Introduction to Convex Sets with Applications to Economics

Debasis Mishra*

March 21, 2011

1 Convex Sets

A set $C \subseteq \mathbb{R}^n$ is called **convex** if for all $x, y \in C$, we have $\lambda x + (1-\lambda)y \in C$ for all $\lambda \in [0, 1]$. The definition says that for any two points in set C, all points on the segment joining these two points must lie in C for C to be convex. Figure 1 shows two sets which are convex and two sets which are not convex.

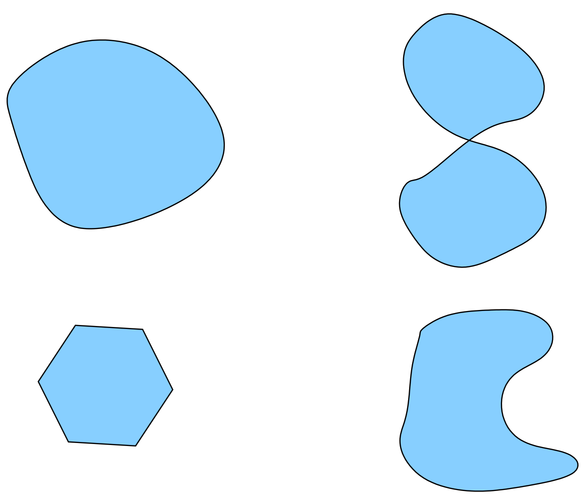

Figure 1: Sets on left are convex, but sets on right are not

Examples of convex sets:

- $C = \{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_1 + 2x_2 - x_3 = 4\}$. This is the equation of a plane in $\mathbb{R}^3$. In general, a **hyperplane** is defined as $C = \{x \in \mathbb{R}^n : p \cdot x = \alpha\}$, where $\alpha \in \mathbb{R}$ and

*Planning Unit, Indian Statistical Institute, 7 Shahid Jit Singh Marg, New Delhi 110016, India, E-mail: dmishra@isid.ac.in

$p \in \mathbb{R}^n$, called the normal to the hyperplane. Notation: As before $p \cdot x$ means dot product of p and x, i.e., $\sum_{i=1}^n p_i x_i$. For simplicity, we will sometimes write this as px. We will often denote a hyperplane as (p, α).

A hyperplane is a convex set. Take any $x, y \in C$. Then for any $\lambda \in [0,1]$ define $z = \lambda x + (1-\lambda)y$. Now, $pz = \lambda px + (1-\lambda)py = \lambda\alpha + (1-\lambda)\alpha = \alpha$. Hence $z \in C$.

- $C = \{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_1 + 2x_2 - x_3 \leq 4\}$. These are points on one side of the hyperplane. In general, a **half-space** is defined as $C = \{x \in \mathbb{R}^n : p \cdot x \leq \alpha\}$, where $\alpha \in \mathbb{R}$ and $p \in \mathbb{R}^n$. As in the case of a hyperplane, every half-space is a convex set.

- $C = \{(x_1, x_2) \in \mathbb{R}^2 : x_1^2 + x_2^2 \leq 4\}$. The set C is a circle of radius two with center $(0,0)$.

- $C = \{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_1 + x_2 - x_3 \leq 2, x_1 + 2x_2 - x_3 \leq 4\}$. Set C is the intersection of two half-spaces. In general, intersection of a finite number of half-spaces is called a **polyhedral set**, and is written as $C = \{x : Ax \leq b\}$, where $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$. Here, C is the intersection of m half-spaces. A polyhedral set is convex, because intersection of convex sets is a convex set, which we prove next.

LEMMA **1** *If C_1 and C_2 are two convex sets, then so is $C_1 \cap C_2$.*

Proof: Suppose $x, y \in C_1 \cap C_2$. Define $z = \lambda x + (1-\lambda)y$ for some $\lambda \in [0,1]$. Since $x, y \in C_1$ and C_1 is convex, $z \in C_1$. Similarly, $z \in C_2$. Hence, $z \in C_1 \cap C_2$. ■

Weighted averages of the form $\sum_{i=1}^k \lambda_i x^i$ with $\sum_{i=1}^k \lambda_i = 1$ and $\lambda_i \geq 0$ for all i is called **convex combination** of points $(x^1, \ldots, x^k)$.

DEFINITION **1** *The* **convex hull** *of a set $C \subseteq \mathbb{R}^n$, denoted as $H(C)$, is collection of all convex combinations of C, i.e., $x \in H(C)$ if and only if x can be represented as $x = \sum_{i=1}^k \lambda_i x^i$ with $\sum_{i=1}^k \lambda_i = 1$ and $\lambda_i \geq 0$ for all i and $x^1, \ldots, x^k \in C$ for some integer k.*

Figure 2 shows some sets and their convex hulls.

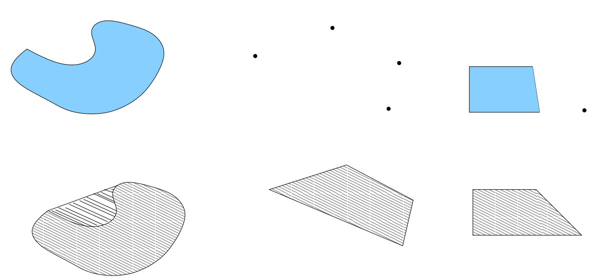

Figure 2: Convex Hulls

The convex hull of a set is a convex set. Take $x, y \in H(C)$. Define $z = \lambda x + (1-\lambda)y$ for any $\lambda \in [0,1]$. This is a convex combination of x and y, which in turn is a convex combination of points in C. Hence, z can be written as convex combination of points in C. In fact, z is a special convex set.

Lemma **2** *Suppose $C \subseteq \mathbb{R}^n$. Then, $H(C)$ is the smallest convex set containing C, i.e., suppose S is a convex set such that $C \subseteq S$, then $H(C) \subseteq S$.*

Proof: Consider any convex set S such that $C \subseteq S$. We will show that every point in $H(C)$ which is a convex combination of k points in C belongs to S. We use induction on k.

The claim holds for $k = 1$ since $C \subseteq S$. Suppose claim holds for $k = m$. We show that the claim holds for $k = m+1$. Let $x^1, \ldots, x^{m+1} \in C$ and $x = \sum_{i=1}^{m+1} \lambda_i x^i$, with usual conditions for convex combination. So, $x \in H(C)$. Now, set $\mu = \sum_{i=1}^{m} \lambda_m$. By our assumption $\mu > 0$ and $\lambda_{m+1} = 1 - \mu$. So, $x = \mu\big[\sum_{i=1}^{m} \frac{\lambda_i}{\mu} x^i\big] + (1-\mu)x^{m+1}$. Now, $\sum_{i=1}^{m} \frac{\lambda_i}{\mu} x^i$ is convex combination of m points in C, and hence must belong to S by our induction hypothesis. Further, $x^{m+1} \in C \subseteq S$. Thus, x is a convex combination of two points in S. Since S is convex, x must belong to S. ■

Hence, the convex hull of a convex set is the same set.

2 Hyperplanes and Separations

In this section, we prove an important result in convexity. It deals with separating a point outside a convex set from the convex set itself by using a hyperplane.

Definition **2** *Let S_1 and S_2 be two non-empty sets in $\mathbb{R}^n$. A hyperplane $H = \{x : px = \alpha\}$ is said to* **separate** *S_1 and S_2 if $px \geq \alpha$ for each $x \in S_1$ and $px \leq \alpha$ for each $x \in S_2$.*

The hyperplane H is said to **strictly separate** *S_1 and S_2 if $px > \alpha$ for all $x \in S_1$ and $px < \alpha$ for all $x \in S_2$.*

The idea of separation is illustrated in Figure 3. Not every pair of sets can be separated. For example, if the interior of two sets intersect, then they cannot be separeted. Figure 4 shows two pairs of sets, one of which can be (strictly) separated but the other pair cannot. So, what kind of sets can be separated. There are various results, all some form of separation theorems, regarding this. We will study a particular result.

Theorem **1** *Let C be a closed convex set in $\mathbb{R}^n$ and $z \in \mathbb{R}^n$ be a point such that $z \notin C$. Then, there exists a hyperplane which strictly separates z and C.*

Before we prove the theorem, it is instructive to look at Figure 5, which gives a geometric interpretation of the proof. We take a point $z \notin C$. Take the closest point to z in C, say y.

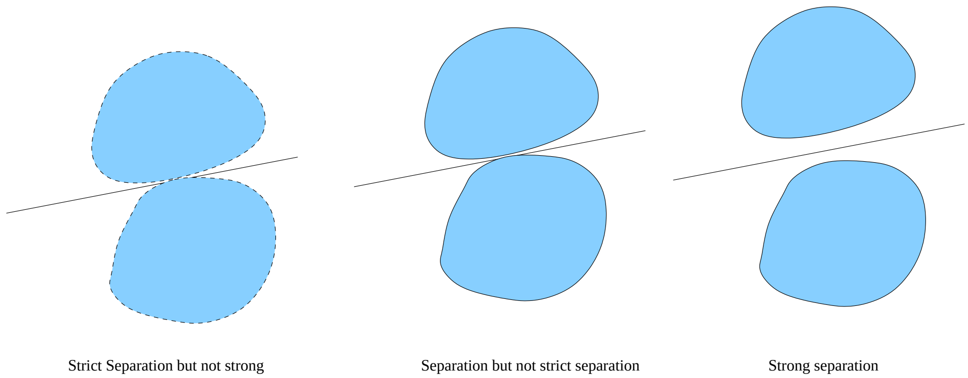

Figure 3: Different types of separation

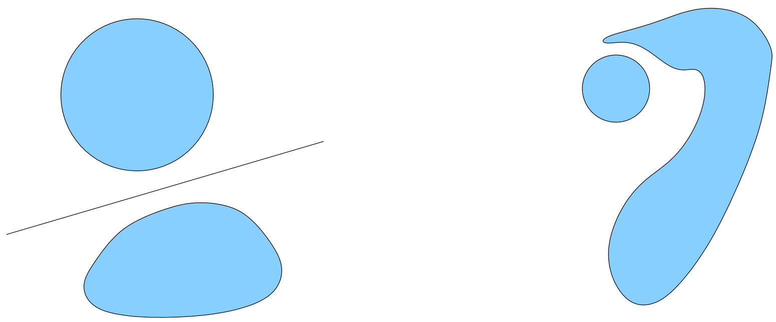

Figure 4: Possibility of separation

Then take the hyperplane passing through the mid-point between z and y and perpendicular to the vector $z - y$. The proof involves showing that this hyperplane separates z and C.

Proof: The proof is trivial if $C = \emptyset$. Suppose $C \neq \emptyset$. Then, consider $z \notin C$. Since $z \notin C$ and C is closed and convex, there exists a point $y \in C$ such that $\|z - y\|$ (i.e., the Euclidean distance between z and y) is minimized. (This claim is trivial geometrically - take any ball $B(z, r)$, which is centered at z and has a radius r, and grow it till it intersects C. Since $C \neq \emptyset$, it will intersect $B(z, r)$ for some r. Also, since C is closed and convex, $B(z, r) \cap C$ is closed and convex. Also, since $B(z, r)$ is bounded, $B(z, r) \cap C$ is compact. So, we need to minimize the continuous function $\|z - y\|$ over a compact set $B(z, r) \cap C$. By Weierstrass' Theorem, a minimum exists of such a function.)

Now set $p = z - y$ and $\alpha = \frac{1}{2}\left[\|z\|^2 - \|y\|^2\right]$. We show that $pz > \alpha$ and $px < \alpha$ for all $x \in C$.

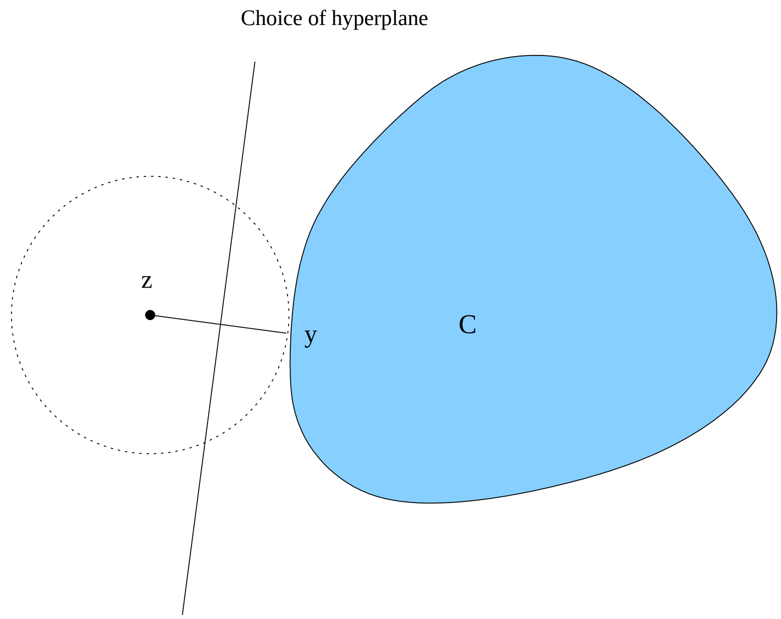

Figure 5: Geometric illustration of separating hyperplane theorem

Now,

$$\begin{aligned} pz - \alpha &= (z-y)z - \frac{1}{2}\left[\|z\|^2 - \|y\|^2\right] \\ &= \frac{1}{2}\left[\|(z-y)\|^2\right] \\ &> 0. \end{aligned}$$

Note that $py < py + \frac{1}{2}\|p\|^2 = (z-y)y + \frac{1}{2}\left[\|z-y\|^2\right] = \alpha$. Assume for contradiction, there exists $x \in C$ such that $px \geq \alpha > py$. Hence, $p(x-y) > 0$. Define

$$\delta = \frac{2p(x-y)}{\|(x-y)\|^2} > 0. \tag{1}$$

Now, choose $1 \geq \lambda > 0$ and $\lambda < \delta$. Such a λ exists because of inequality 1. Define $w = \lambda x + (1-\lambda)y$. Since C is convex, w belongs to C. Now,

$$\begin{aligned} \|(z-w)\|^2 &= \|(z-y) + \lambda(y-x)\|^2 \\ &= \|p - \lambda(x-y)\|^2 \\ &= \|p\|^2 - 2\lambda(x-y)p + \lambda^2\|(x-y)\|^2 \\ &< \|p\|^2 \\ &= \|(z-y)\|^2. \end{aligned}$$

Hence, w nearer to z than y. This is a contradiction. Hence, for all $x \in C$ we should have $px < \alpha$. ∎

There are other generalizations of this theorem, which we will not cover here. For example, if you drop the requirement that the convex set be closed, then we will get weak separation. The separation theorems have a wide variety of applications.

3 Farkas Lemma

We will now state the most important result of this section. The result is called **Farkas Lemma** or **Theorem of Alternatives**.

Suppose $A \in \mathbb{R}^{m \times n}$. Let $(a_1, \ldots, a_n)$ be the columns of A. The set of all non-negative linear combinations of columns of A is called the **cone** of A. Formally,

$$cone(A) = \{b \in \mathbb{R}^m : Ax = b \text{ for some } x \in \mathbb{R}^n_+\}$$

As an example, consider the following 2×3 matrix,

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

If we take $x = (2, 3, 1)$, then we get the following $b \in cone(A)$.

$$b = \begin{bmatrix} 4 + 0 + 1 = 5 \\ 2 + (-6) + (-2) = -6 \end{bmatrix}$$

The $cone(A)$ is depicted in Figure 6. As can be seen from the figure, $cone(A)$ is a convex and closed set.

Lemma **3** *Suppose $A \in \mathbb{R}^{m \times n}$. Then $cone(A)$ is a convex set.*

Proof: Let $b^1, b^2 \in cone(A)$. Define $b = \lambda b^1 + (1 - \lambda) b^2$ for some $\lambda \in (0, 1)$. Let $b^1 = Ax^1$ and $b^2 = Ax^2$ for $x^1, x^2 \in \mathbb{R}^n_+$. Then $b = \lambda A x^1 + (1 - \lambda) A x^2 = A\big[\lambda x^1 + (1 - \lambda) x^2\big]$. Since $\mathbb{R}^n_+$ is convex $x = \lambda x^1 + (1 - \lambda) x^2 \in \mathbb{R}^n_+$. So, $b \in cone(A)$, and hence, $cone(A)$ is convex. ■

We state the following result without a proof.

Lemma **4** *Suppose $A \in \mathbb{R}^{m \times n}$. Then $cone(A)$ is a closed set.*

Theorem **2** *Let $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$. Suppose $F = \{x \in \mathbb{R}^n_+ : Ax = b\}$ and $G = \{y \in \mathbb{R}^m : yA \geq 0, yb < 0\}$. Then, $F \neq \emptyset$ if and only if $G = \emptyset$.*

Before we describe the proof, a word about the notation used in Theorem 2. The inequality $yA \geq 0$ is a system of n inequalities. It means that $ya^j \geq 0$ for every column vector a^j of A. The multiplication yA is a matrix multiplication, where appropriate transpose needs to be taken. We have abused notation to write yA to mean this matrix multiplication.

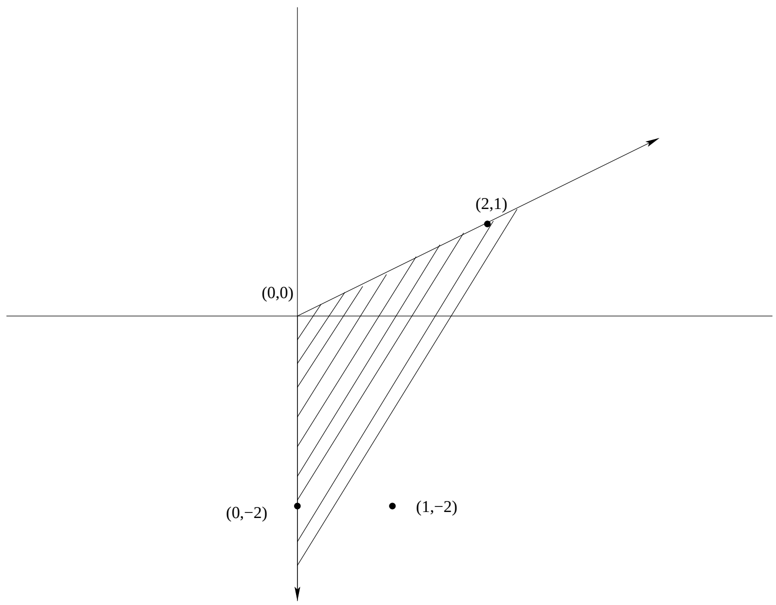

Figure 6: Illustration of cone in $\mathbb{R}^2$

The system of inequalities in G is called the **Farkas alternative**. Let us apply Farkas Lemma to some examples. Does the following system have a non-negative solution?

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

The Farkas alternative for this system is:

$$\begin{aligned} y_1 + y_2 &< 0 \\ 4y_1 + y_2 &\geq 0 \\ y_1 &\geq 0 \\ -5y_1 + 2y_2 &\geq 0. \end{aligned}$$

The last three inequalities can be written in matrix form as follows.

$$\begin{bmatrix} 4 & 1 \\ 1 & 0 \\ -5 & 2 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \geq \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Since $y_1 \geq 0$, from the first inequality, we get $y_2 < 0$. This contradicts the last inequality. Hence, Farkas alternative have no solution, implying that the original system of equations have a solution.

Proof: Suppose $F \neq \emptyset$. Let $x \in F$. Choose $y \in \mathbb{R}^m$ such that $yA \geq 0$. Then, $yb = y(Ax) = (yA)x \geq 0$. Hence, $G = \emptyset$.

Suppose $G = \emptyset$, and assume for contradiction $F = \emptyset$. This means $\{x \in \mathbb{R}^n_+ : Ax = b\} = \emptyset$, i.e., $b \notin cone(A)$. Since $cone(A)$ is closed and convex, we can separate it from b by a hyperplane (y, α) such that $yb < \alpha$ and $yz > \alpha$ for all $z \in cone(A)$. Notice that $0 \in cone(A)$. Hence, $\alpha < 0$. So, we get $yb < 0$.

Let a^j be a column vector of A. We will show that $ya^j \geq 0$. Consider any $\lambda > 0$. By definition $\lambda a^j \in cone(A)$. Hence, $y(\lambda a^j) > \alpha$. Assume for contradiction $ya^j < 0$. Since $\lambda > 0$ can be chosen arbitrarily large, we can choose it such that $y(\lambda a^j) < \alpha$. This is a contradiction. Hence, $ya^j \geq 0$. Thus, we get $yA \geq 0$ and $yb < \alpha < 0$. Hence $G \neq \emptyset$. ■

Does the following system have a non-negative solution?

$$\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \\ 2 \\ 1 \end{bmatrix}$$

The farkas alternative is:

$$\begin{bmatrix} 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} \geq \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$2y_1 + 2y_2 + 2y_3 + y_4 < 0.$$

One possible solution to the Farkas alternative is $y_1 = y_2 = y_3 = \frac{-1}{2}$ and $y_4 = 1$. Hence, the original system has no non-negative solution. Intuitively, it says if we multiply $\frac{-1}{2}$ to the first three equations in the original system and multiply 1 to the last equation, and add all of them up, we will get a contradiction: $0 = -2$.

Often we come across systems of equations and inequalities, with variables that are **free** (no non-negative constraints) and variables that are constrained to be non-negative. Farkas Lemma can be easily generalized to such systems.

THEOREM **3** *Let $A \in \mathbb{R}^{m \times n}$, $B \in \mathbb{R}^{m \times t}$, $b \in \mathbb{R}^m$, $C \in \mathbb{R}^{k \times n}$, $D \in \mathbb{R}^{k \times t}$, and $d \in \mathbb{R}^k$. Suppose $F = \{x \in \mathbb{R}^n_+, x' \in \mathbb{R}^t : Ax + Bx' = b, Cx + Dx' \leq d\}$ and $G = \{y \in \mathbb{R}^m, y' \in \mathbb{R}^k_+ : yA + y'C \geq 0, yB + y'D = 0, yb + y'd < 0\}$. Then, $F \neq \emptyset$ if and only if $G = \emptyset$.*

Proof: Consider the system of inequalities $Cx + Dx' \leq d$. This can be converted to a system of equations by introducing **slack** variables for every inequality. In particular, consider variables $s \in \mathbb{R}^k_+$ such that $Cx + Dx' + s = d$.

Now, the vectors $x' \in \mathbb{R}^t$ can be written as $x' = x^+ - x^-$, where $x^+, x^- \in \mathbb{R}^t_+$. This is because every real number can be written as difference of two non-negative real numbers. So, the set F can be rewritten as $F = \{x \in \mathbb{R}^n_+, x^+, x^- \in \mathbb{R}^t_+, s \in \mathbb{R}^k_+ : Ax + Bx^+ - Bx^- + 0.s = b, Cx + Dx^+ - Dx^- + Is = d\}$, where I is the identity matrix. In matrix form, this looks as follows:

$$\begin{bmatrix} A & B & -B & 0 \\ C & D & -D & I \end{bmatrix} \begin{bmatrix} x \\ x^+ \\ x^- \\ s \end{bmatrix} = \begin{bmatrix} b \\ d \end{bmatrix}$$

The Farkas alternative for this system of equations has two sets of variables $y \in \mathbb{R}^m$ and $y' \in \mathbb{R}^k$. One inequality is $by + dy' < 0$. The other inequalities are $Ay + Cy \geq 0$, $By + Dy' \geq 0$, $-By - Dy' \geq 0$, and $y' \geq 0$. This simplifies to $Ay + Cy' \geq 0$ and $By + Dy' = 0$ with $y' \in \mathbb{R}^k_+$. From Farkas Lemma (Theorem 2), the result now follows. ■

Here is an example of how to write Farkas alternative for general system of constraints. Consider the following set of constraints.

$$\begin{aligned} x_1 - 3x_2 + x_3 &\leq -3 \\ x_1 + x_2 - x_3 &\geq 2 \\ x_1 + 2x_2 + 3x_3 &= 5 \\ x_1 - x_2 &= 2 \\ x_1, x_2 &\geq 0 \end{aligned}$$

Here, x_3 is a *free* variable (without the non-negativity constraint). The first step is to conver the set of constraints into the form in Theorem 3. For this, we need to convert the second constraint into $\leq$ form.

$$\begin{aligned} x_1 - 3x_2 + x_3 &\leq -3 \\ -x_1 - x_2 + x_3 &\leq -2 \\ x_1 + 2x_2 + 3x_3 &= 5 \\ x_1 - x_2 &= 2 \\ x_1, x_2 &\geq 0 \end{aligned}$$

For writing the Farkas alternative, we first associate a variable with every constraint: $y = (y_1, y_2, y_3, y_4)$ for four constraints. Out of this, first and second constraints are inequalities, so corresponding variables (y_1, y_2) are non-negative, while variables (y_3, y_4) corresponding to equations are free.

Now, the strict inequality in the Farkas alternative is easy to write: $-3y_1-2y_2+5y_3+2y_4 < 0$. There will be four constraints in the Farkas alternative, each corresponding to the variables in the original system. The constraints corresponding to non-negative variables are weak inequalities, while the constraints corresponding to free variables are equations. For example, the inequality corresponding to variable x_1 is $y_1 - y_2 + y_3 + y_4 \geq 0$. Similarly, the inequality corresponding to x_2 is $-3y_1 - y_2 + 2y_3 - y_4 \geq 0$. Since the variable x_3 is free, the constraint corresponding to it is an equation: $y_1 + y_2 + 3y_3 = 0$. Hence, the Farkas alternative is:

$$\begin{aligned} -3y_1 - 2y - 2 + 5y_3 + 2y_4 &< 0 \\ y_1 - y_2 + y_3 + y_4 &\geq 0 \\ -3y_1 - y_2 + 2y_3 - y_4 &\geq 0 \\ y_1 + y_2 + 3y_3 &= 0 \\ y_1, y_2 &\geq 0. \end{aligned}$$

4 Application: Core of Cooperative Games

A cooperative game (with transferrable utility) is defined by a set of n players N and a value function $v : 2^N \rightarrow \mathbb{R}$ which represents the value or worth of a coalition (subset) of players. For every coalition $S \subseteq N$, a value $v(S)$ is attached. The exact method of finding the value function depends on the problems. The tuple (N, v) defines a cooperative game. We had already defined cooperative games using cost function. Here is an example with value function.

Sale of an item. Consider the sharing of an item between two buyers and a seller (who owns the item). The set of players can be denoted as $N = \{1, 2, s\}$, where s is the seller. The valuation of the item to buyers (i.e., the utility a buyer gets by getting the item) are: 5 and 10. The seller has no value for the item. The cooperative game can be defined as follows: $v(\emptyset) = v(\{s\}) = 0$; $v(\{1\}) = v(\{2\}) = v(\{1, 2\}) = 0$; $v(\{1, s\}) = 5, v(\{2, s\}) = v(\{1, 2, s\}) = 10$ (by assigning the item to the highest valued buyer).

The definition of a cooperative game says nothing about how the value of a game should be divided between players. The cooperative game literature deals with such issues in details.

A vector $x \in \mathbb{R}^n$ is called an **imputation** if $\sum_{i \in N} x_i = v(N)$ and $x_i \geq v(\{i\})$. One can think of an imputation as a division of $v(N)$ that gives every player at least as much as he will get himself. When we generalize this to every coalition of agents, we get the notion of *core*.

DEFINITION **3** *The* **core** *of a game* (N, v) *is the set* $C(N, v) = \Big\{x \in \mathbb{R}^n : \sum_{i \in N} x_i = v(N),\ \sum_{i \in S} x_i \geq v(S)\ \forall\ S \subsetneq N\Big\}$.

The core constraints are stability condition. It stipulates that every coalition of agents must get at least as they will get if they form their own coalition. Otherwise, such a coalition may break from the grand coalition.

In the example above, the set of inequalities for core are

$$\begin{aligned} x_1 + x_2 + x_s &= 10 \\ x_1 + x_2 &\geq 0 \\ x_1 + x_s &\geq 5 \\ x_2 + x_s &\geq 10 \\ x_i &\geq 0 \; \forall \; i \in \{1, 2, s\} \end{aligned}$$

Simple substitutions give, $x_2 \leq 5$ and $x_1 \leq 0$. $x_1 \geq 0$ gives $x_1 = 0$, and thus $x_s \geq 5$. So, $x_1 = 0, 0 \leq x_2 \leq 5$, and $5 \leq x_s \leq 10$ with $x_2 + x_s = 10$ constitutes the core of this game.

But not all games have a core. For example, consider a game with two agents $\{1, 2\}$ with $v(\{1\}) = 1 = v(\{2\})$ but $v(\{1, 2\}) = 3$. This game has an empty core since no (x_1, x_2) can satisfy $x_1 \leq 1, x_2 \leq 1$, and $x_1 + x_2 = 3$.

A necessary and sufficient condition can be found by using Farkas Lemma.

Let $B(N)$ be the set of feasible solutions to the following system:

$$\begin{aligned} \sum_{S \subsetneq N : i \in S} y_S &= 1, \qquad \forall \; i \in N, \\ y_S &\geq 0, \qquad \forall \; S \subsetneq N. \end{aligned}$$

y_S can be thought as the weight given to coalition S. It is easy to verify that $B(N) \neq \emptyset$. For example, by setting $y_S = 1$ for all S with $|S| = 1$ and setting $y_S = 0$ otherwise gives a feasible $y \in B(N)$.

Theorem 4 **(Bondareva-Shapley)** *$C(N, v) \neq \emptyset$ if and only if*

$$v(N) \geq \sum_{S \subsetneq N} v(S) y_S, \qquad \forall \; y \in B(N).$$

(If a game satisfies this condition then it is called a **balanced** *game, i.e., the core of a game is non-empty if and only if it is balanced).*

Proof: Consider the following system of constraints corresponding to core of a game:

$$\begin{aligned} &\sum_{i \in N} x_i = v(N) && \textbf{(CORE)} \\ &\sum_{i \in S} x_i \geq v(S) \qquad \forall \; S \subsetneq N \\ &x_i \text{ free} \qquad \forall \; i \in N. \end{aligned}$$

The Farkas alternative for (**CORE**) is

$$v(N)y_N - \sum_{S \subsetneq N} v(S)y_S < 0 \qquad \textbf{(BAL)}$$
$$y_N - \sum_{S \subsetneq N : i \in S} y_S = 0 \qquad \forall\, i \in N$$
$$y_S \geq 0 \qquad \forall\, S \subsetneq N$$
$$y_N \text{ free.}$$

Now, suppose $C(N,v) \neq \emptyset$. Then (**CORE**) has a solution, and Farkas alternative (**BAL**) has no solution. Now, consider any $y \in B(N)$. For any $y \in B(N)$, we let $y_N = 1$ and the final two constraints of (**BAL**) are satisfied. Since (**BAL**) has no solution, we must have $v(N) - \sum_{S \subsetneq N} v(S)y_S \geq 0$. This implies that the game is balanced.

Now, suppose the game is balanced. Then for all $y \in B(N)$, we have $v(N) \geq \sum_{S \subsetneq N} v(S)y_S$. We will show that (**BAL**) has no solution, and hence, (**CORE**) has a solution and $C(N,v) \neq \emptyset$. Assume for contradiction, (**BAL**) has a solution y. Clearly, $y_N \neq 0$ (else every $y_S = 0$, and this will contradict the first inequality in (**BAL**)). Now, define, $y'_S = \frac{y_S}{y_N}$ for all $S \subseteq N$. Since y is a solution to (**BAL**), we get

$$v(N) < \sum_{S \subsetneq N} v(S)y'_S$$
$$\sum_{S \subsetneq N : i \in S} y'_S = 1 \qquad \forall\, i \in N$$
$$y'_S \geq 0 \qquad \forall\, S \subsetneq N.$$

Hence, $y' \in B(N)$. But the first inequality contradicts the fact that the game is balanced. ■

Let us verify that the game in our earlier example (of sale of an item) is balanced. Notice that in that game the only coalitions, besides the grand coalition, having positive values are $\{1, s\}$ and $\{2, s\}$. So, we need to show for every $y \in B(N)$, we have

$$v(\{1,2,s\}) = 10 \geq v(\{1,s\})y_{\{1,s\}} + v(\{2,s\})y_{\{2,s\}} = 5y_{\{1,s\}} + 10y_{\{2,s\}}.$$

But $y \in B(N)$ implies that $y_{\{1,s\}} + y_{\{2,s\}} \leq 1$. Since $y_{\{2,s\}} \leq 1$, we get the desired balanced game.

Indeed, it is easy to state a result for general cooperative "market game". A market game is defined by a seller s and a set of buyers B. So, the set of players is $N = B \cup \{s\}$. The key feature of the market game is the special type of value function. In particular, for every $S \subseteq N$ we define $v(S)$ to be the value of the market with player set (coalition) S. The restriction we put is $v(S)=0$ if $s \notin S$. Now, call a market game **monotonic** if $v(S) \le v(T)$ for all $S \subseteq T \subseteq N$.

THEOREM **5** *Every monotonic market game is balanced.*

Proof: Pick any $y \in B(N)$. Now,

$$\begin{aligned}\sum_{S \subsetneq N} v(S)y_S &= \sum_{S \subsetneq N: s \in S} v(S)y_S \\ &\le v(N) \sum_{S \subsetneq N: s \in S} y_S,\end{aligned}$$

where the inequality uses the fact that market game is monotonic. But since $y \in B(N)$, we get that $\sum_{S \subsetneq N: s \in S} y_S = 1$. This shows that $\sum_{S \subsetneq N} v(S)y_S \le v(N)$. So, the monotonic market game is balanced. ■

Indeed, a trivial element in the core of a monotonic market game is $x_s = v(N)$ and $x_i = 0$ if $i \neq s$.

Note on cost games. If the cooperative game is defined by a cost function c instead of a value function v, then the core constraints change in the following manner:

$$\begin{aligned}&\sum_{i \in N} x_i = c(N) \\ &\sum_{i \in S} x_i \le c(S) \qquad \forall\ S \subsetneq N \\ &x_i \text{ free} \qquad \forall\ i \in N.\end{aligned}$$

It is easy to verify that the corresponding Farkas alternative gives the following balancedness condition.

$$\sum_{S \subsetneq N} c(S)y_S \ge c(N) \qquad \forall\ y \in B(N).$$

5 Carathéodory Theorem

This section is devoted to an improtant result in the theory of convex sets. The result says that if we choose any point in the convex hull of an arbitrary set $S \subseteq \mathbb{R}^n$, it can be expressed as convex combination of at most $n+1$ points in S. To prove this result, we start with reviewing some basic definitions in linear algebra.

For any set of points $x_1, \ldots, x_k \in \mathbb{R}^n$ and $\lambda_1, \ldots, \lambda_k \in \mathbb{R}$ the point $\sum_{i=1}^{k} \lambda_i x_i$ is called

- a **linear combination** of $x_1, \ldots, x_k$,

- an **affine combination** of $x_1, \ldots, x_k$ if $\sum_{i=1}^{k} \lambda_i = 1$,
- a **convex combination** of $x_1, \ldots, x_k$ if $\sum_{i=1}^{k} \lambda_i = 1$ and $\lambda_1, \ldots, \lambda_k \geq 0$.

Note that if $x \in \mathbb{R}^n$ is a convex combination of points $x_1, \ldots, x_k \in \mathbb{R}^n$ then it is also an affine combination and linear combination of these points. Similarly, if $x \in \mathbb{R}^n$ is an affine combination of $x_1, \ldots, x_k \in \mathbb{R}^n$ then it is also a linear combination of these points.

A set of points $x_1, \ldots, x_k \in \mathbb{R}^n$ are **linearly independent** if none of them can be expressed as a linear combination of others. In other words, if $x_1, \ldots, x_k$ are linearly independent then $\sum_{i=1}^{k} \lambda_i x_i = 0$ implies that $\lambda_i = 0$ for all $i \in \{1, \ldots, k\}$. If $x_1, \ldots, x_k$ are not linearly independent then they called **linearly dependent**.

Here are some examples of linearly independent vectors.

$$\begin{aligned} (4, -90) \\ (1, 0, 5), (2, 5, 2) \\ (1, 4), (3, -3). \end{aligned}$$

Here are some examples of linearly dependent vectors.

$$\begin{aligned} (1, -2), (-2, 4) \\ (0, 1, 0), (2, -3, 5), (2, -2, 5). \end{aligned}$$

Similarly, we can define the notion of affine independence. A set of points $x_1, \ldots, x_k \in \mathbb{R}^n$ are **affinely independent** if none of them can be expressed as affine combination of others. A set of points $x_1, \ldots, x_k \in \mathbb{R}^n$ are **affinely dependent** if they are not affinely independent. The alternate definition of affine independence is here.

LEMMA **5** *A set of point $x_1, \ldots, x_k \in \mathbb{R}^n$ are affinely independent if and only if $x_2 - x_1, x_3 - x_1, \ldots, x_k - x_1$ are linearly independent.*

Proof: Suppose $x_1, \ldots, x_k$ are affinely independent. Assume for contradiction $x_2 - x_1, x_3 - x_1, \ldots, x_k - x_1$ are linearly dependent. Then, there exists $\lambda_2, \ldots, \lambda_k$, not all of them zero, and $x_j - x_1$ such that such that

$$\sum_{i=2, i \neq j}^{k} \lambda_i (x_i - x_1) = x_j - x_1.$$

This implies that

$$\sum_{i=2, i \neq j}^{k} \lambda_i x_i + [1 - \sum_{i=1, i \neq j}^{k} \lambda_i] x_1 = x_j.$$

This shows that $x_1, \ldots, x_k$ are affinely dependent. This is a contradiction.

Now, suppose $x_2 - x_1, x_3 - x_1, \ldots, x_k - x_1$ are linearly independent. Assume for contradiction $x_1, \ldots, x_k$ are affinely dependent. Then, some point, say x_j, can be expressed as affine combination of others. Hence, for some λs with $\sum_{i=1,i\neq j}^{k} \lambda_i = 1$

$$x_j = \sum_{i=1,i\neq j}^{k} \lambda_i x_i$$
$$\Rightarrow x_j - x_1 = \sum_{i=1,i\neq j}^{k} \lambda_i x_i - \sum_{i=1,i\neq j}^{k} \lambda_i x_1 = \sum_{i=2,i\neq j}^{k} \lambda_i [x_i - x_1].$$

This is a contradiction to the fact that $x_2 - x_1, \ldots, x_k - x_1$ are linearly independent. ■

This shows that if $x_1, \ldots, x_k$ are affinely dependent then $x_2 - x_1, \ldots, x_k - x_1$ must be linearly dependent. The maximum number of linearly independent points in $\mathbb{R}^n$ is n. We state this as a lemma.

LEMMA **6** *The maximum number of linearly independent points in $\mathbb{R}^n$ is n.*

Proof: The proof is left as an exercise. It has to do with solving a system of n equations with n variables. ■

The idea in the previous lemma is extended to any arbitrary set to define the dimension of a set. Consider any set S. The maximum number of linearly independent points in S is called the **dimension** of S. In that sense, the dimension of $\mathbb{R}^n$ is n.

THEOREM **6 (Carathéodory Theorem)** *Let $S \subseteq \mathbb{R}^n$ be an arbitrary set. If $x \in H(S)$, then there exists $x_1, \ldots, x_{n+1} \in S$ such that $x \in H(\{x_1, \ldots, x_{n+1}\})$.*

In words, any point in the convex hull of a set in $\mathbb{R}^n$ can be written as the convex combination of at most $n + 1$ points in that set.
Proof: Pick $x \in H(S)$. Let $x = \sum_{i=1}^{k} \lambda_i x_i$ with $\lambda_i > 0$ for all $i \in \{1, \ldots, k\}$ and $\sum_{i=1}^{k} \lambda_i = 1$. If $k \leq n + 1$, then we are done. Suppose $k > n + 1$. Then, $x_2 - x_1, \ldots, x_k - x_1$ are more than n points in $\mathbb{R}^n$. Hence, they are linearly dependent. Hence, there exists $\mu_2, \ldots, \mu_k$, not all of them equal to zero, such that $\sum_{i=2}^{k} \mu_i(x_i - x_1) = 0$. Let $\mu_1 = -\sum_{i=2}^{k} \mu_i$. Thus, $\sum_{i=1}^{k} \mu_i x_i = 0$ with $\sum_{i=1}^{k} \mu_i = 0$ and at least one μ_i positive. Thus for any $\alpha > 0$ we have

$$x = \sum_{i=1}^{k} \lambda_i x_i = \sum_{i=1}^{k} \lambda_i x_i - \alpha \sum_{i=1}^{k} \mu_i x_i = \sum_{i=1}^{k} (\lambda_i - \alpha\mu_i) x_i.$$

Choose α as follows:

$$\alpha = \min_{1 \le i \le k} \{\frac{\lambda_i}{\mu_i} : \mu_i > 0\} = \frac{\lambda_j}{\mu_j}.$$

Note that $\alpha > 0$ since $\lambda_j > 0$. Further, for any $i \in \{1, \ldots, k\}$, $\lambda_i - \alpha\mu_i > 0$ if $\mu_i < 0$ and if $\mu_i > 0$, then $\lambda_i - \alpha\mu_i = \lambda_i - \lambda_j \frac{\mu_i}{\mu_j} \ge 0$. Also, note that $\sum_{i=1}^{k}(\lambda_i - \alpha\mu_i) = 1$. Hence, x is a convex combination of $x_1, \ldots, x_k$ but with $\lambda_j - \alpha\mu_j = 0$. Hence, x can be expressed as convex combination of $k-1$ points in S. The process can be repeated till we have $n+1$ points. ■

A consequence of Carathéodory Theorem is the following result. We give an outline of the proof below.

Corollary 1 *The convex hull of a compact set is a compact set.*

Proof: Let S be a compact set. Since S is bounded, $H(S)$ is also bounded. Now, take a sequence $\{x^k\}$ in $H(S)$. By Carathéodory theorem, each x^k can be represented as convex combination of $n+1$ points in S. This gives a sequence in αs and xs, where α belongs to a bounded set and $x \in S$ belongs to a bounded set. Hence, both of them must have limit points. Since S is closed, the result follows. ■