Comput Optim Appl (2010) 46: 347–368
DOI 10.1007/s10589-009-9265-2

Inexact Josephy–Newton framework for generalized equations and its applications to local analysis of Newtonian methods for constrained optimization

A.F. Izmailov · M.V. Solodov

Received: 21 October 2008 / Published online: 13 June 2009

Abstract We propose and analyze a perturbed version of the classical Josephy–Newton method for solving generalized equations. This perturbed framework is convenient to treat in a unified way standard sequential quadratic programming, its stabilized version, sequential quadratically constrained quadratic programming, and linearly constrained Lagrangian methods. For the linearly constrained Lagrangian methods, in particular, we obtain superlinear convergence under the second-order sufficient optimality condition and the strict Mangasarian–Fromovitz constraint qualification, while previous results in the literature assume (in addition to second-order sufficiency) the stronger linear independence constraint qualification as well as the strict complementarity condition. For the sequential quadratically constrained quadratic programming methods, we prove primal-dual superlinear/quadratic convergence under the same assumptions as above, which also gives a new result.

Keywords Newton method · Josephy–Newton method · Generalized equation · Variational problem · Linearly constrained Lagrangian method · (Stabilized) sequential quadratic programming

Research of the first author is supported by the Russian Foundation for Basic Research Grants 07-01-00270, 07-01-00416 and 08-01-90001-Bel, and by RF President's Grant NS-693.2008.1 for the support of leading scientific schools. The second author is supported in part by CNPq Grants 300513/2008-9 and 471267/2007-4, by PRONEX–Optimization, and by FAPERJ.

A.F. Izmailov
Moscow State University, Faculty of Computational Mathematics and Cybernetics, Department of Operations Research, Leninskiye Gory, GSP-2, 119992 Moscow, Russia
e-mail: izmaf@ccas.ru

M.V. Solodov (✉)
IMPA—Instituto de Matemática Pura e Aplicada, Estrada Dona Castorina 110, Jardim Botânico, Rio de Janeiro, RJ 22460-320, Brazil
e-mail: solodov@impa.br

1 Introduction

In this paper, we are interested in Newton and Newton-related methods for solving generalized equations (GE) of the form

$$\Phi(z) + N(z) \ni 0, \tag{1.1}$$

where $\Phi : \mathbf{R}^s \to \mathbf{R}^s$ is a smooth mapping, and $N : \mathbf{R}^s \to 2^{\mathbf{R}^s}$ is a set-valued mapping (i.e., for each $z \in \mathbf{R}^s$, $N(z)$ is a subset of $\mathbf{R}^s$). The so-called *Josephy–Newton method* (JNM) for (1.1) goes back to [18, 19]; it is the following iterative procedure. For the current iterate $z^k \in \mathbf{R}^s$, the next iterate z^{k+1} is computed as a solution of the (partially) linearized GE at z^k:

$$\Phi(z^k) + \Phi'(z^k)(z - z^k) + N(z) \ni 0. \tag{1.2}$$

A well-known and computationally important example of JNM is the sequential quadratic programming (SQP) method for optimization, e.g., [6]. To this end, consider the mathematical programming (MP) problem

$$\begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & F(x) = 0, \qquad G(x) \le 0, \end{array} \tag{1.3}$$

where $f : \mathbf{R}^n \to \mathbf{R}$ is a smooth function and $F : \mathbf{R}^n \to \mathbf{R}^l$ and $G : \mathbf{R}^n \to \mathbf{R}^m$ are smooth mappings. Stationary points of MP problem (1.3) and the associated Lagrange multipliers are characterized by the Karush–Kuhn–Tucker (KKT) optimality system

$$\frac{\partial L}{\partial x}(x, \lambda, \mu) = 0, \qquad F(x) = 0, \qquad \mu \ge 0, \qquad G(x) \le 0, \qquad \langle \mu, G(x)\rangle = 0, \tag{1.4}$$

where

$$L : \mathbf{R}^n \times \mathbf{R}^l \times \mathbf{R}^m \to \mathbf{R}, \quad L(x, \lambda, \mu) = f(x) + \langle \lambda, F(x)\rangle + \langle \mu, G(x)\rangle$$

is the Lagrangian function of MP problem (1.3).

KKT system (1.4) can be written in the form of GE (1.1) with $s = n + l + m$, $z = (x, \lambda, \mu) \in \mathbf{R}^s = \mathbf{R}^n \times \mathbf{R}^l \times \mathbf{R}^m$,

$$\Phi(z) = \left(\frac{\partial L}{\partial x}(x, \lambda, \mu), F(x), G(x)\right), \tag{1.5}$$

and

$$N(z) = N(\mu) = \begin{cases} \{0\} \times \{0\} \times \{y \in \mathbf{R}^m_+ \mid \langle \mu, y\rangle = 0\}, & \text{if } \mu \ge 0; \\ \emptyset, & \text{otherwise.} \end{cases} \tag{1.6}$$

In the case of GE associated to a KKT system, for a given $z^k = (x^k, \lambda^k, \mu^k) \in \mathbf{R}^s$, the iteration system (1.2) of JNM takes the form

$$\frac{\partial L}{\partial x}(x^k, \lambda^k, \mu^k) + \frac{\partial^2 L}{\partial x^2}(x^k, \lambda^k, \mu^k)(x - x^k)$$
$$+ (F'(x^k))^{\mathrm{T}}(\lambda - \lambda^k) + (G'(x^k))^{\mathrm{T}}(\mu - \mu^k) = 0,$$

$$F(x^k) + F'(x^k)(x - x^k) = 0,$$
$$\mu \geq 0, \qquad G(x^k) + G'(x^k)(x - x^k) \leq 0, \qquad \langle \mu, G(x^k) + G'(x^k)(x - x^k) \rangle = 0,$$

where the unknown is $z = (x, \lambda, \mu) \in \mathbf{R}^s$. As is easily seen, the latter is the KKT optimality system for the quadratic programming (QP) problem

$$\begin{aligned} &\text{minimize} \quad \langle f'(x^k), x - x^k \rangle + \frac{1}{2} \frac{\partial^2 L}{\partial x^2}(x^k, \lambda^k, \mu^k)[x - x^k, x - x^k] \\ &\text{subject to} \quad F(x^k) + F'(x^k)(x - x^k) = 0, \qquad G(x^k) + G'(x^k)(x - x^k) \leq 0, \end{aligned}$$

which is precisely the iteration/subproblem of SQP. It is interesting to point out that the strongest local convergence results for SQP that are currently available in the literature, follow from the analysis of general JNM in [7]; more on this in the sequel.

Let us go back to the general problem (1.1). The sharpest existing theory of local convergence and rate of convergence of JNM to a given solution $\bar{z}$ of GE (1.1) is developed in [7]. This theory relies on the following two notions.

Definition 1.1 Solution $\bar{z}$ of GE (1.1) is said to be *semistable* if for any $r \in \mathbf{R}^s$ close enough to zero, any solution $z(r)$ of the perturbed GE

$$\Phi(z) + N(z) \ni r \tag{1.7}$$

close enough to $\bar{z}$ satisfies the estimate

$$\|z(r) - \bar{z}\| = O(\|r\|).$$

If JNM generates a well-defined trajectory $\{z^k\}$ convergent to a semistable solution $\bar{z}$, the rate of convergence is necessarily superlinear. However, semistability does not guarantee solvability of subproblems (1.2) for z^k arbitrarily close to $\bar{z}$. To this end, the following notion comes into play.

Definition 1.2 Solution $\bar{z}$ of GE (1.1) is said to be *hemistable* if for any $z \in \mathbf{R}^s$ close enough to $\bar{z}$, GE

$$\Phi(z) + \Phi'(z)\zeta + N(z + \zeta) \ni 0 \tag{1.8}$$

has a solution $\zeta(z)$ such that $\zeta(z) \to 0$ as $z \to \bar{z}$.

Convergence results for JNM then affirm the following.

Theorem 1.1 ([7]) *Let $\Phi : \mathbf{R}^s \to \mathbf{R}^s$ be differentiable near a point $\bar{z} \in \mathbf{R}^s$, and suppose that the derivative of Φ is continuous at $\bar{z}$. Assume that $\bar{z}$ is a semistable and hemistable solution of GE (1.1).*

Then there exists $\delta > 0$ such that for any starting point $z^0 \in \mathbf{R}^s$ close enough to $\bar{z}$, there exists a trajectory $\{z^k\} \subset \mathbf{R}^s$ such that z^{k+1} is a solution of GE (1.2) for each $k = 0, 1, \ldots,$ satisfying $\|z^{k+1} - z^k\| \leq \delta$; any such trajectory converges to $\bar{z}$, and the *rate of convergence is superlinear. Moreover, the rate of convergence is quadratic provided the derivative of* Φ *is locally Lipschitz-continuous with respect to* $\bar{z}$*, i.e.*,

$$\|\Phi'(z) - \Phi'(\bar{z})\| = O(\|z - \bar{z}\|).$$

Generally, none of the two properties of semistability and hemistability is implied by the other. However, both are evidently implied by Robinson's strong regularity of the solution $\bar{z}$ [30]. Moreover, strong regularity implies that for any $z \in \mathbf{R}^s$ close enough to $\bar{z}$, GE (1.8) has a unique solution near $\bar{z}$.

Let us now mention what the general convergence result for JNM given in Theorem 1.1 means for some special cases of GEs. In the case of the usual equation

$$\Phi(z) = 0, \tag{1.9}$$

(i.e., when $N(\cdot) \equiv \{0\}$), both semistability and strong regularity are equivalent to saying that $\Phi'(\bar{z})$ is a nonsingular matrix. Moreover, in this case JNM is, of course, the usual *Newton method* (NM) for (1.9), with its iteration defined by the linear system

$$\Phi(z^k) + \Phi'(z^k)(z - z^k) = 0.$$

In this case, the assertion of Theorem 1.1 can be sharpened as follows: $\delta > 0$ can be taken arbitrarily, and for any $z^0 \in \mathbf{R}^s$ close enough to $\bar{z}$, the corresponding trajectory $\{z^k\}$ is unique. With these clarifications, Theorem 1.1 contains the standard local convergence and rate of convergence result for the classical Newton method.

Let us now consider again the optimization case and show how sharp convergence results for SQP are obtained from Theorem 1.1 for JNM.

Let $\bar{x} \in \mathbf{R}^n$ be a stationary point of MP problem (1.3), and let $(\bar{\lambda}, \bar{\mu}) \in \mathbf{R}^n \times \mathbf{R}^m$ be a Lagrange multiplier associated with $\bar{x}$, that is, the triple $(\bar{x}, \bar{\lambda}, \bar{\mu})$ satisfies (1.4). We assume that f, F and G are twice differentiable near $\bar{x}$, and their second derivatives are continuous at $\bar{x}$. Recall that the Mangasarian–Fromovitz constraint qualification (MFCQ) at $\bar{x}$ consists of saying that rank $F'(\bar{x}) = l$ and there exists $\bar{\xi} \in \ker F'(\bar{x})$ such that $G'_{A(\bar{x})}(\bar{x})\bar{\xi} < 0$, where $A(\bar{x}) = \{i = 1, \ldots, m \mid G_i(\bar{x}) = 0\}$ is the set of constraints active at $\bar{x}$. For MP problems, MFCQ is another name for the so-called Robinson's constraint qualification [8, Definition 2.86]. The strict Mangasarian–Fromovitz constraint qualification (SMFCQ) for $(\bar{\lambda}, \bar{\mu})$ consists of saying that

$$\operatorname{rank} \begin{pmatrix} F'(\bar{x}), \\ G'_{A_+(\bar{x},\bar{\mu})}(\bar{x}) \end{pmatrix} = l + |A_+(\bar{x}, \bar{\mu})|$$

and there exists $\bar{\xi} \in \ker F'(\bar{x})$ such that

$$G'_{A_+(\bar{x},\bar{\mu})}(\bar{x})\bar{\xi} = 0, \qquad G'_{A(\bar{x})\setminus A_+(\bar{x},\bar{\mu})}(\bar{x})\bar{\xi} < 0,$$

where $A_+(\bar{x}, \bar{\mu}) = \{i \in A(\bar{x}) \mid \bar{\mu}_i > 0\}$ is the set of strongly active constraints. SMFCQ is actually equivalent to the combination of MFCQ with the requirement that $(\bar{\lambda}, \bar{\mu})$ be the unique Lagrange multiplier associated with $\bar{x}$.

The following key facts were established in [7]:

- If the solution $\bar{z} = (\bar{x}, \bar{\lambda}, \bar{\mu})$ of GE (1.1) with $\Phi(\cdot)$ and $N(\cdot)$ defined according to (1.5) and (1.6), respectively, is semistable then $\bar{x}$ necessarily satisfies SMFCQ for $(\bar{\lambda}, \bar{\mu})$.
- If SMFCQ holds at $\bar{x}$ for $(\bar{\lambda}, \bar{\mu})$, and if the second-order sufficient optimality condition (SOSC) holds as well, then $\bar{z}$ is semistable. Here, SOSC means that

$$\frac{\partial L}{\partial x}(\bar{x}, \bar{\lambda}, \bar{\mu})[\xi, \xi] > 0 \quad \forall \xi \in C(\bar{x}) \setminus \{0\}, \tag{1.10}$$

where

$$C(\bar{x}) = \{\xi \in \mathbf{R}^n \mid F'(\bar{x})\xi = 0,\ G'_{A(\bar{x})}(\bar{x})\xi \le 0,\ \langle f'(\bar{x}), \xi\rangle \le 0\}$$

is the critical cone of the MP problem (1.3) at $\bar{x}$.
- If $\bar{x}$ is a local solution of MP problem (1.3) then SOSC (1.10) is also necessary for semistability of $\bar{z}$ and, moreover, semistability implies hemistability of $\bar{z}$.

It thus follows immediately from Theorem 1.1 (and the facts stated above) that local superlinear convergence of SQP method for MP problem (1.3) is guaranteed under SMFCQ and SOSC. We emphasize that this consequence of the analysis of general JNM in [7] is the sharpest known local convergence result for SQP. In addition to SOSC, other results in the literature assume the stronger linear independence constraint qualification instead of SMFCQ (e.g., [9, Theorem 15.4]) and sometimes also the strict complementarity condition $\bar{\mu}_{A(\bar{x})} > 0$ (e.g., [5, pp. 252–256] and [9, Theorem 15.2], or the original work [28]).

The discussion above puts in evidence that (exact) JNM for GE is a convenient and fruitful tool for analyzing SQP for optimization. In what follows, we shall extend the framework for dealing with Newton-related algorithms that can be regarded as *inexact JNM* (iJNM). These will include the stabilized version of SQP [12, 13, 16, 33–35], sequential quadratically constrained quadratic programming [2, 11, 15, 31], and linearly constrained Lagrangian methods [14, 24, 27]. Formally, instead of (1.2), the next iterate z^{k+1} would now satisfy the (perturbed) GE

$$\Phi(z^k) + \Phi'(z^k)(z - z^k) + \omega^k + N(z) \ni 0, \tag{1.11}$$

where $\omega^k \in \mathbf{R}^s$ is a perturbation term. This term may have different forms and meanings, may play various roles, and may conform to different sets of assumptions. This would depend on particular algorithms at hand, and on the particular purposes of the analysis. In Sect. 2, we prove convergence results for general iJNM. In Sect. 3, we show that the linearly constrained Lagrangian methods fall into our general framework. As a consequence, we obtain local superlinear/quadratic convergence under assumptions that are strictly weaker than those in the literature [14, 27]. In Sect. 4, we consider the sequential quadratically constrained quadratic programming method, and obtain a new convergence result (which improves [11] in the case of optimization, while being neither weaker nor stronger than the results in [2] and [15]). Finally, the stabilized version of SQP is considered in Sect. 5.

Before proceeding, we note that our development for iJNM was motivated, in part, by [13], which considers a different generalized perturbed version of JNM. The framework of [13] assumes a condition weaker than semistability, allowing the case of nonisolated solutions. In fact, the original version of our local convergence result for iJNM (1.11) could have been derived from [13]. However, the refined Theorem 2.1, given in the present version, cannot be obtained using [13]. Some further comments on this follow the proof of Theorem 2.1. Concerning specific optimization algorithms considered here, only stabilized SQP is analyzed in [13]; this application would be discussed in Sect. 5.

Our notation is mostly standard. By $\langle\cdot,\cdot\rangle$ we denote the Euclidian inner product, and by $\|\cdot\|$ the associated norm (the space is always clear from the context). For MP problem (1.3), we denote by $\sigma(\cdot)$ the natural residual function for its KKT system (1.4), i.e.,

$$\sigma:\mathbf{R}^n\times\mathbf{R}^l\times\mathbf{R}^m\to\mathbf{R},\quad \sigma(x,\lambda,\mu)=\left\|\left(\frac{\partial L}{\partial x}(x,\ \lambda,\ \mu),\ F(x),\min\{\mu,-G(x)\}\right)\right\|, \tag{1.12}$$

where the minimum is applied componentwise. Note that under our assumptions, $\sigma(\cdot)$ is Lipschitz-continuous near $\{\bar{x}\}\times\mathbf{R}^l\times\mathbf{R}^m$, and since $\sigma(\bar{x},\bar{\lambda},\bar{\mu})=0$, it holds that

$$\sigma(x,\lambda,\mu)=O(\|x-\bar{x}\|+\|\lambda-\bar{\lambda}\|+\|\mu-\bar{\mu}\|). \tag{1.13}$$

2 Inexact Josephy–Newton method for generalized equations

Similarly to other inexact frameworks in the literature, theoretical results regarding iJNM may be rather different in nature. For example, *a posteriori* analysis does not consider solvability of subproblems: it subsumes that the sequences $\{z^k\}\subset\mathbf{R}^s$ and $\{\omega^k\}\subset\mathbf{R}^s$ are given, with z^{k+1} satisfying (1.11) for each $k=0,1,\ldots,$ and one is interested under which assumptions (in particular, regarding $\{\omega^k\}$) convergence of $\{z^k\}$ to $\bar{z}$ is preserved, and the convergence rate estimates can be given. One example of this type of result is the famous Dennis–Moré Theorem [10] for usual equations, which is the theoretical basis for quasi-Newton methods. In contrast, *a priori* analysis does not subsume $\{z^k\}$ to be given: the role of $\{\omega^k\}$ is now primary with respect to $\{z^k\}$, and the question of solvability of subproblems (1.11) is an important part of the analysis. Of course, this can be possible only introducing some structure for the perturbation terms.

Speaking about inexact NM (iNM) for usual equations, *a priori* analysis is often possible and natural. However, for more complex GEs solvability of iJNM subproblems can generally be impossible to establish, at least without stronger assumptions which could be avoided otherwise. In such cases, the results which are in a sense intermediate between *a priori* and *a posteriori* can be more appropriate, with solvability of subproblems still being assumed, having in mind that this can be verified separately for some specific algorithms and/or problem classes. Theorem 2.1 below is an intermediate result of this kind, and it is a generalization of Theorem 1.1, allowing for inexactness. Solvability of subproblems will be treated later on, separately for specific algorithms.

It is convenient to start with an *a posteriori* result regarding the superlinear rate of convergence, assuming convergence itself.

Proposition 2.1 *Let a mapping $\Phi : \mathbf{R}^s \to \mathbf{R}^s$ be differentiable in a neighborhood of $\bar{z} \in \mathbf{R}^s$, with its derivative being continuous at $\bar{z}$. Let $\bar{z}$ be a semistable solution of GE (1.1). Let a sequence $\{z^k\} \subset \mathbf{R}^s$ be convergent to $\bar{z}$, and assume that z^{k+1} satisfies (1.11) for each $k = 0, 1, \ldots,$ with some $\omega^k \in \mathbf{R}^s$ such that*

$$\|\omega^k\| = o(\|z^{k+1} - z^k\| + \|z^k - \bar{z}\|). \tag{2.1}$$

Then the rate of convergence of $\{z^k\}$ is superlinear. Moreover, the rate of convergence is quadratic provided the derivative of Φ is locally Lipschitz-continuous with respect to $\bar{z}$, and provided

$$\|\omega^k\| = O(\|z^{k+1} - z^k\|^2 + \|z^{k+1} - z^k\|\|z^k - \bar{z}\| + \|z^k - \bar{z}\|^2). \tag{2.2}$$

Proof If $z^k = \bar{z}$ for some k, semistability of $\bar{z}$ implies that $z^k = \bar{z}$ for all subsequent values of k, and the assertions hold trivially. We therefore assume that $z^k \neq \bar{z}\ \forall\, k$.

For each k, z^{k+1} is a solution of GE (1.7) with

$$r = r^k = \Phi(z^{k+1}) - \Phi(z^k) - \Phi'(z^k)(z^{k+1} - z^k) - \omega^k, \tag{2.3}$$

and according to the Mean-Value Theorem (of the form in [20, Chap. XVII, §1.3]) and (2.1),

$$\begin{aligned} \|r^k\| &\le \sup\{\|\Phi'(tz^{k+1} + (1-t)z^k) - \Phi'(z^k)\| \mid t \in [0, 1]\}\|z^{k+1} - z^k\| + \|\omega^k\| \\ &= o(\|z^{k+1} - z^k\| + \|z^k - \bar{z}\|). \end{aligned} \tag{2.4}$$

Then, by semistability of $\bar{z}$, it holds that

$$\begin{aligned} \|z^{k+1} - \bar{z}\| &= O(\|r^k\|) \\ &= o(\|z^{k+1} - z^k\|) + o(\|z^k - \bar{z}\|) \\ &= o(\|z^{k+1} - \bar{z}\| + \|z^k - \bar{z}\|), \end{aligned}$$

i.e.,

$$\begin{aligned} 0 &= \lim_{k\to\infty} \frac{\|z^{k+1} - \bar{z}\|}{\|z^{k+1} - \bar{z}\| + \|z^k - \bar{z}\|} \\ &= \lim_{k\to\infty} \frac{1}{1 + \|z^k - \bar{z}\|/\|z^{k+1} - \bar{z}\|}. \end{aligned}$$

The latter relation implies that

$$\frac{\|z^k - \bar{z}\|}{\|z^{k+1} - \bar{z}\|} \to \infty \quad \text{as } k \to \infty,$$

i.e.,

$$\|z^{k+1} - \bar{z}\| = o(\|z^k - \bar{z}\|),$$

which shows the superlinear convergence rate of $\{z^k\}$.

Furthermore, if the derivative of Φ is locally Lipschitz-continuous with respect to $\bar{z}$, from (2.2) and (2.3) it follows that the estimate (2.4) can be sharpened as follows:

$$\|r^k\| = O(\|z^{k+1} - z^k\|^2 + \|z^{k+1} - z^k\|\|z^k - \bar{z}\| + \|z^k - \bar{z}\|^2).$$

Then by semistability of $\bar{z}$, we obtain

$$\begin{aligned}\|z^{k+1} - \bar{z}\| &= O(\|r^k\|)\\ &= O(\|z^{k+1} - z^k\|^2) + O(\|z^k - \bar{z}\|^2)\\ &= O(\|z^{k+1} - \bar{z}\|^2 + \|z^{k+1} - \bar{z}\|\|z^k - \bar{z}\| + \|z^k - \bar{z}\|^2),\end{aligned}$$

which means that the quantities

$$\begin{aligned}&\frac{\|z^{k+1} - \bar{z}\|}{\|z^{k+1} - \bar{z}\|^2 + \|z^{k+1} - \bar{z}\|\|z^k - \bar{z}\| + \|z^k - \bar{z}\|^2}\\ &\quad = \frac{1}{\|z^{k+1} - \bar{z}\| + \|z^k - \bar{z}\| + \|z^k - \bar{z}\|^2/\|z^{k+1} - \bar{z}\|}\end{aligned}$$

form a bounded sequence. But the latter is possible only when there exists $\gamma > 0$ such that

$$\frac{\|z^k - \bar{z}\|^2}{\|z^{k+1} - \bar{z}\|} \geq \gamma \quad \forall k,$$

i.e.,

$$\|z^{k+1} - \bar{z}\| \leq \frac{1}{\gamma}\|z^k - \bar{z}\|^2 \quad \forall k,$$

which gives the quadratic convergence rate of $\{z^k\}$. □

As an immediate application of Proposition 2.1, we can recover the result of [7, Theorem 2.1] concerning the quasi-Newton version of JNM. Specifically, let $\{J_k\} \subset \mathbf{R}^{n\times n}$ be a sequence of matrices satisfying the Dennis–Moré condition

$$(J_k - \Phi'(z^k))(z^{k+1} - z^k) = o(\|z^{k+1} - z^k\|). \tag{2.5}$$

For a current $z^k \in \mathbf{R}^n$, the next iterate z^{k+1} is computed as a solution of GE

$$\Phi(z^k) + J_k(z - z^k) + N(z) \ni 0, \tag{2.6}$$

which can be interpreted as (1.11) with

$$\omega^k = (J_k - \Phi'(z^k))(z^{k+1} - z^k).$$

Then, by (2.5), it follows that

$$\omega^k = o(\|z^{k+1} - z^k\|),$$

and hence, Proposition 2.1 implies the following *a posteriori* result for quasi-Newton JNM.

Corollary 2.1 *Let a mapping $\Phi : \mathbf{R}^n \to \mathbf{R}^n$ be differentiable in a neighborhood of $\bar{z} \in \mathbf{R}^n$, with its derivative being continuous at $\bar{z}$. Let $\bar{z}$ be a semistable solution of GE (1.1). Let $\{J_k\} \subset \mathbf{R}^{n\times n}$ be a sequence of matrices, and let $\{z^k\} \subset \mathbf{R}^n$ be a sequence convergent to $\bar{z}$ and such that z^{k+1} satisfies (2.6) for all k large enough. Assume, finally, that condition (2.5) holds.*

Then the rate of convergence of $\{z^k\}$ is superlinear.

We next present our main convergence result concerning iJNM. In what follows, we shall deal with iJNM with iteration subproblem of the form

$$\Phi(z^k) + \Phi'(z^k)(z - z^k) + \Omega(z^k, z - z^k) + N(z) \ni 0, \tag{2.7}$$

where $\Omega : \mathbf{R}^s \times \mathbf{R}^s \to 2^{\mathbf{R}^s}$ is a given multifunction. In other words, the perturbation term appearing in (1.11) must satisfy the inclusion $\omega^k \in \Omega(z^k, z^{k+1} - z^k)$.

Theorem 2.1 *Let $\Phi : \mathbf{R}^s \to \mathbf{R}^s$ be differentiable near a point $\bar{z} \in \mathbf{R}^s$, and suppose that the derivative of Φ is continuous at $\bar{z}$. Assume that $\bar{z}$ is a semistable solution of GE (1.1). Let $\Omega : \mathbf{R}^s \times \mathbf{R}^s \to 2^{\mathbf{R}^s}$ be a multifunction satisfying the following assumptions:*

(iJNM1) *For each $z \in \mathbf{R}^s$ close enough to $\bar{z}$, the GE*

$$\Phi(z) + \Phi'(z)\zeta + \Omega(z, \zeta) + N(z + \zeta) \ni 0 \tag{2.8}$$

has a solution $\zeta(z)$ such that $\zeta(z) \to 0$ as $z \to \bar{z}$.

(iJNM2) *The estimate*

$$\|\omega\| = o(\|\zeta\| + \|z - \bar{z}\|) \tag{2.9}$$

holds uniformly for $\omega \in \Omega(z, \zeta)$, $z \in \mathbf{R}^s$ and $\zeta \in \mathbf{R}^s$ close enough to zero and satisfying

$$\Phi(z) + \Phi'(z)\zeta + \omega + N(z + \zeta) \ni 0. \tag{2.10}$$

Then there exists $\delta > 0$ such that for any starting point $z^0 \in \mathbf{R}^s$ close enough to $\bar{z}$, there exists a trajectory $\{z^k\} \subset \mathbf{R}^s$ such that z^{k+1} is a solution of GE (2.7) for each $k = 0, 1, \ldots,$ satisfying

$$\|z^{k+1} - z^k\| \le \delta; \tag{2.11}$$

any such trajectory converges to $\bar{z}$, and the rate of convergence is superlinear. Moreover, the rate of convergence is quadratic provided the derivative of Φ is locally Lipschitz-continuous with respect to $\bar{z}$, and provided (2.9) in Assumption (iJNM2) *can be replaced by the estimate*

$$\|\omega\| = O(\|\zeta\|^2 + \|\zeta\|\|z - \bar{z}\| + \|z - \bar{z}\|^2). \tag{2.12}$$

Proof Semistability of $\bar{z}$ implies the existence of $\delta_1 > 0$ and $M > 0$ such that for any $r \in \mathbf{R}^s$ and any solution $z(r)$ of GE (1.7), satisfying $\|z(r) - \bar{z}\| \le \delta_1$, it holds that

$$\|z(r) - \bar{z}\| \le M\|r\|. \tag{2.13}$$

Fix any $\delta_2 \in (0, \delta_1]$. According to Assumption (iJNM1), there exists $\delta \in (0, 3\delta_2/5]$ such that the inequality $\|z^k - \bar{z}\| \leq 2\delta/3$ implies the existence of a solution z^{k+1} of GE (2.7) such that $\|z^{k+1} - \bar{z}\| \leq \delta_2$. Then z^{k+1} is a solution of GE (1.7) with $r = r^k$ defined in (2.3), and with some $\omega^k \in \Omega(z^k, z^{k+1} - z^k)$, and the inequality in (2.4) and condition (2.9) imply that

$$\|r^k\| \leq \frac{1}{5M}(\|z^{k+1} - z^k\| + \|z^k - \bar{z}\|), \tag{2.14}$$

perhaps for a smaller value of δ_2 (and hence, also of δ). Since $\delta_2 \leq \delta_1$, (2.13) holds with $r = r^k$ for $z^{k+1} = z(r^k)$. Hence, taking into account (2.14), we obtain

$$\begin{aligned}\|z^{k+1} - \bar{z}\| &\leq \frac{1}{5}\|z^{k+1} - z^k\| + \frac{1}{5}\|z^k - \bar{z}\| \\ &\leq \frac{1}{5}\|z^{k+1} - \bar{z}\| + \frac{2}{5}\|z^k - \bar{z}\|.\end{aligned}$$

This implies the inequality

$$\|z^{k+1} - \bar{z}\| \leq \frac{1}{2}\|z^k - \bar{z}\|, \tag{2.15}$$

which, in turn, implies that

$$\|z^{k+1} - \bar{z}\| \leq \frac{1}{3}\delta. \tag{2.16}$$

Hence,

$$\begin{aligned}\|z^{k+1} - z^k\| &\leq \|z^{k+1} - \bar{z}\| + \|z^k - \bar{z}\| \\ &\leq \delta.\end{aligned}$$

We thus proved that if $\|z^k - \bar{z}\| \leq 2\delta/3$ then GE (2.7) has a solution z^{k+1} satisfying (2.11).

Suppose now that $\|z^k - \bar{z}\| \leq 2\delta/3$, and z^{k+1} is any solution of GE (2.7) satisfying (2.11). Then

$$\begin{aligned}\|z^{k+1} - \bar{z}\| &\leq \|z^{k+1} - z^k\| + \|z^k - \bar{z}\| \\ &\leq \frac{5}{3}\delta \\ &\leq \delta_2 \\ &\leq \delta_1.\end{aligned}$$

Thus, $z^{k+1} = z(r^k)$ satisfies (2.13) with $r = r^k$, and by the same argument as above, the latter implies (2.15) and (2.16).

Therefore, if $\|z^0 - \bar{z}\| \leq 2\delta/3$ then the next iterate z^1 can be chosen in such a way that (2.11) would hold with $k = 0$, and any such choice will give (2.15) and (2.16) with $k = 0$. The latter implies that $\|z^1 - \bar{z}\| \leq 2\delta/3$. Hence, the next iterate z^2 can be chosen in such a way that (2.11) would hold with $k = 1$, and any such choice will give (2.15) and (2.16) with $k = 1$. Continuing this argument, we obtain that there exists a trajectory $\{z^k\}$ such that for each k, z^{k+1} is a solution of GE (2.7) satisfying (2.11), and for any such trajectory (2.15) is valid for all k. But the latter implies that $\{z^k\}$ converges to $\bar{z}$.

To complete the proof (with respect to the rate of convergence), it remains to invoke Proposition 2.1. □

For exact JNM (i.e., when $\Omega(\cdot) \equiv \{0\}$), Theorem 2.1 reduces to Theorem 1.1 (in particular, Assumption (iJNM1) reduces to hemistability).

We note that results *related* to Theorem 2.1 can be obtained from the framework of [13], adding to the latter the assumption of semistability. To see that Theorem 2.1 does not follow from [13], it is enough to note the difference between the "localization" condition (2.11) above and

$$\|z^{k+1} - z^k\| \leq \Delta \|z^k - \bar{z}\|$$

for some $\Delta > 0$, used in [13]. Since under our assumptions solutions of iJNM subproblems need not be unique (even locally), this difference means that Theorem 2.1 and [13] may refer, in principle, to different iterative sequences. Condition (2.11) appears somewhat more "practical", as it does not involve the unknown solution and merely requires the next iterate to be within some fixed distance to the previous one (a numerically natural assumption).

3 Linearly constrained Lagrangian methods

Linearly constrained Lagrangian (LCL) methods are traditionally stated for MP problems with equality constraints and bound constraints (inequality constraints are reformulated introducing slacks); see [14, 24, 27]. We therefore consider MP problem (1.3) with bound constraints given by $G(x) = -x$, $x \in \mathbf{R}^n$. For the current primal-dual iterate $(x^k, \lambda^k, \mu^k) \in \mathbf{R}^n \times \mathbf{R}^l \times \mathbf{R}^n$, the next primal iterate x^{k+1} of LCL method is computed as a stationary point of the subproblem of minimizing the (augmented) Lagrangian subject to bounds and linearized equality constraints:

$$\begin{array}{ll} \text{minimize} & f(x) + \langle \lambda^k, F(x)\rangle + \dfrac{c_k}{2}\|F(x)\|^2 \\ \text{subject to} & F(x^k) + F'(x^k)(x - x^k) = 0, \qquad x \geq 0, \end{array} \tag{3.1}$$

where $c_k \geq 0$ is the penalty parameter. The next dual iterate $(\lambda^{k+1}, \mu^{k+1})$ is of the form $(\lambda^k + u^k, \mu^{k+1})$, where (u^k, μ^{k+1}) is a Lagrange multiplier of problem (3.1), associated to the stationary point x^{k+1}.

Local superlinear convergence of LCL methods had been previously established under SOSC, the linear independence constraint qualification, and the strict complementarity condition (see [14, 27]). In what follows, we obtain a stronger result, replacing the linear independence constraint qualification by the weaker SMFCQ and removing strict complementarity.

Assuming that $c_k = c\ \forall k$ (which is natural for asymptotic analysis, see the discussion in [14]), the KKT system of subproblem (3.1) has the form

$$\begin{aligned} &f'(x) + (F'(x))^{\mathrm{T}}\lambda^k + c(F'(x))^{\mathrm{T}}F(x) + (F'(x^k))^{\mathrm{T}}u - \mu = 0,\\ &F(x^k) + F'(x^k)(x - x^k) = 0,\\ &\mu \geq 0, \qquad x \geq 0, \qquad \langle \mu, x\rangle = 0. \end{aligned} \tag{3.2}$$

Observe that the last two lines of system (3.2), associated to the bound and linearized constraints, are exactly the same as in the SQP iteration for MP problem (1.3) at hand. Structural perturbation that defines LCL within iJNM framework is therefore given by the first line of (3.2). In particular, exact LCL method is a special case of iJNM (2.7) for GE (1.1) with $\Phi(\cdot)$ and $N(\cdot)$ defined according to (1.5) and (1.6), respectively, with the perturbation term being $\Omega(z, \zeta) = \{\omega(z, \zeta)\}$, with

$$\begin{aligned} \omega(z, \zeta) = \bigg(&\frac{\partial L}{\partial x}(x + \xi, \lambda, \mu) - \frac{\partial L}{\partial x}(x, \lambda, \mu)\\ &- \frac{\partial^2 L}{\partial x^2}(x, \lambda, \mu)\xi + c(F'(x + \xi))^{\mathrm{T}}F(x + \xi), 0, 0\bigg), \end{aligned} \tag{3.3}$$

where $z = (x, \lambda, \mu) \in \mathbf{R}^s$, $\zeta = (\xi, u, v) \in \mathbf{R}^s$.

We are now in position to state our convergence result. The proof is by verifying Assumptions (iJNM1) and (iJNM2), thus obtaining the assertions of Theorem 2.1.

Corollary 3.1 *Let $f : \mathbf{R}^n \to \mathbf{R}$ and $F : \mathbf{R}^n \to \mathbf{R}^l$ be twice differentiable near a point $\bar{x} \in \mathbf{R}^n$, and suppose that the second derivatives of f and F are continuous at $\bar{x}$. Assume that $\bar{x}$ is a local solution of MP problem (1.3) with $G(x) = -x$, $x \in \mathbf{R}^n$, satisfying SMFCQ and SOSC (1.10) for the associated Lagrange multiplier $(\bar{\lambda}, \bar{\mu}) \in \mathbf{R}^l \times \mathbf{R}^m$.*

Then for any fixed $c \geq 0$, there exists $\delta > 0$ such that for any starting point $(x^0, \lambda^0, \mu^0) \in \mathbf{R}^n \times \mathbf{R}^l \times \mathbf{R}^m$ close enough to $(\bar{x}, \bar{\lambda}, \bar{\mu})$, there exists a trajectory $\{(x^k, \lambda^k, \mu^k)\} \subset \mathbf{R}^n \times \mathbf{R}^l \times \mathbf{R}^m$ such that for each $k = 0, 1, \ldots,$ the triple $(x^{k+1}, \lambda^{k+1} - \lambda^k, \mu^{k+1})$ satisfies the system (3.2), and also satisfies

$$\|(x^{k+1} - x^k, \lambda^{k+1} - \lambda^k, \mu^{k+1} - \mu^k)\| \leq \delta; \tag{3.4}$$

any such trajectory converges to $(\bar{x}, \bar{\lambda}, \bar{\mu})$, and the rate of convergence is superlinear. Moreover, the rate of convergence is quadratic provided the second derivatives of f and of F are locally Lipschitz-continuous with respect to $\bar{x}$.

Proof We start with (iJNM1). For each fixed $(x, \lambda) \in \mathbf{R}^n \times \mathbf{R}^l$ regarded as a parameter, consider the MP problem

$$\begin{aligned} \text{minimize}_\xi \quad & f(x + \xi) + \langle \lambda, F(x + \xi)\rangle + \frac{c}{2}\|F(x + \xi)\|^2\\ \text{subject to} \quad & F(x) + F'(x)\xi = 0,\\ & x + \xi \geq 0. \end{aligned} \tag{3.5}$$

It can be easily checked that for the base parameter value $(x, \lambda) = (\bar{x}, \bar{\lambda})$, $\bar{\xi} = 0$ is a stationary point for this problem, and $(0, \bar{\mu})$ is an associated Lagrange multiplier. Furthermore, SMFCQ and SOSC (1.10) hold for this problem at this point for this multiplier. Hence, $\bar{\xi} = 0$ is a strict local minimizer of the problem associated to the base parameter value and it satisfies MFCQ. It then follows (e.g., from Robinson's stability theorem [29] and, e.g., from [3, Theorem 3.1]) that for each $(x, \lambda) \in \mathbf{R}^n \times \mathbf{R}^l$ close enough to $(\bar{x}, \bar{\lambda})$, the problem (3.5) has a local solution $\xi(x, \lambda) \in \mathbf{R}^n$ such that $\xi(x, \lambda) \to 0$ as $(x, \lambda) \to (\bar{x}, \bar{\lambda})$. Since MFCQ is stable under small perturbations (see, e.g., [8, Remark 2.88]), we conclude that for $(x, \lambda) \in \mathbf{R}^n \times \mathbf{R}^l$ close enough to $(\bar{x}, \bar{\lambda})$, MFCQ holds at $\xi(x, \lambda)$. Hence, by the standard first-order necessary optimality conditions (see, e.g., [8, Theorem 3.9]) for problem (3.5), there exists $(u(x, \lambda), \tilde{\mu}(x, \lambda)) \in \mathbf{R}^l \times \mathbf{R}^n$ such that the triple $(\xi(x, \lambda), u(x, \lambda), \tilde{\mu}(x, \lambda))$ satisfies the KKT system

$$
\begin{aligned}
&f'(x+\xi) + (F'(x+\xi))^{\mathrm{T}}\lambda + c(F'(x+\xi))^{\mathrm{T}}F(x+\xi) + (F'(x))^{\mathrm{T}}u - \tilde{\mu} = 0,\\
&F(x) + F'(x)\xi = 0, \qquad\qquad (3.6)\\
&\tilde{\mu} \geq 0, \qquad x+\xi \geq 0, \qquad \langle \tilde{\mu}, x+\xi \rangle = 0.
\end{aligned}
$$

Since under SMFCQ and SOSC primal-dual solutions of parametric KKT systems have locally upper Lipschitzian behaviour (see [8, Theorem 5.9]), we obtain the estimate

$$
\|\xi(x, \lambda)\| + \|u(x, \lambda)\| + \|\tilde{\mu}(x, \lambda) - \bar{\mu}\| = O(\|x - \bar{x}\| + \|\lambda - \bar{\lambda}\|), \qquad (3.7)
$$

and in particular, $(\xi(x, \lambda), u(x, \lambda), \tilde{\mu}(x, \lambda)) \to (0, 0, \bar{\mu})$ as $(x, \lambda) \to (\bar{x}, \bar{\lambda})$. Fixing an arbitrary pair $(\lambda, \mu) \in \mathbf{R}^l \times \mathbf{R}^n$, and denoting $v = \tilde{\mu} - \mu$, we can re-write the KKT system (3.6) as follows:

$$
\begin{aligned}
&\frac{\partial L}{\partial x}(x, \lambda, \mu) + \frac{\partial^2 L}{\partial x^2}(x, \lambda, \mu)\xi + (F'(x))^{\mathrm{T}}u - v\\
&\quad + \frac{\partial L}{\partial x}(x+\xi, \lambda, \mu) - \frac{\partial L}{\partial x}(x, \lambda, \mu)\\
&\quad - \frac{\partial^2 L}{\partial x^2}(x, \lambda, \mu)\xi + c(F'(x+\xi))^{\mathrm{T}}F(x+\xi) = 0,\\
&F(x) + F'(x)\xi = 0,\\
&\mu + v \geq 0, \qquad x+\xi \geq 0, \qquad \langle \mu + v, x+\xi \rangle = 0.
\end{aligned}
$$

The latter system is precisely (2.8) for LCL method, with $z = (x, \lambda, \mu)$ and $\zeta = (\xi, u, v)$. Denoting $v(x, \lambda, \mu) = \tilde{\mu}(x, \lambda) - \mu$, and setting $\zeta(z) = (\xi(x, \lambda), u(x, \lambda), v(x, \lambda, \mu))$, we finally obtain from (3.7) that

$$
\begin{aligned}
\|\zeta(z)\| &\leq \|\xi(x, \lambda)\| + \|u(x, \lambda)\| + \|\tilde{\mu}(x, \lambda) - \bar{\mu}\| + \|\mu - \bar{\mu}\|\\
&= O(\|x - \bar{x}\| + \|\lambda - \bar{\lambda}\|) + \|\mu - \bar{\mu}\|\\
&= O(\|z - \bar{z}\|), \qquad (3.8)
\end{aligned}
$$

and in particular $\zeta(z) \to 0$ as $z \to \bar{z}$, with $\bar{z} = (\bar{x}, \bar{\lambda}, \bar{\mu})$. Hence, Assumption (iJNM1) is verified.

We next check Assumption (iJNM2). According to the argument above, estimate (3.8) is valid for $z = (x, \lambda, \mu) \in \mathbf{R}^s$ and for any solution $\zeta(z) = (\xi(x, \lambda), u(x, \lambda), v(x, \lambda, \mu))$ of GE (2.10), such that $\zeta(z)$ is close enough to zero. Moreover, for any such z and $\zeta(z)$, it holds that $F(x) + F'(x)\xi(x, \lambda) = 0$, and hence, employing (3.3) and the Mean-Value Theorem, we derive the estimate

$$\begin{aligned}
\|\omega(z, \zeta(z))\| &\leq \left\| \frac{\partial L}{\partial x}(x + \xi(x, \lambda), \lambda, \mu) - \frac{\partial L}{\partial x}(x, \lambda, \mu) - \frac{\partial^2 L}{\partial x^2}(x, \lambda, \mu)\xi(x, \lambda) \right\| \\
&\quad + c\|(F'(x + \xi(x, \lambda)))^{\mathrm{T}}(F(x + \xi(x, \lambda)) - F(x) - F'(x)\xi(x, \lambda))\| \\
&= o(\|\xi(x, \lambda)\|) \\
&= o(\|\zeta(z)\|) \\
&= o(\|z - \bar{z}\|), \qquad (3.9)
\end{aligned}$$

which verifies Assumption (iJNM2). Moreover, if the second derivatives of f and F are locally Lipschitz-continuous with respect to $\bar{x}$, the right-hand side of (3.9) can be replaced by $O(\|z - \bar{z}\|^2)$, so that (iJNM2) holds with (2.9) replaced by (2.12).

All the assertions now follow applying Theorem 2.1. □

We emphasize that the result above is stronger than previous local convergence results for LCL methods. Apart from SOSC, all the existing results of this kind (see [14, 27]) assume the linear independence constraint qualification (stronger than SMFCQ) and in addition strict complementarity, and they assert R-superlinear (rather than Q-superlinear) convergence.

Taking into account that (3.1) is not a QP, it is difficult to expect that the subproblems can be solved exactly. Motivated by this, apart from interpreting exact LCL as iJNM, let us introduce an extra perturbation associated to inexact solution of LCL subproblems. Specifically, it seems natural to consider the *truncated LCL* (tLCL) method, where (3.2) is replaced by a version where the parts corresponding to general nonlinearities are relaxed:

$$\begin{aligned}
&\|f'(x) + (F'(x))^{\mathrm{T}}\lambda^k + c(F'(x))^{\mathrm{T}}F(x) + (F'(x^k))^{\mathrm{T}}u - \mu\| \leq \varphi(\sigma(x^k, \lambda^k, \mu^k)), \\
&\|F(x^k) + F'(x^k)(x - x^k)\| \leq \varphi(\sigma(x^k, \lambda^k, \mu^k)), \qquad (3.10) \\
&\mu \geq 0, \qquad x \geq 0, \qquad \langle \mu, x \rangle = 0,
\end{aligned}$$

where $\sigma(\cdot)$ is the natural residual (1.12) measuring violation of optimality conditions (1.4) for the MP problem (1.3), and $\varphi : \mathbf{R} \to \mathbf{R}_+$ is some forcing function (specific conditions on φ will be imposed later on). Note that the first line of (3.10) can be re-written as follows:

$$\varphi(\sigma(x^k, \lambda^k, \mu^k)) \geq \left\| \frac{\partial L}{\partial x}(x, \lambda^k, \mu^k) + (F'(x^k))^{\mathrm{T}}u - (\mu - \mu^k) + c(F'(x))^{\mathrm{T}}F(x) \right\|$$

$$= \left\| \frac{\partial L}{\partial x}(x^k, \lambda^k, \mu^k) + \frac{\partial^2 L}{\partial x^2}(x^k, \lambda^k, \mu^k)(x - x^k) \right.$$
$$+ (F'(x^k))^{\mathrm{T}} u - (\mu - \mu^k)$$
$$+ \left(\frac{\partial L}{\partial x}(x, \lambda^k, \mu^k) - \frac{\partial L}{\partial x}(x^k, \lambda^k, \mu^k) \right.$$
$$- \frac{\partial^2 L}{\partial x^2}(x^k, \lambda^k, \mu^k)(x - x^k)$$
$$\left. \left. + c(F'(x))^{\mathrm{T}} F(x) \right) \right\|.$$

We therefore can define the total perturbation (which combines that with respect to SQP with that with respect to exact LCL) as

$$\Omega(z, \zeta) = \omega(z, \zeta) + \Theta(z),$$

with $\omega(z, \zeta)$ defined according to (3.3), and with

$$\Theta(z) = \{\theta \in \mathbf{R}^s \mid \|\theta\| \leq \varphi(\sigma(x, \lambda, \mu))\}, \tag{3.11}$$

where $z = (x, \lambda, \mu) \in \mathbf{R}^s$, $\zeta = (\xi, u, v) \in \mathbf{R}^s$. In particular, tLCL method is a special case of iJNM (2.7) for GE (1.1) with $\Phi(\cdot)$ and $N(\cdot)$ defined according to (1.5) and (1.6), respectively. Separating the perturbation into the single-valued part $\omega(z, \zeta)$ and the set valued-part $\Theta(z)$ is instructive, because the two parts correspond to inexactness of different kind: $\omega(z, \zeta)$ stands for structural inexactness of LCL methods with respect to SQP, while $\Theta(z)$ stands for additional inexactness allowed when solving LCL subproblems.

Corollary 3.2 *Under the assumptions of Corollary 3.1, for any fixed $c \geq 0$, there exists $\delta > 0$ such that for any function $\varphi : \mathbf{R}_+ \to \mathbf{R}_+$ such that $\varphi(t) = o(t)$, and for any starting point $(x^0, \lambda^0, \mu^0) \in \mathbf{R}^n \times \mathbf{R}^l \times \mathbf{R}^m$ close enough to $(\bar{x}, \bar{\lambda}, \bar{\mu})$, there exists a trajectory $\{(x^k, \lambda^k, \mu^k)\} \subset \mathbf{R}^n \times \mathbf{R}^l \times \mathbf{R}^m$ such that for each $k = 0, 1, \ldots,$ the triple $(x^{k+1}, \lambda^{k+1} - \lambda^k, \mu^{k+1})$ satisfies the system (3.10) and (3.4); any such trajectory converges to $(\bar{x}, \bar{\lambda}, \bar{\mu})$, and the rate of convergence is superlinear. Moreover, the rate of convergence is quadratic provided the second derivatives of f and F are locally Lipschitz-continuous with respect to $\bar{x}$, and provided $\varphi(t) = O(t^2)$.*

Proof Assumption (iJNM1) was established in the proof of Corollary 3.1.

Assuming that $z = (x, \lambda, \mu) \in \mathbf{R}^s$ and $\zeta = (\xi, u, v)$ satisfy GE (2.10) with $\omega \in \Omega(z, \zeta)$, we have that, in particular,

$$\|F(x) + F'(x)\xi\| \leq \varphi(\sigma(x, \lambda, \mu)).$$

Thus, by (3.3), employing the Mean-Value Theorem and (1.13), we obtain the estimate

$$\|\omega(z, \zeta)\| \leq \left\| \frac{\partial L}{\partial x}(x + \xi, \lambda, \mu) - \frac{\partial L}{\partial x}(x, \lambda, \mu) - \frac{\partial^2 L}{\partial x^2}(x, \lambda, \mu)\xi \right\|$$

$$
\begin{aligned}
&+ c\|(F'(x+\xi))^{\mathrm{T}}(F(x+\xi) - F(x) - F'(x)\xi)\| + O(\varphi(\sigma(x,\lambda,\mu))) \\
&= o(\|\xi\|) + o(\sigma(x,\lambda,\mu)) \\
&= o(\|\zeta\|) + o(\|z-\bar{z}\|).
\end{aligned}
$$

Furthermore, by (1.13) and (3.11), for each $\theta \in \Theta(z,\zeta)$ it holds that

$$
\begin{aligned}
\|\theta\| &\le \varphi(\sigma(x,\lambda,\mu)) \\
&= o(\sigma(x,\lambda,\mu)) \\
&= o(\|z-\bar{z}\|).
\end{aligned}
$$

Combining the latter two estimates, we obtain that for each $\omega(= \omega(z,\zeta) + \theta) \in \Omega(z,\zeta)$, the estimate (2.9) holds, and hence, Assumption (iJNM2) is satisfied.

Moreover, assuming that the second derivatives of f and F are locally Lipschitz-continuous with respect to $\bar{x}$, and that $\varphi(t) = O(t^2)$, it can be easily checked that (2.9) can be replaced by the stronger estimate (2.12).

The assertions now follow applying Theorem 2.1. □

4 Sequential quadratically constrained quadratic programming

The sequential quadratically constrained quadratic programming (SQCQP) method is the following iterative procedure. For the current primal iterate $x^k \in \mathbf{R}^n$, the next iterate x^{k+1} is computed as a stationary point of the subproblem

$$
\begin{aligned}
\text{minimize} \quad & \langle f'(x^k), x - x^k\rangle + \frac{1}{2} f''(x^k)[x-x^k, x-x^k] \\
\text{subject to} \quad & F(x^k) + F'(x^k)(x-x^k) + \frac{1}{2}F''(x^k)[x-x^k, x-x^k] = 0, \\
& G(x^k) + G'(x^k)(x-x^k) + \frac{1}{2}G''(x^k)[x-x^k, x-x^k] \le 0,
\end{aligned}
\tag{4.1}
$$

with quadratic objective function and quadratic constraints. The dual sequence $\{(\lambda^k, \mu^k)\}$ is given by Lagrange multipliers associated to the primal sequence. We note, in the passing, that SQCQP is in principle a primal algorithm, since dual variables are not used to formulate the subproblems. Nevertheless, dual behaviour is certainly important, as most standard stopping criteria for the MP problem (1.3) are based on some primal-dual measure of optimality, such as the natural residual (1.12).

As some previous work on SQCQP and related methods, we mention [2, 11, 15, 21, 26, 31, 32]. In the convex case, subproblem (4.1) can be cast as a second-order cone program [23, 25], which can be solved efficiently by interior-point algorithms. Another possibility for the convex case is [17]. In [2], nonconvex subproblems were also handled quite efficiently by using other computational techniques. In the nonconvex case, one might also use [1, 4] for solving the subproblems.

We next discuss previous local convergence results for SQCQP (we note that they concern inequality-constrained problems, i.e., there is no F part in the MP problem (1.3) and in the iteration (4.1)). In [2], local primal superlinear rate of convergence of a trust-region SQCQP method was obtained under the MFCQ and a certain quadratic growth condition (under MFCQ, the latter is equivalent to SOSC, see [8, Theorem 3.70]). Quadratic convergence of the primal-dual sequence was obtained in [15] under the convexity assumptions on f and on G, MFCQ, and strong SOSC. In [11], primal-dual quadratic convergence was established under the linear independence constraint qualification, SOSC, and strict complementarity. In what follows, we obtain primal-dual quadratic convergence under SMFCQ and SOSC, thus improving the result in [11] (in the case of optimization, as [11] deals with the more general variational problems). We note that this new result is complementary to those in [2, 15], in the sense that neither of the three results implies any of the others.

First note that the KKT system of subproblem (4.1) can be written in the following form:

$$\begin{aligned}
&\frac{\partial L}{\partial x}(x^k,\lambda^k,\mu^k)+\frac{\partial^2 L}{\partial x^2}(x^k,\lambda^k,\mu^k)(x-x^k)\\
&\quad+(F'(x^k))^{\mathrm{T}}(\lambda-\lambda^k)+(F''(x^k)[x-x^k])^{\mathrm{T}}(\lambda-\lambda^k)\\
&\quad+(G'(x^k))^{\mathrm{T}}(\mu-\mu^k)+(G''(x^k)[x-x^k])^{\mathrm{T}}(\mu-\mu^k)=0,\\
&F(x^k)+F'(x^k)(x-x^k)+\frac{1}{2}F''(x^k)[x-x^k,x-x^k]=0,\\
&\mu\geq 0,\qquad G(x^k)+G'(x^k)(x-x^k)+\frac{1}{2}G''(x^k)[x-x^k,x-x^k]\leq 0,\\
&\left\langle \mu, G(x^k)+G'(x^k)(x-x^k)+\frac{1}{2}G''(x^k)[x-x^k,x-x^k]\right\rangle=0.
\end{aligned}$$

By setting $\Omega(z,\zeta)=\{\omega(z,\zeta)\}$, with

$$\omega(z,\zeta)=\left((F''(x)[\xi])^{\mathrm{T}}u+(G''(x)[\xi])^{\mathrm{T}}v,\ \frac{1}{2}F''(x)[\xi,\xi],\ \frac{1}{2}G''(x)[\xi,\xi]\right),\tag{4.2}$$

where $z=(x,\lambda,\mu)\in\mathbf{R}^s$, $\zeta=(\xi,u,v)\in\mathbf{R}^s$, we observe that SQCQP method is a particular instance of iJNM (2.7) for GE (1.1) with $\Phi(\cdot)$ and $N(\cdot)$ defined according to (1.5) and (1.6), respectively.

Our proof of convergence is again via verification of the assumptions of Theorem 2.1 for iJNM.

Corollary 4.1 *Let $f:\mathbf{R}^n\to\mathbf{R}$, $F:\mathbf{R}^n\to\mathbf{R}^l$ and $G:\mathbf{R}^n\to\mathbf{R}^m$ be twice differentiable near a point $\bar{x}\in\mathbf{R}^n$, and suppose that the second derivatives of f, F and G are continuous at $\bar{x}$. Assume that $\bar{x}$ is a local solution of MP problem (1.3), satisfying SMFCQ and SOSC (1.10) for the associated Lagrange multiplier $(\bar{\lambda},\bar{\mu})\in\mathbf{R}^l\times\mathbf{R}^m$.*

Then there exists $\delta>0$ such that for any starting point $(x^0,\lambda^0,\mu^0)\in\mathbf{R}^n\times\mathbf{R}^l\times\mathbf{R}^m$ close enough to $(\bar{x},\bar{\lambda},\bar{\mu})$, there exists a trajectory $\{(x^k,\lambda^k,\mu^k)\}\subset\mathbf{R}^n\times\mathbf{R}^l\times\mathbf{R}^m$ such that x^{k+1} is a stationary point of problem (4.1) for each $k=0,1,\ldots,$ and $(\lambda^{k+1},\mu^{k+1})$ is an associated Lagrange multiplier, satisfying (3.4); any such trajectory converges to $(\bar{x},\bar{\lambda},\bar{\mu})$, and the rate of convergence is superlinear. Moreover, the rate of convergence is quadratic provided the second derivatives of f, F and G are locally Lipschitz-continuous with respect to $\bar{x}$.

Proof For each fixed $x \in \mathbf{R}^n$ regarded as a parameter, consider the MP problem

$$\begin{aligned} \text{minimize}_\xi \quad & \langle f'(x), \xi\rangle + \frac{1}{2} f''(x)[\xi, \xi] \\ \text{subject to} \quad & F(x) + F'(x)\xi + \frac{1}{2} F''(x)[\xi, \xi] = 0, \\ & G(x) + G'(x)\xi + \frac{1}{2} G''(x)[\xi, \xi] \le 0. \end{aligned} \tag{4.3}$$

For the base parameter value $x = \bar{x}$, the point $\bar{\xi} = 0$ is stationary for this problem, with the associated Lagrange multiplier $(\bar{\lambda}, \bar{\mu})$. It can be easily checked that SMFCQ and SOSC (1.10) hold for this problem, at this point, for this multiplier. Repeating the argument from the proof of Corollary 3.1, we obtain that for each $x \in \mathbf{R}^n$ close enough to $\bar{x}$, there exists $(\tilde{\lambda}(x), \tilde{\mu}(x)) \in \mathbf{R}^l \times \mathbf{R}^m$ such that the triple $(\xi(x), \tilde{\lambda}(x), \tilde{\mu}(x))$ satisfies the KKT system

$$\begin{aligned} & f'(x) + f''(x)\xi + (F'(x))^{\mathrm{T}}\tilde{\lambda} + (F''(x)[\xi])^{\mathrm{T}}\tilde{\lambda} + (G'(x))^{\mathrm{T}}\tilde{\mu} + (G''(x)[\xi])^{\mathrm{T}}\tilde{\mu} = 0, \\ & F(x) + F'(x)\xi + \frac{1}{2} F''(x)[\xi, \xi] = 0, \\ & \tilde{\mu} \ge 0, \qquad G(x) + G'(x)\xi + \frac{1}{2} G''(x)[\xi, \xi] \le 0, \\ & \left\langle \tilde{\mu}, G(x) + G'(x)\xi + \frac{1}{2} G''(x)[\xi, \xi] \right\rangle = 0 \end{aligned}$$

of problem (4.3), and

$$\|\xi(x)\| + \|\tilde{\lambda}(x) - \bar{\lambda}\| + \|\tilde{\mu}(x) - \bar{\mu}\| = O(\|x - \bar{x}\|). \tag{4.4}$$

In particular, $(\xi(x), \tilde{\lambda}(x), \tilde{\mu}(x)) \to (0, \bar{\lambda}, \bar{\mu})$ as $x \to \bar{x}$. Fixing an arbitrary pair $(\lambda, \mu) \in \mathbf{R}^l \times \mathbf{R}^m$, and denoting $(u, v) = (\tilde{\lambda} - \lambda, \tilde{\mu} - \mu)$, we can re-write the latter KKT system as follows:

$$\begin{aligned} & \frac{\partial L}{\partial x}(x, \lambda, \mu) + \frac{\partial^2 L}{\partial x^2}(x, \lambda, \mu)\xi + (F'(x))^{\mathrm{T}}u + (F''(x)[\xi])^{\mathrm{T}}u \\ & \quad + (G'(x))^{\mathrm{T}}v + (G''(x)[\xi])^{\mathrm{T}}v = 0, \\ & F(x) + F'(x)\xi + \frac{1}{2} F''(x)[\xi, \xi] = 0, \\ & \mu + v \ge 0, \qquad G(x) + G'(x)\xi + \frac{1}{2} G''(x)[\xi, \xi] \le 0, \\ & \left\langle \mu + v, G(x) + G'(x)\xi + \frac{1}{2} G''(x)[\xi, \xi] \right\rangle = 0. \end{aligned}$$

The latter system is precisely (2.8) for the SQCQP method, with $z = (x, \lambda, \mu)$ and $\zeta = (\xi, u, v)$. Using the notation $(u(x, \lambda), v(x, \mu)) = (\tilde{\lambda}(x) - \lambda, \tilde{\mu}(x) - \mu)$, and setting $\zeta(z) = (\xi(x), u(x, \lambda), v(x, \mu))$, we finally obtain from (4.4) that

$$\begin{aligned}\|\zeta(z)\| &\leq \|\xi(x)\| + \|\tilde{\lambda}(x) - \bar{\lambda}\| + \|\tilde{\mu}(x) - \bar{\mu}\| + \|\lambda - \bar{\lambda}\| + \|\mu - \bar{\mu}\| \\ &= O(\|x - \bar{x}\|) + \|\lambda - \bar{\lambda}\| + \|\mu - \bar{\mu}\| \\ &= O(\|z - \bar{z}\|), \end{aligned} \tag{4.5}$$

and in particular $\zeta(z) \to 0$ as $z \to \bar{z}$, with $\bar{z} = (\bar{x}, \bar{\lambda}, \bar{\mu})$. This verifies Assumption (iJNM1).

Furthermore, according to the argument above, estimate (4.5) is valid for $z = (x, \lambda, \mu) \in \mathbf{R}^s$ and for any solution $\zeta(z) = (\xi(x), u(x, \lambda), v(x, \mu))$ of GE (2.10), such that $\zeta(z)$ is close enough to zero. Employing (4.2), we now obtain the estimate

$$\begin{aligned}\|\omega(z, \zeta(z))\| &= O(\|\zeta(z)\|^2) \\ &= O(\|z - \bar{z}\|^2).\end{aligned}$$

Hence, Assumption (iJNM2), with (2.9) replaced by (2.12), is verified.

The assertions now follow applying Theorem 2.1. □

5 Stabilized SQP

The stabilized SQP (sSQP) method is the following iterative procedure. For the current primal-dual iterate $(x^k, \lambda^k, \mu^k) \in \mathbf{R}^n \times \mathbf{R}^l \times \mathbf{R}^m$, the next iterate $(x^{k+1}, \lambda^{k+1}, \mu^{k+1})$ is computed as a stationary point of the QP subproblem

$$\begin{aligned} \text{minimize}_{(x,\lambda,\mu)} \quad & \langle f'(x^k), x - x^k\rangle + \frac{1}{2}\frac{\partial^2 L}{\partial x^2}(x^k, \lambda^k, \mu^k)[x - x^k, x - x^k] \\ & + \frac{\sigma_k}{2}(\|\lambda\|^2 + \|\mu\|^2) \\ \text{subject to} \quad & F(x^k) + F'(x^k)(x - x^k) - \sigma_k(\lambda - \lambda^k) = 0, \\ & G(x^k) + G'(x^k)(x - x^k) - \sigma_k(\mu - \mu^k) \leq 0, \end{aligned} \tag{5.1}$$

with $\sigma_k > 0$ (to be specific, we shall take $\sigma_k = \sigma(x^k, \lambda^k, \mu^k)$, where $\sigma(\cdot)$ is the natural residual defined in (1.12)).

The sSQP method was proposed in [33] (for inequality-constrained problems) and further studied in [12, 13, 16, 22, 34, 35]. In [34, 35], superlinear convergence of sSQP was established under MFCQ and the strong SOSC assumed for all multipliers. Superlinear convergence had also been shown under the sole assumption of strong SOSC for some multiplier $\bar{\mu}$, provided that μ^0 is close enough to such $\bar{\mu}$ [16]; see also [13]. In what follows, we shall establish superlinear convergence under SMFCQ and SOSC, which gives a result neither stronger nor weaker than the above. It should be noted that, in principle, sSQP had been introduced for the purposes of dual stabilization, when multiplier associated to a solution is not unique. In this sense, SMFCQ is not the most natural assumption in this context. Nevertheless, whether the multiplier is unique or not is often unknown (a priori, at the time of applying the algorithm).

For this reason, one would still like to know how the algorithm behaves in different situations. Finally, it should also be noted that recently convergence of sSQP had been shown under SOSC only [12], which gives the strongest result so far. Even though the result of this section is weaker than this most recent one, we chose to include a brief exposition as an application of iJNM framework. It is also worthwhile to point out the following difference between "localization" conditions in our Corollary 5.1 below and previous convergence results for sSQP (recall also related comments concerning general iJNM after the proof of Theorem 2.1). In [12, 13], superlinear convergence is affirmed for primal-dual iterates satisfying

$$\|(x^{k+1}-x^k, \lambda^{k+1}-\lambda^k, \mu^{k+1}-\mu^k)\| \le \Delta \|(x^k-\bar{x}, \lambda^k-\bar{\lambda}, \mu^k-\bar{\mu})\|$$

for some $\Delta > 0$, while the condition in Corollary 5.1 is (3.4) for some $\delta > 0$. The latter appears somewhat more "practical", as it does not involve the unknown solution.

Note first that the KKT system of subproblem (5.1) can be written in the following form:

$$\begin{aligned}
&\frac{\partial L}{\partial x}(x^k, \lambda^k, \mu^k) + \frac{\partial^2 L}{\partial x^2}(x^k, \lambda^k, \mu^k)(x-x^k) + (F'(x^k))^{\mathrm{T}}(\lambda-\lambda^k) \\
&\quad + (G'(x^k))^{\mathrm{T}}(\mu-\mu^k) = 0, \\
&F(x^k) + F'(x^k)(x-x^k) - \sigma_k(\lambda-\lambda^k) = 0, \\
&\mu \ge 0, \qquad G(x^k) + G'(x^k)(x-x^k) - \sigma_k(\mu-\mu^k) \le 0, \\
&\langle \mu, G(x^k) + G'(x^k)(x-x^k) - \sigma_k(\mu-\mu^k)\rangle = 0.
\end{aligned}$$

By setting $\Omega(z, \zeta) = \{\omega(z, \zeta)\}$ with

$$\omega(z, \zeta) = (0, -\sigma(x, \lambda, \mu)u, -\sigma(x, \lambda, \mu)v), \tag{5.2}$$

where $z = (x, \lambda, \mu) \in \mathbf{R}^s$, $\zeta = (\xi, u, v) \in \mathbf{R}^s$, we observe that sSQP is a particular instance of iJNM (2.7) for GE (1.1) with $\Phi(\cdot)$ and $N(\cdot)$ defined according to (1.5) and (1.6), respectively.

Corollary 5.1 *Under the assumptions of Corollary 4.1, there exists $\delta > 0$ such that for any starting point $(x^0, \lambda^0, \mu^0) \in \mathbf{R}^n \times \mathbf{R}^l \times \mathbf{R}^m$ close enough to $(\bar{x}, \bar{\lambda}, \bar{\mu})$, there exists a trajectory $\{(x^k, \lambda^k, \mu^k)\} \subset \mathbf{R}^n \times \mathbf{R}^l \times \mathbf{R}^m$ such that $(x^{k+1}, \lambda^{k+1}, \mu^{k+1})$ is a stationary point of QP problem (5.1) for each $k = 0, 1, \ldots,$ with $\sigma_k = \sigma(x^k, \lambda^k, \mu^k)$ defined according to (1.12), satisfying (3.4); any such trajectory converges to $(\bar{x}, \bar{\lambda}, \bar{\mu})$, and the rate of convergence is superlinear. Moreover, the rate of convergence is quadratic provided the second derivatives of f, F and G are locally Lipschitz-continuous with respect to $\bar{x}$.*

Proof Under SOSC (1.10), Assumption (iJNM1) with $\bar{z} = (\bar{x}, \bar{\lambda}, \bar{\mu})$ for sSQP follows from [12, Theorem 3].

Regarding Assumption (iJNM2), by (1.13) and (5.2) we obtain that the estimate

$$\begin{aligned}
\|\omega(z, \zeta)\| &= \sigma(x, \lambda, \mu)\|(u, v)\| \\
&= O(\|\zeta\|\|z-\bar{z}\|)
\end{aligned}$$

holds for $z = (x, \lambda, \mu) \in \mathbf{R}^s$ and $\zeta = (\xi, u, v) \in \mathbf{R}^s$. This estimate implies Assumption (iJNM2) with (2.9) replaced by (2.12).

The assertion now follows from Theorem 2.1. □

References

1. An, L.T.H.: An efficient algorithm for globally minimizing a quadratic function under convex quadratic constraints. Math. Program. **87**, 401–426 (2000)
2. Anitescu, M.: A superlinearly convergent sequential quadratically constrained quadratic programming algorithm for degenerate nonlinear programming. SIAM J. Optim. **12**, 949–978 (2002)
3. Arutyunov, A.V., Izmailov, A.F.: Sensitivity analysis for cone-constrained optimization problems under the relaxed constraint qualifications. Math. Oper. Res. **30**, 333–353 (2005)
4. Audet, C., Hansen, P., Jaumard, B., Savard, G.: A branch and cut algorithm for nonconvex quadratically constrained quadratic programming. Math. Program. **87**, 131–152 (2000)
5. Bertsekas, D.P.: Constrained Optimization and Lagrange Multiplier Methods. Academic Press, New York (1982)
6. Boggs, B.T., Tolle, J.W.: Sequential quadratic programming. Acta Numer. **4**, 1–51 (1996)
7. Bonnans, J.F.: Local analysis of Newton-type methods for variational inequalities and nonlinear programming. Appl. Math. Optim. **29**, 161–186 (1994)
8. Bonnans, J.F., Shapiro, A.: Perturbation Analysis of Optimization Problems. Springer, New York (2000)
9. Bonnans, J.F., Gilbert, J.Ch., Lemaréchal, C., Sagastizábal, C.: Numerical Optimization: Theoretical and Practical Aspects, 2nd edn. Springer, Berlin (2006)
10. Dennis, J.E., Moré, J.J.: A characterization of superlinear convergence and its application to quasi-Newton methods. Math. Comput. **28**, 549–560 (1974)
11. Fernández, D., Solodov, M.: On local convergence of sequential quadratically-constrained quadratic-programming type methods, with an extension to variational problems. Comput. Optim. Appl. **39**, 143–160 (2008)
12. Fernández, D., Solodov, M.: Stabilized sequential quadratic programming for optimization and a stabilized Newton-type method for variational problems. Math. Program. (2009). DOI 10.1007/s10107-008-0255-4
13. Fischer, A.: Local behavior of an iterative framework for generalized equations with nonisolated solutions. Math. Program. **94**, 91–124 (2002)
14. Friedlander, M.P., Saunders, M.A.: A globally convergent linearly constrained Lagrangian method for nonlinear optimization. SIAM J. Optim. **15**, 863–897 (2005)
15. Fukushima, M., Luo, Z.-Q., Tseng, P.: A sequential quadratically constrained quadratic programming method for differentiable convex minimization. SIAM J. Optim. **13**, 1098–1119 (2003)
16. Hager, W.W.: Stabilized sequential quadratic programming. Comput. Optim. Appl. **12**, 253–273 (1999)
17. Huang, Z.-H., Sun, D., Zhao, G.: A smoothing Newton-type algorithm of stronger convergence for the quadratically constrained convex quadratic programming. Comput. Optim. Appl. **35**, 199–237 (2006)
18. Josephy, N.H.: Newton's method for generalized equations. Technical Summary Report no. 1965, Mathematics Research Center, University of Wisconsin, Madison (1979)
19. Josephy, N.H.: Quasi-Newton methods for generalized equations. Technical Summary Report no. 1966, Mathematics Research Center, University of Wisconsin, Madison (1979)
20. Kantorovich, L.V., Akilov, G.P.: Functional Analysis, 2nd edn. Pergamon, Oxford (1982)
21. Kruk, S., Wolkowicz, H.: Sequential, quadratically constrained, quadratic programming for general nonlinear programming. In: Wolkowicz, H., Saigal, R., Vandenberghe, L. (eds.) Handbook of Semidefinite Programming, pp. 563–575. Kluwer Academic, Dordrecht (2000)
22. Li, D.-H., Qi, L.: Stabilized SQP method via linear equations. Appl. Math. Techn. Rept. AMR00/5, Univ. New South Wales, Sydney (2000)
23. Lobo, M.S., Vandenberghe, L., Boyd, S., Lebret, H.: Applications of second-order cone programming. Linear Algebra Appl. **284**, 193–228 (1998)
24. Murtagh, B.A., Saunders, M.A.: A projected Lagrangian algorithm and its implementation for sparse nonlinear constraints. Math. Program. Study **16**, 84–117 (1982)

25. Nesterov, Y.E., Nemirovskii, A.S.: Interior Point Polynomial Methods in Convex Programming: Theory and Applications. SIAM, Philadelphia (1993)
26. Panin, V.M.: A second-order method for discrete min-max problem. USSR Comput. Math. Math. Phys. **19**, 90–100 (1979)
27. Robinson, S.M.: A quadratically convergent algorithm for general nonlinear programming problems. Math. Program. **3**, 145–156 (1972)
28. Robinson, S.M.: Perturbed Kuhn–Tucker points and rates of convergence for a class of nonlinear-programming algorithms. Math. Program. **7**, 1–16 (1974)
29. Robinson, S.M.: Stability theorems for systems of inequalities, Part II: Differentiable nonlinear systems. SIAM J. Numer. Anal. **13**, 497–513 (1976)
30. Robinson, S.M.: Strongly regular generalized equations. Math. Oper. Res. **5**, 43–62 (1980)
31. Solodov, M.V.: On the sequential quadratically constrained quadratic programming methods. Math. Oper. Res. **29**, 64–79 (2004)
32. Wiest, E.J., Polak, E.: A generalized quadratic programming-based phase-I–phase-II method for inequality-constrained optimization. Appl. Math. Optim. **26**, 223–252 (1992)
33. Wright, S.J.: Superlinear convergence of a stabilized SQP method to a degenerate solution. Comput. Optim. Appl. **11**, 253–275 (1998)
34. Wright, S.J.: Modifying SQP for degenerate problems. SIAM J. Optim. **13**, 470–497 (2002)
35. Wright, S.J.: Constraint identification and algorithm stabilization for degenerate nonlinear programs. Math. Program. **95**, 137–160 (2003)