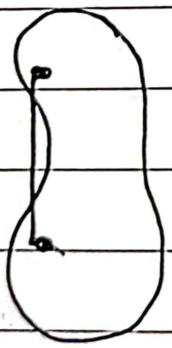
# Convex optimization

What is a convex opt problem? But before that, convex set.


Round from the outside and no bumps to the inside then it is a convex set.

A set is convex if we take any two points in the set, connect these points by a straight line and the ~~straight~~ straight line is still contained within the set.

Eg:-



[Potato with no bumps example]

Not convex

(def ~~AD~~ ANY 2 points should satisfy condition)

Date / /

Notes

Formal definition :-

A subset S of a vector space is called convex if $\forall x, y \in S$ and $\forall t \in [0,1]$

$$tx + (1-t)y \in S$$

t => paramter that drives one from x to y

Beware! x, y are vectors here.

$t = 0$ — point y

$t = 1$ — point x

Any other value lies on the line between x and y.

That line should be inside S.

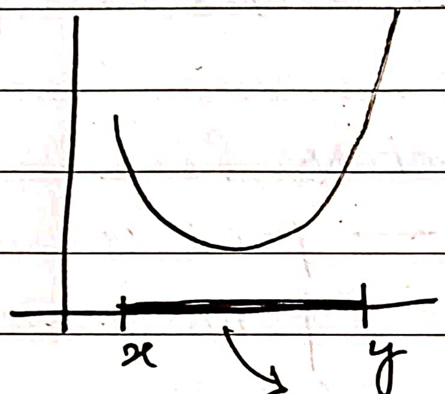
What is a convex function?

We took at the graph of the function. and take any 2 points. Then connect them. This line needs to be above the graph of the function.

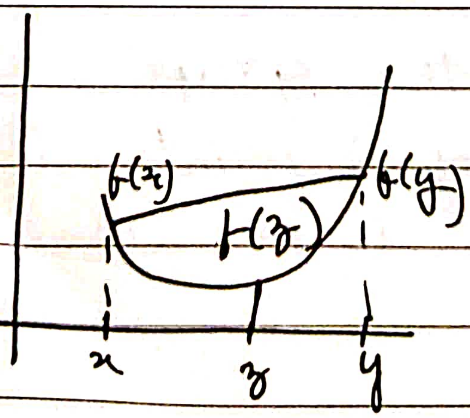
Formal definition

$f: S \longrightarrow \mathbb{R}$ that is defined on a convex domain $S$ is called convex $\forall x, y \in S$ and $t \in [0, 1]$

$$f(tx + (1-t)y) \leq t f(x) + (1-t) f(y)$$




on input domain
but function domain
is supposed to be
convex




Pick a point $z$

$tx + (1-t)y$

Notes Date / /

$f(z) = f(tx + (1-t)y)$

$\leq$ line between function values

$\leq tf(x) + (1-t)f(y)$

This is def of a convex function.

Concave would be: replace '$\leq$' with '$\geq$'
Some functions might neither be concave nor convex.

~~Examples:~~

Level curves — sort of Contour lines which ~~Sho~~ tell us how a function grows or shrinks.

Finally, we come back to

What is Convex opt?

This is an opt problem:-

→ minimize $f$

→ subject to $g_i(x) \leq 0$ $(i=1,\dots,k)$

This optimization problem is called convex if the functions $f$ and $g_i$ are convex.

Basically we want to find the min of convex function $f$ where

$g_1(x) \leq 0$

$g_2(x) \leq 0$

$\vdots$

$g_k(x) \leq 0$

Notes

Date / /

$f, g_i$ $(i = 1, \ldots, k)$ are convex.

Sometimes, we also consider some equality constraints

$h_j = 0$ instead of $g_i(x) \leq 0$ condition.

Sometimes, ~~inequality~~ equality ~~conditions~~ constraints $h_j \leq 0$ and $-h_j \leq 0$.

Why are we interested ?

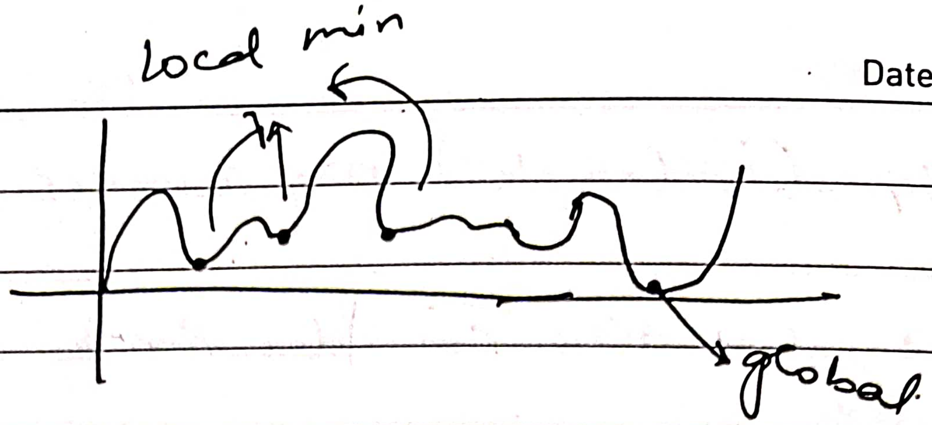
Convex opt problems have a property "any local minimum is already a global minimum".

Some terms :-

→ $f$ ⇒ objective function

→ $g_i$ ⇒ constrains $(i = \{1, \ldots, k\})$

→ Set of points where all constraints are satisfied called ⇒ feasible set.




In normal, we would ~~fo~~ go gradient descent and eventually find min, but we will find a local minimum. Not optimal!

Objective function should be convex.

Problem defined as

$f_0(\bar{x}) \longrightarrow$ Convex function to be minimized.

$f_i(\bar{x}) \leq 0 \quad i \in [1, m]$

$\longrightarrow$ convex inequality

$h_j(\bar{x}) = 0$

$\longrightarrow$ affine equality

Affine functions:-

$$h(\bar{x}) = a^T \bar{x} + b$$

Notes

Date / /

Equality constraints

$\bar{x} = x, y$

$h(\bar{x}) = h(x, y)$

$\star\ h(x, y) = 0$

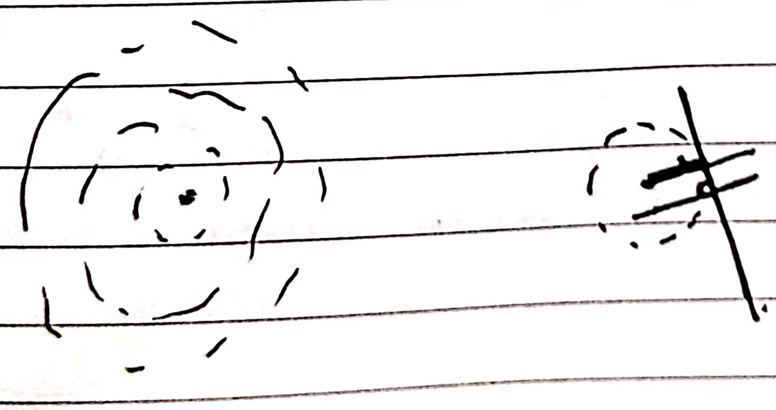
$f(x, y) \Longrightarrow$ function to minimize.

The point $(x, y)$ when $f$ is min and $h(x, y) = 0$ is the point where

$h(x, y)$ is a tangent to $f(x, y)$ at the minima.

Circle eq:-

If $f(x, y)$ is $x^2 + y^2$ level curves will be concentric circles



Normal of level curve of $f$

$\nabla_x f(x)$

$x^*, y^*$ be our minima

$$\nabla_{x,y} f(x,y)\Big|_{x^*, y^*} = \lambda \nabla_{x,y} h(x,y)$$

Cubic

$h_1(x,y,z) = 0$ and $h_2(x,y,z) = 0$
minimize $f$

Solution point will lie on both planes simultaneously and the curve $f$.

$f$ is tangential to intersection of planes $h_1(x,y)$ & $h_2(x,y)$.

Normal to $f$ at $(x^*, y^*, z^*)$
is $\parallel$ to ~~both~~ linear combination of normals to the planes.

$$\nabla_{x,y,z} f(x,y,z) = \lambda_1 \nabla_{x,y,z} h_1(x,y,z)$$
$$+ \lambda_2 \nabla_{x,y,z} h_2(x,y,z)$$

Problem

Notes

$g(x_1, x_2) = x_1 + x_2 - 1$ } Solve

Date / /

Langrangian to generalize.

→ kinda like a recipe to ~~soft~~ solve convex opt problems.

$$L(\bar{x}, \lambda) = f_0(\bar{x}) + \sum_{i=0}^{n} \lambda_i h_i(\bar{x})$$

each $\lambda_i$ => Lagrangian multiplier.

Imp Opt value of $x^*$

$\Longrightarrow$ critical points of $L$

$$\nabla_{x,\lambda} L(\bar{x}, \lambda)\Big|_{x^* \lambda^*} = 0.$$

Example problem:-

$$f_0(x) = x^2 + y^2 + z^2$$

$h_1$ — $x + y = 3$

$h_2$ — $x - y = 3$

$n$ is no. of constraints.

Apply Lagrangian

$$L(\bar{x}, \bar{\lambda}) = f_0(\bar{x}) + \sum_{i=1}^{n} \lambda_i h_i(\bar{x})$$

Notes Date / /

$\mathcal{L}(x, y, z, \lambda_1, \lambda_2)$

$$= x^2+y^2+z^2+\lambda_1(x+y-3) + \lambda_2(x-y-3)$$

$$\frac{\delta \mathcal{L}}{\delta x} = 2x + \lambda_1 + \lambda_2$$

$$\frac{\delta \mathcal{L}}{\delta y} = 2y + \lambda_1 - \lambda_2$$

$$\frac{\delta \mathcal{L}}{\delta z} = 2z$$

$$\frac{\delta \mathcal{L}}{\delta \lambda_1} = x+y-3$$

$$\frac{\delta \mathcal{L}}{\delta \lambda_2} = x-y-3$$

All ~~de~~rivatives are zero.

$$\frac{\delta \mathcal{L}}{\delta\, \cancel{\;}}\ \cancel{=0}$$

Notes Date / /

$2x + \lambda_1 + \lambda_2 = 0$

$2z = 0 \quad \boxed{z = 0}$

$2y + \lambda_1 - \lambda_2 = 0$

$x + y = 3$

$x - y = 3$

$2x = 6$

$\boxed{x = 3}$

$y = 3 - 3$

$\boxed{y = 0}$

$2x + \lambda_1 + \lambda_2 = 0$

$2y + \lambda_1 - \lambda_2 = 0$

$2x + 2y + 2\lambda_1 = 0$

$\cancel{2x}$ .

$6 - 3 + \lambda_2 = 0$

$\boxed{\lambda_2 = -3}$

$6 + 2\lambda_1 = 0$

$\boxed{\lambda_1 = -3}$

Inequality constraints:-

$p^* = f_0(x^*)$

set of all values $x \in S$ such that all constrains are satisfied is feasible region or set!

$p^* \Rightarrow$ Solution of primal problem.

→ For each eq constraint, we introduce $\nu_j \in \mathbb{R}$

→ For each ineq constraint, we introduce $\lambda_i \geq 0$

Lagrangian

$$\mathcal{L}(\bar{x}, \lambda, \nu) = f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{j=1}^{n} \nu_j h_j(x)$$

Next we define the dual function:-

$$g : \mathbb{R}^m \times \mathbb{R}^n \longrightarrow \mathbb{R}$$

$$g(\lambda, \nu) = \inf_x \mathcal{L}(x, \lambda, \nu)$$

So dual function, we get rid of the x and have a function g in terms of $\lambda, \nu$.

$g \Rightarrow$ is ALWAYS concave

Notes

We constrain ourselves to

$$\lambda_i \geq 0 \quad \forall\, i > 0$$

$$f_i(x) \leq 0 \qquad \lambda_i f_i(x) \leq 0 \quad \forall\, i > 0$$

$$\mathcal{L}(x, \lambda, \nu) \leq f_0(x) \quad \forall\, x \in S$$

$$\lambda \geq 0$$

Dual function is a lower bound on primal

$$g(\lambda, \nu) \leq \mathcal{L}(x, \lambda, \nu)$$

$$g(\lambda, \nu) \leq f_0(x) \qquad \forall\, x \in D$$

$$g(\lambda, \nu) \leq f_0(x^*)$$

$$\underset{\text{at } x = x^*}{g(\lambda, \nu)} \leq p^* \qquad [\lambda \geq 0]$$

Dual optimal is $\leq$ primal optimal.

$$\underbrace{d^* \leq p^*}_{\text{weak duality}}$$

Notes Date / /

$(p^* - d^*) \Rightarrow$ duality gap.

We say that strong duality holds when $p^* = d^*$.

$$d^* = \max_{\lambda \geq 0} \left( g(\lambda^*, \gamma^*) \right)$$

$$= \mathcal{L}(\bar{x}_{d^*}, \lambda^*, \gamma^*)$$

$$\Rightarrow d^*$$

① $\mathcal{L}(\bar{x}_d^*, \lambda^*, \gamma^*) \leq \mathcal{L}(\bar{x}^*, \lambda^*, \gamma^*)$

$\because g(\lambda^*, \gamma^*) \leq \mathcal{L}(x, \lambda^*, \gamma^*) \ \forall x \in D$

② $\mathcal{L}(\bar{x}^*, \lambda^*, \gamma^*) \leq f_0(\bar{x}^*)$

$f_0(\bar{x}^*) \rightarrow p^*$

from ① and ②

$$d^* \leq \mathcal{L}(\bar{x}^*, \lambda^*, \gamma^*) \leq p^*$$

$\therefore$

$$d^* = p^*$$

$$p^* = \mathcal{L}(\bar{x}^*, \lambda^*, \gamma^*)$$

$$f_0(\bar{x}^*) = f_0(\bar{x}^*) + \sum_{i=1}^{m} \lambda_i f_i(\bar{x}^*) + \sum_{j=1}^{n} \gamma_i h_i(\bar{x}^*)$$

eq constraints $\Rightarrow 0$

$\Rightarrow \sum_{i=1}^{m} \lambda_i f_i(\overline{x}^*) = 0$

$\Rightarrow \lambda \geq 0$ and $f_i(x) \leq 0 \ \forall i > 0$

$\rightarrow \lambda_i^* f_i(x^*) \leq 0 \quad \forall i > 0$

$\Rightarrow \lambda_i^* f_i(x^*) = 0 \quad \forall i > 0$

This property in case of strong duality is called

"Complementary slackness"

Finally KKT conditions:-

Given $f_0(x) \Longrightarrow$ convex fn

ineq $f_i(x) \leq 0 \quad i \in \{1, 2, \cdots, m\}$

eq $h_j(x) = 0 \quad j \in \{1, 2, \ldots, n\}$

$L(x, \lambda, \nu) = f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{j=1}^{n} \nu_j h$

$g(\lambda, \nu) = \inf_{x \in D} L(x, \lambda, \nu)$

If $f_0(x^*) = p^*$
and $g(\lambda^*, \nu^*) = d^*$

KKT (Karush-Kuhn-Tucker) Conditions will be:-

→ $f_i(x^*) \leq 0 \quad \forall\, i \in \{1, \ldots, m\}$
(Primal feasibility, 1)

→ $h_j(x^*) = 0 \quad \forall\, i \in \{1, \ldots, n\}$
(Primal feasibility, 2)

→ $\lambda^* \geq 0$ (Dual feasibility)

→ $\lambda_i^* \cdot f_i(x^*) = 0 \quad \forall\, i = \{1, 2, \ldots, n\}$
(Complementary Slackness)

→ $\nabla_x L(x, \lambda, \nu) \Big|_{x^*, \lambda^*, \nu^*} = 0$

(Stationary)

Then $p^* = d^*$ and they are optimal values for both the primal and dual problem.

Notes Example 2.  Date / /

Minimize

$$f_0(x, y) = x^2 + y^2$$

$$x + y - 1 \leq 0$$

$$x - y + 2 \leq 0$$

$$f_1(x, y) \Rightarrow x + y - 1$$

$$f_2(x, y) \Rightarrow x - y + 2$$

No eq constraints.

Start with complementary slackness

$$\lambda_i f_i(\bar{x}^*) = 0$$

$$\lambda_1 f_1(x, y) = 0 \qquad \lambda_2 f_2(x, y) = 0$$

$$\lambda_1 (x + y - 1) = 0 \qquad \lambda_2 (x - y + 2) = 0$$

Case 1 :-

Notes Date / /

Case 1:- $f_1$ and $f_2$ are 0.

$x+y-1=0$ $x-y+2=0$

No inequalities then

~~L(x,y,~~

Equality case it becomes

$$L = x^2+y^2+r_1(x+y-1)+r_2(x-y+2)$$

$$\frac{\delta L}{\delta x} = 2x+r_1+r_2=0$$

$$\frac{\delta L}{\delta y} = 2y+r_1-r_2=0$$

$$\frac{\delta L}{\delta r_1} = x+y-1=0$$

$$\frac{\delta L}{\delta r_2} = x-y+2=0$$

$$\boxed{x=-\frac{1}{2} \quad y=\frac{3}{2}}$$

$$-1+3+2r_1=0$$

$$2r_1=-2$$

$$\boxed{r_1=-1}$$

$$\boxed{r_2=2}$$

Case 2:- ~~$f_1$~~ ~~$f_2$~~

$f_1 \neq 0$ $f_2=0$

$f_1(x,y) \leq 0$ $f_2=0$

1 ineq and 1 eq

~~$\mathcal{L} = x^2+y^2+\lambda f_1(x,y) + \gamma f_2(x,y)$~~

~~$x^2+y^2+\lambda$~~

If $f_1 \neq 0$ $\lambda_1 = 0$

$$\mathcal{L} = x^2+y^2+\gamma(x-y+2)$$

$$\frac{\delta \mathcal{L}}{\delta x} = 2x+\gamma = 0$$

$$\frac{\delta \mathcal{L}}{\delta y} = 2y-\gamma = 0$$

$$\frac{\delta \mathcal{L}}{\delta \gamma} = x-y+2 = 0$$

$2x+2y=0$

$x=-y$

$-2y=-2$

$\boxed{\gamma = 2}$ $\boxed{y=1}$

$\boxed{x=-1}$

Case 3:-

$f_1(x,y)=0$ $f_2(x,y) \neq 0$

$\therefore$ $x-y+2<0$

$\lambda_2 = 0$

$\mathcal{L} = x^2 + y^2 + \lambda(x + y - 1)$.

$2x + \lambda = 0$

$2y + \lambda = 0$

$x + y = 1$

$2x - 2y = 0$

$x = y$

$\boxed{x = y = \frac{1}{2}}$

$\lambda = -2x$

$\boxed{\lambda = -1}$

case 4:- $f_1 \neq 0$ $f_2 \neq 0$

$x + y - 1 < 0$ $x - y + 2 < 0$

$\lambda_1 = 0$ $\lambda_2 = 0$

No eq cons

$L = x^2 + y^2$

$\frac{\delta L}{\delta x} = 2x = 0$ $\boxed{x = 0}$

$\frac{\delta L}{\delta y} = 2y = 0$ $\boxed{y = 0}$