Convex Sets and Convex Functions

1 Convex Sets,

In this section, we introduce one of the most important ideas in economic modelling, in the theory of optimization and, indeed in much of modern analysis and computational mathematics: that of a convex set. Almost every situation we will meet will depend on this geometric idea. As an independent idea, the notion of convexity appeared at the end of the 19th century, particularly in the works of Minkowski who is supposed to have said:

“Everything that is convex interests me.”

We discuss other ideas which stem from the basic definition, and in particular, the notion of a convex and concave functions which which are so prevalent in economic models.

The geometric definition, as we will see, makes sense in any vector space. Since, for the most of our work we deal only with \( \mathbb{R}^n \), the definitions will be stated in that context. The interested student may, however, reformulate the definitions either, in an abstract setting or in some concrete vector space as, for example, \( C([0, 1]; \mathbb{R}) \).

Intuitively if we think of \( \mathbb{R}^2 \) or \( \mathbb{R}^3 \), a convex set of vectors is a set that contains all the points of any line segment joining two points of the set (see the next figure).

Here is the definition.

**Definition 1.1** Let \( u, v \in V \). Then the set of all convex combinations of \( u \) and \( v \) is the set of points

\[
\{ w_\lambda \in V : w_\lambda = (1 - \lambda)u + \lambda v, 0 \leq \lambda \leq 1 \}. \quad (1.1)
\]

In, say, \( \mathbb{R}^2 \) or \( \mathbb{R}^3 \), this set is exactly the line segment joining the two points \( u \) and \( v \). (See the examples below.)

Next, is the notion of a convex set.

---

\( ^1 \)This symbol stands for the vector space of continuous, real-valued functions defined on the closed interval \([0, 1]\)
Definition 1.2 Let $K \subset V$. Then the set $K$ is said to be \textbf{convex} provided that given two points $u, v \in K$ the set (1.1) is a subset of $K$.

We give some simple examples:

Examples 1.3 (a) An interval of $[a, b] \subset \mathbb{R}$ is a convex set. To see this, let $c, d \in [a, b]$ and assume, without loss of generality, that $c < d$. Let $\lambda \in (0, 1)$. Then,

\[
\begin{align*}
    a & \leq c = (1 - \lambda)c + \lambda c < (1 - \lambda)c + \lambda d \\
    & < (1 - \lambda)d + \lambda d = d \\
    & \leq b.
\end{align*}
\]

(b) A disk with center $(0, 0)$ and radius $c$ is a convex subset of $\mathbb{R}^2$. This may be easily checked (Exercise!) by using the usual distance formula in $\mathbb{R}^2$ namely $\|x - y\| = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2}$ and the triangle inequality $\|u + v\| \leq \|u\| + \|v\|$.

(c) In $\mathbb{R}^n$ the set $H := \{x \in \mathbb{R}^n : a_1x_1 + \ldots + a_nx_n = c\}$ is a convex set. For any particular choice of constants $a_i$ it is a hyperplane in $\mathbb{R}^n$. Its defining equation is a generalization of the usual equation of a plane in $\mathbb{R}^3$, namely the equation $ax + by + cz + d = 0$, and so the set $H$ is called a \textbf{hyperplane}.

To see that $H$ is a convex set, let $x^{(1)}, x^{(2)} \in H$ and define $z \in \mathbb{R}^3$ by $z := (1 - \lambda)x^{(1)} + \lambda x^{(2)}$. Then
\[ z = \sum_{i=1}^{n} a_i[(1 - \lambda)x_i^{(1)} + \lambda x_i^{(2)}] = \sum_{i=1}^{n} (1 - \lambda)a_i x_i^{(1)} + \lambda a_i x_i^{(2)} \]

\[ = (1 - \lambda) \sum_{i=1}^{n} a_i x_i^{(1)} + a_i x_i^{(2)} = (1 - \lambda)c + \lambda c \]

\[ = c. \]

Hence \( z \in H \).

(d) As a generalization of the preceding example, let \( A \) be an \( m \times n \) matrix, \( b \in \mathbb{R}^m \), and let \( S := \{ x \in \mathbb{R}^n : Ax = b \} \). (The set \( S \) is just the set of all solutions of the linear equation \( Ax = b \).) Then the set \( S \) is a convex subset of \( \mathbb{R}^n \). Indeed, let \( x^{(1)}, x^{(2)} \in S \). Then

\[ A \left( (1 - \lambda)x^{(1)} + \lambda x^{(2)} \right) = (1 - \lambda)A \left( x^{(1)} \right) + \lambda A \left( x^{(2)} \right) = (1 - \lambda)b + \lambda b = b. \]

(e) There are always two, so-called trivial examples. These are the empty set \( \emptyset \), and the entire space \( \mathbb{R}^n \). Note also that a singleton \( \{ x \} \) is convex. In this latter case, as in the case of the empty set, the definition is satisfied vacuously.

**Note:** In (c) above, we can take \( A = (a_1, a_2, \ldots, a_n) \) so that with this choice, the present example is certainly a generalization.

One important fact about \( \mathbb{R}^n \) is that the unit ball

\[ B_1 = \{ c \in \mathbb{R}^n | \| x \| \leq 1 \} \]

is a convex set. This follows from the triangle equality for norms: for any \( x, y \in B_1 \) and any \( \lambda \in [0, 1] \) we have

\[ \| (1 - \lambda) x + \lambda y \| \leq (1 - \lambda) \| x \| + \lambda \| y \| \leq (1 - \lambda) + \lambda = 1. \]

Now the ball \( B_1 \) is a closed set. It is easy to see that if we take its interior

\[ \overset{\circ}{B_1} = \{ x \in \mathbb{R}^n | \| x \| < 1 \}, \]

this set is also convex. This gives us a hint regarding our next result.

**Proposition 1.4** If \( C \subset \mathbb{R}^n \) is convex, the set \( \text{cl}(C) \), the closure of \( C \), is also convex.
Proof: Suppose \( x, y \in \text{cl}(C) \). Then there exist sequences \( \{x_n\}_{n=1}^{\infty} \) and \( \{y_n\}_{n=1}^{\infty} \) in \( C \) such that \( x_n \to x \) and \( y_n \to y \) as \( n \to \infty \). For some \( \lambda, 0 \leq \lambda \leq 1 \), define \( z_n := (1 - \lambda) x_n + \lambda y_n \). Then, by convexity of \( C \), \( z_n \to C \) as \( n \to \infty \). Moreover \( z_n \to (1 - \lambda) x + \lambda y \) as \( n \to \infty \). Hence this latter point lies in \( \text{cl}(C) \).

The simple example of the two intervals \([0, 1]\) and \([2, 3]\) on the real line shows that the union of two sets is not necessarily convex. On the other hand, we have the result:

**Proposition 1.5** The intersection of any number of convex sets is convex.

**Proof:** Let \( \{K_\alpha\}_{\alpha \in A} \) be a family of convex sets, and let \( K := \bigcup_{\alpha \in A} K_\alpha \). Then, for any \( x, y \in K \) by definition of the intersection of a family of sets, \( x, y \in K_\alpha \) for all \( \alpha \in A \) and each of these sets is convex. Hence for any \( \alpha \in A \), \( (1 - \lambda) x + \lambda y \in K_\alpha \). Hence \( (1 - \lambda) x + \lambda y \in K \).

While, by definition, a set is convex provided all convex combinations of two points in the set is again in the set, it is a simple matter to check that we can make a more general statement. This statement is the content of the following proposition. Notice the way in which the proof is constructed; it is often very useful in computations!

**Proposition 1.6** Let \( K \) be a convex set and let \( \lambda_1, \lambda_2, \ldots, \lambda_p \geq 0 \) and \( \sum_{i=1}^{p} \lambda_i = 1 \). If \( x_1, x_2, \ldots, x_p \in K \) then

\[
\sum_{i=1}^{p} \lambda_i x_i \in K.
\]

**Proof:** We prove the result by induction. Since \( K \) is convex, the result is true, trivially, for \( p = 1 \) and by definition for \( p = 2 \). Suppose that the proposition is true for \( p = r \) (induction hypothesis!) and consider the convex combination \( \lambda_1 x_1 + \lambda_2 x_2 + \ldots + \lambda_{r+1} \).

Define \( \Lambda := \sum_{i=1}^{r} \lambda_i \). Then since \( 1 - \Lambda = \sum_{i=1}^{r+1} \lambda_i - \sum_{i=1}^{r} \lambda_i = \lambda_{r+1} \), we have

\[
\left( \sum_{i=1}^{r} \lambda_i x_i \right) + \lambda_{r+1} x_{r+1} = \Lambda \left( \sum_{i=1}^{r} \frac{\lambda_i}{\Lambda} x_i \right) + (1 - \Lambda) x_{r+1}.
\]

Note that \( \sum_{i=1}^{r} \left( \frac{\lambda_i}{\Lambda} \right) = 1 \) and so, by the induction hypothesis, \( \sum_{i=1}^{r} \left( \frac{\lambda_i}{\Lambda} \right) x_i \in K \). Since \( x_{r+1} \in K \) it follows that the right hand side is a convex combination of two points of \( K \) and hence lies in \( K \).

Relative to the vector space operations, we have the following result:
Proposition 1.7 Let $C, C_1$ and $C_2$ be convex sets in $\mathbb{R}^n$ and let $\beta \in \mathbb{R}$. Then

(a) The set $\beta C := \{z \in \mathbb{R}^n \mid z = \beta x, x \in C\}$ is convex.

(b) The set $C_1 + C_2 := \{x \in \mathbb{R}^n \mid z = x_1 + x_2, x_1 \in C_1, x_2 \in C_2\}$ is convex.

Proof: For part (a), let $z_1$ and $z_2$ both be elements of $\beta C$. Then there exists points $x_1, x_2 \in C$ such that $z_1 = \beta x_1$ and $z_2 = \beta x_2$. Choose any $\lambda \in [0, 1]$ and form the convex combination

$$z = (1 - \lambda) z_1 + \lambda z_2.$$

But then

$$z = (1 - \lambda) \beta x_1 + \lambda \beta x_2 = \beta [(1 - \lambda) x_1 + \lambda x_2].$$

But $C$ is convex so that $(1 - \lambda) x_1 + \lambda x_2 \in C$ and hence $z \in \beta C$. This proves part (a).

Part (b) is proved by a similar argument by simply noting that

$$(1 - \lambda) (x_1 + x_2) + \lambda (y_1 + y_2) = (1 - \lambda) x_1 + (1 - \lambda) x_2 + \lambda y_1 + \lambda y_2.$$

For any given set which is not convex, we often want to find a set which is convex and which contains the set. Since the entire vector space $V$ is obviously a convex set, there is always at least one such convex set containing the given one. In fact, there are infinitely many such sets. We can make a more economical choice if we start with the following fact.

Intuitively, given a set $C \subset V$, the intersection of all convex sets containing $C$ is the “smallest” subset containing $C$. We make this into a definition.

Definition 1.8 The convex hull of a set $C$ is the intersection of all convex sets which contain the set $C$. We denote the convex hull by $\text{co} (C)$.

We illustrate this definition in the next figure where the dotted line together with the original boundaries of the set for the boundary of the convex hull.
Examples 1.9  (a) Suppose that \([a, b]\) and \([c, d]\) are two intervals of the real line with \(c < b\) so that the intervals are disjoint. Then the convex hull of the set \([a, b] \cup [c, d]\) is just the interval \([a, d]\).

(b) In \(\mathbb{R}^2\) consider the elliptical annular region \(\mathcal{E}\) consisting of the disk \(\{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq R\}\) for a given positive number \(R\) without the interior points of an elliptical region as indicated in the next figure.

Figure 3: The Formation of a Convex Hull

Figure 4: The Elliptically Annular Region

Clearly the set \(\mathcal{E}\) is not convex for the line segment joining the indicated points \(P\) and \(Q\) has points lying in the “hole” of region and hence not in \(\mathcal{E}\). Indeed, this is the case for any line segment joining two points of the region which are,
say, symmetric with respect to the origin. Clearly the entire disk of radius \( R \) is convex and indeed is the convex hull, \( \text{co}(\mathcal{E}) \).

These examples are typical. In each case, we see that the convex hull is obtained by adjoining all linear combinations of points in the original set. This is indeed a general result.

**Theorem 1.10** Let \( C \subset V \). Then the set of all convex combinations of points of the set \( C \) is exactly \( \text{co}(C) \).

**Proof:** Let us denote the set of all convex combinations of points of \( C \) by \( L(C) \). It is clear from Proposition 1.6 that \( L(C) \supset \text{co}(C) \). To see that the opposite inclusion holds, simply observe that if \( K \supset C \) is convex, then it must contain all the convex combinations of points of \( C \) and hence \( L(C) \subset K \). From this it follows that \( L(C) \) is a convex set, containing \( C \) and contained in every convex subset that contains \( C \), hence \( L(C) \subset \text{co}(C) \).

Convex sets in \( \mathbb{R}^n \) have a very nice characterization discovered by Carathéodory.

**Theorem 1.11** Let \( C \) be a subset of \( \mathbb{R}^n \). Then every element of \( \text{co}(C) \) can be represented as a convex combination of no more than \( (n + 1) \) elements of \( C \).

**Proof:** Let \( x \in \text{co}(C) \). Then \( x \) is a convex combination of points of \( C \), and we write

\[
x = \sum_{i=1}^{m} \alpha_i x^{(i)}, \quad x^{(i)} \in C, \quad \alpha_i \geq 0, \quad \sum_{i=1}^{m} \alpha_i = 1.
\]

Let us assume that \( m \) is the *minimal* number of vectors for which such a representation is possible. In particular, this means that, for all \( i = 1, \ldots, m \) we have \( \alpha_i > 0 \), since if not, the number \( m \) would not be minimal.

Now, if \( m \leq n+1 \) there is nothing to prove. On the other hand, suppose that \( m > n+1 \). Then the vectors \( x^{(i)} - x^{(1)}, \ i = 2, \ldots, m \), must be linearly dependent since \( m - 1 > n \). This means that there are scalars \( \beta_i, i = 2, \ldots, m \), not all zero, such that

\[
\sum_{i=2}^{m} \beta_i (x^{(i)} - x^{(1)}) = 0.
\]

Now define \( \beta_1 := -\sum_{i=2}^{m} \beta_i \). Then,
\[ \sum_{i=1}^{m} \beta_i x^{(i)} = \beta_1 x^{(1)} + \sum_{i=2}^{m} \beta_i x^{(i)} = -\sum_{i=2}^{m} \beta_i x^{(1)} + \sum_{i=2}^{m} \beta_i x^{(i)} = 0, \]

and \( \sum \beta_i = 0 \). Hence at least one \( \beta_i > 0 \) since this is a sum of numbers which are not all zero.

Then, introducing a real parameter \( \lambda \),

\[ x = \sum_{i=1}^{m} \alpha_i x^{(i)} - \lambda \sum_{i=1}^{m} \beta_i x^{(i)} = \sum_{i=1}^{m} (\alpha_i - \lambda \beta_i) x^{(i)}. \]

which we can do since \( \sum_{i=1}^{m} \beta_i x^{(i)} = 0 \).

Now, recalling that all the \( \alpha_i > 0 \), let \( \hat{\lambda} \) be given by

\[ \hat{\lambda} := \min_{1 \leq i \leq m} \left\{ \frac{\alpha_i}{\beta_i} \mid \beta_i > 0 \right\} = \frac{\alpha_j}{\beta_j}. \]

From the definition, \( \hat{\lambda} > 0 \), and, for every \( i, \ 1 \leq i \leq m \), we have \( \alpha_i - \hat{\lambda} \beta_i \geq 0 \), with, for \( i = j \), \( \alpha_j - \hat{\lambda} \beta_j = 0 \).

Therefore,

\[ x = \sum_{i=1}^{m} (\alpha_i - \hat{\lambda} \beta_i) x^{(i)}, \]

where, for every \( i, (\alpha_i - \hat{\lambda} \beta_i) \geq 0 \) and

\[ \sum_{i=1}^{m} (\alpha_i - \hat{\lambda} \beta_i) = \left( \sum_{i=1}^{m} \alpha_i \right) - \hat{\lambda} \left( \sum_{i=1}^{m} \beta_i \right) = \sum_{i=1}^{m} \alpha_i = 1, \]

so that, since one of these (the \( j^{th} \)) vanishes, we have a convex combination of fewer than \( m \) points which contradicts the minimality of \( m \).

\[ \square \]

Carathéodory’s Theorem has the nature of a representation theorem somewhat analogous to the theorem which says that any vector in a vector space can be represented as a linear combination of the elements of a basis. One thing both theorems do, is to give a finite and minimal representation of all elements of an infinite set. The drawback of Carathéodory’s Theorem, unlike the latter representation, is that the choice of elements used to represent the point is neither uniquely determined for that point, nor does the theorem guarantee that the same set of vectors in \( C \) can be used to
represent all vectors in $C$; the representing vectors will usually change with the point being represented. Nevertheless, the theorem is useful in a number of ways we will see presently. First, a couple of examples.

**Examples 1.12**  
(a) Recalling Example 1.7(a), in $\mathbb{R}$, consider the interval $[0, 1]$ and the subinterval $\left(\frac{1}{4}, \frac{3}{4}\right)$. Then $\text{co} \left(\frac{1}{4}, \frac{3}{4}\right) = \left[\frac{1}{4}, \frac{3}{4}\right]$. If we take the point $x = \frac{1}{2}$, then we have both

$$x = \frac{1}{2} \left(\frac{3}{8}\right) + \frac{1}{2} \left(\frac{5}{8}\right) \quad \text{and} \quad x = \frac{3}{4} \left(\frac{7}{16}\right) + \frac{1}{4} \left(\frac{11}{16}\right).$$

So that certainly there is no uniqueness in the representation of $x = \frac{1}{2}$.

(b) In $\mathbb{R}^2$ we consider the two triangular regions $T_1$, $T_2$, joining the points $(0, 0)$, $(1, 4)$, $(0, 2)$, $(3, 4)$ and $(4, 0)$ as pictured in the next figures. The second of the figures indicates that joining the apexes of the triangles forms a trapezoid which is a convex set. It is the convex hull of the set $T_1 \cup T_2$.

![Figure 5: The Triangles $T_1$ and $T_2$](image1)  
![Figure 6: The Convex Hull of $T_1 \cup T_2$](image2)

Again, it is clear that two points which both lie in one of the original triangles have more than one representation. Similarly, if we choose two points, one from $T_1$ and one from $T_2$, say the points $(1, 2)$ and $(3, 2)$, the point

9
\[ \frac{1}{2} \begin{pmatrix} 1 \\ 2 \end{pmatrix} + \begin{pmatrix} 3 \\ 2 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \end{pmatrix} \]
does not lie in the original set \( T_1 \cup T_2 \), but does lie in the convex hull. Moreover, this point can also be represented by

\[ \frac{1}{3} \begin{pmatrix} 3 \\ \frac{3}{2} \end{pmatrix} + \frac{2}{3} \begin{pmatrix} 9 \\ \frac{7}{4} \end{pmatrix} \]
as can easily be checked.

The next results depend on the notion of norm in \( \mathbb{R}^n \) and on the convergence of a sequence of points in \( \mathbb{R}^n \). In particular, it relies on the fact that, in \( \mathbb{R}^n \), or for that matter in any complete metric space, Cauchy sequences converge.

Recall that a set of points in \( \mathbb{R}^n \) is called **compact** provided it is closed and bounded. One way to characterize such a set in \( \mathbb{R}^n \) is that if \( C \subset \mathbb{R}^n \) is compact, then, given any sequence \( \{x^{(1)}, x^{(2)}, \ldots, x^{(n)}, \ldots \} \subset C \) there is a subsequence which converges,

\[ \lim_{k \to \infty} x^{(n_k)} = x^o \in C. \]
As a corollary to Carathéodory’s Theorem, we have the next result about compact sets:

**Corollary 1.13** The convex hull of a compact set in \( \mathbb{R}^n \) is compact.

**Proof:** Let \( C \subset \mathbb{R}^n \) be compact. Notice that the simplex

\[ \sigma := \left\{ (\lambda_1, \lambda_2, \ldots, \lambda_n) \in \mathbb{R}^n : \sum_{i=1}^{n} \lambda_i = 1 \right\} \]
is also closed and bounded and is therefore compact. (Check!) Now suppose that \( \{v^{(j)}\}_{j=1}^{\infty} \subset \text{co}(C) \). By Carathéodory’s Theorem, each \( v^{(j)} \) can be written in the form

\[ v^{(k)} = \sum_{i=1}^{n+1} \lambda_{k,i} x^{(k,i)}, \text{ where } \lambda_{k,i} \geq 0, \sum_{i=1}^{n+1} \lambda_{k,i} = 1, \text{ and } x^{(k,i)} \in C. \]

Then, since \( C \) and \( \sigma \) are compact, there exists a sequence \( k_1, k_2, \ldots \) such that the limits \( \lim_{j \to \infty} \lambda_{k_j,i} = \lambda_i \) and \( \lim_{j \to \infty} x^{(k_j,i)} = x^{(i)} \) exist for \( i = 1, 2, \ldots, n + 1 \). Clearly \( \lambda_i \geq 0, \sum_{i=1}^{n+1} \lambda_i = 1 \) and \( x_i \in C \).

Thus, the sequence \( \{v^{(k)}\}_{k=1}^{\infty} \) has a subsequence, \( \{v^{(k_j)}\}_{j=1}^{\infty} \) which converges to a point of \( \text{co}(C) \) which shows that this latter set is compact. \( \square \)
The next result shows that if \( C \) is closed and convex (but perhaps not bounded) is has a smallest element in a certain sense. It is a simple result from analysis that involves the facts that the function \( x \rightarrow \|x\| \) is a continuous map from \( \mathbb{R}^n \rightarrow \mathbb{R} \) and, again, that Cauchy sequences in \( \mathbb{R}^n \) converge. It also relies heavily on the parallelogram law for norms.

**Theorem 1.14** Every closed convex subset of \( \mathbb{R}^n \) has a unique element of minimum norm.

**Proof:**

Let \( K \) be such a set and note that \( \iota := \inf_{x \in K} \|x\| \geq 0 \) so that the function \( x \rightarrow \|x\| \) is bounded below on \( K \). Let \( x^{(1)}, x^{(2)}, \ldots \) be a sequence of points of \( K \) such that

\[
\lim_{i \to \infty} \|x^{(i)}\| = \iota. \tag{2}
\]

Then, by the parallelogram law,

\[
\|x^{(i)} - x^{(j)}\|^2 = 2\|x^{(i)}\|^2 + 2\|x^{(j)}\|^2 - 4\left\| \frac{1}{2} (x^{(i)} + x^{(j)}) \right\|^2.
\]

Since \( K \) is convex, \( \frac{1}{2} (x^{(i)} + x^{(j)}) \in K \) so that \( \| \frac{1}{2} (x^{(i)} + x^{(j)}) \| \geq \iota^2 \). Hence

\[
\|x^{(i)} - x^{(j)}\|^2 \leq 2\|x^{(i)}\|^2 + 2\|x^{(j)}\|^2 - 4\iota^2.
\]

As \( i, j \to \infty \), we have \( 2\|x^{(i)}\|^2 + 2\|x^{(j)}\|^2 - 4\iota \to 0 \). Thus, \( \{x^{(j)}\}_{j=1}^{\infty} \) is a Cauchy sequence and has a limit point \( x \). Since \( K \) is closed, \( x \in K \). Moreover, since the function \( x \rightarrow \|x\| \) is a continuous function from \( \mathbb{R}^n \rightarrow \mathbb{R} \),

\[
\iota = \lim_{j \to \infty} \|x^{(j)}\| = \|x\|.
\]

In order to show uniqueness of the point with minimal norm, suppose that there were two points, \( x, y \in K, x \neq y \) such that \( \|x\| = \|y\| = \iota \). Then by the parallelogram law,

\[
0 < \|x - y\|^2 = 2\|x\|^2 + 2\|y\|^2 - 4\left\| \frac{1}{2} (x^{(i)} + x^{(j)}) \right\|^2
\]

\[
= 2\iota^2 + 2\iota^2 - 4\left\| \frac{1}{2} (x^{(i)} + x^{(j)}) \right\|^2
\]

so that \( 4\iota > 4\left\| \frac{1}{2} (x^{(i)} + x^{(j)}) \right\|^2 \) or \( \frac{1}{2} (x^{(i)} + x^{(j)}) \| > \iota \) which would give a vector in \( K \) of norm less than the infimum \( \iota \).

\[\square\]

\[\text{Here, and throughout this course, we shall call such a sequence a minimizing sequence.}\]
Example 1.15 It is easy to illustrate the statement of this last theorem in a concrete case. Suppose that we define three sets in $\mathbb{R}^2$ by $H^+_1 := \{(x,y) \in \mathbb{R}^2 : 5x - y \geq 1\}$, $H^+_2 := \{(x,y) \in \mathbb{R}^2 : 2x + 4y \geq 7\}$ and $H^+_3 := \{(x,y) \in \mathbb{R}^2 : 2x + 2y \geq 6\}$ whose intersection (the intersection of half-spaces) forms a convex set illustrated below. The point of minimal norm is the closest point in this set to the origin. From the projection theorem in $\mathbb{R}^2$, that point is determined by the intersection of the boundary line $2x + 4y = 6$ with a line perpendicular to it and which passes through the origin as illustrated here.

\[\|z\|\] is a minimum. Let $u \in U$. By the convexity of the convex hull of $U$,

2 Separation Properties

There are a number of results which are of crucial importance in the theory of convex sets, and in the theory of mathematical optimization, particularly with regard to the development of necessary conditions, as for example in the theory of Lagrange multipliers. They also play a central role in the treatment of equilibrium models, in game theory, and in the analysis of production plans.
The results we refer to are usually lumped together under the rubric of *separation theorems.* We discuss three of these results which will be useful to us. We will confine our attention to the finite dimensional case.\(^3\)

Our proof of the Separation Theorem (and its corollaries) depends on the idea of the projection of a point onto a convex set. We begin by proving that result which is similar to the problem of finding a point of minimum norm in a closed convex set.

The statement of the Projection Theorem is as follows:

**Theorem 2.1** Let \( C \subset \mathbb{R}^n \) be a closed, convex set. Then

(a) For every \( x \in \mathbb{R}^n \) there exists a unique vector \( z^* \in C \) that minimizes \( \|z-x\| \) over all \( z \in C \). We call \( z^* \) the projection of \( x \) onto \( C \).

(b) \( z^* \) is the projection of \( x \) onto \( C \) if and only if

\[
\langle y - z^*, x - z^* \rangle \leq 0, \text{ for all } y \in C.
\]

**Proof:** Fix \( x \in \mathbb{R}^n \) and let \( w \in C \). Then minimizing \( \|x-z\| \) over all \( z \in C \) is equivalent to minimizing the same function over the set \( \{ z \in C : \|x-z\| \leq \|x-w\| \} \). This latter set is both closed and bounded and therefore the continuous function \( g(z) := \|z-x\| \), according to the theorem of Weierstrass, takes on its minimum at some point of the set.

We use the parallelogram identity to prove uniqueness as follows. Suppose that there are two distinct points, \( z_1 \) and \( z_2 \), which both minimize \( \|z-x\| \) and denote this minimum by \( \iota \). Then we have

\[
0 < \|(z_1 - x) - (z_2 - x)\|^2 = 2\|z_1 - x\|^2 + 2\|z_2 - x\|^2 - 4\left\| \frac{1}{2}[(z_1 - x) + (z_2 - x)] \right\|^2
\]

\[
= 2\|z_1 - x\|^2 + 2\|z_2 - x\|^2 - 4\left\| \frac{z_1 + z_2}{2} - x \right\|^2 = 2\iota^2 + 2\iota^2 - 4\|\hat{z} - x\|^2,
\]

where \( \hat{z} = (z_1 + z_2)/2 \in C \) since \( C \) is convex. Rearranging, and taking square roots, we have

\[
\|\hat{z} - x\| < \iota
\]

\(^3\)In a general inner product space or in Hilbert space the basic theorem is called the Hahn-Banach theorem and is one of the central results of functional analysis.
which is a contradiction of the fact that \( z_1 \) and \( z_2 \) give minimal values to the distance. Thus uniqueness is established.

To prove the inequality in part (b), and using \( \langle \cdot, \cdot \rangle \) for the inner product, we have, for all \( y, z \in C \), the inequality

\[
\|y - x\|^2 = \|y - z\|^2 + \|z - x\|^2 - 2 \langle (y - z), (x - z) \rangle \\
\geq \|z - x\|^2 - 2 \langle (y - z), (x - z) \rangle .
\]

Hence, if \( z \) is such that \( \langle (y - z), (x - z) \rangle \leq 0 \) for all \( y \in C \), then \( \|y - x\|^2 \geq \|z - x\|^2 \) for all \( y \in C \) and so, by definition \( z = z^* \).

To prove the necessity of the condition, let \( z^* \) be the projection of \( x \) onto \( C \) and let \( y \in C \) be arbitrary. For \( \alpha > 0 \) define \( y_\alpha = (1 - \alpha)z^* + \alpha y \) then

\[
\|x - y_\alpha\|^2 = \|(1 - \alpha)(x - z^*) + \alpha(x - y)\|^2 \\
= (1 - \alpha)^2 \|x - z^*\|^2 + \alpha^2 \|x - y\|^2 + 2 (1 - \alpha) \alpha \langle (x - z^*), (x - y) \rangle .
\]

Now consider the function \( \varphi(\alpha) := \|x - y_\alpha\|^2 \). Then we have from the preceding result

\[
\left. \frac{\partial \varphi}{\partial \alpha} \right|_{\alpha=0} = -2 \|x - z^*\|^2 + 2 \langle (x - z^*), (x - y) \rangle = -2 \langle (y - z^*), (x - z^*) \rangle .
\]

Therefore, if \( \langle (y - z^*), (x - z^*) \rangle > 0 \) for some \( y \in C \), then

\[
\left. \frac{\partial}{\partial \alpha} \left\{ \|x - y_\alpha\|^2 \right\} \right|_{\alpha=0} < 0
\]

and, for positive but small enough \( \alpha \), we have \( \|x - y_\alpha\| < \|x - z^*\| \). This contradicts the fact that \( z^* \) is the projection of \( x \) onto \( C \) and shows that \( \langle (y - z^*), (x - z^*) \rangle \leq 0 \) for all \( y \in C \).

In order to study the separation properties of convex sets we need the notion of a hyperplane.

**Definition 2.2** Let \( a \in \mathbb{R}^n \) and \( b \in \mathbb{R} \) and assume \( a \neq 0 \). Then the set

\[
H := \{ x \in \mathbb{R}^n \mid \langle a, x \rangle = b \},
\]

is called a hyperplane with normal vector \( a \).
We note, in passing, that hyperplanes are convex sets, a fact that follows from the bilinearity of the inner product. Indeed, if \( x, y \in H \) and \( 0 \leq \lambda \leq 1 \) we have

\[
\langle a, (1 - \lambda) x + \lambda y \rangle = (1 - \lambda) \langle a, x \rangle + \lambda \langle a, y \rangle = (1 - \lambda) b + \lambda b = b.
\]

Each such hyperplane defines the half-spaces

\[
H^+ := \{ x \in \mathbb{R}^n \mid \langle a, x \rangle \geq b \}, \quad \text{and} \quad H^- := \{ x \in \mathbb{R}^n \mid \langle a, x \rangle \leq b \}.
\]

Note that \( H^+ \cap H^- = H \). These two half-spaces are closed sets, that is, they contain all their limit points. Their interiors \(^4\) are given by

\[
H^o^+ := \{ x \in \mathbb{R}^n \mid \langle a, x \rangle > b \}, \quad \text{and} \quad H^o^- := \{ x \in \mathbb{R}^n \mid \langle a, x \rangle < b \}.
\]

Whether closed or open, these half-spaces are said to be generated by the hyperplane \( H \). Each is a convex set.

We can now describe what properties, relative to a convex set, the hyperplanes are to have.

**Definition 2.3** Let \( S, T \subset \mathbb{R}^n \) and let \( H \) be a hyperplane. Then \( H \) is said to separate \( S \) from \( T \) if \( S \) lies in one closed half-space determined by \( H \) while \( T \) lies in the other closed half-space. In this case \( H \) is called a separating hyperplane. If \( S \) and \( T \) lie in the open half-spaces, then \( H \) is said to strictly separate \( S \) and \( T \). If \( H \) contains one or more boundary points of a convex set and all other points lie in one of the two half-spaces, then \( H \) is said to support the convex set at those points.

We are now in a position to prove three theorems, what we will call the Basic Separation Theorem, the Supporting Hyperplane Theorem, and the Strict Separation Theorem. In particular we will prove the Basic Separation Theorem in \( \mathbb{R}^n \) using the result concerning the projection of a point on a convex set. The second of the theorems, the Basic Support Theorem, is a corollary of the first.

**Theorem 2.4** Let \( C \subset \mathbb{R}^n \) be convex and suppose that \( y \notin \text{cl}(C) \). Then there exist an \( a \in \mathbb{R}^n \) and a number \( \gamma \in \mathbb{R} \) such that \( \langle a, x \rangle > \gamma \) for all \( x \in C \) and \( \langle a, y \rangle \leq \gamma \).

\(^4\)Recall that a point \( x_o \) is an interior point of a set \( S \subset \mathbb{R}^n \) provided there is a ball with center \( x_o \) of sufficiently small radius, which contains only points of \( S \). The interior of a set consists of all the interior points (if any) of that set.
Proof: Let \( \hat{c} \) be the projection of \( y \) on \( \ell(C) \), and let \( \gamma = \inf_{x \in C} \| x - y \| \), i.e., \( \gamma \) is the distance from \( y \) to its projection \( \hat{c} \). Note that since \( y \notin \ell(C) \), \( \gamma > 0 \).

Now, choose an arbitrary \( x \in C \) and \( \lambda \in (0, 1) \) and form \( x_\lambda := (1 - \lambda)\hat{c} + \lambda x \). Since \( x \in C \) and \( \hat{c} \in \ell(C) \), \( x_\lambda \in \ell(C) \). So we have, by definition of the projection,

\[
\| x_\lambda - y \|^2 = \| (1 - \lambda)\hat{c} + \lambda x - y \|^2 = \| \hat{c} + \lambda(x - \hat{c}) - y \|^2 \\
\geq \| \hat{c} - y \|^2 > 0.
\]

But we can write \( \| \hat{c} + \lambda(x - \hat{c}) - y \|^2 = \| (\hat{c} - y) + \lambda(x - \hat{c}) \|^2 \) and we can expand this latter expression using the rules of inner products.

\[
0 < \| \hat{c} - y \|^2 \leq \| (\hat{c} - y) + \lambda(x - \hat{c}) \|^2 = \langle (\hat{c} - y) + \lambda(x - \hat{c}), (\hat{c} - y) + \lambda(x - \hat{c}) \rangle \\
= \langle (\hat{c} - y), (\hat{c} - y) \rangle + \langle (\hat{c} - x), \lambda(x - \hat{c}) \rangle + \langle \lambda(x - \hat{c}), (\hat{c} - y) \rangle + \langle \lambda(x - \hat{c}), \lambda(x - \hat{c}) \rangle,
\]

and from this inequality, we deduce that

\[
2\lambda \langle (\hat{c} - y), (x - \hat{c}) \rangle + \lambda^2 \langle (x - \hat{c}), (x - \hat{c}) \rangle \geq 0,
\]

since \( \| \hat{c} - y \| > 0 \).

From this last inequality, dividing both sides by \( 2\lambda \) and taking a limit as \( \lambda \to 0^+ \), we obtain

\[
\langle \hat{c} - y, x - \hat{c} \rangle \geq 0.
\]

Again, we can expand the last expression \( \langle \hat{c} - y, x - \hat{c} \rangle = \langle \hat{c} - y, x \rangle + \langle \hat{c} - y, -\hat{c} \rangle \geq 0 \).

By adding and subtracting \( y \) and recalling that \( \| \hat{c} - y \| > 0 \), we can make the following estimate,

\[
\langle \hat{c} - y, x \rangle \geq \langle \hat{c} - y, \hat{c} \rangle = \langle \hat{c} - y, y - y + \hat{c} \rangle \\
= \langle \hat{c} - y, y \rangle + \langle \hat{c} - y, \hat{c} - y \rangle = \langle \hat{c} - y, y \rangle + \| \hat{c} - y \|^2 > \langle \hat{c} - y, y \rangle.
\]

In summary, we have \( \langle \hat{c} - y, x \rangle > \langle \hat{c} - y, y \rangle \) for all \( x \in C \). Now, we need only define \( a := \hat{c} - y \). Then this last inequality reads

\[
\langle a, x \rangle > \langle a, y \rangle \quad \text{for all} \quad x \in C.
\]
Before proving the Basic Support Theorem, let us recall some terminology. Suppose that \( S \subset \mathbb{R}^n \). Then a point \( s \in S \) is called a boundary point of \( S \) provided that every ball with \( s \) as center intersects both \( S \) and its complement \( \mathbb{R}^n \setminus S := \{ x \in \mathbb{R}^n : x \not\in S \} \). Note that every boundary point is a limit point of the set \( S \), but that the converse is not true.

**Theorem 2.5** Let \( C \) be a convex set and let \( y \) be a boundary point of \( C \). Then there is a hyperplane containing \( y \) and containing \( C \) in one of its half spaces.

**Proof:** Let \( \{ y^{(k)} \}_{k=1}^{\infty} \) be a sequence in \( \mathbb{R}^n \setminus c\ell(C) \) with \( y^{(k)} \to y \) as \( k \to \infty \). Let \( \{ a^{(k)} \}_{k=1}^{\infty} \) be the sequence of vectors constructed in the previous theorem and define \( \hat{a}^{(k)} := a^{(k)}/\|a^{(k)}\| \). Then, for each \( k \), \( \langle \hat{a}^{(k)}, y^{(k)} \rangle < \inf_{x \in C} \langle \hat{a}^{(k)}, x \rangle \).

Since the sequence \( \{ \hat{a}^{(k)} \}_{k=1}^{\infty} \) is bounded, it contains a convergent subsequence \( \{ \hat{a}^{(k_j)} \}_{j=1}^{\infty} \) which converges to a limit \( \hat{a}^{(o)} \). Then, for any \( x \in C \),

\[
\langle a^{(o)}, y \rangle = \lim_{j \to \infty} \langle \hat{a}^{(k_j)}, y^{(k_j)} \rangle \leq \lim_{j \to \infty} \langle \hat{a}^{(k_j)}, x \rangle = \langle \hat{a}^{(o)}, x \rangle.
\]

Before proving the last of our separation results, we need a lemma concerning Minkowski sums. It is not true, in general, that the sum of two closed, convex sets is closed.

**Lemma 2.6** The Minkowski sum of two closed, convex sets at least one of which is compact, is a closed convex set.

**Proof:** Let \( C_1 \) and \( C_2 \) be closed, convex sets and suppose that \( C_2 \) is compact. We have seen that the Minkowski sum \( C_1 + C_2 \) is convex. Here we need only prove that it is closed.

We will show that every convergent sequence in the Minkowski sum converges to a point of \( C_1 + C_2 \). Hence this set is closed. To this end, suppose \( \{ x_1^{(k)} + x_2^{(k)} \}_{k=1}^{\infty} \) is a sequence in the Minkowski sum with \( \{ x_1^{(k)} \}_{k=1}^{\infty} \subset C_1 \) and \( \{ x_2^{(k)} \}_{k=1}^{\infty} \subset C_2 \). Since \( C_2 \) is compact the sequence \( \{ x_2^{(k)} \}_{k=1}^{\infty} \) is bounded. Then, since \( \{ x_1^{(k)} + x_2^{(k)} \}_{k=1}^{\infty} \) converges, the sequence \( \{ x^{(k)} \}_{k=1}^{\infty} \) must also be bounded. Hence the sequence of pairs \( \{ x_1^{(k)} m x_2^{(k)} \}_{k=1}^{\infty} \) is bounded and hence has a limit point \( (\bar{x}_1, \bar{x}_2) \) with \( \bar{x}_1 \in C_1 \) and \( \bar{x}_2 \in C_2 \) since both of the sets are closed. The vector \( \bar{x}_1 + \bar{x}_2 \) is the limit of the sequence \( \{ x_1^{(k)} + x_2^{(k)} \}_{k=1}^{\infty} \) and hence \( \bar{x}_1 + \bar{x}_2 \in C_1 + C_2 \) which was what we were to prove. \( \square \)

We now prove the Strict Separation Theorem.
**Theorem 2.7** If $C_1$ and $C_2$ are two, non-empty and disjoint convex sets with $C_1$ closed and $C_2$ compact, then there exists a hyperplane that strictly separates them. In other words there is a vector $a \neq 0$ and $b \in \mathbb{R}$ such that

$$\langle a, x_1 \rangle < b < \langle a, x_2 \rangle \text{ for all } x_1 \in C_1 \text{ and } x_2 \in C_2.$$ 

**Proof:** Consider the problem

$$\min \|x_1 - x_2\| \text{ subject to } x_1 \in C_1, \, x_2 \in C_2$$

From the lemma, the set $C = C_1 - C_2$ is convex and closed. So the minimization problem is equivalent to finding the closest point of $C$ to the origin. Let $(x^*_1 - x^*_2)$ be the point of minimal norm with $x^*_1 \in C_1$ and $x^*_2 \in C_2$. Now set

$$a = \frac{x^*_2 - x^*_1}{2}, \quad x^* = \frac{x^*_1 + x^*_2}{2}, \quad b = \langle a, x^* \rangle.$$ 

Then $a \neq 0$ since the sets $C_1$ and $C_2$ are disjoint and $x^*_1 \in C_1$ and $x^*_2 \in C_2$.

Then the hyperplan $H$ given by $H = \{x \in \mathbb{R}^n \mid \langle a, x \rangle = b\}$ contains the point $x^*$. Then $x^*_1$ is the projection of $x^*$ onto $C_1$ and $x^*_2$ is the projection of $x^*$ onto $C_2$. Then, by the second part of the Projection Theorem,

$$\langle (x^* - x^*_1), (x_1 - x^*_1) \rangle \leq 0 \text{ for all } x_1 \in C_1,$$

or equivalently, since $x^* - x^*_1 = a$,

$$\langle a, x_1 \rangle \leq \langle a, x^*_1 \rangle + \langle a, x^*_1 - x^* \rangle = b - \|a\|^2 < b$$

for all $x_1 \in C_1$. Hence the left-hand side of the conclusion of the theorem is established. The right-hand side is established similarly. \qed

Finally, we have a result of Minkowski that characterises closed convex sets.

**Corollary 2.8** Every closed non-trivial convex subset of $\mathbb{R}^n$ is the intersection of all half-spaces containing it.

**Proof:** Let $C$ be such a set and let $\{H_\alpha\}_{\alpha \in A}$ be the family of all half-spaces containing the set $C$. Since $C \neq \mathbb{R}^n$ this family is non-empty and moreover neither is the intersection of all the family since $C$ is in each. So we need only show

$$\bigcap_{\alpha \in A} H_\alpha \subset C.$$
To see this, suppose that \( x \notin C \). Then the set \( \{x\} \) is closed bounded and convex so we may apply the strict separation result. Hence there exists a half-space containing \( C \) but not containing \( x \). Hence if \( x \notin C \), then \( x \) cannot be in the intersection of all the half-spaces containing \( C \).

We conclude this section by making some remarks about this last result. Up to this point, we have concentrated on what may be called an “intrinsic” definition of a convex set. Thus, for example, we describe a convex set as the convex hull of some set of points, or we think of a given convex set as generated by convex combinations of points in the set as described by the result of Carathéodory. Here, in the result of Minkowski, we see what is called a dual description of a convex set as the intersection of all half spaces that contain the set. In some sense this is an “extrinsic” characterization of the set in question. This is the first introduction of the notion of duality, a notion that plays a crucial role in linear programming, in convex analysis, as well as in the theory of infinite dimensional spaces. We will meet duality presently.

### 3 The Farkas-Minkowski Theorem

The results presented below, the first of which appeared in 1902, are concerned with the existence of non-negative solutions of the linear system

\[
\begin{align*}
Ax &= b, \\
x &\geq 0,
\end{align*}
\]  

(3.2)

(3.3)

where \( A \) is an \( m \times n \) matrix with real entries, \( x \in \mathbb{R}^n, b \in \mathbb{R}^m \). Here is a basic statement of the theorem which is due to Farkas and Minkowski.

**Theorem 3.1** A necessary and sufficient condition that (3.2-3.3) has a solution is that, for all \( y \in \mathbb{R}^m \) with the property that \( A^\top y \geq 0 \), we have \( \langle b, y \rangle \geq 0 \).

This theorem may be reformulated as an alternative theorem, in other words, as a theorem asserting that one set of equalities and/or inequalities has a solution if and only if another set does not. It is easy to see that the following statement is equivalent to the first:

**Theorem 3.2** The system (3.2-3.3) has a solution if and only if the system
\[ A^\top y \geq 0, \]
\[ \langle b, y \rangle < 0, \]

has no solution.

There are a number of ways to prove Theorem 3.1. One way is to use the duality theorem of linear programming. Since the Farkas-Minkowski Theorem is used in some discussions of linear programming, it is useful to have an independent proof even if it may be less elementary in the sense that it uses a separation theorem. This is the proof which we will present below. Once established, it can then be used to prove the Duality Theorem of Linear Programming.

Before starting, it is useful to recall some facts about cones in \( \mathbb{R}^n \).

The definition of a cone shows that it is a collection of half-lines emanating from the origin. The origin itself may, or may not be in the cone. It is an elementary fact, as we have seen, that an arbitrary intersection of cones is again a cone.

Now, let \( x^{(1)}, \ldots, x^{(k)} \) be any \( k \) elements of \( \mathbb{R}^n \). We look at all vectors of the form \( \sum_{i=1}^{k} \mu_i x^{(i)} \) where for each \( i, \mu_i > 0 \). This set is clearly a cone (in fact it is even convex) and is called the cone generated by the vectors \( x^{(i)}, i = 1, \ldots, k \). It is useful to recast the definition of this particular cone as follows. Take \( A \) to be an \( n \times k \) matrix whose columns are the vectors \( x^{(i)} \). Then the cone generated by these vectors is the set \( \{ z \in \mathbb{R}^n | z = A\mu, \mu \geq 0 \} \).

Before proving the theorem, we need to prove an auxiliary result regarding this particular cone which is generated by the columns of \( A \).

**Lemma 3.3** Let \( A \) be an \( m \times n \) matrix. Then the set \( \mathcal{R} = \{ z \in \mathbb{R}^m | z = A\mu, \mu \geq 0 \} \) is a closed subset of \( \mathbb{R}^m \).

**Proof:** We first recognize that it is possible to write elements of the set \( \mathcal{R} \) as a positive linear combination of \( n \) vectors, namely the columns \( a_j \) of the matrix \( A \). That is

\[ \mathcal{R} := \left\{ z \in \mathbb{R}^m \mid z = \sum_{j=1}^{n} \mu_j a_j, \mu_j \geq 0 \right\}. \]
Notice that the set $\mathcal{R}$ is a cone. We will show that it is closed by an induction argument based on the number of vectors $a_j$, $j = 1, \ldots, k$.

When $k = 1$, $\mathcal{R}$ is either $\{0\}$ if $\mu_1 = 0$ or is a half-line and is therefore, in either case, is a closed set. Now, suppose that for some $k \geq 0$ all sets of the form

$$\mathcal{R}_{k-1} := \left\{ z \mid z = \sum_{j=1}^{k-1} \mu_j a_j, \mu_j \geq 0 \right\},$$

are closed. We then consider a set of the form

$$\mathcal{R}_k := \left\{ z \mid z = \sum_{j=1}^{k} \mu_j a_j, \mu_j \geq 0 \right\},$$

and show that this latter set is also closed. There are two cases. First, suppose that $\mathcal{R}_k$ contains the vectors $-a_1, -a_2, \ldots, -a_k$. Then $\mathcal{R}_k$ is a subspace of dimension not exceeding $k$ so it is closed.

The second possibility is that $\mathcal{R}_k$ does not contain one of the $-a_i$. Without loss of generality, we may suppose that it does not contain $-a_k$ (renumber if necessary). Then, every $y \in \mathcal{R}_k$ has the form $y = y_{k-1} + \alpha a_k$ where $y_{k-1}$ is in the set $\mathcal{R}_{k-1}$. To show that $\mathcal{R}_k$ is closed, suppose that $z^{(o)}$ is a limit point. Then there exists a sequence $\{z^{(n)}\}_{n=1}^{\infty}$ in $\mathcal{R}_k$ such that $z^{(n)} \to z^{(o)}$ as $n \to \infty$ where the $z^{(n)}$ have the form

$$z^{(n)} = y^{(n)}_{k-1} + \alpha_n a_k, \alpha_n \geq 0.$$

Let us suppose, for the moment, that the sequence $\{\alpha_n\}_{n=1}^{\infty}$ is bounded. Then, without loss of generality, we may assume that the sequence converges to a limit $\alpha$ as $n \to \infty$. Then, $z^{(n)} - \alpha_n a_k \in \mathcal{R}_{k-1}$ and this latter set is closed. Therefore

$$z - \alpha a_k = \lim_{n \to \infty} (z^{(n)} - \alpha_n a_k) = \lim_{n \to \infty} y^{(n)}_{k-1} := \bar{y} \in \mathcal{R}_{k-1}.$$

We may conclude that

$$z = \bar{y} + \alpha a_k \in \mathcal{R}_k.$$

Hence this latter set is closed.

It remains for us to prove that the sequence $\{\alpha_n\}_{n=1}^{\infty}$ is a bounded sequence. Assume the contrary, namely that $\alpha_n \to \infty$ as $n \to \infty$. Then since the $z^{(n)}$ converge, they form a
bounded sequence. Hence \((1/\alpha_n)z^{(n)} \to 0\) as \(n \to \infty\). It follows that \((1/\alpha_n)y_{k-1}^{(n)} + a_k \to 0\) as \(n \to \infty\). Therefore \(\lim_{n \to \infty}(1/\alpha_n)y_{k-1}^{(n)} = -a_k\). But since \(\mathcal{R}_{k-1}\) is closed, this means that \(-a_k \in \mathcal{R}_{k-1}\) which is a contradiction.

Having this lemma in hand, we may turn to the proof of Theorem 3.1.

**Proof:** First, it is easy to see that the condition is necessary. Indeed, if the system (3.2-3.3) has a non-negative solution \(\mathbf{x} \geq 0\), then, for all \(\mathbf{y} \in \mathbb{R}^m\) such that \(A^\top \mathbf{y} \geq 0\), we have\(^5\)

\[
\langle \mathbf{y}, \mathbf{b} \rangle = \langle \mathbf{y}, A\mathbf{x} \rangle = \langle A^\top \mathbf{y}, \mathbf{x} \rangle \geq 0,
\]

since all terms in the inner product are products of non-negative real numbers.

To see that the condition is sufficient we assume that the system (3.2-3.3) has no solution and show that there is some vector \(\mathbf{y}\) such that \(A^\top \mathbf{y} \geq 0\) and \(\langle \mathbf{b}, \mathbf{y} \rangle < 0\).

In order to do this, we will apply the basic separation theorem. Consider the set

\[
\mathcal{R} := \{z \in \mathbb{R}^m \mid z = Ax, \mathbf{x} \geq 0\}.
\]

Clearly this set is convex and, by the preceding lemma, it is closed. To say that the system (3.2-3.3) has no solution says that \(\mathbf{b} \notin \mathcal{R}\). Observe that the set \(\{\mathbf{b}\}\) is closed, bounded and convex. Hence, by the strict separation theorem, there exists a vector \(\mathbf{a} \in \mathbb{R}^m, \mathbf{a} \neq 0\) and a scalar \(\alpha\), such that

\[
\langle \mathbf{a}, \mathbf{y} \rangle < \alpha \leq \langle \mathbf{a}, \mathbf{b} \rangle, \quad \text{for all } \mathbf{y} \in \mathcal{R}.
\]

Since \(0 \in \mathcal{R}\) we must have \(\alpha > 0\). Hence \(\langle \mathbf{a}, \mathbf{b} \rangle > 0\). Likewise, \(\langle \mathbf{a}, A\mathbf{x} \rangle \leq \alpha\) for all \(\mathbf{x} \geq 0\). From this it follows that \(A^\top \mathbf{a} \leq 0\). Indeed, if the vector \(\mathbf{w} = A^\top \mathbf{a}\) were to have a positive component, say \(w_j\), then we can take \(\hat{x} = (0, 0, \ldots, 0, M, 0, \ldots, 0)^\top\) where \(M > 0\) appears in the \(j^{th}\) position. Then certainly \(\hat{x} \geq 0\) and

\[
\langle A^\top \mathbf{a}, \hat{x} \rangle = w_j M,
\]

which can be made as large as desired by choosing \(M\) sufficiently large. In particular, if we choose \(M > \alpha/w_j\) then the bound \(\langle \mathbf{a}, A\mathbf{x} \rangle \leq \alpha\) is violated. This shows that

\(^5\)Recall that, for the Euclidean inner product \(\langle A\mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{x}, A^\top \mathbf{y} \rangle\). Indeed \(\langle \mathbf{y}, A\mathbf{x} \rangle = \sum_{i=1}^n \sum_{k=1}^n a_{i,k}x_ky_i = \sum_{k=1}^n \left(\sum_{i=1}^n a_{i,k}y_i\right) = \langle A^\top \mathbf{y}, \mathbf{x} \rangle\).
\[ A^\top a \leq 0 \] and completes the proof. Indeed, we simply set \( y = -a \) to get the required result. \( \square \)

There are a number of variants each of which is just a reworking of the basic theorem. Two of them are particularly useful:

(a) The system

\[
\begin{align*}
Ax & \leq b, \\
x & \geq 0,
\end{align*}
\]

has a solution if and only if, for all \( y \geq 0 \) such that \( A^\top y \geq 0 \) we have \( \langle y, b \rangle \geq 0 \).

(b) The system

\[
Ax \leq b,
\]

has a solution if and only if for all \( y \geq 0 \) such that \( Ay = 0 \) the inner product \( \langle y, b \rangle \geq 0 \).

There are also a number of closely related results. Here is one.

**Theorem 3.4** (Gordon) Let \( A \) be an \( m \times n \) real matrix, \( x \in \mathbb{R}^n \) and \( y \in \mathbb{R}^m \). Then one and only one of the following conditions holds:

1. There exists and \( x \in \mathbb{R}^n \) such that \( Ax < 0 \);

2. There exists a \( y \in \mathbb{R}^m \), \( y \neq 0 \) such that \( A^\top y = 0 \) and \( y \geq 0 \).

**Proof:** Let \( \hat{e} = (1, 1, \ldots, 1)^\top \in \mathbb{R}^m \). Then the first condition is equivalent to saying that \( Ax \leq -\hat{e} \) has a solution. By Theorem 3.1 this is equivalent to the statement that if \( y \geq 0 \) and \( A^\top y = 0 \) then \( \langle -y, \hat{e} \rangle \geq 0 \). Hence there is no \( y \neq 0 \) such that \( A^\top y = 0 \) and \( y \geq 0 \).

Conversely, if there is a \( y \neq 0 \) such that \( A^\top y = 0 \) and \( y \geq 0 \), then the condition of the Farkas-Minkowski Theorem does not hold and hence \( Ax < 0 \) has no solution. \( \square \)

We now turn to a basic solvability theorem from Linear Algebra which is sometimes called the **Fredholm Alternative Theorem**. While the usual proofs of this result do not use the Farkas-Minkowski result, we do so here.
Theorem 3.5 Either the system of equations $Ax = b$ has a solution, or else there is a vector $y$ such that $A^\top y = 0$ and $y^\top b \neq 0$.

Proof: In order to use the Farkas-Minkowski Theorem, we need to rewrite the equation $Ax = b$ which has no sign constraint, as an equivalent system which does have such a constraint. We do this by introducing two new variables $u \geq 0$ and $v \geq 0$ with $x + v = u$. The equation then is $A(u - v) = b$ or, in partitioned form

$$[A, -A] \begin{pmatrix} u \\ v \end{pmatrix} = b,$$

with non-negative solution $(u, v)^\top \geq 0$. The Farkas alternative is then that $y^\top [A, -A] \geq 0$, $y^\top b < 0$ has a solution $y$. These inequalities mean, in particular, that $y^\top A \geq 0$ and $-y^\top A \geq 0$ which together imply that $y^\top A = 0$. So the Farkas alternative says $y^\top A = 0, y^\top b < 0$ has a solution.

Without loss of generality, we may replace $y$ by $-y$ so that we now have $y^\top A = 0$ and $y^\top b \neq 0$. This is exactly the Fredholm Alternative. $\square$

The Fredholm Alternative is often stated in a slightly different way which leads some people to say that “uniqueness implies existence”, probably not a good way to think of the result.

Corollary 3.6 If the homogeneous equation $y^\top A = 0$ has a unique solution then there exists a solution to the system $Ax = b$.

Let us illustrate the application of the Minkowski-Farkas result in two simple cases.

Example 3.7 Does the algebraic system

$$\begin{pmatrix} 4 & 1 & -5 \\ 1 & 0 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

have a non-negative solution?

The Farkas Alternative for this system is the dot product $\langle (1, 1), (y_1, y_2) \rangle < 0$ together with
\[
\begin{pmatrix}
4 & 1 \\
170 & 2 \\
-5 & 2
\end{pmatrix}
\begin{pmatrix}
y_1 \\
y_2
\end{pmatrix}
\geq
\begin{pmatrix}
0 \\
0 \\
0
\end{pmatrix}.
\]

Since \( y_1 \geq 0 \) from the second row, the dot product \( \langle B, y \rangle \) requires that \( y_2 < 0 \). But then the inequality from the last row is violated. This shows that there are no solutions to the alternative system, hence the original system has a non-negative solution.

**Example 3.8** We look for non-negative solutions of the equation

\[
\begin{pmatrix}
1 & 1 & 0 \\
0 & 1 & 1 \\
1 & 0 & 1 \\
1 & 1 & 1
\end{pmatrix}
\begin{pmatrix}
x_1 \\
x_2 \\
x_3
\end{pmatrix}
= 
\begin{pmatrix}
2 \\
2 \\
2 \\
1
\end{pmatrix}.
\]

The Farkas Alternative is

\[
\begin{pmatrix}
1 & 0 & 1 & 1 \\
1 & 1 & 0 & 1 \\
0 & 1 & 1 & 1
\end{pmatrix}
\begin{pmatrix}
y_1 \\
y_2 \\
y_3 \\
y_4
\end{pmatrix}
\geq
\begin{pmatrix}
0 \\
0 \\
0
\end{pmatrix}
\]

together with

\[2y_1 + 2y_2 + 2y_3 + y_4 < 0.\]

Since this system has the solution \( y_1 = y_2 = y_3 = -1/2 \) and \( y_4 = 1 \), there are no non-negative solutions of the original equation.

Finally, we look at an application of the Fredholm Alternative.

**Example 3.9** Consider the set of equations

\[
\begin{align*}
2x + 3y &= 1 \\
x - 3y &= 1 \\
-x + y &= 0
\end{align*}
\]
Now look at the transposed system

\[
\begin{align*}
2u + v - w &= 0 \\
3u - v + w &= 0 \\
\end{align*}
\]

together with \( u + v \neq 0 \).

Since this is a homogeneous system, we may well normalize this last requirement to \( u + v = 1 \). Then look at the matrix

\[
\begin{pmatrix}
2 & 1 & -1 \\
3 & -3 & 1 \\
1 & 1 & 0
\end{pmatrix}
\]

It is easy to check that this matrix has row-eschelon form

\[
\begin{pmatrix}
2 & 1 & -1 & 0 \\
0 & -9/5 & 5/2 & 0 \\
0 & 0 & 7/9 & 1
\end{pmatrix}
\]

and hence the non-homogeneous system has a non-trivial solution. Hence the original system has no solution.

### 3.1 Applications to Linear Programming

As we mentioned in the beginning, the Farkas-Minkowski Alternative is often proved, not by recourse to a separation theorem but from the Duality Theorem of Linear Programming. However, the Farkas-Minkowski theorem leads to a particularly simple proof of the Duality Theorem. Let us explain.

Linear programming deals with problems of the form

\[
\begin{align*}
\text{minimize} & \quad \langle c, x \rangle \\
\text{subject to} & \quad Ax = b, \ x \geq 0.
\end{align*}
\]

Where \( x, c \in \mathbb{R}^n, b \in \mathbb{R}^m \) and \( A \) is \( m \times n \) matrix.

This problem is called the \textbf{primal} problem. The dual problem to this primal is
maximize \( \langle b, y \rangle \)
subject to \( A^\top y \leq c \).

It is clear that the feasible region, that is the set of points that satisfy the constraints in either program is a convex set. These convex sets are called the feasible sets for the two problems. The first thing that we can say concerning the feasible points, while simple, is of great use.

**Theorem 3.10 (The Weak Duality Theorem)**

Let \( x \) be a feasible point for the primal problem, and \( y \) a feasible point for its dual, then \( \langle c, x \rangle \geq \langle b, y \rangle \).

**Proof:**

\[
\langle c, x \rangle - \langle b, y \rangle = \langle c, x \rangle - \langle Ax, y \rangle = \langle x, c \rangle - \langle x, A^\top y \rangle = \langle x, c - A^\top y \rangle \geq 0,
\]

since, both arguments in the last inner product are vectors with positive components. Hence we have the inequality \( \langle c, x \rangle \geq \langle b, y \rangle \). \( \square \)

We call the value \( \langle c, x \rangle - \langle b, y \rangle \) the **duality gap**.

This theorem shows that a feasible solution to either problem yields a bound on the value of the other. From that fact we can say a number of things. First, suppose that the primal problem is unbounded, i.e. \( \inf \langle c, x \rangle = -\infty \). Then the set of feasible points for the dual must be empty since, if there were a feasible point \( y \) then \( \langle b, y \rangle \) would be a lower bound for \( \langle c, x \rangle \) which is a contradiction. Similarly, if the dual is unbounded, then the primal problem is infeasible.

Also, if we have feasible solution for both problems for which the duality gap is zero, then these feasible points are optimal solutions for their respective problems. Indeed, since a dual feasible solution gives an upper bound to the cost functional of the primal, if they are equal then the primal has achieved its maximum value. Likewise for the dual problem.

We are now in a position to state and prove the **strong duality property** which, when established, shows that we may solve the dual problem in place of the primal and this will lead to a solution of the primal itself.

**Theorem 3.11 (Strong Duality)**

Suppose that both the primal and the dual problems have feasible solutions. Then \( x \) is optimal for the primal problem if and only if
(a) \( \bar{x} \) is feasible for the primal.

(b) there is a feasible point for the dual, \( \bar{y} \), that satisfies

\[
\langle c, \bar{x} \rangle = \langle b, \bar{y} \rangle
\]

Proof: Since, by weak duality, we always have \( \langle c, x \rangle \geq \langle b, y \rangle \) for all feasible points \( x \) for the primal and \( y \) for the dual, we need to show that there is a feasible primal/dual pair \( \{\bar{x}, \bar{y}\} \) for which the reverse inequality is true.

Let \( \epsilon > 0 \) and consider the system

\[
\begin{align*}
-\langle c, x \rangle & \leq -\langle c, \bar{x} \rangle - \epsilon \\
A x &= *b \\
x & \geq 0
\end{align*}
\]

By definition of optimality for the primal problem, this system has no solution. Hence, by the Farkas Theorem the system

\[
\begin{align*}
-\lambda c + A^\top y & \geq 0 \\
\lambda(-\langle c, \bar{x} \rangle - \epsilon) + \langle y, b \rangle & < 0 \\
\lambda & \geq 0
\end{align*}
\]

has a solution. Denote that solution by \( (\lambda^*, y^*) \). Now, \( \lambda^* \geq 0 \), but if equality were to hold, then the last system of inequalities would reduce to

\[
\begin{align*}
A^\top y^* & \geq 0 \\
\langle y^*, b \rangle & < 0.
\end{align*}
\]

But then the Farkas alternative implies that the system \( A x = 0 \) has no non-negative solution. Hence the primal problem is infeasible which violates the original assumption. Hence we must have \( \lambda^* > 0 \).

Now, we define \( \lambda' = y^*/\lambda^* \) and note that

\[
A^\top y' = \frac{A^\top y^*}{\lambda^*} \geq c
\]
which means that \( y' \) is a feasible point for the dual problem. Moreover \( \langle y', b \rangle < \langle c, \overline{x} \rangle - \epsilon \)
as a simple computation shows. Since \( y' \) is feasible for the dual problem, we then have

\[
\langle c, \overline{x} \rangle \leq \langle y^*, b \rangle \leq \langle y', b \rangle < \langle c, \overline{x} \rangle + \epsilon .
\]

Since \( \epsilon > 0 \) is arbitrary, it follows that the optimal values of the primal and dual are the same: \( \langle c, \overline{x} \rangle = \langle y^*, b \rangle . \)

\[\Box\]

### 3.2 De Finetti’s Arbitrage Theorem

Arbitrage is the practice of buying one commodity at a low price in one market and making a risk-free profit by taking advantage of a higher price in another market. As a trivial example, suppose that the market near your home is selling bananas for $0.29/pound while the market near your mother-in-law’s house is selling them for $0.40/pound. This price differential can support a more harmonious family, since you are motivated to buy, say 10 pounds of bananas at your local market, and sell them to the market near the in-laws. Hence your purchase price is $2.90 and your selling price is $4.00 for a profit of $1.10.

You can imagine doing this often and continuing, at least for a while, to make the profit. However, the elementary laws of supply and demand will tell us that if you make this trade often enough, the supply of bananas will decrease at your neighborhood market and the price will go up, while the price at the distant store, where the supply will increase, will tend to decrease. Eventually a price equilibrium will be established and the market in bananas will be “arbitrage free”.

The first mathematical result was given by Augustin Cournot in 1838 in his book *Researches on the Mathematical Principles of the Theory of Wealth*. In that book he stated what might be called the first arbitrage theorem:

There exists a system of absolute prices for commodities such that the exchange rates (or relative prices) are price ratios if and only if the exchange rates are arbitrage-free.

In these notes, we want to look at the closely related statement due to De Finetti. To set the stage, suppose that there is an experiment having \( m \) possible outcomes for which there are \( n \) possible wagers. That is, if you bet the amount \( y \) on wager \( i \) you win the amount \( y r_i(j) \) if the outcome of the experiment is \( j \). Here \( y \) can be positive, negative, or zero. A betting strategy is a vector \( \mathbf{y} = (y_1, \ldots, y_n)^\top \) which means that
you simultaneously bet the amount $y_i$ on wager $i$ for $i = 1, \ldots, n$. So if the outcome of the experiment is $j$, your gain (or loss) from the strategy $y$ is given by

$$\sum_{j=1}^{n} y_i r_j(j).$$

The result of De Finetti is the following:

**Theorem 3.12** Exactly one of the following is true: Either

- there exists a probability vector $p = (p_1, \ldots, p_m)$ with
  $$\sum_{j=1}^{m} r_i(j) p_j = 0,$$
  or

- there exists a betting strategy $y$ such that
  $$\sum_{i=1}^{n} y_i r_i(j) > 0, \quad \text{for all } j = 1, 2, \ldots, m.$$

In other words, either there exists a probability distribution on the outcome under which all bets have expected gain equal to zero, or else there is a betting strategy which always results in a positive win.

**Proof:** Consider the matrix

$$A := \begin{pmatrix}
  r_1(1) & \cdots & r_1(m) \\
  \vdots & \ddots & \vdots \\
  r_n(1) & \cdots & r_n(m) \\
  -1 & \cdots & -1
\end{pmatrix}$$

and the vector $b := (0, 0, \cdots, 0, -1)^\top$. Then consider the system

$$A p = b, \quad p \geq 0, \quad p = (p_1, \ldots, p_m)^\top,$$

which reads

30
\[ \sum_{j=1}^{m} r_i(j) p_j = 0, \ i = 1, \ldots, n \]

while the last row, the \( n + 1^{st} \) row reads

\[ \sum_{j=1}^{m} -p_j = -1. \]

Now the Farkas-Minkowski Theorem tells us that either this system has a solution \( p \) or the system

\[ y^\top A \geq 0, \ y^\top b < 0 \]

has a solution \( y \in \mathbb{R}^n \) but not both.

But with our choice of matrix \( A \), the inequality \( y^\top A \geq 0 \) reads (note that \( A \) is an \( m \times (n + 1) \)-matrix)

\[ \sum_{i=1}^{n} x_i r_i(j) - y_{n+1} \geq 0, \ j = 1, \ldots, m, \]

while the second inequality \( y^\top b < 0 \) reads just \( -y_{n+1} < 0 \). Thus we have

\[ \sum_{i=1}^{n} x_i r_i(j) \geq y_{n+1} > 0, \ j = 1, \ldots, m. \]

This is the second possibility of the theorem. \( \square \)

### 3.3 European Call Options and Fair Prices

A call option is a contract between two parties, the seller and the buyer, that the buyer has the right to buy a certain commodity or certain financial instrument from the seller at a certain agreed upon price at some time in the future. The buyer has no obligation to buy, but the seller does have the obligation to sell at the agreed upon price (usually called the strike price) if the call is exercised. In exchange for this obligation, the buyer usually pays a premium or option price as the price of the option when it is originally purchased.

The two parties are making two different bets; the seller is betting that the future price of the commodity is less than the strike price. In this case, the seller collects the
premium and does not receive any gain if the commodity prices rises above the strike price. The buyer is betting that the future price will be higher than the strike price, so that he can buy the commodity at a price lower than the prevailing one and then sell, thereby making a profit. While the buyer may make a great deal of profit, his loss is limited by the premium that is paid.

The most common types of options are the American Option and the European Option. The former allows the option to be exercised at any time during the life of the option while the latter allows it to be exercised only on the expiration date of the option. Here, we are going to look at a simple example of a one period European call option and ask what a “fair” price for the option is.

We assume that there are only two possible outcomes, so that \( m = 2 \) and the matrix \( A \) is given by

\[
A = \begin{pmatrix} r(1) & r(2) \\ -1 & -1 \end{pmatrix}.
\]

If \( r(1) \neq r(2) \) then the matrix \( A \) is non-singular and has inverse

\[
A^{-1} = \frac{1}{r(2) - r(1)} \begin{pmatrix} -1 & -r(2) \\ 1 & r(1) \end{pmatrix},
\]

so that, with \( b \) as in the preceding section \( b = (0, -1)^\top \) and

\[
A^{-1} b = \frac{1}{r(2) - r(1)} \begin{pmatrix} -1 & -r(2) \\ 1 & r(1) \end{pmatrix} \begin{pmatrix} 0 \\ -1 \end{pmatrix} = \frac{1}{r(2) - r(1)} \begin{pmatrix} r(2) \\ -r(1) \end{pmatrix}.
\]

So if \( r(2) > 0 \) and \( r(1) < 0 \) then the right-hand side of the above relation can be interpreted as a probability vector

\[
p = \begin{pmatrix} 1 - p \\ p \end{pmatrix} = \frac{1}{r(2) - r(1)} \begin{pmatrix} r(2) \\ -r(1) \end{pmatrix}.
\]

Of course, if both are positive, any wager which assigns a positive bet to both is a guaranteed win, and if both are negative, any wager that assigns a negative bet to both is a guaranteed win.

Now suppose that we allow another wager. This new wager has the return \( \alpha \) if the outcome is 1 and \( \beta \) if the outcome is 2. Then, according to De Finetti’s theorem, unless

\[
\alpha (1 - p) + \beta p = 0,
\]
there will be some combination of the two wagers that has a guaranteed profit.

we now assume that there is a certain stock which at time \( t = 0 \) has the price \( S(0) \). Suppose that at time \( t = 1 \) there are two possibilities, either the value of the stock can go up by a fixed factor \( u \) so that the price at \( t = 1 \) is \( S(1) = uS(0) \) or it can go down by a fixed factor \( d \) so that \( S(1) = dS(0) \), \( u > 1 > d \). Moreover, suppose that the interest rate on a deposit in a bank for this period is \( r \) so that the current value of \( M \) is \( M(1 + r)^{-1} \).

So the profit per dollar if the stock goes up is \( r(2) = (u/r + 1) - 1 \), while if the stock goes down it is \( r(1) = (d/1 + r) - 1 \). In terms of the probability, \( p \), defined above, we have

\[
1 - p = \frac{u - 1 - r}{u - d},
\]

\[
p = \frac{1 + r - d}{u - d}.
\]

We now want to decide what is a “fair price” for a European call option. Let \( K \) be strike price at the end of one time period. Let \( C \) be the current price of the call option. Define the parameter \( k \) by

\[
K = k S(0).
\]

Suppose for the moment that the stock goes up. Since you can buy the stock at time \( t = 1 \) for a price of \( kS(0) \) (the strike price) and then sell it immediately at the price \( uS(0) \), and since the price of the option today is \( $C \), the gain per unit purchased is

\[
\frac{u - k}{1 + r} - C.
\]

Of course, if the stock goes down by a factor of \( d \), you lose \( $C^6 \).

Now, De Finetti’s theorem says that, unless

\[
0 = -(1 - p)C + p \left( \frac{u - k}{1 + r} - C \right) = p \cdot \frac{u - k}{1 + r} - C
\]

\[
= \frac{1 + r - d}{u - d} \cdot \frac{u - k}{1 + r} - C,
\]

\[\text{This is under the assumption that } d \leq k \leq u.\]
there is a mixed strategy of buying or selling the stock and buying and selling the option with a sure profit. Thus the price of the option should be set at the “fair price” given by

\[ C = \frac{1 + r - d}{u - d} \cdot \frac{u - k}{1 + r}. \]

This is the fair price of the option in the sense that if the option were priced differently, an arbitrageur could make a guaranteed profit.

### 3.4 Extreme Points of Convex Sets

**Definition 3.13** A point \( x \) in a convex set \( C \) is said to be an extreme point of \( C \) provided there are not two distinct points \( x_1, x_2 \in C, x_1 \neq x_2 \) such that \( x = (1 - \lambda) x_1 + \lambda x_2 \), for some \( \lambda \in (0, 1) \).

Notice that the corner points of a polygon satisfy this condition and are therefore extreme points. In the case of a polygon there are only finitely many. However, this may not be the case. We need only to think of the convex set in \( \mathbb{R}^2 \) consisting of the unit disk centered at the origin. In that case, the extreme points are all the points of the unit circle, i.e., the boundary points of the disk.

There are several questions that we wish to answer: (a) Do all convex sets have an extreme point? (\textbf{Ans.:} NO!); (b) Are there useful conditions which guarantee that an extreme point exists? (\textbf{Ans.} YES!); (c) What is the relationship, under the conditions of part (b) between the convex set, \( C \), and the set of its extreme points, ext \( (C) \)? (\textbf{Ans.}: \( \text{cl} \{\text{co ext} \ (C)\} = C \)), and, (d) What implications do these facts have with respect to optimization problems?

It is very easy to substantiate our answer to (a) above. Indeed, if we look in \( \mathbb{R}^2 \) at the set \( \{(x, y) \in \mathbb{R}^2 : 0 \leq x \leq 1, y \in \mathbb{R} \} \) (draw a picture!) we have an example of a convex set with no extreme points. In fact this example is in a way typical of sets without extreme points, but to understand this statement we need a preliminary result.

**Lemma 3.14** Let \( C \) be a convex set, and \( H \) a supporting hyperplane of \( C \). Define the (convex) set \( T := C \cap H \). Then every extreme point of \( T \) is also an extreme point of \( C \).

**Remark:** The set \( T \) is illustrated in the next figure. Note that the intersection of \( H \) with \( C \) is not just a single point. It is, nevertheless, closed and convex since both \( H \) and \( C \) enjoy those properties.
Figure 9: The set $T = H \cap C$

**Proof:** Suppose $\tilde{x} \in T$ is *not* an extreme point of $C$. Then we can find a $\lambda \in (0,1)$ such that $\tilde{x} = (1 - \lambda)x_1 + \lambda x_2$ for some $x_1, x_2 \in C$, $x_1 \neq x_2$. Assume, without loss of generality, that $H$ is described by $\langle x, a \rangle = c$ and that the convex set $C$ lies in the positive half-space determined by $H$. Then $\langle x_1, a \rangle \geq c$ and $\langle x_2, a \rangle \geq c$. But since $\tilde{x} \in H$,

$$c = \langle \tilde{x}, a \rangle = (1 - \lambda) \langle x_1, a \rangle + \lambda \langle x_2, a \rangle,$$

and thus $x_1$ and $x_2$ lie in $H$. Hence $x_1, x_2 \in T$ and hence $\tilde{x}$ is not an extreme point of $T$. \hfill \Box

We can now prove a theorem that guarantees the existence of extreme points for certain convex subsets of $\mathbb{R}^n$.

**Theorem 3.15** A non-empty, closed, convex set $C \subset \mathbb{R}^n$ has at least one extreme point if and only if it does not contain a line, that is, a set $L$ of the form $L = \{x + \alpha d : \alpha \in \mathbb{R}, d \neq 0\}$.

**Proof:** Suppose first that $C$ has an extreme point $x^{(e)}$ and contains a line $L = \{\overline{x} + \alpha d : \alpha \in \mathbb{R}, d \neq 0\}$. We will see that these two assumptions lead to a contradiction. For each $n \in \mathbb{N}$, the vector

$$x^{(n)}_{\pm} := \left(1 - \frac{1}{n}\right)x^{(e)} + \frac{1}{n}(\overline{x}_{\pm} n d) = x^{(e)} + d + \frac{1}{n}(\overline{x} - x^{(e)})$$

lies in the line segment connecting $x^{(e)}$ and $\overline{x} \pm n d$, and so it belongs to $C$. Since $C$ is closed, $\overline{x} \pm d = \lim_{n \to \infty} x^{(n)}_{\pm}$ must also belong to $C$. It follows that the three vectors
$\pi - d$, $x^{(e)}$, and $\pi + d$, all belong to $C$ contradicting the hypothesis that $x^{(e)}$ is an extreme point of $C$.

To prove the converse statement, we will use induction on the dimension of the space. Suppose then that $C$ does not contain a line. Take, first, $n = 1$. The statement is obviously true in this case for the only closed convex sets which are not all of $\mathbb{R}$ are just closed intervals.

We assume, as induction hypothesis, that the statement is true for $\mathbb{R}^{n-1}$. Now, if a nonempty, closed, convex subset of $\mathbb{R}^n$ contains no line, then it must have boundary points. Let $\overline{x}$ be such a point and let $H$ be a supporting hyperplane of $C$ at $\overline{x}$. Since $H$ is an $(n - 1)$-dimensional manifold, the set $C \cap H$ lies in an $(n - 1)$-dimensional subspace and contains no line, so by the induction hypothesis, $C \cap H$ must have an extreme point. By the preceding lemma, this extreme point of $C$.

From this result we have the important representation theorem:

**Theorem 3.16** A closed, bounded, convex set $C \subset \mathbb{R}^n$ is equal to $\text{cl} [ \text{co ext } (C)]$.

**Proof:** We begin by observing that since the set $C$ is bounded, it can contain no line. Moreover, the smallest convex set that contains the non-empty set $\text{ext } (C)$ is just the convex hull of this latter set. So we certainly have that $C \supset \text{cl}[ \text{co ext } (C)] \neq \emptyset$. Denote the closed convex hull of the extreme points by $K$. We remark that, since $C$ is bounded, $K$ is necessarily bounded.

We need to show that $C \subset K$. Assume the contrary, namely that there is a $y \in C$ with $y \not\in K$. Then by the first separation theorem, there is a hyperplane $H$ separating $y$ and $K$. Thus, for some $a \neq y$, $\langle y, a \rangle < \inf_{x \in K} \langle x, a \rangle$. Let $c_o = \inf_{x \in C} \langle x, a \rangle$. The number $c_o$ is finite and there is and $\hat{x} \in C$ such that $\langle \hat{x}, a \rangle = c_o$ because, by the theorem of Weierstraß, a continuous function (in this case $x \mapsto \langle x, a \rangle$) takes on its minimum value over any closed bounded set.$^7$

It follows that the hyperplane $H = \{ x \in \mathbb{R}^n : \langle x, a \rangle = c_o \}$ is a supporting hyperplane to $C$. It is disjoint from $K$ since $c_o < \inf_{x \in K} \langle x, a \rangle$. The preceding two results then show that, the set $H \cap C$ has an extreme point which is also a fortiori an extreme point of $C$ and which cannot then be in $K$. This is a contradiction. \[\square\]

Up to this point, we have answered the first three of the questions raised above. Next, we want to explore the relationship of these results to the linear programming problem.

$^7$The continuity of this map follows immediately from the Cauchy-Schwarz Inequality!
Consider the linear program (P):

\[
\begin{align*}
\text{minimize} & \quad c^\top \cdot x \\
\text{subject to} & \quad A x = b \\
& \quad x \geq 0,
\end{align*}
\]

where \( A \) is an \( m \times n \) matrix of rank \( m \).

Recall the following definitions:

**Definition 3.17** A feasible solution is an element \( x \in \mathbb{R}^n \) which satisfies the constraints \( A x = b \), and \( x \geq 0 \).

Among all solutions of the equation \( A x = b \), certain ones are called basic.

**Definition 3.18** Let \( B \) be any \( m \times m \) non-singular submatrix of \( A \) consisting of linearly independent columns of \( A \). Then if the \( n - m \) components of a solution \( x \) corresponding to the columns of \( A \) which do not appear in \( B \) are set equal to zero, the solution, \( x_B \), of the resulting set of equations is said to be a basic feasible solution with respect to the basis consisting of columns of \( B \). The components of \( x_B \) associated with the columns of \( B \) are called the basic are called basic variables.

Note that the equation \( A x = b \) may not have any basic solutions in the general case. In order to insure that basic solutions exist, it is usual to make certain assumptions: (a) that \( n > m \); (b) that the rows of \( A \) are linearly independent; (c) that the rank of \( A \) is \( m \). These conditions are sufficient for the existence of at least one basic solution.

It may well occur that some components of a basic solution are zero.

**Definition 3.19** If one or more basic variables in a basic solution has the value zero, then that solution is said to be a degenerate basic solution.

We shall refer to a feasible solution which is also basic as a basic feasible solution. A feasible solution that achieves the minimum value of the cost functional is said to be an optimal feasible solution. If it is also basic, then it is an optimal basic feasible solution.

Let us return to the linear programming problem \( \mathcal{P} \). The fundamental result is that we need only search among the basic feasible solutions for an optimal solution. Indeed, that is what the Simplex Method actually does.
Theorem 3.20 (The Fundamental Theorem of Linear Programming)

Given the linear programming problem $P$, where $A$ is an $m \times n$ matrix of rank $m$:

1. If there is any feasible solution, then there is a basic feasible solution.
2. If there is any optimal solution, then there is a basic optimal solution.

Proof: Suppose that a feasible solution exists. Choose any feasible solution among those with the fewest non-zero components. If there are no non-zero components, then $x = 0$ and $x$ is a basic solution by definition. Otherwise, take the index set $J := \{j_1, j_2, \ldots, j_r\}$ with elements corresponding to those $x_{j_i} > 0$. Then if we denote the corresponding columns by $\{a^{(j_1)}, a^{(j_2)}, \ldots, a^{(j_r)}\}$ there are two possibilities:

1. The set of columns is linearly independent. In this case, we certainly have $r \leq m$. If $r = m$ the corresponding solution is basic and the proof is complete. If $r < m$ then, since $A$ has rank $m$, we choose $m - r$ vectors from the remaining $n - r$ columns of $A$ so that the resulting set of $m$ column vectors is linearly independent. Assigning the value zero to the corresponding $m - r$ variables yields a (degenerate) basic feasible solution.

2. The set of columns $\{a^{(j_1)}, a^{(j_2)}, \ldots, a^{(j_r)}\}$ is linearly dependent. Then there is a choice of scalars $\alpha_{j_i}$, not all zero, such that

$$\alpha_{j_1} a^{(j_1)} + \alpha_{j_2} a^{(j_2)} + \ldots + \alpha_{j_r} a^{(j_r)} = 0. \quad (3.4)$$

Without loss of generality, we may take $\alpha_{j_i} \neq 0$ and, indeed, $\alpha_{j_i} > 0$ (otherwise multiply (3.4) by $-1$).

Now, since $x_{j_i} > 0$, the corresponding feasible solution $x_J$ is just a linear combination of the columns $a^{(j_i)}$. Hence

$$Ax = \sum_{i=1}^{r} x_{j_i} a^{(j_i)} = b.$$ 

Now multiplying the dependence relation (3.4) by a real number $\lambda$ and subtracting, we have

$$A(x - \lambda \alpha) = \sum_{i=1}^{r} (x_{j_i} - \lambda \alpha_{j_i}) a^{(j_i)} = b,$$

and this equation holds for every $\lambda$ although one or more components, $x_k - \lambda \alpha_{j_k}$ may violate the non-negativity condition. Now let $\alpha = (\alpha_1, \ldots, \alpha_n)^T$ where

$$\alpha_k := \begin{cases} \alpha_{j_i}, & k = j_i \\ 0, & k \neq j_i \end{cases}$$

38
Then, for each value of $\lambda$ the vector $x - \lambda \alpha$ is a solution of the constraint equations. Note that, for the special value $\lambda = 0$, this vector is the original feasible solution. Now, as $\lambda$ increases from 0, the components of $x - \lambda \alpha$ will individually change; each will either increase, decrease, or stay constant, depending on whether the corresponding $\alpha_i$ is positive, negative, or zero.

Suppose, for some $i_o$, $\alpha_{i_o} \leq 0$. Then $(x_{i_o} - \lambda \alpha_{i_o}) \geq 0$ for all $\lambda > 0$. On the other hand, if for some $i_o$, $\alpha_{i_o} > 0$, then $(x_{i_o} - \lambda \alpha_{i_o})$ remains greater than 0 only for sufficiently small $\lambda > 0$. Now, we take

$$\tilde{\lambda} := \min \left\{ \frac{x_i}{\alpha_i} \mid x_i > 0, \alpha_i > 0 \right\}.$$

Then $\tilde{\lambda}$ corresponds to the first value of $\lambda$ for which one or more non-zero components of $x - \lambda \alpha$ becomes 0. For this value of $\lambda$ the vector $x - \tilde{\lambda} \alpha$ is feasible and has at most $r - 1$ positive components. Repeating this process if necessary, we can obtain a feasible solution with non-zero components corresponding to linearly independent columns. In this situation, the previous alternative applies.

This completes the proof of the first part of the theorem.

Now assume that $x^*$ is an optimal solution. There is no guarantee that this optimal solution is unique. In fact, we have seen cases where there is no uniqueness. Some of these solutions may have more positive components than others. Without loss of generality, we assume that $x^*$ has a minimal number of positive components. If $x^* = 0$ then $x^*$ is basic and the cost is zero. If $x^* \neq 0$ and if $J$ is the corresponding index set, then there are two cases as before. The proof of the first case in which the corresponding columns are linearly independent is exactly as in the previous proof of this case.

In the second case, we proceed just as with the second case above. To do this, however, we must show that, for any $\lambda$, the vector $x^* - \lambda \alpha$ is optimal. To show this, we note that the associated cost is

$$\langle c, x^* - \lambda \alpha \rangle = \langle c, x^* \rangle - \lambda \langle c, \alpha \rangle$$

Then, for sufficiently small $|\lambda|$, the vector $x^* - \lambda \alpha$ is a feasible solution for positive or negative values of $\lambda$. Hence, we conclude that

$$\langle c, \alpha \rangle = 0$$

since, were it not, then we could determine a small $\lambda$ of the proper sign, to make
\[ \langle c, x^* - \lambda \bar{\alpha} \rangle < \langle c, x^* \rangle, \]

which would violate the assumption of the optimality of \( x^* \).

Having established that the new feasible solution with fewer positive components is also optimal, the remainder of the proof may be completed exactly as in case 2 above. \( \square \)

**REMARK:** If we are dealing with a problem in \( n \) variables and \( m \) constraints, we can produce an upper bound for the number of basic feasible solutions. Specifically, we must choose \( m \) columns from a set of \( n \) columns and there are at most

\[
\binom{n}{m} = \frac{n!}{m!(n-m)!}.
\]

ways to make the selection. Of course, some choices may not lead to a linearly independent set of \( m \) columns so that this number is an upper bound.

In Linear Programming, the feasible region in \( \mathbb{R}^n \) is defined by \( P := \{x \in \mathbb{R}^n | Ax = b, x \geq 0\} \). The set \( P \), as we have seen, is a convex subset of \( \mathbb{R}^n \). It is called a convex polytope. The term convex polyhedron refers to convex polytope which is bounded. Polytopes in two dimensions are often called polygons. Recall that the vertices of a convex polytope are what we called extreme points of that set.

Recall that extreme points of a convex set are those which cannot be represented as a proper convex combination of two other (distinct) points of the convex set. It may, or may not be the case that a convex set has any extreme points as shown by the example in \( \mathbb{R}^2 \) of the strip \( S := \{(x, y) \in \mathbb{R}^2 | 0 \leq x \leq 1, y \in \mathbb{R}\} \). On the other hand, the square defined by the inequalities \( |x| \leq 1, |y| \leq 1 \) has exactly four extreme points, while the unit disk described by the inequality \( x^2 + y^2 \leq 1 \) has infinitely many. These examples raise the question of finding conditions under which a convex set has extreme points. The answer in general vector spaces is answered by one of the “big theorems” called the Krein-Milman Theorem. However, as we will see presently, our study of the linear programming problem actually answers this question for convex polytopes without needing to call on that major result.

The algebraic characterization of the vertices of the feasible polytope confirms the observation that we made by following the steps of the Simplex Algorithm in our introductory example. Some of the techniques used in proving the preceding theorem come into play in making this characterization as we will now discover.

**Theorem 3.21**
The set of extreme points, \( E \), of the feasible region \( P \) is exactly the set, \( B \) of all basic feasible solutions of the linear programming problem.

**Proof:** We wish to show that \( E = B \) so, as usual, we break the proof into two parts.

(a) \( B \subset E \).

Suppose that \( x^{(b)} \in B \). Then, for the index set \( J(x^{(b)}) \subset \{1,2,\ldots,n\} \) is defined by \( j \in J(x^{(b)}) \) if and only if \( x_j^{(b)} > 0 \). Now suppose that \( x^{(b)} \notin E \). Then there exist two distinct feasible points \( y, z \in P \) and a \( \lambda \in (0,1) \) for which \( x^{(b)} = (1-\lambda)y + \lambda z \). Observe that for any integer \( k \notin J \), it must be true that \( (1-\lambda)y_k + \lambda z_k = 0 \). Since \( 0 < \lambda < 1 \) and \( y, z \geq 0 \), this implies that, for all such indices \( k \), \( y_k = z_k = 0 \).

Now \( x^{(b)} \) is basic, so that the columns of \( A \) corresponding to the non-zero components form a linearly independent set in \( \mathbb{R}^n \). Since the only non-zero components of \( y \) and \( z \) have the same indices as the non-zero components of \( x^{(b)} \), we have \( \text{span}\{a^{(j)}\}, j \in J(x^{(b)}) \) contains \( x^{(b)}, y \) and \( z \). Moreover, since \( y \) and \( z \) are feasible, we have \( Ay = b \) and \( Az = b \) so that

\[
b = (1-\lambda)b + \lambda b = (1-\lambda)Ay + \lambda Az.
\]

Now the system \( Ax = b \) is uniquely solvable on the set \( \{ x \in \mathbb{R}^n | x_i = 0, i \notin J(x^{(b)}) \} \), so that we must have \( Ax^{(b)} = Ay = Az \) and hence \( x^{(b)} \) cannot be written as a proper convex combination of two other distinct points of \( P \), which means that \( x^{(b)} \) is an extreme point of \( P \).

(b) \( E \subset B \)

If \( x^{(e)} \) is an extreme point of \( P \), let us assume that the columns associated with non-zero components correspond to columns \( a^{(j_i)} \), \( i = 1, \ldots, k \). Then, as before, we have

\[
x_{j_1}^{(e)}a^{(j_1)} + x_{j_2}^{(e)}a^{(j_2)} + \ldots + x_{j_k}^{(e)}a^{(j_k)} = b,
\]

where the \( x_{j_i}^{(e)} > 0 \), \( i = 1, \ldots, k \). To show that \( x^{(e)} \) is a basic feasible solution, we must show that the set of vectors \( \{a^{(j_1)}, a^{(j_2)}, \ldots, a^{(j_k)}\} \) is a linearly independent set. Let us suppose that this is not the case. Then there are constants \( \alpha_{j_1}, \ldots, \alpha_{j_k} \), not all zero, such that

\[
\alpha_{j_1}a^{(j_1)} + \alpha_{j_2}a^{(j_2)} + \alpha_{j_k}a^{(j_k)} = 0.
\]
We set \( \mathbf{\alpha} = \text{col} (\alpha_\ell) \) where \( \alpha_\ell = 0 \) if \( \ell \neq j_i \) for some \( i \), and \( \alpha_\ell = \alpha_{j_i} \) if \( \ell = j_i \) for some \( i \). Then, since all the non-zero components of \( \mathbf{x}^{(e)} \) are strictly positive, it is possible to find an \( \epsilon > 0 \) such that

\[
\mathbf{x}^{(e)} + \epsilon \mathbf{\alpha} \geq 0, \quad \text{and} \quad \mathbf{x}^{(e)} - \epsilon \mathbf{\alpha} \geq 0.
\]

Note that each satisfies \( A\mathbf{x} = \mathbf{b} \) since \( A(\epsilon \mathbf{\alpha}) = \epsilon A(\mathbf{\alpha}) = 0 \) by the dependence relation.

We then have

\[
\mathbf{x}^{(e)} = \frac{1}{2} (\mathbf{x}^{(e)} + \epsilon \mathbf{\alpha}) + \frac{1}{2} (\mathbf{x}^{(e)} - \epsilon \mathbf{\alpha}),
\]

which expresses \( \mathbf{x}^{(e)} \) as a proper convex combination of two distinct vectors in \( P \). But this is impossible since \( \mathbf{x}^{(e)} \) is an extreme point of \( P \). It follows that the set \( \{ \mathbf{a}^{(j_1)}, \mathbf{a}^{(j_2)} , \ldots , \mathbf{a}^{(j_k)} \} \) is a linearly independent set and so \( \mathbf{x}^{(e)} \) is a basic feasible solution. \( \square \)

There are several results that are useful to state that follow immediately from this result.

**Corollary 3.22** If the convex set \( P \) is non-empty, then it has an extreme point.

Indeed, this fact follows from the Fundamental Theorem of LP and this last equivalence theorem.

**Corollary 3.23** If there is a finite optimal solution for the problem, then there is a finite optimal solution which is an extreme point.

**Corollary 3.24** The constraint set \( P \) contains at most a finite number of extreme points.

**Corollary 3.25** If the set \( P \) is bounded, then \( P \) consists of points that are finite linear combinations of a finite number of points of \( P \).

Finally, we have a direct proof of the following result:
Theorem 3.26 If $P$ is bounded, then a linear cost function $\langle c, x \rangle$ achieves its minimum on $P$ at an extreme point of $P$.

Proof: Let $x_1, x_2, \ldots, x_k$ be the extreme points of $P$. Then any point $x \in P$ can be written as a convex combination

$$x = \sum_{i=1}^{k} \lambda_i x_i, \quad \text{where} \quad \lambda_i \geq 0, \ i = 1, 2, \ldots, k, \quad \text{and} \quad \sum_{i=1}^{k} \lambda_i = 1.$$ 

Then

$$\langle c, x \rangle = \sum_{i=1}^{k} \lambda_i \langle c, x_i \rangle.$$ 

Now let $\iota := \min \{ \langle c, x_i \rangle | i = 1, 2, \ldots, k \}$. Then from the representation of $x$ in terms of the extreme points we have

$$\langle c, x \rangle = \sum_{i=1}^{k} \lambda_i \langle c, x_i \rangle \geq (\lambda_1 + \lambda_2 + \ldots + \lambda_k) \iota = \iota,$$

and so the minimum of the cost functional is $\iota$ which is a value taken on at one of the extreme points. \hfill $\square$

4 Convex Functions

Our next topic is that of convex functions. Again, we will concentrate on the context of a map $f : \mathbb{R}^n \to \mathbb{R}$ although the situation can be generalized immediately by replacing $\mathbb{R}^n$ with any real vector space $V$. We will state many of the definitions below in this more general setting.

We will also find it useful, and in fact modern algorithms reflect this usefulness, to consider functions $f : \mathbb{R}^n \to \mathbb{R}^*$ where $\mathbb{R}^*$ is the set of extended real numbers introduced earlier$^8$.

Before beginning with the main part of the discussion, we want to keep a couple of examples in mind.

$^8$See p. 10 of the handout on Preliminary Material.
The primal example of a convex function is \( x \mapsto x^2, \ x \in \mathbb{R} \). As we learn in elementary calculus, this function is infinitely often differentiable and has a single critical point at which the function in fact takes on, not just a relative minimum, but an absolute minimum.

![Figure 10: The Generic Parabola](image)

A critical point is, by definition, the solution of the equation \( \frac{d}{dx} x^2 = 2x \) or \( 2x = 0 \). We can apply the second derivative test at the point \( x = 0 \) to determine the nature of the critical point and we find that, since \( \frac{d^2}{dx^2} (x^2) = 2 > 0 \), the function is "concave up" and the critical point is indeed a point of relative minimum. That this point gives an absolute minimum to the function, we need only remark that the function values are bounded below by zero since \( x^2 > 0 \) for all \( x \neq 0 \).

We can give a similar example in \( \mathbb{R}^2 \).

**Example 4.1** We consider the function

\[
(x, y) \mapsto \frac{1}{2} x^2 + \frac{1}{3} y^2 := z,
\]

whose graph appears in the next figure.

This is an elliptic paraboloid. In this case we expect that, once again, the minimum will occur at the origin of coordinates and, setting \( f(x, y) = z \), we can compute
Figure 11: An Elliptic Parabola

\[ \text{grad } (f)(x, y) = \begin{pmatrix} x \\ \frac{2}{3} y \end{pmatrix}, \quad \text{and } H(f(x, y)) = \begin{pmatrix} 1 & 0 \\ 0 & \frac{2}{3} \end{pmatrix}. \]

Notice that, in our terminology, the Hessian matrix \( H(f) \) is positive definite at all points \((x, y) \in \mathbb{R}^2\). Here the critical points are exactly those for which \( \text{grad } [f(x, y)] = 0 \) whose only solution is \( x = 0, y = 0 \). The second derivative test is just that

\[ \det H(f(x, y)) = \left( \frac{\partial^2 f}{\partial x^2} \right) \left( \frac{\partial^2 f}{\partial y^2} \right) - \frac{\partial^2 f}{\partial x \partial y} > 0 \]

which is clearly satisfied.

Again, since for all \((x, y) \neq (0, 0), z > 0\), the origin is a point where \( f \) has an absolute minimum.

As the idea of convex set lies at the foundation of our analysis, we want to describe the notion of convex functions in terms of convex sets. We recall that, if \( A \) and \( B \) are two non-empty sets, then the Cartesian product of these two sets \( A \times B \) is defined as the set of ordered pairs \( \{(a, b) : a \in A, b \in B\} \). Notice that order does matter here and that \( A \times B \neq B \times A! \) Simple examples are

1. Let \( A = [-1, 1], B = [-1, 1] \) so that \( A \times B = \{(x, y) : -1 \leq x \leq 1, -1 \leq y \leq 1\} \) which is just the square centered at the origin, of side two.
2. \( \mathbb{R}^2 \) itself can be identified (and we usually do!) with the Cartesian product \( \mathbb{R} \times \mathbb{R} \).

3. let \( C \subset \mathbb{R}^2 \) be convex and let \( S := \mathbb{R}^+ \times C \). Then \( S \) is called a right cylinder and is just \( \{(z, x) \in \mathbb{R}^3 : z > 0, x \in C \} \). If, in particular \( C = \{(u, v) \in \mathbb{R}^2 : u^2 + v^2 \leq 1 \} \), then \( S \) is the usual right circulinder lying above the \( x, y \)-plane (without the bottom!).

This last example shows us a situation where \( A \times B \) is convex. In fact it it a general result that if \( A \) and \( B \) are two non-empty convex sets in a vector space \( V \), then \( A \times B \) is likewise a convex set in \( V \times V \).

**Exercise 4.2** Prove this last statement. (Take \( V = \mathbb{R}^n \) if you must!)

Previously, we introduced the idea of the epigraph of a function \( f : X \rightarrow \mathbb{R} \) where \( X \subset \mathbb{R}^n \). For convenience, we repeat the definition here.

**Definition 4.3** Let \( X \subset \mathbb{R}^n \) be a non-empty set. If \( f : X \rightarrow \mathbb{R} \) then \( \text{epi} (f) \) is defined by

\[
\text{epi} (f) := \{(x, z) \in X \times \mathbb{R} | z \geq f(x)\}.
\]

Convex functions are defined in terms of their epigraphs:

**Definition 4.4** Let \( C \subset \mathbb{R}^n \) be convex and \( f : C \rightarrow \mathbb{R}^* \). Then the function \( f \) is called a convex function provided \( \text{epi} (f) \subset \mathbb{R} \times \mathbb{R}^n \) is a convex set.

We emphasize that this definition has the advantage of directly relating the theory of convex sets to the theory of convex functions. However, a more traditional definition is that a function is convex provided that, for any \( x, y \in C \) and any \( \lambda \in [0, 1] \)

\[
f ( (1 - \lambda) x + \lambda y) \leq (1 - \lambda) f(x) + \lambda f(y),
\]

which is sometimes referred to as Jensen’s inequality.

In fact, these definitions turn out to be equivalent. Indeed, we have the following result.

**Theorem 4.5** Let \( C \subset \mathbb{R}^n \) be convex and \( f : C \rightarrow \mathbb{R}^* \). Then the following are equivalent:
(a) \( \text{epi}(f) \) is convex.

(b) For all \( \lambda_1, \lambda_2, \ldots, \lambda_n \) with \( \lambda_i \geq 0 \) and \( \sum_{i=1}^{n} \lambda_i = 1 \), and points \( x^{(i)} \in C, i = 1, 2, \ldots, n \), we have

\[
f \left( \sum_{i=1}^{n} \lambda_i x^{(i)} \right) \leq \sum_{i=1}^{n} \lambda_i f(x^{(i)}).
\]

(c) For any \( x, y \in C \) and \( \lambda \in [0, 1] \),

\[
f \left( (1 - \lambda)x + \lambda y \right) \leq (1 - \lambda)f(x) + \lambda f(y).
\]

**Proof:** To see that (a) implies (b) we note that, if for all \( i = 1, 2, \ldots, n \), \( x^{(i)}, f(x^{(i)}) \in \text{epi}(f) \), then since this latter set is convex, we have

\[
\sum_{i=1}^{n} \lambda_i (x^{(i)}, f(x^{(i)})) = \left( \sum_{i=1}^{n} \lambda_i x^{(i)}, \sum_{i=1}^{n} \lambda_i f(x^{(i)}) \right) \in \text{epi}(f),
\]

which, in turn, implies that

\[
f \left( \sum_{i=1}^{n} \lambda_i x^{(i)} \right) \leq \sum_{i=1}^{n} \lambda_i f(x^{(i)}).
\]

This establishes (b). It is obvious that (b) implies (c). So it remains only to show that (c) implies (a) in order to establish the equivalence.

To this end, suppose that \( (x^{(1)}, z_1), (x^{(2)}, z_2) \in \text{epi}(f) \) and take \( 0 \leq \lambda \leq 1 \). Then

\[
(1 - \lambda) (x^{(1)}, z_1) + \lambda (x^{(2)}, z_2) = \left( (1 - \lambda) x^{(1)} + \lambda x^{(2)}, (1 - \lambda) z_1 + \lambda z_2 \right),
\]

and since \( f(x^{(1)}) \leq z_1 \) and \( f(x^{(2)}) \leq z_2 \) we have, since \( (1 - \lambda) > 0 \), and \( \lambda > 0 \), that

\[
(1 - \lambda) f(x^{(1)}) + \lambda f(x^{(2)}) \leq (1 - \lambda) z_1 + \lambda z_2.
\]

Hence, by the assumption (c), \( f \left( (1 - \lambda) x^{(1)} + \lambda x^{(2)} \right) \leq (1 - \lambda) z_1 + \lambda z_2 \), which shows the point \( (1 - \lambda) x^{(1)} + \lambda x^{(2)}, (1 - \lambda) z_1 + \lambda z_2 \) is in \( \text{epi}(f) \). \( \square \)

We pause to remark that some authors, particularly in applications to Economics, discuss **concave functions**. These latter functions are simply related to convex functions. Indeed a function \( f \) is concave if and only if the function \(-f\) is convex\(^9\). Again, we take as our basic object of study the class of **convex** functions.

\(^9\)This will also be true of quasi-convex and quasi-concave functions which we will define below.
We can see another connection between convex sets and convex functions if we introduce the indicator function, $\psi_K$ of a set $K \subset \mathbb{R}^n$. Indeed, $\psi_K : \mathbb{R}^n \rightarrow \mathbb{R}^*$ is defined by

$$
\psi_K(x) = \begin{cases} 
0 & \text{if } x \in K, \\
+\infty & \text{if } x \notin K.
\end{cases}
$$

**Proposition 4.6** A non-empty subset $D \subset \mathbb{R}^n$ is convex if and only if its indicator function is convex.

**Proof:** The result follows immediately from the fact that $\text{epi}(\psi_D) = D \times \mathbb{R}_{\geq 0}$.

Certain simple properties follow immediately from the analytic form of the definition (part (c) of the equivalence theorem above). Indeed, it is easy to see, and we leave it as an exercise for the reader, that if $f$ and $g$ are convex functions defined on a convex set $C$, then $f + g$ is likewise convex on $C$ provided there is no point for which $f(x) = +\infty$ and $g(x) = -\infty$. The same is true if $\beta \in \mathbb{R}, \beta > 0$ and we consider $\beta f$.

Moreover, we have the following simple result which is useful.

**Proposition 4.7** Let $f : \mathbb{R}^n \rightarrow \mathbb{R}$ be given, $x^{(1)}, x^{(2)} \in \mathbb{R}^n$ be fixed and define a function $\varphi : [0, 1] \rightarrow \mathbb{R}$ by $\varphi(\lambda) := f((1 - \lambda)x^{(1)} + \lambda x^{(2)})$. Then the function $f$ is convex on $\mathbb{R}^n$ if and only if the function $\varphi$ is convex on $[0, 1]$.

**Proof:** Suppose, first, that $f$ is convex on $\mathbb{R}^n$. Then it is sufficient to show that $\text{epi}(\varphi)$ is a convex subset of $\mathbb{R}^2$. To see this, let $(\lambda_1, z_1), (\lambda_2, z_2) \in \text{epi}(\varphi)$ and let

$$
\begin{align*}
\hat{y}^{(1)} &= \lambda_1 x^{(1)} + (1 - \lambda_1) x^{(2)}, \\
\hat{y}^{(2)} &= \lambda_2 x^{(1)} + (1 - \lambda_2) x^{(2)}.
\end{align*}
$$

Then

$$
f(\hat{y}^{(1)}) = \varphi(\lambda_1) \leq z_1 \quad \text{and} \quad f(\hat{y}^{(2)}) = \varphi(\lambda_2) \leq z_2.
$$

Hence $(\hat{y}^{(1)}, z_1) \in \text{epi}(f)$ and $(\hat{y}^{(2)}, z_2) \in \text{epi}(f)$. Since $\text{epi}(f)$ is a convex set, we also have $(\mu \hat{y}^{(1)} + (1 - \mu) \hat{y}^{(2)}, \mu z_1 + (1 - \mu) z_2) \in \text{epi}(f)$ for every $\mu \in [0, 1]$. It follows that $f(\mu \hat{y}^{(1)} + (1 - \mu) \hat{y}^{(2)}) \leq \mu z_1 + (1 - \mu) z_2)$.
Now

\[\begin{align*}
\mu \textbf{y}^{(1)} + (1 - \mu) \textbf{y}^{(2)} &= \mu (\lambda_1 \textbf{x}^{(1)} + (1 - \lambda_1) \textbf{x}^{(2)}) + (1 - \mu) (\lambda_2 \textbf{x}^{(1)} + (1 - \lambda_2) \textbf{x}^{(2)}) \\
&= (\mu \lambda_1 + (1 - \mu) \lambda_2) \textbf{x}^{(1)} + \mu (1 - \lambda_1) + (1 - \mu) (1 - \lambda_2) \textbf{x}^{(2)} ,
\end{align*}\]

and since

\[1 - [\mu \lambda_1 + (1 - \mu) \lambda_2]] = [\mu + (1 - \mu)] - [\mu \lambda_1 + (1 - \mu) \lambda_2]]
\]

\[= \mu (1 - \lambda_1) + (1 - \mu) (1 - \lambda_2) ,
\]

we have from the definition of \(\varphi\) that \(f(\mu \textbf{y}^{(1)} + (1 - \mu) \textbf{y}^{(2)}) = \varphi(\mu \lambda_1 + (1 - \mu) \lambda_2)\) and so \((\mu \lambda_1 + (1 - \mu) \lambda_2, \mu z_1 + (1 - \mu) z_2) \in \text{epi}(\varphi)\) i.e., \(\varphi\) is convex.

We leave the proof of the converse statement as an exercise. \(\square\)

We also point out that if \(f : \mathbb{R}^n \rightarrow \mathbb{R}\) is a linear or affine, then \(f\) is convex. Indeed, suppose that for a vector \(\textbf{a} \in \mathbb{R}^n\) and a real number \(b\), the function \(f\) is given by \(f(\textbf{x}) = \langle \textbf{a}, \textbf{x} \rangle + b\). Then we have, for any \(\lambda \in [0, 1]\),

\[f((1 - \lambda) \textbf{x} + \lambda \textbf{y}) = \langle \textbf{a}, (1 - \lambda) \textbf{x} + \lambda \textbf{y} \rangle + b
\]

\[= (1 - \lambda) \langle \textbf{a}, \textbf{x} \rangle + \lambda \langle \textbf{a}, \textbf{y} \rangle + (1 - \lambda) b + \lambda b
\]

\[= (1 - \lambda) \langle \textbf{a}, \textbf{x} \rangle + b) + \lambda (\langle \textbf{a}, \textbf{y} \rangle + b) = (1 - \lambda) f(\textbf{x}) + \lambda f(\textbf{y}) ,
\]

and so \(f\) is convex, the weak inequality being an equality in this case.

In the case that \(f\) is linear, that is \(f(\textbf{x}) = \langle \textbf{a}, \textbf{x} \rangle\) for some \(\textbf{a} \in \mathbb{R}^n\) then it is easy to see that the map \(\varphi : \textbf{x} \rightarrow [f(\textbf{x})]^2\) is also convex. Indeed, if \(\textbf{x}, \textbf{y} \in \mathbb{R}^n\) then, setting \(\alpha = f(\textbf{x})\) and \(\beta = f(\textbf{y})\), and taking \(0 < \lambda < 1\) we have

\[(1 - \lambda) \varphi(\textbf{x}) + \lambda \varphi(\textbf{y}) - \varphi((1 - \lambda) \textbf{x} + \lambda \textbf{y})
\]

\[= (1 - \lambda) \alpha^2 + \lambda \beta^2 - ((1 - \lambda) \alpha + \lambda \beta)^2
\]

\[= (1 - \lambda) \lambda (\alpha - \beta)^2 \geq 0 .
\]

Note, that in particular for the function \(f : \mathbb{R} \rightarrow \mathbb{R}\) given by \(f(x) = x\) is linear and that \([f(x)]^2 = x^2\) so that we have a proof that the function that we usually write \(y = x^2\) is a convex function.

The next result expands our repertoire of convex functions.
Proposition 4.8  
(a) If $A : \mathbb{R}^m \to \mathbb{R}^n$ is linear and $f : \mathbb{R}^n \to \mathbb{R}^*$ is convex, then $f \circ A$ is convex as a map from $\mathbb{R}^m$ to $\mathbb{R}$.

(b) If $f$ is as in part (a) and $\varphi : \mathbb{R} \to \mathbb{R}$ is convex and non-decreasing, then $\varphi \circ f : \mathbb{R}^n \to \mathbb{R}$ is convex.

(c) Let $\{f_\alpha\}$ be a family of functions $f_\alpha : \mathbb{R}^n \to \mathbb{R}^*$ then its upper envelope $\sup_{\alpha \in A} f_\alpha$ is convex.

Proof: To prove (a) we use Jensen’s inequality:

Given any $x, y \in \mathbb{R}^n$ and $\lambda \in [0,1]$ we have

$$(f \circ A)((1 - \lambda)x + \lambda y) = f((1 - \lambda)(Ax) + \lambda(Ay)) \leq (1 - \lambda)f(Ax) + \lambda f(Ay)$$

$$= (1 - \lambda)(f \circ A)(x) + \lambda(f \circ A)(y).$$

For part (b), again we take $x, y \in \mathbb{R}^n$ and $\lambda \in [0,1]$. Then

$$(\varphi \circ f)[(1 - \lambda)x + \lambda y] \leq \varphi[(1 - \lambda)f(x) + \lambda f(y)]$$

$$\leq (1 - \lambda)\varphi(f(x)) + \lambda\varphi(f(y)) = (1 - \lambda)(\varphi \circ f)(x) + \lambda(\varphi \circ f)(y),$$

where the first inequality comes from the convexity of $f$ and the monotonicity of $\varphi$ and the second from the convexity of this later function. This proves part (b).

To establish part (c) we note that, since the arbitrary intersection of convex sets is convex, it suffices to show that

$$\text{epi} \left( \sup_{\alpha \in A} f_\alpha \right) = \bigcup_{\alpha \in A} \text{epi} (f_\alpha).$$

To check the equality of these two sets, start with a point

$$(x, z) \in \text{epi} \left( \sup_{\alpha \in A} f_\alpha \right).$$

Then $z \geq \sup_{\alpha \in A} f_\alpha(x)$ and so, for all $\beta \in A$, $z \geq f_\beta(x)$. Hence, by definition, $(x, z) \in \text{epi} f_\beta$ for all $\beta$ and therefore
\[(x, z) \in \bigcap_{\alpha \in A} \text{epi } (f_\alpha).\]

Conversely, suppose \((x, z) \in \text{epi } (f_\alpha)\) for all \(\alpha \in A\). Then \(z \geq f_\alpha(x)\) for all \(\alpha \in A\) and hence \(z \geq \sup_{\alpha \in A} f_\alpha\). But this, by definition, implies \((x, z) \in \text{epi } (\sup_{\alpha \in A} f_\alpha)\). This completes the proof of part (c) and the proposition. \(\qed\)

Next, we introduce the definition:

**Definition 4.9** Let \(f : \mathbb{R}^n \rightarrow \mathbb{R}^*\), and \(\alpha \in \mathbb{R}\). Then the sets

\[S(f, \alpha) := \{x \in \mathbb{R}^n \mid f(x) < \alpha\}\]

and \(\overline{S}(f, \alpha) := \{x \in \mathbb{R}^n \mid f(x) \leq \alpha\}\),

are called lower sections of the function \(f\).

**Proposition 4.10** If \(f : \mathbb{R} \rightarrow \mathbb{R}^*\) is convex, then its lower sections are likewise convex.

The proof of this result is trivial and we omit it.

The converse of this last proposition is false as can be easily seen from the function \(x \mapsto x^{\frac{1}{2}}\) from \(\mathbb{R}_{>0}\) to \(\mathbb{R}\). However, the class of functions whose lower level sets \(\overline{S}(f, \alpha)\) (or equivalently the sets \(S(f, \alpha)\)) are all convex is likewise an important class of functions and are called quasi-convex. These functions appear in game theory nonlinear programming (optimization) problems and mathematical economics. For example, quasi-convex utility functions imply that consumers have convex preferences. They are obviously generalizations of convex functions since every convex function is clearly quasi-convex. However they are not as easy to work with. In particular, while the sum of two convex functions is convex, the same is not true of quasi-convex functions as the following example shows.

**Example 4.11** Define

\[f(x) = \begin{cases} 0 & x \leq -2 \\ -(x + 2) & -2 < x \leq -1 \\ x & -1 < x \leq 0 \\ 0 & x > 0 \end{cases} \text{ and } g(x) = \begin{cases} 0 & x \leq 0 \\ -x & 0 < x \leq 1 \\ x - 2 & 1 < x \leq 2 \\ 0 & x > 2 \end{cases} .\]

Here, the functions are each concave, the level sections are convex for each function so that each is quasi-convex, and yet the level section corresponding to \(\alpha = -1/2\) for the sum \(f + g\) is not convex. Hence the sum is not quasi-convex.
It is useful for applications to have an analytic criterion for quasi-convexity. This is the content of the next result.

**Proposition 4.12** A function \( f : \mathbb{R}^n \to \mathbb{R}^* \) is quasi-convex if and only if, for any \( x, y \in \mathbb{R}^n \) and any \( \lambda \in [0, 1] \) we have

\[
f((1 - \lambda) x + \lambda y) \leq \max\{f(x), f(y)\}.
\]

**Proof:** Suppose that the sets \( \overline{S}(f, \alpha) \) are convex for every \( \alpha \). Let \( x, y \in \mathbb{R}^n \) and let \( \bar{\alpha} := \max\{f(x), f(y)\} \). Then \( \overline{S}(f, \bar{\alpha}) \) is convex and, since both \( f(x) \leq \bar{\alpha} \) and \( f(y) \leq \bar{\alpha} \), we have that both \( x \) and \( y \) belong to \( \overline{S}(f, \bar{\alpha}) \). Since this latter set is convex, we have

\[
(1 - \lambda) x + \lambda y \in \overline{S}(f, \bar{\alpha}) \text{ or } f((1 - \lambda) x + \lambda y) \leq \bar{\alpha} = \max\{f(x), f(y)\}.
\]

As we have seen above, the sum of two quasi-convex functions may well not be quasi-convex. With this analytic test for quasi-convexity, we can check that there are certain operations which preserve quasi-convexity. We leave the proof of the following result to the reader.

**Proposition 4.13** (a) If the functions \( f_1, \ldots, f_k \) are quasi-convex and \( \alpha_1, \ldots, \alpha_k \) are non-negative real numbers, then the function \( f := \max\{\alpha_1 f_1, \ldots, \alpha_k f_k\} \) is quasi-convex.

(b) If \( \varphi : \mathbb{R} \to \mathbb{R} \) is a non-decreasing function and \( f : \mathbb{R}^n \to \mathbb{R} \) is quasi-convex, then the composition \( \varphi \circ f \) is a quasi-convex function.

We now return to the study of convex functions.

A simple sketch of the parabola \( y = x^2 \) and any horizontal cord (which necessarily lies above the graph) will convince the reader that all points in the domain corresponding to the values of the function which lie below that horizontal line, form a convex set in the domain. Indeed, this is a property of convex functions which is often useful.

**Proposition 4.14** If \( C \subset \mathbb{R}^n \) is a convex set and \( f : C \to \mathbb{R} \) is a convex function, then the level sets \( \{x \in C \mid f(x) \leq \alpha\} \) and \( \{x \in C \mid f(x) < \alpha\} \) are convex for all scalars \( \alpha \).
Proof: We leave this proof as an exercise.

Notice that, since the intersection of convex sets is convex, the set of points simultaneously satisfying $m$ inequalities $f_1(x) \leq c_1, f_2(x) \leq c_2, \ldots, f_m(x) \leq c_m$ where each $f_i$ is a convex function, defines a convex set. In particular, the polygonal region defined by a set of such inequalities when the $f_i$ are affine is convex.

From this result, we can obtain an important fact about points at which a convex function attains a minimum.

**Proposition 4.15** Let $C \subset \mathbb{R}$ be a convex set and $f : C \rightarrow \mathbb{R}$ a convex function. Then the set of points $M \subset C$ at which $f$ attains its minimum is convex. Moreover, any relative minimum is an absolute minimum.

**Proof:** If the function does not attain its minimum at any point of $C$, then the set of such points in empty, which is a convex set. So, suppose that the set of points at which the function attains its minimum is non-empty and let $m$ be the minimal value attained by $f$. If $x, y \in M$ and $\lambda \in [0, 1]$ then certainly $(1 - \lambda)x + \lambda y \in C$ and so

$$m \leq f((1 - \lambda)x + \lambda y) \leq (1 - \lambda)f(x) + \lambda f(y) = m,$$

and so the point $(1 - \lambda)x + \lambda y \in M$. Hence $M$, the set of minimal points, is convex.

Now, suppose that $x^* \in C$ is a relative minimum point of $f$, but that there is another point $\hat{x} \in C$ such that $f(\hat{x}) < f(x^*)$. On the line $(1 - \lambda)\hat{x} + \lambda x^*, 0 < \lambda < 1,$ we have

$$f((1 - \lambda)\hat{x} + \lambda x^*) \leq (1 - \lambda)f(\hat{x}) + \lambda f(x^*) < f(x^*),$$

contradicting the fact that $x^*$ is a relative minimum point. \hfill \Box

Again, the example of the simple parabola, shows that the set $M$ may well contain only a single point, i.e., it may well be that the minimum point is unique. We can guarantee that this is the case for an important class of convex functions.

**Definition 4.16** A real-valued function $f$, defined on a convex set $C \subset \mathbb{R}$ is said to be strictly convex provided, for all $x, y \in C, x \neq y$ and $\lambda \in (0, 1)$, we have

$$f((1 - \lambda)x + \lambda y) < (1 - \lambda)f(x) + \lambda f(y).$$
Proposition 4.17 If $C \subset \mathbb{R}^n$ is a convex set and $f : C \rightarrow \mathbb{R}$ is a strictly convex function then $f$ attains its minimum at, at most, one point.

Proof: Suppose that the set of minimal points $M$ is not empty and contains two distinct points $x$ and $y$. Then, for any $0 < \lambda < 1$, since $M$ is convex, we have $(1 - \lambda)x + \lambda y \in M$. But $f$ is strictly convex. Hence

$$m = f((1 - \lambda)x + \lambda y) < (1 - \lambda)f(x) + \lambda f(y) = m,$$

which is a contradiction. \qed

If a function is differentiable then, as in the case in elementary calculus, we can give characterizations of convex functions using derivatives. If $f$ is a continuously differentiable function defined on an open convex set $C \subset \mathbb{R}^n$ then we denote its gradient at $x \in C$, as usual, by $\nabla f(x)$. The excess function

$$E(x, y) := f(y) - f(x) - \langle \nabla f(x), y - x \rangle$$

is a measure of the discrepancy between the value of $f$ at the point $y$ and the value of the tangent approximation at $x$ to $f$ at the point $y$. This is illustrated in the next figure.

Figure 12: The Tangent Approximation

Now we introduce the notion of a monotone derivative

54
Definition 4.18 The map $x \mapsto \nabla f(x)$ is said to be monotone on $C \subset \mathbb{R}^n$ provided

$$\langle \nabla f(y) - \nabla f(x), y - x \rangle \geq 0,$$

for all $x, y \in C$.

We can now characterize convexity in terms of the function $E$ and the monotonicity concept just introduced. However, before stating and proving the next theorem, we need a lemma.

Lemma 4.19 Let $f$ be a real-valued, differentiable function defined on an open interval $I \subset \mathbb{R}$. Then if the first derivative $f'$ is a non-decreasing function on $I$, the function $f$ is convex on $I$.

Proof: Choose $x, y \in I$ with $x < y$, and for any $\lambda \in [0, 1]$, define $z_\lambda := (1 - \lambda)x + \lambda y$. By the Mean Value Theorem, there exist $u, v \in \mathbb{R}, x \leq v \leq z_\lambda \leq u \leq y$ such that

$$f(y) = f(z_\lambda) + (y - z_\lambda) f'(u), \quad \text{and} \quad f(z_\lambda) = f(x) + (z_\lambda - x) f'(v).$$

But, $y - z_\lambda = y - (1 - \lambda)x - \lambda y = (1 - \lambda)(y - x)$ and $z_\lambda - x = (1 - \lambda)x + \lambda y - x = \lambda(y - x)$ and so the two expressions above may be rewritten as

$$f(y) = f(z_\lambda) + \lambda (y - x) f'(u), \quad \text{and} \quad f(z_\lambda) = f(x) + \lambda (y - x) f'(v).$$

Since, by choice, $v < u$, and since $f'$ is non-decreasing, this latter equation yields

$$f(z_\lambda) \leq f(x) + \lambda (y - x) f'(u).$$

Hence, multiplying this last inequality by $(1 - \lambda)$ and the expression for $f(y)$ by $-\lambda$ and adding we get

$$(1 - \lambda) f(z_\lambda) - \lambda f(y) \leq (1 - \lambda) f(x) + \lambda(1 - \lambda)(y - x)f'(u) - \lambda f(z_\lambda) - \lambda(1 - \lambda)(y - x)f'(u),$$

which we then rearrange to yield

$$(1 - \lambda) f(z_\lambda) + \lambda f(z_\lambda) = f(z_\lambda) \leq (1 - \lambda) f(x) + \lambda f(y),$$

and this is just the condition for the convexity of $f$. \qed
We can now prove a theorem which gives three different characterizations of convexity for continuously differentiable functions.

**Theorem 4.20** Let $f$ be a continuously differentiable function defined on an open convex set $C \subset \mathbb{R}^n$. Then the following are equivalent:

(a) $E(x, y) \geq 0$ for all $x, y \in C$;
(b) the map $x \mapsto \nabla f(x)$ is monotone in $C$;
(c) the function $f$ is convex on $C$.

**Proof:** Suppose that (a) holds, i.e. $E(x, y) \geq 0$ on $C \times C$. Then we have both

$$f(y) - f(x) \geq \langle \nabla f(x), y - x \rangle,$$

and

$$f(x) - f(y) \geq \langle \nabla f(y), x - y \rangle = -\langle \nabla f(y), y - x \rangle.$$

Then, from the second inequality, $f(y) - f(x) \leq \langle \nabla f(y), x - y \rangle$, and so

$$\langle \nabla f(y) - \nabla f(x), y - x \rangle = \langle \nabla f(y), y - x \rangle - \langle \nabla f(x), y - x \rangle \geq (f(y) - f(x)) - (f(y) - f(x)) = 0.$$

Hence, the map $x \mapsto \nabla f(x)$ is monotone in $C$.

Now suppose the map $x \mapsto \nabla f(x)$ is monotone in $C$, and choose $x, y \in C$. Define a function $\varphi : [0, 1] \rightarrow \mathbb{R}$ by $\varphi(t) := f(x + t(y - x))$. We observe, first, that if $\varphi$ is convex on $[0, 1]$ then $f$ is convex on $C$. To see this, let $u, v \in [0, 1]$ be arbitrary. On the one hand,

$$\varphi((1-\lambda)u+\lambda v) = f(x + ((1-\lambda)u+\lambda v)(y-x)) = f\left((1-[(1-\lambda)u+\lambda v])x+(1-\lambda)u+\lambda v\right) y),$$

while, on the other hand,

$$\varphi((1-\lambda)u+\lambda v) \leq (1-\lambda) f(x + u(y - x)) + f(x + v(y - x)).$$

Setting $u = 0$ and $v = 1$ in the above expressions yields
\[ f((1-\lambda)x + \lambda y) \leq (1-\lambda)f(x) + \lambda f(y). \]

so the convexity of \( \varphi \) on \([0,1]\) implies the convexity of \( f \) on \( C \).

Now, choose any \( \alpha, \beta, 0 \leq \alpha < \beta \leq 1 \). Then

\[
\varphi'(\beta) - \varphi'(\alpha) = \langle (\nabla f(x + \beta(y-x)) - \nabla f(x + \alpha(y-x)), y - x \rangle.
\]

Setting \( u := x + \alpha(y-x) \) and \( v := x + \beta(y-x) \)\(^{10}\) we have \( v - u = (\beta - \alpha)(y-x) \) and so

\[
\varphi'(\beta) - \varphi'(\alpha) = \langle (\nabla f(v) - \nabla f(u), v - u \rangle \geq 0.
\]

Hence \( \varphi' \) is non-decreasing, so that the function \( \varphi \) is convex.

Finally, if \( f \) is convex on \( C \), then, for fixed \( x, y \in C \) define

\[
h(\lambda) := (1-\lambda)f(x) + \lambda f(y) - f((1-\lambda)x + \lambda y) \]

Then \( \lambda \mapsto h(\lambda) \) is a non-negative, differentiable function on \([0,1]\) and attains its minimum at \( \lambda = 0 \). Therefore \( 0 \leq h'(0) = E(x, y) \), and the proof is complete. \( \square \)

As an immediate corollary, we have

**Corollary 4.21** Let \( f \) be a continuously differentiable convex function defined on a convex set \( C \). If there is a point \( x^* \in C \) such that, for all \( y \in C \), \( \langle \nabla f(x^*), y - x^* \rangle \geq 0 \), then \( x^* \) is an absolute minimum point of \( f \) over \( C \).

**Proof:** By the preceeding theorem, the convexity of \( f \) implies that

\[
f(y) - f(x^*) \geq \langle \nabla f(x^*), y - x^* \rangle,
\]

and so, by hypothesis,

\[
f(y) \geq f(x^*) + \langle \nabla f(x^*), y - x^* \rangle \geq f(x^*).
\]

\( \square \)

The inequality \( E(x, y) \geq 0 \) shows that local information about a convex function, given in terms of the derivative at a point) gives us global information in terms of a global

\(^{10}\)Note that \( u \) and \( v \) are convex combinations of points in the convex set \( C \) and so \( u, v \in C \).
underestimator of the function $f$. In a way, this is the key property of convex functions. For example, suppose that $\nabla f(x) = 0$. Then, for all $y \in \text{dom}(f)$, $f(y) \geq f(x)$ so that $x$ is a global minimizer of the convex function $f$.

It is also important to remark that the hypothesis that the convex function $f$ is defined on a convex set is crucial, both for the first order conditions as well as for the second order conditions. Indeed, if we consider the function $f(x) = 1/x^2$ with domain $\{x \in \mathbb{R} | x \neq 0\}$. The usual second order condition $f''(x) > 0$ for all $x \in \text{dom}(f)$ yet $f$ is not convex there so that the second order test fails.

The condition $E(x, y) \geq 0$ can be given an important geometrical interpretation in terms of epigraphs. Indeed if $f$ is convex and $x, y \in \text{dom}(f)$ then for $(x, z) \in \text{epi}(f)$, then

$$z \geq f(y) \geq f(x) + \nabla f(x) \top (y - x),$$

can be expressed as

$$
\begin{bmatrix}
\nabla f(x) \\
-1
\end{bmatrix} \top \begin{bmatrix}
y \\
z
\end{bmatrix} - \begin{bmatrix}
x \\
f(x)
\end{bmatrix} \leq 0.
$$

This shows that the hyperplane defined by $(\nabla f(x), -1)^\top$ supports $\text{epi}(f)$ at the boundary point $(x, f(x))$.

We now turn to so-called second order criteria for convexity. The discussion involves the Hessian matrix of a twice continuously differentiable function, and depends on the question of whether this matrix is positive semi-definite or even positive definite (for strict convexity)\textsuperscript{11}. Let us recall some definitions.

**Definition 4.22** A real symmetric $n \times n$ matrix $A$ is said to be

(a) **Positive definite** provided $x^\top A x > 0$ for all $x \in \mathbb{R}^n$, $x \neq 0$.

(b) **Negative definite** provided $x^\top A x < 0$ for all $x \in \mathbb{R}^n$, $x \neq 0$.

(c) **Positive semidefinite** provided $x^\top A x \geq 0$ for all $x \in \mathbb{R}^n$, $x \neq 0$.

(d) **Negative semidefinite** provided $x^\top A x \leq 0$ for all $x \in \mathbb{R}^n$, $x \neq 0$.

(e) **Indefinite** provided $x^\top A x$ takes on values that differ in sign.

\textsuperscript{11}Notice that the smoothness assumption on the function $f$ is sufficient to insure that the Hessian matrix of second partial derivatives is symmetric.
It is important to be able to determine if a matrix is indeed positive definite. In order to do this, a number of criteria have been developed. Perhaps the most important characterization is in terms of the eigenvalues.

**Theorem 4.23** Let $A$ be a real symmetric $n \times n$ matrix. Then $A$ is positive definite if and only if all its eigenvalues are positive.

**Proof:** If $A$ is positive definite and $\lambda$ is an eigenvalue of $A$, then, for any eigenvector $x$ belonging to $\lambda$

$$x^\top A x = \lambda x^\top x = \lambda \|x\|^2, .$$

Hence

$$\lambda = \frac{x^\top A x}{\|x\|^2} > 0$$

Conversely, suppose that all the eigenvalues of $A$ are positive. Let $\{x_1, \ldots, x_n\}$ be an orthonormal set of eigenvectors of $A$.\(^\text{12}\) Hence any $x \in \mathbb{R}^n$ can be written as

$$x = \alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_n x_n$$

with

$$\alpha_i = x^\top x_i \text{ for } i = 1, 2, \ldots, n, \text{ and } \sum_{i=1}^n \alpha_i^2 = \|x\|^2 > 0.$$ 

It follows that

$$x^\top A x = (\alpha_1 x_1 + \cdots + \alpha_n x_n)^\top (\alpha_1 \lambda_1 x_1 + \cdots + \alpha_n \lambda_n x_n)$$

$$= \sum_{i=1}^n \alpha_i^2 \lambda_i \geq (\min_i \{\lambda_i\}) \|x\| > 0.$$ 

Hence $A$ is positive definite. $\square$

In simple cases where we can compute the eigenvalues easily, this is a useful criterion.

\(^{12}\text{The so-called Spectral Theorem for real symmetric matrices states that such a matrix can be diagonalized, and hence has } n \text{ linearly independent eigenvectors. These can be replaced by a set of } n \text{ orthonormal eigenvectors.}\)
**Example 4.24** Let

\[
A = \begin{pmatrix} 2 & -2 \\ -2 & 5 \end{pmatrix}.
\]

Then the eigenvalues are the roots of

\[
\det (A - \lambda I) = (2 - \lambda) (5 - \lambda) - 4 = (\lambda - 1) (\lambda - 6).
\]

Hence the eigenvalues are both positive and hence the matrix is positive definite. In this particular case it is easy to check directly that \( A \) is positive definite. Indeed

\[
(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}) \begin{pmatrix} 2 & -2 \\ -2 & 5 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = (x_1, x_2) \begin{pmatrix} 2x_1 - 2x_2 \\ -2x_1 + 5x_2 \end{pmatrix} = 2x_2^2 - 4x_1x_2 + 5x_2^2 = 2[x_1^2 - 2x_1x_2 + x_2^2] + 4x_2^2 = 2(x_1 - x_2)^2 + 4x_2^2 > 0.
\]

This last theorem has some immediate useful consequences. First, if \( A \) is positive definite, then \( A \) must be nonsingular, since singular matrices have \( \lambda = 0 \) as an eigenvalue. Moreover, since we know that the \( \det(A) \) is the product of the eigenvalues, and since each eigenvalue is positive, then \( \det(A) > 0 \). Finally, we have the following result which depends on the notion of **leading principle submatrices**.

**Definition 4.25** Given any \( n \times N \) matrix \( A \), let \( A_r \) denote the matrix formed by deleting the last \( n - r \) rows and columns of \( A \). Then \( A_r \) is called the **leading principal submatrix** of \( A \).

**Proposition 4.26** If \( A \) is a symmetric positive definite matrix then the leading principal submatrices \( A_1, A_2, \ldots, A_n \) of \( A \) are all positive definite. In particular, \( \det(A_r) > 0 \).

**Proof:** Let \( \mathbf{x}_r = (x_1, x_2, \ldots, x_r)^\top \) be any non-zero vector in \( \mathbb{R}^r \). Set

\[
\mathbf{x} = (x_1, x_2, \ldots, x_r, 0, \ldots, 0)^\top.
\]

Since \( \mathbf{x}_r^\top A_r \mathbf{x}_r = \mathbf{x}^\top A \mathbf{x} > 0 \), it follows that \( A_r \) is positive definite, by definition. □

This proposition is half of the famous criterion of Sylvester for positive definite matrices.
**Theorem 4.27** A real, symmetric matrix $A$ is positive definite if and only if all of its leading principle minors are positive definite.

We will not prove this theorem here but refer the reader to his or her favorite treatise on linear algebra.

**Example 4.28** Let

$$
A = \begin{pmatrix}
2 & -1 & 0 \\
-1 & 2 & -1 \\
0 & -1 & 2
\end{pmatrix}
$$

Then

$$
A_2 = (2), \quad A_2 = \begin{pmatrix}
2 & -1 \\
-1 & 2
\end{pmatrix}, \quad A_3 = A.
$$

Then

$$
det A_1 = 2, \quad det A_2 = 4 - 1 = 3, \quad and \quad det A = 4.
$$

Hence, according to Sylvester’s criterion, the matrix $A$ is positive definite.

Now we are ready to look at second order conditions for convexity.

**Proposition 4.29** Let $D \subset \mathbb{R}^n$ be an open convex set and let $f : D \rightarrow \mathbb{R}$ be twice continuously differentiable in $D$. Then $f$ is convex if and only if the Hessian matrix of $f$ is positive semidefinite throughout $D$.

**Proof:** By Taylor’s Theorem we have

$$
f(y) = f(x) + \langle \nabla f(x), y - x \rangle + \frac{1}{2} \langle y - x, \nabla^2 f(x + \lambda (y - x))(y - x) \rangle,
$$

for some $\lambda \in [0, 1]$. Clearly, if the Hessian is positive semi-definite, we have

$$
f(y) \geq f(x) + \langle \nabla f(x), y - x \rangle,
$$

which in view of the definition of the excess function, means that $E(x, y) \geq 0$ which implies that $f$ is convex on $D$. 

61
Conversely, suppose that the Hessian is not positive semi-definite at some point \( x \in D \). Then, by the continuity of the Hessian, there is a \( y \in D \) so that, for all \( \lambda \in [0, 1] \),

\[
\langle y - x, \nabla^2 f(x + \lambda(y - x))(y - x) \rangle < 0,
\]

which, in light of the second order Taylor expansion implies that \( E(x, y) < 0 \) and so \( f \) cannot be convex. \( \square \)

Let us consider, as an example, the quadratic function \( f : \mathbb{R}^n \to \mathbb{R} \) with \( \text{dom}(f) = \mathbb{R}^n \), given by

\[
f(x) = \frac{1}{2} x^\top Q x + q^\top x + r,
\]

with \( Q \) and \( n \times n \) symmetric matrix, \( q \in \mathbb{R}^n \) and \( r \in \mathbb{R} \). Then since as we have seen previously, \( \nabla^2 f(x) = Q \), the function \( f \) is convex if and only if \( Q \) is positive semidefinite. Strict convexity of \( f \) is likewise characterized by the positive definiteness of \( Q \).

These first and second-order necessary conditions give us methods of showing that a given function is convex. Thus, we either check the definition, Jensen’s inequality, using the equivalence that is given by Theorem 2.1.3, or showing that the Hessian is positive semi-definite. Let us look at some simple examples.

**Example 4.30**  (a) The function real-valued function defined on \( \mathbb{R}^+ \) \(^{13}\) given by \( x \mapsto x \ln(x) \). Then, since this function \( C^2(\mathbb{R}^+) \) and \( f'(x) = \ln(x) + 1 \) and \( f''(x) = 1/x > 0 \), we see that \( f \) is (even strictly) convex.

(b) The max function \( f(x) = \max\{x_1, \ldots, x_n\} \) is convex on \( \mathbb{R}^n \). Here we can use Jensen’s inequality. Let \( \lambda \in [0, 1] \) then

\[
f((1 - \lambda) x + \lambda y) = \max_{1 \leq i \leq n} (\lambda x_i + \lambda y_i) \leq \lambda \max_{1 \leq i \leq n} x_i + (1 - \lambda) \max_{1 \leq i \leq n} y_i
\]

\[
= (1 - \lambda) f(x) + \lambda f(y).
\]

(c) The function \( q : \mathbb{R} \times \mathbb{R}^+ \to \mathbb{R} \) given by \( q(x, y) = x^2/y \) is convex. In this case, \( \nabla q(x, y) = (2x/y, -x^2/y^2)^\top \) while an easy computation shows

\[
\nabla^2 q(x, y) = \frac{2}{y^3} \begin{pmatrix} y^2 & -xy \\ -xy & x^2 \end{pmatrix}.
\]

\(^{13}\)Recall that by \( \mathbb{R}^+ \) we mean the set \( \{ x \in \mathbb{R} | x > 0 \} \).
Since $y > 0$ and
\[
(u_1, u_2) \begin{pmatrix} y^2 & -xy \\ -xy & x^2 \end{pmatrix} (u_1, u_2)^T = (u_1 y - u_2 x)^2 \geq 0,
\]
the Hessian of $q$ is positive definite and the function is convex.

5 Appendix