ISS0031 Modeling and Identification

Juri Belikov

Department of Computer Control,
Tallinn University of Technology

jbelikov@cc.ic.ee

September 19, 2014

Let $S \neq \emptyset, S \subset \mathbb{R}^n$ and $x_1, x_2 \in S$.

Definition

The set $[x_1, x_2] = \{x | x = \lambda x_1 + (1-\lambda)x_2, 0 \leq \lambda \leq 1\}$ is called a **line segment** with the endpoints x_1, x_2.

Example: Let $x_1(2,1)$ and $x_2(4,3)$. Next, using the formula from the above definition, we get

$$x_1 = 2 \cdot \lambda + (1-\lambda) \cdot 4$$
$$x_2 = 1 \cdot \lambda + (1-\lambda) \cdot 3$$
$$0 \leq \lambda \leq 1.$$

Let $S \neq \emptyset, S \subset \mathbb{R}^n$ and $x_1, x_2 \in S$.

Definition

The set $[x_1, x_2] = \{x | x = \lambda x_1 + (1-\lambda)x_2, 0 \leq \lambda \leq 1\}$ is called a **line segment** with the endpoints x_1, x_2.

Example: Let $x_1(2, 1)$ and $x_2(4, 3)$. Next, using the formula from the above definition, we get

$$x_1 = 2 \cdot \lambda + (1-\lambda) \cdot 4$$
$$x_2 = 1 \cdot \lambda + (1-\lambda) \cdot 3$$
$$0 \leq \lambda \leq 1.$$

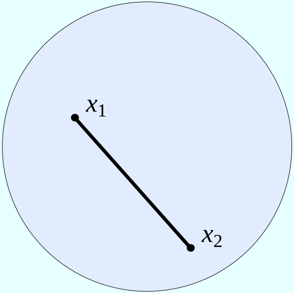
Convex Set

Definition

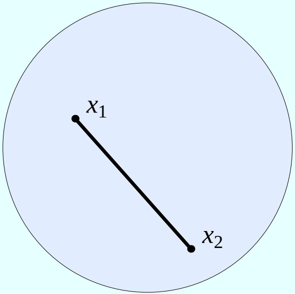
A set S in a vector space over $\mathbb{R}$ is called a **convex set** if the line segment joining any pair of points $x_1, x_2 \in S$ lies entirely in S.

Convex Set

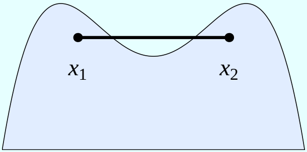
Illustrative examples I



(a) convex



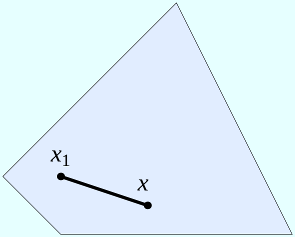
(c) convex



(d) non-convex

Convex Set

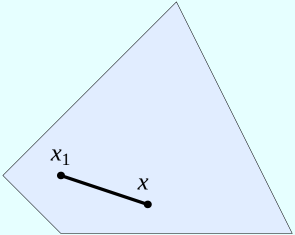
Illustrative examples II



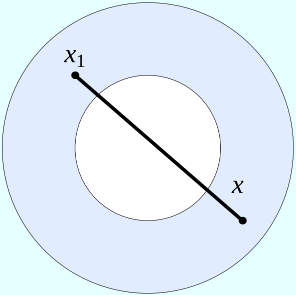
(e) convex

Convex Set

Illustrative examples II



(g) convex



(h) non-convex

Proposition

A solution set $\mathbb{L}$ for the linear inequality $a_1x_1 + a_2x_2 + \cdots + a_nx_n \leq b$ is a convex set.

Proof: Let the points $C_1(c_1^1, c_2^1, \ldots, c_n^1)$ and $C_2(c_1^2, c_2^2, \ldots, c_n^2)$ be solutions of the given inequality. Then,

$$a_1c_1^1 + a_2c_2^1 + \cdots + a_nc_n^1 \leq b$$
$$a_1c_1^2 + a_2c_2^2 + \cdots + a_nc_n^2 \leq b$$

Next, we multiply the first inequality by λ, the second inequality by $1-\lambda$ and add results

$$a_1(\lambda c_1^1 + (1-\lambda)c_1^2) + a_2(\lambda c_2^1 + (1-\lambda)c_2^2) + \cdots + a_n(\lambda c_n^1 + (1-\lambda)c_n^2) \leq \lambda b + (1-\lambda)b = b.$$

Using the obtained result, we can conclude that the point

$$C(\lambda c_1^1 + (1-\lambda)c_1^2, \lambda c_2^1 + (1-\lambda)c_2^2, \ldots, \lambda c_n^1 + (1-\lambda)c_n^2) = \lambda C_1 + (1-\lambda)C_2 \in \mathbb{L}. \quad \blacksquare$$

Proposition

A solution set $\mathbb{L}$ for the linear inequality $a_1x_1 + a_2x_2 + \cdots + a_nx_n \leq b$ is a convex set.

Proof: Let the points $C_1(c_1^1, c_2^1, \ldots, c_n^1)$ and $C_2(c_1^2, c_2^2, \ldots, c_n^2)$ be solutions of the given inequality. Then,

$$a_1c_1^1 + a_2c_2^1 + \cdots + a_nc_n^1 \leq b$$
$$a_1c_1^2 + a_2c_2^2 + \cdots + a_nc_n^2 \leq b$$

Next, we multiply the first inequality by λ, the second inequality by $1-\lambda$ and add results

$$a_1(\lambda c_1^1 + (1-\lambda)c_1^2) + a_2(\lambda c_2^1 + (1-\lambda)c_2^2) + \cdots + a_n(\lambda c_n^1 + (1-\lambda)c_n^2) \leq \lambda b + (1-\lambda)b = b.$$

Using the obtained result, we can conclude that the point

$$C(\lambda c_1^1 + (1-\lambda)c_1^2, \lambda c_2^1 + (1-\lambda)c_2^2, \ldots, \lambda c_n^1 + (1-\lambda)c_n^2) = \lambda C_1 + (1-\lambda)C_2 \in \mathbb{L}. \qquad \blacksquare$$

Proposition

A solution set $\mathbb{L}$ for the linear inequality $a_1x_1 + a_2x_2 + \cdots + a_nx_n \leq b$ is a convex set.

Proof: Let the points $C_1(c_1^1, c_2^1, \ldots, c_n^1)$ and $C_2(c_1^2, c_2^2, \ldots, c_n^2)$ be solutions of the given inequality. Then,

$$a_1c_1^1 + a_2c_2^1 + \cdots + a_nc_n^1 \leq b$$
$$a_1c_1^2 + a_2c_2^2 + \cdots + a_nc_n^2 \leq b$$

Next, we multiply the first inequality by λ, the second inequality by $1-\lambda$ and add results

$$a_1(\lambda c_1^1 + (1-\lambda)c_1^2) + a_2(\lambda c_2^1 + (1-\lambda)c_2^2) + \cdots + a_n(\lambda c_n^1 + (1-\lambda)c_n^2) \leq \lambda b + (1-\lambda)b = b.$$

Using the obtained result, we can conclude that the point

$$C(\lambda c_1^1 + (1-\lambda)c_1^2, \lambda c_2^1 + (1-\lambda)c_2^2, \ldots, \lambda c_n^1 + (1-\lambda)c_n^2) = \lambda C_1 + (1-\lambda)C_2 \in \mathbb{L}. \qquad \blacksquare$$

Proposition

A solution set $\mathbb{L}$ for the linear inequality $a_1x_1 + a_2x_2 + \cdots + a_nx_n \leq b$ is a convex set.

Proof: Let the points $C_1(c_1^1, c_2^1, \dots, c_n^1)$ and $C_2(c_1^2, c_2^2, \dots, c_n^2)$ be solutions of the given inequality. Then,

$$a_1c_1^1 + a_2c_2^1 + \cdots + a_nc_n^1 \leq b$$
$$a_1c_1^2 + a_2c_2^2 + \cdots + a_nc_n^2 \leq b$$

Next, we multiply the first inequality by λ, the second inequality by $1 - \lambda$ and add results

$$a_1(\lambda c_1^1 + (1-\lambda)c_1^2) + a_2(\lambda c_2^1 + (1-\lambda)c_2^2) + \cdots + a_n(\lambda c_n^1 + (1-\lambda)c_n^2) \leq \lambda b + (1-\lambda)b = b.$$

Using the obtained result, we can conclude that the point

$$C(\lambda c_1^1 + (1-\lambda)c_1^2, \lambda c_2^1 + (1-\lambda)c_2^2, \dots, \lambda c_n^1 + (1-\lambda)c_n^2) = \lambda C_1 + (1-\lambda)C_2 \in \mathbb{L}. \qquad \blacksquare$$

Proposition

The intersection of any finite number of convex sets is a convex set.

Proof: Suppose $S_1, S_2, \ldots, S_n$ are convex sets. Then their intersection $\bigcap_{i=1}^{n} S_i = \{x : x \in S_i, \forall i = 1, \ldots, n\}$ is also a convex set. To see this, consider $x_1, x_2 \in \bigcap_{i=1}^{n} S_i$ and $0 \leq \lambda \leq 1$. Then, $\lambda x_1 + (1-\lambda)x_2 \in S_i$ for $i = 1, \ldots, n$ by definition of convex set. Therefore, $\lambda x_1 + (1-\lambda)x_2 \in \bigcap_{i=1}^{n} S_i$. ∎

Proposition

The intersection of any finite number of convex sets is a convex set.

Proof: Suppose $S_1, S_2, \ldots, S_n$ are convex sets. Then their intersection $\bigcap_{i=1}^{n} S_i = \{x : x \in S_i, \forall i = 1, \ldots, n\}$ is also a convex set. To see this, consider $x_1, x_2 \in \bigcap_{i=1}^{n} S_i$ and $0 \leq \lambda \leq 1$. Then, $\lambda x_1 + (1-\lambda)x_2 \in S_i$ for $i = 1, \ldots, n$ by definition of convex set. Therefore, $\lambda x_1 + (1-\lambda)x_2 \in \bigcap_{i=1}^{n} S_i$. ■

Corollary

The solution set of a **system of linear inequalities** *is a convex set.*

Corollary

The solution set of **linear equations** *is a convex set.*

Corollary

The solution set of **constraints** *for linear programming problem (set of feasible solutions) is a convex set.*

Convex Set

Corollary

The solution set of a **system of linear inequalities** *is a convex set.*

Corollary

The solution set of **linear equations** *is a convex set.*

Corollary

The solution set of **constraints** *for linear programming problem (set of feasible solutions) is a convex set.*

Corollary

The solution set of a **system of linear inequalities** *is a convex set.*

Corollary

The solution set of **linear equations** *is a convex set.*

Corollary

The solution set of **constraints** *for linear programming problem (set of feasible solutions) is a convex set.*

Next, we generalize the definition of convex set.

Definition

Given a finite number of points $x_1, x_2, \ldots, x_n$ in a real vector space, a **convex combination** of these points is a point of the form $\alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_n x_n$, where the real numbers $\alpha_i \geq 0$ and $\alpha_1 + \alpha_2 + \cdots + \alpha_n = 1$.

Let $n = 2$, then a convex combination of points x_1, x_2 is in the form

$$x = \lambda_1 x_1 + \lambda_2 x_2,$$

where $\lambda_1 + \lambda_2 = 1$ and $\lambda_1, \lambda_2 \geq 0$. Denote by $\lambda_1 = \lambda$, then $\lambda_2 = 1 - \lambda$ and we get that the convex combination of two points is

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

Let $n = 3$, then

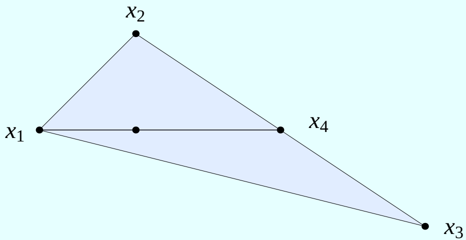
$$x_4 = \alpha_2 x_2 + \alpha_3 x_3,$$

where $\alpha_2 + \lambda_3 = 1$ and $\alpha_2, \alpha_3 \geq 0$; $x = \beta_1 x_1 + \beta_4 x_4$, where $\beta_1 + \beta_4 = 1$ and $\beta_1, \beta_4 \geq 0$. Then, we get that the convex combination of 3 points is:

$$x = \beta_1 x_1 + \beta_4(\alpha_2 x_2 + \alpha_3 x_3) = \beta_1 x_1 + \beta_4 \alpha_2 x_2 + \beta_4 \alpha_3 x_3 = \lambda_1 x_1 + \lambda_2 x_2 + \lambda_3 x_3,$$

where

$$\lambda_1 + \lambda_2 + \lambda_3 = \beta_1 + (\beta_4 \alpha_2 + \beta_4 \alpha_3) = \beta_1 + \beta_4(\alpha_2 + \alpha_3) = \beta_1 + \beta_4 = 1.$$



Definition

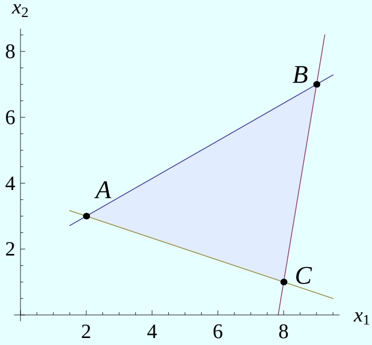
Let x be a **convex combination** of points from the set S. Then, S is called convex if $x \in S$.

Verify that the point $P(6,3)$ is an interior point of the set

$$
\begin{aligned}
-4x_1 + 7x_2 &\leq 13 \\
6x_1 - x_2 &\leq 47 \\
x_1 + 3x_2 &\geq 11
\end{aligned}
$$

and **express** P as a convex combination of the vertices of the solutions of these system.

Substitute coordinates of P to each inequality $\Rightarrow$ all of them hold $\Rightarrow$ $P(6,3)$ is the interior point of the corresponding polytope.



The obtained polytope has 3 vertices. Let us find coordinates of the point A.

$$-4x_1 + 7x_2 = 13$$
$$x_1 + 3x_2 = 11$$

One method for solving such a system is as follows. First, solve the second equation for x_1 in terms of x_2 as

$$x_1 = 11 - 3x_2.$$

Now, substitute this expression for x_1 into the first equation as

$$-4(11 - 3x_2) + 7x_2 = 13.$$

This results in a single equation involving only the variable x_2. Solving gives $x_2 = 3$, and substituting this into the equation for x_1 yields $x_1 = 2$. Therefore, $A(2, 3)$.

Similarly, we can calculate that $B(9, 7)$ and $C(8, 1)$.

Next, according to the definition of convex combination, we get

$$X = \alpha A + \beta B + \gamma C$$

with $\alpha + \beta + \gamma = 1$ and $\alpha, \beta, \gamma \geq 0$. Thus, we can construct the following system of equations.

$$\begin{aligned} 2\alpha + 9\beta + 8\gamma &= 6 \\ 3\alpha + 7\beta + \gamma &= 3 \\ \alpha + \beta + \gamma &= 1 \end{aligned}$$

A solution to the system above is given by $\alpha = 7/19, \beta = 8/19, \gamma = 4/19$. Finally, substituting the obtained solution to the expression for X, we get

$$X = \frac{7}{19}A + \frac{4}{19}B + \frac{8}{19}C.$$

Definition

A real valued function $f : S \to \mathbb{R}$ defined on a convex set S in a vector space is called **convex** or **concave** if, for any two points x_1 and x_2 in S and any $0 \leq \lambda \leq 1$, $f(\lambda x_1 + (1-\lambda)x_2) \leq \lambda f(x_1) + (1-\lambda)f(x_2)$ or $f(\lambda x_1 + (1-\lambda)x_2) \geq \lambda f(x_1) + (1-\lambda)f(x_2)$.

Example:

- $f(x) = x^2$ is convex on $\mathbb{R}$;
- $f(x) = \log x$ is concave on $\mathbb{R}^+$;
- $f(x) = \dfrac{1}{x}$ is convex on $\mathbb{R}^+$ and concave on $\mathbb{R}^-$;
- $f(x) = x^3 - x$ is neither convex nor concave on $\mathbb{R}$.

Proposition

A linear function $f = a_1x_1 + a_2x_2 + \cdots + a_nx_n$ *is both* **convex** *and* **concave**.

Proof: Let $X_1(x_1', x_2', \ldots, x_n')$ and $X_2(x_1'', x_2'', \ldots, x_n'')$. Then,

$$\lambda X_1 + (1-\lambda)X_2 = (\lambda x_1' + (1-\lambda)x_1'', \ldots, \lambda x_n' + (1-\lambda)x_n'')$$

and

$$\begin{aligned} f(\lambda X_1 + (1-\lambda)X_2) &= a_1(\lambda x_1' + (1-\lambda)x_1'') + \cdots + a_n(\lambda x_n' + (1-\lambda)x_n'') = \\ &= \lambda(a_1x_1' + a_2x_2' + \cdots + a_nx_n') + (1-\lambda)(a_1x_1'' + a_2x_2'' + \cdots + a_nx_n'') = \\ &= \lambda f(X_1) + (1-\lambda)f(X_2). \end{aligned}$$

Hence, we can conclude that f is convex and concave. ■

Proposition

A linear function $f = a_1x_1 + a_2x_2 + \cdots + a_nx_n$ *is both* **convex** *and* **concave**.

Proof: Let $X_1(x_1', x_2', \ldots, x_n')$ and $X_2(x_1'', x_2'', \ldots, x_n'')$. Then,

$$\lambda X_1 + (1-\lambda)X_2 = (\lambda x_1' + (1-\lambda)x_1'', \ldots, \lambda x_n' + (1-\lambda)x_n'')$$

and

$$\begin{aligned} f(\lambda X_1 + (1-\lambda)X_2) &= a_1(\lambda x_1' + (1-\lambda)x_1'') + \cdots + a_n(\lambda x_n' + (1-\lambda)x_n'') = \\ &= \lambda(a_1x_1' + a_2x_2' + \cdots + a_nx_n') + (1-\lambda)(a_1x_1'' + a_2x_2'' + \cdots + a_nx_n'') = \\ &= \lambda f(X_1) + (1-\lambda)f(X_2). \end{aligned}$$

Hence, we can conclude that f is convex and concave. ■

Definition

A function $f(x)$ is said to have a **local** maximum (minimum) at x_0 if there exists an interval I around x_0 such that $f(x_0) \geq f(x)$ $(f(x_0) \leq f(x))$ for all $x \in I$.

Definition

We say that the function $f(x)$ has a **global** maximum (minimum) at $x = x_0$ on the interval I, if $f(x_0) \geq f(x)$ $(f(x_0) \leq f(x))$ for all $x \in I$.

Note that if $f(x)$ is a continuous function on a closed bounded interval $[a, b]$, then $f(x)$ will have a global maximum and a global minimum on $[a, b]$. On the other hand, if the interval is not bounded or closed, then there is no guarantee that a continuous function $f(x)$ will have global extremum.

Example:

- $f(x) = x^2$ does not have a global maximum on the interval $[0, \infty)$;
- function $f(x) = -\dfrac{1}{x}$ does not have a global minimum on the interval $(0, 1)$.

Convex Optimization Problem

Definition

A **convex optimization problem** is a problem where all of the constraints are convex functions, and the objective is a convex function if minimizing, or a concave function if maximizing.

Theorem

For a convex optimization problem all locally optimal points are globally optimal.

The linear programming problem can be stated as follows:

$$z = c_1x_1 + c_2x_2 + \cdots + c_nx_n \to \min$$

subject to the constraints

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &\leq b_1 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &\leq b_m \end{aligned}$$

and

$$x_1 \geq 0, x_2 \geq 0, \ldots, x_n \geq 0.$$

Theorem

If a linear programming problem has a solution, then it must **occur at a vertex**, *or corner point, of the feasible set S, associated with the problem. Furthermore, if the objective function z is optimized at two adjacent vertices of S, then it is optimized at every point on the line segment joining these two vertices, in which case there are infinitely many solutions to the problem.*

Proof: The proof is by contradiction. Suppose that the optimal solution x^* is an interior point of the feasible set S. Since the set is convex, then there exist two points $x_1, x_2 \in S$ such that $x^* \in [x_1, x_2]$, i.e.,

$$x^* = \lambda x_1 + (1 - \lambda) x_2.$$

We know that x^* is optimal solution, then denoting

$$f(x) := c_1 x_1 + \cdots + c_n x_n,$$

we get

$$\begin{aligned} f(x^*) &\geq f(x_1), \\ f(x^*) &\geq f(x_2). \end{aligned} \tag{1}$$

Since $f(x)$ is linear (convex) function

$$f(x^*) = f(\lambda x_1 + (1 - \lambda x_2)) = \lambda f(x_1) + (1 - \lambda) f(x_2). \tag{2}$$

Substituting (2) to (1), we get

$$f(x_1) \leq \lambda f(x_1) + (1 - \lambda) f(x_2)$$
$$f(x_2) \leq \lambda f(x_1) + (1 - \lambda) f(x_2)$$

or after simple transformations $f(x_1) = f(x_2)$. From (2) it follows that

$$f(x^*) = \lambda f(x_1) + (1 - \lambda) f(x_2) = f(x_1).$$

Hence, we get

$$f(x_1) = f(x_2) = f(x^*) = z_0 \in \mathbb{R}.$$

As a result, points x_1, x_2, x^* are in the hyperplane $f(x) = z_0$. We know that the point x^* defines this hyperplane; however, the end points of the line segment $[x_1, x_2]$ are free to choose. Therefore, points x_1, x_2 may not necessarily belong to this hyperplane. This contradicts our assumption, showing that x^* has to be on the boundary of the set S. ■

Corollary

Above theorem tells us that our search for the solution(s) to a linear programming problem may be restricted to the **examination** of the set of vertices of the feasible set S associated with the problem. Since a feasible set S has finitely many vertices, the theorem suggest that the solution(s) may be found by inspecting the values of the objective function z at these vertices.