Properties of systems of inequalities and their influence in convex optimization

Marco A. López

University of Alicante

One World Optimization Seminar (17-05-2021)

Semi-infinite programming (SIP) problem

- It is an optimization problem with *finitely many variables* $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ and a feasible set, $\mathbb{F}$, described by *infinitely many constraints*:

$$(\mathcal{P}): \qquad \text{Min } g(x) \qquad \text{s.t.} \quad f_t(x) \leq 0 \quad t \in T,$$

where T is an infinite *index set.*
An important **extension** is the *generalized SIP* (GSIP), for which the index set $T = T(x)$ depends on x, i.e.

$$(\mathcal{P}): \qquad \text{Min } g(x) \qquad \text{s.t.} \quad f_t(x) \leq 0 \quad t \in T(x).$$

- In the last decades years Semi-infinite Programming has known a tremendous development. More than **1300 articles** and **10 books** on theory, numerical methods and applications of SIP.

Convex SIP problem

Today we are dealing with the *convex optimization problem*

$$(\mathcal{P}) \quad \text{Min} \quad g(x) \quad \text{s.t.} \quad f_t(x) \leq 0,\ t \in T,\ x \in \mathbb{C},$$

where $\mathbb{C}$ is a (non-empty) closed convex set in $\mathbb{R}^n$, and $g, f_t : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\},\ t \in T,$ are proper lower semicontinuous (lsc) convex functions .

We say that the *constraint system*

$$\tau := \{f_t(x) \leq 0,\ t \in T;\ x \in \mathbb{C}\}, \tag{1}$$

is *consistent* when $\mathbb{F} \neq \emptyset$. By v and $\mathbb{S}$ we represent the *optimal value* and the *optimal set* of problem $(\mathcal{P})$, respectively.

In *linear SIP* , the explicit constraints are *affine* and $\mathbb{C} = \mathbb{R}^n$, i.e.

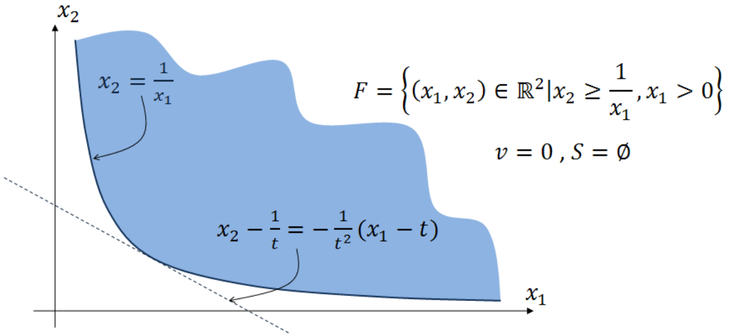
$$\tau = \{\langle a_t, x\rangle \leq b_t,\ t \in T\}. \tag{2}$$

LSIP is very different to ordinary LP!

Example

Let us consider the LSIP problem, in $\mathbb{R}^2$, with $T := \,]0, +\infty[$,

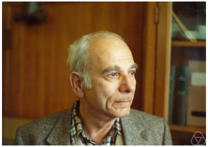
$$(\mathcal{P}): \qquad \text{Min}_{x_1, x_2} \quad x_1$$
$$s.t. \quad -t^{-2}x_1 - x_2 \leq -2t^{-1}, \quad t \in \,]0, +\infty[$$



Some drops of history...

- Prehistory of SIP is related to Chebyshev approximation, the classical work of *Alfréd Haar* on linear semi-infinite systems [Ha'24], and the *Fritz John* optimality condition [Jo'48].

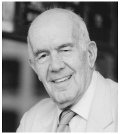
Haar



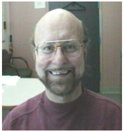
Fritz John

- The term SIP was coined in 1962 by *Charnes, Cooper* and *Kortanek* [ChCoKo'62].

Charnes



Cooper



Kortanek

But ..., are there real-world optimization problems with **infinitely many constraints**?

- *George Dantzig* (1991): "(...) One of the results, published jointly with Abraham Wald, was on the Neyman-Pearson Lemma. In today's terminology, this part of my thesis was on the existence of Lagrange multipliers (or dual variables) for a semi-infinite linear program (...)".
- Other significant applications in statistics: *optimal experimental design* in regression, constrained multinomial *maximum-likelihood estimation*, *robustness* in Bayesian statistics, *actuarial risk theory*, etc.
- In the review papers HeKo'93 and Po'87, as well as in Go-ML'98, the reader will find many applications of SIP in different fields such as *functional approximation, robotics, pde's, engineering design, optimal control,* transportation problems, fuzzy sets, cooperative games, robust optimization, etc.
- Reputed optimization books devoted some chapters to SIP *e.g., Krabs* [Kr'79], *Anderson-Nash* [AnNa'87], *Guddat* et al. [Gu'90], *Bonnans-Shapiro* [BoSh'00], and *Polak* [Po'97].

Polynomial approximation

- We want to *approximate* the function $f \in \mathcal{C}([a,b])$ by a *polynomial* of degree at most $n-1$

$$p(x,t) := x_0 + x_1 t + ... + x_{n-1} t^{n-1} \text{ with } t \in [a,b].$$

- If we use the ∞-norm (Chebyshev-norm), minimizing the approximation error $z := \|f - p(x,\cdot)\|_\infty$ is a problem which is equivalently expressed as the linear SIP:

$$\operatorname*{Min}_{x,z} \; z \text{ s.t. } \begin{cases} f(t) - x_0 - x_1 t - ... - x_{n-1} t^{n-1} - z \leq 0 & t \in [a,b], \\ -(f(t) - x_0 - x_1 t - ... - x_{n-1} t^{n-1}) - z \leq 0 & t \in [a,b]. \end{cases}$$

Theorem (Chebyshev, 1874)

A polynomial $p(x^{opt}, t)$ is a best approximation of f on $[a,b]$ if and only if there exist $n+1$ points $a \leq t_1 < ... < t_{n+1} \leq b$ and a number $\sigma \in \{-1, 1\}$ such that

$$(-1)^i \sigma \{f(t_i) - p(x^{opt}, t_i)\} = \|f - p(x^{opt}, \cdot)\|_\infty = z^{opt}.$$

The Minimum-Volume Circumscribed Ellipsoid Problem

- This is the problem of **finding a minimum-volume ellipsoid covering a convex body** $K \subset \mathbb{R}^n$.
- Fritz John [FJ'47] proved that such an ellipsoid, denoted by E^K, *exists* and is *unique*.He was a pioneer in *semi-infinite programming* .
  John F. , Extremum problems with inequalities as subsidiary conditions. In: Studies and Essays, Courant Anniversary Volume , New York: Interscience (1948), 187-204.
- In some *convex programming algorithms*, including the *ellipsoid method*, the exact or nearly exact ellipsoid needs to be computed. For K sufficientely simple, E^K can be obtained analytically. In more general cases, *interior-point-type algorithms* are used to approximate E^K.
- Some historic works dealing with this topic via SIP are:
  Juhnke F., Embedded maximal ellipsoids and semi-infinite optimization. Beiträge Algebra Geom. 35 (1994), 163–171.
  Juhnke F., Polarity of embedded and circumscribed ellipsoids. Beiträge Algebra Geom. 36 (1995), 17–24.

- If $X \in \mathcal{P}_n$, the convex (solid) cone of all the $n \times n$ symmetric positive definite matrices, and $c \in \mathbb{R}^n$, the set

$$E(X,c) := \{x \in \mathbb{R}^n : \ (x-c)^\top X(x-c) \leq 1\}$$

is an *ellipsoid with center* c, whose *volume* is given by

$$\mathrm{vol}\, E(X,c) = (\det X)^{-1/2}\,\mathrm{vol}(\mathbb{B}) = (\det X)^{-1/2}\frac{\pi^{n/2}}{\Gamma((n/2)+1)},$$

where $\mathbb{B}$ is the closed unit ball in $\mathbb{R}^n$.

- The *associated SIP* problem is

$$\min_{X,c} \mathrm{vol}\, E(X,c) \quad \text{s.t. } (z-c)^\top X(z-c) \leq 1 \ \ \forall z \in K,$$

or equivalently (taking logarithms)

$$\min_{X,c} \ -\lg(\det X) \quad \text{s.t. } (z-c)^\top X(z-c) \leq 1 \ \ \forall z \in K.$$

- The function

$$X \in \mathcal{P}_n \mapsto -\lg(\det X), \qquad (3)$$

is *strictly convex* and *differentiable* on $\mathcal{P}_n$, with $\nabla(-\lg(\det X)) = -X^{-1}$.

- Moreover, $-\lg(\det X) \to +\infty$ when X approaches a matrix $Z \in \operatorname{bd} \mathcal{P}_n$ since $\det Z = 0$, entailing that the function (3) is *coercive* on $\mathcal{P}_n$, and the problem

$$(\mathcal{P}): \quad \min_{\substack{X \in \mathcal{P}_n \\ c \in \mathbb{R}^n}} \; -\lg(\det X) \quad \text{s.t. } (z-c)^\top X (z-c) \leq 1 \;\; \forall z \in K,$$

has *a unique optimal solution.*.

- If σ_A denotes the *support function* of A (i.e., $\sigma_A(u) := \sup_{x \in A} \langle u, x \rangle$),

$$K \subset E(X,c) \iff \sigma_K(u) \leq \sigma_{E(X,c)}(u) = \langle u, c \rangle + \sqrt{u^\top X^{-1} u}, \;\; \forall u \in \mathbb{R}^n.$$

Since the support function is positively homogeneous, and $-\lg(\det X) = \lg(\det X^{-1})$, we get the SIP reformulation

$$(\mathcal{P}): \quad \min_{\substack{X \in \mathcal{P}_n \\ c \in \mathbb{R}^n}} \; \lg(\det X^{-1}) \quad \text{s.t. } \sigma_K(u) \leq \langle u, c \rangle + \sqrt{u^\top X^{-1} u}, \;\; \forall u \in \operatorname{bd} \mathbb{B}.$$

Theorem 1 in Juhnke'95 caracterizes the unique optimal solution of $(\mathcal{P})$.

Today, our aims are:

1. To study some properties of the **infinite convex systems**, as the *Farkas-Minkowski property* (FM, in brief) and the *local Farkas-Minkowski property* (LFM, in short), which give rise to weak *CQ's* in SIP.
2. To provide *optimality conditions* by appealing to the properties of the *supremum function* of an infinite family of convex functions and the characterizations of its subdifferential.
3. To show how useful is *convex analysis* in the mathematical developments...

- 927 publications according in *MathSciNet* for "Review Text=(optimality conditions).
- 69 publications according in *MathSciNet* for "Review Text=(optimality conditions AND semi-infinite programming

Summary

- 1st step in formulating optimality conditions
- Notation and basic tools
- KKT'1 optimality conditions - FM property
- KKT'2 optimality conditions - LFM property
- KKT'3 asymptotic optimality conditions
- KKT'4 conditions for SIP under compacity/(upper)continuity
- Bibliographical comments and references

1st step in formulating optimality conditions

1. If $\bar{x}$ is *optimal* for $(\mathcal{P})$, then it is optimal for the unconstrained problem

$$\underset{x \in \mathbb{R}^n}{\mathrm{Min}} \;\; \varphi(x) := \sup\{g(x) - g(\bar{x});\; f_t(x),\; t \in T;\; \mathrm{I}_{\mathbb{C}}(x)\},$$

where $\mathrm{I}_{\mathbb{C}}$ is the *indicator function* of $\mathbb{C}$. Obviously, for $x \in \mathbb{F}$ one has $g(x) - g(\bar{x}) \geq 0$, and so $\varphi(x) \geq \varphi(\bar{x}) = 0$. Hence,

$$\bar{x} \text{ is optimal for } (\mathcal{P}) \implies 0_n \in \partial\varphi(\bar{x}).$$

We need: a) To express $\partial\varphi(z)$ in terms of the *approximate/exact subdifferentials* of g and the f_t's. (φ is a supremum function!)

2. A second obvious fact is

$$\bar{x} \text{ is optimal for } (\mathcal{P}) \iff 0_n \in \partial(g + \mathrm{I}_{\mathbb{F}})(\bar{x}),$$

and

$$\partial(g + \mathrm{I}_{\mathbb{F}})(\bar{x}) \supset \partial g(\bar{x}) + \partial \mathrm{I}_{\mathbb{F}}(\bar{x}) = \partial g(\bar{x}) + \mathrm{N}_{\mathbb{F}}(\bar{x}),$$

where $\mathrm{N}_{\mathbb{F}}(\bar{x})$ is the normal cone to $\mathbb{F}$ at $\bar{x}$.
We need: b) To provide conditions (*CQs*) ensuring that the inclusion "$\supset$" become an equality.
c) To express $\mathrm{N}_{\mathbb{F}}(\bar{x})$ in terms of the functions $f_t,\; t \in T$, and $\mathrm{I}_{\mathbb{C}}$.

Notations and basic tools

- **Given** $A, B \subset \mathbb{R}^p$, we consider the *Minkowski sum*: $A + B := \{a + b \mid a \in A,\ b \in B\}$, $A + \emptyset = \emptyset + A = \emptyset$.
- $\mathrm{co}\, A$ is the *convex hull* of A, $\mathrm{cone}\, A$ is the *convex cone generated by A*.
- $\mathrm{int}\, A$ is the *interior* of A, $\mathrm{cl}\, A$ and $\overline{A}$ denote indistinctly the *closure* of A; $\mathrm{rint}\, A$ is the topological *relative interior* of A.
- **Given** $h : \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$, $\mathrm{dom}\, h$ and $\mathrm{epi}\, h$ represent its *(effective) domain* and *epigraph* , respectively.
- h is *proper* if $\mathrm{dom}\, h \neq \emptyset$ and $h(x) > -\infty\ \forall x \in \mathbb{R}^n$; it is *convex* if $\mathrm{epi}\, h$ is convex, and $h \in \Gamma_0(\mathbb{R}^n)$ if it is proper, lower semicontinuous and convex.
- $\overline{\mathrm{co}} h$ represents the *lsc convex hull* of h; i.e., $\mathrm{epi}(\overline{\mathrm{co}} h) = \overline{\mathrm{co}}\,(\mathrm{epi}\, h)$.
- The *$\varepsilon-$subdifferential* of h at $x \in h^{-1}(\mathbb{R})$, $\varepsilon \geq 0$, is the closed convex set

$$\partial_\varepsilon h(x) := \{u \in \mathbb{R}^n \mid h(y) - h(x) \geq \langle u, y - x\rangle - \varepsilon,\ \forall y \in \mathbb{R}^n\}.$$

The *Fenchel conjugate* of h is the lsc convex function $h^* : \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ given by

$$h^*(u) := \sup\{\langle u, x\rangle - h(x) \mid x \in \mathbb{R}^n\}.$$

We have $h^* = (\operatorname{cl} h)^* = (\overline{\operatorname{co}} h)^*$. Moreover, the *Fenchel-Young inequality* establishes

$$u \in \partial h(x) \Leftrightarrow h(x) + h^*(u) \leq \langle u, x\rangle \Leftrightarrow h(x) + h^*(u) = \langle u, x\rangle .$$

The *support* and the *indicator* functions of $A \neq \emptyset$ are respectively

$$\begin{aligned} \sigma_A(u) &: = \sup\{\langle u, a\rangle \mid a \in A\}, \text{ for } u \in \mathbb{R}^n, \\ \mathrm{I}_A(x) &: = \begin{cases} 0, & \text{if } x \in A, \\ +\infty, & \text{if } x \in \mathbb{R}^n \setminus A. \end{cases} \end{aligned}$$

σ_A is sublinear, lsc, and satisfies $\sigma_A = \sigma_{\overline{\operatorname{co}} A} = \mathrm{I}^*_{\overline{\operatorname{co}} A}$. Therefore, $\operatorname{epi} \sigma_A$ is a closed convex cone.

For every family of functions f_i, $i \in I$, (I arbitrary), we have

$$(\inf_{i\in I} f_i)^* = \sup_{i\in I} f_i^*. \tag{4}$$

If $\{f_i, i \in I\} \subset \Gamma_0(\mathbb{R}^n)$ and $\sup_{i\in I} f_i$ is proper, then

$$(\sup_{i\in I} f_i)^* = \overline{\mathrm{co}}(\inf_{i\in I} f_i^*). \tag{5}$$

For $f, g \in \Gamma_0(X)$ such that $\mathrm{dom}\, f \cap \mathrm{dom}\, g \neq \emptyset$, it is well known that

$$(f \Box g)^* = f^* + g^*,\ (f+g)^* = \mathrm{cl}(f^* \Box g^*). \tag{6}$$

Clearly, (6) and (4) imply that

$$\mathrm{epi}(f+g)^* = \mathrm{cl}(\mathrm{epi}\, f^* + \mathrm{epi}\, g^*), \tag{7}$$

and

$$\mathrm{epi}(\sup_{i\in I} f_i)^* = \overline{\mathrm{co}}(\cup_{i\in I}\, \mathrm{epi}\, f_i^*).$$

The cl in the first equation is superfluous if one of f and g is *continuous* at some point of $\mathrm{dom}\, f \cap \mathrm{dom}\, g$ (then, $\mathrm{epi}\, f^* + \mathrm{epi}\, g^*$ is closed).

Subdifferential calculus rules for the sum

- First results for the sum:
  1) Suppose that one of the following conditions holds:
     a) $\mathrm{rint}(\mathrm{dom}\, f) \cap \mathrm{rint}(\mathrm{dom}\, g) \neq \emptyset$,
     b) f is continuous at some point of $\mathrm{dom}\, g$.

  Then

$$\partial(f+g)(x) = \partial f(x) + \partial g(x).$$

  2) If $f, g \in \Gamma_0(\mathbb{R}^n)$ one has (*Hiriart-Urruty, Phelps' 93*)

$$\partial(f+g)(x) = \bigcap_{\varepsilon>0} \mathrm{cl}(\partial_\varepsilon f(x) + \partial_\varepsilon g(x)).$$

  3) If $f, g \in \Gamma_0(\mathbb{R}^n)$ and $(\mathrm{dom}\, g) \cap \mathrm{rint}(\mathrm{dom}\, f) \neq \emptyset$, then Th.12 in *Correa, Hantoute, ML '16* yields

$$\partial(f+g)(x) = \bigcap_{\varepsilon>0} \mathrm{cl}(\partial f(x) + \partial g_\varepsilon(x)).$$

KKT'1 optimality conditions - FM property

Definition

We call *characteristic cone* of $\tau = \{f_t(x) \leq 0,\ t \in T;\ x \in \mathbb{C}\}$ to the convex cone

$$\mathbb{K} := \text{cone}\left\{\bigcup_{t \in T} \text{epi}\, f_t^* \cup \text{epi}\, \sigma_{\mathbb{C}}\right\} = \text{cone}\left\{\bigcup_{t \in T} \text{epi}\, f_t^*\right\} + \text{epi}\, \sigma_{\mathbb{C}}. \qquad (8)$$

For the linear system (2),

$$\text{epi}\, f_t^* = (a_t, b_t) + \mathbb{R}_+(0_n, 1),\ t \in T,$$

and

$$\text{epi}\, \sigma_{\mathbb{R}^n} = \mathbb{R}_+(0_n, 1).$$

Hence,

$$\mathbb{K} = \text{cone}\{(a_t, b_t),\ t \in T;\ (0_n, 1)\}. \qquad (9)$$

Lemma

If $\mathbb{F} = \{x \in \mathbb{C} : \ f_t(x) \leq 0, \ t \in T\} \neq \emptyset$, *then*

$$\operatorname{epi} \sigma_{\mathbb{F}} = \operatorname{cl} \mathbb{K} = \operatorname{cl} \operatorname{cone} \left\{ \bigcup_{t \in T} \operatorname{epi} f_t^* \cup \operatorname{epi} \sigma_{\mathbb{C}} \right\}.$$

Proof [sketch] If $h := \sup\{f_t, \ t \in T; \ \mathrm{I}_{\mathbb{C}}\}$, we have

$$x \in \mathbb{F} \Leftrightarrow h(x) \leq 0 \Leftrightarrow h(x) = 0.$$

Then, by (5),

$$\begin{aligned} \operatorname{epi} h^* &= \operatorname{epi} \left(\{\sup \{f_t, \ t \in T; \ \mathrm{I}_{\mathbb{C}}\}\}^* \right) \\ &= \overline{\operatorname{co}} \left\{ \bigcup_{t \in T} \operatorname{epi} f_t^* \cup \operatorname{epi} \sigma_{\mathbb{C}} \right\}, \end{aligned}$$

and

$$\operatorname{epi} \sigma_{\mathbb{F}} \overset{(*)}{=} \overline{\operatorname{cone}}(\operatorname{epi} h^*) = \operatorname{cl} \mathbb{K}.$$

$(*)$ follows from Lemma 3.1(b) in *Jey'03*.

Theorem (generalized Farkas)

Let $\varphi, \psi \in \Gamma_0(\mathbb{R}^n)$. Then $\varphi(x) \leq \psi(x)$ for all $x \in \mathbb{F}$, assumed non-empty, if and only if

$$\operatorname{epi}\varphi^* \subset \operatorname{cl}\left(\operatorname{epi}\psi^* + \mathbb{K}\right). \tag{10}$$

Proof.

$$\begin{aligned} \varphi(x) &\leq \psi(x)\ \forall x \in \mathbb{F} \iff \varphi \leq \psi + \mathrm{I}_{\mathbb{F}} \\ &\Leftrightarrow (\psi + \mathrm{I}_{\mathbb{F}})^* \leq \varphi^* \\ &\Leftrightarrow \operatorname{epi}\varphi^* \subset \operatorname{epi}(\psi + \mathrm{I}_{\mathbb{F}})^*, \end{aligned}$$

but applying (7), the previous lemma, and $\operatorname{cl}(A + B) = \operatorname{cl}(A + \operatorname{cl} B)$:

$$\begin{aligned} \operatorname{epi}(\psi + \mathrm{I}_{\mathbb{F}})^* &= \operatorname{cl}(\operatorname{epi}\psi^* + \operatorname{epi}\sigma_{\mathbb{F}}) \\ &= \operatorname{cl}(\operatorname{epi}\psi^* + \operatorname{cl}\mathbb{K}) = \operatorname{cl}(\operatorname{epi}\psi^* + \mathbb{K}). \end{aligned}$$

□

Corollary

Given $(a, \alpha) \in \mathbb{R}^{n+1}$, the inequality $\langle a, x\rangle \leq \alpha$ holds for all $x \in \mathbb{F}$, assumed non-empty (i.e., $\langle a, x\rangle \leq \alpha$ is a linear consequence of τ), if and only if

$$(a, \alpha) \in \operatorname{cl} \mathbb{K}.$$

Proof.

Apply the generalized Farkas theorem with $\varphi = \langle a, \cdot\rangle - \alpha$ and $\psi \equiv 0$. Then, $\langle a, x\rangle \leq \alpha$ holds for all $x \in \mathbb{F}$ if and only if

$$\begin{aligned} (a, \alpha) + \mathbb{R}_+(0_n, 1) &= \operatorname{epi} \varphi^* \\ &\subset \operatorname{cl}(\operatorname{epi} \psi^* + \mathbb{K}) = \operatorname{cl}(\mathbb{R}_+(\theta, 1) + \mathbb{K}) = \operatorname{cl} \mathbb{K}. \end{aligned}$$

In other words, $\langle a, x\rangle \leq \alpha$ is a consequence of τ if and only if $(a_t, \alpha) \in \operatorname{cl} \mathbb{K}$. □

The following property is crucial in getting optimality conditions for $(\mathcal{P})$.

Definition

We say that the *consistent* system $\tau = \{f_t(x) \le 0,\ t \in T;\ x \in \mathbb{C}\})$ is *Farkas-Minkowski* (*FM*, in brief) if $\mathbb{K}$ is closed.

Theorem

If τ is FM and consistent, then very linear consequence $\langle a, x\rangle \le \alpha$ is a cosequence of τ if it is also consequence of a finite subsystem

$$\tau_S := \{f_t(x) \le 0,\ t \in S;\ x \in \mathbb{C}\},\ \textit{with } S \subset T \textit{ and } |S| < \infty.$$

The converse statement holds if τ is linear.

Theorem

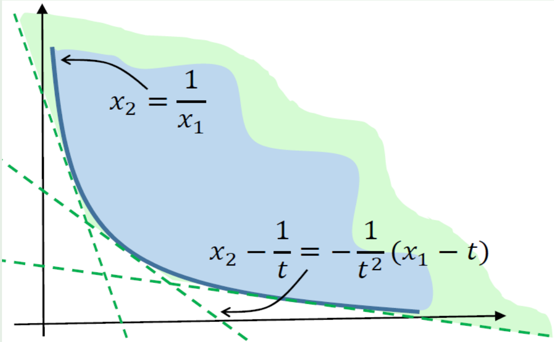
For the linear SIP, if T is compact, the functions $a_{(\cdot)} : T \to \mathbb{R}^n$ and $b_{(\cdot)} : T \to \mathbb{R}$ are continuous, and there exists a point $\widehat{x}$ (Slater point) satisfying $\langle a_t, \widehat{x}\rangle < b_t,\ \forall t \in T$, then τ is FM.

Example

The constraint system τ of the following problem *is not FM* as $-x_1 \leq 0$ is a consequence of τ but not of any finite subsystem

$$(\mathcal{P}): \quad \text{Min} \quad x_1$$
$$\text{s.t.} \quad -x_1 - t^2 x_2 \leq -2t, \; t \in T =]0, \infty[.$$

Observe that $v = 0$ and $\mathbb{S} = \emptyset$, and that the optimal value of any finite subproblem is $v(S) = -\infty$.



The following theorem (*Dinh, Goberna, ML, Son' 07*) provides *non-asymptotic* KKT-type optimality conditions for the problem

$$(\mathcal{P}) \quad \text{Min} \quad g(x)$$
$$\text{s.t.} \quad f_t(x) \leq 0, t \in T, \quad x \in \mathbb{C},$$

with a non-empty feasible set $\mathbb{F}$.

Theorem (KKT'1)

Given the problem $(\mathcal{P})$, *assume that* τ *is FM and that one of the following conditions holds:*
a) g *is continuous at some point of* $\mathbb{F}$,
b) $(\mathrm{rint}\,\mathbb{F}) \cap \mathrm{rint}(\mathrm{dom}\,g) \neq \emptyset$.
Then $\overline{x} \in \mathbb{F} \cap \mathrm{dom}\,g$ *is a global minimum of* $(\mathcal{P})$ *if and only if there exists* $\lambda \in \mathbb{R}_+^{(T)}$ *such that* $\partial f_t(\overline{x}) \neq \emptyset$, $\forall t \in \mathrm{supp}\,\lambda$, *and the KKT conditions*

$$0_n \in \partial g(\overline{x}) + \sum_{t \in T} \lambda_t \partial f_t(\overline{x}) + \mathrm{N}_{\mathbb{C}}(\overline{x}) \text{ and } \lambda_t f_t(\overline{x}) = 0, \ \forall t \in T, \qquad \text{(KKT'1)}$$

hold.

Here $\mathbb{R}_+^{(T)}$ is the convex cone of functions $\lambda : T \longrightarrow \mathbb{R}_+$ which vanishes at every point of T except at finitely many.

Proof.

[Proof of KKT'1 (sketch)] The point $\overline{x} \in \mathbb{F} \cap \operatorname{dom} g$ is a minimizer of $(\mathcal{P})$ if and only if

$$0_n \in \partial(g + \mathrm{I}_{\mathbb{F}})(\overline{x}) \overset{(*)}{=} \partial g(\overline{x}) + \partial \mathrm{I}_{\mathbb{F}}(\overline{x}) = \partial g(\overline{x}) + \mathrm{N}_{\mathbb{F}}(\overline{x}); \qquad (11)$$

i.e., if and only if there exists $u \in \partial g(\overline{x})$ such that $\langle -u, x\rangle \leq \langle -u, \overline{x}\rangle$ is consequence of τ.

$(^*)$ Thanks to the assumption: a) or b) holds.

$(\Rightarrow)$ If $\overline{x}$ is a minimizer of $(\mathcal{P})$, since τ is FM we have

$$-(u, \langle u, \overline{x}\rangle) \in \operatorname{cl} \mathbb{K} = \mathbb{K} = \operatorname{cone} \left\{ \bigcup_{t \in T} \operatorname{epi} f_t^* \right\} + \operatorname{epi} \sigma_{\mathbb{C}},$$

and $\exists\ \lambda \in \mathbb{R}_+^{(T)}$, $u_t \in \operatorname{dom} f_t^*$, $\alpha_t \geq 0$, $\forall t \in T$, $v \in \operatorname{dom} \sigma_{\mathbb{C}}$, $\beta \geq 0$, satisfying

$$-(u, \langle u, \overline{x}\rangle) = \sum_{t \in T} \lambda_t \left(u_t, f_t^*\left(u_t\right) + \alpha_t\right) + \left(v, \sigma_{\mathbb{C}}\left(v\right) + \beta\right),$$

leading to (KKT'1) by the relationship between the subdifferential and the conjugate.

$(\Leftarrow)$ Straightforward (standard argument). □

KKT'2 optimality conditions - LFM property

Let us introduce a *weaker* CQ. Given $z \in \mathbb{F}$, the set of indices corresponding to the *active constraints at z* is $T(z) := \{t \in T : \ f_t(z) = 0\}$. It is easily verified that

$$\mathrm{N}_{\mathbb{C}}(z) + \operatorname{cone}\left(\bigcup_{t \in T(z)} \partial f_t(z)\right) \subseteq \mathrm{N}_{\mathbb{F}}(z). \tag{12}$$

Definition

The consistent constraint system τ is *locally Farkas-Minkowski (LFM*, in short*) at* $z \in \mathbb{F}$ if

$$\mathrm{N}_{\mathbb{F}}(z) \subseteq \mathrm{N}_{\mathbb{C}}(z) + \operatorname{cone}\left(\bigcup_{t \in T(z)} \partial f_t(z)\right). \tag{13}$$

τ is said to be *LFM* if it is *LFM at every feasible point* $z \in \mathbb{F}$.

In LSIP ($\mathbb{C} = \mathbb{R}^n, \ f_t(x) = \langle a_t, x\rangle - b_t, \ t \in T$), (13) becomes

$$\mathrm{N}_{\mathbb{F}}(z) \subseteq \operatorname{cone}\{a_t, t \in T(z)\}.$$

The LFM property is closely related to the so-called *basic constraint qualification at z*. In fact, LFM and BCQ are equivalent under the continuity of the function $f := \sup_{t \in T} f_t$ at the reference point z and $z \in \operatorname{int} \mathbb{C}$.

The following proposition is a LFM counterpart of a similar property for FM systems.

Theorem

Let $z \in \mathbb{F}$. If τ is LFM at z and for certain $a \in \mathbb{R}^n$ we have

$$\langle a, x\rangle \leq \langle a, z\rangle, \text{ for all } x \in \mathbb{F},$$

then $\langle a, x\rangle \leq \langle a, z\rangle$ is also a consequence of a finite subsystem of τ. The converse statement holds provided that τ is linear.

For general convex systems, it can be proved that

$$\tau \text{ is } FM \Rightarrow \tau \text{ is } LFM \text{ at any } z \in \mathbb{F}.$$

Example

The constraint system $-x_1 - t^2 x_2 \leq -2t,\ t \in T =]0, \infty[$, of the example is *LFM*.

The following theorem provides a *second* KKT-type optimality conditions for the problem

$$(\mathcal{P}) \quad \mathrm{Min}\, g(x) \quad \text{s.t. } f_t(x) \leq 0,\ t \in T, \quad x \in \mathbb{C}.$$

Theorem (KKT'2)

Given the problem $(\mathcal{P})$ *and* $\overline{x} \in \mathbb{F} \cap \operatorname{dom} g$, *assume that* τ *is LFM at* $\overline{x}$, *and that* g *is continuous at some point of* $\mathbb{F}$. *Then* $\overline{x}$ *is a global minimum of* $(\mathcal{P})$ *if and only if there exists* $\lambda \in \mathbb{R}_+^{(T)}$ *such that* $\partial f_t(\overline{x}) \neq \emptyset$, $\forall t \in \operatorname{supp} \lambda$, *and the KKT conditions hold*

$$\theta \in \partial g(\overline{x}) + \sum_{t \in T} \lambda_t \partial f_t(\overline{x}) + \mathrm{N}_{\mathbb{C}}(\overline{x}) \text{ and } \lambda_t f_t(\overline{x}) = 0,\ \forall t \in T. \qquad \text{(KKT'2)}$$

Proof.

According to *Pshenichnyi-Rockafellar theorem* (e.g. *Zal'02 [Th. 2.9.1]*),

$$\begin{aligned} \overline{x} \text{ is optimal for } (\mathcal{P}) \quad &\Leftrightarrow \quad \partial g(\overline{x}) \cap (-\mathrm{N}_{\mathbb{F}}(\overline{x})) \neq \emptyset \\ &\Leftrightarrow \quad \theta \in \partial g(\overline{x}) + \mathrm{N}_{\mathbb{F}}(\overline{x}) \\ &\overset{LFM}{\Leftrightarrow} \theta \in \partial g(\overline{x}) + \sum_{t \in T(\overline{x})} \lambda_t \partial f_t(\overline{x}) + \mathrm{N}_{\mathbb{C}}(\overline{x}). \end{aligned}$$

KKT'3 asymptotic optimality conditions

Theorem (KKT'3)

Given the problem $(\mathcal{P})$, *assume that* τ *is FM and* $(\operatorname{rint}\mathbb{F}) \cap \operatorname{dom} g \neq \emptyset$. *Then,* $\overline{x} \in (\operatorname{dom} g) \cap \mathbb{F}$ *is optimal for* $(\mathcal{P})$ *if and only if, for each fixed* $\varepsilon > 0$, *there exists* $\lambda^{\varepsilon} \in \mathbb{R}_{+}^{(T)}$ *such that* $\operatorname{supp} \lambda^{\varepsilon} \subset T(\overline{x})$ *and the following condition holds:*

$$\theta \in \partial_{\varepsilon} g(\overline{x}) + \sum_{\operatorname{supp} \lambda^{\varepsilon}} \lambda_t^{\varepsilon} \partial f_t(\overline{x}) + \mathrm{N}_{\mathbb{C}}(\overline{x}) + \varepsilon \mathbb{B}, \tag{14}$$

where $\mathbb{B}$ *is the closed unite ball (centered at* 0_n*).*

Proof.

[Sketch of the proof] $(\Rightarrow)$ Since $(\operatorname{rint}\mathbb{F}) \cap \operatorname{dom} g \neq \emptyset$, Th. 12 in *Correa, Hantoute, ML '16* yields

$$\partial(g + \mathrm{I}_{\mathbb{F}})(\overline{x}) = \bigcap_{\varepsilon > 0} \operatorname{cl}(\partial g_{\varepsilon}(\overline{x}) + \mathrm{N}_{\mathbb{F}}(\overline{x})).$$

Then,

$$\overline{x} \text{ is optimal for } (\mathcal{P}) \Leftrightarrow 0_n \in \bigcap_{\varepsilon > 0} \operatorname{cl}(\partial g_{\varepsilon}(\overline{x}) + \mathrm{N}_{\mathbb{F}}(\overline{x})).$$

Proof.

[Sketch of the proof - cont'd]
Thus,

$$\overline{x} \text{ is optimal for } (\mathcal{P}) \Leftrightarrow 0_n \in \partial g_\varepsilon(\overline{x}) + \mathrm{N}_{\mathbb{F}}(\overline{x}) + \rho\mathbb{B}, \text{ for every positive } \varepsilon \text{ and } \rho.$$

If we take $\rho = \varepsilon$, for every $\varepsilon > 0$, there exists $u_\varepsilon^* \in \mathrm{N}_{\mathbb{F}}(\overline{x})$ such that

$$0_n \in \partial g_\varepsilon(\overline{x}) + u_\varepsilon^* + \varepsilon\mathbb{B}.$$

Since $u_\varepsilon^* \in \mathrm{N}_{\mathbb{F}}(\overline{x})$ is equivalent to say that $\langle u_\varepsilon^*, x\rangle \leq \langle u_\varepsilon^*, \overline{x}\rangle$ is a consequence of the FM system τ, we conclude the existence of $\lambda^\varepsilon \in \mathbb{R}_+^{(T)}$, $\operatorname{supp}\lambda^\varepsilon \subset T(\overline{x})$, such that

$$u_\varepsilon^* \in \sum_{\operatorname{supp}\lambda^\varepsilon} \lambda_t^\varepsilon \partial f_t(\overline{x}) + \mathrm{N}_{\mathbb{C}}(\overline{x}).$$

The necessity is proved.
$(\Leftarrow)$ Straightforward (standard arguments). □

KKT'4 conditions for SIP under compacity/continuity

Consider the convex SIP problem

$$(\mathcal{P}) \quad \mathrm{Min}\, g(x) \quad \text{s.t. } f_t(x) \leq 0,\ t \in T, \quad x \in \mathbb{C}.$$

Theorem

Suppose that, for a given $\overline{x} \in \mathbb{F}$, *and some* $\varepsilon_0 > 0$
(i) $T_{\varepsilon_0}(x)$ *is compact,*
(ii) $\forall z \in \mathrm{dom}\, f$, $t \to f_t(z)$ *is usc on* $T_{\varepsilon_0}(x)$,
(iii) $\exists\ x_0 \in \mathbb{D}$ *such that* $\sup_{t \in T(\overline{x})} f_t(x_0) < 0$ *(Slater point),*
then for some $\lambda_1, \cdots, \lambda_k > 0$ *and* $t_1, \cdots, t_k \in T(\overline{x})$

$$0_n \in \partial g(\overline{x}) + \textstyle\sum_{i=1}^{k} \lambda_i \partial f_{t_i}(\overline{x}) + \mathrm{N}_{\mathbb{D}}(\overline{x}),$$

where $\mathbb{D} := \mathbb{C} \cap \mathrm{dom}\, g \cap \mathrm{dom}(\sup_{t \in T} f_t)$.

Proof Based on the equality (see Th. 3 in *Correa, Hantoute, ML '19*):

$$\partial f(x) = \mathrm{co}\left(\textstyle\bigcup_{t \in T(x)} \partial (f_t + \mathrm{I}_{\mathrm{dom}\, f})(x) \right).$$

Bibliographic comments

- The closedness of $\mathbb{K}$ was introduced in *Charnes, Cooper, Kortanek'65* as a general assumption for the duality theory in LSIP.
- The FM property for convex systems was first studied in *Jeyakumar, Lee, Dinh'04* , with X being Banach and all the functions finite valued, under the name of *closed cone constraint qualification.* In the framework of optimality conditions for the convex SIP was first considered in *ML-Vercher'83.*
- The LFM property, under the name of *basic constraint qualification (BCQ),* appeared in *Hiriart-Urruty, Lemaréchal'93*. It was extended in *Puente, Vera de Serio'99* to linear semi-infinite systems. The consequences of its extension to convex semi-infinite systems were analyzed in *Fajardo, López'99.*
- For a deep analysis of BCQ and related conditions see also *Li, Nahak, Singer'00* and *Li, Ng'05.* An extensive comparative analysis of constraints qualifications for $(\mathcal{P})$ is also given in *Li, Ng, Pong'08*.

References

Charnes A. Cooper W.W., K.O. Kortanek K.O., *On representations of semi-infinite programs which have no duality gaps*, Management Science 12 (1965) 113–121.

Correa R., Hanoute A., LÓpez M.A., *Moreau-Rockafellar-type formulas for the subdifferential of the supremum function*, SIAM Journal on Optimization 29 (2019)

N. Dinh N., Goberna M.A., López M.A., *From linear to convex systems: consistency, Farkas' lemma and applications*. J. Convex Anal. 13 (2006), 113–133.

Dinh N., Goberna M.A., López M.A., Son T.Q., *New Farkas-type constraint qualifications in convex infinite programming*, ESAIM: COCV 13 (2007)

M.D. Fajardo M.D., López M.A., *Locally Farkas-Minkowski systems in convex semi-infinite programming*. Journal of Optimization Theory and Applications **103** (1999), 313-335.

- GOBERNA M.A., LÓPEZ M.A., *Linear Semi-infinite Optimization*. J. Wiley, Chichester (1998).
- JEYAKUMAR V., *Characterizing set containments involving infinite convex constraints and reverse-convex constraints*, SIAM J. Optimization **13** (2003) 947-959.
- JEYAKUMAR V., LEE G.M., DINH N., *A new closed cone constraint qualification for convex Optimization,* Applied Mathematics Research Report AMR04/8, University of New South Wales, Sydney, Australia (2004)
- HIRIART-URRUTY J.-B., PHELPS R.R.: *Subdifferential calculus using epsilon-subdifferentials*, J. Funct. Anal., 118 (1993)
- HIRIART-URRUTY J.-B., LEMARÉCHAL C.: *Convex Analysis and Minimization Algorithms I, II,* Springer-Verlarg, Berlin, 1993.

Li W., Nahak C., Singer I., *Constraint qualification for semi-infinite systems of convex inequalities*. SIAM J. Optimization 11 (2000) 31-52.

Li Ch., Ng K.F., *On constraint qualification for an infinite system of convex inequalities in a Banach space*. SIAM J. Optimization 15 (2005) 488-512.

Li Ch., Ng K.F., Pong T.K., *Constraint qualifications for convex inequality systems with applications in constrained optimization*. SIAM J. Optim. 19 (2008), 163–187.

Lopez M.A., Vercher E., *Optimality conditions for nondifferentiable convex semi-infinite programming*. Math. Programming 27 (1983), 307-319.

Puente R., Vera de Serio V.N., *Locally Farkas-Minkowski linear semi-infinite systems*. *TOP* 7 (1999) 103-121.

Zalinescu, C. *Convex analysis in general vector spaces*. World Scientific Publishing Co., Inc., River Edge, NJ, 2002.

Thank you for your attention