Pré-Publicações do Departamento de Matemática
Universidade de Coimbra
Preprint Number 05–03

LOCAL ANALYSIS OF THE FEASIBLE PRIMAL-DUAL INTERIOR-POINT METHOD

R. SILVA, J. SOARES AND L. N. VICENTE

ABSTRACT: In this paper we analyze the rate of local convergence of the Newton primal-dual interior-point method when the iterates are kept strictly feasible with respect to the inequality constraints.

It is shown under the classical conditions that the rate is q–quadratic when the functions associated to the inequality constraints define a locally concave feasible region. In the nonconcave case, the q–quadratic rate is achieved provided the step in the primal variables does not become asymptotically orthogonal to any of the gradients of the binding inequality constraints.

KEYWORDS: Interior-point methods, strict feasibility, centrality, local convergence.
AMS SUBJECT CLASSIFICATION (2000): 90C30, 90C51.

1. Introduction

The local convergence theory of (infeasible) primal-dual interior-point methods for nonlinear programming was developed in the papers by El-Bakry *et al.* [4] and Yamashita and Yabe [12]. These papers show a q–quadratic rate of local convergence under the classical assumptions (second order sufficient optimality conditions, linear independence of the gradients of functions defining the binding constraints (LICQ), and strict complementarity). The study of q–superlinear convergence for quasi-Newton updates is reported in [8] and [12]. Furthermore, Vicente and Wright [10] proved a q–quadratic rate of convergence for a variant of the primal-dual interior-point method under degeneracy (replacing the LICQ by the Mangasarian–Fromowitz constraint qualification). In these approaches, the corresponding primal-dual interior-point method deals with the multipliers associated to both equality and inequality constraints as independent variables, and the primal-dual step is a Newton step for a perturbation of the first order necessary conditions for optimality. These approaches are infeasible since feasibility, corresponding to equality and, more importantly, to inequality constraints (rather than simple bounds), is only achieved asymptotically. Other rates of convergence

Received March 18, 2005.

Support for this work was provided by Centro de Matemática da Universidade de Coimbra and by FCT under grant POCTI/59442/MAT/2004.

for different interior-point methods for nonlinear programming have been established in [1], [3], [7], [9], and [11].

Gould, Orban, Sartenaer, and Toint [6] investigated the rate of convergence of primal-dual logarithmic barrier interior-point methods for linear equality constraints and general inequalities. The log-barrier approach maintains the iterates strictly feasible with respect to the inequality constraints, and the multipliers corresponding to the equalities are treated implicitly as dependent variables. The authors proved q–superlinear convergence, with a rate that may be chosen arbitrarily close to quadratic. Basically, they studied conditions under which a single primal-dual Newton step is strictly feasible and satisfies appropriate log-barrier subproblem termination criteria.

In this paper we analyze the rate of local convergence of the feasible primal-dual interior-point method along the lines of the analyses in [4] and [12]. The aspect considered is that inequality constraints are not converted into equalities using slack variables. The method keeps strict feasibility with respect to the inequality constraints. The other components of the primal-dual interior-point method remain essentially the same: the primal-dual step is a Newton step on the perturbed KKT first order optimality conditions and the various parameters are updated appropriately to induce a q–quadratic rate on the sequence of primal-dual iterates. To our knowledge this is the first time quadratic convergence is proved for interior-point methods for general nonlinear programming when inequalities are treated directly without the introduction of slack variables.

Our work on the feasible primal-dual interior-point method was motivated by an application to nonlinear programming problems arising from convex disjunctive programming [2]. In these problems, some of the functions might not be smooth or well-defined if certain inequalities are not satisfied strictly.

The material of this paper is organized in the following way. In Section 2, we describe the feasible primal-dual interior-point method in detail. The method is analyzed in Section 3, where it is shown that the iterates converge locally with a q–quadratic rate in the case of concave inequalities. The analysis includes the case where the step length is computed inexactly. The nonconcave case is discussed in Section 4. The rate remains q–quadratic for nonconcave inequalities as long as the primal component of the step is asymptotically nonorthogonal to the gradients of the (nonconcave) functions defining the binding inequalities. The paper is concluded in Section 5 with remarks about the interest and limitation of the analyzed approach.

2. The feasible primal-dual interior-point method

We consider the general nonlinear programming problem written in the form

$$\begin{array}{ll} \min & f(x), \\ \text{s.t.} & h(x) = 0, \\ & g(x) \leq 0, \end{array} \tag{1}$$

where $f : \mathbb{R}^n \longrightarrow \mathbb{R}$, $h : \mathbb{R}^n \longrightarrow \mathbb{R}^{m_h}$, and $g : \mathbb{R}^n \longrightarrow \mathbb{R}^{m_g}$. The assumptions on the differentiability of the functions f, g, and h will be stated later. The numbers m_h and m_g are assumed to be positive integers. The material of this paper remains valid in the case where there are no equality constraints ($m_h = 0$).

The Lagrangean function for problem (1) is $\ell : \mathbb{R}^{n+m_h+m_g} \longrightarrow \mathbb{R}$ defined by

$$\ell(x, y, z) = f(x) + h(x)^\top y + g(x)^\top z,$$

where x are the primal variables and the pair (y, z) represents the dual variables (or Lagrange multipliers). The gradient and the Hessian of ℓ with respect to the primal variables are given by

$$\begin{array}{rcl} \nabla_x \ell(x, y, z) & = & \nabla f(x) + \nabla h(x)\, y + \nabla g(x)\, z, \\ \nabla^2_{xx} \ell(x, y, z) & = & \nabla^2 f(x) + \sum_{j=1}^{m_h} y_j \nabla^2 h_j(x) + \sum_{j=1}^{m_g} z_j \nabla^2 g_j(x), \end{array}$$

whenever f, g, and h are continuously differentiable at x.

The Karush–Kuhn–Tucker (KKT) first order (necessary optimality) conditions for problem (1) are described by

$$\begin{array}{l} F_0(x, y, z) \ \stackrel{\text{def}}{=} \ \begin{pmatrix} \nabla_x \ell(x, y, z) \\ h(x) \\ -G(x) z \end{pmatrix} \ = \ 0, \\ g(x) \leq 0, \quad z \ \geq \ 0, \end{array} \tag{2}$$

where $G(x) = \mathrm{diag}(g(x))$. As we will see later, the primal-dual interior-point method is based on a perturbation of the conditions (2), given by

$$\begin{array}{l} F_\mu(x, y, z) \ \stackrel{\text{def}}{=} \ \begin{pmatrix} \nabla_x \ell(x, y, z) \\ h(x) \\ -G(x) z - \mu e \end{pmatrix} \ = \ 0, \\ g(x) \leq 0, \quad z \ \geq \ 0, \end{array}$$

where μ is a positive scalar and e is a vector of ones of dimension m_g. Note that, for $\hat{e} = (0, 0, e^\top)^\top \in \mathbb{R}^{n+m_h+m_g}$,

$$F_\mu(x, y, z) \;=\; F_0(x, y, z) - \mu \hat{e}. \tag{3}$$

The main part of the iterative step of the primal-dual interior-point method consists of the linearization of the perturbed KKT system. One computes a primal-dual step $\Delta w = (\Delta x, \Delta y, \Delta z)$, by solving the linear system of equations

$$F'_\mu(w)\, \Delta w = -F_\mu(w), \tag{4}$$

for fixed $w = (x, y, z)$ and $\mu > 0$, where $F'_\mu(w)$ is the Jacobian of $F_\mu(w)$. Notice that, from (3), $F'_\mu(w)$ is also the Jacobian of $F_0(w)$. The primal-dual system (4) can be written by blocks in the form

$$\begin{pmatrix} \nabla^2_{xx}\ell(x, y, z) & \nabla h(x) & \nabla g(x) \\ \nabla h(x)^\top & 0 & 0 \\ -Z\nabla g(x)^\top & 0 & -G(x) \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta y \\ \Delta z \end{pmatrix} = - \begin{pmatrix} \nabla_x \ell(x, y, z) \\ h(x) \\ -G(x)z - \mu e \end{pmatrix}, \tag{5}$$

where $Z = \mathrm{diag}(z)$.

Most variants of the primal-dual interior-point method keep positive all the variables subject to nonnegativity constraints. In our case, it means keeping the multipliers z positive. The parameter μ is driven to zero asymptotically. Since we are looking at the feasible variant of the primal-dual interior-point method, we must also keep $g(x)$ negative throughout the iterations. The main steps of this feasible variant are described below in Algorithm 2.1. For the purpose of analyzing local convergence, we do not include any stopping criterion.

Algorithm 2.1 (Feasible primal-dual interior-point method.).
Choose an initial point $w_0 = (x_0, y_0, z_0)$ with $g(x_0) < 0$ and $z_0 > 0$.

For $k = 0, 1, 2, \ldots$

Step 1.: Choose the parameter μ_k.

Step 2.: Compute the solution $\Delta w_k = (\Delta x_k, \Delta y_k, \Delta z_k)$ of the system (5), for $x = x_k$, $y = y_k$, and $z = z_k$.

Step 3.: Compute a positive step length α_k such that

$$g(x_k + \alpha_k \Delta x_k) \;<\; 0 \quad \text{and} \quad z_k + \alpha_k \Delta z_k \;>\; 0. \tag{6}$$

Step 4.: Define the next iterate $w_{k+1} = (x_{k+1}, y_{k+1}, z_{k+1})$ according to:

$$w_{k+1} = w_k + \alpha_k \Delta w_k. \tag{7}$$

Since the step size α_k must satisfy (6) throughout the iterations, we will impose that

$$\alpha_k \;=\; \min\left\{1,\ \tau_k \min_{i=1,\ldots,m_g}\left\{-\frac{(z_k)_i}{(\Delta z_k)_i}\ :\ (\Delta z_k)_i < 0\right\},\ \tau_k \min_{i=1,\ldots,m_g} \bar{\alpha}_k^i\right\}, \qquad (8)$$

where

$$\bar{\alpha}_k^i \;\leq\; \min\left\{\alpha :\ g_i(x_k + \alpha\Delta x_k) = 0,\ \alpha > 0\right\}, \qquad i = 1, \ldots, m_g. \qquad (9)$$

Whenever the minimum is not achieved, it is assumed by convention that it is set to $+\infty$.

We observe first that when the inequality constraints are of the simple bound type ($-x \leq 0$), the choice for α_k is of the type given above with the inequalities in (9) satisfied as equalities. In general, when the functions defining the inequality constraints are nonlinear, it might be computationally expensive to actually determine the step lengths $\bar{\alpha}_k^i$ such that

$$\bar{\alpha}_k^i \;=\; \min\left\{\alpha :\ g_i(x_k + \alpha\Delta x_k) = 0,\ \alpha > 0\right\}, \quad i = 1, \ldots, m_g. \qquad (10)$$

On the other hand, to get a fast rate of local convergence one cannot compute step lengths $\bar{\alpha}_k^i$ that differ too much from (10). However, it is possible to allow a certain inexactness in this computation. Let us define the residuals

$$r_k^i \;\stackrel{\mathrm{def}}{=}\; g_i(x_k + \bar{\alpha}_k^i \Delta x_k), \quad i = 1, \ldots, m_g.$$

We will show that the feasible primal-dual interior-point method will retain local q–quadratic convergence as long as the residuals r_k^i satisfy the condition

$$-r_k^i \;\leq\; \min\left\{\sigma(-g_i(x_k)), \frac{-g_i(x_k)c_1\|\Delta w_k\|}{1 + c_1\|\Delta w_k\|}\right\}, \quad i = 1, \ldots, m_g, \qquad (11)$$

where $\sigma \in (0,1)$ and $c_1 > 0$ are chosen independently of the iteration counter k.

Moreover, to achieve a q–quadratic rate of local convergence, the feasible primal-dual interior-point method must update the parameters τ_k and μ_k satisfying the classical conditions

$$1 - \tau_k \;\leq\; c_2\|F_0(w_k)\|, \qquad (12)$$

$$\mu_k \;\leq\; c_3\|F_0(w_k)\|^2, \qquad (13)$$

where c_2 and c_3 are constants independent of k. Vector and matrix norms in this paper are chosen to be the Euclidean ones.

3. Analysis of local convergence

We start by defining the concept of a strictly feasible neighborhood around a feasible point x_*. Our goal is to state the convergence results and develop the corresponding analysis in the smallest possible region of local smoothness allowed by the inequality constraints.

As we have pointed out before, the present work was actually motivated by the application of the primal-dual interior-point method to nonlinear programming problems arising from convex disjunctive programming [2]. These problems exhibit a local smoothness pattern which is similar to the one analyzed in this paper. Basically, the region of smoothness is locally confined to a strictly feasible neighborhood.

Definition 3.1. Given a feasible point $x_* \in \mathbb{R}^n$ ($h(x_*) = 0$ and $g(x_*) \leq 0$), we say that $\mathcal{N}(x_*; \epsilon)$ is a strictly feasible neighborhood of size $\epsilon > 0$ centered at x_* if

$$\mathcal{N}(x_*; \epsilon) \; = \; \{x_*\} \; \cup \; \{x \in \mathbb{R}^n : \; \|x - x_*\| < \epsilon \; \text{ and } \; g(x) < 0\}.$$

The local convergence of the feasible primal-dual interior-point method is analyzed at a point x_* satisfying the following assumptions.

(A1) The second order partial derivatives of these functions exist at x_*. There exists an $\epsilon > 0$, such that the functions f, g, and h are twice continuously differentiable in the set $\mathcal{N}(x_*; \epsilon) \backslash \{x_*\}$. Moreover, the second order partial derivatives of f, g, and h are Lipschitz continuous in $\mathcal{N}(x_*; \epsilon)$.

(A2) The point x_* is feasible and the gradients of the active constraints are linearly independent at x_*.

(A3) There exist Lagrange multipliers y_* and z_* such that $w_* = (x_*, y_*, z_*)$ satisfies the first order KKT conditions and the second order sufficient conditions and such that the pair $(-g(x_*), z_*)$ satisfies the strict complementarity condition $(-g(x_*) + z_* > 0)$.

Assumptions A2-A3 are the classical (nondegenerate) assumptions used to locally analyze interior-point methods. The difference from what is classically assumed is given in Assumption A1. We have restricted the smoothness of the functions defining the problem to a strictly feasible neighborhood centered at the point x_*, since the application of a feasible method might be motivated from the fact that the functions might not be defined or might not be smooth outside such set.

It results from Assumption A3 that the multipliers associated to the inequalities are nonnegative ($z_* \geq 0$) and also that

$$F_0(w_*) \;=\; 0. \tag{14}$$

We recall now the basic smoothness results that are required in the proof of the local convergence of the primal-dual interior-point method. Since we are interested in the feasible variant of this method, we present these results only for points in the strictly feasible neighborhood of x_* assured by Assumption A1. We extend now this neighborhood to all variables (primal x_* and dual (y_*, z_*)), defining the set

$$\mathcal{N}(w_*; \epsilon) = \{w_*\} \cup \{w = (x, y, z) \in \mathbb{R}^{n+m_h+m_g} : \; \|w - w_*\| < \epsilon \text{ and } g(x) < 0\},$$

where $w_* = (x_*, y_*, z_*)$. Note that the existence of the multipliers y_* and z_* is guaranteed by Assumptions A2-A3. Clearly, if $w = (x, y, z) \in \mathcal{N}(w_*; \epsilon)$ then $x \in \mathcal{N}(x_*; \epsilon)$.

Lemma 3.1. *Let x_* be a point for which Assumptions A1–A3 hold and $w_* = (x_*, y_*, z_*)$. Then, there exists a positive constant γ such that*

$$\begin{aligned} \|F_0(w^1) - F_0(w^2)\| &\leq \gamma \|w^1 - w^2\|, & (15)\\ \|F_0'(w^1) - F_0'(w^2)\| &\leq \gamma \|w^1 - w^2\|, & \\ \|F_0(w^1) - F_0(w^2) - F_0'(w^2)(w^1 - w^2)\| &\leq \tfrac{1}{2}\gamma \|w^1 - w^2\|^2, & (16)\end{aligned}$$

for all w^1 and w^2 in $\mathcal{N}(w_*; \epsilon)$.

The next lemma states that the primal-dual matrix is nonsingular around w_*, in the sense that is of interest to us. For a proof see, for instance, [5].

Lemma 3.2. *Let x_* be a point for which Assumptions A1–A3 hold and $w_* = (x_*, y_*, z_*)$. Then the following holds:*

(i) *$F_0'(w_*)$ is nonsingular;*

(ii) *$F_0'(w)$ is nonsingular for w in $\mathcal{N}(w_*; \epsilon_{ns})$, for some ϵ_{ns} satisfying $0 < \epsilon_{ns} < \epsilon$.*

From this lemma, it is assured the existence of a constant $\zeta > 0$ such that

$$\|F_0'(w)^{-1}\| \leq \zeta, \tag{17}$$

for all w in $\mathcal{N}(w_*; \epsilon_{ns})$. For such points w, the primal-dual step Δw given by the solution of the system (4) is well-defined and is equal to

$$\Delta w = -F_\mu'(w)^{-1} F_\mu(w). \tag{18}$$

The local asymptotic behavior of the feasible primal-dual interior-point method is studied first for concave binding inequalities.

(A4) The functions g_i, for $i \in \{1, \ldots, m_g\}$ such that $g_i(x_*) = 0$, are concave.

The main part of the analysis is spent proving a lower bound for the length of the step size parameter α_k. For this result we do not need τ_k and μ_k to satisfy the precise orders of magnitude given in the bounds (12) and (13), but rather being able to choose μ_k sufficiently small and τ_k sufficiently close to one.

Lemma 3.3. *Let x_* be a point for which Assumptions A1–A4 hold and $w_* = (x_*, y_*, z_*)$. Consider a sequence $\{w_k = (x_k, y_k, z_k)\}$ generated by the feasible primal-dual interior-point method described in Algorithm 2.1. If α_k satisfies (8)-(9) and (11) and τ_k and μ_k satisfy (12) and (13), then there exist positive constants ε and κ independent of k such that, when*

$$w_0 \in \mathcal{N}(w_*; \varepsilon), \tag{19}$$

the bound

$$1 - \alpha_k \leq (1 - \tau_k) + \kappa\zeta(\|F_0(w_k)\| + \mu_k\|\hat{e}\|), \tag{20}$$

holds for all iterates k.

Proof: First we have to set $\varepsilon = \epsilon_{ns}$, where ϵ_{ns} is given in Lemma 3.2.

Using (18), (3), (14), (17), (15), and (19) sequentially, it is easily derived the following bound for the primal-dual step:

$$\begin{aligned} \|\Delta w_k\| &= \|F'_\mu(w_k)^{-1} F_\mu(w_k)\| \\ &\leq \|F_\mu(w_k)^{-1}\|(\|F_0(w_k)\| + \mu_k\|\hat{e}\|) \\ &\leq \zeta(\gamma\|w_k - w^*\| + \mu_k\|\hat{e}\|) \\ &\leq \zeta(\gamma\varepsilon + \mu_k\|\hat{e}\|). \end{aligned}$$

Thus, from the condition (13) on the size of μ_k, and given a constant $\eta > 0$, one can reduce ε if necessary such that

$$\|\Delta w_k\| \leq \eta. \tag{21}$$

In particular, it is possible to choose a sufficiently small ε such that

$$\kappa\|\Delta w_k\| \leq \tau_k, \tag{22}$$

where κ is defined by

$$\kappa \;\stackrel{\mathrm{def}}{=}\; \max\left\{\frac{\kappa_2}{1-\sigma}, \kappa_1 + \kappa_1 c_1 \eta + c_1\right\}.$$

The constants κ_1 and κ_2 are given by

$$\kappa_1 \;=\; 2\max\left\{\frac{1}{(z_*)_i} \,:\, (z_*)_i > 0,\; i \in \{1,\ldots,m_g\}\right\}$$

and

$$\kappa_2 \;=\; 2M_{\nabla g}\max\left\{-\frac{1}{g_i(x_*)} \,:\, g_i(x_*) < 0,\; i \in \{1,\ldots,m_g\}\right\},$$

where $M_{\nabla g}$ is an upper bound on the size of ∇g in $\mathcal{N}(x_*;\epsilon)$.

We divide the proof in two separate cases: the case where the step length is defined by a multiplier and the case where the step length is defined by an inequality.

Case where step length is defined by a multiplier. In this first case we assume that there exists an index $i \in \{1,\ldots,m_g\}$ for which $(\Delta z_k)_i < 0$ and

$$\alpha_k \;=\; -\tau_k \frac{(z_k)_i}{(\Delta z_k)_i}.$$

If i is such that $(z_*)_i > 0$ then, from the definition of κ and from (22),

$$\alpha_k \;=\; \tau_k \frac{(z_k)_i}{-(\Delta z_k)_i} \;\geq\; \frac{\tau_k}{\kappa\|\Delta w_k\|} \;\geq\; 1.$$

When $(z_*)_i = 0$ (and $g_i(x_*) < 0$), we make use of the primal-dual block equation (see (5))

$$-Z_k \nabla g(x_k)^\top \Delta x_k - G(x_k)\Delta z_k \;=\; G(x_k) z_k + \mu_k e,$$

to write

$$-(z_k)_i \nabla g_i(x_k)^\top \Delta x_k - g_i(x_k)(\Delta z_k)_i \;=\; g_i(x_k)(z_k)_i + \mu_k,$$

or equivalently,

$$-\frac{(\Delta z_k)_i}{(z_k)_i} \;=\; 1 + \frac{\mu_k}{g_i(x_k)(z_k)_i} + p_k^i$$

with

$$p_k^i \;=\; \frac{\nabla g_i(x_k)^\top \Delta x_k}{g_i(x_k)} \;\leq\; \frac{|\nabla g_i(x_k)^\top \Delta x_k|}{-g_i(x_k)} \;\leq\; \kappa\|\Delta w_k\|.$$

Thus, since $\mu_k/(g_i(x_k)(z_k)_i) < 0$,

$$-\frac{(\Delta z_k)_i}{(z_k)_i} \;\leq\; 1 + \kappa\|\Delta w_k\|$$

and

$$\alpha_k \;=\; \tau_k \frac{(z_k)_i}{-(\Delta z_k)_i} \;\geq\; \frac{\tau_k}{1+\kappa\|\Delta w_k\|} \;\geq\; \tau_k(1-\kappa\|\Delta w_k\|).$$

Case where step length is defined by an inequality. Now we are interested in the case

$$\alpha_k \;=\; \tau_k \bar{\alpha}_k^i,$$

for some index $i \in \{1,\ldots,m_g\}$. By applying the mean value theorem, we have

$$r_k^i - g_i(x_k) \;=\; g_i(x_k + \bar{\alpha}_k^i \Delta x_k) - g_i(x_k) \;=\; \bar{\alpha}_k^i \nabla g_i(x_k + t_k^i \bar{\alpha}_k^i \Delta x_k)^\top \Delta x_k,$$

for some $t_k^i \in (0,1)$, and the step length $\bar{\alpha}_k^i$ can be written as

$$\bar{\alpha}_k^i \;=\; \frac{r_k^i - g_i(x_k)}{\nabla g_i(x_k + t_k^i \bar{\alpha}_k^i \Delta x_k)^\top \Delta x_k}.$$

Since $-r_k^i \leq \sigma(-g_i(x_k))$, both the numerator and the denominator in this expression for $\bar{\alpha}_k^i$ are positive.

If i is such that $g_i(x_*) < 0$ then, from the definitions of κ_2 and κ and from (22),

$$\begin{aligned}
\alpha_k \;=\; \tau_k \bar{\alpha}_k^i \;&\geq\; \tau_k \frac{(1-\sigma)(-g_i(x_k))}{\nabla g_i(x_k + t_k^i \bar{\alpha}_k^i \Delta x_k)^\top \Delta x_k} \\
&\geq\; \tau_k \frac{(1-\sigma)(-g_i(x_k))}{\|\nabla g_i(x_k + t_k^i \bar{\alpha}_k^i \Delta x_k)\|\,\|\Delta x_k\|} \\
&\geq\; \tau_k \frac{(1-\sigma)}{\kappa_2\|\Delta w_k\|} \\
&\geq\; \frac{\tau_k}{\kappa\|\Delta w_k\|} \\
&\geq\; 1.
\end{aligned}$$

When $g_i(x_*) = 0$ (and $(z_*)_i > 0$), we must first add and subtract

$$r_k^i \frac{(\Delta z_k)_i}{(z_k)_i} + r_k^i + r_k^i \frac{\mu_k}{g_i(x_k)(z_k)_i}$$

to the right hand side in the primal-dual equation

$$-\nabla g_i(x_k)^\top \Delta x_k \;=\; g_i(x_k) \frac{(\Delta z_k)_i}{(z_k)_i} + g_i(x_k) + g_i(x_k) \frac{\mu_k}{g_i(x_k)(z_k)_i}.$$

After division by $g_i(x_k) - r_k^i$, this results in

$$\begin{aligned} -\frac{\nabla g_i(x_k)^\top \Delta x_k}{g_i(x_k)^\top - r_k^i} \;&=\; \frac{(\Delta z_k)_i}{(z_k)_i} + 1 + \frac{\mu_k}{g_i(x_k)(z_k)_i} \\ &\quad + \frac{r_k^i}{g_i(x_k) - r_k^i} \frac{(\Delta z_k)_i}{(z_k)_i} + \frac{r_k^i}{g_i(x_k) - r_k^i} \\ &\quad + \frac{\mu_k r_k^i}{g_i(x_k)(z_k)_i (g_i(x_k) - r_k^i)}. \end{aligned}$$

Since the third and the sixth terms in the right hand side of this equality are negative and since, from (11),

$$\frac{r_k^i}{g_i(x_k) - r_k^i} \;\le\; c_1 \|\Delta w_k\|,$$

we obtain, from (21),

$$\begin{aligned} -\frac{\nabla g_i(x_k)^\top \Delta x_k}{g_i(x_k)^\top - r_k^i} \;&\le\; 1 + \kappa_1 \|\Delta w_k\| + \kappa_1 c_1 \|\Delta w_k\|^2 + c_1 \|\Delta w_k\| \\ &\le\; 1 + (\kappa_1 + \kappa_1 c_1 \eta + c_1) \|\Delta w_k\| \\ &\le\; 1 + \kappa \|\Delta w_k\|. \end{aligned}$$

Now, from the concavity of g_i, we derive

$$\begin{aligned} &-\frac{\nabla g_i(x_k + t_k^i \bar{\alpha}_k^i \Delta x_k)^\top \Delta x_k}{g_i(x_k) - r_k^i} \\ &= \; -\frac{\nabla g_i(x_k + t_k^i \bar{\alpha}_i \Delta x_k)^\top \Delta x_k}{g_i(x_k) - r_k^i} + \frac{\nabla g_i(x_k)^\top \Delta x_k}{g_i(x_k) - r_k^i} - \frac{\nabla g_i(x_k)^\top \Delta x_k}{g_i(x_k) - r_k^i} \\ &= \; \frac{[\nabla g_i(x_k + t_k^i \bar{\alpha}_k^i \Delta x_k) - \nabla g_i(x_k)]^\top \Delta x_k}{r_k^i - g_i(x_k)} - \frac{\nabla g_i(x_k)^\top \Delta x_k}{g_i(x_k) - r_k^i} \end{aligned}$$

$$\begin{aligned} &\leq\ -\frac{\nabla g_i(x_k)^\top \Delta x_k}{g_i(x_k) - r_k^i} \\ &\leq\ 1 + \kappa\|\Delta w_k\| \end{aligned} \tag{23}$$

and

$$\tau_k \bar{\alpha}_k^i \ = \ \tau_k \frac{r_k^i - g_i(x_k)}{\nabla g_i(x_k + t_k^i \bar{\alpha}_k^i \Delta x_k)^\top \Delta x_k} \ \geq \ \frac{\tau_k}{1 + \kappa\|\Delta w_k\|} \ \geq \ \tau_k(1 - \kappa\|\Delta w_k\|).$$

Conclusion. Combining all the four bounds derived for α_k (two in each case considered), one obtains

$$\alpha_k \ \geq \ \min\{1, \tau_k(1 - \kappa\|\Delta w_k\|)\} \ = \ \tau_k(1 - \kappa\|\Delta w_k\|) \ \geq \ \tau_k - \kappa\|\Delta w_k\|.$$

The last inequality above is based on the fact that $\tau_k < 1$ for all k provided ε is chosen small enough. Finally, from this lower bound on α_k, we get

$$0 \ \leq \ 1 - \alpha_k \ \leq \ (1 - \tau_k) + \kappa\|\Delta w_k\| \ \leq \ (1 - \tau_k) + \kappa\zeta(\|F_0(w_k)\| + \mu_k\|\hat{e}\|),$$

which concludes the proof of the lemma. ■

We can state now the q–quadratic rate of local convergence of Algorithm 2.1. The proof can be found in [12] and we describe it briefly for completeness. It is in this part of the theory that one uses the fact that τ_k and μ_k satisfy the orders of magnitude given in the bounds (12) and (13).

Theorem 3.1. *Let x_* be a point for which Assumptions A1–A4 hold and $w_* = (x_*, y_*, z_*)$. Consider a sequence $\{w_k = (x_k, y_k, z_k)\}$ generated by the feasible primal-dual interior-point method described in Algorithm 2.1. If α_k satisfies (8)-(9) and (11) and τ_k and μ_k satisfy (12) and (13), then there exists a positive constant ε independent of k such that, when*

$$w_0 \ \in \ \mathcal{N}(w_*; \varepsilon),$$

the sequence $\{w_k\}$ is well defined and converges to w_*. Moreover, we have

$$\|w_{k+1} - w_*\| \leq \nu\|w_k - w_*\|^2, \tag{24}$$

for all iterates k, where ν is a positive constant independent of k.

Proof: Let us assume that $\|w_k - w_*\| < \varepsilon$. By applying (7), (18), (3), and (14), we obtain

$$\begin{aligned} w_{k+1} - w_* \ &= \ (w_k - w_*) + \alpha_k \Delta w_k \\ &= \ (1 - \alpha_k)(w_k - w_*) \end{aligned}$$

$$+\, \alpha_k F'_{\mu_k}(w_k)^{-1}[F_0(w_*) - F_0(w_k) - F'_{\mu_k}(w_k)(w_* - w_k) + \mu_k \hat{e}].$$

Now, using (20), (16), (17), and $\alpha_k \leq 1$, we have, for a sufficiently small ε,

$$\begin{aligned} \|w_{k+1} - w_*\| &\leq (1-\alpha_k)\|w_k - w_*\| \\ &\quad + \alpha_k \|F'_{\mu_k}(w_k)^{-1}\| \, \|F_0(w_*) - F_0(w_k) - F'_{\mu_k}(w_k)(w_* - w_k)\| \\ &\quad + \alpha_k \mu_k \|F'_{\mu_k}(w_k)^{-1}\| \, \|\hat{e}\| \\ &\leq [(1-\tau_k) + \kappa\zeta \|F_0(w_k)\| + \kappa\zeta\|\hat{e}\|\mu_k] \, \|w_k - w_*\| \\ &\quad + \frac{\zeta\gamma}{2}\|w_k - w_*\|^2 + \|\hat{e}\|\zeta\mu_k. \end{aligned}$$

We also known that τ_k and μ_k satisfy (12) and (13). Thus, using the fact that $\|F_0(w_k)\| = \|F_0(w_k) - F_0(w_*)\| \leq \gamma\|w_k - w_*\|$, we assure the existence of a constant $\nu > 0$ independent of the iterates such that (24) holds. It is possible to prove by induction that $\{w_k\}$ converges to w_*, if ε is chosen sufficiently small. The inequality (24) shows that the local convergence rate is q–quadratic. ■

4. The nonconcave case

The concavity of the inequality constraint functions was required in (23) when the binding constraint function g_i was responsible for the step size α_k. However, one can see that the method retains a q–quadratic rate in the nonconcave case as long as there exists a positive constant β such that

$$r_k^i - g_i(x_k) \;=\; g_i(x_k + \bar{\alpha}_k^i \Delta x_k) - g_i(x_k) \;\geq\; \beta\bar{\alpha}_k^i \|\Delta x_k\| \tag{25}$$

for all k and all indices i corresponding to $g_i(x_*) = 0$. In fact, one would get

$$\frac{[\nabla g_i(x_k + t_k^i \bar{\alpha}_k^i \Delta x_k) - \nabla g_i(x_k)]^\top \Delta x_k}{r_k^i - g_i(x_k)} \;\leq\; \frac{L_{\nabla g_i}}{\beta}\|\Delta x_k\|,$$

where $L_{\nabla g_i}$ is the Lipschitz constant of the function ∇g_i in $\mathcal{N}(x_*;\epsilon)$. Then,

$$\begin{aligned} \frac{1}{\bar{\alpha}_k^i} \;=\; -\frac{\nabla g_i(x_k + t_k^i \bar{\alpha}_k^i \Delta x_k)^\top \Delta x_k}{g_i(x_k) - r_k^i} &\leq 1 + [(\kappa_1 + \kappa_1 c_1 \eta + c_1) + L_{\nabla g_i}/\beta)] \, \|\Delta w_k\| \\ &\leq 1 + \kappa\|\Delta w_k\|, \end{aligned}$$

after an appropriate redefinition of κ.

The bound (25) is satisfied as long as

$$\liminf_{k \longrightarrow +\infty} \nabla g_i(x_k)^\top \frac{\Delta x_k}{\|\Delta x_k\|} \;=\; 2\beta \;>\; 0. \tag{26}$$

To see why this is true let us expand $g_i(x_k + \bar{\alpha}_k^i \Delta x_k)$ around x_k:

$$g_i(x_k + \bar{\alpha}_k^i \Delta x_k) - g_i(x_k) \;=\; \bar{\alpha}_k^i \nabla g_i(x_k)^\top \Delta x_k + \frac{(\bar{\alpha}_k^i)^2}{2} \Delta x_k^\top \nabla^2 g_i(x_k + s_k^i \bar{\alpha}_k^i \Delta x_k) \Delta x_k,$$

for some $s_k^i \in (0,1)$. Since we are only looking at the cases where $\bar{\alpha}_k^i \leq 1$, one can see that

$$\frac{g_i(x_k + \bar{\alpha}_k^i \Delta x_k) - g_i(x_k)}{\bar{\alpha}_k^i \|\Delta x_k\|} \;\geq\; \beta$$

as long as

$$\|\Delta x_k\| \;\leq\; \|\Delta w_k\| \;\leq\; \frac{2\beta}{M_{\nabla^2 g_i}},$$

where $M_{\nabla^2 g_i}$ is an upper bound on the size of the Hessian $\nabla^2 g_i$ in $\mathcal{N}(x_*; \epsilon)$, requiring again a redefinition of κ.

In our preliminary numerical testing with the feasible primal-dual interior-point method we have clearly observed how condition (26) for small values of β deteriorated the convergence of the method. Decreasing values of β caused the method to take longer to achieve the region of fast local convergence.

Condition (26) has no influence if the constraint is concave because, even when this condition is not satisfied, the concavity of the function defining the constraint allows a locally full step ($\alpha_k = 1$) with respect to that constraint.

5. Concluding remarks

Keeping the iterates strictly feasible with respect to the inequality constraints in the way required for the feasible primal-dual interior-point method is a numerically expensive procedure. The exact computation of the step size would require the solution of a number of nonlinear equations per iteration. The inexact requirements (11) stated for the size of the residual of the equation solvers do alleviate the computational burden, but not by much. The proposed method would only become efficient for a small number of inequality constraints or when the structure of the functions defining the inequality constraints eases the step size calculation considerably.

However, the primal-dual interior-point method could be applied in a strictly feasible mode to only a subset of the problem inequalities, the remaining being treated by slack variables. The inequalities imposed strictly could be those for which the evaluation of the objective function or other constraint functions require their strict feasibility.

Despite the numerical considerations, looking at the infeasible variant of the primal-dual interior-point method is of interest on itself. It is worth pointing out that the approach presented in this paper covers the infeasible case ([4, 12]) since simple bounds of the type $x \geq 0 \Longleftrightarrow -x \leq 0$ correspond to concave inequalities. Linear inequality constraints are also concave and can be treated without slack variables for achieving the purpose of fast local convergence. Finally, the observation that the q–quadratic rate is retained in the general nonconcave case provided the angle between the primal step and the gradients of the binding constraints is kept away from ninety degrees, see (26), fits well into the theory of interior-point methods since it corresponds to the notion of centrality.

References

[1] R. H. Byrd, G. Liu, and J. Nocedal, *On the local behaviour of an interior point method for nonlinear programming*, in Numerical analysis 1997 (Dundee), vol. 380 of Pitman Res. Notes Math. Ser., Longman, Harlow, 1998, pp. 37–56.

[2] S. Ceria and J. Soares, *Convex programming for disjunctive convex optimization*, Math. Programming, 86 (1999), pp. 595–614.

[3] T. F. Coleman and Y. Li, *On the convergence of interior–reflective Newton methods for nonlinear minimization subject to bounds*, Math. Programming, 67 (1994), pp. 189–224.

[4] A. S. El-Bakry, R. A. Tapia, T. Tsuchiya, and Y. Zhang, *On the formulation and theory of the Newton interior–point method for nonlinear programming*, J. Optim. Theory Appl., 89 (1996), pp. 507–541.

[5] A. V. Fiacco and G. P. McCormick, *Nonlinear Programming: Sequential Unconstrained Minimization Techniques*, SIAM, Philadelphia, 1990.

[6] N. I. M. Gould, D. Orban, A. Sartenaer, and P. L. Toint, *Superlinear convergence of primal-dual interior point algorithms for nonlinear programming*, SIAM J. Optim., 11 (2001), pp. 974–1002.

[7] M. Heinkenschloss, M. Ulbrich, and S. Ulbrich, *Superlinear and quadratic convergence of affine-scaling interior-point Newton methods for problems with simple bounds without strict complementarity assumption*, Math. Programming, 86 (1999), pp. 615–635.

[8] H. J. Martinez, Z. Parada, and R. A. Tapia, *On the characterization of q-superlinear convergence of quasi-Newton interior-point methods for nonlinear programming*, Boletin de la Sociedad Matematica Mexicana, 1 (1995), pp. 1–12.

[9] L. N. Vicente, *Local convergence of the affine-scaling interior-point algorithm for nonlinear programming*, Comput. Optim. and Appl., 17 (2000), pp. 23–35.

[10] L. N. Vicente and S. J. Wright, *Local convergence of a primal-dual method for degenerate nonlinear programming*, Comput. Optim. and Appl., 22 (2002), pp. 311–328.

[11] A. Wchter and L. T. Biegler, *Line search filter methods for nonlinear programming: Local convergence*, SIAM J. Optim., (to appear).

[12] H. Yamashita and H. Yabe, *Superlinear and quadratic convergence of some primal-dual interior point methods for constrained optimization*, Math. Programming, 75 (1996), pp. 377–397.

R. Silva
PhD student at Departamento de Matemática, Universidade de Coimbra (renata@mat.uc.pt), supported by a scholarship from Fundação para a Ciência e a Tecnologia (FCT).

J. Soares
Departamento de Matemática, Universidade de Coimbra, 3001-454 Coimbra, Portugal (jsoares@mat.uc.pt).

L. N. Vicente
Departamento de Matemática, Universidade de Coimbra, 3001-454 Coimbra, Portugal (lnv@mat.uc.pt).