manuscript No.
(will be inserted by the editor)

Best approximate solutions of inconsistent linear inequality systems*

Miguel A. Goberna[1], Jean-Baptiste Hiriart-Urruty[2], Marco A. López[3]**

[1] Department of Mathematics, University of Alicante, Spain, e-mail: mgoberna@ua.es
[2] Institute of Mathematics, Paul Sabatier University, Toulouse, France, e-mail: jbhu@math.univ-toulouse.fr
[3] Department of Mathematics, University of Alicante, Spain and Federation University of Australia, Ballarat, Australia, e-mail: marco.antonio@ua.es

The date of receipt and acceptance will be inserted by the editor

To Michel Théra in occasion of his 70th anniversary

Abstract This paper is intended to characterize three types of best approximate solutions for inconsistent linear inequality systems with an arbitrary number of constraints. It also gives conditions guaranteeing the existence of best uniform solutions and discusses potential applications.

Key words Linear inequality systems – best uniform solutions – Best L_1 solutions – best least squares solutions
MSC2010 Subject Classification: 15A39, 65F20, 90C25, 90C34

Mathematics Subject Classification (2010) 15A39, 65F20, 90C25, 90C34

1 Introduction

This paper deals with approximate solutions of inconsistent linear systems of the form

$$\left\{ a(t)^\top x \leq b(t),\ t \in T \right\}, \tag{1}$$

where T is a (possibly infinite) index set, $a : T \longrightarrow \mathbb{R}^n$ and $b : T \longrightarrow \mathbb{R}$ are given mappings, and $a(t)^\top$ represents the transpose of $a(t)$. The system in

* This work was partially supported by the MINECO of Spain and ERDF of EU, Grant MTM2014-59179-C2-1-P, and by the Australian Research Council, Project DP160100854.

** Corresponding author.

(1) is called *finite* (respectively, *semi-infinite*) whenever the index set T is finite (infinite). In the particular case of finite linear systems, it is customary to write $a_t \in \mathbb{R}^n$ (respectively, $b_t \in \mathbb{R}$) instead of $a(t)$ $(b(t))$ for all $t \in T$. We will use both notations at our convenience.

The constraint systems of linear optimization problems (also called programs) with finite dimensional decision space are expressed as (1). These problems are called either *finite* or *semi-infinite* according to the nature of its constraint system. In many (possibly inconsistent) linear semi-infinite programs arising in practice, T is either an interval of time, or a region in the plane (or in the space) on which the functions a and b are continuous (see, *e.g.*, [9] and [11]). In particular, the constraint systems of the linear semi-infinite programs arising in the best uniform and L_1 approximation of real-valued functions by polynomials are always consistent [9, Chapter 1]) but, in general, the consistency or not of the system in (1) depends on the data, *i.e,* the mappings a and b. In the linear semi-infinite programming (SIP in short) literature, the programs, as well as their corresponding constraint systems, are said to be *discrete* whenever T is a topological space without accumulation points (*e.g.*, a finite set equipped with the discrete topology) and *continuous* whenever T is a compact Hausdorff space and the functions a and b are continuous on T.

In this paper we evaluate the infeasibility of a given $x \in \mathbb{R}^n$ with respect to a given inconsistent system $\left\{a(t)^\top x \leq b(t),\ t \in T\right\}$ by means of three norms defined on certain spaces of functions. Since two of these definitions involve integrals, on T, of functions depending on a and b, we must impose suitable conditions on a, b and T guaranteeing integrability. Moreover, in order to characterize the best approximation for one of these norms, we must apply the Leibnitz integral rule for derivation under the integral sign, which has been established for boxes, that is, cartesian products of (possibly improper) compact intervals in $\mathbb{R}$. For this reason, we consider in this paper linear systems such that T is the union of two disjoint subspaces of some Euclidean space $\mathbb{R}^m$ (at least one of them nonempty): a finite set and a finite union of pairwise disjoint compact Hausdorff sets on which the mappings a and b are continuous. More precisely, we assume that the index set in (1) can be expressed as

$$T = \left\{t^1, ..., t^q\right\} \cup \left(\bigcup_{j=1}^{p} T_j\right) \subset \mathbb{R}^m,$$

where T_j is a box, $a_{|T_j} \in \mathcal{C}(T_j, \mathbb{R})^n$, $b_{|T_j} \in \mathcal{C}(T_j, \mathbb{R})$, $j = 1, ..., p$, where $\mathcal{C}(T_j, \mathbb{R})$ denotes the space of real-valued continuous functions on T_j, and the sets $\left\{t^1, ..., t^q\right\}, T_1, ..., T_p$ are pairwise disjoint.

Obviously, $\left\{a_t^\top x \leq b_t, t \in T\right\}$ is called *discrete* when $\bigcup_{j=1}^{p} T_j = \emptyset$ (*i.e.*, $p = 0 \neq q$) and *continuous* when $\left\{t^1, ..., t^q\right\} = \emptyset$ (*i.e.*, $q = 0 \neq p$). We say that $\left\{a_t^\top x \leq b_t, t \in T\right\}$ is a *mixed* system otherwise (*i.e.*, when $p \neq 0 \neq q$).

Examples 1, 2, and 5 below deal with inconsistent discrete, continuous, and mixed systems, respectively.

The *residual* of $x \in \mathbb{R}^n$ is the nonzero real-valued continuous function

$$T \ni t \mapsto \left(a_t^\top x - b_t\right)_+ := \max\left\{0, a_t^\top x - b_t\right\},$$

whose size can be measured in different ways, *e.g.*, by the L_∞, L_1, and L_2 norms in the space $\mathcal{C}(T, \mathbb{R})$ of real-valued continuous functions on T. We consider the problem of computing the best approximate solution of $\left\{a_t^\top x \le b_t, t \in T\right\}$ for these norms (*i.e.*, finding those $x \in \mathbb{R}^n$ minimizing the corresponding norm of the residual).

The *best uniform solutions* to $\left\{a_t^\top x \le b_t,\ t \in T\right\}$ are the optimal solutions to the problem

$$P_0: \ \operatorname{Min}_{x\in\mathbb{R}^n} f_0(x) = \max_{t\in T}\left(a_t^\top x - b_t\right)_+ = \max_{t\in T}\left(a_t^\top x - b_t\right),$$

the *best* L_1*−solutions* are the optimal solutions to the problem

$$P_1: \ \operatorname{Min}_{x\in\mathbb{R}^n} f_1(x) = \sum_{j=1}^{p} \int_{T_j} \left(a_t^\top x - b_t\right)_+ dt_1 ... dt_m + \sum_{k=1}^{q} \left(a_{t^k}^\top x - b_{t^k}\right)_+,$$

and, finally, the *best least squares solutions* are the optimal solutions to

$$P_2: \ \operatorname{Min}_{x\in\mathbb{R}^n} \ f_2(x) = \frac{1}{2}\sum_{j=1}^{p} \int_{T_j} \left[\left(a_t^\top x - b_t\right)_+\right]^2 dt_1 ... dt_m + \frac{1}{2}\sum_{k=1}^{q} \left[\left(a_{t^k}^\top x - b_{t^k}\right)_+\right]^2.$$

(We take the square of the Euclidean norm for the sake of smoothness of the objective function and divide by 2 to simplify the expressions of the gradient of f_2.)

We illustrate these unconstrained convex optimization problems with a simple example.

Example 1 Consider the discrete system, with $n = 2$, $p = 0$ and $q = 3$,

$$\{x_1 \le -1, -x_1 \le -1,\ x_2 \le 1\},$$

whose inequalities are indexed with $t = 1, 2, 3$. It is easy to see that

$$f_0(x) = \max\{|x_1| + 1,\ x_2 - 1\},$$

$$f_1(x) = \begin{cases} x_1 + x_2, & x \in [1, +\infty[\times [1, +\infty[, \\ x_2 + 1, & x \in [-1, 1] \times [1, +\infty[, \\ -x_1 + x_2, & x \in]-\infty, -1] \times [1, +\infty[, \\ -x_1 + 1, & x \in]-\infty, -1] \times]-\infty, 1], \\ 2, & x \in [-1, 1] \times]-\infty, 1], \\ x_1 + 1, & x \in [1, +\infty[\times]-\infty, 1], \end{cases}$$

and

$$f_2(x) = \frac{1}{2}\left\{\left[(x_1+1)_+\right]^2 + \left[(-x_1+1)_+\right]^2 + \left[(x_2-1)_+\right]^2\right\}.$$

One immediately realizes that the set of minimizers of f_0 is $\{0\} \times]-\infty, 2]$, that the set of minimizers of f_1 is $[-1,1] \times]-\infty, 1]$, after solving six very simple linear programs, and verifies, with more effort (see [5, Example 2.2.1]), that the set of minimizers of f_2 is $\{0\} \times]-\infty, 1]$. Thus, the three problems, P_0, P_1 and P_2 have multiple optimal solutions and none of the objective functions, f_0, f_1 and f_2, is coercive.

Obviously, the functions f_0, f_1 and f_2 are bounded from below, but their infimum values can be unattainable; in other words, it may happen that the sets of best uniform, L_1 and least squares solutions be empty as the following example shows:

Example 2 Consider the system $\sigma_1 = \left\{-t^2x \leq -2t,\ t \in [0,1]\right\}$ (with $n = m = p = 1$ and $q = 0$). Since $a_t x - b_t = t(2 - tx)$, we have

$$a_t x - b_t \geq 0 \Longleftrightarrow \begin{cases} t \in \left[0, \frac{2}{x}\right], & \text{if } x > 0, \\ t \in \mathbb{R}_+, & \text{if } x = 0, \\ t \in \left]-\infty, \frac{2}{x}\right] \cup \mathbb{R}_+, & \text{if } x < 0. \end{cases}$$

So,

$$\{t \in \mathbb{R} : a_t x - b_t \geq 0\} = \begin{cases} \left[0, \frac{2}{x}\right], & \text{if } x > 0, \\ \mathbb{R}_+, & \text{if } x = 0, \\ \left]-\infty, \frac{2}{x}\right] \cup \mathbb{R}_+, & \text{if } x < 0, \end{cases}$$

and

$$\{t \in [0,1] : a_t x - b_t \geq 0\} = \begin{cases} [0,1], & \text{if } x \leq 2, \\ \left[0, \frac{2}{x}\right], & \text{else.} \end{cases}$$

Consequently,

$$f_0(x) = \begin{cases} 2 - x, & \text{if } x \leq 1, \\ \frac{1}{x}, & \text{else,} \end{cases}$$

$$f_1(x) = \begin{cases} 1 - \frac{x}{3}, & \text{if } x \leq 2, \\ \frac{4}{3x^2}, & \text{else,} \end{cases}$$

and

$$f_2(x) = \begin{cases} \frac{2}{3} - \frac{x}{2} + \frac{x^2}{10}, & \text{if } x \leq 2, \\ \frac{8}{15x^3}, & \text{else,} \end{cases}$$

so that the three functions are decreasing and differentiable, and their infimum (zero) is not attained.

Observe that the existence of linear systems such that f_0, f_1 and f_2 are decreasing is independent of the dimension of the space of variables (consider, *e.g.*, the system $\left\{-t^2 x_n \leq -2t, t \in [0,1]\right\}$ in $\mathbb{R}^n$). Consequently, regarding the existence of approximate solutions for continuous and mixed systems, our main objective will be to give sufficient conditions as general as possible.

There exists a stream of works dealing with *discrete* inconsistent linear systems with different purposes:

1. Correcting the data, *i.e.*, determining the smallest perturbation of the perturbable data of an inconsistent system $Ax \leqq b$ providing a consistent one (see, *e.g.*, [1], [16], and [4], whose results also apply to continuous inconsistent systems). The analyzed perturbations affect either the pair (A, b), or the matrix A, or the column vector b, and are measured with a variety of norms.

2. Calculating error bounds, *i.e.* positive scalars that multiplied by the norm of the residual of any $x \in \mathbb{R}^n$ provide bounds for the distance between x and the set of minimizers of that norm ([18]).

3. Numerical methods to compute best least squares solutions (see, *e.g.*, the recent paper [22] on variants of Han's algorithm [12], whose fundamentals are the existence and characterization theorems for discrete inconsistent system proved in the latter paper; the recent work [5] provides two new proofs of Han's existence theorem; two additional works, [6] and [17], on the so-called hybrid algorithm). To the best of our knowledge, no extension of these results and methods to continuous and mixed linear inconsistent systems is still available.

The paper is organized as follows: Section 2 provides necessary or sufficient conditions for the existence of best uniform solutions, which are also characterized; Sections 3 and 4 are concerned with the characterization of best L_1 and L_2 solutions; Section 5 contains some comments on the possible use of the presented results in applications; finally, Section 6 presents some conclusions.

2 Best uniform solutions

We now introduce the necessary notation. The zero vector in $\mathbb{R}^n$ is denoted by 0_n. Given a set $X \subset \mathbb{R}^n$, we denote by $\operatorname{conv} X$, $\operatorname{cone} X = \mathbb{R}_+ \operatorname{conv} X$, $\operatorname{int} X$, $\operatorname{cl} X$, $\operatorname{bd} X$, and $\operatorname{rint} X$ the convex hull of X, the convex conical hull of X, the interior of X, the closure of X, the boundary of X , and the relative interior of X, respectively. The *indicator function* of X is $I_X : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$, defined by $I_X(x) = 0$, if $x \in X$, and $I_X(x) = +\infty$, otherwise. Given a function $f : \mathbb{R}^n \to \overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty, -\infty\}$, its *domain* is $\operatorname{dom} f = \{x \in \mathbb{R}^n : f(x) < +\infty\}$, and its *epigraph* is

$$\operatorname{epi} f = \left\{(x, \alpha) \in \mathbb{R}^{n+1} : f(x) \leq \alpha\right\}.$$

If $\operatorname{dom} f \neq \emptyset$ and $-\infty \notin f(\mathbb{R}^n)$ the function f is called *proper*, and if $\operatorname{epi} f$ is closed we say that f is a *closed function.*

The *closure* of a proper convex function f is the closed proper function $\operatorname{cl} f$ such that

$$\operatorname{epi}(\operatorname{cl} f) = \operatorname{cl}(\operatorname{epi} f).$$

The *Legendre-Fenchel conjugate* of f is the function $f^* : \mathbb{R}^n \to \overline{\mathbb{R}}$ defined by

$$f^*(u) = \sup_{x \in \mathbb{R}^n} \left(u^\top x - f(x) \right).$$

Given a family of proper convex functions on $\mathbb{R}^n$, the *convex hull* of these functions is the convex function $\operatorname{conv}_{t \in T} f_t$ defined by

$$(\operatorname*{conv}_{t \in T} f_t)(x) = \inf \left\{ \mu : \ (x, \alpha) \in \operatorname{conv} \left(\bigcup_{t \in T} \operatorname{epi} f_t \right) \right\}.$$

It is well-known (*e.g.* [23, Theorem 16.5]) that if $\{f_t, \ t \in T\}$ is a family of closed proper convex functions on $\mathbb{R}^n$, then

$$\left(\sup_{t \in T} f_t \right)^* = \operatorname{cl} \left(\operatorname{conv} \{ f_t^*, \ t \in T \} \right). \tag{2}$$

In this section the special structure of T is of no use, except the fact that it is a compact set. In other words, we are considering the system

$$\sigma := \{ a_t^\top x \leq b_t, \ t \in T \}$$

with T compact and the functions $t \mapsto a_t$ and $t \mapsto b_t$ continuous. The inconsistency of σ entails

$$f_0(x) = \max\{ a_t^\top x - b_t, \ t \in T \} > 0, \text{ for all } x \in \mathbb{R}^n.$$

Now we derive the conjugate of f_0. If

$$f_t(x) := a_t^\top x - b_t, \quad t \in T,$$

since

$$f_t^*(\cdot) = I_{\{a_t\}}(\cdot) + b_t, \ t \in T,$$

(2) and the compactness of the set

$$C := \operatorname{conv}\{ (a_t, b_t) : \ t \in T \},$$

yield

$$\begin{aligned} \operatorname{epi} f_0^* &= \operatorname{cl} \left(\operatorname{conv} \left\{ \bigcup_{t \in T} \{ (a_t, b_t) + \mathbb{R}_+ (0_n, 1) \} \right\} \right) \\ &= C + \mathbb{R}_+ (0_n, 1), \end{aligned}$$

and

$$f_0^*(u) = \min \left\{ \sum_{t \in T} \lambda_t b_t \ \middle| \ \begin{array}{l} \lambda_t \geq 0, \text{ only finitely many positive,} \\ \sum_{t \in T} \lambda_t = 1, \text{ and } \sum_{t \in T} \lambda_t a_t = u \end{array} \right\}. \tag{3}$$

According to Theorem 11.8(d) in [24], the function f_0 never is 1-coercive since $\operatorname{dom} f_0^* \neq \mathbb{R}^n$.

Corollary 3.1.1 in [9] establishes that $\sigma = \left\{a_t^\top x \leq b_t, t \in T\right\}$ is inconsistent if and only if

$$\begin{pmatrix} 0_n \\ -1 \end{pmatrix} \in \operatorname{cl}\operatorname{cone}\left\{\begin{pmatrix} a_t \\ b_t \end{pmatrix} : t \in T\right\}. \tag{4}$$

From now on we only deal with an inconsistent system $\sigma = \left\{a_t^\top x \leq b_t, t \in T\right\}$. The following proposition gives a necessary condition *(N)*, which is independent of the right-hand side b, and a sufficient condition *(S)* for the existence of a best uniform solution for σ.

Proposition 1 (Existence of best uniform solutions) *Given an inconsistent system $\sigma = \left\{a_t^\top x \leq b_t, t \in T\right\}$, the following statements hold:*
(S) If σ is discrete or satisfies

$$0_n \in \operatorname{rint}\operatorname{conv}\left\{a_t : t \in T\right\}, \tag{5}$$

then there exists a best uniform solution. When the convex hull in (5) is full dimensional, then the set of best uniform solutions is bounded.
(N) If there exists a best uniform solution, then

$$\begin{pmatrix} 0_n \\ -1 \end{pmatrix} \in \operatorname{cone}\left\{\begin{pmatrix} a_t \\ b_t \end{pmatrix} : t \in T\right\}. \tag{6}$$

Proof Proof: It is based on well-known results of convex analysis.
(S) If σ is an inconsistent discrete system, f_0 is a polyhedral function which is bounded from below, so its minimum is attained.

Suppose now that $\bigcup_{j=1}^{p} T_j \neq \emptyset$ and that σ is inconsistent. Then $\overline{x}$ will be a best uniform solution of σ if and only if $0_n \in \partial f_0(\overline{x})$, but this happens if and only if $\overline{x} \in \partial f_0^*(0_n)$; *i.e.* $\partial f_0^*(0_n)$ is the set of best uniform solutions of σ. Since $\operatorname{dom} f_0^* = \operatorname{conv}\{a_t : \ t \in T\}$, we have

$$0_n \in \operatorname{rint}\operatorname{conv}\{a_t : \ t \in T\} \Longrightarrow \partial f_0^*(0_n) \neq \emptyset,$$

and

$$0_n \in \operatorname{int}\operatorname{conv}\{a_t : \ t \in T\} \Longrightarrow \partial f_0^*(0_n) \text{ is non-empty and compact.}$$

(N) If there exists a best uniform solution of σ, $\overline{x}$, we have $\overline{x} \in \partial f_0^*(0_n)$, and

$$f_0(\overline{x}) + f_0^*(0_n) = 0_n^\top \overline{x} = 0.$$

Since σ is inconsistent, $f_0(\overline{x}) > 0$, entailing $f_0^*(0_n) < 0$. Then, according to the expression of f_0^* given in (3), there will exist scalars $\overline{\lambda}_t \geq 0$, only finitely many positive, and such that $\sum_{t\in T} \overline{\lambda}_t = 1$,

$$\sum_{t\in T} \overline{\lambda}_t a_t = 0_n, \text{ and } f_0^*(0_n) = \sum_{t\in T} \overline{\lambda}_t b_t.$$

Therefore

$$\sum_{t\in T} \frac{\overline{\lambda}_t}{|f_0^*(0_n)|} \begin{pmatrix} a_t \\ b_t \end{pmatrix} = \begin{pmatrix} 0_n \\ -1 \end{pmatrix},$$

and we get the aimed necessary condition (6).

Remark 1 It is also possible to prove Proposition 1 by using linear SIP theory as P_0 is equivalent to the linear SIP problem

$$\begin{array}{lll} P_0' : & \mathrm{Min}_{(x,x_{n+1})\in\mathbb{R}^{n+1}} & x_{n+1} \\ & \text{subject to} & a_t^\top x - x_{n+1} \leq b_t, \ t \in T. \end{array}$$

To prove (S), one can assume that $\bigcup_{j=1}^{p} T_j \neq \emptyset$. By [9, Theorem 8.1(v)], P_0' is solvable if and only if

$$\begin{pmatrix} 0_n \\ -1 \end{pmatrix} \in \operatorname{rint} \operatorname{cone} \left\{ \begin{pmatrix} a_t \\ -1 \end{pmatrix} : t \in T \right\},$$

which is equivalent to (5). On the other hand, by [9, Theorem 8.1(vi)], the optimal set of P_0' is bounded if and only if

$$\begin{pmatrix} 0_n \\ -1 \end{pmatrix} \in \operatorname{int} \operatorname{cone} \left\{ \begin{pmatrix} a_t \\ -1 \end{pmatrix} : t \in T \right\},$$

which is equivalent to $0_n \in \operatorname{int} \operatorname{conv} \{a_t : t \in T\}$.
Regarding (N), under the assumptions on the system in (1), the dual problem of P_0' in Haar's sense is also solvable with the same optimal value $f_0(\overline{x}) > 0$ (see, *e.g.*, [9, Theorem 8.1]). This means that there exist scalars $\overline{\lambda}_t \geq 0$, only finitely many positive such that

$$\sum_{t\in T} \overline{\lambda}_t \begin{pmatrix} -a_t \\ 1 \end{pmatrix} = \begin{pmatrix} 0_n \\ -1 \end{pmatrix} \text{ and } -\sum_{t\in T} \overline{\lambda}_t b_t = f_0(\overline{x}).$$

The rest of the proof is as above.

Example 1 shows that condition (5) is not necessary for the existence of a best uniform solution, as

$$0_2 = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \in \operatorname{bd} \operatorname{conv} \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}.$$

Moreover,

$$\operatorname{argmax}\left\{a_t^\top \begin{pmatrix} 0 \\ x_2 \end{pmatrix} - b_t : t \in T\right\} = \{1,2,3\},$$

with $0_2 \in \operatorname{conv}\{a_1, a_2, a_3\}$, while, for any $x_2 < 2$,

$$\operatorname{argmax}\left\{a_t^\top \begin{pmatrix} 0 \\ x_2 \end{pmatrix} - b_t : t \in T\right\} = \{1,2\},$$

with $0_2 \in \operatorname{conv}\{a_1, a_2\}$, confirming that any element of $\{0\} \times]-\infty, 2]$ is a best uniform solution (observe that (8) given later fails at all these points). The next example shows that, in a similar way, condition (6) meaning that σ contains a finite inconsistent subsystem does not guarantee the existence of uniform solutions to σ.

Example 3 Replacing the right-hand side in the system σ_1 (see Example 2) $-2t$, by $-2t-1$, for the new system, say σ_2,

$$f_0(x) = \begin{cases} 3 - x, \text{ if } x \leq 1, \\ 1 + \frac{1}{x}, \text{ else,} \end{cases}$$

so that its infimum, now 1, is still unattainable. The difference is that σ_2 satisfies the necessary condition for the existence of uniform solution as $(a_0, b_0) = (0, -1)$ while σ_1 does not. Thus, the necessary condition given in Proposition 1 is not sufficient.

Proposition 2 (Characterization of best uniform solutions) *A given $\overline{x} \in \mathbb{R}^n$ is a best uniform solution if and only if*

$$0_n \in \operatorname{conv}\left\{a_t : t \in \operatorname{argmax}\{a_t^\top \overline{x} - b_t : t \in T\}\right\}. \tag{7}$$

In particular, if

$$0_n \in \operatorname{int}\operatorname{conv}\left\{a_t : t \in \operatorname{argmax}\{a_t^\top \overline{x} - b_t : t \in T\}\right\}, \tag{8}$$

then $\overline{x}$ is the unique best uniform solution.

Proof The first statement comes from $0_n \in \partial f_0(\overline{x})$ and the Valadier formula (*e.g.* [13, VI Corollary 4.4.4]). The second statement is a consequence of the so-called Polyak's condition for the existence of sharp minimum [21].

Remark 2 The proof of Proposition 2 based on linear SIP theory is not so simple. Let $(\overline{x}, \overline{x}_{n+1}) \in \mathbb{R}^{n+1}$ with set of active indices $T(\overline{x}, \overline{x}_{n+1}) := \left\{t \in T : a_t^\top \overline{x} - \overline{x}_{n+1} = b_t\right\}$.
On the one hand, by [9, Theorem 7.1], $(\overline{x}, \overline{x}_{n+1}) \in \mathbb{R}^{n+1}$ is an optimal solution of P_0' if and only if

$$\begin{pmatrix} 0_n \\ -1 \end{pmatrix} \in \operatorname{cone}\left\{\begin{pmatrix} a_t \\ -1 \end{pmatrix} : t \in T(\overline{x}, \overline{x}_{n+1})\right\},$$

i.e.,

$$0_n \in \operatorname{conv}\{a_t : t \in T(\overline{x}, \overline{x}_{n+1})\}. \tag{9}$$

In this case, $\overline{x}_{n+1} = \max_{t\in T}\left(a_t^\top \overline{x} - b_t\right)$ and so $T\left(\overline{x}, \overline{x}_{n+1}\right) = \operatorname{argmax}\{a_t^\top \overline{x} - b_t : t \in T\}$ which replaced in (9) yields (7).
On the other hand, if (10) holds, *i.e.*,

$$\begin{pmatrix} 0_n \\ -1 \end{pmatrix} \in \operatorname{int}\operatorname{cone}\left\{\begin{pmatrix} a_t \\ -1 \end{pmatrix} : t \in \operatorname{argmax}\{a_t^\top \overline{x} - b_t : t \in T\}\right\}, \tag{10}$$

then, by [9, Theorem 10.6], $\left(\overline{x}, \max_{t\in T}\left(a_t^\top \overline{x} - b_t\right)\right)$ is a strongly unique optimal solution of P_0'.

The next example illustrates the application of Proposition 2 to a continuous inconsistent system.

Example 4 Consider the deterministic counterpart of the uncertain system in $\mathbb{R}^2$ $\{-2x_1 \leq 0,\ 2x_1 \leq 0\}$, with uncertainty intervals of the form $\alpha \pm \varepsilon$ for each coefficient α, with $\varepsilon > 0$, formulated as P_R in (17). This counterpart is the inconsistent continuous system $\left\{a_t^\top x \leq b_t,\ t \in T\right\}$, where $T = T_1 \cup T_2$, with $T_j = \left((-1)^j 2, 0, 0\right) + C$, $j = 1, 2$, $C = [-\varepsilon, \varepsilon]^3$, $a_t = (t_1, t_2)^\top$ and $b_t = t_3$ for all $t = (t_1, t_2, t_3) \in T$. Let us analyze the particular case $\varepsilon = 1$. The problem to be solved is

$$\begin{array}{ll} P_0' : \operatorname{Min}_{(x_1,x_2,x_3)\in\mathbb{R}^3} & x_3 \\ \quad\ \text{subject to} & t_1x_1 + t_2x_2 - x_3 \leq t_3,\ t \in T_1 \cup T_2, \end{array}$$

where all constraints of $t_1x_1 + t_2x_2 - x_3 \leq t_3,\ t \in T_j$, are consequences of the inequalities corresponding to the extreme points of T_j. Observe also that, for these points, fixed t_1 and t_2, t_3 takes values ± 1, with $t_1x_1 + t_2x_2 - x_3 \leq 1$ being a consequence of $t_1x_1 + t_2x_2 - x_3 \leq -1$. So, P_0' is equivalent to the problem P_0'' obtained by replacing the constraints corresponding to indices $t \in T_1 \cup T_2$ with the subsystem formed by the following eight inequalities:

$$\begin{aligned} -x_1 + x_2 - x_3 &\leq -1, && \text{(a)} \\ -3x_1 + x_2 - x_3 &\leq -1, && \text{(b)} \\ -x_1 - x_2 - x_3 &\leq -1, && \text{(c)} \\ -3x_1 - x_2 - x_3 &\leq -1, && \text{(d)} \\ x_1 + x_2 - x_3 &\leq -1, && \text{(e)} \\ 3x_1 + x_2 - x_3 &\leq -1, && \text{(f)} \\ x_1 - x_2 - x_3 &\leq -1, && \text{(g)} \\ 3x_1 - x_2 - x_3 &\leq -1. && \text{(h)} \end{aligned}$$

If we sum term by term the inequalities (a), (c), (e) and (g), we conclude that $-x_3 \leq -1$ is a consequence of the inequality system above and so, an optimal solution of P_0' is $(0, 0, 1)$, which means that $\overline{x} = (0, 0)$ minimizes f_0 with $f_0\left(0, 0\right) = 1$.

We now observe that $a_t^\top \overline{x} - b_t = -t_3$ attains its maximum on T at the union of the lower facets of T_1 and T_2, *i.e.*,

$$\begin{aligned}\operatorname{argmax}\left\{a_t^\top \overline{x} - b_t : t \in T\right\} = &([-3,-1] \times [-1,1] \times \{-1\}) \\ &\cup ([1,3] \times [-1,1] \times \{-1\}).\end{aligned}$$

Hence $\overline{x}$ satisfies (7) and (8), *i.e.*, it is the unique best uniform solution.

3 Best L_1 solutions

In the discrete case P_1 consists in minimizing a nonnegative (convex) piecewise linear function. This allows us to assert the existence of solutions as in Proposition 1 (without any assumption on the data (a_{t^k}, b_{t^k}), $k = 1, ..., q$). Next we focus on the characterization of optimal solution in the mixed case (the corresponding to the discrete and the continuous cases can be seen as particular cases).

Proposition 3 (Characterization of L_1-solutions) *A given $\overline{x} \in \mathbb{R}^n$ is a best L_1-solution if and only if*

$$0_n \in \sum_{j=1}^{p} \int_{T_j} A_t^j \, dt + \sum_{\substack{1 \le k \le q: \\ a_{t^k}^\top \overline{x} = b_{t^k}}} \operatorname{conv}\{a_{t^k}, 0_n\} + \sum_{\substack{1 \le k \le q: \\ a_{t^k}^\top \overline{x} > b_{t^k}}} a_{t^k}, \qquad (11)$$

where

$$A_t^j := \begin{cases} \{a_t\}, & \text{if } a_t^\top \overline{x} - b_t > 0, \\ \operatorname{conv}\{a_t, 0_n\}, & \text{if } a_t^\top \overline{x} - b_t = 0, \\ \{0_n\}, & \text{if } a_t^\top \overline{x} - b_t < 0, \end{cases}$$

and

$$\int_{T_j} A_t^j dt = \left\{ z = \int_{T_j} u(t) dt : \ u(.) \text{ is a Lebesgue-integrable selection in } A_{(.)}^j \right\}.$$

Proof In this mixed setting, the L_1 solutions are those vectors $\overline{x}$ satisfying

$$0_n \in \partial \left\{ \sum_{j=1}^{p} \int_{T_j} \left(a_t^\top \overline{x} - b_t\right)_+ dt_1 ... dt_m + \sum_{k=1}^{q} \left(a_{t^k}^\top \overline{x} - b_{t^k}\right)_+ \right\}. \qquad (12)$$

Taking into account that the function $t \mapsto \left(a_t^\top \overline{x} - b_t\right)_+$ is continuous on each T_j, that $x \mapsto \left(a_t^\top x - b_t\right)_+$ is convex and finite-valued for each $x \in \mathbb{R}^n$, and so it is a normal convex integrand (see, e.g. [24, Proposition 14.39]), we can write (applying e.g. [14, Theorem 4, §8.3] and observing that all the integrals are finite-valued convex functions)

$$0_n \in \sum_{j=1}^{p} \int_{T_j} \partial \left(a_t^\top \overline{x} - b_t\right)_+ dt + \sum_{k=1}^{q} \partial \left(a_{t^k}^\top \overline{x} - b_{t^k}\right)_+, \qquad (13)$$

where

$$\int_{T_j} \partial \left(a_t^\top \overline{x} - b_t\right)_+ dt = \left\{ z = \int_{T_j} u(t)\ dt \;\middle|\; \begin{array}{c} u(.) \text{ is Lebesgue-integrable in } T_j \\ \text{and } u(t) \in \partial \left(a_t^\top \overline{x} - b_t\right)_+,\ t \in T_j \end{array} \right\}.$$

4 Best least squares solutions

Concerning the existence, Han's original proof in [12] as also the new proofs in [5] (all for the discrete case) are not easily adaptable to continuous systems. We therefore just propose a characterization of solutions.

Proposition 4 (Characterization of L_2 solutions) *A given $\overline{x} \in \mathbb{R}^n$ is a least squares solution if and only if*

$$\sum_{j=1}^{p} \int_{T_j} \left(a_t^\top \overline{x} - b_t\right)_+ a_t dt_1 ... dt_m + \sum_{k=1}^{q} \left(a_{t^k}^\top \overline{x} - b_{t^k}\right)_+ a_{t^k} = 0_n. \tag{14}$$

Proof We can assume without loss of generality that $p \geq 1$ and $q \geq 1$. Denoting by $p_+ : \mathbb{R} \to \mathbb{R}$ the *positive part function*, *i.e.*, $p_+(y) = \max\{y, 0\}$, and by $h_t : \mathbb{R}^n \to \mathbb{R}$ the affine function such that $h_t(x) = a_t^\top x - b_t$, $t \in T$, we can write

$$f_2(x) = \frac{1}{2} \sum_{j=1}^{p} \int_{T_j} \left(p_+^2 \circ h_t\right)(x)\, dt_1 ... dt_m + \frac{1}{2} \sum_{k=1}^{q} \left(p_+^2 \circ h_{t^k}\right).$$

Obviously, p_+^2 is convex and differentiable, with $\frac{dp_+^2(y)}{dy} = 2p_+(y)$ for all $y \in \mathbb{R}$ while $\nabla h_t(x) = a_t$ for all $x \in \mathbb{R}^n$.

Let $t \in T_j$, with $j \in \{1, ..., p\}$. Then $x \mapsto \displaystyle\int_{T_j} \left(p_+^2 \circ h_t\right)(x)\, dt_1 ... dt_m$ is convex and differentiable, with gradient

$$\nabla \left(\int_{T_j} \left(p_+^2 \circ h_t\right) \right)(x) = 2 \int_{T_j} \left(a_t^\top x - b_t\right)_+ a_t dt_1 ... dt_m, \tag{15}$$

provided by the Leibniz integral rule for derivation under the integral sign, taking into account that the partial derivatives or gradients are continuous and bounded by integrable functions (see, *e.g.*, [19, Section 14.2]).

Analogously,

$$\nabla \left(p_+^2 \circ h_{t^k}\right)(x) = 2 \left(a_{t^k}^\top x - b_{t^k}\right)_+ a_{t^k},\ k = 1, ..., q. \tag{16}$$

The conclusion follows from (15), (16), and the well-known coincidence of global minima and critical points for convex differentiable functions.

5 Applications

Linear SIP problems frequently arise in applications, in many cases with constraint systems which can be consistent or not depending on the data (see, *e.g.*, [11] and references therein). One of the fields where inconsistent linear systems arise more frequently in practice is robust linear optimization, which provides a deterministic framework for uncertain problems, as large uncertainty sets may provide inconsistent robust counterparts. The next examples illustrates this situation in robust production planning.

Example 5 The basic production planning model consists of maximizing the cash-flow $c(x_1, ..., x_n)$ of the total production, with x_i denoting the production level of the i-th commodity, and the decision vector $x = (x_1, ..., x_n)$ must satisfy p linear constraints $a_j^\top x \leq b_j$, where the components of a_j are the technological coefficients while b_j represents the available amount of the j-th resource. In practical situations the coefficients of the constraints (except the positivity constraints $x_i \geq 0$) are uncertain while the objective function c is deterministic. The robust optimization approach provides a deterministic framework for uncertain problems (see, e.g., [2], [3] and references therein). Following this conservative approach, the input data, (a_j, b_j), $j \in J$, are uncertain vectors and $(a_j, b_j) \in T_j \subset \mathbb{R}^{n+1}$, where the sets T_j are specified uncertainty sets. For the sake of simplicity, assume that all the uncertainty sets are boxes, i.e., that each coefficient of $a_j^\top x \leq b_j$ takes values in a given interval in $\mathbb{R}$. By enforcing the constraints for all possible uncertainties within T_j, $j \in J$, the uncertain production planning problem is captured by the so-called *robust counterpart*

$$\begin{aligned} P_R : \text{Max}_{x\in\mathbb{R}^n} \quad & c(x) \\ \text{subject to } & t_1x_1 + ... + t_nx_n \leq t_{n+1}, \quad t \in \bigcup_{j=1}^{p} T_j, \\ & -x_k \leq 0, k = 1, ..., n, \end{aligned} \tag{17}$$

which is a linear SIP problem whose mixed constraint system can be written as $\left\{a(t)^\top x \leq b(t)\,,\ t \in T\right\}$ in (1), with $T = \bigcup_{j=1}^{p} T_j \cup \left\{t^1, ..., t^n\right\} \subset \mathbb{R}^{n+1}$, $t^k := (-e_k, 0)$, where e_k denotes the $k-$th element of the canonical basis of $\mathbb{R}^n$, $k = 1, ..., n$, $a(t) = (t_1, ..., t_n)^\top$, and $b(t) = t_{n+1}$ for all $t = (t_1, ..., t_{n+1}) \in T$. We may assume that the boxes $T_1, ..., T_p$ are pairwise disjoint and

$$\left\{t^1, ..., t^n\right\} \cap \left(\bigcup_{j=1}^{p} T_j\right) = \emptyset \tag{18}$$

(otherwise, replace T_j with $T_j \times \{j\}$, $j = 1, ..., p$, and t^k with $\left(t^k, p + k\right)$, $k = 1, ..., n$). When the the union of $\left\{t^1, ..., t^n\right\}$ with the boxes T_j, $j =$

$1, ..., p$, is too large, we may have

$$\begin{pmatrix} 0_n \\ -1 \end{pmatrix} \in \operatorname{cone}\left(\left\{t^1, ..., t^n\right\} \cup \bigcup_{j=1}^{p} T_j\right),$$

so that, by (4), P_R is inconsistent, and the decision maker has two options: either reducing the length of the uncertainty intervals (bounds for these lengths can be found in [8] and references therein) or, assuming a minimum risk, select a best approximate solution for the constraint system of P_R.

The practical application of the previous results depends on the tractability of the problems P_0, P_1, and P_2.

When the given inconsistent system is discrete, P_0 can be reformulated as a linear program P_0' that can efficiently be solved by any of the well-known simplex-like or interior-point methods; when the system is continuous, the linear semi-infinite program P_0' can be solved by simplex-like, cutting-plane, and grid discretization methods; finally, when the system is mixed, only grid discretization methods are viable, taking into account that $\left\{t^1, ..., t^q\right\}$ should be part of any grid (a brief survey of these methods can be found in [10, Chapter 1]). In many economic problems (production planning, allocation of resources, portfolio, etc.) with interval uncertain constraints $a_j^\top x \leq b_j$, all the sets $\left\{(a_t, b_t),\ t \in T_j\right\}$, $j = 1, ..., p$, are boxes, in which case P_0' can be reformulated as an ordinary linear program following the same strategy as in Example 4 (by eliminating redundant constraints).

The main difficulty with solving P_1 and P_2 analytically, in the continuous and mixed cases, is that one can hardly get explicit formulas for the integrals of $\left(a_t^\top x - b_t\right)_+$ and its square in terms of the variables $x_1, ..., x_n$, although they can be easily evaluated for particular values of $x = (x_1, ..., x_n)$ by getting a convenient representation of the polytope $\left\{t \in T_j : a_t^\top x - b_t \geq 0\right\}$ once checked that $a_t^\top x - b_t \geq 0$ for some extreme point of T_j; otherwise

$$\int_{T_j} \left(a_t^\top x - b_t\right)_+ dt_1...dt_m = \int_{T_j} \left[\left(a_t^\top x - b_t\right)_+\right]^2 dt_1...dt_m = 0.$$

For instance, in Example 4, for $x = (0, 2)$, $a_t^\top x - b_t = 2t_2 - t_3 \geq 0$ for $(1, 1, -1) \in T_1$,

$$\left\{t \in T_1 : a_t^\top x - b_t \geq 0\right\} = [-3, -1] \times \left\{(t_2, t_3) \in [-1, 1]^2 : t_2 \geq \frac{t_3}{2}\right\},$$

and

$$\begin{aligned} \int_{T_1} \left(a_t^\top x - b_t\right)_+ dt_1 dt_2 dt_3 &= \int_{-3}^{-1} dt_1 \int_{-1}^{1} dt_3 \int_{\frac{t_3}{2}}^{1} (2t_2 - t_3)\, dt_2 \\ &= \int_{-3}^{-1} dt_1 \int_{-1}^{1} \left(\tfrac{t_3^2}{4} - t_3 + 1\right) dt_3 = \tfrac{13}{3}. \end{aligned}$$

The advantage of P_2 over P_1 consists of the differentiability of the objective function f_2 and the fact that its gradient (the function in the left-hand side of (14) is Lipschitz continuous, property guaranteeing the convergence of the steepest descent method, the quadratic convergence of Newton's method, etc.

6 Conclusions

We have analyzed in this paper different aspects of the best approximation problem for inconsistent continuous and mixed linear systems when the infeasibility is measured with either the uniform norm, or the Euclidean norm, or the L_1 norm, which are now compared with different criteria:

– Robustness (in the sense of sensitivity with respect to error data): Due to the presence of square terms in f_2, least squares solutions are more sensible to error data (the coefficients of the systems) than best uniform and L_1 solutions. This is the same situation as in regression analysis and other branches of statistics.
– Tractability: Thanks to the differentiability of f_2, least squares solutions are more easily computable by the numerical methods for unconstrained programs than best uniform and L_1 solutions. However, best uniform solutions can be approximated by using efficient linear SIP numerical methods when m is sufficiently small.
– Characterization of solutions: We have provided characterizations of the best approximate solutions for the three norms in terms of the data, but the simplest one corresponds to best uniform solutions. So, best approximate solutions are preferable from the stopping rule perspective.
– Existence of solutions: We have characterized the existence of best uniform solutions; providing existence theorems for least squares solutions and best L_1 solutions remains an open problem.
– Generality: The results on best uniform solutions are valid for systems with an arbitrary index set T (not necessarily the union of a finite set with a family of boxes).

References

1. Amaral, P., Júdice, J., Sherali, H.D.: A reformulation-linearization-convexification algorithm for optimal correction of an inconsistent system of linear constraints. Comput. Oper. Res. **35**, 1494-1509 (2008)
2. Ben-Tal, A., El Ghaoui, L., Nemirovski, A.: Robust Optimization. Princeton U.P. (2009)
3. Bertsimas, D., Brown, D.B., Caramanis, C.: Theory and applications of robust optimization. SIAM Rev. **53**, 464–501 (2011)
4. Cánovas, M.J., López, M.A. Parra, J., Toledo, F.J.: Distance to ill-posedness for linear inequality systems under block perturbations: convex and infinite-dimensional cases. Optimization **60**, 925-946 (2011)

5. Contesse, L., Hiriart-Urruty, J.-B., Penot, J.-P.: Least squares solutions of linear inequality systems: a pedestrian approach. RAIRO. Published on line, June 2016. DOI: http://dx.doi.org/10.1051/ro/2016042
6. Dax, A.: A hybrid algorithm for solving linear inequalities in a least squares sense. Numer. Algor. **50**, 97-114 (2009)
7. Folland, G.B.: Real Analysis: Modern Techniques and Their Applications (2nd ed.). Wiley (1999)
8. Goberna, M.A., Jeyakumar, V., Li, G., Vicente-Pérez, J.: Robust solutions to multi-objective linear programs with uncertain data. European J. Oper. Res. **242**, 730-743 (2015)
9. Goberna, M.A., López, M.A.: Linear Semi-Infinite Optimization. Wiley (1998)
10. Goberna, M.A., López, M.A.: Post-Optimal Analysis in Linear Semi-Infinite Optimization. Springer (2014)
11. Goberna, M.A., López, M.A.: Recent contributions to linear semi-infinite optimization. 4OR, to appear.
12. Han, S.P.: Least-squares solution of linear inequalities. Tech. Rep. TR-2141, Mathematics Research Center, University of Wisconsin-Madison (1980)
13. Hiriart Urruty, J.-B., Lemaréchal, C.: Convex Analysis and Minimization Algorithms I, Springer (1993)
14. Ioffe, A.D., Tihomirov V.M.: Theory of Extremal Problems, North-Holland (1979)
15. Jones, D.S.: The Theory of Generalized Functions (2nd. ed.), Cambridge U.P. (1982)
16. Le, N.Z.: Correction of inconsistent systems of linear inequalities with matrices of block structure by the minimax criterion (Russian). Vestnik Moskov. Univ. Ser. XV Vychisl. Mat. Kibernet **55**, 18-25 (2011); translation in Moscow Univ. Comput. Math. Cybernet. **3**5, 167-175 (2011)
17. Lei, Y.: The inexact fixed matrix iteration for solving large linear inequalities in a least squares sense. Numer. Algor. **69**, 227-251 (2015)
18. Mangasarian, O.L.: Error bounds for inconsistent linear inequalities and programs. Oper. Res. Lett. **15**, 187-192 (1994)
19. Mawhin, J., Analyse: Fondements, Techniques, Évolution (in French). De Boeck Université (1992)
20. Munkres, J.R.: Analysis on Manifolds. Westview Press (1991)
21. Polyak, B.T.: Sharp minima. Institute of Control Sciences, USSR, Lecture Notes. Presented at the IIASA Workshop on Generalized Lagrangians and their Applications, IIASA, Laxenburg, Austria (1979)
22. Popa, C., Şerban, C.: Han-type algorithms for inconsistent systems of linear inequalities-a unified approach. Appl. Math. Comput. **246**, 247-256 (2014)
23. Rockafellar, R.T.: Convex Analysis. Princeton U.P. (1970)
24. Rockafellar, R.T., Wets, R.J.-B.: Variational Analysis. Springer (1998)