Appendix A

Linear algebra and Convex sets

The simplex method described in Chapter 2 is a method of algebraic nature and consists of solving systems of linear equations and determining the solution that optimizes a previously determined objective function. In order to understand the algebraic procedure, it is instructive to study its underlying geometry.

In this Appendix, we review some basic linear algebra and convexity results needed throughout the preceding chapters.

A.1 Matrices and vectors

Let $\mathbb{R}$ be the field of real numbers. The elements of $\mathbb{R}$ are called scalars. A *matrix* $\mathbf{A}$ is a rectangular array of scalars. If $\mathbf{A} \in \mathbb{R}^{m \times n}$, the matrix has m rows and n columns whose components are elements of $\mathbb{R}$, and it is called an $m \times n$ matrix (reads “m by n):

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix},$$

where a_{ij} denotes the entry in row i and column j. The matrix can also be denoted by $\mathbf{A} = (a_{ij})$. If a matrix has just one column, that is, an $m \times 1$ matrix, it is said to be a *column vector*.

$$\mathbf{a} = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix}$$

Therefore, $\mathbf{a}^T$ will denote a *row vector*.

Examples.

1. The following array of scalars is a 3×4 matrix:

$$\mathbf{A} = \begin{pmatrix} 3 & -2 & \frac{2}{5} & 1 \\ 0 & 7 & 2 & \frac{1}{2} \\ 1 & 0 & 1 & 1 \end{pmatrix}$$

2. The following array of scalars is a 3×1 vector:

$$\mathbf{a} = \begin{pmatrix} 3 \\ 0 \\ 1 \end{pmatrix}$$

□

A.1.1 Matrix operations

Addition of matrices

Given two $m \times n$ matrices, $\mathbf{A} = (a_{ij})$ and $\mathbf{B} = (b_{ij}) \in \mathbb{R}^{m \times n}$, the *addition* of $\mathbf{A}$ and $\mathbf{B}$, denoted by $\mathbf{A} + \mathbf{B}$, is the matrix $\mathbf{C} = (c_{ij}) \in \mathbb{R}^{m \times n}$ obtained by performing the addition componentwise:

$$c_{ij} = a_{ij} + b_{ij}, \quad i = 1, \ldots, m, \ j = 1, \ldots, n.$$

Note that, the addition of two matrices is only possible if they are both of the same size. The resulting matrix will also be of the same size.

The addition of vectors is a particular case of addition of matrices. Therefore, it is defined in the same way. We only have to take into account that a vector is a one-column matrix.

Examples.

1. Consider the vectors $\mathbf{a} = \begin{pmatrix} 1 \\ 3 \end{pmatrix}$ and $\mathbf{b} = \begin{pmatrix} 1 \\ 4 \end{pmatrix}$,

$$\mathbf{a} + \mathbf{b} = \begin{pmatrix} 1 \\ 3 \end{pmatrix} + \begin{pmatrix} 1 \\ 4 \end{pmatrix} = \begin{pmatrix} 2 \\ 7 \end{pmatrix}.$$

2. Consider $\mathbf{A} = \begin{pmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & -1 & 0 \\ 2 & 0 & 1 & -3 \end{pmatrix}$ and $\mathbf{B} = \begin{pmatrix} 1 & 4 & 2 & 0 \\ 0 & 0 & -1 & 0 \\ 1 & 1 & 2 & -1 \end{pmatrix}$,

$$\mathbf{A}+\mathbf{B} = \begin{pmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & -1 & 0 \\ 2 & 0 & 1 & -3 \end{pmatrix} + \begin{pmatrix} 1 & 4 & 2 & 0 \\ 0 & 0 & -1 & 0 \\ 1 & 1 & 2 & -1 \end{pmatrix} = \begin{pmatrix} 2 & 4 & 3 & 1 \\ 0 & 1 & -2 & 0 \\ 3 & 1 & 3 & -4 \end{pmatrix}.$$

$\square$

Properties

1. The addition of matrices is an inner operation in $\mathbb{R}^{m\times n}$.

$$\mathbf{A}, \mathbf{B} \in \mathbb{R}^{m\times n} \Rightarrow \mathbf{A} + \mathbf{B} \in \mathbb{R}^{m\times n}.$$

2. The addition of matrices is commutative.

$$\forall \mathbf{A}, \mathbf{B} \in \mathbb{R}^{m\times n} \quad \mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}.$$

3. The addition of matrices is associative.

$$\forall \mathbf{A}, \mathbf{B}, \mathbf{C} \in \mathbb{R}^{m\times n} \quad (\mathbf{A} + \mathbf{B}) + \mathbf{C} = \mathbf{A} + (\mathbf{B} + \mathbf{C}).$$

4. There exists a neutral element for the addition. $\forall \mathbf{A} \in \mathbb{R}^{m\times n}$ there exists a neutral element denoted by $\mathbf{0} \in \mathbb{R}^{m\times n}$ such that

$$\mathbf{A} + \mathbf{0} = \mathbf{0} + \mathbf{A} = \mathbf{A}.$$

5. There exist opposite elements for the addition. $\forall \mathbf{A} \in \mathbb{R}^{m\times n}$ there exists an element denoted by $-\mathbf{A} \in \mathbb{R}^{m\times n}$ such that

$$\mathbf{A} + (-\mathbf{A}) = (-\mathbf{A}) + \mathbf{A} = \mathbf{0}.$$

The addition of vectors satisfies the same properties as the addition of matrices.

Scalar multiplication

Let $\alpha \in \mathbb{R}$ be a scalar and $\mathbf{A} = (a_{ij}) \in \mathbb{R}^{m\times n}$ be a matrix. The operation of multiplying $\mathbf{A}$ by α is represented by $\alpha \cdot \mathbf{A}$ and is performed componentwise, that is, multiplying every element of $\mathbf{A}$ by α. The result of this multiplication is a matrix $\mathbf{B} = (b_{ij}) \in \mathbb{R}^{m\times n}$ of the same size as $\mathbf{A}$, such that

$$b_{ij} = \alpha \cdot a_{ij}, \quad i = 1, \ldots, m, \; j = 1, \ldots, n.$$

Examples.

1. Let us consider $\mathbf{A} = \begin{pmatrix} 0 & -2 \\ 1 & 2 \\ 1 & 1 \end{pmatrix}$ and $\alpha = -2$. In the scalar multiplication,

$$\alpha \cdot \mathbf{A} = -2 \cdot \begin{pmatrix} 0 & -2 \\ 1 & 2 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 4 \\ -2 & -4 \\ -2 & -2 \end{pmatrix}.$$

2. Let us consider $\mathbf{a} = \begin{pmatrix} 1 \\ 3 \\ -5 \end{pmatrix}$ and $\alpha = \frac{1}{2}$. In the scalar multiplication,

$$\alpha \cdot \mathbf{a} = \frac{1}{2} \cdot \begin{pmatrix} 1 \\ 3 \\ -5 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} \\ \frac{3}{2} \\ -\frac{5}{2} \end{pmatrix}.$$

□

Inner product

Let $\mathbf{a}^T = (a_1 \ \cdots \ a_n) \in \mathbb{R}^{1\times n}$ be a row vector and $\mathbf{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} \in \mathbb{R}^n$ a column vector. The result of multiplying $\mathbf{a}^T$ and $\mathbf{b}$ is called the *inner product* of the two vectors, and is denoted by $\mathbf{a}^T \cdot \mathbf{b}$. The result of this multiplication is a real number, and is obtained multiplying componentwise the two vectors and adding the results in the following way:

$$\mathbf{a}^T \cdot \mathbf{b} = (a_1 \ \cdots \ a_n). \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = \sum_{i=1}^{n} a_i \cdot b_i.$$

The inner product of two vectors is only possible if both are of the same size.

$$\mathbf{a} \in \mathbb{R}^n, \mathbf{b} \in \mathbb{R}^n \Rightarrow \mathbf{a}^T \cdot \mathbf{b} \in \mathbb{R}.$$

Example. Let $\mathbf{a} = \begin{pmatrix} 4 \\ 2 \\ 7 \end{pmatrix}$ and $\mathbf{b} = \begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix} \in \mathbb{R}^3$. The inner product:

$$\mathbf{a}^T \cdot \mathbf{b} = (4\ \ 2\ \ 7) \cdot \begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix} = 7 \in \mathbb{R}\,.$$

□

Matrix multiplication

Let $\mathbf{A} \in \mathbb{R}^{m\times n}$ be an $m \times n$ matrix and $\mathbf{B} \in \mathbb{R}^{n\times p}$ an $n \times p$ matrix. The *product* $\mathbf{A} \cdot \mathbf{B}$ is defined as the $m \times p$ matrix $\mathbf{C} = \mathbf{A} \cdot \mathbf{B} \in \mathbb{R}^{m\times p}$, where the entry (i, j) in $\mathbf{C}$ is the inner product of the ith row of $\mathbf{A}$ and the jth column of $\mathbf{B}$.

Example. Consider matrices $\mathbf{A} = \begin{pmatrix} 4 & -3 \\ 1 & 0 \\ 1 & 1 \end{pmatrix}$ and $\mathbf{B} = \begin{pmatrix} 1 & 1 & 0 & 1 \\ 1 & 2 & 7 & 4 \end{pmatrix}$.

$$\mathbf{C} = \mathbf{A} \cdot \mathbf{B} = \begin{pmatrix} 1 & -2 & -21 & -8 \\ 1 & 1 & 0 & 1 \\ 2 & 3 & 7 & 5 \end{pmatrix}.$$

□

Properties

1. The product of matrices is associative.

$$\forall \mathbf{A} \in \mathbb{R}^{m\times n},\ \forall \mathbf{B} \in \mathbb{R}^{n\times p},\ \forall \mathbf{C} \in \mathbb{R}^{p\times q}, \quad (\mathbf{A} \cdot \mathbf{B}) \cdot \mathbf{C} = \mathbf{A} \cdot (\mathbf{B} \cdot \mathbf{C}).$$

2. The product of matrices is distributive with respect to the sum.

$$\forall \mathbf{A}, \mathbf{B} \in \mathbb{R}^{m\times n},\ \forall \mathbf{C} \in \mathbb{R}^{n\times p}, \quad (\mathbf{A} + \mathbf{B}) \cdot \mathbf{C} = \mathbf{A} \cdot \mathbf{C} + \mathbf{B} \cdot \mathbf{C}.$$

$$\forall \mathbf{A} \in \mathbb{R}^{m\times n},\ \forall \mathbf{B}, \mathbf{C} \in \mathbb{R}^{n\times p}, \quad \mathbf{A} \cdot (\mathbf{B} + \mathbf{C}) = \mathbf{A} \cdot \mathbf{B} + \mathbf{A} \cdot \mathbf{C}.$$

3. $\forall \mathbf{A} \in \mathbb{R}^{m\times n}, \quad \mathbf{A}\cdot \mathbf{0}_{n\times p} = \mathbf{0}_{m\times p}, \quad \mathbf{0}_{q\times m}\cdot \mathbf{A} = \mathbf{0}_{q\times n}.$

4. $\forall \mathbf{A} \in \mathbb{R}^{m\times n}, \quad \mathbf{I}_m \cdot \mathbf{A} = \mathbf{A}\cdot \mathbf{I}_n = \mathbf{A}.$

5. $\forall \alpha \in \mathbb{R},\ \forall \mathbf{A} \in \mathbb{R}^{m\times n},\ \forall \mathbf{B} \in \mathbb{R}^{n\times p}, \quad \alpha\cdot(\mathbf{A}\cdot\mathbf{B}) = (\alpha\cdot\mathbf{A})\cdot\mathbf{B} = \mathbf{A}\cdot(\alpha\cdot\mathbf{B}).$

A.1.2 The rank of a matrix

Given a matrix $\mathbf{A} \in \mathbb{R}^{m\times n}$, it can be reduced to the form $\mathbf{U}$ by performing elementary row operations through the Gaussian elimination. The number of pivot elements in $\mathbf{U}$ is the number of nonzero rows. Zero is never allowed as a pivot.

Example. Consider the following matrix:

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 3 & -4 & 1 \\ 1 & 2 & 2 & 5 & 4 \\ 3 & 2 & -5 & 2 & 4 \\ 2 & 0 & -6 & 9 & 7 \end{pmatrix}.$$

Performing Gaussian elimination, we obtain the upper triangular matrix $\mathbf{U}$.

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 3 & -4 & 1 \\ 1 & 2 & 2 & 5 & 4 \\ 3 & 2 & -5 & 2 & 4 \\ 2 & 0 & -6 & 9 & 7 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 2 & 3 & -4 & 1 \\ 0 & 0 & -1 & 9 & 3 \\ 0 & -4 & -14 & 14 & 1 \\ 0 & -4 & -12 & 17 & 5 \end{pmatrix}$$

$$\rightarrow \begin{pmatrix} 1 & 2 & 3 & -4 & 1 \\ 0 & -4 & -14 & 14 & 1 \\ 0 & 0 & -1 & 9 & 3 \\ 0 & -4 & -12 & 17 & 5 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 2 & 3 & -4 & 1 \\ 0 & -4 & -14 & 14 & 1 \\ 0 & 0 & -1 & 9 & 3 \\ 0 & 0 & 2 & 3 & 4 \end{pmatrix}$$

$$\rightarrow \begin{pmatrix} 1 & 2 & 3 & -4 & 1 \\ 0 & -4 & -14 & 14 & 1 \\ 0 & 0 & -1 & 9 & 3 \\ 0 & 0 & 0 & 21 & 10 \end{pmatrix} = \mathbf{U}.$$

□

Definition A.1.1 *Let* $\mathbf{A} \in \mathbb{R}^{m\times n}$ *be a matrix, and let* $\mathbf{U}$ *be the resulting matrix after performing Gaussian elimination. The* rank *of matrix* $\mathbf{A}$ *is denoted by rank* $\mathbf{A}$ *and is equal to the number of pivot elements of matrix* $\mathbf{U}$.

The rank of the matrix in the preceding example is 4, the same as the number of pivot elements. Notice that the rank coincides with the number of nonzero rows in $\mathbf{U}$.

A.2 Systems of linear equations

Let us consider a system with m linear equations of n variables

$$\mathbf{A}\mathbf{x} = \mathbf{b},$$

where $\mathbf{A} \in \mathbb{R}^{m\times n}$, rank $\mathbf{A} = r$ and $\mathbf{b} \in \mathbb{R}^m$. We shall solve the system by performing Gaussian elimination. The following cases may arise:

- rank $\mathbf{A} \neq$ rank $(\mathbf{A}\ \mathbf{b})$. Then, the system has no solution. It is said that the system is inconsistent.
- rank $\mathbf{A} =$ rank $(\mathbf{A}\ \mathbf{b}) = r$. Then, there exists at least one solution to the system. It is said that the system is consistent.
  - $r =$ number of variables. Then, there is one and only one solution. It is said that the solution of the system is unique.
  - $r <$ number of variables. Then, there exists an infinite number of solutions.

Example. Let us consider the following system of linear equations:

$$\begin{aligned} 2x_1 - x_2 + 3x_3 &= 2 \\ x_1 + x_2 - x_3 &= 4 \\ 3x_1 \qquad + 2x_3 &= 5 \end{aligned}$$

In order to determine whether the system is consistent or inconsistent, we will compute the rank of matrices $\mathbf{A}$ and $(\mathbf{A}\ \mathbf{b})$.

$$\mathbf{A} = \begin{pmatrix} 2 & -1 & 3 \\ 1 & 1 & -1 \\ 3 & 0 & 2 \end{pmatrix}, \qquad (\mathbf{A}\ \mathbf{b}) = \begin{pmatrix} 2 & -1 & 3 & 2 \\ 1 & 1 & -1 & 4 \\ 3 & 0 & 2 & 5 \end{pmatrix}.$$

Performing the Gaussian elimination, it results:

$$\begin{pmatrix} 2 & -1 & 3 & 2 \\ 1 & 1 & -1 & 4 \\ 3 & 0 & 2 & 5 \end{pmatrix} \rightarrow \begin{pmatrix} 2 & -1 & 3 & 2 \\ 0 & \frac{3}{2} & -\frac{5}{2} & 3 \\ 0 & \frac{3}{2} & -\frac{5}{2} & 2 \end{pmatrix} \rightarrow \begin{pmatrix} 2 & -1 & 3 & 2 \\ 0 & \frac{3}{2} & -\frac{5}{2} & 3 \\ 0 & 0 & 0 & -1 \end{pmatrix}.$$

We conclude that rank $\mathbf{A} = 2 < 3 =$ rank $(\mathbf{A}\ \mathbf{b})$. Thus, the system is inconsistent.

□

Example. Let us consider the following system of linear equations:

$$\begin{aligned} 2x_1 + x_2 &= 3 \\ x_1 + x_2 &= 4 \end{aligned}$$

Performing the Gaussian elimination, we obtain:

$$\begin{pmatrix} 2 & 1 & 3 \\ 1 & 1 & 4 \end{pmatrix} \rightarrow \begin{pmatrix} 2 & 1 & 3 \\ 0 & \frac{1}{2} & \frac{5}{2} \end{pmatrix},$$

from where we conclude that rank $\mathbf{A} =$ rank $(\mathbf{A}\ \mathbf{b}) = 2 =$ number of variables. Thus, there exists a unique solution to the system: $x_1 = -1, x_2 = 5$. □

Example. Let us consider the following system of linear equations:

$$
\begin{aligned}
2x_1 - x_2 + 3x_3 &= 2 \\
x_1 + x_2 - x_3 &= 4 \\
3x_1 \qquad + 2x_3 &= 6
\end{aligned}
$$

Performing the Gaussian elimination,

$$
\begin{pmatrix} 2 & -1 & 3 & 2 \\ 1 & 1 & -1 & 4 \\ 3 & 0 & 2 & 6 \end{pmatrix} \rightarrow \begin{pmatrix} 2 & -1 & 3 & 2 \\ 0 & \frac{3}{2} & -\frac{5}{2} & 3 \\ 0 & \frac{3}{2} & -\frac{5}{2} & 3 \end{pmatrix} \rightarrow \begin{pmatrix} 2 & -1 & 3 & 2 \\ 0 & \frac{3}{2} & -\frac{5}{2} & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}.
$$

We conclude that rank $\mathbf{A}$ = rank $(\mathbf{A}\ \mathbf{b}) = 2 <$ number of variables. Thus, there exists an infinite number of solutions to the system. Observe that the system can be written in the following way:

$$
\begin{aligned}
2x_1 - x_2 &= 2 - 3x_3 \\
\frac{3}{2}x_2 &= 3 + \frac{5}{2}x_3
\end{aligned}
$$

The variables x_1 and x_2 are called *basic variables* and x_3 is a *nonbasic variable*. The infinite solutions of the system can be described as:

$$
x_1 = 2 - \frac{2}{3}x_3, \quad x_2 = 2 + \frac{5}{3}x_3, \quad x_3 \in \mathbb{R}\,.
$$

□

A.2.1 Basic solutions

Let $\mathbf{Ax} = \mathbf{b}$ be a system, where $\mathbf{A} \in \mathbb{R}^{m\times n}$, $m < n$, and rank $\mathbf{A}$ = rank $(\mathbf{A}\ \mathbf{b}) = m$. Assume that the first m columns of $\mathbf{A}$ are linearly independent. Let $\mathbf{B}$ be the $m \times m$ submatrix of $\mathbf{A}$ formed by its first m columns. Let $\mathbf{N}$ denote the last $n-m$ columns of $\mathbf{A}$. We can write the system $\mathbf{Ax} = \mathbf{b}$ in the following way:

$$
(\mathbf{B}\ \ \mathbf{N}) \begin{pmatrix} \mathbf{x}_B \\ \mathbf{x}_N \end{pmatrix} = \mathbf{b},
$$

or

$$\mathbf{B}\mathbf{x}_B + \mathbf{N}\mathbf{x}_N = \mathbf{b}.$$

The variables in $\mathbf{x}_B$ are the *basic variables*, and the ones in $\mathbf{x}_N$ are the *nonbasic variables*. Then, we can solve $\mathbf{x}_B$ in terms of $\mathbf{x}_N$,

$$\mathbf{B}\mathbf{x}_B = \mathbf{b} - \mathbf{N}\mathbf{x}_N.$$

We observe that there exists an infinite number of solutions to the system, which can be computed assigning arbitrary values to the nonbasic vector $\mathbf{x}_N$. The unique solution obtained setting the nonbasic variables equal to 0, that is setting $\mathbf{x}_N = \mathbf{0}$, satisfies:

$$\mathbf{B}\mathbf{x}_B = \mathbf{b},$$

and is called a *basic solution* of the system.

Example. Consider the following system of linear equations:

$$\begin{aligned} 2x_1 - x_2 + 3x_3 &= 2 \\ x_1 + x_2 - x_3 &= 4 \\ 3x_1 \qquad + 2x_3 &= 6 \end{aligned}$$

By Gaussian elimination we obtain the equivalent system:

$$\begin{aligned} 2x_1 - x_2 &= 2 - 3x_3 \\ \frac{3}{2}x_2 &= 3 + \frac{5}{2}x_3 \end{aligned}$$

There exists an infinite number of solutions to the system. We can solve $x_1 = 2 - \frac{2}{3}x_3$ and $x_2 = 2 + \frac{5}{3}x_3$, in terms of the nonbasic variable x_3. If we assign $x_3 = 0$, we obtain the basic solution $x_1 = 2$, $x_2 = 2$.

There are different ways of extracting a submatrix $\mathbf{B}$ out of $\mathbf{A}$, and computing the corresponding basic solution setting the nonbasic variables equal to 0. The maximum number of basic solutions is bounded by the number of ways that m columns can be extracted out of n to form the basis:

$$\begin{pmatrix} n \\ m \end{pmatrix} = \frac{n!}{m!\,(n-m)!}$$

In this example, the number of basic solutions is less than or equal to:

$$\begin{pmatrix} 3 \\ 2 \end{pmatrix} = \frac{3!}{2!\,(3-2)!} = 3.$$

□

A.3 Vector spaces

Let us consider the vector space $\mathbb{R}^m$.

Definition A.3.1 (Linear combination) *A linear combination of vectors* $\mathbf{v}_1$, $\mathbf{v}_2$, $\dots$, $\mathbf{v}_n$ *of* $\mathbb{R}^m$ *is:*

$$\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \cdots + \alpha_n\mathbf{v}_n,$$

where $\alpha_1, \dots, \alpha_n \in \mathbb{R}$ *are real numbers.*

Example. Let $\mathbf{v}_1$ and $\mathbf{v}_2$ be the vectors:

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

- The following is a linear combination of $\mathbf{v}_1$ and $\mathbf{v}_2$:

$$2\begin{pmatrix} 1 \\ 0 \end{pmatrix} + 5\begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

- The next expression represents all possible linear combinations of the vectors $\mathbf{v}_1$ and $\mathbf{v}_2$:

$$\alpha_1\begin{pmatrix} 1 \\ 0 \end{pmatrix} + \alpha_2\begin{pmatrix} 1 \\ -1 \end{pmatrix},$$

  for every $\alpha_1, \alpha_2 \in \mathbb{R}$.

$\square$

A.3.1 Linear dependence and independence

Definition A.3.2 *The vectors* $\mathbf{v}_1, \dots, \mathbf{v}_n \in \mathbb{R}^m$ *are said to be linearly independent if for every linear combination such that*

$$\alpha_1\mathbf{v}_1 + \cdots + \alpha_n\mathbf{v}_n = \mathbf{0}$$

it implies that

$$\alpha_1 = \cdots = \alpha_n = 0.$$

Definition A.3.3 *The vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n \in \mathbb{R}^m$ are said to be linearly dependent if there exist $\alpha_1, \cdots, \alpha_n \in \mathbb{R}$ not all of them zero, such that*

$$\alpha_1 \mathbf{v}_1 + \cdots + \alpha_n \mathbf{v}_n = \mathbf{0}.$$

Obviously, they are not linearly independent.

Example.

1. Let us consider the vectors $\begin{pmatrix} 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 4 \\ -4 \end{pmatrix}$.

   As,
   $$4 \begin{pmatrix} 1 \\ -1 \end{pmatrix} + (-1) \begin{pmatrix} 4 \\ -4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$
   we conclude that the two vectors are linearly dependent.

2. Let us consider the vectors $\begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}, \begin{pmatrix} 3 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 9 \\ -3 \\ 5 \end{pmatrix}$,

   and the following linear combination:
   $$\alpha_1 \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix} + \alpha_2 \begin{pmatrix} 3 \\ 0 \\ -1 \end{pmatrix} + \alpha_3 \begin{pmatrix} 9 \\ -3 \\ 5 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$
   By Gaussian elimination we obtain:
   $$\begin{pmatrix} 1 & 3 & 9 \\ -1 & 0 & -3 \\ 2 & -1 & 5 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 3 & 9 \\ 0 & 3 & 6 \\ 0 & -7 & -13 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 3 & 9 \\ 0 & 3 & 6 \\ 0 & 0 & 1 \end{pmatrix}.$$

   As there are three pivot elements in the reduced matrix, the unique solution of the system is: $\alpha_1 = \alpha_2 = \alpha_3 = 0$. This implies that the three vectors given are linearly independent. □

A.3.2 Basis and dimension

Definition A.3.4 *A set of vectors $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_p\} \subseteq \mathbb{R}^m$ is said to be a spanning set for $\mathbb{R}^m$ if every vector $\mathbf{v} \in \mathbb{R}^m$ can be represented as a linear combination of the vectors of S, that is, if there exist $\alpha_1, \ldots, \alpha_p \in \mathbb{R}$ such that*

$$\mathbf{v} = \alpha_1 \mathbf{v}_1 + \cdots + \alpha_p \mathbf{v}_p.$$

Example. Let S be the following set of vectors, and $\mathbf{v}$ be a vector of $\mathbb{R}^3$:

$$S = \left\{ \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}, \begin{pmatrix} 3 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ -2 \\ 2 \end{pmatrix} \right\}, \quad \mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}.$$

We will see that there exist $\alpha_1, \alpha_2, \alpha_3, \alpha_4 \in \mathbb{R}$ such that the following system is consistent:

$$\alpha_1 \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} + \alpha_2 \begin{pmatrix} 3 \\ 1 \\ 0 \end{pmatrix} + \alpha_3 \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} + \alpha_4 \begin{pmatrix} 2 \\ -2 \\ 2 \end{pmatrix} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}$$

By Gaussian elimination we obtain:

$$\begin{pmatrix} 1 & 3 & 2 & 2 & v_1 \\ -1 & 1 & 1 & -2 & v_2 \\ 1 & 0 & 1 & 2 & v_3 \end{pmatrix} \to \begin{pmatrix} 1 & 3 & 2 & 2 & v_1 \\ 0 & 4 & 3 & 0 & v_2 + v_1 \\ 0 & -3 & -1 & 0 & v_3 - v_1 \end{pmatrix} \to$$

$$\to \begin{pmatrix} 1 & 3 & 2 & 2 & v_1 \\ 0 & 4 & 3 & 0 & v_2 + v_1 \\ 0 & 0 & \frac{5}{4} & 0 & v_3 - \frac{1}{4}v_1 + \frac{3}{4}v_2 \end{pmatrix}.$$

As rank $\mathbf{A}$ = rank $(\mathbf{A}\ \mathbf{b}) = 3$, the system is consistent. Therefore, S is a spanning set for $\mathbb{R}^3$. □

Definition A.3.5 *A collection of vectors $B = \{\mathbf{v}_1, \ldots, \mathbf{v}_m\} \subseteq \mathbb{R}^m$ forms a* basis *of $\mathbb{R}^m$ if the following conditions hold:*

- *The vectors of B are linearly independent.*
- *B is a spanning set for $\mathbb{R}^m$.*

There are infinite bases in a vector space. However, all of them contain the same number of vectors. This number is the *dimension* of the vector space.

Example. Prove that the following set of vectors B forms a basis for $\mathbb{R}^3$.

$$B = \left\{ \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} \right\}.$$

In order to prove that the three vectors are linearly independent, we solve the system:

$$\alpha_1 \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} + \alpha_2 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + \alpha_3 \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

By Gaussian elimination,

$$\begin{pmatrix} 1 & 0 & 1 \\ 2 & 1 & 1 \\ 0 & 0 & 2 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 2 \end{pmatrix}.$$

The system has a unique solution: $\alpha_1 = \alpha_2 = \alpha_3 = 0$. Therefore, the vectors of B are linearly independent.

In order to prove that B is a spanning set for $\mathbb{R}^3$, we solve the system:

$$\alpha_1 \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} + \alpha_2 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + \alpha_3 \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}.$$

It is easy to see that the system is consistent. Thus, B is a spanning set for $\mathbb{R}^3$.

Therefore, B forms a basis in $\mathbb{R}^3$. $\square$

Theorem A.3.1 *Let $B = \{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$ be a basis in $\mathbb{R}^m$. Then, every vector $\mathbf{v} \in \mathbb{R}^m$ can be expressed as a linear combination of $\mathbf{v}_1, \ldots, \mathbf{v}_m$, and the coefficients of that linear combination are unique.*

The unique coefficients of the linear combination of Theorem A.3.1 are the *coordinates* of the vector $\mathbf{v}$.

Theorem A.3.2 *Given a basis B for $\mathbb{R}^m$ and a vector $\mathbf{v} \in \mathbb{R}^m$, $\mathbf{v} \notin B$ and $\mathbf{v} \neq \mathbf{0}$, it is always possible to form another basis, by replacing a vector in B by the vector $\mathbf{v}$.*

This result is crucial in the development of the linear programming. In fact, the simplex algorithm starts with a basic feasible solution and moves to a better one, by replacing a vector in the basis, as described in the previous theorem. A condition must be satisfied in order to guarantee that, once a vector in the basis is replaced by another vector, the new set of vectors will still form a basis. We will illustrate the condition with an example.

Example. Consider a basis B and a vector $\mathbf{v}$ in $\mathbb{R}^3$.

$$B = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\} = \left\{ \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} \right\}, \quad \mathbf{v} = \begin{pmatrix} 3 \\ 1 \\ 0 \end{pmatrix}.$$

We represent $\mathbf{v}$ as a linear combination of the vectors of B:

$$\begin{pmatrix} 3 \\ 1 \\ 0 \end{pmatrix} = \alpha_1 \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} + \alpha_2 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + \alpha_3 \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix},$$

solve the system and obtain the coordinates: $\alpha_1 = 3$, $\alpha_2 = -5$, $\alpha_3 = 0$. Since $\alpha_1 \neq 0$ and $\alpha_2 \neq 0$, vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ can be substituted by $\mathbf{v}$ to obtain the following bases:

$$B' = \{\mathbf{v}, \mathbf{v}_2, \mathbf{v}_3\} \ , \ B'' = \{\mathbf{v}_1, \mathbf{v}, \mathbf{v}_3\}.$$

However, the set of vectors generated by substituting $\mathbf{v}_3$ by $\mathbf{v}$ does not form a basis, because $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}\}$ are linearly dependent. This happens because $\alpha_3 = 0$. Therefore, the vector $\mathbf{v}_3$ cannot be substituted by $\mathbf{v}$.

$\square$

A.4 Convex sets

The *Euclidean plane* is the set of all ordered pairs of real numbers

$$\mathbb{R}^2 = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \text{ where } x_1 \text{ and } x_2 \text{ are real numbers} \right\}.$$

Geometrically, $\mathbb{R}^2$ can be represented as in Figure A.1:

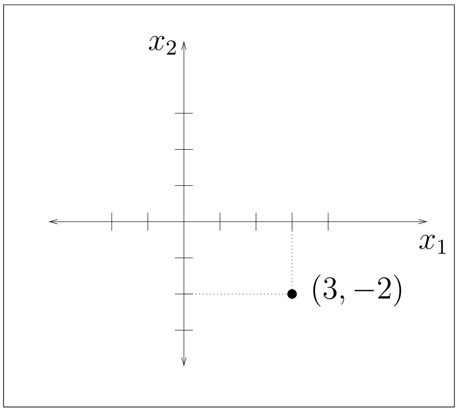

Figure A.1: Euclidean plane

In $\mathbb{R}^2$ an equation of the form $a_1x_1+a_2x_2 = c$, where a_1, a_2 and c are constants in $\mathbb{R}$, represents a *straight line*. For example, the equation $2x_1 + 3x_2 = 6$ is the line drawn in Figure A.2.

An inequality of the form $a_1x_1 + a_2x_2 \leq c$ is the set of all of the points lying on the line $a_1x_1 + a_2x_2 = c$, together with all those points lying to one side of the line. For example, $2x_1 + 3x_2 \leq 6$ is the set of all points in the shaded region in Figure A.3.

A *half-space* of $\mathbb{R}^2$ is the set of all points of $\mathbb{R}^2$ which satisfies an inequality of the form $a_1x_1 + a_2x_2 \leq c$ or $a_1x_1 + a_2x_2 \geq c$, where at least one of the constants a_1 or a_2 is nonzero.

The 3-dimensional *Euclidean space* is the set of all ordered triplets,

$$\mathbb{R}^3 = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \text{ where } x_1, x_2 \text{ and } x_3 \text{ are real numbers} \right\}.$$

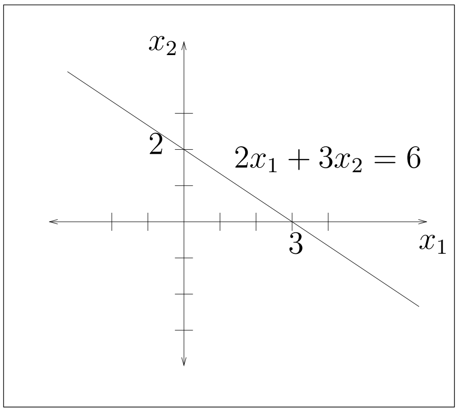

Figure A.2: A straight line in the plane

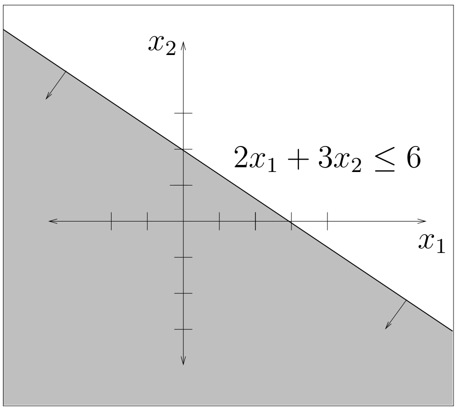

Figure A.3: An inequality in the plane

In $\mathbb{R}^3$ the equation of the form $a_1x_1 + a_2x_2 + a_3x_3 = c$ where a_1, a_2, a_3 and c are constants in $\mathbb{R}$, represents a *plane*. For instance, $3x_1 - x_2 + 4x_3 = 6$ is a plane.

A *half-space* of $\mathbb{R}^3$ is the set of all points of $\mathbb{R}^3$ which satisfies an inequality of the form $a_1x_1 + a_2x_2 + a_3x_3 \leq c$ or $a_1x_1 + a_2x_2 + a_3x_3 \geq c$.

We can generalize these definitions to an n-dimensional Euclidean space

$$\mathbb{R}^n = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \text{ where } x_1, x_2, \ldots x_n \text{ are real numbers} \right\}$$

In $\mathbb{R}^n$ the equation $a_1x_1 + a_2x_2 + \cdots + a_nx_n = c$ where $a_1, \ldots, a_n, c \in \mathbb{R}$ are constants, represents a *hyperplane*.

A *half-space* of $\mathbb{R}^n$ is the set of the points of $\mathbb{R}^n$ which satisfies an inequality of the form $a_1x_1 + a_2x_2 + \cdots + a_nx_n \leq c$ or $a_1x_1 + a_2x_2 + \cdots + a_nx_n \geq c$.

Definition A.4.1 *A subset C of $\mathbb{R}^n$ is a* convex set *if C is empty, if C contains a single point, or if for every two distinct points in C, the line segment connecting them lies entirely in C.*

The sets (a), (b) and (c) in Figure A.4 are convex. The set (d) is not convex.

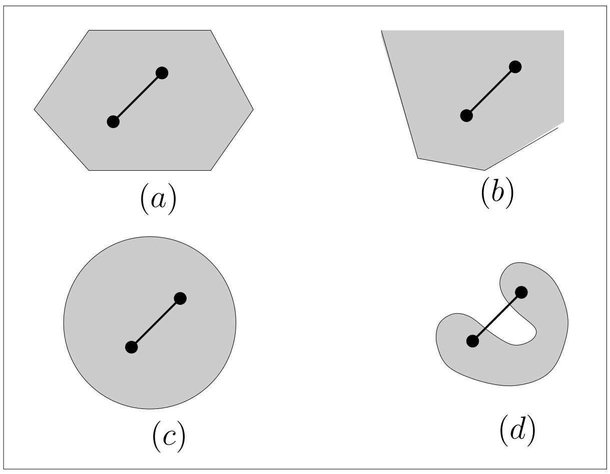

Figure A.4: Convex sets: (a), (b), (c). The set (d) is not convex.

The following results can be proved:

- A hyperplane is a convex set.

- A half-space is a convex set.
- The intersection of a finite number of convex sets is a convex set.

In linear programming, convex sets such as hyperplanes, half-spaces and the intersection of a finite number of convex sets are of special significance, because they appear in the study of linear models. The intersection of a finite number of half-spaces is a convex set of the form (a), where the vertices of the set are called *extreme points*.

A.5 Extreme points and basic feasible solutions

A set of linear inequations can be converted into a set of linear equations by adding variables. We shall now show that if the variables are restricted to take on values greater than or equal to zero, then by converting a set of linear inequations into a set of linear equations, we can find a correspondence between basic feasible solutions of linear equations and the extreme points of the set of inequations.

Consider the following set of inequations:

$$\begin{aligned} -x_1 + 4x_2 &\leq 4 \\ x_1 - x_2 &\leq 3 \end{aligned}$$

Figure A.5 represents the points which satisfy the two inequalities and the nonnegativity constraints $x_1 \geq 0$ and $x_2 \geq 0$. We can see that the intersection of the two half-spaces, together with $x_1 \geq 0$ and $x_2 \geq 0$, is a convex set; a polygon in this case. The polygon has a finite number of vertices, which are the extreme points of the set.

The point O is the origin of the coordinate system. The point A is the intersection of $-x_1 + 4x_2 = 4$ and the x_2 axis. The point B is the intersection of $-x_1 + 4x_2 = 4$ and $x_1 - x_2 = 3$. The point C is the intersection of $x_1 - x_2 = 3$ and the x_1 axis.

$$O = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad A = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad B = \begin{pmatrix} \frac{16}{3} \\ \frac{7}{3} \end{pmatrix}, \quad C = \begin{pmatrix} 3 \\ 0 \end{pmatrix}$$

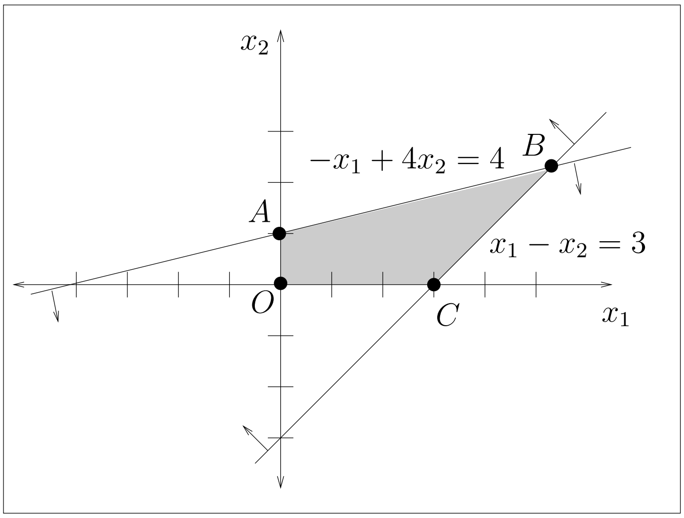

Figure A.5: A convex set and the extreme points

In general, as the example shows, in the Euclidean plane, the intersection of a finite number of half-spaces is a convex set, that is, either it is the empty set, or a set with a unique point, or a polygon with a finite number of extreme points. In the 3-dimensional Euclidean space, the intersection of a finite number of half-spaces is also a convex set, that is, either it is the empty set, it is a set with a unique point, or it is a polyhedron with a finite number of extreme points. In the n-dimensional Euclidean space, the intersection of a finite number of half-spaces is a convex set called polytope.

We can convert inequations into equations by adding nonnegative variables x_3 and x_4. We get the following system of linear equations:

$$\begin{aligned} -x_1 + 4x_2 + x_3 \quad &= 4 \\ x_1 - x_2 \quad + x_4 &= 3 \end{aligned}$$

There exists an infinite number of solutions to the system, specifically, the set of all points in the shaded region of Figure A.5. We can compute the basic solutions and choose the ones with all the variables greater than or equal to zero, so that we can confirm that they correspond to the extreme points of the polygon in Figure A.5.

1. Choosing the first and second columns in the system of equations, which are linearly independent, and setting $x_3 = x_4 = 0$, we obtain the following system:

$$\begin{aligned} -x_1 + 4x_2 &= 4 \\ x_1 - x_2 &= 3 \end{aligned}$$

   The solution of the system is $x_1 = \frac{16}{3}$, $x_2 = \frac{7}{3}$, which corresponds to the extreme point B (see Figure A.5).

2. Choosing the first and third columns in the system of equations, which are linearly independent, and setting $x_2 = x_4 = 0$, we obtain the following system:

$$\begin{aligned} -x_1 + x_3 &= 4 \\ x_1 \qquad &= 3 \end{aligned}$$

   The solution of the system is $x_1 = 3$, $x_3 = 7$, which corresponds to the extreme point C (see Figure A.5).

3. Choosing the first and fourth columns in the system of equations, which are linearly independent, and setting $x_2 = x_3 = 0$, we obtain the following system:

$$\begin{aligned} -x_1 \qquad &= 4 \\ x_1 + x_4 &= 3 \end{aligned}$$

   The solution of the system is $x_1 = -4$, $x_4 = 7$ which does not correspond to any extreme point, because it has a negative component; it violates the nonnegativity constraint.

4. Choosing the second and third columns in the system of equations, which are linearly independent, and setting $x_1 = x_4 = 0$, we obtain the following system:

$$\begin{aligned} 4x_2 + x_3 &= 4 \\ -x_2 \qquad &= 3 \end{aligned}$$

The solution of the system is $x_2 = -3$, $x_3 = 16$ which does not correspond to any extreme point, because it has a negative component; it violates the nonnegativity constraint.

5. Choosing the second and fourth columns in the system of equations, which are linearly independent, and setting $x_1 = x_3 = 0$, we obtain the following system:

$$\begin{aligned} 4x_2 \qquad &= 4 \\ -x_2 + x_4 &= 3 \end{aligned}$$

The solution of the system is $x_2 = 1$, $x_4 = 4$, which corresponds to the extreme point A (see Figure A.5).

6. Choosing the third and fourth columns in the system of equations, which are linearly independent, and setting $x_1 = x_2 = 0$, we obtain the following system:

$$\begin{aligned} x_3 \qquad &= 4 \\ x_4 &= 3 \end{aligned}$$

The solution of the system is $x_3 = 4$, $x_4 = 3$, which corresponds to the extreme point O (see Figure A.5). $\square$

The procedure used in the preceding example leads to the computation of the extreme points of the convex set in Figure A.5; the intersection of a finite number of half-spaces where all variables are restricted to take on values greater than or equal to zero. The procedure can be generalized to n-dimensional vector spaces, $n > 2$. Observe that the extreme points can be computed without representing geometrically the convex set.