Non-Linear Programming

Convex Programming Problems

Anand Jacob Abraham

Department of Industrial & Systems Engineering
IIT Kharagpur

Table of contents

Convex set and convex/concave functions

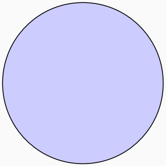
Convex sets

Definition (Convex set)

The set S is convex if $\alpha \boldsymbol{x} + (1-\alpha)\boldsymbol{y} \in S \forall \alpha \in (0,1)$ and $\forall \boldsymbol{x}, \boldsymbol{y} \in S$.

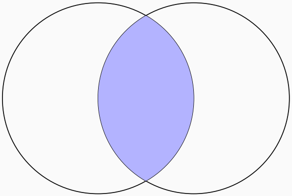
- Basic geometric intuition of a convex set is that if you pick any distinct vectors in the set, it is possible to draw a straight line segment which can be drawn between these points without the line segment ever leaving the set.

Convex sets examples

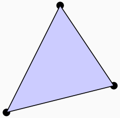


(a) Convex set (disk)

(b) Half-space



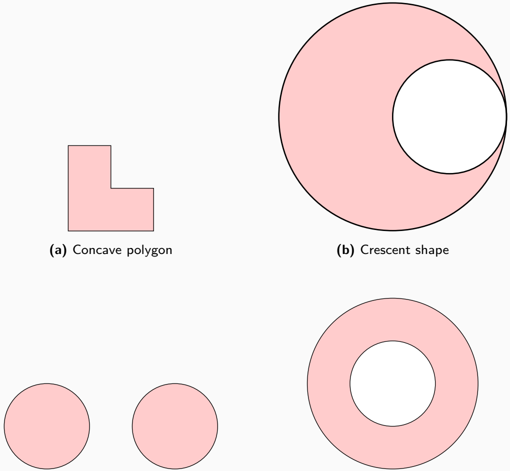
(c) Intersection of convex sets



(d) Convex hull

Figure 1: Examples of convex sets

Non-convex sets examples



(a) Concave polygon

(b) Crescent shape

(c) Disconnected set

(d) Annulus (ring)

Figure 2: Examples of non-convex sets

Hypograph and Epigraph

The pair $(\boldsymbol{x}, l(\boldsymbol{x}))$ is called the graph of a function.

Definition (Hypograph)

The epigraph of a function $l : \mathcal{D} \mapsto \mathbb{R}$ is the locus of points/vectors $[\boldsymbol{x}, \beta]^T \in \mathcal{D} \times \mathbb{R} : \beta \leq l(\boldsymbol{x})$. Mathematically denoted by hyp(l).

Definition (Epigraph)

The epigraph of a function $l : \mathcal{D} \mapsto \mathbb{R}$ is the locus of points/vectors $[\boldsymbol{x}, \beta]^T \in \mathcal{D} \times \mathbb{R} : \beta \geq l(\boldsymbol{x})$. Mathematically, denoted by epi(l)

Hypograph and Epigraph: Geometrically

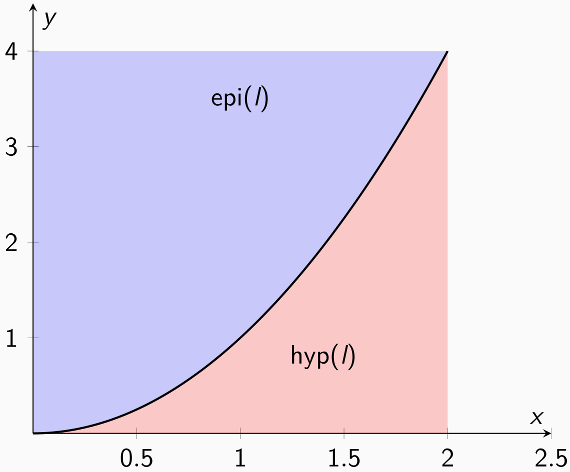


Figure 3: Graph, epi(l), hyp(l) of x^2 on [0,2]

Convex and concave functions

Definition (Convex function)

A function $l : \mathcal{D} \mapsto \mathbb{R}$ is convex if $\mathrm{epi}(l)$ is a convex set.

Definition (Concave function)

A function $l : \mathcal{D} \mapsto \mathbb{R}$ is convex if its $\mathrm{hyp}(l)$ is a convex set.

Properties of convex function

1. $l : \mathcal{D} \mapsto \mathbb{R}$ is a convex function if $f(\alpha \boldsymbol{x} + (1-\alpha)\boldsymbol{y}) \leq \alpha f(\boldsymbol{x}) + (1-\alpha) f(\boldsymbol{y}) \forall \boldsymbol{x}, \boldsymbol{y} \in \mathcal{D}$.
2. $l : \mathcal{D} \mapsto \mathbb{R}$ is a strict convex function if $f(\alpha \boldsymbol{x} + (1-\alpha)\boldsymbol{y}) < \alpha f(\boldsymbol{x}) + (1-\alpha) f(\boldsymbol{y}) \forall \boldsymbol{x}, \boldsymbol{y} \in \mathcal{D}$.
3. $l : \mathcal{D} \mapsto \mathbb{R}$ is a concave function if $f(\alpha \boldsymbol{x} + (1-\alpha)\boldsymbol{y}) \geq \alpha f(\boldsymbol{x}) + (1-\alpha) f(\boldsymbol{y}) \forall \boldsymbol{x}, \boldsymbol{y} \in \mathcal{D}$.
4. $l : \mathcal{D} \mapsto \mathbb{R}$ is a strict concave function if $f(\alpha \boldsymbol{x} + (1-\alpha)\boldsymbol{y}) > \alpha f(\boldsymbol{x}) + (1-\alpha) f(\boldsymbol{y}) \forall \boldsymbol{x}, \boldsymbol{y} \in \mathcal{D}$.
5. If f_1, f_2 are concave functions, then $f_1 + f_2$ is also concave.
6. If $l : \mathcal{D} \mapsto \mathbb{R}$ is strictly convex, then $-l$ is strictly concave.
7. If a function $l : \mathcal{D} \mapsto \mathbb{R}$ is convex/concave, then the set $\mathcal{D}$ is a convex set.

Alternate definitions for concave function

- $l : \mathcal{D} \mapsto \mathbb{R}$ is a concave function if and only if the Hessian of the function is negative semi-definite.
- $l : \mathcal{D} \mapsto \mathbb{R}$ is a concave function if and only if $f(\alpha \boldsymbol{u} + (1-\alpha)\boldsymbol{v}) \geq \alpha f(\boldsymbol{u}) + (1-\alpha) f(\boldsymbol{v}) \forall \boldsymbol{u}, \boldsymbol{v} \in \mathcal{D}$
- Similarly, we can get results for convex functions also.

Convex Programming Problems and its Properties

Convex Programming Problem

Definition (Convex Programming Problem)

A Convex Programming Problem (CPP) is an optimization problem $\max_{\boldsymbol{x} \in S} f(\boldsymbol{x})$ where f is concave or convex and S is a convex set.

- If the objective function is concave and the problem is of maximization type with convex constraint set, some books/literature calls this type of problems as *Concave Maximization Problems* (CMP).

We now show three important results (theorems) for CPPs/CMP. There are some intermediate Lemmas as well (used to derive other main results).

Result 1: Local maximizers are global

Theorem

$\boldsymbol{x}^*$ is a local maximizer to a CPP $\max_{\boldsymbol{x} \in S} f(\boldsymbol{x})$ if and only if it is a global maximizer.

Proof.

Suppose $\boldsymbol{x}^*$ is a local maximizer which is not a global maximizer.

$$\exists \boldsymbol{y} \in S : f(\boldsymbol{y}) > f(\boldsymbol{x}^*)$$

, then as $y \in S$ and due to convexity of S and concaveness of f, we may write

$$f(\alpha \boldsymbol{y} + (1-\alpha)\boldsymbol{x}^*) \geq \alpha f(\boldsymbol{y}) + (1-\alpha) f(\boldsymbol{x}^*) \forall \alpha \in (0,1)$$

$$\implies f(\alpha \boldsymbol{y} + (1-\alpha)\boldsymbol{x}^*) \geq \alpha f(\boldsymbol{y}) + (1-\alpha) f(\boldsymbol{x}^*) > f(\boldsymbol{x}^*) \forall \alpha \in (0,1)$$

$$\implies f(\alpha \boldsymbol{y} + (1-\alpha)\boldsymbol{x}^*) > f(\boldsymbol{x}^*) : \alpha > 0, \alpha \to 0$$

This means that there exists a vector $\alpha \boldsymbol{y} + (1-\alpha)\boldsymbol{x}^*$ which in the neighborhood of $\boldsymbol{x}^*$ which is better than $\boldsymbol{x}^*$[1]. This contradicts the local optimality of $\boldsymbol{x}^*$. Which means that if $\boldsymbol{x}^*$ is locally optimal then it also has to be globally optimal in a CPP.

The only if part of the proof is trivial. □

[1]This point can be seen visualized as a point on the line connecting $\boldsymbol{x}^*$ and $\boldsymbol{y}$ on the set S. Every point on this line guarantees a better objective function value owing to

A less important short result

Lemma

If $f : \mathcal{D} \mapsto \mathbb{R}$ *is concave then* $\forall \boldsymbol{u}, \boldsymbol{v} \in \mathcal{D}$

$$f(\boldsymbol{v}) \leq f(\boldsymbol{u}) + (\boldsymbol{v} - \boldsymbol{u})^T \nabla f(\boldsymbol{u})$$

Proof.

As the function is concave in $\mathcal{D}$, we have

$$f(\alpha \boldsymbol{v} + (1 - \alpha)\boldsymbol{u}) \geq \alpha f(\boldsymbol{u}) + (1 - \alpha) f(\boldsymbol{v}) \forall \alpha \in (0, 1).$$

On simplifying the above expression, we get the following.

$$f(\boldsymbol{u} + \alpha(\boldsymbol{v} - \boldsymbol{u})) \geq f(\boldsymbol{u}) + \alpha(f(\boldsymbol{v}) - f(\boldsymbol{u}))$$

Upon rearrangement we get

$$f(\boldsymbol{u}) + \frac{f(\boldsymbol{u} + \alpha(\boldsymbol{v} - \boldsymbol{u})) - f(\boldsymbol{u})}{\alpha} \geq f(\boldsymbol{v})$$

Applying the limit $\alpha \to 0$, we get the following required result.

$$f(\boldsymbol{u}) + (\boldsymbol{v} - \boldsymbol{u})^T \nabla f(\boldsymbol{u}) \geq f(\boldsymbol{v})$$

□

Result 2: Sufficiency of first order condition in Unconstrained optimization

Theorem

If $\boldsymbol{x}^*$ is a stationary point for the optimization problem $\max_{\boldsymbol{x}\in\mathbb{R}^n} f(\boldsymbol{x})$, then $\boldsymbol{x}^*$ is a global maximizer.

Proof.

As $\boldsymbol{x}^*$ is a stationary point $\nabla f(\boldsymbol{x}^*) = \boldsymbol{0}$. Assume that $\boldsymbol{x}^*$ is not the global maximizer, then $\exists \boldsymbol{y} \in \mathbb{R}^n : f(\boldsymbol{y}) > f(\boldsymbol{x}^*)$. Using the previous lemma 13 we can write the following.

$$f(\boldsymbol{y}) \leq f(\boldsymbol{x}^*) + (\boldsymbol{y} - \boldsymbol{x}^*)^T \nabla f(\boldsymbol{x}^*)$$

As $\boldsymbol{x}^*$ is a stationary point $\nabla f(\boldsymbol{x}^*) = \boldsymbol{0}$ we get the following $f(\boldsymbol{y}) \leq f(\boldsymbol{x}^*)$. However, this contradicts our earlier assumption that $\boldsymbol{x}^*$ is not a global maximizer. Therefore $\boldsymbol{x}^*$ is a global maximizer. $\square$

Result 3: Sufficiency of KKT

Theorem

If $f : \mathbb{R}^n \mapsto \mathbb{R}$ *is concave over the convex set* $S = \{\boldsymbol{x} : \boldsymbol{h}(\boldsymbol{x}) = \boldsymbol{0}, \boldsymbol{g}(\boldsymbol{x}) \leq \boldsymbol{0}\}$ *and* $\boldsymbol{x}^* \in S$ *satisfies the KKT conditions below*

1. $\nabla f(\boldsymbol{x}^*) = Dh(\boldsymbol{x}^*)^T \boldsymbol{\lambda}^* + Dg(\boldsymbol{x}^*)^T \boldsymbol{\mu}^*.$
2. $\boldsymbol{\mu}^* \boldsymbol{g}(\boldsymbol{x}) = 0.$
3. $\boldsymbol{\mu}^* \geq \boldsymbol{0}.$

then $\boldsymbol{x}^*$ *is a global maximizer.*