \neq 2(x_3 + y_3) + 1$$

□

Theorem 1.2.2: Let V be a vector space. The sets $\{\underline{0}\}$ and V are always subspaces of V.

1.3. Bases and Dimension

Let $\left\{\underline{v_1}, \underline{v_2}, \ldots, \underline{v_n}\right\}$ be a set of n vectors of a vector space V, and let $\lambda_1, \lambda_2, \ldots, \lambda_n$ be n real numbers (not necessarely distinct). Every summation defined as such:

$$\sum_{i=1}^{n} \lambda_i \underline{v_i} = \lambda_1 \underline{v_1} + \lambda_2 \underline{v_2} + \ldots + \lambda_i \underline{v_i} + \ldots + \lambda_n \underline{v_n}$$

Is called a **linear combination** of the vectors $\{\underline{v_1}, \underline{v_2}, \ldots, \underline{v_n}\}$, with $\lambda_1, \lambda_2, \ldots, \lambda_n$ as **coefficients**.
A set of vectors $\{\underline{v_1}, \underline{v_2}, \ldots, \underline{v_n}\}$ is said to be **linearly independent** if the only linear combination of such vectors that equals the null vector is the one that has 0 as every coefficient. If there exists a linear combination of such vectors that is equal to the null vector and that has at least a non-zero coefficient, those vectors are said to be **linearly independent**.

Theorem 1.3.1: Let $\{\underline{v_1}, \underline{v_2}, \ldots, \underline{v_n}\}$ be a set of n vectors of a vector space V. If those vectors are linearly dependent, there exists at least one vector of such set that can be expressed as a linear combination of the remaining vectors, and vice versa.

Proof: If such set of vectors is linearly dependent, there must exist a linear combination of the set that equals the null vector, be it:

$$\sum_{i=1}^{n} \lambda_i \underline{v_i} = \lambda_1 \underline{v_1} + \lambda_2 \underline{v_2} + \ldots + \lambda_j \underline{v_j} + \ldots + \lambda_n \underline{v_n} = \underline{0}$$

Where there's at least a non-zero coefficient. Let's assume, without loss of generality, that the j-th coefficient is non-zero (whether such coefficient is the only one to be non-zero is irrelevant). The product of such coefficient with the corresponding vector can be moved to the right side of the equation like so:

$$\lambda_1 \underline{v_1} + \lambda_2 \underline{v_2} + \ldots + \lambda_n \underline{v_n} = -\lambda_j \underline{v_j}$$

Dividing both sides by $-\lambda_j$ gives:

$$-\frac{\lambda_1}{\lambda_j} \underline{v_1} - \frac{\lambda_2}{\lambda_j} \underline{v_2} - \ldots - \frac{\lambda_n}{\lambda_j} \underline{v_n} = \underline{v_j}$$

Each $-\frac{\lambda_i}{\lambda_j}$ is itself a real number, and therefore the expression above is a linear combination that is equal to a vector of the set.
On the other hand, assume that the j-th vector of the set is equal to a linear combination of the remaining vectors like so:

$$\lambda_1 \underline{v_1} + \lambda_2 \underline{v_2} + \ldots + \lambda_n \underline{v_n} = \underline{v_j}$$

Moving $\underline{v_j}$ to the left gives:

$$\lambda_1 \underline{v_1} + \lambda_2 \underline{v_2} + \ldots + (-1) \underline{v_j} + \ldots + \lambda_n \underline{v_n} = \underline{0}$$

Since -1 is a real number, the expression on the left side of the equation is indeed a linear combination of the whole set, that is equal to the null vector. □

Exercise 1.3.1: Consider the vector space $\mathbb{R}^2$. Check if the vectors $\underline{v_1} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\underline{v_2} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ are linearly independent or linearly dependent.

Solution: Consider such linear combination:

$$\lambda_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

To test whether $\underline{v_1}$ and $\underline{v_2}$ are linearly dependent or independent, it suffices to solve this system of equations:

$$\lambda_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Rightarrow \begin{pmatrix} \lambda_1 \\ 0 \end{pmatrix} + \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Rightarrow \begin{pmatrix} 2\lambda_1 \\ \lambda_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Rightarrow \begin{cases} 2\lambda_1 = 0 \\ \lambda_2 = 0 \end{cases} \Rightarrow \begin{cases} \lambda_1 = 0 \\ \lambda_2 = 0 \end{cases}$$

Since the only solution to such system is $\lambda_1 = 0, \lambda_2 = 0$, $\underline{v_1}$ and $\underline{v_2}$ are linearly independent. □

Exercise 1.3.2: Consider the vector space $\mathbb{R}^2$. Check if the vectors $\underline{v_1} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $\underline{v_2} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and $\underline{v_3} = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$ are linearly independent or linearly dependent.

Solution: Consider such linear combination:

$$\lambda_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \lambda_3 \begin{pmatrix} 2 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

To test whether $\underline{v_1}$, $\underline{v_2}$ and $\underline{v_3}$ are linearly dependent or independent, it suffices to solve this system of equations:

$$\lambda_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \lambda_3 \begin{pmatrix} 2 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Rightarrow \begin{cases} \lambda_1 + \lambda_2 + 2\lambda_3 = 0 \\ \lambda_2 = 0 \end{cases} \Rightarrow \begin{cases} \lambda_1 = -2\lambda_3 \\ \lambda_2 = 0 \end{cases}$$

Since there are infinite solutions to such system, including non-zero solutions, $\underline{v_1}$, $\underline{v_2}$ and $\underline{v_3}$ are linearly dependent. For example, setting $\lambda_1 = 1$ results in:

$$1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + 0 \begin{pmatrix} 1 \\ 1 \end{pmatrix} - \frac{1}{2} \begin{pmatrix} 2 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Which is, indeed, a correct identity. □

A set of vectors $S = \left\{\underline{s_1}, ..., \underline{s_n}\right\}$ of a vector space V is said to **generate** V if every vector of V can be written as a linear combination of the vectors in S. That is to say, S generates V if for every $\underline{v} \in V$ there exist a set of coefficients $\lambda_1, ..., \lambda_n$ such that:

$$\underline{v} = \sum_{i=1}^{n} \lambda_i \underline{s_i} = \lambda_1 \underline{s_1} + ... + \lambda_n \underline{s_n}$$

Theorem 1.3.2: Let S be a set of vectors of a vector space V that can generate V. Let $\underline{w} \in V$ be a random vector of V. The set of vectors $S \cup \{\underline{w}\}$ is linearly dependent.

Proof: If S can generate V and $\underline{w}$ belongs to V, there exists a linear combination of the vectors in S such that:

$$\underline{w} = \sum_{i=1}^{n} \lambda_i \underline{s_i} = \lambda_1 \underline{s_1} + ... + \lambda_n \underline{s_n}$$

Moving $\underline{w}$ to the right side of the equation gives:

$$\underline{0} = (-1)\underline{w} + \lambda_1 \underline{s_1} + ... + \lambda_n \underline{s_n}$$

The expression on the right side of the equation is indeed a linear combination of $S \cup \{\underline{w}\}$, that is equal to the null vector. Since at least -1 is a non zero coefficient, such set is linearly dependent. □

A set of vectors that can generate a vector space and is itself linearly independent is called a **basis** for such vector space. The cardinality of a basis is called the **dimension** of the corresponding vector space. If a vector space contains just the null vector, such vector space is said to have dimension 0.

Theorem 1.3.3: A basis of a vector space has the minimum cardinality out of every set of vectors that can generate it. In other words, if a basis of a vector space has cardinality n, at least n vectors are needed to generate such space.

Exercise 1.3.3: Consider the vector space $\mathbb{P}_2$. Knowing that the sets $\mathcal{B}_1 = \{1, x, x^2\}$ and $\mathcal{B}_2 = \{(x+1), (x-1), x^2\}$ are both bases for $\mathbb{P}_2$, write the polynomial $p(x) = 3x^2 + 2x - 5$ as a linear combination of each.

Solution: It's trivial to see that, for the first basis, such linear combination is $p(x)$ itself:

$$\lambda_0 1 + \lambda_1 x + \lambda_2 x^2 = -5 + 2x + 3x^2 \Rightarrow \begin{cases} \lambda_0 = -5 \\ \lambda_1 = 2 \\ \lambda_2 = 3 \end{cases}$$

Regarding the second basis, it can be rewritten as:

$$\lambda_0(x+1) + \lambda_1(x-1) + \lambda_2 x^2 \Rightarrow \lambda_0 x + \lambda_0 + \lambda_1 x - \lambda_1 + \lambda_2 x^2 \Rightarrow (\lambda_0 - \lambda_1) + (\lambda_0 + \lambda_1)x + \lambda_2 x^2$$

Equating it term by term:

$$(\lambda_0 - \lambda_1) + (\lambda_0 + \lambda_1)x + \lambda_2 x^2 = -5 + 2x + 3x^2 \Rightarrow \begin{cases} \lambda_0 - \lambda_1 = -5 \\ \lambda_0 + \lambda_1 = 2 \\ \lambda_2 = 3 \end{cases} \Rightarrow \begin{cases} \lambda_0 = -\frac{3}{2} \\ \lambda_1 = \frac{7}{2} \\ \lambda_2 = 3 \end{cases}$$

Therefore:

$$3x^2 + 2x - 5 = -\frac{3}{2}(x+1) + \frac{7}{2}(x-1) + 3x^2$$

□

The basis of a vector space that renders calculations the most "comfortable" is called the **canonical basis** for such vector space. Such basis is different from vector space to vector space.

Exercise 1.3.4: Determine the dimension of $\mathbb{R}^n$

Solution: Consider any n-dimensional vector of coordinates $a_1, a_2, ..., a_n$. It's easy to see that such vector is equal to the following linear combination:

$$\begin{pmatrix} a_1 \\ a_2 \\ \dots \\ a_n \end{pmatrix} = a_1 \begin{pmatrix} 1 \\ 0 \\ \dots \\ 0 \end{pmatrix} + a_2 \begin{pmatrix} 0 \\ 1 \\ \dots \\ 0 \end{pmatrix} + ... + a_n \begin{pmatrix} 0 \\ 0 \\ \dots \\ 1 \end{pmatrix}$$

A set containing such vectors in linearly independent. Infact:

$$\lambda_1 \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} + ... + \lambda_n \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \Rightarrow \begin{cases} \lambda_1 + 0 + ... + 0 = 0 \\ 0 + \lambda_2 + ... + 0 = 0 \\ \vdots \\ 0 + 0 + ... + \lambda_n = 0 \end{cases} \Rightarrow \begin{cases} \lambda_1 = 0 \\ \lambda_2 = 0 \\ \vdots \\ \lambda_n = 0 \end{cases}$$

This set of vectors is linearly independent and can generate $\mathbb{R}^n$, therefore it's a basis for $\mathbb{R}^n$. The dimension of $\mathbb{R}^n$ is then n, since such set has cardinality n. In particular, this specific basis is the canonical basis for $\mathbb{R}^n$. $\square$

Exercise 1.3.5: Determine the dimension of $\mathbb{P}_n$

Solution: Consider any polynomial of degree at most n having real coefficients:

$$a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0$$

Each monomial of such polynomial are itself polynomials of degree at most n having real coefficients. Therefore, the polynomial itself can be seen as a linear combination of the polynomials $\{x^n, x^{n-1}, \dots, x^1, x^0\}$ with coefficients $a_n, a_{n-1}, \dots, a_1, a_0$.
Such set of vectors is linearly independent. Infact:
This set of vectors is linearly independent and can generate $\mathbb{P}_n$, therefore it's a basis for $\mathbb{P}_n$. The dimension of $\mathbb{P}_n$ is then $n+1$, since such set has cardinality $n+1$. In particular, this specific basis is the canonical basis for $\mathbb{P}_n$. $\square$

Consider a vector space V and two bases $\mathcal{B} = \{\underline{b_1}, \underline{b_2}, \dots, \underline{b_n}\}$ and $\mathcal{B}' = \{\underline{b'_1}, \underline{b'_2}, \dots, \underline{b'_n}\}$. A vector $\underline{x} \in V$ can be represented with respect to both bases, non necessarely equivalent composant by composant:

$$\underline{x} \Leftrightarrow \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}_{\mathcal{B}} \qquad \underline{x} \Leftrightarrow \begin{pmatrix} x'_1 \\ x'_2 \\ \vdots \\ x'_n \end{pmatrix}_{\mathcal{B}'}$$

Being both bases constituted by vectors of the same vector space, it is possible to express the elements of $\mathcal{B}'$ as linear combinations of the elements of $\mathcal{B}$:

$$\begin{cases} \underline{b'_1} = p_{1,1}\underline{b_1} + p_{1,2}\underline{b_2} + \dots + p_{1,n}\underline{b_n} = \sum_{j=1}^n p_{1,j}\underline{b_j} \\ \underline{b'_2} = p_{2,1}\underline{b_1} + p_{2,2}\underline{b_2} + \dots + p_{2,n}\underline{b_n} = \sum_{j=1}^n p_{2,j}\underline{b_j} \\ \vdots \\ \underline{b'_n} = p_{n,1}\underline{b_1} + p_{n,2}\underline{b_2} + \dots + p_{n,n}\underline{b_n} = \sum_{j=1}^n p_{n,j}\underline{b_j} \end{cases}$$

Therefore:

$$\underline{x} = \sum_{j=1}^n x'_j \underline{b'_j} = \sum_{j=1}^n x'_j \sum_{i=1}^n p_{j,i}\underline{b_i} = \sum_{i=1}^n x_i \underline{b_i}$$

By comparing the third and fourth members of the equality term by term:

$$x_i = \sum_{j=1}^n p_{i,j} x'_j \Rightarrow \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} p_{1,1} & p_{1,2} & \cdots & p_{1,n} \\ p_{2,1} & p_{2,2} & \cdots & p_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ p_{n,1} & p_{n,2} & \cdots & p_{n,n} \end{pmatrix} \begin{pmatrix} x'_1 \\ x'_2 \\ \vdots \\ x'_n \end{pmatrix} = P \begin{pmatrix} x'_1 \\ x'_2 \\ \vdots \\ x'_n \end{pmatrix}$$

Of course, it is also possible to go the other way around, expressing the elements of $\mathcal{B}$ as linear combinations of the elements of $\mathcal{B}'$:

$$\begin{cases} \underline{b_1} = q_{1,1}\underline{b'_1} + q_{1,2}\underline{b'_2} + \dots + q_{1,n}\underline{b'_n} = \sum_{j=1}^n q_{1,j}\underline{b'_j} \\ \underline{b_2} = q_{2,1}\underline{b'_1} + q_{2,2}\underline{b'_2} + \dots + q_{2,n}\underline{b'_n} = \sum_{j=1}^n q_{2,j}\underline{b'_j} \\ \vdots \\ \underline{b_n} = q_{n,1}\underline{b'_1} + q_{n,2}\underline{b'_2} + \dots + q_{n,n}\underline{b'_n} = \sum_{j=1}^n q_{n,j}\underline{b'_j} \end{cases}$$

Therefore:

$$\underline{x} = \sum_{j=1}^{n} x_j \underline{b_j} = \sum_{j=1}^{n} x_j \sum_{i=1}^{n} q_{j,i} \underline{b'_i} = \sum_{i=1}^{n} x'_i \underline{b'_i}$$

By comparing the third and fourth members of the equality term by term:

$$x'_i = \sum_{j=1}^{n} q_{i,j} x_j \Rightarrow \begin{pmatrix} x'_1 \\ x'_2 \\ \vdots \\ x'_n \end{pmatrix} = \begin{pmatrix} q_{1,1} & q_{1,2} & \cdots & q_{1,n} \\ q_{2,1} & q_{2,2} & \cdots & q_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ q_{n,1} & q_{n,2} & \cdots & q_{n,n} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = Q \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

This means that to change the representation of a vector with respect to a given basis to the representation with respect to a different basis it suffices to perform a matrix multiplication. But an even more interesting result can be obtained by substituting the one in the expression for the other:

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = P \begin{pmatrix} x'_1 \\ x'_2 \\ \vdots \\ x'_n \end{pmatrix} = PQ \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \qquad \begin{pmatrix} x'_1 \\ x'_2 \\ \vdots \\ x'_n \end{pmatrix} = Q \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = QP \begin{pmatrix} x'_1 \\ x'_2 \\ \vdots \\ x'_n \end{pmatrix}$$

Since the two matrices on the side of the equalities are the same, for these equalities to hold both matrix products PQ and QP must be equal to the identity matrix. In other words, P and Q are the inverse of the other.

1.4. Linear Transformations

A transformation $\phi : V \mapsto W$, with both V and W being vector spaces, is called a **linear transformation** if and only if:

$$\phi\left(\underline{v_1} + \underline{v_2}\right) = \phi\left(\underline{v_1}\right) + \phi\left(\underline{v_2}\right) \ \forall \underline{v_1}, \underline{v_2} \in V \qquad \phi(\lambda \underline{v}) = \lambda \phi(\underline{v}) \ \forall \underline{v} \in V, \lambda \in \mathbb{R}$$

In particular, if $V = W$, the transformation ϕ is said to be an **endomorphism**.

Exercise 1.4.1: Consider the vector space $\mathbb{R}$ (that is, the set of real numbers). Check whether the transformations $\phi_1(x) = 2x$ and $\phi_2(x) = x + 1$ are linear or not.

Solution:

- The transformation $\phi_1(x) = 2x$ is linear. Infact, given two real numbers a and b, is indeed true that $2(a + b) = 2a + 2b$, since the product between real numbers has the distributive property. Similarly, given a real number a and a real number λ, it is true that $2(\lambda a) = 2\lambda a$, since the product between real numbers has the associative property;
- The transformation $\phi_2(x) = x + 1$ is not linear. Given two real numbers a and b, it results in $\phi_2(a + b) = (a + b) + 1 = a + b + 1$, while $\phi_2(a) + \phi_2(b) = a + 1 + b + 1 = a + b + 2$.

□

It can be shown that a linear transformation is equivalent to a manipulation of matrices.
Let $\phi : V \mapsto W$ be a linear transformation between two vector space V and W. Let $B = \left\{\underline{b_1}, ..., \underline{b_n}\right\}$ be a basis for V and $C = \left\{\underline{c_1}, ..., \underline{c_m}\right\}$ a basis for W. Each vector $\underline{x} \in V$ can be written as a linear combination of the vectors of B:

$$\underline{x} = \sum_{i=1}^{n} \lambda_i \underline{b_i}$$

Applying ϕ to $\underline{x}$ gives:

$$\phi(\underline{x}) = \phi\left(\sum_{i=1}^{n} \lambda_i \underline{b_i}\right) = \sum_{i=1}^{n} \phi\left(\lambda_i \underline{b_i}\right) = \sum_{i=1}^{n} \lambda_i \phi\left(\underline{b_i}\right)$$

The two rightmost equalities stem from the fact that ϕ is linear.
Each $\phi(b_i)$ is a vector of W, since it's the result of applying ϕ to an vector of V. This means that each $\phi(b_i)$ can itself be written as a linear combination of elements of C:

$$\phi(b_i) = \sum_{j=1}^{m} \gamma_{j,i} c_j$$

Substituting it back in the previous expression gives:

$$\phi(\underline{x}) = \sum_{i=1}^{n} \lambda_i \left(\sum_{j=1}^{m} \gamma_{j,i} c_j\right) = \sum_{i,j=1}^{n,m} \lambda_i \gamma_{j,i} c_j$$

This means that, fixed a given basis B, to know all the relevant information regarding a vector $\underline{x}$ of V it suffices to "store" the λ coefficients of its linear combination with respect to B in a (column) vector.
In a similar fashion, to know all the relevant information of its image $\phi(\underline{x})$ it suffices to store the $\sum_{j=1}^{m} \lambda_i \gamma_{j,i}$ coefficients of its linear combination with respect to C in a (column) vector:

$$\underline{x} \Longleftrightarrow \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} \qquad\qquad \phi(\underline{x}) \Longleftrightarrow \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix}$$

Where, for clarity, each $\sum_{j=1}^{m} \lambda_i \gamma_{j,i}$ has been written simply as μ_i.
It is then possible to describe the application of the transformation ϕ as the following product of matrices:

$$\begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix} = \begin{pmatrix} \gamma_{1,1} & \gamma_{1,2} & \cdots & \gamma_{1,n} \\ \gamma_{2,1} & \gamma_{2,2} & \cdots & \gamma_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_{m,1} & \gamma_{m,2} & \cdots & \gamma_{m,n} \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}$$

Exercise 1.4.2: Consider the linear application $T : \mathbb{R}^3 \mapsto \mathbb{R}^2$ defined below. Express it as a matrix multiplication with respect to the two bases $\mathcal{B}_1$ and $\mathcal{B}_2$.

$$T\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x - z \\ y + z \end{pmatrix} \qquad \mathcal{B}_1 = \left\{ \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ -1 \end{pmatrix}, \begin{pmatrix} -2 \\ 1 \\ 4 \end{pmatrix} \right\} \qquad \mathcal{B}_2 = \left\{ \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \end{pmatrix} \right\}$$

Solution: The first step is to express the vectors of $\mathcal{B}_1$ for $\mathbb{R}^3$ evaluated in T as linear combinations of the vectors of $\mathcal{B}_2$ for $\mathbb{R}^2$:

$$T\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \lambda_{1,1}\begin{pmatrix} 1 \\ 2 \end{pmatrix} + \lambda_{2,1}\begin{pmatrix} 2 \\ 1 \end{pmatrix} \quad T\begin{pmatrix} 2 \\ 1 \\ -1 \end{pmatrix} = \lambda_{1,2}\begin{pmatrix} 1 \\ 2 \end{pmatrix} + \lambda_{2,2}\begin{pmatrix} 2 \\ 1 \end{pmatrix} \quad T\begin{pmatrix} -2 \\ 1 \\ 4 \end{pmatrix} = \lambda_{1,3}\begin{pmatrix} 1 \\ 2 \end{pmatrix} + \lambda_{2,3}\begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

□

Let $T : V \mapsto W$ a linear transformation between vector spaces V and W. The set of all vectors of W that have a correspondant in V through T is called the **image** of the transformation T, and is denoted as $\mathfrak{I}(T)$. It may or may not coincide with W.

$$\mathfrak{I}(T) = \{\underline{w} \in W : \exists \underline{v} \in V \quad \text{s.t.} \;\; T(\underline{v}) = \underline{w}\}$$

The notion of image is present in every transformations, not just linear ones, but images of linear transformations possess properties that images of generic transformations don't.

Theorem 1.4.1: Let $T : V \mapsto W$ be a linear transformation between vector spaces V and W. $\mathfrak{I}(W)$ is a subspace of W.

Proof: By Theorem 1.2.1, it suffices to prove that $\underline{w_1} + \underline{w_2} \in \mathfrak{I}(W)$ holds for all $\underline{w_1}, \underline{w_2} \in \mathfrak{I}(W)$ and that $\lambda\underline{w} \in \mathfrak{I}(W)$ holds for all $\underline{w} \in \mathfrak{I}(W)$ and $\lambda \in \mathbb{R}$.
By definition, if $\underline{w} \in \mathfrak{I}(W)$ then there exists $\underline{v} \in V$ such that $T(\underline{v}) = \underline{w}$. Therefore:

$$\underline{w_1} + \underline{w_2} = T\left(\underline{v_1}\right) + T\left(\underline{v_2}\right) \qquad \lambda\underline{w} = \lambda T(\underline{v})$$

By virtue of T being linear:

$$\underline{w_1} + \underline{w_2} = T\left(\underline{v_1}\right) + T\left(\underline{v_2}\right) = T\left(\underline{v_1} + \underline{v_2}\right) \qquad \lambda\underline{w} = \lambda T(\underline{v}) = T(\lambda\underline{v})$$

In both cases, there exists a vector in V such that the application of T gives such vector, therefore $\mathfrak{I}(W)$ is algebraically closed with respect to the operations defined for W. $\square$

Let $T : V \mapsto W$ a linear transformation between vector spaces V and W. The set of all vectors of V such that the application of T to those vectors gives the null vector (of W) is called the **kernel** of T, and is denoted as $\ker(T)$.

$$\ker(T) = \{\underline{v} \in V : T(\underline{v}) = \underline{0}\}$$

Theorem 1.4.2: Let $T : V \mapsto W$ be a linear transformation between vector spaces V and W. $\ker(V)$ is a subspace of V.

Proof: By Theorem 1.2.1, it suffices to prove that $\underline{v_1} + \underline{v_2} \in \ker(V)$ holds for all $\underline{v_1}, \underline{v_2} \in \ker(V)$ and that $\lambda\underline{v} \in \ker(V)$ holds for all $\underline{v} \in \ker(V)$ and $\lambda \in \mathbb{R}$.
By definition, if $\underline{v} \in \ker(V)$ holds, then $T(\underline{v}) = 0$. By virtue of T being linear:

$$T\left(\underline{v_1} + \underline{v_2}\right) = T\left(\underline{v_1}\right) + T\left(\underline{v_2}\right) = \underline{0} + \underline{0} = \underline{0} \qquad T(\lambda\underline{v}) = \lambda T(\underline{v}) = \lambda(\underline{0}) = \underline{0}$$

$\square$

Let $T : V \mapsto W$ be a linear transformation between vector spaces V and W. The dimension of the image of T is called the **rank** of T, and denoted as $\operatorname{rank}(T)$, while the dimension of the kernel of T is called the **nullity** of T, and denoted as $\operatorname{null}(T)$.

Theorem 1.4.3 (Rank-nullity theorem): Let $T : V \mapsto W$ be a linear transformation between vector spaces V and W. The dimension of V is given by the sum of the rank of T and the nullity of T:

$$\dim(V) = \operatorname{rank}(T) + \operatorname{null}(T) = \dim(\ker(T)) + \dim(\mathfrak{I}(T))$$

Let $T : V \mapsto W$ be a linear transformation between vector spaces V and W. The linear transformation $T^{-1} : W \mapsto V$ is said to be the **inverse** of T if:

$$T^{-1}(T(\underline{v})) = T(T^{-1}(\underline{v})) = \underline{v}, \ \forall \underline{v} \in V$$

As for any function, a linear transformation T has an inverse if and only if it is both injective and subjective. A linear transformation that has an inverse is said to be **invertible**.

Theorem 1.4.4: Let $T : V \mapsto W$ be a linear transformation. If T is injective, then its nullity is 0.

Proof: If T is injective then, for any distinct $\underline{v_1}, \underline{v_2} \in V$, $T\left(\underline{v_1}\right) \neq T\left(\underline{v_2}\right)$, which is to say $T\left(\underline{v_1}\right) - T\left(\underline{v_2}\right) \neq \underline{0}$. But T is linear by definition, therefore $T\left(\underline{v_1}\right) - T\left(\underline{v_2}\right) = T\left(\underline{v_1} - \underline{v_2}\right)$. Being V a vector space, it algebraically closed with respect to the sum of vectors, therefore $\left(\underline{v_1} - \underline{v_2}\right)$ is itself a member of V distinct from $\underline{0}$, be it $\underline{v}$. In other words, if T is injective, $T(\underline{v})$ has to be different from $\underline{0}$ for any $\underline{v} \in V$, that isn't the null vector, that is to say that the kernel is only composed of the null vector, which is the definition of the nullity of a linear transformation to be 0. □

Theorem 1.4.5: Let $T : V \mapsto W$ be a linear transformation. If T is invertible, then V and W have the same dimension.

Proof: By Theorem 1.4.3, $\dim(V) = \dim(\ker(T)) + \dim(\mathfrak{I}(T))$. Being T invertible, the dimension of the image equals the dimension of the codomain W. By Theorem 1.4.4, $\dim(\ker(T)) = 0$. Therefore, $\dim(V) = 0 + \dim(\mathfrak{I}(T)) = \dim(W)$. □

As stated before, every result concerning linear transformations can be formulated very naturally as a result concerning matrices.

Theorem 1.4.6: Let A be the $m \times n$ matrix associated to the invertible linear application $T : V \mapsto W$ with respect to two bases $\mathcal{B}$ and $\mathcal{C}$. Then, m and n are equal (that is, A is a square matrix).

Proof: By Theorem 1.4.5, if T is an invertible linear transformation, $\dim(V) = \dim(W)$. Since the dimensions of A are $\dim(V)$ and $\dim(W)$ respectively, $m = n$. □

Indeed, it is possible to define a kernel and an image for an invertible matrix. Consider the linear transformation $T : V \mapsto W$, with respect to whom a $n \times n$ matrix A of real values can be associated. Therefore, any vector $\underline{w} \in W$ can be written as $A\underline{v}$, where $\underline{v}$ is a vector in V. Writing v as a linear combination of the canonical basis of V gives:

$$\underline{w} = A\underline{v} = A\left(\lambda_1 \underline{e_1} + \lambda_2 \underline{e_2} + \dots + \lambda_n \underline{e_n}\right) = A\lambda_1 \underline{e_1} + A\lambda_2 \underline{e_2} + \dots + A\lambda_n \underline{e_n}$$

But multiplying the matrix A by the canonical vector $\underline{e_i}$ simply returns the i-th column of A:

$$A\underline{e_i} = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n} \end{pmatrix} \begin{pmatrix} 0 \\ \vdots \\ i \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} a_{1,i} \\ a_{2,i} \\ \vdots \\ a_{n,i} \end{pmatrix}$$

Which means that any image can be written as a linear combination of the columns of A, taken as vectors:

$$\underline{w} = A\underline{v} = A\lambda_1 \underline{e_1} + A\lambda_2 \underline{e_2} + \dots + A\lambda_n \underline{e_n} = \lambda_1 \begin{pmatrix} a_{1,1} \\ a_{2,1} \\ \vdots \\ a_{n,1} \end{pmatrix} + \lambda_2 \begin{pmatrix} a_{1,2} \\ a_{2,2} \\ \vdots \\ a_{n,2} \end{pmatrix} + \dots + \lambda_n \begin{pmatrix} a_{1,n} \\ a_{2,n} \\ \vdots \\ a_{n,n} \end{pmatrix}$$

Being $\underline{w}$ a generic vector, this must mean that the columns of A, taken as vectors, is a set that can generate W. To know the dimension of $\mathfrak{I}(V)$, that is to say, the rank of A, it suffices to find the smallest number of column-vectors of A that is linearly independent.

By Theorem 1.4.3, the dimension of $\mathfrak{I}(V)$, which is equal to $\text{rank}(A)$, has to be equal to the dimension of the domain, which in this case is just $\mathbb{R}^n$ having dimension n. Therefore, for a matrix to be invertible (that is, to be the matrix associated to an invertible linear transformation), its rank has to be equal to the number of its columns or, equivalently, if its columns form a linearly independent set. If this happens, such a matrix is said to have **full rank**.

Theorem 1.4.7: A matrix is invertible if and only if it has full rank.

1.5. Eigenvalues and eigenvectors

Let A be an $n \times n$ square matrix, and let λ be a real value. The n-dimensional vector $\underline{x}$ is said to be an **eigenvector** of A if it's not null and if:

$$A\underline{x} = \lambda\underline{x}$$

Where λ is the corresponding **eigenvalue** of A.
Retrieving the eigenvectors of a matrix A by applying such definition is not possible, since the information contained in the equation is insufficient. Infact:

$$A\underline{x} = \lambda\underline{x} \Rightarrow \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \dots & a_{n,n} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \lambda \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \Rightarrow \begin{cases} a_{1,1} \cdot x_1 + a_{1,2} \cdot x_2 + \dots + x_n a_{1,n} = \lambda x_1 \\ a_{2,1} \cdot x_1 + a_{2,2} \cdot x_2 + \dots + x_n a_{2,n} = \lambda x_2 \\ \vdots \\ a_{n,1} \cdot x_1 + a_{n,2} \cdot x_2 + \dots + x_n a_{n,n} = \lambda x_n \end{cases}$$

Even assuming the A matrix to be known, this system of equation has n equations but $n+1$ unknowns (the n components of $\underline{x}$ and λ). It is still possible to retrieve the eigenvectors of a matrix by following a different approach, by first retrieving its eigenvalues and then applying such definition.
Given a square matrix A and a real value λ, the **characteristic polynomial** of A is defined as:

$$p_A(\lambda) = \det(A - \lambda I) = c_0 + c_1\lambda + c_2\lambda^2 + \dots + c_{n-1}\lambda^{n-1} + (-1)^n \lambda^n$$

Where:

$$c_0 = \det(A) \qquad\qquad c_{n-1} = (-1)^{n-1} \operatorname{tr}(A)$$

Theorem 1.5.1: A real value is an eigenvalue for a given matrix if and only if it is a root of its characteristic polynomial.

Proof: First, suppose that $\lambda \in \mathbb{R}$ is an eigenvalue for a $n \times n$ square matrix A. By definition of eigenvalue, there must exist a non-null vector $\underline{x}$ such that $A\underline{x} = \lambda\underline{x}$. Then:

$$A\underline{x} = \lambda\underline{x} \Rightarrow A\underline{x} = \lambda I\underline{x} \Rightarrow A\underline{x} - \lambda I\underline{x} = \underline{0} \Rightarrow (A - \lambda I)\underline{x} = \underline{0}$$

This means that $\underline{x}$ is a vector that belongs to the kernel of the matrix $(A - \lambda I)$. Therefore, the nullity of $(A - \lambda I)$ can't be zero.
By Theorem 1.4.3, $\dim(A - \lambda I) = \operatorname{rank}(A - \lambda I) + \operatorname{null}(A - \lambda I)$. But $(A - \lambda I)$ and A have the same dimension, therefore $n = \operatorname{rank}(A - \lambda I) + \operatorname{null}(A - \lambda I)$. Since $\operatorname{null}(A - \lambda I)$ is non zero, for this equality to hold the rank of $(A - \lambda I)$ has to be less than n. By Theorem 1.4.7, the matrix $(A - \lambda I)$ cannot be invertible, and by Theorem 1.1.1 this must mean that the determinant of $(A - \lambda I)$ is 0.
Suppose then that λ is a root for the characteristic polynomial of A. This means that $\det(A - \lambda I)$ is equal to 0. By Theorem 1.1.1, this must mean that $(A - \lambda I)$ is not invertible, which in turn by Theorem 1.4.7 must mean that the rank of $(A - \lambda I)$ is less than n. By Theorem 1.4.3, $n = \operatorname{rank}(A - \lambda I) + \operatorname{null}(A - \lambda I)$, and being the rank less than n in turn implies that the kernel of $(A - \lambda I)$ does not contain just the null vector. This means that it exists a vector $\underline{x}$ such that $(A - \lambda I)\underline{x} = \underline{0}$. But then:

$$(A - \lambda I)\underline{x} = \underline{0} \Rightarrow A\underline{x} - \lambda I\underline{x} = \underline{0} \Rightarrow A\underline{x} = \lambda I\underline{x} \Rightarrow A\underline{x} = \lambda\underline{x}$$

Which is the definition of eigenvalue. □

Knowing how to compute eigenvalues, it is then possible to solve the aforementioned equation and retrieve the eigenvectors.

Eigenvectors and eigenvalues can be defined with respect to linear transformations as well. Given a linear transformation $T : V \mapsto V$, a vector $\underline{v} \in V$ is an eigenvector for T if $T\underline{v} = \lambda\underline{v}$, where λ is an eigenvalue for T. Notice how it has been imposed that the transformation T is an endomorphism, since otherwise mirroring the definition of eigenvector for matrices could not have been possible.
As stated in Theorem 1.5.1, to compute the eigenvalues of a matrix, it suffices to compute its characteristic polynomial. But any matrix can be associated to a linear transformation and vice versa, therefore to compute the eigenvalues of a linear transformation it suffices to compute the characteristic polynomial of the associated matrix of the linear transformation.
Recall that to construct the associated matrix two bases ought to be fixed, and changing the bases results in a different matrix. But since the determinant is invariant with respect to different bases for the associated matrix, it does not matter which bases were chosen. Therefore, the characteristic polynomial can be said to refer to the linear transformation itself, and not to a certain associated matrix.
Suppose that to a certain linear transformation $T : V \mapsto V$ it is possible to associate a certain matrix A with respect to a basis $\mathcal{B}$ and a different matrix A' with respect to a different basis $\mathcal{B}'$. Recall that the formula to pass from a matrix to the other is $A = PA'P^{-1}$. Then:

$$p_A(\lambda) = \det(A - \lambda I) = \det\big(PA'P^{-1} - \lambda I\big) = \det\big(PA'P^{-1} - \lambda PIP^{-1}\big) = \det\big(P(A'P^{-1} - \lambda IP^{-1})\big) =$$
$$\det\big(P(A' - \lambda I)P^{-1}\big) = \cancel{\det(P)}\det(A' - \lambda I)\cancel{\det(P^{-1})} = \det(A' - \lambda I) = p_{A'}(\lambda)$$