A Reformulation-Linearization-Convexification Algorithm for Optimal Correction of an Inconsistent System of Linear Constraints

P. Amaral^a * J. Júdice^b H. D. Sherali^c

Submitted: November 30, 2004 Revised: May 15, 2006

Abstract

In this paper, an algorithm is introduced to find an optimal solution for an optimization problem that arises in total least squares with inequality constraints, and in the correction of infeasible linear systems of inequalities. The stated problem is a nonconvex program with a special structure that allows the use of a reformulation—linearization—convexification technique for its solution. A branch—and—bound method for finding a global optimum for this problem is introduced based on this technique. Some computational experiments are included to highlight the efficacy of the proposed methodology.

keywords: Infeasible systems; Nonlinear programming; Global optimization.

1 Introduction

The problem we address in this paper arises in the context of correcting infeasible linear systems of inequalities, such as $Ax \leq b, x \in \mathbb{R}^n$, where $A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m$.

In linear programming and in constraint satisfaction, one is often confronted with an empty set of solutions due to many causes, such as the lack of communication between different decision makers, update of old models, or integration of partial models. In post-infeasibility analysis, several attempts are made in order to retrieve valuable information regarding the inconsistency of a given model [11],[15],[16],[17],[19], such as the identification of conflicting sets of constraints [1],[2],[10],[12],[14] and irreducible inconsistent sets (IIS) of constraints, for both continuous [13],[18],[23],[31] and mixed-integer [20] problems. This information can be used to reformulate the model, either by removing constraints or slightly changing the coefficients of the constraints. In [21], the authors proposed a method based on an hierarchical classification of constraints to remove constraints in order to obtain a feasible set. This procedure is, however, inadequate in cases where the physical events that the constraints seek to prevent cannot be ignored, or when we are working with only approximate data. In the context of linear programming, some theoretical results regarding the distance to infeasibility of a linear system are presented in [25] and [26], and techniques for deciding about the existence of solutions for approximate data are explored in [28].

Although the problem of inconsistency in linear models has attracted some attention, not much has been done concerning the development of exact algorithms for finding minimal corrections

 ^a Departamento de Matemática, Universidade Nova de Lisboa, Monte da Caparica, 2829-516 Caparica, Portugal
 ^b Departamento de Matemática, Universidade de Coimbra and Instituto Telecomunicações, Polo II, 3030-290
 Coimbra, Portugal

^c Department of Industrial & Systems Engineering, Virginia Polytechnic Institute & State University, 547 Whittemore Hall, Blacksburg, VA 24061, USA

^{*}Corresponding author: E-mail address: paca@fct.unl.pt, Fax: +351-21-2948391

(according to some criteria) of the coefficients of an infeasible linear system of inequalities. The perturbation of the vector b alone is a less difficult problem and has been considered in [27]. The correction of both A and b is a more challenging task, due to the introduction of nonlinearity, but is more adequate and advisable in practice. Let us consider, for instance, a production model where x represents the production of some products, the matrix A represents both the consumption of raw material per unit and certain cost and sale price coefficients, and b is the amount of available raw material together with a desired level of profit. In such a situation it is plausible in the case of inconsistency, to not only to rectify the amount of available raw material and the target profit (right-hand-side values), but also the specifications of the production process and the sale price. Even if our primary goal is not to derive a modification of the model, it can be useful to find a feasible model that is closest to the infeasible one in some sense. Furthermore, if the original system describes some ideal properties of a new product or the behavior of some structure, in case of inconsistency of the model, it would be of importance to understand how the paradigms we are seeking must be minimally adapted for the purpose of an implementation. The solution of a feasible perturbed model can be insightful. The smaller the difference between the infeasible and the modified feasible model, the greater the possibility that the suggested changes can be realistically implemented. The knowledge of a feasible solution for a corrected model can provide insights into the nature of the infeasibility and the way to overcome it. We can even analyze only a subset of constraints, as for example, an IIS or a set of soft-constraints in such an analysis to either focus attention on different subsets of inconsistent constraints or on manipulating parameters of flexible relationships. In such contexts, a useful piece of information (among others) to provide the decision-maker would be an optimally perturbed feasible model, in the sense of minimizing a measure of distance between the set of parameters of the infeasible and corrected constraints.

These considerations lead to a general formulation of the following optimization problem that seeks the optimal correction p and H, respectively, of the vector b and the matrix A of the given linear system of inequalities $Ax \leq b$:

$$Minimize \varphi(H, p) (1)$$

subject to
$$(A+H)x \le b+p$$
 (2)

$$x \in \mathbf{X}, H \in \mathbb{R}^{m \times n}, p \in \mathbb{R}^m, \tag{3}$$

where $\mathbf{X} \subseteq \mathbb{R}^n$ is a convex set and φ is an appropriate matrix norm. For $\varphi = \|\cdot\|_{l_1}$ and $\varphi = \|\cdot\|_{l_\infty}$ (the generalization for matrices of the respective norms l_1 and l_∞), Vatolin [30] proved that it is possible to find an optimal correction by solving a set of linear programming problems. Later in [3], it was shown that this approach is also applicable for the ∞ -norm l_∞ , and that the number of linear programming problems to be solved is 2n+1 for the l_1 and l_∞ norms, and l_∞ for the l_∞ norm. Furthermore, it was required that an optimal correction should involve changes in only one column of l_∞ in the case of norms l_1 or l_∞ , while for the norm l_∞ , the perturbation of the coefficients of every row should only differ in sign. This introduced a fixed pattern for the correction matrix, which turns out to be quite unnatural in practical situations where a free pattern is more suitable. These conclusions have motivated us to study the case of finding an optimal correction with respect to the Frobenius norm l_∞ , that is, to consider the following optimization problem:

1
$$||A||_{l_1} = \sum_{ij} |a_{ij}| , ||A||_{l_{\infty}} = \max_{ij} |a_{ij}|$$
 2
$$||A||_{\infty} = \max_i \sum_j |a_{ij}|$$

(P) Minimize $||[H,p]||_F^2$

subject to
$$(A+H)x \le b+p$$

 $H \in \mathbb{R}^{m \times n}, \quad p \in \mathbb{R}^m, \quad x \in \mathbf{X}.$ (4)

The interest here is not as much to directly diagnose what was the underlying problem in formulating the infeasible model, but rather, to provide this insight indirectly through the solution of problem (4), which yields the least Frobenius norm correction to the constraint matrix and right—hand—side in order to attain feasibility.

It is interesting to note that for $\mathbf{X} = \mathbb{R}^n$ Problem (4) may fail to have a solution, and that a local minimizer exists iff the correction corresponds to an application of the total least squares (TLS) method to the set of active constraints [6]. Some algorithms for finding such a local minimizer are discussed in [4]. In [7], a tree search procedure based on the enumeration of the active set of constraints was proposed, where some reduction tests were implemented in order to reduce the tree search. Although problems of small size have been efficiently solved, the overall effort required for finding a global minimum is usually too high for medium-scale problems. In practice, it is important to define the set \mathbf{X} such that the existence of an optimal solution is guaranteed, and that the solution of the corrected system is in a certain domain of interest. The most general choice for \mathbf{X} corresponds to

$$\mathbf{X} \subseteq \{x \in \mathbb{R}^n : l \le x \le u\},\tag{5}$$

where l and u are fixed vectors. Then, \mathbf{X} is compact and the optimization problem (P) has a global minimum. This is the choice we adopt in this paper.

This paper is organized as follows. In the next section, we present some important results that lead to the development of the main algorithm to be introduced in Section 3. Section 4 includes the report of some experiments with the algorithm for a number of test instances. Some conclusions and recommendations for future research are provided in the closing section of the paper.

2 Preliminary results

In [6] it was shown that for $\mathbf{X} = \mathbb{R}^n$, the problem (4) is equivalent to the unconstrained nonlinear and nonconvex problem:

(P)
$$\min_{x \in \mathbf{X}} \frac{\|(Ax - b)^{+}\|^{2}}{1 + \|x\|^{2}},$$
 (6)

where $(\cdot)^+$ denotes a component-wise application of the operator $\max\{0,\cdot\}$ and $\|\cdot\|$ represents the Euclidean norm. This equivalence still holds for

$$\mathbf{X} \subseteq \{x : l \le x \le u\}.$$

Once Problem (6) is solved, we can directly obtain an optimal correction matrix [H, p] for Problem (4) via

$$[H, p] = -\lambda^*[x^{*T}, 1],$$
 (7)

where x^* is the optimal solution found for Problem (6) and

$$\lambda^* = \frac{1}{1 + \|x^*\|^2} (Ax^* - b)^+. \tag{8}$$

The proof can be found in [6]. The next theorem presents another formulation for the original problem (4) that is exploited in this paper.

Proposition 2.1. The formulation (6) is equivalent to:

$$(P1): Minimize \frac{\beta}{1+\|x\|^2}$$
 (9)

subject to
$$\beta \ge ||v||^2$$
 (10)

$$v \ge Ax - b \tag{11}$$

$$v \ge 0 \tag{12}$$

$$x \in \mathbf{X}$$
. (13)

Proof: It is sufficient to show that for any feasible solution to one problem, there exists a feasible solution to the other problem having at least as good an objective value.

Given x feasible to (P), it is easy to see that there exists (x, v, β) feasible to (P1) with the same objective value. In fact for $v = (Ax - b)^+$ and $\beta = ||v||^2$, we get

$$\frac{\beta}{1+\|x\|^2} = \frac{\|(Ax-b)^+\|^2}{1+\|x\|^2}.$$

Now, for (x, v, β) feasible to (P1), let $\overline{v} = (Ax - b)^+$ and $\overline{\beta} = ||\overline{v}||^2$. Then $(x, \overline{v}, \overline{\beta})$ is feasible to (P1) and

$$\frac{\overline{\beta}}{1+\|x\|^2} = \frac{\|\overline{v}\|^2}{1+\|x\|^2} = \frac{\|(Ax-b)^+\|^2}{1+\|x\|^2}
\leq \frac{\|v\|^2}{1+\|x\|^2} \leq \frac{\beta}{1+\|x\|^2}.$$
(14)

The first inequality in (14) follows because $Ax - b \le v$ and $v \ge 0$ imply that $0 \le (Ax - b)^+ \le v$ and so, $||(Ax - b)^+||^2 \le ||v||^2$. Also, x is feasible to (P) and (14) shows that it has at least as good an objective value in Problem P as (x, v, β) does in Problem P1. This completes the proof. \square

In (P1), upon using the substitution

$$||x||^2 = \alpha,$$

and considering that $\mathbf{X} \subseteq \{x : 0 \le l \le x \le u\}$, we immediately obtain the following result.

Corollary 2.1. If

$$\alpha_l = ||l||^2 \text{ and } \alpha_u = ||u||^2,$$
 (15)

then the problem

$$(P2) \quad \textit{Minimize} \quad \frac{\beta}{1+\alpha} \tag{16}$$

subject to
$$\beta \ge ||v||^2$$
 (17)

$$v \ge Ax - b \tag{18}$$

$$||x||^2 = \alpha \tag{19}$$

$$x \in \mathbf{X} \subseteq \{x : 0 \le l \le x \le u\} \tag{20}$$

$$v \ge 0 \tag{21}$$

$$\alpha_l \le \alpha \le \alpha_u, \tag{22}$$

is equivalent to (P1) and, consequently, to Problem (4).

Now, consider the nonlinear programming relaxation to Problem P2:

$$(RP2)$$
 Minimize $\frac{\beta}{1+\alpha}$ (23)

subject to
$$\beta \ge ||v||^2$$
 (24)

$$v \ge Ax - b \tag{25}$$

$$||x||^2 \le \alpha \tag{26}$$

$$x \in \mathbf{X} \subseteq \{x : 0 \le l \le x \le u\} \tag{27}$$

$$v \ge 0 \tag{28}$$

$$\alpha_l \le \alpha \le \alpha_u. \tag{29}$$

Then the feasible region of (RP2) is convex and the objective function is pseudoconvex over this set [8]. Therefore, any KKT (Karush-Kuhn-Tucker) solution to this problem is also a global minimum of (RP2). Furthermore, any such solution is a global minimum of the problem (P2) if and only if $||x||^2 = \alpha$. It might therefore seem that there could exist some cases where the solution to (P2) can be recovered by simply finding a global minimum for (RP2). Unfortunately, this is not usually the case, as the inequality $||x||^2 \le \alpha$ is in general inactive at such a global minimum to (RP2). This is quite understandable, as α tends to increase as much as possible in order to minimize the objective function of (RP2). So (P2) needs to be processed by a global optimization algorithm. In this paper, we propose a branch-and-bound algorithm for (P2) that is based on the idea of partitioning the set

$$\Omega = \{x : 0 \le l \le x \le u\}.$$

Each node k of the enumeration tree in this process is associated with a proper subset of Ω identified as:

$$\Omega^k = \{x : l_i^k \le x_i \le u_i^k, \text{ for } i = 1, \dots, n\},\tag{30}$$

along with the following corresponding node subproblem:

$$(P2^{k}) \quad \text{Minimize} \quad \frac{\beta}{1+\alpha}$$

$$\text{subject to} \quad \beta \geq ||v||^{2}$$

$$v \geq Ax - b$$

$$||x||^{2} = \alpha$$

$$x \in \Omega^{k}$$

$$v \geq 0$$

$$\alpha_{k}^{l} \leq \alpha \leq \alpha_{v}^{k},$$

$$(31)$$

where $\alpha_l^k = ||l^k||^2$ and $\alpha_u^k = ||u^k||^2$. At each node, instead of solving $(P2^k)$ directly, we obtain a lower bound for the optimal value of $(P2^k)$ by solving a special convex problem. To construct such a program, we can simply replace the equality $||x||^2 = \alpha$ by the inequality constraint (26). This nonlinear relaxation is denoted by $(RP2^k)$ and is obtained from (RP2) by requiring x to belong to the set Ω^k and by constraining α accordingly (as in (31)). Alternatively, we can exploit the so-called Reformulation-Linearization-Convexification Technique (RLT), as described in [29],

and consider the following relaxation

$$LB(P2^k)$$
: Minimize $\frac{\beta}{1+\alpha}$ (32)

subject to
$$\beta > ||v||^2$$
 (33)

$$v \ge Ax - b \tag{34}$$

$$\sum_{i=1}^{n} y_i = \alpha \tag{35}$$

$$x_i^2 \le y_i, \forall i = 1, \dots, n \tag{36}$$

$$[(x_i - l_i^k)(u_i^k - x_i)]_L \ge 0, \forall i = 1, \dots, n$$
(37)

$$x \in \mathbf{X} \tag{38}$$

$$x \in \Omega^k \tag{39}$$

$$\alpha_l^k \le \alpha \le \alpha_u^k,\tag{40}$$

where $[\cdot]_L$ denotes the linearization of the product term $[\cdot]$ under the substitution

$$y_i = x_i^2, \ \forall i = 1, \dots, n. \tag{41}$$

Note that in (37), known as bound-factor-constraints [29],

$$[(x_i - l_i^k)(u_i^k - x_i)]_L = [x_i u_i^k - l_i^k u_i^k + l_i^k x_i - x_i^2]_L$$

= $x_i u_i^k - l_i^k u_i^k + l_i^k x_i - y_i$.

Problem $LB(P2^k)$ is a convex nonlinear program with a very special structure that can be efficiently solved by a nonlinear programming algorithm. It should be added that the constraint $\alpha = ||x||^2$ has been convexified by introducing new variables y_i and additional constraints (35)-(37), where (36) and (37) approximate the nonconvex relationship $y_i = x_i^2$, $\forall i = 1, ..., n$. The following result holds in regard to $LB(P2^k)$.

Proposition 2.2. If $(\overline{x}, \overline{v}, \overline{\beta}, \overline{\alpha}, \overline{y})$ solves Problem $LB(P2^k)$ with objective value $\nu(LB(P2^k))$, then $\nu(LB(P2^k))$ is a lower bound for the optimal value of $(P2^k)$. Moreover, if $\overline{x}_i = l_i^k$ or $\overline{x}_i = u_i^k$, for each $i = 1, \ldots, n$, then $\overline{y}_i = \overline{x}_i^2$, $\forall i = 1, \ldots, n$.

It is also important to add that the relaxation $LB(P2^k)$ provides tighter lower bounds than the previous one $(RP2^k)$. In fact, from (35) and (36), we get that for any feasible solution to $LB(P2^k)$.

$$||x||^2 = \sum_{i=1}^n x_i^2 \le \sum_{i=1}^n y_i = \alpha.$$

Therefore the feasible region of the problem $LB(P2^k)$ projected onto the (x, v, β, α) -space is included in that for $(RP2^k)$. For this reason, in this paper, we use the relaxation $LB(P2^k)$ instead of $(RP2^k)$.

3 Overall algorithm

At each node of the proposed branch-and-bound procedure, starting with $\Omega^0 = \Omega$, the problem $LB(P2^k)$ is solved to derive a lower bound on the node subproblem. Let $(\overline{x}, \overline{v}, \overline{\beta}, \overline{\alpha}, \overline{y})$ be the optimal solution obtained and let $\nu(LB(P2^k))$ be its optimal value. If

$$\nu(LB(P2^k)) \ge UB(1 - \epsilon) \tag{42}$$

for some tolerance $\epsilon \geq 0$, we fathom this node. Otherwise, denoting K as the total number of nodes enumerated thus far, we partition this node into subproblems $(P2^{K+1})$ and $(P2^{K+2})$, based on the corresponding partition of Ω^k into Ω^{K+1} and Ω^{K+2} as follows:

$$\Omega^{K+1} = \{x : l_i^{K+1} \le x_i \le u_i^{K+1}, \text{ for } i = 1, \dots, n\}$$

and

$$\Omega^{K+2} = \{x : l_i^{K+2} \le x_i \le u_i^{K+2}, \text{ for } i = 1, \dots, n\},\$$

where the bounds describing Ω^{K+1} and Ω^{K+2} are discussed next, within the following proposed branching strategy.

3.1 Branching variable selection scheme

Let

$$p \in \arg\max_{i=1,\dots,n} \{\theta_i\}$$
 where $\theta_i = \overline{y}_i - \overline{x}_i^2$, for $i = 1,\dots,n$. (43)

If $\theta_p > 0$, then

$$\Omega^{K+1} = \{ x : l_i^k \le x_i \le u_i^k, i = 1, \dots, n, i \ne p, l_p^k \le x_p \le \overline{x}_p \},$$
(44)

$$\Omega^{K+2} = \{x : l_i^k \le x_i \le u_i^k, i = 1, \dots, n, i \ne p, \overline{x}_p \le x_p \le u_p^k \}.$$
(45)

If $\max_{i=1,\dots,n}\{\theta_i\}=0$, then by Proposition 2.2, $(\overline{x},\overline{v},\overline{\beta},\overline{\alpha})$ is a feasible solution for $(P2^k)$, whose value equals $\nu(LB(P2^k))$, a lower bound that is achieved, and hence, is an optimal solution to $(P2^k)$. Therefore, we update the incumbent solution (along with UB), if necessary, and fathom the node in this case.

3.2 Computing Upper Bounds

Given the solution $(\overline{x}, \overline{v}, \overline{\beta}, \overline{\alpha}, \overline{y})$ to $LB(P2^k)$, we compute an upper bound UB by setting

$$UB \leftarrow \min\{UB, f(\overline{x}), \phi(\overline{x}, \overline{\beta}, \overline{v})\}$$
 (46)

where

$$f(\overline{x}) = \frac{\|(A\overline{x} - b)^+\|^2}{1 + \|\overline{x}\|^2},$$

and

$$\phi(\overline{x}, \overline{\beta}, \overline{v}) = \frac{\widehat{\beta}}{1 + \|\widehat{x}\|^2},$$

where $(\widehat{x}, \widehat{\beta}, \widehat{v})$ is a stationary point to (P1) that is obtained by starting at the initial point $(\overline{x}, \overline{\beta}, \overline{v})$ and applying a nonlinear programming algorithm, (we used GAMS-MINOS [9],[24] for this purpose).

Remark 3.1. As is reported in the Section 4, it is further advisable to compute an upper-bound at the initial root node by obtaining a stationary point to the following equivalent representation of the original problem

(P3): Minimize
$$\frac{\|v\|^2}{1+\|x\|^2}$$
 (47)

subject to
$$v \ge Ax - b$$
 (48)

$$v \ge 0 \tag{49}$$

$$x \in \mathbf{X}.\tag{50}$$

3.3 Computing an initial feasible solution

In order to induce a faster convergence toward optimality for each subproblem $LB(P2^k)$, it is important to start with an initial point, $(\hat{x}, \hat{v}, \hat{\beta}, \hat{\alpha}, \hat{y})$, say, that is feasible for $LB(P2^k)$. To do this, we solve the following program:

$$\hat{x} = \arg \min_{x \in \mathbf{X} \bigcap \{x: l^k \le x \le u^k\}} \|x - x^*\|_{l_1}$$
(51)

where x^* is the current best known solution. Note that for **X** polyhedral (including **X** = \mathbb{R}^n), (51) can be solved via the LP:

Minimize
$$\left\{ \sum_{i=1}^{n} z_i : z_i \ge x_i - x_i^* \text{ and } z_i \ge x_i^* - x_i, \ \forall i = 1, \dots, n, \ x \in \mathbf{X}, l^k \le x \le u^k \right\}.$$
 (52)

After obtaining \hat{x} , we then compute the remainder of the initial solution as

$$\hat{v} = \max\{0, A\hat{x} - b\},
\hat{\beta} = \|\hat{v}\|^2,
\hat{y}_i = \hat{x}_i^2, \forall i = 1, ..., n,
\hat{\alpha} = \|\hat{x}\|^2.$$
(53)

The following result holds.

Proposition 3.1. The solution $(\hat{x}, \hat{v}, \hat{\beta}, \hat{\alpha}, \hat{y})$ as given by (51)-(53) is feasible to $LB(P2^k)$.

Proof Feasibility to (33),(34),(35),(36),(38),(39), and (40) is evident, by the construction of the solution (51)-(53) and the bounds derived based on $l^k \leq x \leq u^k$. Moreover, feasibility to the RLT constraints (37) follows by construction, since the product relationships in (41) are satisfied via (53). This completes the proof.

3.4 Algorithm and Convergence Theorem

In this subsection, we describe the main steps of the algorithm. To do this, we start by describing the following parameters:

k- index for the current subproblem under analysis;

K- total number of nodes enumerated (in addition to the root node);

UB- best known upper bound;

 x_{inc} - incumbent solution;

L- queue of indices of subproblems created but not expanded;

 $LB(P2^k)$ - lower bounding problem as described by (32)-(40);

 $(x^k, v^k, \beta^k, \alpha^k, y^k)$ - optimal solution obtained for $LB(P2^k)$;

 Ω^k - as defined in (30);

 ϵ - optimality tolerance;

 $\nu(.)$ - optimal value of problem (.).

Algorithm RLT-BB

- **0-(Initialization)** Let K = k = 0, $L = \emptyset$, $UB = \infty$, and $\epsilon \ge 0$ (we chose $\epsilon = 10^{-6}$). Solve Problem $LB(P2^k)$. Update UB through (46) and set x_{inc} equal to the corresponding best x-solution found.
- **1-(Pick next node)** If $L = \emptyset$ then stop; otherwise, find $k \in \arg\min\{\nu(LB(P2^t)) : t \in L\}$.
- **2-(Dequeue)** Set $L \leftarrow L \{k\}$.
- **3-(Branching rule)** Find a branching index p via (43). If $\theta_p > 0$, go to Step 4. Otherwise, update UB and x_{inc} using the solution to $LB(P2^k)$, remove any indices t from L for which $\nu(LB(P2^t)) \geq UB(1-\epsilon)$, and go to Step 1.
- 4-(Solve, Update, and Queue) Set i = 1.
 - **4.1** Define Ω^{K+i} according to (44)-(45). Solve Problem $LB(P2^{K+i})$.
 - **4.2** If $\nu(LB(P2^{K+i})) < UB(1-\epsilon)$ then go to Step 4.3; otherwise, go to Step 4.5.
 - **4.3** Update UB according to (46). If UB was updated remove all indices $t \in L$ for which $\nu(LB(P2^t)) \geq UB(1-\epsilon)$ and put x_{inc} equal to the revised incumbent x-solution found.
 - **4.4** Set $L \leftarrow L \cup \{K + i\}$.
 - **4.5** If i=2, set $K \leftarrow K+2$ and go to Step 1; otherwise, let i=2 and go to Step 4.1.

To complete this section we present the convergence theorem for this algorithm.

Proposition 3.2. The algorithm RLT-BB, when run with $\epsilon \equiv 0$, either terminates finitely with a global optimum to the problem, or else an infinite branch-and-bound tree is generated, such that any accumulation point of the relaxation lower-bounding problem solution along any infinite branch of the enumeration tree is a global optimum for problem (P2).

Proof The case of finite convergence is obvious from the validity of the bounds derived by the algorithm. Now, suppose that an infinite branch-and-bound tree is generated. Then there exists a branching index p that is selected infinitely often along an infinite branch of this tree. Let S_1 be an infinite sequence of nested nodes generated by the sequence of such branchings over p, and let $LB(P2^k)$ and $(x^k, v^k, \beta^k, \alpha^k, y^k)$ for $k \in S_1$ be the corresponding lower bounding problems and optimal solutions, respectively. Over some convergent subsequence $(k \in S_2 \subseteq S_1)$ suppose that

$$(x^k, v^k, \beta^k, \alpha^k, y^k, l^k, u^k) \longrightarrow (x^*, v^*, \beta^*, \alpha^*, y^*, l^*, u^*).$$

$$k \in S_1$$

Then, using the proof in [29], we see that in the limit $x_p^* = l_p^*$ or $x_p^* = u_p^*$. Moreover, by Proposition 2.2, this yields $\theta_p = 0$, in the limit, where θ_i is defined in (43). Furthermore, again by (43), this gives $\theta_i = 0, \forall i = 1, \ldots, n$, in the limit. Hence, since the limiting solution $(x^*, v^*, \beta^*, \alpha^*)$ is feasible to Problem P2 with objective value V^* , we get

$$V^* \ge \nu(P2). \tag{54}$$

However, the least lower bound node selection criterion ensures that $\nu(LB(P2^k)) \leq \nu(P2), \forall k \in S_2$. In the limit, this yields $V^* \leq \nu(P2)$. This, together with (54), yields $V^* = \nu(P2)$ and so, $(x^*, v^*, \beta^*, \alpha^*)$ solves Problem P2. This completes the proof.

3.5 An example

Consider the following inconsistent system of inequalities, illustrated in Figure (1):

$$\begin{cases}
-x_1 - x_2 \leq -7 \\
x_2 \leq 3 \\
2x_1 - x_2 \leq -2.
\end{cases}$$

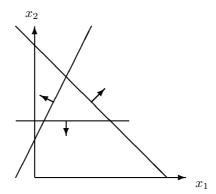


Figure 1: Inconsistent system for the illustrative example.

Suppose that we seek an optimal correction on the domain X defined by

$$1 \le x_i \le 5 \text{ for } i = 1, 2.$$

To apply the algorithm, we have:

$$l = [1, 1]$$
 and $u = [5, 5]$.

Hence,

$$\Omega^0 = \{x : 1 \le x_1 \le 5, 1 \le x_2 \le 5\},\$$

$$\alpha_l^0 = 2.0000$$
, $\alpha_u^0 = 50.0000$.

The solution to $LB(P2^0)$ is as follows:

$$x^0 = (1.6079, 4.6618)$$
 $y^0 = (4.6473, 22.9709)$ $\alpha^0 = 27.6182$ $\beta^0 = 3.6018$ $v^0 = (0.7303, 1.6618, 0.5539),$

and the lower bound is given by

$$\nu(LB(P2^0)) = 0.1259.$$

The upper bound can be updated to 0.1423 according to (46), and the corresponding incumbent solution is $x_{inc} = x^0$. In order to apply the branching rule to partition Ω^0 we get:

$$\begin{array}{rcl} \theta_1 & = & \overline{y}_1 - \overline{x}_1^2 = 2.0620, \\ \theta_2 & = & \overline{y}_2 - \overline{x}_2^2 = 1.2385, \\ \\ \max_{i=1,2} \{\theta_i\} = 2.0620 \text{ and } p = \arg\max_{i=1,2} \{\theta_i\} = 1. \end{array}$$

We thus obtain Ω^1 and Ω^2 based on the partition of $[l_1, u_1] = [1, 5]$ into $[l_1^1, u_1^1] = [1, 1.6079]$ and $[l_1^2, u_1^2] = [1.6079, 5]$:

$$\Omega^1 = \{x : 1 \le x_1 \le 1.6079, 1 \le x_2 \le 5\},\$$

$$\Omega^2 = \{x : 1.6079 \le x_1 \le 5, 1 \le x_2 \le 5\}.$$

Now, solving $LB(P2^1)$, we get:

and the upper bound UB is updated to 0.1422, with $x_{inc} = x^1$.

Likewise, $LB(P2^2)$ leads to the following solution:

The upper bound is updated to 0.1421, with $x_{inc} = x^2$. Now, K = 2, $L = \{1, 2\}$ and since

$$k \in argmin\{\nu(LB(P2^t)) : t \in L\} = 1,$$

the first subproblem is chosen to apply the branching rule and to partition the corresponding hyperrectangle Ω^1 . This yields:

$$\max_{i=1,2} \{\theta_i\} = 1.2523 \text{ and } p = \arg\max_{i=1,2} \{\theta_i\} = 2.$$

Accordingly, Ω^3 and Ω^4 are constructed based on the partition of $[l_2^1, u_2^1] = [1, 5]$ into $[l_2^3, u_2^3] = [1, 4.6576]$ and $[l_2^4, u_2^4] = [4.6576, 5]$, respectively:

$$\Omega^{3} = \{x : 1 \le x_{1} \le 1.6079, 1 \le x_{2} \le 4.6576\},$$

$$\Omega^{4} = \{x : 1 \le x_{1} \le 1.6079, 4.6576 \le x_{2} \le 5\}.$$

The solution to $LB(P2^3)$ is as follows:

Consequently, the node corresponding to subproblem $LB(P2^3)$ is fathomed. Hence, L is updated to $\{2\}$ and $LB(P2^4)$ is solved, producing the following solution:

$$\begin{array}{rclcrcl} x^4 & = & (1.5880, 4.7563) & y^4 & = & (2.5335, 22.6460) \\ \alpha^4 & = & 25.1795 & \beta^4 & = & 3.6906 \\ v^4 & = & (0.6557, 1.7563, 0.4198) & \nu(LB(P2^4)) & = & 0.1410. \end{array}$$

The upper bound is updated to 0.1412 with $x_{inc} = x^4$. Thus, K = 4, $L = \{2, 4\}$, and $LB(P2^2)$ is the next subproblem that is selected to continue the search, because

$$\begin{array}{rcl} \nu(LB(P2^2)) & = & \min\{\nu(LB(P2^2)), \nu(LB(P2^4))\} \\ & = & \min\{0.1356, 0.1410\} \\ & = & 0.1356. \end{array}$$

Now, $\max_{i=1,2} \{\theta_i\} = 1.1810$ and p=2, and proceeding as above, the partitioning of Ω^2 results in

$$\Omega^5 = \{x : 1.6079 \le x_1 \le 5, 1 \le x_2 \le 4.6790\},\$$

$$\Omega^6 = \{x : 1.6079 \le x_1 \le 5, 4.6790 \le x_2 \le 5\}.$$

Since $\nu(LB(P2^5)) = 0.1418$ is greater than the upper bound, this node is fathomed. The optimal value to $LB(P2^6)$ is $\nu(LB(P2^6)) = 0.1408$. Hence, K = 6, L is updated to $L = \{4, 6\}$. The next

subproblem to be picked for branching is $LB(P2^6)$. Now max $\{\theta_i\} = 0.1356$, p = 1 and from the partition of Ω^6 , we obtain:

$$\Omega^7 = \{x : 1.6079 \le x_1 \le 1.6483, 4.6790 \le x_2 \le 5\},\$$

$$\Omega^8 = \{x : 1.6483 \le x_1 \le 5, 4.6790 \le x_2 \le 5\}.$$

The optimal values of $\nu(LB(P2^7))$ and $\nu(LB(P2^8))$ are respectively 0.1411 and 0.1415. Since this last value is greater than the upper bound, we only include node 7 in L, which yields $L=\{4,7\}$ with K=8. This process continues until a global minimum is found. The search tree in Figure 2 indicates how the algorithm has performed in order to find such a global minimum. For each node, the optimal value of the lower bound problem (LB), the upper bound (UB), when it is updated at that node, and the optimal solution $x=(x_1,x_2)$ of the corresponding relaxation are shown. In the right-upper corner of each box (node), we indicate the order in which each node is selected from the queue L. The number appearing above each box gives the node number in the order in which it is generated. The value of $\theta=\max_{i=1,2}\{\theta_i\}$ is given to the right of each box.

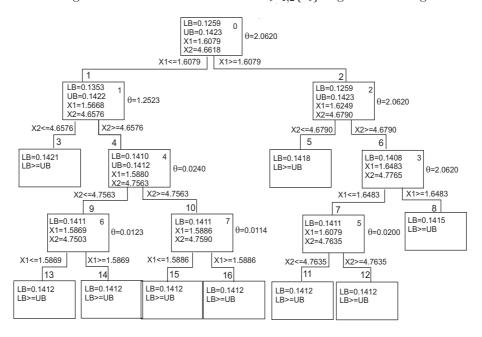


Figure 2: Graphical representation for the solution of the illustrative example.

The search inspects 16 nodes, but only 7 are introduced in the queue L for branching. The variable that induced the partition of Ω^k for each branch is also depicted in Figure 2. The optimal solution obtained is the incumbent solution corresponding to the last update of the upper bound (0.1412), which is given by $x^4 = (1.5880, 4.7563)$. The optimal correction of the matrix [A, -b], as given by (7) and (8), is

$$[H,p] = -\frac{1}{26.1441} \left[\begin{array}{c} 0.0251 \\ 0.0672 \\ 0.0161 \end{array} \right] \left[\begin{array}{cccc} 1.5880 & 4.75633 & 1 \end{array} \right] = \left[\begin{array}{ccccc} -0.03983 & -0.1193 & 0.0251 \\ -0.1067 & -0.3195 & 0.0672 \\ -0.0255 & -0.0764 & 0.0161 \end{array} \right].$$

Thus, the corrected linear system is

As expected, upon the substitution of $(x_1, x_2) = (1.5880, 4.7563)$ the inequalities are verified to be satisfied as equalities.

4 Computational experience

In order to test the performance of the algorithm we report some computational results for a set of infeasible linear systems of the type $\{x \in \mathbb{R}^n : Ax \leq b, l \leq x \leq u\}$, where A is a real matrix of order $m \times n$ and b is a real vector of size m. We consider two sets of problems. The first group is taken from a set of infeasible linear programming problems a selected from Netlib a. For the application of the algorithm in its current version, we have added lower and upper bounds (a in a and a includes the number of constraints a and variables a of the chosen Netlib problems.

Name	m	n
Galenet	10	8
Itest2	9	4
Itest6	13	8
Bgprtr	34	34
Forest6	96	95
Klein1	54	54
Woodinfe	70	89

Table 1: Caracteristics of test problems from Netlib.

The second set of test problems consists of matrices A and vectors b, such that $A^T = [A_1^T, A_2^T]$, $b^T = [b_1^T, b_2^T]^T$, so that $Ax \leq b$ is infeasible and has an optimal correction given by the Total Least Squares solution vector x_{TLS} [22], of the system $A_1x = b_1$. To construct these problems ([3], [6]) a nonsingular matrix B and a vector \hat{b} are first considered such that $Bx \leq \hat{b}$ has at least a feasible solution. Then a new constraint $\gamma^T x \leq \delta$ is added in order to render the following system infeasible:

$$A_1 x \le b_1 \Leftrightarrow \left[\begin{array}{c} B \\ \gamma^T \end{array} \right] x \le \left[\begin{array}{c} \hat{b} \\ \delta \end{array} \right].$$

A simple choice for this vector γ and scalar δ is as follows:

$$\left\{ \begin{array}{lcl} \gamma & = & B^T \psi \\ \delta & < & \psi^T b \end{array} \right.$$

where ψ is a negative random vector. It is now possible to find the Total Least Squares (TLS) solution, x_{TLS} , of the system $A_1x = b_1$. The matrix A and the vector b are then constructed by augmenting A_1 and b_1 respectively by a matrix $A_2 \in \Re^{(m-(n+1))\times n}$ and a vector $b_2 \in \Re^{m-(n+1)}$, such that x_{TLS} is feasible for the set of constraints $A_2x \leq b_2$.

Again for the application of the current version of the algorithm, we have introduced for each problem a set of finite lower and upper bounds ($l_i = 1$ and $u_i = 5$, for all \underline{i}). Table 2 summarizes the dimensions of these problems.

All the tests have been performed on a Pentium IV (Intel) with Hyperthreading, CPU 3.0 Ghz, 2Gb RAM, and operating system Linux. The method was implemented in the General Algebraic Modeling System (GAMS) language (Rev 118 Linux/Intel) and the NLP solver MINOS (version 5.5) has been used to compute the lower bounds required at each node and for computing ϕ in (46), as well as for determining a stationary point to the nonlinear program (47)-(50) that provides an upper bound at the root node.

Table 3 reports the following information for each test problem:

³Test problems collected by John W. Chinneck.

⁴The Netlib repository contains freely available software, documents, and databases. The repository is maintained by AT&T Bell Laboratories, the University of Tennessee and Oak Ridge National Laboratory among individual contributions.

Name	m	n
Pro6 to Pro10	20	10
Pro11 to Pro15	30	15
Pro16 to Pro20	40	20

Table 2: Dimension of test problems.

ND: total number of nodes in the tree;

CPU: total CPU time in seconds.

ITER: total number of MINOS iterations;

INITUB: initial upper bound (obtained at the root node);

VALOPT: optimal value;

NDOPT: node at which the optimal solution was obtained.

NUPDUB: - number of upper bound updates.

	RTLmin							
Problems	ND	CPU	ITER	INITUB	VALOPT	NDOPT	NUPDUB	
Galenet	1	0	133	3.7313	3.7313	1	0	
Itest2	27	0	751	0.4257	0.4257	1	0	
Itest6	1	0	37	82654535.9118	82654535.9118	1	0	
Bgprtr	1	0.01	194	1264.5915	1264.5915	1	0	
Forest	1	0.08	471	3458.7896	3458.7896	1	0	
Klein1	221	4.62	23919	34.6664	34.6664	1	0	
Woodinfe	1	0	8193	0.0019	0.0019	1	0	
Prob4	5	0	217	157.6815	157.6815	1	0	
Prob5	4	0	187	262.8295	262.8295	1	0	
Prob6	5	0.01	211	315.1673	315.1673	1	0	
Prob7	68	0	1196	214.8726	214.8726	1	0	
Prob8	1	0	103	2187.0132	2187.0132	1	0	
Prob9	20	0.01	522	149.3484	149.3484	1	0	
Prob10	29	0	604	250.8560	250.8560	1	0	
Prob11	26	0.01	500	396.0405	396.0405	1	0	
Prob12	3	0.01	184	1130.5302	1130.5302	1	0	
Prob13	12	0.02	489	679.4526	679.4526	1	0	
Prob14	6	0.02	382	756.9513	756.9513	1	0	
Prob15	476	1.08	23552	365.3541	364.4484	424	13	
Prob16	26	0.12	2052	1121.5359	1121.5359	1	0	
Prob17	27	0.13	1980	1092.9802	1092.9802	1	0	
Prob18	9	0.04	623	1750.9055	1738.2852	8	4	
Prob19	25	0.17	1999	1104.5614	1104.5614	1	0	
Prob20	199	1.26	18410	944.7827	944.7827	1	0	

Table 3: Computational Results

In all the runs, we have used the stopping criteria (42), with $\epsilon = 10^{-6}$. We have successfully found an optimal solution for all the test problems by using this tolerance, with the exception of

the problems Klein1, Woodinfe and Prob15. For these three problems, the fathoming criterion used was

$$UB - \nu(LB(P2^k)) \le \widetilde{\epsilon} \max\{1, UB\}.$$

The smallest values of $\tilde{\epsilon}$ that led to a successful termination of the algorithm (in fewer than 500 nodes) are given bellow.

- $\tilde{\epsilon} = 10^{-2}$ for Problem Klein1
- $\tilde{\epsilon} = 10^{-3}$ for Problem Woodinfe
- $\tilde{\epsilon} = 10^{-6}$ for Problem Prob15.

The results show that the optimal global solution is almost always (except for two of the 24 instances) the upper bound computed at the root node. This is a key feature of the tree search and substantially reduces the computational time. There are three problems for which the algorithm has been unable to terminate with the normal stopping criterion as discussed above. These three test problems indicate the need for computing better lower bounds for reducing the search. Also, the algorithm appears to solve medium-scale instances with a reasonable computational effort. Although we address in this paper a problem that is different from the one solved in [5], because no lower and upper bounds were included in the set **X** in [5], and we have used a different computational platform, we can still make some comparisons with the enumerative procedure discussed there: the algorithm introduced in this paper performs much better in terms of computational effort to find a global optimum.

5 Conclusions and future work

In this paper, we have proposed a method for obtaining an optimal solution for a nonlinear nonconvex program that arises in a total least squares approach for finding a correction of an infeasible linear system of inequalities. Using an equivalent reformulation, we solved this problem to global optimization via a new branch-and-bound algorithm. This procedure exploits the Reformulation-Linearization Convexification Technique (RLT) [29] to convexify a relaxation for deriving lower bounds on the optimal value of the original problem. Together with a framework to obtain upper bounds, we developed a tree search procedure based on a partitioning of the domain of the original variables, and established global convergence of the proposed algorithm. Computational experience reveals that the approach is suitable for handling problems having m and n ranging to about 100. This is not too restrictive, especially considering the application of this theory in the context of infeasible problem corrections. In many of these cases, we are required to maintain some constraints unchanged, and so, the nonlinear problem we formulate is defined only over a particular subset of constraints chosen by the user or by an expert system [12],[13]. Also, this formulation could be applied to Irreducible Inconsistent Systems (IIS) as identified in postoptimality analyses [23],[31], where each IIS might involve only a relatively small subset of the original LP constraints and variables.

We would also like to point out the importance of finding global optimal corrections, for instance, in the framework of Constraint Satisfaction Techniques, as in other contexts. When dealing with real models, it is essential to make as small changes as possible, in order to mitigate the risk of invalidating the corrected model. Our approach provides such a facility of determining a minimally perturbed feasible model in a least squares sense.

The solution of problems of this type with other definitions for the compact set \mathbf{X} is a useful topic for future investigation. We also recommend considering a partitioning of the constraints into two groups, namely, soft- and hard-constraints, where the set of hard-constraints is assumed to be invariant and cannot be corrected. This situation is typical in the analysis of problems that arise in real-life applications.

Acknowledgments: This research has been partially supported by the *National Science Foundation*, under Grant Number DMI 0094462, by project *FCT-POCTI/*35059/MAT/2000, Portugal, and by the *Centro de Matemática e Aplicações (CMA) FCT UNL* Portugal. We would also like to acknowledge Luís M. Fernandes for his help on the computational experience and the referees for their valuables comments and suggestions.

References

- [1] Amaldi E,Kann V. The complexity and approximability of finding maximum feasible subsystems of linear relations. Theoretical Computer Science 1995; 147:181–210.
- [2] Amaldi E, Pfetsch M E, Trotter J L E. On the maximum feasible subsystems, IISs and IIS-hypergraphs. Mathematical Series A 2003; 95:533–54.
- [3] Amaral P. Contribuições para o estudo de sistemas lineares inconsistentes. PhD Disertation, Faculty of Science and Technology, UNL, Lisbon, Portugal, 2001 (in Portuguese).
- [4] Amaral P, Barahona P. About infeasibility in the constraints of a linear model. Ricerca Operativa 1999; 92:49–67.
- [5] Amaral P, Barahona P. On optimal correction of inconsistent linear constraints. In: Hentenryck PV (Ed). Principles and Practice of Constraint Programming, CP'2002 (Procs.), Lecture Notes in Computer Science 2002. 2470: 33–46.
- [6] Amaral P, Barahona P. Connections between the Total Least Squares and the correction of an infeasible system of Linear Inequalities. Linear Algebra and Applications 2005; 395: 191-210.
- [7] Amaral P, Barahona P. A framework for optimal correction of inconsistent linear constraints. Constraints 2005; 10: 67-86.
- [8] Bazaraa M S, Sherali H D, Shetty C M. Nonlinear Programming: Theory and Algorithms. John Wiley & Sons, New York, third edition 2006.
- [9] Brooke, Kendrick, Meeraus, Raman. GAMS A User's Guide. http://www.gams.com/docs/document.htm.
- [10] Chakravarti N. Some results concerning post-infeasibility analysis. European Journal of Operations Research 1994; 73:139–43.
- [11] , Chinneck J W. Minos(ii):Infeasibility analysis using minos. Computers and Operations Research 1994; 21:1–9.
- [12] Chinneck J W. Finding a useful subset of constraints for analysis in an infeasible linear program. INFORMS Journal on Computing 1997; 9: 164–74.
- [13] Chinneck J W, Dravnieks E W. Locating minimal infeasible constraint sets in linear programs. ORSA Journal on Computing 1991; 3: 157–68.
- [14] Chinneck J W, Saunders M A. Minos(iis) version 4.2: Analysing infeasibilities in linear programs. European Journal of Operations Research 1995; 81: 217–8.
- [15] Greenberg H J. Computer-assisted analysis for diagnosing infeasible or unbounded linear programs. Mathematical Programming 1987; Study 31:79–97.
- [16] Greenberg H J. Consistency, redundancy and implied equalities in linear systems. Annals of Mathematics and Artificial Intelligence 1993; 17: 37–83.

- [17] Greenberg H J. Enhancements of analyse: A computer–assisted analysis system for mathematical programming models and solutions. ACM Transactions on Mathematical Software 1993; 19:233–56.
- [18] Greenberg H J. How to analyze the results of linear programs- Part 3: Infeasibility Diagnoses. Interfaces 1993; 23: 120–39.
- [19] Greenberg H J, Murphy F H. Approaches to diagnosing infeasible linear programs. ORSA Journal on Computing 1991; Study 31: 79–97.
- [20] Guieu O, Chinneck J W. Analyzing infeasible mixed-integer and integer linear programs; INFORMS Journal on Computing 1999; 11: 63–77.
- [21] Holzbaur C, Menezes F, Barahona P. Defeasibility in CLP(Q) through generalised slack variables. In: Freuder E C (Ed.). Proceedings of CP '96, 2nd Int. Conf.in Principles and Practice of Constraint Programming. Lecture Notes in Computer Science vol. 1118, Springer-Verlag, 1996. p 209–23.
- [22] Huffel S V. The total least squares problem: computational aspects and analysis. Frontiers in Applied Mathematics, SIAM, 1991.
- [23] Loon J N M. Irreduciblyinconsistent systems of linear inequalities. European Journal of Operations Research 1981; 8: 282–8.
- [24] Bruce A. Murtagh et al. MINOS- NLP solver from Stanford University. http://www.gams.com/docs/document.htm.
- [25] Pena J. Understanding the geometry of infeasible perturbations of a conic linear system. SIAM Journal on Optimization 2000; 10:534–50.
- [26] Renegar J. Some perturbation theory for linear programming. Mathematical Programming 1994; 65:73–91.
- [27] Roodman G M. Post-infeasibility analysis in linear programming. Management Science 1979; 25: 916–22.
- [28] Vera j R. Ill-posedness and the complexity of deciding existence of solutions to linear programs. SIAM Journal on Optimization 1996; 6:549–569.
- [29] Sherali H D, Tuncbilek C H. A global optimization algorithm for polynomial programming problems using a reformulation-linearization technique. Journal of Global Optimization 1992; 2:101-12.
- [30] Vatolin A A. An lp-based algorithm for the correction of inconsistent linear equation and inequality systems. Optimization 1992; 24:157-64.
- [31] Wang H F, Huang C S. Inconsistent structures of linear systems. International Journal of General Systems 1992; 21:65-81.