FAQs on Convex Optimization

1. What is a convex programming problem?

   A convex programming problem is the minimization of a convex function on a convex set, i.e.

$$\min f(x)$$
$$X \in C$$

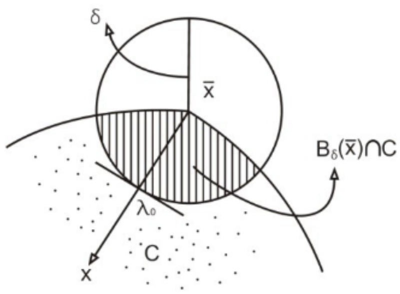
where f: $R^n \to R$ and C R^n. f is a convex function and c a convex set. Usually C is described as follows

$$C = \{ x: \; g_i \; (x) \leq 0, i=1.......m, \; h_j \; (\hat{})=0, j=1.....m\}$$

where $g_i's$ are convex function and $hj's$ are affine function.

2. What is the importance of convex optimization problems?

   The major importance of convex programming or convex optimization arises from the fact that every local minimum is a global minimum.



Let us consider minimizing f: $R^n \to R$ or C R^n where f is a convex function and C is a convex set.

Let $\acute{x}$ be a local minimum of f on C. thus ∃δ>0 such that ∀ z ∈ ($\acute{x}$)∩C, f(z)≥f($\acute{x}$). Let X ∈ C(take it outside B_δ ($\acute{x}$)∩C). Join x & B_δ ($\acute{x}$)∩C)using a line segment. Let

$$Z_\lambda = \lambda x + (1- \lambda) \; \acute{x}$$

Thus ∃ $\lambda_0 \in$ (o,1) such that ∀∈ (o, λ_0 , $¿ Z_\lambda \in B_\delta$ ($\acute{x}$)∩C

Thus for λ ∈ (o, $\lambda_0 ¿$

$$f(z\lambda) \geq (\; \acute{x} \;)$$

$$f(\lambda x + (1-\lambda)\acute{x}) \geq f(\acute{x})$$

By convexity of

$$\lambda f(x) + (1-\lambda) f() \geq f(\acute{x})$$

$$\Rightarrow \quad \lambda(f(x) - f(\acute{x})) \geq o$$

$$\Rightarrow \quad (f(x) - f(\acute{x})) \geq o, \text{ as } \lambda > 0$$

Since x is arbitrary we have as $\acute{x}$ the global minimum.

3. What can we tell about the continuity and differentiality of a convex function?

- If f: $R^n \to R$ is convex then f is continuous and even locally Lipschitz, i.e; for any x $\in R^n$ and K≥ 0 such that for all y,z $\in B_\delta$ (x) we have

  $$|f(y) - f(z)| \leq K\| y-z\|,$$

- If f: C → R is convex and C is a closed convex set then, f is continuous on the interior of C.

- If f: $R^n \to R$ is convex, then it is differentiable almost everywhere, i.e.; the set of points in R^n at which f is not differentiable forms a set of measure zero.

- A differentiable function f: $R^n \to R$ is convex if and only if; for all x, y in R^n.

  $$f(y) - f(x) \geq < \langle \nabla f(x), y-x \rangle$$

  Thus if (x $\in R^n$) be such that $\nabla f = 0,$ then f has a global minimizer at x.

4. If f: $R^n \to R$ is differentiable then can we detect it.

If f is twice continuously differentiable then there is at least a theoretical way to detect it.

- A function f is convex if and only if the Hersian matrix ∇^2 f(x) is positive for all x $\in R^n$ semi-definitely.
- If ∇^2 f(x) is positive definite for all x $\in R^n$, then f is strictly converse. The converse need not be true.
  Example : f(x) = X^4, X $\in R$
- If f is strongly convex then ∇^2 f(x) is always positive definite.

Let f be a p-strongly convex function. since f is twice continuously differentiable, it is differentiable and hence

$$f(y) - f(x) \quad \langle \nabla f(x), y-x \rangle + p \quad \|y-x\|^2, \; P>0$$

Now by Taylor's theorem for any $\lambda > 0,$ & w $\in R^n$

$$f(x+\lambda w) = f(x) + \lambda \langle \nabla f(x), w \rangle + \frac{1}{2} \lambda^2 \langle w, \nabla^2 f(x)\, w \rangle + 0(\lambda^2)$$

Now by strong convexity

$$\frac{1}{2} \lambda^2 \langle w, \nabla^2 f(x)\, w \rangle + 0(\lambda^2) \geq P\lambda^2 \|w\|^2$$

$$\Rightarrow \quad \frac{1}{2} \langle w, \nabla^2 f(x)\, w \rangle + 0 \frac{(\lambda^2)}{\lambda^2} \geq P \|w\|^2$$

Now as $\lambda \downarrow 0$ (i.e; $\lambda \to 0$) we have

$$\frac{1}{2} \langle w, \nabla^2 f(x)\, w \rangle \geq P \|w\|^2$$

i.e; $\langle w, \nabla^2 f(x)\, w \rangle \geq 2P \|w\|^2$

Thus $\nabla^2 f(x)$ is positive definite.

5. What are the major classes of convex optimization problems?

a) Linear Programming problem
b) Conic Programming problem
c) Semi-definite Programming
d) Quadratic convex programming under linear constraints
e) Quadratic convex programming under quadratic constraint

- <u>Linear Programming</u> : min < ax >
  subject to
  Ax = b
  x ≥ o

  where C $\in R^n$, A is a m $\times$ n matrix, b $\in R^m$, & x≥ 0 $\Rightarrow$ x $\in R^n$

  This is called linear programming in the standard form.
  Important feature: If a lower bound exists a minimizer exists.

- <u>Conic Programming</u> :

  min < ax >

subject to

$Ax = b$

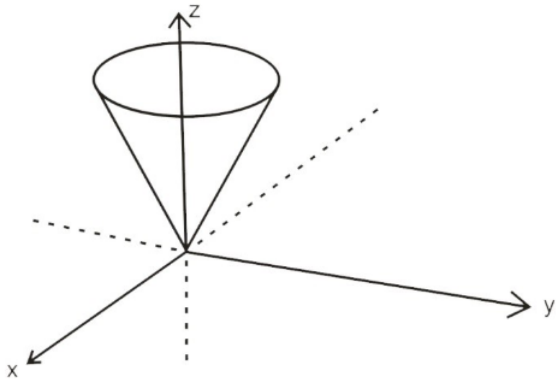
$x \in K$

where K is a pointed convex cone. The cone is called pointed if K ∩ (-K) = {0}

K for example could be the ice-cream cone or Lorenz-cone.

K= { x $\in R^n$ } : $\sqrt{x_1^2+x_2^2+\ldots\ldots\ldots+x_{n-1}^2}$ .... $\leq$ $x_n; x_n \geq 0$ case the above conic problem is called the second-order conic programming problem (SOCP for short).

- Lorenz cone:



Lorenz cone is not a polyhedral cone.

- Semi- definite Programming :

S^n : set of nXn systematic matrices

S_+^n : set of nXn, systematic and positive semidefinite matrices

S^n_{++} : { X $\nabla \in$ S^n_+ : X is positive definite}

S_+^n is a convex cone but not polyhedral

Inner product in S^n : $\langle X, Y \rangle$ trace (X,Y)

min $\langle C, X \rangle$

$\langle A_i, X \rangle = b_i$

$X \in S_+^n$

- Semi definite programming or SDD for short is not a linear programming problem in matrices.

Quadratic convex programming with linear constraints.

$$\min \frac{1}{2}<x,Qx \text{ ¿ } + \text{ ¿}c,x>\text{¿} + d$$

subject to $Ax = b$

$x \geq 0$

$Q \in S^n_+$, $c \in R^n$, $d \in R$, A is a $m \times n$ matrix and $x \geq 0 \Leftrightarrow x \in R_{¿}^{+¿^n}$

<u>Important fact :</u> If a lower bound exists, then a minimizer exists. This is the celebrated Frank-Wolfe theorem.

- Quadratic convex programming with linear constraints

$$\min \frac{1}{2} \text{ ¿}x,Q_0 \; x>\text{¿} + \text{ ¿}C_0, \; x>\text{¿} + d_0$$

subject to

$$\frac{1}{2} \text{ ¿}x,Q_i \; x>\text{¿} + \text{ ¿ } C_i, \; x>\text{¿} + d_i \leq 0$$

i=1,............,m

where $Q_0, Q_1, \ldots, Q_m$ are positive semi-definite matrices, C_0, C_1, C_m are vectors in R^n and $d_0, d_1 \ldots d_m$ are elements in R.

6. What are saddle point conditions?

Consider the convex optimization problem (CP)

min f(x)

subject to

$g_i(x) \leq 0$, i-1,2,........m

Construct the Lagrangian as follows

$$L(x, \lambda) = f(x) + \lambda_1 g_1(x) + \lambda_2 g_2(x) + \ldots\ldots\ldots + \lambda_m g_m(x)$$

where $\lambda = \begin{pmatrix} \lambda 1 \\ \lambda m \end{pmatrix} \in R_{¿}^{+¿^m}$ i.e; $\lambda_i \geq 0$, for all i=1,.......m

A vector is $(\acute{x}, \acute{\lambda}) \in R^n \times R_{¿}^{+¿^m}$ is called a saddle point if

$L(\acute{x}, \lambda) \leq (\acute{x}, \acute{\lambda}) \leq L(x, \acute{\lambda})$, for all $x \in R^n$, and $\lambda \in R_{¿}^{+¿^m}$

If solves convex optimization problem and slater condition holds, i.e; there exists $\hat{x} \in R^n$ s.t.

$¿ g_i(\hat{x})$ $¿ 0.$ $\forall$ i=1,......,m then there exists $\acute{\lambda} \in R_¿^{+¿^m}$ s.t.

i) L ($\acute{x}, \lambda$) ≤ ($\acute{x}, \acute{\lambda}$) ≤ L (x, $\acute{\lambda}$), for all x $\in R^n$, and λ $\in R_¿^{+¿^m}$

ii) $\acute{\lambda} g_i$ ($\acute{x}$) = 0, i=1,.......,m

If there exists a pair of ($\acute{x},$ $\acute{\lambda}$) $\in R^n \times R_¿^{+¿^m}$ such that i) & ii) hold then $\acute{x}$ solves (CP).