# Optimal conditions for mathematical optimization problem

Eladio OCAÑA [1]     Pedro GAJARDO [2]

January 2013

[1] IMCA-FC, Universidad Nacional de Ingeniería, Lima, Perú
[2] Universidad Técnica Federico Santa María, Valparaíso, Chile

# Contents:

1. A brief introduction of convex set and convex function,
2. Mathematical formulation of optimization problem,
3. Optimality conditions.

**Definition (Convex set)**

*Given a subset $C$ of $\mathbb{R}^n$, it is said to be* **convex** *if for all $x, y \in C$ and all $\lambda \in [0,1] \subset \mathbb{R}$ one has*

$$\lambda x + (1-\lambda) y \in C.$$

Geometrically it says that all segments with endpoints in $C$ are contained in $C$.

## Example (Examples of convex sets)

- *The whole space* $\mathbb{R}^n$*;*
- *The intervals* $[a, b], [a, b), \cdots$ *of* $\mathbb{R}$*;*
- *The balls open and closed of* $\mathbb{R}^n$*;*
- *For any matrix* $A \in \mathbb{R}^{p\times n}$ *and any vector* $a \in \mathbb{R}^p$*,*

$$C = \{x \in \mathbb{R}^n : Ax \leq a\};$$

- *etc*

## Definition (Convex function)

*A function* $f : \mathbb{R}^n \to \mathbb{R}$ *is said to be* **convex** *if*

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y) \;\; \forall\, x, y \in \mathbb{R}^n, \; \forall\, \lambda \in [0,1].$$

## Example (Examples of convex functions)

- *Given* $a \in \mathbb{R}^n$ *and* $\alpha \in \mathbb{R}$*, then* $f(x) = \langle a, x\rangle + \alpha$ *is convex;*
- *Given* $A \in \mathbb{R}^{n\times n}$ *symmetric,* $a \in \mathbb{R}^n$ *and* $\alpha \in \mathbb{R}$*, then*

  $$f(x) := \langle Ax, x\rangle + \langle a, x\rangle + \alpha$$

  *is convex* **if and only if** *A is positive semi-definite (psd);*
- *etc.*

# Some properties of convex functions

1. Let $f, g : \mathbb{R}^n \to \mathbb{R}$ convex and $\lambda \geq 0$, then $f + \lambda g$ is convex;
2. Let $f : \mathbb{R}^n \to \mathbb{R}$ convex and $k : \mathbb{R} \to \mathbb{R}$ increasing and convex, then the composition $k \circ f : \mathbb{R}^n \to \mathbb{R}$ is convex;
3. Let $\{f_i\}_{i \in I}$ be a family of convex functions defined on $\mathbb{R}^n$, then the function

$$f(x) := \max_{i \in I} f_i(x) \ \text{ for all } \ x \in \mathbb{R}^n$$

is convex;
4. Let $\varphi : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ convex, then the marginal function

$$h(x) := \min_{y \in \mathbb{R}^m} \varphi(x, y) \ \text{ for all } x \in \mathbb{R}^n$$

is convex.

# Convexity of functions in terms of its derivative

Let $f : \mathbb{R}^n \to \mathbb{R}$ be a differentiable function. Then,

1. $f$ is convex if and only if

$$f(x) + \langle \nabla f(x), y - x \rangle \leq f(y) \text{ for all } x, y \in \mathbb{R}^n;$$

2. $f$ is convex if and only if

$$\langle \nabla f(x) - \nabla f(y), x - y \rangle \geq 0 \text{ for all } x, y \in \mathbb{R}^n;$$

3. $f$ is convex if and only if the Hessian matrix $\nabla^2 f(x)$ is psd for all $x \in \mathbb{R}^n$.

# The optimization problem

The general formulation of a mathematical minimization problem is as follows:

$$m = \min_{x} [\, f(x) \; : \; x \in C \,] \qquad (P)$$

where $f : \mathbb{R}^n \to \mathbb{R}$ is the **objective function**, $\emptyset \neq C \subset \mathbb{R}^n$ is the **admissible set** and $m$ is the **optimal value.**
Frequently, $C$ is of the form

$$C = \{x \in \mathbb{R}^n \; : \; g_i(x) \leq 0, \, i = 1, \cdots, p \, ; \;\; h_j(x) = 0, \, j = 1, \cdots, q\}$$

for some functions $g_i, h_j : \mathbb{R}^n \to \mathbb{R}$.

# Global and local minimum

## Definition

*Given a point $\bar{x} \in C$, it is a*

- ***global minimum*** *of $(P)$ if $f(\bar{x}) \leq f(x)$ for all $x \in C$,*
- ***local minimum*** *of $(P)$ if exists a neighborhood $V$ of $\bar{x}$ such that*

$$f(\bar{x}) \leq f(x) \quad \text{for all } x \in C \cap V.$$

**Remark.** Global minimum implies local minimum.

# Optimality conditions: The unconstrained case

Here $C = \mathbb{R}^n$ and so the problem $(P)$ is

$$m = \min_x [\, f(x) \;:\; x \in \mathbb{R}^n \,].$$

## Proposition (First order optimality condition)

1. *If $\bar{x}$ is a local minimum of $(P) \Longrightarrow \nabla f(\bar{x}) = 0$;*
2. *If $\nabla f(\bar{x}) = 0$ and $f$ convex $\Longrightarrow \bar{x}$ is a global minimum of $(P)$.*

# Optimality conditions: The unconstrained case

Proposition (Second order optimality condition)

1. *If* $\bar{x}$ *is a local minimum of* $(P) \implies \nabla f(\bar{x}) = 0$ *and* $\nabla^2 f(\bar{x})$ *is psd;*
2. *If* $\nabla f(\bar{x}) = 0$ *and* $\nabla^2 f(\bar{x})$ *is pd* $\implies \bar{x}$ *is a local minimum of* $(P)$.

**Remark.** The converse are not true: consider the real functions defined on $\mathbb{R}$,

- $f(x) = x^3$ in the first one, and
- $f(x) = x^4$ in the second one.

# Optimality conditions: Convexity on the admissible set

Let $\bar{x} \in C$ and assume $C$ convex.

**Proposition (First order optimality condition)**

1. *If $\bar{x}$ is a local minimum of $(P)$,*

$$\implies \langle \nabla f(\bar{x}), x - \bar{x} \rangle \geq 0 \ \ \forall x \in C.$$

2. *If $\langle \nabla f(\bar{x}), x - \bar{x} \rangle \geq 0 \ \ \forall x \in C$ and $f$ convex,*

$$\implies \bar{x} \text{ is a global minimum of } (P).$$

# Optimality conditions: ‘The general case"

Let us consider $C$ of the form

$$C = \{x \in \mathbb{R}^n \,:\, g_i(x) \leq 0,\ i = 1, \cdots, p\}$$

where $g_i : \mathbb{R}^n \to \mathbb{R}$ are arbitrary functions.
In connection with $C$ and $\bar{x} \in C$, let us denote

$$I(\bar{x}) = \{i = 1, 2, \cdots, p \,:\, g_i(x) = 0\}$$

called (eventually empty) the set of all active indices of the constrains.
In order to extend the previous conditions (necessary and sufficient conditions) to this general case, let us introduce the following set

$$L(\bar{x}, C) = \{d \in \mathbb{R}^n : \langle \nabla g_i(\bar{x}), d\rangle \leq 0 \ \ \forall\, i \in I(\bar{x})\,\}$$

called the **linearized tangent cone** of $C$ at $\bar{x}$.

# Optimality conditions: “The general case"

Under some “regularity condition" on $C$, for example: i) each $g_i$ affine, ii) slater's condition, iii) Mangasarian-Fromowitz condition, $\cdots\cdots\cdots$

we have the following result:

## Theorem (First order optimality condition)

*Let $\bar{x}$ be a point of $C$ .*

1. *If $\bar{x}$ is a local minimum of $(P)$*

$$\Longrightarrow \langle \nabla f(\bar{x}), d\rangle \geq 0 \ \ \forall\, d \in L(\bar{x}, C). \tag{1}$$

2. *If $\langle \nabla f(\bar{x}), d\rangle \geq 0 \; \forall\, d \in L(\bar{x}, C)$ with $C$ convex and $f$ convex*

$$\Longrightarrow \bar{x} \text{ is a global minimum of } (P).$$

# KKT optimality condition: “The general case"

The optimality condition (1) can be written as

$$\langle \nabla g_i(\bar{x}), d\rangle \leq 0 \;\; \forall\, i \in I(\bar{x}) \quad \Longrightarrow \quad \langle \nabla f(\bar{x}), d\rangle \geq 0. \qquad (2)$$

The following elementary property allow us to rewrite (2) as an alternative way.

## Theorem (Farkas' Lemma)

*Let $A \in \mathbb{R}^{m\times n}$ and $c \in \mathbb{R}^n$. The two following conditions are equivalents:*

1. $Ad \leq 0 \implies \langle c, d\rangle \geq 0.$
2. $\exists\, u \in \mathbb{R}^m_+$ *such that* $c + A^t u = 0.$

# KKT optimality condition: "The general case"

By denoting:

- matrix $A$ whose rows are the $\nabla g_i(\bar{x})$ for $i \in I(\bar{x})$, and
- $c = \nabla f(\bar{x})$.

Then, from Farkas' lemma, condition (2) is equivalent to the existence of $\bar{u} \in \mathbb{R}^m_+$ (with $m = \operatorname{card}(I(\bar{x}))$) such that

$$\nabla f(\bar{x}) + \sum_{i \in I(\bar{x})} \bar{u}_i \nabla g_i(\bar{x}) = 0.$$

# KKT optimality condition: "The general case"

Now, because $I(\bar{x}) = \{i = 1, 2, \cdots, p \,:\, g_i(\bar{x}) = 0\}$, the previous condition is equivalent to the existence of $\bar{u} \in \mathbb{R}^p_+$ such that

$$\nabla f(\bar{x}) + \sum_{i=1}^{p} \bar{u}_i \nabla g_i(\bar{x}) = 0,$$

$$g_i(\bar{x}) \leq 0, \quad \bar{u}_i g_i(\bar{x}) = 0 \quad \forall\, i = 1, \cdots, p.$$

# KKT optimality condition: “The general case"

In resume we have the well known KKT optimality condition

Theorem (Karush-Kuhn-Tucker condition)

*If* $\bar{x} \in C$ *is a local minimum of* $(P)$, *then there exist* $\bar{u} \in \mathbb{R}^p_+$ *such that*

$$\nabla f(\bar{x}) + \sum_{i=1}^{p} \bar{u}_i \nabla g_i(\bar{x}) = 0,$$

$$g_i(\bar{x}) \le 0, \quad \bar{u}_i g_i(\bar{x}) = 0 \quad \forall\, i = 1, \cdots, p.$$

**Remark.** The scalars $\bar{u}_i$ are called **KKT multipliers.**

# KKT optimality condition: “The general case"

In the convex case the KKT-condition is also **sufficient.**

## Theorem

*Assume $f$, $g_i$ are convex for all $i = 1, \cdots, p$. If there exist $\bar{u} \in \mathbb{R}^p_+$ such that*

$$\nabla f(\bar{x}) + \sum_{i=1}^{p} \bar{u}_i \nabla g_i(\bar{x}) = 0,$$

$$g_i(\bar{x}) \leq 0, \quad \bar{u}_i g_i(\bar{x}) = 0 \quad \forall\, i = 1, \cdots, p.$$

*Then $\bar{x}$ is a global solution of $(P)$.*