CO 250: Introduction to Optimization

Module 6: Nonlinear Programs (Convexity)

Definition

A Nonlinear Program (NLP) is a problem of the form:

$$\boxed{\begin{array}{lll} \min & f(x) & \\ \text{s.t.} & & \\ & g_i(x) \le 0 & (i = 1, \ldots, k) \end{array}}$$

where

$f : \Re^n \to \Re$, and

$g_i : \Re^n \to \Re$ for $i = 1, \ldots, k$.

Remark

There aren't any restrictions regarding the type of functions.

This is a very general model, but NLPs can be very hard to solve!

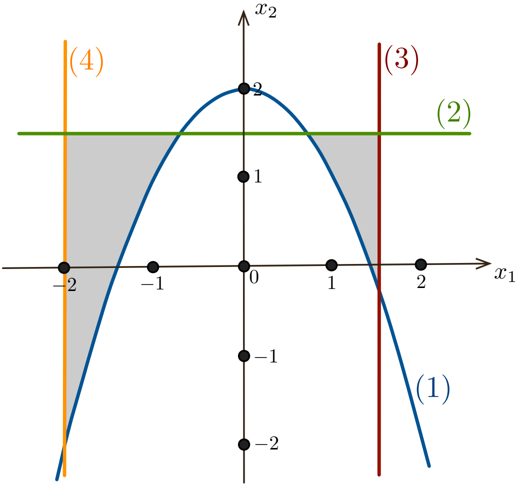
Definition

A Nonlinear Program (NLP) is a problem of the form:

$$\boxed{\begin{array}{lll} \min & f(x) & \\ \text{s.t.} & & \\ & g_i(x) \leq 0 & (i = 1, \ldots, k) \end{array}}$$

$$\boxed{\begin{array}{lrcl} \min & x_2 & & \\ \text{s.t.} & & & \\ & -x_1^2 - x_2 + 2 & \leq & 0 \\ & x_2 - \frac{3}{2} & \leq & 0 \\ & x_1 - \frac{3}{2} & \leq & 0 \\ & -x_1 - 2 & \leq & 0 \end{array}}$$

$$\begin{array}{lrcl} \min & x_2 & & \\ \text{s.t.} & & & \\ & -x_1^2 - x_2 + 2 & \leq & 0 \\ & x_2 - 3/2 & \leq & 0 \\ & x_1 - 3/2 & \leq & 0 \\ & -x_1 - 2 & \leq & 0 \end{array}$$



FEASIBLE REGION

(1) $x_2 \geq 2 - x_1^2$.
(2) $x_2 \leq \frac{3}{2}$.
(3) $x_1 \leq \frac{3}{2}$.
(4) $x_1 \geq 2$.

Definition

A Nonlinear Program (NLP) is a problem of the form:

$$\boxed{\begin{array}{lll} \min & f(x) & \\ \text{s.t.} & & \\ & g_i(x) \leq 0 & (i = 1, \ldots, k) \end{array}} \qquad \text{(P)}$$

Remark

We may assume $f(x)$ is a linear function, i.e., $f(x) = c^\top x$.

We can rewrite (P) as

$$\boxed{\begin{array}{lll} \min & \lambda & \\ \text{s.t.} & & \\ & \lambda \geq f(x) & \\ & g_i(x) \leq 0 & (i = 1, \ldots, k) \end{array}} \qquad \text{(Q)}$$

The optimal solution to (Q) will have $\lambda = f(x)$.

Nonlinear Programs Generalize Linear Programs

$$
\begin{array}{ll}
\max & x_1 + x_2 \\
\text{s.t.} & \\
& \begin{array}{rcl} 2x_1 - x_2 & \geq & 3 \\ x_1 - x_2 & = & 4 \end{array} \\
& x_1, x_2 \geq 0
\end{array}
$$

$$
\begin{array}{ll}
\min & -x_1 - x_2 \\
\text{s.t.} & \\
& \begin{array}{rcl} -2x_1 + x_2 + 3 & \leq & 0 \\ x_1 - x_2 - 4 & \leq & 0 \\ -x_1 + x_2 + 4 & \leq & 0 \\ -x_1 & \leq & 0 \\ -x_2 & \leq & 0 \end{array}
\end{array}
$$

Nonlinear Programs can also generalize Integer Programs!

Nonlinear Programs Generalize Integer Programs

$$\begin{array}{ll} \max & c^\top x \\ \text{s.t.} & \\ & Ax \leq b \\ & x_j \in \{0,1\} \quad (j = 1, \ldots, n) \end{array}$$

$0, 1$ IP

Idea

$x_j \in \{0,1\} \quad \Longleftrightarrow \quad x_j(1 - x_j) = 0$

$$\begin{array}{ll} \min & -c^\top x \\ \text{s.t.} & \\ & \begin{array}{rcll} Ax & \leq & b & \\ x_j(1-x_j) & \leq & 0 & (j = 1, \ldots, n) \\ -x_j(1-x_j) & \leq & 0 & (j = 1, \ldots, n) \end{array} \end{array}$$

Quadratic NLP

Remark

$0, 1$ IPs are hard to solve; thus, quadratic NLPs are also hard to solve.

Nonlinear Programs Generalize Integer Programs

$$\begin{aligned} \max \quad & c^\top x \\ \text{s.t.} \quad & \\ & Ax \le b \\ & x_j \text{ integer} \quad (j = 1, \ldots, n) \end{aligned}$$

pure IP

Idea

x_j integer $\quad \Longleftrightarrow \quad \sin(\pi x) = 0.$

$$\begin{aligned} \min \quad & -c^\top x \\ \text{s.t.} \quad & \\ & \begin{array}{rcll} Ax & \le & b & \\ \sin(\pi x) & = & 0 & (j = 1, \ldots, n) \end{array} \end{aligned}$$

Remark

IPs are hard to solve; thus, NLPs are also hard to solve.

Question

What makes solving an NLP hard?

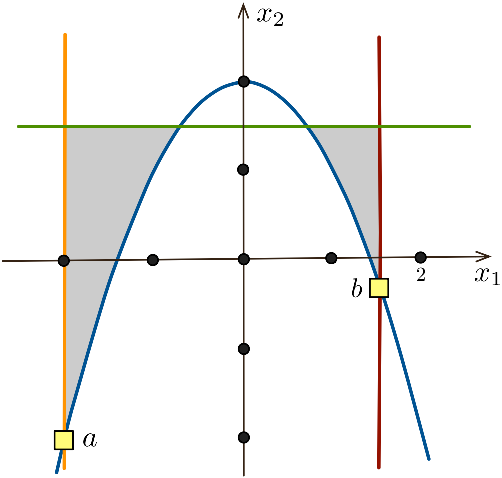
META STRATEGY FOR SOLVING AN OPTIMIZATION PROBLEM

- Find a feasible solution x.
- If x is optimal, STOP.
- Starting with x, find a "better" feasible solution.

Meta strategy for solving an optimization problem

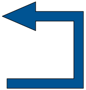
- Find a feasible solution x.
- If x is optimal, STOP.
- Starting with x, find a "better" feasible solution.

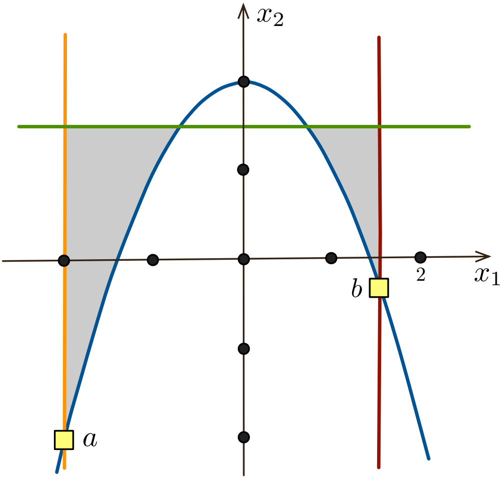
$$
\begin{array}{lrcl}
\min & x_2 & & \\
\text{s.t.} & & & \\
& -x_1^2 - x_2 + 2 & \leq & 0 \\
& x_2 - 3/2 & \leq & 0 \\
& x_1 - 3/2 & \leq & 0 \\
& -x_1 - 2 & \leq & 0
\end{array}
$$



Meta strategy for solving an optimization problem

- Find a feasible solution x.
- If x is optimal, STOP.
- Starting with x, find a "better" feasible solution.





a is an optimal solution

there is no better solution around b

b is a local optimum

Definition

Consider

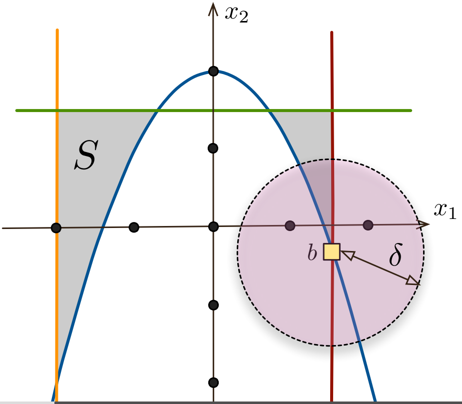
$$\min\{f(x) : x \in S\}. \tag{P}$$

$x \in S$ is a local optimum if there exists $\delta > 0$ such that

$$\forall x' \in S \quad \text{where} \quad ||x' - x|| \le \delta \quad \text{and we have} \quad f(x) \le f(x').$$

$\min\{x_2 : x \in S\}$

b is a local optimum



Convexity Helps

Definition

Consider

$$\min\{f(x) : x \in S\}. \tag{P}$$

$x \in S$ is a local optimum if there exists $\delta > 0$ such that

$$\forall x' \in S \quad \text{where} \quad ||x' - x|| \leq \delta \quad \text{and we have} \quad f(x) \leq f(x').$$

Proposition

Consider

$$\min\left\{c^\top x : x \in S\right\}. \tag{P}$$

If S is convex and x is a local optimum, then x is optimal.

Proposition

Consider

$$\min \left\{c^\top x : x \in S\right\}. \qquad \text{(P)}$$

If S is convex and x is a local optimum, then x is optimal.

Proof

Suppose $\exists x' \in S$ with $c^\top x' < c^\top x$.

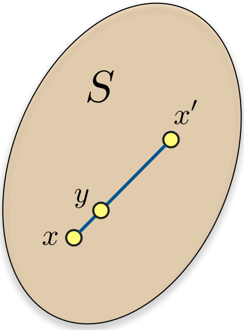
Let $y = \lambda x' + (1-\lambda)x$ for $\lambda > 0$ small.

Since S is convex, $y \in S$.

As λ small $||y - x|| \leq \delta$.

$$\begin{aligned} c^\top y &= c^\top(\lambda x' + (1-\lambda)x) \\ &= \underbrace{\lambda}_{\geq 0}\,\underbrace{c^\top x'}_{< c^\top x} + \underbrace{(1-\lambda)}_{\geq 0}\, c^\top x \\ &< \lambda c^\top x + (1-\lambda) c^\top x \\ &= c^\top x \end{aligned}$$

This is a contradiction.



Consider

$$\boxed{\begin{array}{lll} \min & c^\top x & \\ \text{s.t.} & & \\ & g_i(x) \le 0 & (i = 1, \ldots, k) \end{array}} \qquad \text{(P)}$$

Goal: Study the case where the feasible region of (P) is convex.

- We will define convex functions
- We will prove

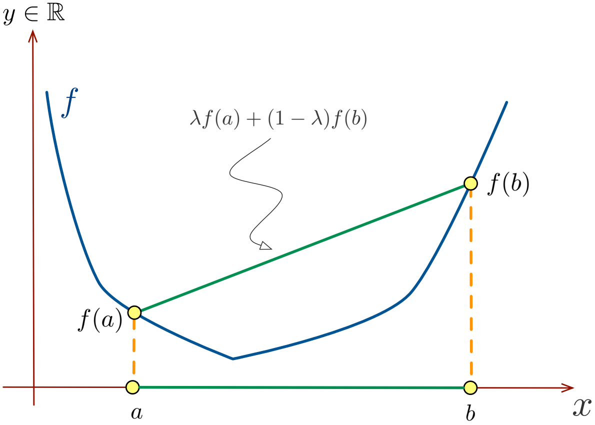
Proposition

If $g_1, \ldots, g_k$ are all convex, then the feasible region of (P) is convex.

Definition

Function $f : \Re^n \to \Re$ is convex if for all $a, b \in \Re^n$,

$$f\big(\lambda a + (1-\lambda)b\big) \leq \lambda f(a) + (1-\lambda) f(b) \quad \text{for all} \quad 0 \leq \lambda \leq 1.$$

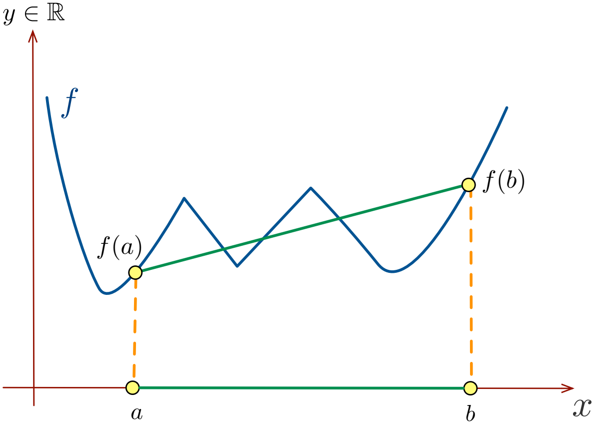


CONVEX FUNCTION!

Definition

Function $f : \Re^n \to \Re$ is convex if for all $a, b \in \Re^n$,

$$f\big(\lambda a + (1-\lambda)b\big) \leq \lambda f(a) + (1-\lambda) f(b) \quad \text{for all} \quad 0 \leq \lambda \leq 1.$$

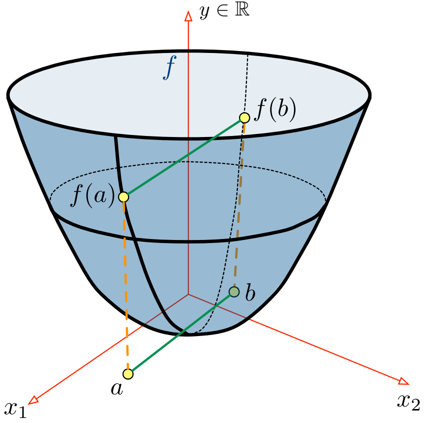


Not a convex function!

Definition

Function $f : \Re^n \to \Re$ is convex if for all $a, b \in \Re^n$,

$$f\big(\lambda a + (1-\lambda)b\big) \leq \lambda f(a) + (1-\lambda)f(b) \quad \text{for all} \quad 0 \leq \lambda \leq 1.$$



Convex function!

Example

We claim that $f(x) = x^2$ is convex.

Pick $a, b \in \Re$ and pick λ where $0 \leq \lambda \leq 1$.

To check:

$$[\lambda a + (1-\lambda)b]^2 \stackrel{?}{\leq} \lambda a^2 + (1-\lambda)b^2.$$

We may assume that $\lambda \neq 0, 1$.

After simplifying

$$\lambda(1-\lambda)2ab \stackrel{?}{\leq} \lambda(1-\lambda)(a^2+b^2),$$

or, equivalently, as $\lambda, (1-\lambda) > 0$,

$$a^2 + b^2 - 2ab \stackrel{?}{\geq} 0,$$

which is the case as $a^2 + b^2 - 2ab = (a-b)^2 \geq 0$.

Why Do We Care About Convex Functions?

Proposition

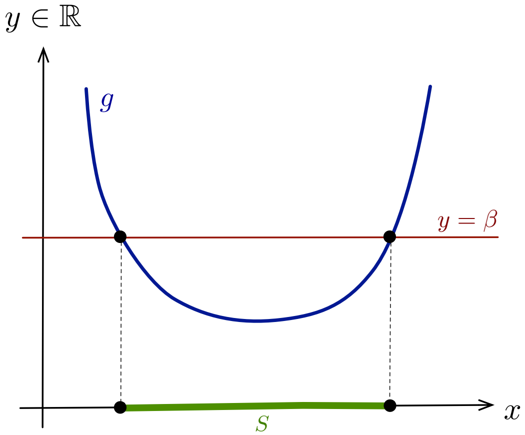
Let $g : \Re^n \to \Re$ be a convex function and $\beta \in \Re$.

It follows that $S = \{x \in \Re^n : g(x) \leq \beta\}$ is a convex set.

Proposition

Let $g : \Re^n \to \Re$ be a convex function and $\beta \in \Re$.

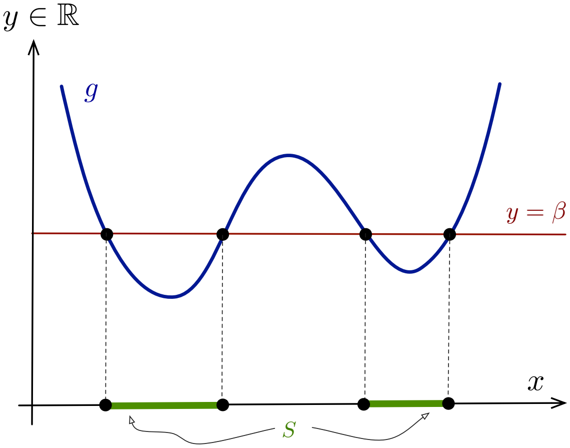
It follows that $S = \{x \in \Re^n : g(x) \leq \beta\}$ is a convex set.



Proposition

Let $g : \Re^n \rightarrow \Re$ be a convex function and $\beta \in \Re$.

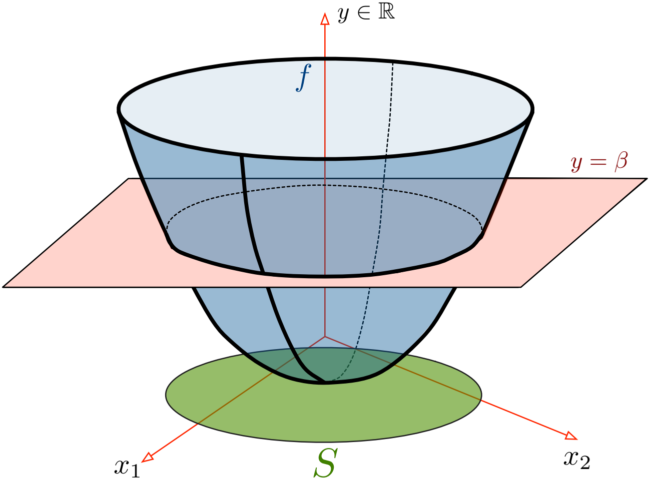
It follows that $S = \{x \in \Re^n : g(x) \leq \beta\}$ is a convex set.



Proposition

Let $g : \Re^n \to \Re$ be a convex function and $\beta \in \Re$.

It follows that $S = \{x \in \Re^n : g(x) \le \beta\}$ is a convex set.



Proposition

Let $g : \Re^n \to \Re$ be a convex function and $\beta \in \Re$.

It follows that $S = \{x \in \Re^n : g(x) \leq \beta\}$ is a convex set.

Proof

Pick $a, b \in S$.

Pick λ where $0 \leq \lambda \leq 1$.

Let $x = \lambda a + (1 - \lambda)b$.

Our goal is to show that $x \in S$, i.e., that $g(x) \leq \beta$.

$$
\begin{aligned}
g(x) &= g\left(\lambda a + (1-\lambda)b\right) \\
&\leq \underbrace{\lambda}_{\geq 0} \underbrace{g(a)}_{\leq \beta} + \underbrace{(1-\lambda)}_{\geq 0} \underbrace{g(b)}_{\leq \beta} \qquad (\text{since } a, b \in S) \\
&\leq \lambda\beta + (1-\lambda)\beta \\
&= \beta
\end{aligned}
$$

Proposition

$$\boxed{\begin{array}{ll} \min & c^\top x \\ \text{s.t.} & \\ & g_i(x) \leq 0 \qquad (i = 1, \ldots, k) \end{array}} \tag{P}$$

If all functions g_i are convex, then the feasible region of (P) is convex.

Proof

Let $S_i = \{x : g_i(x) \leq 0\}$.

By the previous result, S_i is convex.

The feasible region of (P) is $S_1 \cap S_2 \cap \ldots \cap S_k$.

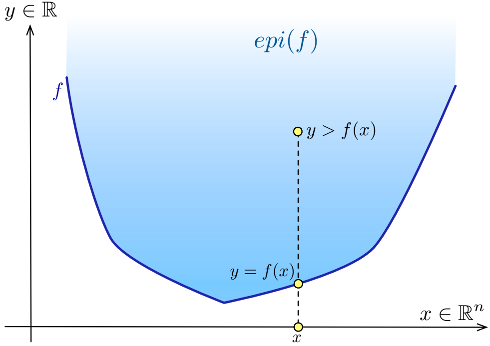
Since the intersection of convex sets is convex, the result follows.

Convex Functions Versus Convex Sets

Definition

Let $f : \Re^n \to \Re$ be a function. The epigraph of f is then given by,

$$epi(f) = \left\{ \begin{pmatrix} y \\ x \end{pmatrix} : y \geq f(x), x \in \Re^n \right\} \subseteq \Re^{n+1}.$$



f is convex

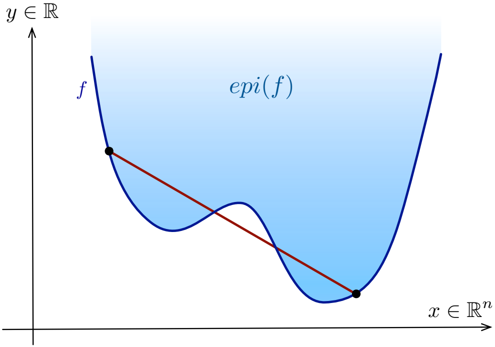
$epi(f)$ is convex

Convex Functions Versus Convex Sets

Definition

Let $f : \Re^n \to \Re$ be a function. The epigraph of f is then given by,

$$epi(f) = \left\{ \begin{pmatrix} y \\ x \end{pmatrix} : y \geq f(x), x \in \Re^n \right\} \subseteq \Re^{n+1}.$$



f is NOT convex

$epi(f)$ is NOT convex

Convex Functions Versus Convex Sets

Definition

Let $f : \Re^n \to \Re$ be a function. The epigraph of f is then given by,

$$epi(f) = \left\{ \begin{pmatrix} y \\ x \end{pmatrix} : y \geq f(x), x \in \Re^n \right\} \subseteq \Re^{n+1}.$$

Proposition

Let $f : \Re^n \to \Re$ be a function. It follows that

1. f is convex $\implies$ $epi(f)$ is convex.
2. $epi(f)$ is convex $\implies$ f is convex.

Proposition

Let $f : \Re^n \to \Re$ be a function. It follows that

1. $\boxed{f \text{ is convex} \quad \Longrightarrow \quad epi(f) \text{ is convex.}}$
2. $epi(f)$ is convex $\quad \Longrightarrow \quad$ f is convex.

Proof

Pick $\begin{pmatrix} \alpha \\ a \end{pmatrix} \begin{pmatrix} \beta \\ b \end{pmatrix} \in epi(f)$. Pick λ where $0 \leq \lambda \leq 1$.

Our goal is to show: $epi(f)$ contains

$$\lambda \begin{pmatrix} \alpha \\ a \end{pmatrix} + (1-\lambda) \begin{pmatrix} \beta \\ b \end{pmatrix} = \begin{pmatrix} \lambda\alpha + (1-\lambda)\beta \\ \lambda a + (1-\lambda) b \end{pmatrix} \tag{$\star$}$$

Consider

$$\begin{aligned} f\left(\lambda a + (1-\lambda)b\right) \quad &\leq \quad (\text{convexity of } f) \\ \underbrace{\lambda}_{\geq 0} \underbrace{f(a)}_{\leq \alpha} + \underbrace{(1-\lambda)}_{\geq 0} \underbrace{f(b)}_{\leq \beta} \quad &\leq \quad \lambda\alpha + (1-\lambda)\beta \end{aligned}$$

Thus, $(\star)$ holds.

Recap

1. NLPs are hard in general.
2. We may assume the objective function of NLPs is linear.
3. Local optimum = optimal sol when the feasible region is convex.
4. We defined convex functions.
5. Convex functions yield a convex feasible region.
6. Convex functions and convex sets are related by epigraphs.