

A NONLINEAR FEASIBILITY PROBLEM HEURISTIC

Sergio Drumond Ventura^{1*}, Angel Ramon Sanchez Delgado¹
and Clóvis Caesar Gonzaga²

Received October 15, 2013 / Accepted June 18, 2014

ABSTRACT. In this work we consider a region $S \subset \mathbb{R}^n$ given by a finite number of nonlinear smooth convex inequalities and having nonempty interior. We assume a point x^0 is given, which is close in certain norm to the analytic center of S , and that a new nonlinear smooth convex inequality is added to those defining S (perturbed region). It is constructively shown how to obtain a shift of the right-hand side of this inequality such that the point x^0 is still close (in the same norm) to the analytic center of this shifted region. Starting from this point and using the theoretical results shown, we develop a heuristic that allows us to obtain the approximate analytic center of the perturbed region. Then, we present a procedure to solve the problem of nonlinear feasibility. The procedure was implemented and we performed some numerical tests for the quadratic (random) case.

Keywords: interior point, analytic center, convex optimization.

1 INTRODUCTION

In Linear Programming (LP), as well as in other areas of optimization, it is desirable to have the possibility of making the post-optimality processing the most efficient way. In particular, if the LP is solved and right after that the original data are altered, one wants to restart the process to find the solution for the new problem from the obtained solution for the original LP. It is thought that in many occasions, to obtain a new solution using the latest information requires less computational effort than to start the process from scratch.

It is widely known that this kind of hot start can be successfully carried out by the Simplex method and its variants for certain perturbations in the data. An important particular case, for its applications, is the addition of a new constraint. Adding a constraint to a LP is equivalent to the addition of a variable in the associated dual problem. If this variable is zero, then the optimal dual solution is feasible for the new problem and then we can start the Dual-Simplex method from

*Corresponding author.

¹Department of Mathematics, Federal Rural University of Rio de Janeiro, BR-465, Km 7, 23890-000 Seropédica, RJ, Brazil. Tel.: +55-21-26814841. E-mails: ventura@ufrj.br; asanchez@ufrj.br

²Department of Mathematics, Federal University of Santa Catarina, Caixa Postal 5210, 88040-970 Florianópolis, SC, Brazil. Tel.: +55-48-37216560. E-mail: clovis@mtm.ufsc.br

this solution to solve the new problem. This often reduces the number of iterations required to resolve the perturbed problem. This idea has an important application in Cutting Planes Methods (CPMs) for Integer Linear Programming (ILP).

In an ILP problem, we want to optimize a linear functional in several variables, subject to linear constraints and such that some or all variables take on integer values, that is, a problem of ILP is a LP problem where some or all variables must be integers. This complication takes the seemingly simple LP problem into a problem of combinatorial nature, whose solution can be difficult to achieve. Many ILP problems, such as the traveling salesman problem, are NP-complete. Traditionally, such problems are solved using heuristics like local search, “Simulated annealing”, genetic algorithms, ant colony, etc., thus obtaining good solutions, and that in some cases can be optimal. However, these techniques do not have the ability to check the optimality the obtained solutions.

On the other hand, classical methods such as “branch and bound”, cutting planes and its hybrids (such as “branch and cut”) use bounds for the optimal value (iteratively updated) to check the optimality of the obtained solution or at least to inform its proximity to the optimal value. Among these techniques, it is interesting the CPM, which is a generic term for optimization methods which iteratively refine a feasible set or objective function by means of linear inequalities, known as cuts. Such procedure are popularly used to find integer solutions to Mixed Integer Linear Programming (MILP) problems, as well as to solve general, not necessarily differentiable convex optimization problems. By considering certain assumptions of the problem, we can prove the convergence of the CPM, what does not mean a reasonable computational behavior. In certain occasions, the time required is so large that in practice it is impossible to obtain a solution. Despite these difficulties, the CPMs still are the best behaved methods in solving specialized problems. In these applications, the Simplex method is used to solve the relaxed LP generated by CPM.

Thirty years after the revolutionary step by N. Karmarkar [11] in the area of optimization, when he introduced a polynomial interior point algorithm with computational efficiency for the solving LP problems, it was natural that some research arose on the application of interior point algorithms in solving post-optimality problems, in particular in solving the LP arising from a relaxed CPM to solve the ILP. One of the early works in this direction was presented in [1, 13, 14], where it is described a CPM for ILP that uses as a subroutine a variant of Karmarkar’s projective algorithm. In each iteration, it is solved the dual problem associated with the relaxed LP using the Karmarkar-variant. As mentioned earlier, when adding a new constraint to the original LP, we easily obtain a dual-feasible solution. The problem is that this solution is not interior. It is obtained by making at least one variable zero. Using heuristics and LP information, the authors provide a direction from which it is possible to obtain an interior solution to restart the projective algorithm or the Karmarkar-variant.

Another way to address the problem of restarting the algorithm when we add a constraint is to use algorithms that work with points that do not necessarily are feasible (see [12, 16]). This methodology achieves both the feasibility and optimality. The problem is that by adding a constraint

to the feasibility region, the iterations of the algorithm tend to move towards the center of the feasible region, thus recovering feasibility but moving away from optimal.

So far one of the best methods that uses interior point in the solution of the relaxed LP in a CPM is the method based on the concept of analytic center. The analytic center for a feasible region is single interior point farthest from the border. Briefly, a CPM can be thought of as a column generation method applied to the dual problem of the relaxed LP. In [18] a polynomial algorithm is proposed with column generation based on interior point and applied to Linear Feasibility Problem (LFP). The LFP is equivalent to a LP problem. In [6] it is presented a polynomial algorithm for column generation, which uses the analytic center and solves the nonlinear convex feasibility problem.

In general, the problem to be solved when using interior point in a CPM is that when adding a new constraint, the current solution is no longer interior, and therefore it is not possible to reset the methodology. In this case, one needs some technique or heuristic that allows the “interior recovery” of the current solution. In this work we present a heuristic to recover the analytic center of a region defined by a finite number of nonlinear inequalities, as soon as we add a new inequality to region. This heuristic is based on the extension of the theoretical results developed in [5].

In the next section we present the theoretical work and it is where the problem to be considered is defined. In the sequence, we present the extended results, as well as the implemented heuristic and its application in solving the nonlinear feasibility problem. Finally, we present the numerical results and the conclusions.

2 THEORETICAL FOUNDATIONS

We consider a region $S \subset \mathbb{R}^n$ given by $S = \{x \in \mathbb{R}^n : f_i(x) \leq 0; i = 1, \dots, m\}$ where $f_i(x)$ are convex functions from \mathbb{R}^n to \mathbb{R} with continuous second order derivatives in $int(S)$ (the interior of S). We will assume that $int(S)$ is nonempty and bounded. The logarithmic barrier function associated to S is given for $P(x) = -\sum_{i=1}^m \ln(-f_i(x))$ and the gradient and the Hessian matrix of $P(x)$ at x are, respectively:

$$g = g(x) = \nabla P(x) = \sum_{i=1}^m \frac{\nabla f_i(x)}{-f_i(x)}, \tag{1}$$

$$H = H(x) = \nabla^2 P(x) = \sum_{i=1}^m \left[\frac{\nabla^2 f_i(x)}{-f_i(x)} + \frac{\nabla f_i(x) \nabla f_i(x)^\top}{f_i(x)^2} \right], \tag{2}$$

where H induces the relative norm $\|x\|_H = \sqrt{x^\top H x}$ for all $x \in \mathbb{R}^n$. We suppose that the Hessian matrices H_i of all constraint functions $f_i(x)$ fulfils a Relative Lipschitz Condition, i.e.,

$$\exists M \geq 0 : \forall w \in \mathbb{R}^n, \forall y \in int(S), \forall h \text{ with } \|h\|_{H_i(y)} \leq 0.5 / \left(1 + M^{1/3}\right) \tag{3}$$

$$\left| w^\top (H_i(y+h) - H_i(y)) w \right| \leq M \|h\|_{H_i(y)} w^\top H_i(y) w. \tag{4}$$

Thus, for each i , the logarithmic barrier function $-\ln(-f_i(x))$ is self-concordant [15]. Also, the logarithmic barrier function $P(x)$ associated to S is a self-concordant function (as the sum of self-concordant functions is again a self-concordant function) with the Hessian matrix being positive-definite on its domain $\text{int}(S)$, and strictly convex. $P(x)$ penalizes points near the boundary of S .

Note that $P(x)$ has a unique minimizer $x_c \in \text{int}(S)$, which is called **analytic center** of S , see e.g. [10], i.e., $x_c = \arg \min\{P(x) : x \in \text{int}(S)\}$. The *proximity* of $x \in \text{int}(S)$ to x_c is defined by $\delta(x) = \|H^{-1}(x)g(x)\|_H = \sqrt{g^T H^{-1}g}$, i.e., the norm (induced by H) of the Newton direction for P at x , or the scaled gradient vector. Then, x will be considered *near* x_c (or ε -approximate center) if $\delta(x) < \varepsilon$ for some $\varepsilon \in (0, 1)$. Clearly, $\delta(x) = 0$ if and only if $x = x_c$.

In this work we suppose a point $x^0 \in \text{int}(S)$ is given such that $\delta = \delta(x^0) < \varepsilon$ for some $\varepsilon \in (0, 1)$, and also, that it is given $f_0(x)$, a convex function from \mathbb{R}^n to \mathbb{R} with continuous second order derivative in $\text{int}(S)$, such that the inequality $f_0(x) \leq 0$ generates a new region S^0 from S :

$$S^0 = \{x \in \mathbb{R}^n : f_0(x) \leq 0; f_i(x) \leq 0; i = 1, \dots, m\}. \tag{5}$$

We are interested in finding a “shift” γ^* of the right-hand side of the new inequality which generates a new region where x^0 is still close, in the same norm, to the corresponding analytic center. This question is conceptually equivalent to the following problem: given an ε -approximate center x^0 of the region S and $f_0(x) \leq \gamma^*$, find an upper bound for δ , depending on f_0 , ε and γ^* , to preserve x^0 to be δ -approximate center of the region

$$S^* = \{x \in \mathbb{R}^n : f_0(x) \leq \gamma^*; f_i(x) \leq 0; i = 1, \dots, m\}.$$

We present results (bounds for γ^*) to preserve centrality after a new constraint $f_0(x) \leq \gamma^*$ is added to S . More accurately, given an ε -approximate center x^0 of the region S , an upper bound for γ^* depending on f_0 and ε , is derived to preserve x^0 to be an $\sqrt{\varepsilon}$ -approximate center of the S^* .

3 RESULTS

With analytic center theory in mind, given $\gamma \in \mathbb{R}$ and $q \geq m$ a positive integer, we start our search for such a region by adding the inequality $f_0(x) \leq \gamma$ repeated q times to those defining S , that is, we set:

$$S_{q,\gamma} = \{x \in \mathbb{R}^n : f_0(x) \leq \gamma \text{ (} q\text{-times)} f_i(x) \leq 0; i = 1, \dots, m\}. \tag{6}$$

Note that the integer q acts as a weight in computing the associated barrier function. Indeed: $P_q(x, \gamma) = -q \ln[\gamma - f_0(x)] + P(x)$. Also, it is a known fact that we only claim that $P_1(x, 0) = P(x, 0) = -\ln[-f_0(x)] + P(x)$. The notation $S_{q,\gamma}$ is used throughout to emphasize that the logarithmic barrier function being considered for the region has the inequality with right-hand side γ repeated q times.

The analytic center of $S_{q,\gamma}$ is $x_c^q(\gamma) = \arg \min\{P_q(x, \gamma) : x \in \text{in}(S_{q,\gamma})\}$. As done previously, the proximity measure to $x_c^q(\gamma)$, given $x \in \text{int}(S_{q,\gamma})$, needs the gradient and the Hessian matrix of $P_q(x, \gamma)$ (with respect to x), which are:

$$g_q(x, \gamma) = \nabla_x P_q(x, \gamma) = \frac{q \nabla f_0(x)}{\gamma - f_0(x)} + g(x), \tag{7}$$

$$H_q(x, \gamma) = \nabla_x^2 P_q(x, \gamma) = \frac{q \nabla^2 f_0(x)}{\gamma - f_0(x)} + \frac{q \nabla f_0(x) \nabla f_0(x)^\top}{[\gamma - f_0(x)]^2} + H(x). \tag{8}$$

If no confusion is possible, we will write, for the sake of shortness, g_q and H_q instead of $g_q(x, \gamma)$ and $H_q(x, \gamma)$. H_q is positive-definite for $x \in \text{int}(S_{q,\gamma})$, and the proximity of x to $x_c^q(\gamma)$ is given by $\delta_q(x, \gamma) = \left\| H_q^{-1}(x, \gamma) g_q(x, \gamma) \right\|_H = \sqrt{g_q^\top H_q^{-1} g_q}$. The following lemma shows the relation between $\delta(x^0)$ and $\delta_q(x^0, \gamma)$, for $\gamma > f_0(x^0)$.

From now on, given two square matrices A and B , by $A \leq B$ we mean that $B - A$ is positive semidefinite.

Lemma 3.1. *Given $\gamma > f_0(x^0)$, $d_0 = \nabla f_0(x^0)$, $Q_0 = \nabla^2 f_0(x^0)$, $\theta = d_0^\top H^{-1}(x^0) d_0 = d_0^\top H_0^{-1} d_0$, $\alpha = d_0^\top H_0^{-1} g(x^0) = d_0^\top H_0^{-1} g_0$ and $\beta_0 = \beta_0(\gamma) = \frac{1}{\gamma - f_0(x^0)}$. Then*

$$\delta_q^2(x^0, \gamma) \leq \delta^2(x^0) + \frac{q^2 \theta \beta_0^2 + 2q \alpha \beta_0}{1 + q \theta \beta_0^2}. \tag{9}$$

Proof. From (7) and (8) we have $g_q(x^0, \gamma) = g_0 + q \beta_0 d_0$, and $H_q(x^0, \gamma) = q \beta_0^2 d_0 d_0^\top + q \beta_0 Q_0 + H_0$. Also, $H_0 + q \beta_0^2 d_0 d_0^\top \leq H_q(x^0, \gamma)$, so that by Corollary 7.7.4 of [9], $H_q^{-1}(x^0, \gamma) \leq (H_0 + q \beta_0^2 d_0 d_0^\top)^{-1}$. As a consequence, it holds that:

$$\begin{aligned} \delta_q^2(x^0, \gamma) &= g_q^\top(x^0, \gamma) H_q^{-1} g_q(x^0, \gamma) \\ &\leq g_q^\top(x^0, \gamma) [H_0 + q \beta_0^2 d_0 d_0^\top]^{-1} g_q(x^0, \gamma) \\ &= (g_0 + q \beta_0 d_0)^\top \left(H_0^{-1} - \frac{q \beta_0^2 H_0^{-1} d_0 d_0^\top H_0^{-1}}{1 + q \theta \beta_0^2} \right) (g_0 + q \beta_0 d_0) \\ &= \delta^2(x^0) + \frac{q \theta \beta_0^2 + 2q \alpha \beta_0 - q \alpha^2 \beta_0^2}{1 + q \theta \beta_0^2} \\ &\leq \delta^2(x^0) + \frac{q \theta \beta_0^2 + 2q \alpha \beta_0}{1 + q \theta \beta_0^2}. \end{aligned} \tag{10}$$

□

The following theorem shows how we can choose γ to make x^0 close to the analytic center of $S_{q,\gamma}$.

Theorem 3.2. Let $d_0 = \nabla f_0(x^0)$, $\theta = d_0^\top H_0^{-1} d_0 > 0$, $\alpha = d_0^\top H_0^{-1} g_0$, $\varepsilon \in (0, 1)$, and $\rho = (\sqrt{\alpha^2 + \theta\varepsilon(1 - \varepsilon)} - \alpha)/(q\theta)$. Then for $\gamma^0 \geq f_0(x^0) + 1/\rho$, if $\delta(x^0) < \varepsilon$ then $\delta_q(x^0, \gamma^0) < \sqrt{\varepsilon}$.

Proof. By Lemma 3.1, $\delta_q^2(x^0, \gamma) \leq \delta^2(x^0) + \frac{q\theta\beta_0^2 + 2q\alpha\beta_0}{1 + q\theta\beta_0^2}$ (where $\beta_0 = 1/(\gamma^0 - f_0(x^0))$). If the numerator of the fraction is negative, we conclude that $\delta_q^2(x^0, \gamma^0) \leq \delta^2(x^0) < \varepsilon^2 < \varepsilon$, and we are done. If it is nonnegative, use the facts that the denominator is greater than 1, and

$$\gamma^0 - f_0(x^0) \geq \frac{q\theta}{\sqrt{\alpha^2 + \theta\varepsilon(1 - \varepsilon)} - \alpha}$$

to get $\beta_0 = \frac{1}{\gamma^0 - f_0(x^0)} \leq \frac{\sqrt{\alpha^2 + \theta\varepsilon(1 - \varepsilon)} - \alpha}{q\theta} = \rho$. On the other hand, $\delta_q^2(x^0, \gamma^0) < \delta^2(x^0) + q^2\theta\beta_0^2 + 2q\alpha\beta_0 = \delta^2(x^0) + q\beta_0(q\theta\beta_0 + 2\alpha) \leq \delta^2(x^0) + q\rho(q\theta\rho + 2\alpha) = \delta^2(x^0) + \varepsilon(1 - \varepsilon) < \varepsilon^2 + \varepsilon(1 - \varepsilon) = \varepsilon$. □

This guarantees that the point x^0 is close to the analytic center of S_{q,γ^0} . Nevertheless, this is not the desired region, since the constraint involving the constraint function $f_0(\cdot)$ appears with weight $q > 1$ in the associated logarithmic barrier function, and we can show that x^0 is close to the analytic center of such a region but with $q = 1$. Recall that the constraint $f_0(x) \leq 0$ was added to S just once.

The following result helps to show that, in general, if a point x is close to the analytic center of $S_{q,\gamma}$, then there exists a “representative” of γ , say γ^* , such that x is close to the analytic center of S_{1,γ^*} .

Theorem 3.3. Given $\gamma \in \mathbb{R}$, $x \in \text{int}(S_{q,\gamma})$, then for $\gamma^* = \frac{\gamma}{q} + \left(1 + \frac{1}{q}\right) f_0(x)$, we have

$$\delta_1(x, \gamma^*) \leq \delta_q(x, \gamma). \tag{11}$$

Proof. Let $\beta = \beta(x, \gamma) = \frac{1}{\gamma - f_0(x)}$, $\beta^* = \beta^*(x, \gamma^*) = \frac{1}{\gamma^* - f_0(x)}$, $d = d(x) = \nabla f_0(x)$, $Q = Q(x) = \nabla^2 f_0(x)$. Clearly $\beta^* = q\beta$. Also, $g_1 = \nabla_x P_1(x, \gamma^*) = \frac{\nabla f_0(x)}{\gamma^* - f_0(x)} + g(x) = q\beta d + g = g_q$. Now,

$$\begin{aligned} H_1 = \nabla_x^2 P(x, \gamma^*) &= \frac{\nabla f_0(x) \nabla f_0(x)^\top}{[\gamma^* - f_0(x)]^2} + \frac{Q}{\gamma^* - f_0(x)} + H \\ &= q^2 \beta^2 \nabla f_0(x) \nabla f_0(x)^\top + q\beta Q + H \\ &= H_q + (q^2 - q)\beta^2 dd^\top \geq H_q, \end{aligned} \tag{12}$$

as $q > 1$. Therefore, using as before, Corollary 7.7.4 of [9], $H_1^{-1} \leq H_q^{-1}$, and $\delta_1^2(x, \gamma^*) = g_1^\top H_1^{-1} g_1 = g_q^\top H_1^{-1} g_q \leq g_q^\top H_q^{-1} g_q = \delta_q^2(x, \gamma)$; or $\delta_1(x, \gamma^*) < \delta_q(x, \gamma)$. □

Corollary 3.4. x^0 is close to the analytic center of $S_{1,\gamma^{0*}}$, where $\gamma^{0*} = \frac{\gamma^0}{q} + \left(1 - \frac{1}{q}\right) f_0(x^0)$, and $\gamma^0 \geq f_0(x^0) + \frac{1}{\rho}$ (see Theorems 3.2 and 3.3). It is clear that $S_{1,\gamma^{0*}}$ is the desired region.

Corollary 3.5. *If $\gamma^{0*} \leq 0$, then $\delta_1(x^0, 0) \leq \sqrt{\varepsilon}$.*

Proof. Let $\gamma^1 = \frac{1}{q}0 + \left(1 - \frac{1}{q}\right) f_0(x^0) = \left(1 - \frac{1}{q}\right) f_0(x^0)$. By hypothesis,

$$\gamma^1 \geq \frac{1}{q}\gamma^{0*} + \left(1 - \frac{1}{q}\right) f_0(x^0) = \gamma^0, \tag{13}$$

and therefore $\gamma^1 \geq f_0(x^0) + \frac{1}{\rho}$. By Theorem 3.2, we have $\delta_q(x^0, \gamma^1) < \sqrt{\varepsilon}$. Note that γ^1 is defined in such a way that its representative regarding x^0 is zero, and that by Theorem 3.3, $\delta_1(x^0, 0) \leq \delta_q(x^0, \gamma^1)$. Combining these two inequalities we obtain $\delta_1(x^0, 0) \leq \sqrt{\varepsilon}$. \square

4 APPLICATIONS (NONLINEAR FEASIBILITY PROBLEM)

4.1 Heuristic

From the literature of mathematical programming, we know that methods of cutting planes [1], using analytic center as testing points, need some clever strategy to “restart” the method of centers, which is used to solve the corresponding partial or relaxed problem (see [2, 3, 4, 7, 17]). Such restarting consists in finding a new point close to the analytic center of a convex perturbed region, right after adding a new constraint. That is, suppose we have a point $x^0 \in \text{int}(S)$ such that $\delta(x^0) < \varepsilon$, for some $\varepsilon \in (0, 1)$. Suppose also that a new convex function $f_0(x)$ is given, along with its second order continuous derivatives, defined in $\text{int}(S)$, and such the inequality $f_0(x) \leq 0$ generates a new region

$$S^0 = \{x \in \mathbb{R}^n : f_0(x) \leq 0, f_i(x) \leq 0, i = 1, \dots, m\}, \tag{14}$$

where $x^0 \notin \text{int}(S^0)$. The problem of restarting in a cutting plane strategy consists in determining a new point close to the analytic center of S^0 . Below we present a heuristic to solve this problem using the results of the previous section.

First of all, as Theorem 3.2 indicates, x^0 is close to the analytic center of S_{q,γ^0} , for $\gamma^0 = f_0(x^0) + \frac{1}{\rho}$; moreover, $\delta_q(x^0, \gamma^0) < \sqrt{\varepsilon}$. Also, by Theorem 3.3, x^0 is close to the analytic center of $S_{1,\gamma^{0*}}$, for $\gamma^{0*} = \frac{1}{q}\gamma^0 + \left(1 - \frac{1}{q}\right) f_0(x^0)$, so that $\delta_1(x^0, \gamma^0) < \sqrt{\varepsilon}$.

Now, if $\gamma^{0*} \leq 0$, then by Corollary 3.5 we have the desired point. Otherwise, that is, if $\gamma^{0*} > 0$ we can initiate the strategy of reducing the value of γ^0 and, simultaneously, change the value of γ^{0*} , trying to make it close to zero. To accomplish this, we will use an algorithm in two phases. The first one is a method of an incomplete method of centers (Algorithm 1), where $(y, z) \in \mathbb{R}^{n+1}$, $f : \mathbb{R}^{n+1} \rightarrow \mathbb{R}$ and τ represents tolerance to the proximity of the analytic center (at the numerical tests, we used $\tau = 0.1$). We want to reduce (not optimize) γ , starting with the value of γ^0 , until that the representative of γ^* becomes negative. The second one is a process of bisection to “approach zero”, where in each iteration we do a centralization, just like in the method of centers.

Algorithm 1: Centering (y, z, f)

Obtain the gradient g and hessian H_f of $f(x)$ in y and z ;
 Solve the linear system $H_f p = -g$;
 Obtain the proximity: $\sigma \leftarrow \|p\|_H$;
while $\sigma > \tau$ **do**
 $\bar{\alpha} \leftarrow \arg \min \{f(y + \alpha p) : y + \alpha p \in S\}$
 $y \leftarrow y + \bar{\alpha} p$
 Obtain the gradient g and hessian H_f of $f(x)$ in y and z ;
 Solve the linear system $H_f p = -g$;
 Obtain the proximity: $\sigma \leftarrow \|p\|_H$;

Remember that the barrier logarithmic function associated to the region S_{q,γ^0} is given by $P_q(x, \gamma^0) = -q \ln(\gamma^0 - f_0(x)) + P(x)$ and corresponds to the auxiliary function used in a method of centers to solve the problem

$$\begin{aligned} \max \quad & -f_0(x) \\ \text{s.t.} \quad & f_i(x) \leq 0, \quad i = 1, \dots, m \end{aligned} \tag{15}$$

Therefore, the phase of reduction (Algorithm 2) can be initiated with the application of the incomplete method of centers starting from x^0 and the bound γ^0 . We update the bound to γ^1 and we obtain a new point x^1 , close to the analytic center of S_{q,γ^1} . Then we evaluate γ^{1*} , the representative of γ^1 , such that x^1 is close to the analytic center of $S_{1,\gamma^{1*}}$, and we check if $\gamma^{1*} \leq 0$ or, at least, if it is close enough to zero. In the first case, the reduction phase ends; in the second one the heuristic stops (the desired point was achieved). If none of the previous conditions are satisfied, then we do another iteration, starting from x^1 and γ^1 , and so on. This phase generates sets $\{\gamma^j\}$, $\{\gamma^{j*}\}$ and $\{x^j\}$ such that each x^j is located close to the analytic center of S_{q,γ^j} and $S_{q,\gamma^{j*}}$. Now, if the method of centers were fully applied, we would achieve the optimal value of (15) in polynomial time. In this way, using an incomplete method of centers, we would achieve in a polynomial time the condition $\gamma^{k*} < 0$, for some k .

Algorithm 2: Reduction ($\gamma^L, y^L, \bar{\gamma}^L$)

while $\bar{\gamma}^L > \delta$ **do**
 $\gamma^R \leftarrow \gamma^L$; $y^R \leftarrow y^L$; $\bar{\gamma}^R \leftarrow \bar{\gamma}^L$;
 $\gamma^L \leftarrow \gamma^L + \theta [f_0(y^L) - \gamma^L]$;
 $y^L \leftarrow \text{Centering}(y^L, -\gamma^L, P_q(x, \gamma^L))$;
 $\bar{\gamma}^L \leftarrow \frac{1}{q}\gamma^L + \left(1 - \frac{1}{q}\right) f_0(y^L)$;
 Return $\gamma^L, y^L, \bar{\gamma}^L, \gamma^R, y^R, \bar{\gamma}^R$;

At the beginning of the second phase we know the iterates $\gamma^{(k-1)}$, $x^{(k-1)}$, $\gamma^{(k-1)*}$, and $\gamma^k, x^k, \gamma^{k*}$ such that $\gamma^{k*} < 0 < \gamma^{(k-1)*}$. This suggests the possibility of beginning bisection (Algorithm 3) over the interval $[\gamma^k, \gamma^{(k-1)*}]$. For simplification, suppose that a left iterate $\gamma^L, x^L, \gamma^{L*}$

and a right iterate $\gamma^R, x^R, \gamma^{R*}$ are given with the property that x^L is close to the analytic center of S_{q,γ^L} and $S_{1,\gamma^{L*}}$, and that x^R is close to the analytic center of S_{q,γ^R} and $S_{1,\gamma^{R*}}$. Also, suppose that $\gamma^{L*} < 0 < \gamma^{R*}$. Regarding the parameters, we used $q = 2n + 1$ and δ represents tolerance to the proximity of the right-hand side of the new added constraint (at the numerical tests, we used $\delta = 0.1$).

Algorithm 3: Bisection($\gamma^L, y^L, \bar{\gamma}^L, \gamma^R, y^R, \bar{\gamma}^R$)

while $\bar{\gamma}^L < -\delta$ and $\bar{\gamma}^R > \delta$ **do**
 $\gamma^M \leftarrow \frac{1}{2}(\gamma^L + \gamma^R)$;
 $y^M \leftarrow \mathbf{Centering}(y^R, -\gamma^M, P_q(x, \gamma^M))$;
 $\bar{\gamma}^M \leftarrow \frac{1}{q}\gamma^M + \left(1 - \frac{1}{q}\right)f_0(y^M)$;
if $\bar{\gamma}^M < 0$ **then**
 $\gamma^L \leftarrow \gamma^M$; $y^L \leftarrow y^M$; $\bar{\gamma}^L \leftarrow \bar{\gamma}^M$;
else
 $\gamma^R \leftarrow \gamma^M$; $y^R \leftarrow y^M$; $\bar{\gamma}^R \leftarrow \bar{\gamma}^M$;
if $\bar{\gamma}^L \geq -\delta$ **then** Return y^L ;
if $\bar{\gamma}^R \leq \delta$ **then** Return y^R ;

As usual in bisection, we evaluate the middle point of the interval $[\gamma^L, \gamma^R]$, that is, $\gamma^M = \frac{\gamma^L + \gamma^R}{2}$. Then, from x^R and γ^M (as bound) we do a centralization to obtain the point x^M , close to the analytic center of S_{q,γ^M} . Then we evaluate γ^{M*} (the representative of γ^M regarding x^M). Clearly, if γ^{M*} is close enough to zero, then the heuristic can stop. Otherwise, there are two possibilities: either $\gamma^{M*} < 0$, in which case $[\gamma^M, \gamma^R]$ is a new interval to consider for bisection, or $\gamma^{M*} > 0$, in which case $[\gamma^L, \gamma^M]$ is a new interval to consider.

Based on the previous procedures, we present the ‘‘Analytic Center Recovery Algorithm’’ – ACRA (Algorithm 4). Such routine consists in finding a new point close to the analytic center of a convex perturbed region, right after adding a new constraint.

Algorithm 4: ACRA(y, S)

Obtain the gradient g and hessian H of $P(x)$ in y^0 ;
 Obtain ρ according to Theorem 3.2
 $\gamma^L \leftarrow f_0(x^0) + \frac{1}{\rho}$; $\bar{\gamma}^L \leftarrow \frac{1}{\rho}\gamma^L + \left(1 - \frac{1}{\rho}\right)f_0(x^0)$;
if $\bar{\gamma}^L \leq 0$ **then**
 $y \leftarrow y^0$;
else
 $y^L \leftarrow y^0$;
 $(\gamma^L, y^L, \bar{\gamma}^L, \gamma^R, y^R, \bar{\gamma}^R) \leftarrow \mathbf{Reduction}(\gamma^L, y^L, \bar{\gamma}^L)$;
 $y \leftarrow \mathbf{Bisection}(\gamma^L, y^L, \bar{\gamma}^L, \gamma^R, y^R, \bar{\gamma}^R)$;

4.2 Nonlinear feasibility problem

Using ACRA heuristic we developed a procedure, called “Convex Feasibility Algorithm” – CFA, to solve the non-linear problem of feasibility in smooth and convex regions (Algorithm 5). The procedure is based on the known methodology of generation of columns (or inequalities) [8]. Our initial problem is to find an interior point close to the analytic center of a given region $S = \{x \in \mathbb{R}^n : f_i(x) \leq 0; i = 1, \dots, m\} \neq \emptyset$. We are supposing that S is a smooth convex region, that is, it satisfies the conditions described at the introduction of this work, so that we can suppose that there exists a box in \mathbb{R}^n , say $T = \{x \in \mathbb{R}^n : \ell \leq x \leq u\}$ containing S , where $\ell, u \in \mathbb{R}^n$ and $\ell_j < u_j$. Therefore, our problem is to find an interior point close to the analytic center of the region $\{x \in \mathbb{R}^n : f_i(x) \leq 0; i = 1, \dots, m, \ell \leq x \leq u\}$. The use of a box containing S , far from complicating the problem, allows us to have an automatic inicialization of the proposed procedure, as we will see below.

Algorithm 5: CFA

```

 $y^0 \leftarrow 0;$ 
for  $j = 1, \dots, m$  do
   $S_j \leftarrow \{x \in \mathbb{R}^n : f_i(x) \leq 0; i = 1, \dots, j \text{ and } -e \leq x \leq e\};$ 
   $y^j \leftarrow \text{ACRA}(y^{j-1}, S_j);$ 

```

To simplify, we can suppose that right after some translation and scaling in the variables, the previous region can be written like $\{x \in \mathbb{R}^n : f_i(x) \leq 0; i = 1, \dots, m, \text{ and } -e \leq x \leq e\}$, where $e = (1, \dots, 1)^T$. In this way, we can initiate the procedure, starting from the exact analytic center of the region $B = \{x \in \mathbb{R}^n : -e \leq x \leq e\}$, which coincides with the geometric center of S when given by $x^0 = 0$. Then, starting with $f_1(x)$, we add this inequality to B and using the proposed heuristic, we obtain a point x^1 , close to the analytic center of the new region. Now, we repeat this process for x^1 and so on, up to the end.

The procedure performs at most m times the ACRA routine, considering that once an inequality is added, this one cannot be violated by any further iterate. At the end, we can withdraw the constraints regarding the unitary box B and obtain the desired point, that is, a point close to the analytic center of S , a key point when one has to work with a methodology of cutting planes that uses like test point, points close to the analytic center of the generated regions during each iteration.

However when working with nonlinear feasibility problem, we consider an implementation of the heuristic that relax the bisection subroutine and we conform to the recovery of the feasibility of the iterated, i.e., we do not try to approach the right-hand side of the added inequality but rather a relaxation of it. In this strategy it is possible, however with little possibility, that the same inequality be violated in different iterations, so preventing convergence. More research is needed to try to determine if there are conditions that guarantee convergence. Aiming to understand the behavior of computational modified heuristic some numerical tests were carried out with promising results.

5 NUMERICAL RESULTS

In this work, we have tried the heuristic with random problems, for the sake of readiness of implementation.

5.1 Generation of random problems

The random convex feasibility problems are generated as follows. We want to generate a region in the form

$$S = \{y \in \mathbb{R}^n : f_1(y) < 0, \dots, f_m(y) < 0 \text{ and } -e < y < e\}, \tag{16}$$

where for each $i = 1, \dots, m$,

$$f_i(y) = \frac{1}{2}y^\top A_i y + b_i^\top y + c_i, \tag{17}$$

and where A_i is symmetric positive-definite.

We start by choosing a point $y_0 \in \{y \in \mathbb{R}^n : -e < y < e\}$, where we randomly obtain each entry according to a uniform distribution in the interval $(-1, 1)$.

Now, for each quadratic function f_i , after randomly generating a symmetric positive-definite matrix A_i , we generate a random direction vector $d \in \mathbb{R}^n$, whose each entry is generated according to a uniform distribution in $(-1, 1)$. Then we normalize d using the A_i -norm, that is, $u_i = d / \|d\|_{A_i}$. Then we set

$$\begin{aligned} c_i &= \frac{1}{2}(y_0 + \alpha u_i)^\top A_i (y_0 + \alpha u_i) + b_i^\top (y_0 + \alpha u_i) \\ &= \frac{1}{2}\alpha^2 + \alpha(A_i y_0 + b_i)^\top u_i + \left[\frac{1}{2}y_0^\top A_i y_0 + b_i^\top y_0 \right]. \end{aligned} \tag{18}$$

Now, if we take $b_i = A_i(y_0 - u_i)$, then

$$c_i = \frac{1}{2}\alpha^2 - \alpha + \left[\frac{1}{2}y_0^\top A_i y_0 + b_i^\top y_0 \right]. \tag{19}$$

So, by randomly choosing $\alpha > 2$, we obtain

$$c_i > \frac{1}{2}y_0^\top A_i y_0 + b_i^\top y_0. \tag{20}$$

that is

$$0 > \frac{1}{2}y_0^\top A_i y_0 + b_i^\top y_0 - c_i. \tag{21}$$

5.2 Results of random problems

We performed four sets of randomly-generated problems. Each set was labeled according to the dimensions of its problems. They were 50×50 , 50×100 , 100×50 and 100×100 , where the

first number is n and the second number is the number m of constraints. The results are shown at Tables 1, 2, 3 and 4, respectively, where we show firstly the results with the bisection routine and secondly the results without the bisection routine. We can see that only 5 test problems are shown at each test set. However, at each test set we had 30 problems and it is worth mentioning that some of the tests were not successful, because the bisection routine achieved its maximum number of iterations (100) and the test was aborted. Nevertheless, the conclusions can be clearly drawn based on these 5 test problems.

Table 1 – Tests with and without the bisection routine for randomly-generated problems with size 50×50 . The average running time was 2 min for the tests with bisection and 50s for those without bisection.

Problem	FEA	RED	BIS	NWT	RED FEA	BIS FEA	NWT RED+BIS	Rel error
1	15	17	117	231	1.13	7.80	1.72	0.109795
2	8	10	72	130	1.25	9.00	1.58	0.180610
3	8	10	89	101	1.25	11.12	1.02	0.394144
4	13	14	110	222	1.07	8.46	1.79	0.120790
5	8	10	80	113	1.25	10.00	1.25	0.140823
1	12	13	0	52	1.08	0.00	4.00	0.042579
2	8	9	0	37	1.12	0.00	4.11	0.062241
3	6	7	0	29	1.16	0.00	4.14	0.385106
4	11	12	0	49	1.09	0.00	4.08	0.072579
5	6	7	0	29	1.16	0.00	4.14	0.362209

Table 2 – Tests with and without the bisection routine for randomly-generated problems with size 50×100 . The average running time was 9min for the tests with bisection and 3min for those without bisection.

Problem	FEA	RED	BIS	NWT	RED FEA	BIS FEA	NWT RED+BIS	Rel error
1	11	13	124	166	1.18	11.27	1.21	0.066950
2	11	13	122	177	1.18	11.09	1.31	0.235775
3	16	17	137	239	1.06	8.56	1.55	0.056632
4	16	16	129	288	1.00	8.06	1.98	0.102728
5	9	11	88	143	1.22	9.77	1.44	0.194135
1	8	10	0	43	1.25	0.00	4.30	0.060965
2	9	10	0	45	1.11	0.00	4.50	0.038112
3	15	16	0	70	1.06	0.00	4.37	0.020806
4	14	14	0	64	1.00	0.00	4.57	0.040943
5	8	9	0	41	1.12	0.00	4.55	0.184487

In the following tables, FEA represents the number of times that ACRA was called. RED stands for the total number of reductions, while BIS stands for the total number of bisections. Also, NWT is the total number of iterations of Newton’s method. Finally, after each test, we obtained the analytic center of the region using MATLAB’s unconstrained minimization routine `fminunc` and compared it with what we obtained after CFA. In this way, the relative error is shown at the last column.

Table 3 – Tests with and without the bisection routine for randomly-generated problems with size 100×50 . The average running time was 7min for the tests with bisection and 2min for those without bisection.

Problem	FEA	RED	BIS	NWT	RED FEA	BIS FEA	$\frac{NWT}{RED+BIS}$	Rel error
1	10	12	136	127	1.20	13.60	0.85	0.161682
2	7	9	86	90	1.28	12.28	0.94	0.440358
3	9	11	119	109	1.22	13.22	0.83	0.168031
4	6	8	71	95	1.33	11.83	1.20	0.258039
5	15	15	145	273	1.00	9.66	1.70	0.052824
1	8	9	0	39	1.12	0.00	4.33	0.029781
2	5	6	0	25	1.20	0.00	4.16	0.442479
3	6	8	0	32	1.33	0.00	4.00	0.190755
4	6	7	0	29	1.16	0.00	4.14	0.220485
5	11	11	0	46	1.00	0.00	4.18	0.124300

Table 4 – Tests with and without the bisection routine for randomly-generated problems with size 100×100 . The average running time was 24min for the tests with bisection and 8min for those without bisection.

Problem	FEA	RED	BIS	NWT	RED FEA	BIS FEA	$\frac{NWT}{RED+BIS}$	Rel error
1	17	17	159	321	1.00	9.35	1.82	0.074167
2	9	11	114	134	1.22	12.66	1.07	0.253716
3	8	10	98	159	1.25	12.25	1.47	0.290124
4	8	10	103	138	1.25	12.87	1.22	0.248952
5	12	14	120	232	1.16	10.00	1.73	0.029818
1	14	14	0	64	1.00	0.00	4.57	0.094816
2	7	8	0	36	1.14	0.00	4.50	0.168162
3	6	7	0	32	1.16	0.00	4.57	0.420553
4	6	7	0	30	1.16	0.00	4.28	0.501800
5	10	11	0	49	1.10	0.00	4.45	0.128808

In Table 1 (problems with size 50×50) it can be seen that problem 3 (with or without bisection) has a high relative error, with a difference of almost 40%. This also applies to problem 5 without bisection. On the other hand, in Table 2 (problems with size 50×100) only problem 2 with bisection shows a high relative error, over 20% of difference. In Table 3 (problems with size 100×50), problem 2 and 4 (with and without bisection) showed the worst relative errors, somewhere around 40% and 20% of difference respectively. Finally, in Table 4 (problems with size 100×100), problems number 2, 3 and 4 (with bisection) also have high relative error, over 20% of difference. It is worth mentioning that for all tables, the number of ACRA calls (FEA) and the number of calls of the reduction procedure (RED) is considerably low when compared to the number of calls of the bisection procedure (BIS). As a final comment, we observe that when we compare the number of calls of the Newton method (NWT) for all the 20 test problems considered, the value obtained *without bisection* is approximately 25% of the value *with bisection*.

All problems were solved by computer codes implemented in MATLAB, running on a computer with an Intel Core 2 Duo – E7500, 2.93 GHz processor, with 3.9 GB of usable RAM.

6 CONCLUSIONS

We extended the known results in [5] for the linear case, using smooth convex regions.

We also presented a heuristic (ACRA) able to recover the approximate analytic center when we perturb (through the addition of new constraints) the initial region. It is worth recalling that this is fundamental when we work inside a cutting plane methodology that uses the approximate analytic center as a testing point.

In the sequence, we presented a procedure (CFA) that solves an important problem of applied mathematics, that is, the feasibility problem, using the heuristic ACRA.

Finally, the numerical results presented solving the feasibility problem for randomly-generated problems, show that the greatest number of iterations, due to the bisection routine, represents nearly ten times the calls to the routine ACRA, when we are interested in obtaining an approximation for the analytic center.

It is important to underline that the number of calls to ACRA in the second case, i.e. without the bisection routine, is smaller than when using the bisection routine. On the other hand, in many problems the number of calls to Newton's method almost duplicates in the second case. Finally, the number of calls to the reduction routine does not show meaningful variation in both cases.

All the tests in this work were carried out using randomly generated problems, mainly because of their easiness and readiness to use. Clearly, every improvement of CFA will depend on more clever ways to define the bisection routine. In the future, with such improvement, we hope to be able to test our heuristic with known library problems.

REFERENCES

- [1] BASESCU VL & MITCHELL JE. 2008. An analytic center cutting plane approach for conic programming. *Mathematics of Operations Research*, **33**: 529–551.
- [2] ENGAU A, ANJOS MF & BOMZE I. 2013. Constraint selection in a build-up interior-point cutting-plane method for solving relaxations of the stable-set problem. *Mathematical Methods of Operation Research*, **78**: 35–59.
- [3] ENGAU A, ANJOS MF & VANNELLI A. 2010. On interior-point warmstarts for linear and combinatorial optimization. *SIAM Journal on Optimization*, **20**(4): 1828–1861.
- [4] ENGAU A. 2012. Recent Progress in Interior-Point Methods: Cutting-Plane Algorithms and Warm Starts. Handbook on Semidefinite, Conic and Polynomial Optimization. *International Series in Operations Research & Management Science*, **166**: 471–498.
- [5] FEIJOO B, SANCHEZ A & GONZAGA CC. 1997. Maintaining closeness to the analytic center of a polytope by perturbing added hyperplanes. *Applied Mathematics & Optimization*, **35**: 139–144.

- [6] GOFFIN JL, LUO Z-Q & YE Y. 1996. Complexity Analysis of an Interior Cutting Plane Method for Convex Feasibility Problems. *SIAM Journal Optimization*, **6**: 638–652.
- [7] GONDZIO J. 2012. Interior Point Methods 25 years later. *European Journal of Operational Research*, **218**: 587–601.
- [8] GONDZIO J, GONZÁLEZ BREVIS P & MUNARI P. 2013. New developments in the primal-dual column generation technique. *European Journal of Operational Research*, **224**: 41–51.
- [9] HORN R & JOHNSON CR. 1985. *Matrix Analysis*. Cambridge University Press.
- [10] HUARD P & LIEU T. 1966. La Méthode des Centres dans un Espace Topologique. *Numerische Mathematik*, **8**: 56–67.
- [11] KARMARKAR N. 1984. A new polynomial-time algorithm for linear programming. *Combinatorica, Berlin*, **4**(2): 373–395.
- [12] LIU YZ & CHEN Y. 2010. *A Full-Newton Step Infeasible Interior-Point Algorithm for Linear Programming Based on a Special Self-Regular Proximity*. The Ninth International Symposium on Operations Research and Its Applications (ISORA'10) Chengdu-Jiuzhaigou, China, pp. 19–23.
- [13] MITCHELL JE. 2003. *Polynomial interior point cutting plane method*. Technical Report, Mathematical Sciences, Rensselaer Polytechnic Institute, Troy, NY, 12180.
- [14] MITCHELL JE & RAMASWAMY S. 2000. A Long-Step, Cutting Plane Algorithm for Linear and Convex Programming. *Annals of OR*, **99**: 95–122.
- [15] NESTEROV YE & NEMIROVSKY AS. 1989. *Self-Concordant Functions and Polynomial Time Methods in Convex Programming*. SIAM Studies in Applied Mathematics.
- [16] ROOS C. 2006. A full-Newton step $O(n)$ infeasible interior-point algorithm for linear optimization. *SIAM J. Optimization*, **16**(4): 1110–1136.
- [17] SKAJAA A, ANDERSEN ED & YE Y. 2013. Warm starting the homogeneous and self-dual interior point method for linear and conic quadratic problems. *Mathematical Programming Computation*, **5**: 1–25.
- [18] YE Y. 1992. A Potential Reduction Algorithm Allowing Column Generation. *SIAM Journal on Optimization*, **2**: 2–7.