Comput Optim Appl
DOI 10.1007/s10589-007-9074-4

Examples of dual behaviour of Newton-type methods on optimization problems with degenerate constraints

A.F. Izmailov · M.V. Solodov

Received: 5 January 2007 / Revised: 30 May 2007
© Springer Science+Business Media, LLC 2007

Abstract We discuss possible scenarios of behaviour of the dual part of sequences generated by primal-dual Newton-type methods when applied to optimization problems with nonunique multipliers associated to a solution. Those scenarios are: (a) failure of convergence of the dual sequence; (b) convergence to a so-called critical multiplier (which, in particular, violates some second-order sufficient conditions for optimality), the latter appearing to be a typical scenario when critical multipliers exist; (c) convergence to a noncritical multiplier. The case of mathematical programs with complementarity constraints is also discussed. We illustrate those scenarios with examples, and discuss consequences for the speed of convergence. We also put together a collection of examples of optimization problems with constraints violating some standard constraint qualifications, intended for preliminary testing of existing algorithms on degenerate problems, or for developing special new algorithms designed to deal with constraints degeneracy.

Keywords Degenerate constraints · Second-order sufficiency · Newton method · SQP

Research of the first author is supported by the Russian Foundation for Basic Research Grants 07-01-00270, 07-01-00416 and 07-01-90102-Mong, and by RF President's Grant NS-9344.2006.1 for the support of leading scientific schools. The second author is supported in part by CNPq Grants 301508/2005-4, 490200/2005-2 and 550317/2005-8, by PRONEX–Optimization, and by FAPERJ Grant E-26/151.942/2004.

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

M.V. Solodov (✉)
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

We consider the problem

$$\begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & F(x) = 0, \qquad G(x) \leq 0, \end{array} \tag{1.1}$$

where $f : \mathbf{R}^n \to \mathbf{R}$ is a smooth function, $F : \mathbf{R}^n \to \mathbf{R}^l$ and $G : \mathbf{R}^n \to \mathbf{R}^m$ are smooth mappings. Stationary points of problem (1.1) and the associated Lagrange multipliers are characterized by the Karush–Kuhn–Tucker (KKT) optimality system

$$\begin{aligned} &\frac{\partial L}{\partial x}(x, \lambda, \mu) = 0, \\ &F(x) = 0, \\ &G(x) \leq 0, \qquad \mu \geq 0, \qquad \langle \mu, G(x) \rangle = 0, \end{aligned} \tag{1.2}$$

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 problem (1.1).

Let $\bar{x} \in \mathbf{R}^n$ be a stationary point of (1.1), where we assume that the set $\mathcal{M}(\bar{x})$ of Lagrange multipliers associated with $\bar{x}$ is nonempty. As is well known, if $\bar{x}$ is a local solution of problem (1.1), then $\mathcal{M}(\bar{x})$ is nonempty if the Mangasarian–Fromovitz constraint qualification (MFCQ) holds at $\bar{x}$, that is,

$$\operatorname{rank} F'(\bar{x}) = l \quad \text{and} \quad \exists \bar{\xi} \in \ker F'(\bar{x}) \quad \text{such that} \quad G'_{I(\bar{x})}(\bar{x})\bar{\xi} < 0,$$

where $I(\bar{x}) = \{i = 1, \ldots, m \mid G_i(\bar{x}) = 0\}$ is the set of indices of inequality constraints active at $\bar{x}$.

We are interested in behaviour of the dual part of sequences generated by primal-dual Newton methods applied to solve (1.1). The case of interest in this work is when $\mathcal{M}(\bar{x})$ is not a singleton. When the multiplier is unique, the dual sequence has only one possible limit (if the method converges in any reasonable sense), which makes such a case trivial from the point of view of dual behaviour. At issue, therefore, are *degenerate* problems, where $\bar{x}$ does not satisfy the so-called strict MFCQ (that is, MFCQ combined with the requirement that the multiplier be unique). Hence, $\bar{x}$ also does not satisfy the stronger linear independence constraint qualification (LICQ), which can be expressed in the form

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

where $|I|$ stands for the cardinality of a finite set I.

The case of violation of classical constraint qualifications has been a subject of considerable interest in the past decade, both in the general case (e.g., [2, 6, 11, 13, 14, 20, 21, 24, 25, 39]) and in the special case of equilibrium or complementarity constraints (e.g., [3, 4, 12, 22, 29, 33–35]).

In addition to constraint qualifications (or lack of them), second-order conditions are important for convergence and rate of convergence of Newton-type methods. Recall that the second-order sufficient condition for optimality (SOSC) holds at $\bar{x}$ with a multiplier $(\bar{\lambda}, \bar{\mu}) \in \mathcal{M}(\bar{x})$, if

$$\left\langle \frac{\partial^2 L}{\partial x^2}(\bar{x}, \bar{\lambda}, \bar{\mu})\xi, \xi \right\rangle > 0 \quad \forall \xi \in C(\bar{x}) \setminus \{0\}, \tag{1.3}$$

where

$$C(\bar{x}) = \{\xi \in \ker F'(\bar{x}) \mid G'_{I(\bar{x})}(\bar{x})\xi \leq 0, \langle f'(\bar{x}), \xi \rangle \leq 0\}$$

is the critical cone of (1.1) at $\bar{x}$. If SOSC holds (with some multipliers) then the point $\bar{x}$ is a strict local minimizer of (1.1) and primal convergence to $\bar{x}$ can be expected for (good implementations of) good algorithms, even in degenerate cases. The speed of convergence, however, is often slow. It has been observed that, when primal convergence is slow, the reason for this is not so much degeneracy as such, but some undesirable behaviour of the dual sequence. Among various scenarios of this behaviour, one of the prominent ones is dual convergence to multipliers violating SOSC (see, e.g., [38, Sect. 6], [25]). Understanding this phenomenon better is one of the purposes for putting together a collection of examples presented in this paper. The other reason is assessing the chances of applicability of various local stabilization/regularization methods that have been proposed to tackle degeneracy [11, 13, 24, 28, 39] (see also [10, 36–38]). Some of those methods do achieve superlinear or quadratic convergence despite degeneracy, if their primal-dual starting point is close to a point satisfying SOSC. To get such a starting point, the issue of dual behaviour of an "outer" (global) phase of the algorithm is again important.

The purpose of this paper, therefore, is to discuss possible scenarios of dual behaviour of Newton methods applied to degenerate optimization problems, and to put together a representative library of small examples that illustrate the possibilities and that can be used for future development of algorithms for degenerate problems. For the case of equality constraints only (i.e., when there are no inequality constraints in (1.1)), we believe that the overall picture is quite clear [25]. The case of mixed constraints, including complementarity constraints, is more complex and requires further study. Examples investigated in the present paper is a step in this direction.

In what follows, the main attention will be paid to the standard linesearch sequential quadratic programming (SQP) algorithm, generating primal-dual trajectories $\{(x^k, \lambda^k, \mu^k)\}$ as follows: $x^{k+1} = x^k + \alpha_k \xi^k$, where $\alpha_k \geq 0$ is the stepsize parameter, ξ^k is a stationary point of the SQP subproblem

$$\begin{aligned} &\text{minimize} \quad \langle f'(x^k), \xi \rangle + \frac{1}{2}\left\langle \frac{\partial^2 L}{\partial x^2}(x^k, \lambda^k, \mu^k)\xi, \xi \right\rangle \\ &\text{subject to} \quad F(x^k) + F'(x^k)\xi = 0, \\ &\qquad\qquad\quad G(x^k) + G'(x^k)\xi \leq 0, \end{aligned} \tag{1.4}$$

and $(\lambda^{k+1}, \mu^{k+1})$ is a Lagrange multiplier associated to ξ^k. The rule we use for choosing α_k in our experiments is as follows: the initial trial value $\alpha_k = 1$ is halved until

$$\varphi_k(x^k + \alpha_k \xi^k) \leq \varphi_k(x^k) + \varepsilon \alpha_k \Delta_k,$$

where $\varepsilon \in (0, 1)$,

$$\varphi_k : \mathbf{R}^n \to \mathbf{R}, \quad \varphi_k(x) = f(x) + c_k \left(\|F(x^k)\|_1 + \sum_{i=1}^m \max\{0, G_i(x^k)\} \right),$$

is the l_1-penalty function with the penalty parameter $c_k > 0$, and

$$\Delta_k = \langle f'(x^k), \xi^k \rangle - c_k \left(\|F(x^k)\|_1 + \sum_{i=1}^m \max\{0, G_i(x^k)\} \right)$$

is an estimate of the directional derivative of φ_k at x^k in the SQP direction ξ^k. In our numerical experiments below we use $\varepsilon = 0.1$ and the following simple update rule for penalty parameters: $c_0 = \|(\lambda^1, \mu^1)\|_\infty + 1$, and then for each $k = 1, 2, \ldots,$ we set $c_k = c_{k-1}$ if $c_{k-1} \geq \|(\lambda^{k+1}, \mu^{k+1})\|_\infty$, and $c_k = \|(\lambda^{k+1}, \mu^{k+1})\|_\infty + 1$ otherwise. The code is written in Matlab, and SQP subproblems (1.4) are solved by the built-in Matlab solver `quadprog`.

Along with SQP, some other Newton-type methods for problem (1.1) (or, more precisely, for KKT system (1.2)) will also be discussed, albeit briefly. Specifically, we shall comment on the so-called semismooth Newton (SNM) method for reformulations of the system (1.2), based on the natural residual complementarity function $\psi(a, b) = \min\{a, b\}$ and on the Fischer–Burmeister complementarity function $\psi(a, b) = \sqrt{a^2 + b^2} - a - b$, $a, b \in \mathbf{R}$. The corresponding versions of SNM will be abbreviated as SNM-NR and SNM-FB, respectively. With these functions, (1.2) can be re-written as the nonsmooth system of equations

$$\Phi(x, \lambda, \mu) = 0,$$

where

$$\Phi : \mathbf{R}^n \times \mathbf{R}^l \times \mathbf{R}^m \to \mathbf{R}^n \times \mathbf{R}^l \times \mathbf{R}^m,$$
$$\Phi(x, \lambda, \mu) = \left(\frac{\partial L}{\partial x}(x, \lambda, \mu), F(x), \psi(\mu, -G(x)) \right),$$

with ψ applied component-wise. SNM-NR and SNM-FB generate primal-dual trajectories $\{(x^k, \lambda^k, \mu^k)\}$ as follows: $(x^{k+1}, \lambda^{k+1}, \mu^{k+1}) = (x^k, \lambda^k, \mu^k) + \alpha_k d^k$, where $\alpha_k \geq 0$ is again the stepsize parameter, and d^k solves the linear system

$$\Phi(x^k, \lambda^k, \mu^k) + \Lambda_k d = 0, \tag{1.5}$$

where $\Lambda_k \in \partial_B \Phi(x^k, \lambda^k, \mu^k)$, and $\partial_B \Phi(x^k, \lambda^k, \mu^k)$ stands for the so-called B-subdifferential of Φ at (x^k, λ^k, μ^k) (see [30, 31]).

Even though SNM methods applied to reformulations of the KKT system are not among common tools to solve optimization problems, we feel that paying some attention to them is justified here, having in mind the issues in consideration. For example, if it were to be discovered that SNM methods exhibit better dual behaviour on some classes of degenerate problems than SQP, this could be useful for obtaining better multiplier estimates. However, in our experience, dual behaviour of SNM is usually quite similar to SQP. For this reason, we shall be commenting on SNM only briefly.

The linesearch procedure we used for SNM-FB is essentially the same as in "General Line Search Algorithm" stated in [8]. Specifically, fix $q \in (0, 1)$, $\varepsilon \in (0, 1)$. For each k, let d^k be defined according to (1.5). We accept $\alpha_k = 1$ if

$$\varphi((x^k, \lambda^k, \mu^k) + \alpha_k d^k) \le q\varphi(x^k, \lambda^k, \mu^k),$$

where

$$\varphi : \mathbf{R}^n \to \mathbf{R}, \quad \varphi(x, \lambda, \mu) = \frac{1}{2}\|\Phi(x, \lambda, \mu)\|^2.$$

Otherwise, if d^k satisfies the descent direction condition

$$\langle \varphi'(x^k, \lambda^k, \mu^k), d^k\rangle \le -\gamma \|d^k\|^\delta,$$

the initial trial value $\alpha_k = 1$ is halved until

$$\varphi((x^k, \lambda^k, \mu^k) + \alpha_k d^k) \le \varphi_k(x^k) + \varepsilon\alpha_k\langle \varphi'(x^k, \lambda^k, \mu^k), d^k\rangle.$$

If d^k does not satisfy the descent direction condition, we re-define $d^k = -\varphi'(x^k, \lambda^k, \mu^k)$ and decrease the initial trial value $\alpha_k = 1$ according to the same rule. In our experiments, we used the following values of parameters: $q = 0.9$, $\delta = 2.1$, $\gamma = 10^{-9}$, $\varepsilon = 10^{-4}$.

For SNM-NR, we use $\alpha_k = 1$ for all k, because this method does not admit a natural and fully satisfactory linesearch procedure.

The stopping criterion for all runs was $\|x^k - \bar{x}\| < 10^{-7}$. Since we deal with degenerate problems, it is natural that many runs of our simple implementations end up in a failure (because of inconsistent or unbounded SQP subproblems (1.4), or inconsistent SNM linear systems (1.5), because of extremely slow convergence, or simply because of the lack of globalization device in the case of SNM-NR). We are not concerned with robustness, but rather with convergence properties, when there is convergence.

2 Critical multipliers

We next recall the notion of *critical multipliers*, which plays a central role in understanding dual behaviour of Newton-type schemes on degenerate problems, see [25]. It is convenient to do this for the case of equality-constrained problems first. For the moment, let the problem be

$$\begin{aligned} &\text{minimize} \quad f(x) \\ &\text{subject to} \quad F(x) = 0 \end{aligned} \tag{2.1}$$

(i.e., in the setting of the definition that follows, there are no G-constraints in (1.1) and no μ-component of a multiplier). Note that for this problem, MFCQ, strict MFCQ and LICQ at $\bar{x}$, all reduce to the following classical regularity condition:

$$\operatorname{rank} F'(\bar{x}) = l. \tag{2.2}$$

Furthermore, in this case the KKT system (1.2) reduces to the Lagrange optimality system

$$\begin{aligned} &\frac{\partial L}{\partial x}(x, \lambda) = 0,\\ &F(x) = 0. \end{aligned} \tag{2.3}$$

Moreover, SQP, SNM-NR and SNM-FB, all reduce to the (damped) Newton–Lagrange method, i.e., the standard Newton method for the system of equations (2.3), with the only possible differences in choosing the stepsize parameters (in the examples presented below, we employ the Newton–Lagrange method with the same line-search procedure as for SQP). Specifically, the system for defining stationary points and multipliers of SQP in (1.4), and the equation of SNM in (1.5), both reduce to the system

$$\begin{aligned} &\frac{\partial L}{\partial x}(x^k, \lambda^k) + \frac{\partial^2 L}{\partial x^2}(x^k, \lambda^k)\xi + (F'(x^k))^{\mathrm{T}}\eta = 0,\\ &F(x^k) + F'(x^k)\xi = 0, \end{aligned}$$

with respect to $(\xi, \eta) \in \mathbf{R}^n \times \mathbf{R}^l$.

The following definition was introduced in [19]; see [25] for detailed treatment.

Definition 2.1 A multiplier $\bar{\lambda} \in \mathcal{M}(\bar{x})$ associated with a stationary point $\bar{x}$ of problem (2.1) is referred to as *critical* if

$$\exists \xi \in \ker F'(\bar{x}) \setminus \{0\} \quad \text{such that} \quad \frac{\partial^2 L}{\partial x^2}(\bar{x}, \bar{\lambda})\xi \in \operatorname{im}(F'(\bar{x}))^{\mathrm{T}},$$

and *noncritical* otherwise.

For problem (2.1), SOSC (1.3) takes the form

$$\left\langle \frac{\partial^2 L}{\partial x^2}(\bar{x}, \bar{\lambda})\xi, \xi \right\rangle > 0 \quad \forall \xi \in \ker F'(\bar{x}) \setminus \{0\}. \tag{2.4}$$

Since $\operatorname{im}(F'(\bar{x}))^{\mathrm{T}} = (\ker F'(\bar{x}))^{\perp}$, it is immediate that criticality implies that

$$\exists \xi \in \ker F'(\bar{x}) \setminus \{0\} \quad \text{such that} \quad \left\langle \frac{\partial^2 L}{\partial x^2}(\bar{x}, \bar{\lambda})\xi, \xi \right\rangle = 0,$$

and in particular, SOSC (2.4) is violated for any critical multiplier $\bar{\lambda}$. If $F'(\bar{x}) = 0$, criticality of $\bar{\lambda}$ reduces to saying that the matrix $\frac{\partial^2 L}{\partial x^2}(\bar{x}, \bar{\lambda})$ is singular. In the general case, criticality can be interpreted as singularity of the so-called reduced Hessian of the Lagrangian; see [25].

Evidently, critical multipliers form a special subclass within the multipliers violating SOSC. As shown in [25], critical multipliers serve as attractors for the dual sequence of Newton-type methods: convergence to such multipliers is typical in a certain sense. In particular, it is something that should be expected to happen when the dual sequence converges and critical multipliers exist. This is quite remarkable, considering that the set of critical multipliers is normally "thin" in $\mathcal{M}(\bar{x})$ (as usual, we are talking about situations where the regularity condition (2.2) does not hold but $\mathcal{M}(\bar{x})$ is nonempty). Moreover, the reason for slow primal convergence is either non-convergence of the dual sequence, or its convergence to critical multipliers, as shown in [25]. If dual sequence were to converge to a noncritical multiplier, primal convergence rate would have been superlinear.

We illustrate the effect of attraction to critical multipliers for equality-constrained problems by the following examples. In those examples, we deal with the pure Newton–Lagrange method (i.e., $\alpha_k = 1$).

Example 2.1 ([6]) The equality-constrained problem

$$\begin{array}{ll} \text{minimize} & -x_1^2 - x_2^2 \\ \text{subject to} & x_1^2 - x_2^2 = 0, \quad x_1 x_2 = 0, \end{array}$$

has the unique feasible point (hence, unique solution) $\bar{x} = 0$, with $\mathcal{M}(\bar{x}) = \mathbf{R}^2$. This solution violates both, the regularity condition (2.2) (since $F'(\bar{x}) = 0$), and SOSC (2.4) for any $\bar{\lambda} \in \mathbf{R}^2$. Critical multipliers are those $\bar{\lambda}$ that satisfy $4\bar{\lambda}_1^2 + \bar{\lambda}_2^2 = 4$.

In Fig. 1, critical multipliers are represented by the thick line (in the form of an oval). Some Newton–Lagrange dual trajectories from primal starting points $x^0 = (1, 2)$ and $x^0 = (2, 1)$ are represented by thin lines, with dots for the dual iterates. It can be shown analytically that for this problem, the dual limit point depends exclusively on x^0 and not on λ^0. Specifically, for any k the primal-dual Newton–Lagrange step is given by

$$x^{k+1} = \frac{1}{2} x^k, \qquad \lambda^{k+1} = \frac{1}{2}\lambda^k + \left(-\frac{(x_2^0)^2 - (x_1^0)^2}{2((x_1^0)^2 + (x_2^0)^2)}, \frac{2x_1^0 x_2^0}{(x_1^0)^2 + (x_2^0)^2} \right),$$

and the dual trajectory converges to

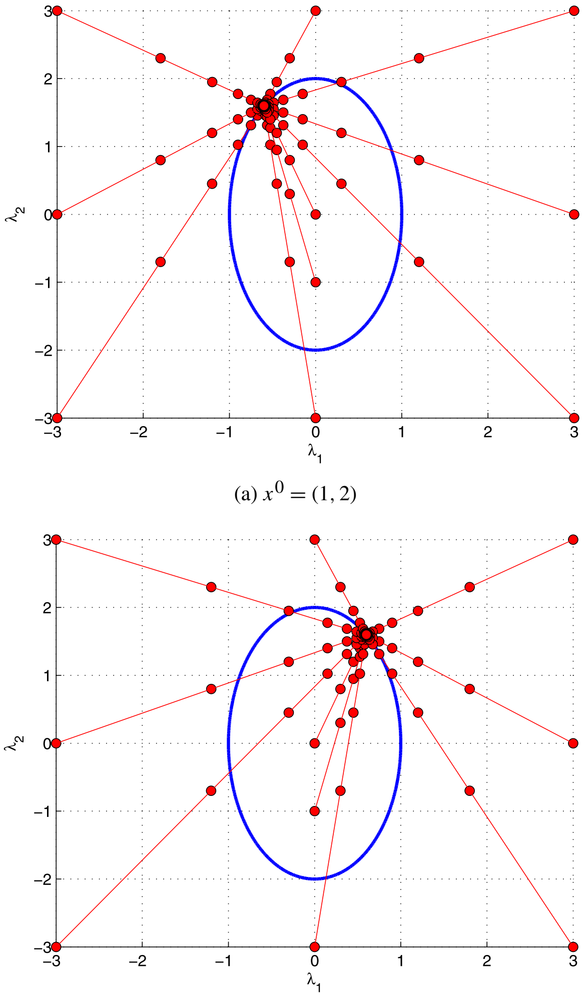
$$\bar{\lambda} = \left(-\frac{(x_2^0)^2 - (x_1^0)^2}{((x_1^0)^2 + (x_2^0)^2)}, \frac{4x_1^0 x_2^0}{(x_1^0)^2 + (x_2^0)^2} \right).$$

Note that the latter is a critical multiplier for any choice of x^0, and that convergence is slow (only linear).

Example 2.2 ([19, Example 7]) The equality-constrained problem

$$\begin{array}{ll} \text{minimize} & x_1^2 - x_2^2 + x_3^2 \\ \text{subject to} & x_1^2 + x_2^2 - x_3^2 = 0, \qquad x_1 x_3 = 0, \end{array}$$

Fig. 1 Dual trajectories for Example 2.1



(a) $x^0 = (1, 2)$

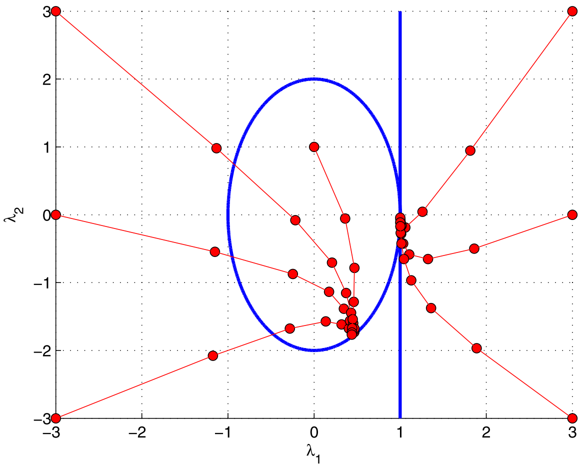
(b) $x^0 = (2, 1)$

has a (non-isolated) solution $\bar{x} = 0$, with $\mathcal{M}(\bar{x}) = \mathbf{R}^2$. This solution violates both, the regularity condition (2.2) (since $F'(\bar{x}) = 0$), and SOSC (2.4) for any $\bar{\lambda} \in \mathbf{R}^2$. Critical multipliers are those $\bar{\lambda}$ that satisfy $\bar{\lambda}_1 = 1$ or $4\bar{\lambda}_1^2 + \bar{\lambda}_2^2 = 4$.

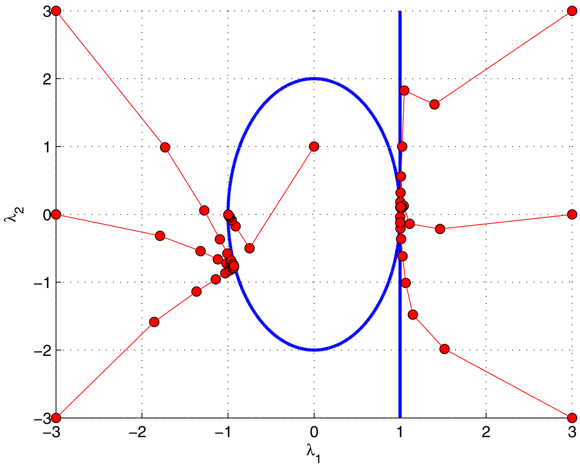
In Figs. 2 and 3, critical multipliers are represented by the thick line (they form an oval and a vertical line). Some Newton–Lagrange dual trajectories from primal starting points $x^0 = (1, 2, 3)$ and $x^0 = (3, 2, 1)$ are shown in Fig. 2. Figure 3 shows distribution of dual iterates at the time of termination of the method according to the stopping criterion, for dual trajectories generated starting from the points on the grid in the domain $[-3, 3] \times [-3, 3]$ (step of the grid is $1/4$).

It is interesting to observe that even though the solution $\bar{x} = 0$ is non-isolated, the Newton–Lagrange primal trajectories are attracted specifically by this (irregular)

Fig. 2 Dual trajectories for Example 2.2



(a) $x^0 = (1, 2, 3)$



(b) $x^0 = (3, 2, 1)$

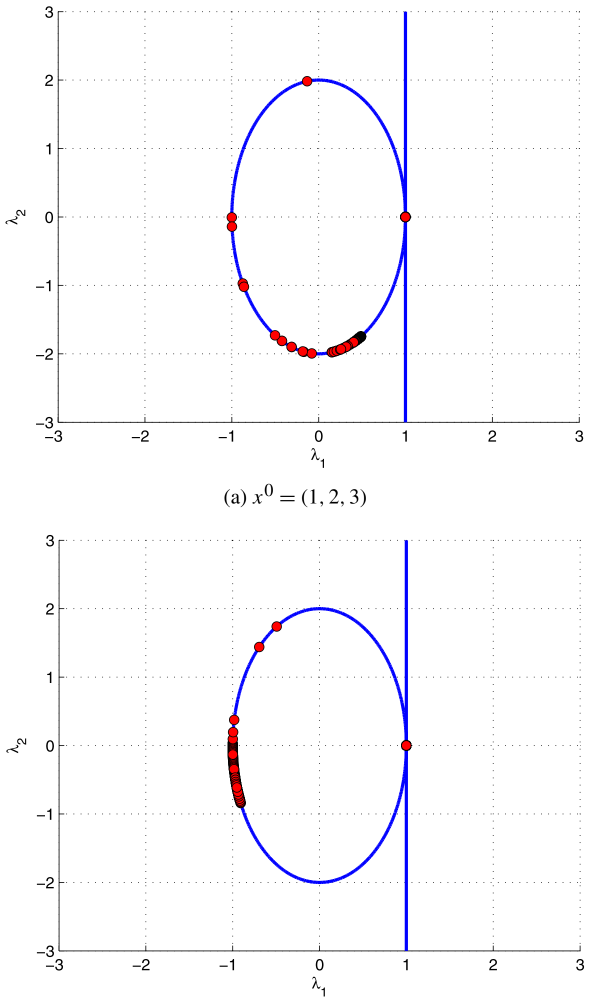
solution, while dual trajectories are attracted by associated critical multipliers. Again, convergence is slow.

Example 2.3 The equality-constrained problem

$$\begin{aligned} &\text{minimize} \quad x_1^2 - x_2^2 + 2x_3^2 \\ &\text{subject to} \quad -\frac{1}{2}x_1^2 + x_2^2 - \frac{1}{2}x_3^2 = 0, \qquad x_1 x_3 = 0, \end{aligned}$$

has the unique solution $\bar{x} = 0$, with $\mathcal{M}(\bar{x}) = \mathbf{R}^2$. This solution violates the regularity condition (2.2) (since $F'(\bar{x}) = 0$) but satisfies SOSC with any $\bar{\lambda}$ such that $\bar{\lambda}_1 \in (0, 1)$,

Fig. 3 Distribution of dual iterates at the time of termination for Example 2.2



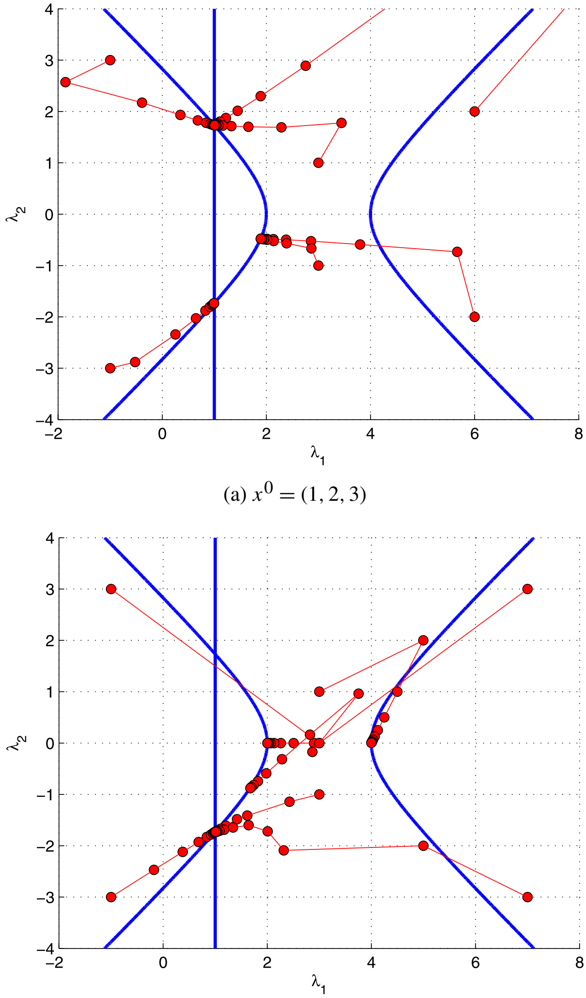
(a) $x^0 = (1, 2, 3)$

(b) $x^0 = (3, 2, 1)$

$(\bar{\lambda}_1 - 3)^2 - \bar{\lambda}_2^2 > 1$. Critical multipliers are those $\bar{\lambda}$ that satisfy $\bar{\lambda}_1 = 1$ or $(\bar{\lambda}_1 - 3)^2 - \bar{\lambda}_2^2 = 1$.

Figures 4 and 5 show the same kind of information for this example as Figs. 2 and 3 for Example 2.2. In particular, thick line represents critical multipliers (they form two branches of a hyperbola and a vertical line). Note that in Fig. 5 (b), there is one dual iterate, which is not close to any critical multiplier at the time of termination. This is a result of non-convergence of the dual trajectory. Figure 6 presents the run that produces this point. Just as an aside, we observe that, even in this case, the set of critical multipliers seems to play an important role for the behaviour of the dual trajectory, as it appears to be moving along this set.

Fig. 4 Dual trajectories for Example 2.3



(a) $x^0 = (1, 2, 3)$

(b) $x^0 = (3, 2, 1)$

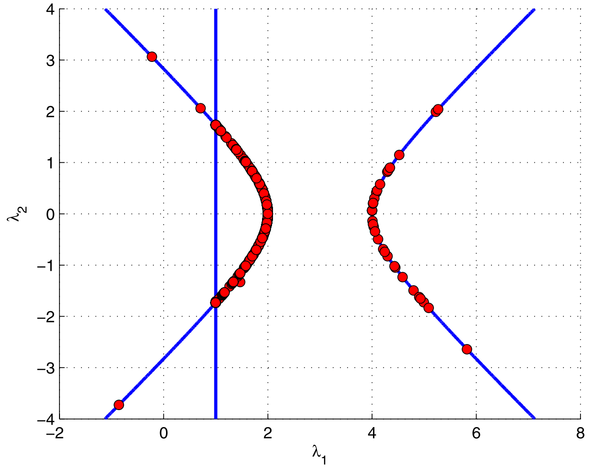
Example 2.4 The equality-constrained problem

$$\begin{aligned} &\text{minimize} \quad x_1^2 + x_2^2 + x_3^2 \\ &\text{subject to} \quad x_1 + x_2 + x_3 + x_1^2 + x_2^2 + x_3^2 = 0, \qquad x_1 + x_2 + x_3 + x_1 x_3 = 0, \end{aligned}$$

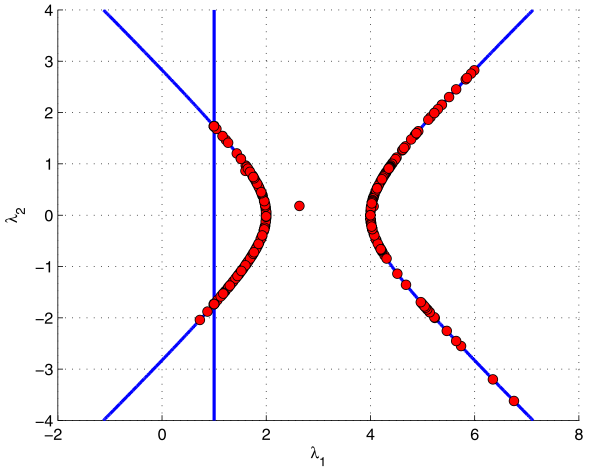
has the unique solution $\bar{x} = 0$, with $\mathcal{M}(\bar{x}) = \{\bar{\lambda} \in \mathbf{R}^2 \mid \bar{\lambda}_1 + \bar{\lambda}_2 = 0\}$. This solution violates the regularity condition (2.2) (since $F'(\bar{x}) = 0$), but satisfies SOSC with any $\bar{\lambda} \in \mathcal{M}(\bar{x})$ such that $\bar{\lambda}_1 > -2/3$. Critical multipliers are $\bar{\lambda} = (-6/5, 6/5)$ and $\bar{\lambda} = (-2/3, 2/3)$.

Figure 7 shows the set of multipliers (thick line) and some Newton–Lagrange dual trajectories for $x^0 = (1, 2, 3)$. Critical multipliers $\bar{\lambda} = (-6/5, 6/5)$ and $\bar{\lambda} =$

Fig. 5 Distribution of dual iterates at the time of termination for Example 2.3



(a) $x^0 = (1, 2, 3)$



(b) $x^0 = (3, 2, 1)$

$(-2/3, 2/3)$ can be barely seen in the figure, because of dual iterates accumulating around those points. As expected, convergence is slow.

Note that for $\lambda^0 = (-1, 0)$, the Newton–Lagrange method fails to make a step.

We next explain how the notion of critical multipliers extends to the case of mixed constraints (1.1).

Let the primal-dual trajectory $\{(x^k, \lambda^k, \mu^k)\}$ be generated by SQP algorithm. Suppose that the primal trajectory $\{x^k\}$ converges to a solution $\bar{x}$ of (1.1). It is quite natural to assume that the set

$$I_k = \{i = 1, \dots, m \mid G_i(x^k) + \langle G_i'(x^k), \xi^k \rangle = 0\}$$

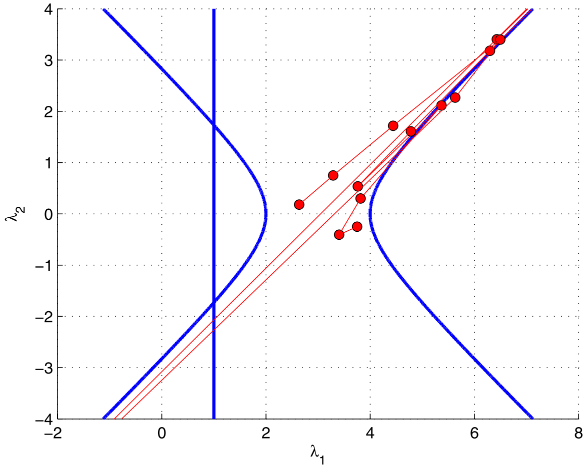


Fig. 6 Dual trajectory for Example 2.3, $x^0 = (3, 2, 1)$, $\lambda^0 = (3.75, -0.25)$

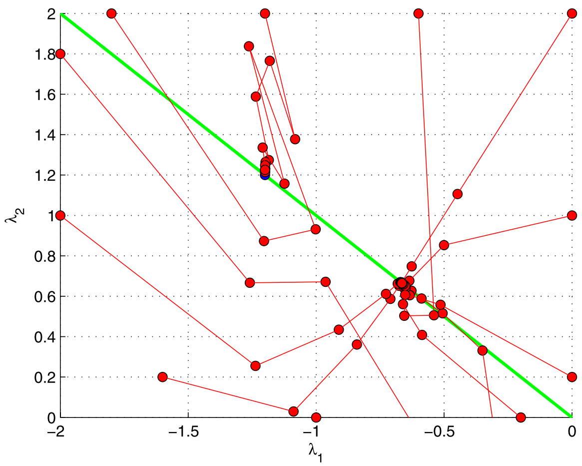


Fig. 7 Dual trajectories for Example 2.4, $x^0 = (1, 2, 3)$

of indices of inequality constraints active at the computed stationary points ξ^k of SQP subproblems remains unchanged for k sufficiently large. This is actually automatic when $\{(\lambda^k, \mu^k)\}$ tends to a multiplier $(\bar{\lambda}, \bar{\mu}) \in \mathcal{M}(\bar{x})$ satisfying strict complementarity, i.e., such that $\bar{\mu}_{I(\bar{x})} > 0$. In other cases, the assumption that the set I_k is asymptotically unchanged may not hold, but this still seems to be reasonable numerical behaviour, which should not be unusual. This is also confirmed by our examples, where we never encountered a situation of I_k changing infinitely often (but, of course, we do not claim that this should be the case always). Note also that if this stabilization property does not hold, one should hardly expect convergence of the dual trajectory, in general. Therefore, it is reasonable to put the case when I_k does not stabilize in the scenario of non-convergence of the dual sequence, while convergence of the latter can be analyzed under the assumption that I_k is stable from some iteration on.

Assuming that $I_k = I$ for all k large enough, we have that $\mu_i^k = 0 \; \forall i \in \{1, \ldots, m\} \setminus I$, for each such k. Then, as is readily seen from (1.4), $(\xi^k, \lambda^{k+1}, \mu^{k+1})$ satisfies the equations

$$
\begin{aligned}
&f'(x^k) + \frac{\partial^2 L_I}{\partial x^2}(x^k, \lambda^k, \mu_I^k)\xi^k + (F'(x^k))^{\mathrm{T}}\lambda^{k+1} + (G_I'(x^k))^{\mathrm{T}}\mu_I^{k+1} = 0, \\
&F(x^k) + F'(x^k)\xi^k = 0, \qquad G_I(x^k) + G_I'(x^k)\xi^k = 0,
\end{aligned}
\tag{2.5}
$$

where we have defined

$$
L_I : \mathbf{R}^n \times \mathbf{R}^l \times \mathbf{R}^{|I|} \to \mathbf{R}, \quad L_I(x, \lambda, \mu_I) = f(x) + \langle \lambda, F(x)\rangle + \langle \mu_I, G_I(x)\rangle.
$$

Note that the latter is the Lagrangian of the following equality-constrained optimization problem:

$$
\begin{aligned}
&\text{minimize} \quad f(x) \\
&\text{subject to} \quad F(x) = 0, \qquad G_I(x) = 0.
\end{aligned}
\tag{2.6}
$$

Assuming that $\{\xi^k\}$ converges to 0 (which is an immediate consequence of the convergence of $\{x^k\}$ if the stepsize α_k stays bounded away from zero), and assuming that $\{(\lambda^k, \mu^k)\}$ is bounded, by passing to the limit in (2.5) (along an appropriate subsequence) we conclude that $\bar{x}$ is a stationary point of (2.6). And, in particular, $I \subset I(\bar{x})$. Moreover, any $(\bar{\lambda}, \bar{\mu}) \in \mathcal{M}(\bar{x})$ satisfying $\bar{\mu}_i = 0\ \forall i \in \{1, \ldots, m\} \setminus I$ corresponds to the multiplier $(\bar{\lambda}, \bar{\mu}_I)$ of (2.6) associated with $\bar{x}$, and according to (2.5), $\{(x^k, \lambda^k, \mu_I^k)\}$ can be thought as generated by the Newton–Lagrange method for (2.6). This motivates the following

Definition 2.2 A multiplier $(\bar{\lambda}, \bar{\mu}) \in \mathcal{M}(\bar{x})$ associated with a stationary point $\bar{x}$ of problem (1.1) is referred to as *critical* with respect to a given index set $I \subset I(\bar{x})$, if $\bar{\mu}_i = 0\ \forall i \in \{1, \ldots, m\} \setminus I$, and the multiplier $(\bar{\lambda}, \bar{\mu}_I)$ associated with stationary point $\bar{x}$ of the equality-constrained problem (2.6) is critical for this problem in the sense of Definition 2.1.

Note that if $\{(\lambda^k, \mu^k)\}$ converges to some $(\bar{\lambda}, \bar{\mu}) \in \mathcal{M}(\bar{x})$ then

$$
I_+(\bar{x}, \bar{\mu}) \subset I \subset I(\bar{x}),
\tag{2.7}
$$

where $I_+(\bar{x}, \bar{\mu}) = \{i \in I(\bar{x}) \mid \bar{\mu}_i > 0\}$. Observe that if strong second-order sufficient condition for optimality (SSOSC)

$$
\left\langle \frac{\partial^2 L}{\partial x^2}(\bar{x}, \bar{\lambda}, \bar{\mu})\xi, \xi \right\rangle > 0 \quad \forall \xi \in (\ker F'(\bar{x}) \cap \ker G'_{I_+(\bar{x}, \bar{\mu})}(\bar{x})) \setminus \{0\}
$$

holds with this $(\bar{\lambda}, \bar{\mu})$, then SOSC holds with $(\bar{\lambda}, \bar{\mu}_I)$ for (2.6) at $\bar{x}$. Hence, in the case of SSOSC, $(\bar{\lambda}, \bar{\mu})$ cannot be a critical multiplier.

3 Scenarios of dual behaviour

We next discuss possible scenarios for dual SQP trajectories $\{(\lambda^k, \mu^k)\}$, assuming that the set I_k of active constraints of SQP subproblems (1.4) stabilizes: $I_k = I$ for all k large enough. Let us discuss separately three possibilities.

Scenario 1: *Dual trajectory does not converge.* According to our numerical experience, and as expected, this usually leads to slow primal convergence. How likely is this scenario? In our experience convergence of the dual sequence is far more typical than non-convergence. However, for equality-constrained problems, when the regularity condition (2.2) does not hold and there are no critical multipliers associated with $\bar{x}$ (i.e., no "natural" attractors), this behaviour is at least possible (if not to say typical). Example 2.3 demonstrates that this behaviour is also possible when critical multipliers do exist (see Fig. 6), although in this case it is definitely not typical (see [25]).

Recall, however, that for equality-constrained problems LICQ and MFCQ are the same, and so the latter cannot hold in the case of non-unique multipliers. Thus, the following question arises: Can the dual trajectory be non-convergent when MFCQ holds (in the case of mixed constraints)? The positive answer to this question is actually very obvious if one would either allow for completely artificial examples (e.g., with duplicated constraints) or consider very specially structured problems (such as MPCC, see below), *and* assume further that the QP solver being used can pick up *any* multiplier of SQP subproblem when there are many. Indeed, consider the problem with two *identical* inequality constraints. These constraints can satisfy or violate MFCQ but, in any case, the two constraints of SQP subproblem will also be identical. Hence, the multipliers associated with a solution of this subproblem will normally not be unique. Then, by picking up appropriate multipliers, one can enforce any kind of limiting behaviour. This is one reason why, in this discussion, we should probably restrict ourselves to the case when QP multipliers are uniquely defined, so that we could say that the trajectory $\{(x^k, \lambda^k, \mu_I^k)\}$ is uniquely defined by the Newton–Lagrange steps for (2.6). In addition, in practice, QP solvers pick up the multipliers (when they are not unique) not arbitrarily but according to some prescribed rules. Thus, when dealing with the case of nonunique QP multipliers one should take these rules into account (as it happens in the case of MPCC; see below).

For these reasons, we feel that the answer to the question above should probably result not from consideration of artificial examples but from computational practice, and be restricted to specific *implementations* of algorithms (rather than algorithms themselves).

We next discuss the case when $\{(\lambda^k, \mu^k)\}$ converges to some $(\bar{\lambda}, \bar{\mu}) \in \mathcal{M}(\bar{x})$.

Scenario 2: *Dual trajectory converges to a noncritical (with respect to the given I) multiplier.* According to existing theoretical and numerical evidence for equality-constrained problems (see [25]), in this case the primal convergence rate is superlinear. This situation is typical if the constraints of (2.6) are regular at $\bar{x}$, which, in turn, may happen when the set I is strictly smaller than $I(\bar{x})$. Note that the latter assumes that the limiting multiplier $(\bar{\lambda}, \bar{\mu})$ violates strict complementarity. Note also that when I is strictly smaller than $I(\bar{x})$, the constraints of (2.6) can be regular at $\bar{x}$ even when standard constraint qualifications do not hold for the original problem (1.1). In this case, the multiplier $(\bar{\lambda}, \bar{\mu}_I)$ associated with stationary point $\bar{x}$ of (2.6) is unique, and this unique multiplier has no special reason to be critical. Note, finally, that in this case, dual convergence is also superlinear.

Example 3.1 ([36]) The inequality-constrained problem

$$\begin{aligned} &\text{minimize} \quad x_1 \\ &\text{subject to} \quad (x_1-2)^2+x_2^2 \le 4, \qquad (x_1-4)^2+x_2^2 \le 16, \end{aligned}$$

has the unique solution $\bar{x}=0$, with $\mathcal{M}(\bar{x})=\{\bar{\mu}\in\mathbf{R}^2 \mid \bar{\mu}_1+2\bar{\mu}_2=1/4, 0\le\bar{\mu}_2\le 1/8\}$. This solution satisfies MFCQ (but not LICQ) and SSOSC holds with any $\bar{\mu}\in\mathcal{M}(\bar{x})$.

In [36], it was demonstrated that for this problem, the full SQP step (with $\alpha_k=1$) may not provide a superlinear decrease of the distance to $\bar{x}$, even from a point arbitrarily close to $\bar{x}$. However, for the *next* step, this will no longer be the case. Numerical experiments show that here $I=\{2\}$, $\{x^k\}$ converges to $\bar{x}$ (apart from a few cases of failure for some starting points), $\{\mu^k\}$ converges to $\bar{\mu}=(0,1/8)$, and the rate of primal and dual convergence is superlinear. Moreover, the unique constraint of the corresponding problem (2.6) is regular at $\bar{x}$, $\bar{\mu}_2=1/8$ is the unique associated multiplier, and SOSC holds for (2.6) at $\bar{x}$ (with this multiplier). Hence, $\bar{\mu}=(0,1/8)$ is a non-strictly complementary and non-critical (with respect to the given I) multiplier of (1.1). (Actually, it can be checked that there are no critical multipliers with respect to any $I\subset I(\bar{x})=\{1,2\}$.)

SNM-FB also usually converges superlinearly for this example. For SNM-NR, on the other hand, the dual trajectory does not converge (it has two distinct accumulation points; the reason for this will be explained below), and the primal trajectory either converges slowly or there is no convergence at all.

The above considerations show that the case when the set I is strictly smaller than $I(\bar{x})$ is possible, in which case the constraints of the associated equality-constrained problem (2.6) can be regular, we may have convergence to a non-critical multiplier, and fast primal-dual convergence as well. However, if the constraints of (2.6) are degenerate at $\bar{x}$ (in particular, if $I=I(\bar{x})$ and standard constraint qualifications do not hold for the original problem (1.1)), convergence to a non-critical multiplier is highly unlikely to occur. The expected behaviour in this case is the following: either $\{(\lambda^k,\mu^k)\}$ does not converge (see Scenario 1) or it converges to a critical multiplier and, in both situations, primal convergence is slow.

Scenario 3: *Dual trajectory converges to a critical multiplier.* As mentioned above, for purely equality-constrained problems this scenario appears to be typical, unless critical multipliers do not exist [25]. More generally, if the limiting multiplier $(\bar{\lambda},\bar{\mu})$ satisfies the strict complementarity condition then, by (2.7), $I=I(\bar{x})$. Thus, according to the discussion above, if standard constraint qualifications do not hold for problem (1.1), it is natural to expect the multiplier $(\bar{\lambda},\bar{\mu})$ to be critical and convergence to be slow. Note that in this case, $\bar{\mu}_i=0\ \forall i\in\{1,\dots,m\}\setminus I$ and $\forall(\bar{\lambda},\bar{\mu})\in\mathcal{M}(\bar{x})$, which means that the corresponding condition in Definition 2.2 holds automatically with this I. Furthermore, in this case, if SOSC holds with $(\bar{\lambda},\bar{\mu})$ for (1.1) then SOSC holds with $(\bar{\lambda},\bar{\mu}_I)$ for (2.6), and in particular, such $(\bar{\lambda},\bar{\mu})$ cannot be critical.

Example 3.2 The inequality-constrained problem

$$\begin{aligned} &\text{minimize} \quad -x^2 \\ &\text{subject to} \quad x^2 \leq 0 \end{aligned}$$

has the unique feasible point (hence, unique solution) $\bar{x} = 0$, with $\mathcal{M}(\bar{x}) = \mathbf{R}_+$. This solution violates MFCQ but satisfies SOSC for any $\bar{\mu} > 1$.

It is easy to check that the SQP step (with $\alpha_k = 1$) is given by $x^{k+1} = x^k/2$, $\mu^{k+1} = 1 - (1 - \mu^k)/2$, and $\{\mu^k\}$ converges to the strictly complementary multiplier $\bar{\mu} = 1$, which is the unique critical (with respect to $I = I(\bar{x}) = \{1\}$) multiplier. Convergence rate is only linear.

SNM-NR and SNM-FB demonstrate similar behaviour for this example.

The next example demonstrates that this scenario is possible also when MFCQ holds at the solution $\bar{x}$.

Example 3.3 ([25, Example 4]) The inequality-constrained problem

$$\begin{aligned} &\text{minimize} \quad -x_1 \\ &\text{subject to} \quad x_1 - x_2^2 \leq 0, \qquad x_1 + x_2^2 \leq 0 \end{aligned}$$

has the unique solution $\bar{x} = 0$, with $\mathcal{M}(\bar{x}) = \{\bar{\mu} \in \mathbf{R}^2_+ \mid \bar{\mu}_1 + \bar{\mu}_2 = 1\}$. This solution satisfies MFCQ (but not LICQ) and SOSC holds for any $\bar{\mu} \in \mathcal{M}(\bar{x})$ such that $\bar{\mu}_1 < \bar{\mu}_2$.

The SQP subproblem (1.4) takes the form

$$\begin{aligned} &\text{minimize} \quad -\xi_1 - (\mu_1^k - \mu_2^k)\xi_2^2 \\ &\text{subject to} \quad x_1^k - (x_2^k)^2 + \xi_1 - 2x_2^k\xi_2 \leq 0, \qquad x_1^k + (x_2^k)^2 + \xi_1 + 2x_2^k\xi_2 \leq 0. \end{aligned}$$

Let, for simplicity, $x_1^k = 0$, $x_2^k \neq 0$. Suppose that μ^k is close enough to $\mathcal{M}(\bar{x})$. Then the point $\xi^{k+1} = (0, -x_2^k/2)$ is stationary in this subproblem, with both constraints being active. Hence, the primal SQP step (with $\alpha_k = 1$) is given by $x^{k+1} = x^k/2 = (0, x_2^k/2)$. In particular, both QP constraints remain active along the primal trajectory, and hence, $I = I(\bar{x}) = \{1, 2\}$.

The multiplier of SQP subproblem is given by

$$\mu^{k+1} = \left(\frac{1}{2} + \frac{1}{4}(\mu_1^k - \mu_2^k), \frac{1}{2} - \frac{1}{4}(\mu_1^k - \mu_2^k) \right).$$

It follows that $\mu_1^{k+1} - \mu_2^{k+1} = \frac{1}{2}(\mu_1^k - \mu_2^k)$, and hence, $(\mu_1^k - \mu_2^k) \to 0$, which implies that $\{\mu^k\} \to (1/2, 1/2)$, which is a strictly complementary multiplier, and the unique critical multiplier (with respect to $I = \{1, 2\}$). Again, convergence rate is only linear.

In our numerical experiments with this problem, SQP either converges slowly to the critical multiplier or fails. SNM-FB usually behaves similarly to SQP, but it is more robust and sometimes converges superlinearly to a noncritical multiplier. SNM-NR either converges slowly to the critical multiplier, or fails (Λ_k becomes almost singular), or sometimes terminates finitely (but still at the critical multiplier!).

Our next example demonstrates how this scenario (convergence to a critical multiplier) can take place with limiting multiplier violating strict complementarity. Note that in the latter case, a multiplier satisfying SOSC but not SSOSC can be critical (see Example 3.5 below).

Example 3.4 ([38, (63)]) The inequality-constrained problem

$$\begin{array}{ll} \text{minimize} & x_1 \\ \text{subject to} & -x_1 \le 0, \qquad (x_1-2)^2 + x_2^2 \le 4, \end{array}$$

has the unique solution $\bar{x} = 0$, with $\mathcal{M}(\bar{x}) = \{\bar{\mu} \in \mathbf{R}^2 \mid \bar{\mu}_1 = 1 - 4\bar{\mu}_2, 0 \le \bar{\mu}_2 \le 1/4\}$. This solution satisfies MFCQ (but not LICQ) and SOSC (and even SSOSC) holds with any $\bar{\lambda} \in \mathcal{M}(\bar{x})$, except for $\bar{\mu} = (1, 0)$, which is the unique critical multiplier for $I = \{1, 2\}$.

The SQP subproblem (1.4) takes the form

$$\begin{array}{ll} \text{minimize} & \xi_1 + \mu_2^k(\xi_1^2 + \xi_2^2) \\ \text{subject to} & -x_1^k - \xi_1 \le 0, \qquad (x_1^k - 2)^2 + (x_2^k)^2 - 4 + 2(x_1^k - 2)\xi_1 + 2x_2^k \xi_2 \le 0. \end{array}$$

Since we expect the dual convergence to $\bar{\mu} = (1, 0)$, we should consider the case when the first constraint of SQP subproblem is active: $\xi_1^k = -x_1^k$, and hence, the full primal SQP step (with $\alpha_k = 1$) gives $x_1^{k+1} = 0$.

Now we can deal solely with the case when $x_1^k = 0$. In this case, the point $\xi^k = (0, -x_2^k/2)$ is stationary in SQP subproblem, with both constraints being active. Hence, the primal SQP step is given by $x^{k+1} = x^k/2 = (0, x_2^k/2)$. In particular, both QP constraints remain active, and hence, $I = I(\bar{x}) = \{1, 2\}$.

The multiplier of SQP subproblem is given by

$$\mu^{k+1} = (1 - 2\mu_2^k, \mu_2^k/2).$$

It follows that $\{\mu^k\} \to \bar{\mu} = (1, 0)$. The constraints of (2.6) are degenerate at $\bar{x}$, and $\bar{\mu} = (1, 0)$ is the unique associated critical multiplier. Hence, this $\bar{\mu}$ is a non-strictly complementary critical (with respect to the taken I) multiplier of (1.1).

Figure 8 presents the set of multipliers (thick line) and some SQP dual trajectories for $x^0 = (1, 2)$. As expected, convergence rate is only linear.

In our numerical experiments with this problem, SQP usually converges slowly to the critical multiplier. Sometimes it terminates finitely (but still at the critical multiplier!). SNM-FB usually behaves similarly to SQP, but sometimes it converges superlinearly to a noncritical multiplier. SNM-NR either converges slowly to the critical multiplier, or fails (Λ_k becomes almost singular).

We proceed with a brief discussion of dual behaviour of SNM-NR and SNM-FB.

Let the current iterate (x^k, λ^k, μ^k) be close enough to $(\bar{x}, \bar{\lambda}, \bar{\mu})$, where $(\bar{\lambda}, \bar{\mu}) \in \mathcal{M}(\bar{x})$. Denote $N = \{1, \ldots, m\} \setminus I(\bar{x})$, $I_+ = I_+(\bar{x}, \bar{\mu})$, $I_0 = I(\bar{x}) \setminus I_+(\bar{x}, \bar{\mu})$. Iteration of SNM-NR can be interpreted as the Newton step for the "branch" equation of the KKT system of the form

$$\Phi_{J_1^k, J_2^k}(x, \lambda, \mu) = 0, \tag{3.1}$$

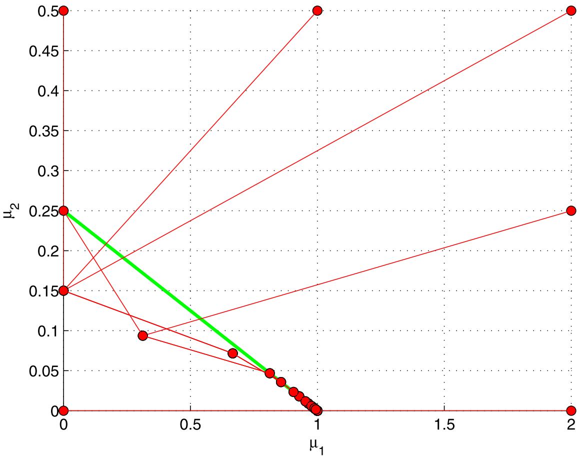


Fig. 8 Dual trajectories for Example 3.4, $x^0 = (1, 2)$

where

$$\Phi_{J_1^k, J_2^k} : \mathbf{R}^n \times \mathbf{R}^l \times \mathbf{R}^m \to \mathbf{R}^n \times \mathbf{R}^l \times \mathbf{R}^m,$$

$$\Phi_{J_1^k, J_2^k}(x, \lambda, \mu) = \left(\frac{\partial L}{\partial x}(x, \lambda, \mu), F(x), G_{I_+}(x), G_{J_1^k}(x), \mu_{J_2^k}, \mu_N \right),$$

and (J_1^k, J_2^k) is a partition of I_0 such that

$$\{i \in I_0 \mid \mu_i^k > G_i(x^k)\} \subset J_1^k, \qquad \{i \in I_0 \mid \mu_i^k < G_i(x^k)\} \subset J_2^k. \tag{3.2}$$

Note that $\Phi_{J_1^k, J_2^k}(x^k, \lambda^k, \mu^k)$ does not depend on a specific choice of a partition (J_1^k, J_2^k) satisfying (3.2), and that every partition gives a valid element of $\partial_B \Phi_{J_1^k, J_2^k}(x^k, \lambda^k, \mu^k)$. Therefore, the presented interpretation of SNM-NR is correct. We refer the reader to [23, Sect. 5] for more details.

Now, suppose that (J_1^k, J_2^k) stabilizes: $(J_1^k, J_2^k) = (J_1, J_2)$ for all k large enough. Then SNM-NR reduces to the Newton–Lagrange method for the equality-constrained problem

$$\begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & F(x) = 0, \qquad G_{I_+ \cup J_1}(x) = 0. \end{array}$$

Hence, under the assumption of stabilization of (J_1^k, J_2^k), all the previous discussion (based on reduction to an equality-constrained problem) readily applies.

However, according to our experience, stabilization of (J_1^k, J_2^k) in SNM-NR appears to be not as natural as stabilization of I_k for SQP. In Example 3.1, dual convergence of SNM-NR fails precisely because the partitions (J_1^k, J_2^k) do not stabilize. See also Examples 4.1 and 4.5 below.

As for SNM-FB, the situation appears to be more subtle. Nevertheless, an iteration of SNM-FB can still be interpreted as *inexact* Newton–Lagrange step for some underlying equality-constrained problem. We skip technical details and instead address the reader to the numerical evidence in Sect. 4.

We complete this section with a special class of degenerate problems, namely, mathematical programs with complementarity constraints (MPCC). Specifically, we consider the problem

$$\begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & H_1(x) \geq 0, \qquad H_2(x) \geq 0, \qquad \langle H_1(x),\ H_2(x)\rangle \leq 0, \end{array} \tag{3.3}$$

where $f: \mathbf{R}^n \to \mathbf{R}$ is a smooth function and $H_1, H_2: \mathbf{R}^n \to \mathbf{R}^m$ are smooth mappings. Standard equality and inequality constraints can be easily added to this problem setting, making obvious (non-essential) modifications in what follows. As is well known, every feasible point of MPCC violates MFCQ. Thus, MPCC is inherently degenerate.

The (standard) Lagrangian of problem (3.3) has the form

$$L(x, \mu) = f(x) - \langle \mu_1, H_1(x)\rangle - \langle \mu_2, H_2(x)\rangle + \mu_3\langle H_1(x), H_2(x)\rangle,$$
$$x \in \mathbf{R}^n, \mu = (\mu_1, \mu_2, \mu_3) \in \mathbf{R}^m \times \mathbf{R}^m \times \mathbf{R}.$$

We next recall terminology which is now standard in MPCC literature. Define the so-called MPCC-Lagrangian of problem (3.3):

$$\mathcal{L}: \mathbf{R}^n \times (\mathbf{R}^m \times \mathbf{R}^m) \to \mathbf{R},$$
$$\mathcal{L}(x, \nu) = f(x) - \langle \nu_1, H_1(x)\rangle - \langle \nu_2, H_2(x)\rangle,\ \nu = (\nu_1, \nu_2).$$

To a feasible point $\bar{x}$ we associate the index sets

$$I_1 = I_1(\bar{x}) = \{i \in I \mid (H_1)_i(\bar{x}) = 0\},$$
$$I_2 = I_2(\bar{x}) = \{i \in I \mid (H_2)_i(\bar{x}) = 0\}, \qquad I_0 = I_1 \cap I_2.$$

A feasible point $\bar{x}$ of (3.3) is said to be a strongly stationary point of this problem if there exists $\bar{\nu} = (\bar{\nu}_1, \bar{\nu}_2) \in \mathbf{R}^m \times \mathbf{R}^m$ satisfying

$$\frac{\partial \mathcal{L}}{\partial x}(\bar{x}, \bar{\nu}) = 0, \qquad (\bar{\nu}_1)_{I_2 \setminus I_1} = 0, \qquad (\bar{\nu}_2)_{I_1 \setminus I_2} = 0, \qquad (\bar{\nu}_1)_{I_0} \geq 0, \qquad (\bar{\nu}_2)_{I_0} \geq 0.$$

Any such $\bar{\nu}$ is called an MPCC-multiplier associated with $\bar{x}$.

Throughout the rest of this section we assume that the so-called MPCC-linear independence constraint qualification (MPCC-LICQ) holds at $\bar{x}$, i.e.,

$$(H_1)_i'(\bar{x}), \quad i \in I_1, \qquad (H_2)_i'(\bar{x}), \quad i \in I_2 \quad \text{are linearly independent.}$$

It was shown in [33] that if MPCC-LICQ holds at a local solution $\bar{x}$ of (3.3), then this point is strongly stationary and the associated MPCC-multiplier $\bar{\nu}$ is unique. Moreover, in this case $\bar{x}$ is a stationary (in the classical sense) point of problem (3.3), and

$$\mathcal{M}(\bar{x}) = \left\{ \mu = (\mu_1, \mu_2, \mu_3) \in \mathbf{R}^m \times \mathbf{R}^m \times \mathbf{R} \left| \begin{array}{l} (\bar{\mu}_1)_i = (\bar{\nu}_1)_i + \bar{\mu}_3 (H_2)_i(\bar{x}), \\ \quad i \in I_1 \setminus I_2, \\ (\bar{\mu}_2)_i = (\bar{\nu}_2)_i + \bar{\mu}_3 (H_1)_i(\bar{x}), \\ \quad i \in I_2 \setminus I_1, \\ (\bar{\mu}_1)_i = (\bar{\nu}_1)_i, i \in I_2, \\ (\bar{\mu}_2)_i = (\bar{\nu}_2)_i, i \in I_1, \\ \bar{\mu}_3 \geq \bar{\gamma} \end{array} \right. \right\},$$

where

$$\bar{\gamma} = \max \left\{ 0, \max_{i \in I_1 \setminus I_2} \left(-\frac{(\bar{\nu}_1)_i}{(H_2)_i(\bar{x})} \right), \max_{i \in I_2 \setminus I_1} \left(-\frac{(\bar{\nu}_2)_i}{(H_1)_i(\bar{x})} \right) \right\}$$

(see, e.g., [12, Proposition 4.1]). Thus, the set $\mathcal{M}(\bar{x})$ is a ray, with its origin corresponding to $\bar{\mu}_3 = \bar{\gamma}$.

It can be easily checked that the (standard) critical cone of problem (3.3) at $\bar{x}$ is given by

$$C(\bar{x}) = \left\{ \xi \in \mathbf{R}^n \left| \begin{array}{l} (H_1)'_{I_1 \setminus I_2}(\bar{x})\xi = 0, (H_2)'_{I_2 \setminus I_1}(\bar{x})\xi = 0, \\ (H_1)'_{I_0}(\bar{x})\xi \geq 0, (H_2)'_{I_0}(\bar{x})\xi \geq 0, \\ \langle f'(\bar{x}), \xi \rangle \leq 0 \end{array} \right. \right\}.$$

We say that MPCC-second-order sufficient condition (MPCC-SOSC) holds at a strongly stationary point $\bar{x}$ of problem (3.3) with the associated MPCC-multiplier $\bar{\nu}$, if

$$\frac{\partial^2 \mathcal{L}}{\partial x^2}(\bar{x}, \bar{\nu})[\xi, \xi] > 0 \quad \forall \xi \in C(\bar{x}) \setminus \{0\}. \tag{3.4}$$

According to [18, Proposition 1], MPCC-SOSC implies the usual SOSC for any $\bar{\mu}$ in the ray $\mathcal{M}(\bar{x})$, including the origin of this ray.

MPCC-SOSC is a rather strong condition. In particular, it cannot be linked to any second-order necessary condition for (3.3). By this we mean that a solution of (3.3) that satisfies MPCC-LICQ (and thus is strongly stationary) does not have to satisfy the condition obtained from (3.4) by replacing the strict inequality by non-strict (as "should" be, by natural analogy with the links between classical necessary conditions and sufficient conditions for optimality). We next state a second-order sufficient condition, which is weaker than MPCC-SOSC, and which is naturally connected to an appropriate second-order necessary condition.

Define the cone

$$C_2(\bar{x}) = \left\{ \xi \in \mathbf{R}^n \left| \begin{array}{l} (H_1)'_{I_1 \setminus I_2}(\bar{x})\xi = 0, (H_2)'_{I_2 \setminus I_1}(\bar{x})\xi = 0, \\ (H_1)'_{I_0}(\bar{x})\xi \geq 0, (H_2)'_{I_0}(\bar{x})\xi \geq 0, \\ \langle (H_1)'_i(\bar{x}), \xi \rangle \langle (H_2)'_i(\bar{x}), \xi \rangle = 0, i \in I_0, \\ \langle f'(\bar{x}), \xi \rangle \leq 0 \end{array} \right. \right\},$$

where the subscript "2" indicates that, unlike $C(\bar{x})$, this set takes into account the second-order information about the last constraint in (3.3). By direct comparison, $C_2(\bar{x}) \subset C(\bar{x})$.

We say that *piecewise SOSC* holds at a strongly stationary point $\bar{x}$ of problem (3.3) with the associated MPCC-multiplier $\bar{\nu}$, if

$$\frac{\partial^2 \mathcal{L}}{\partial x^2}(\bar{x}, \bar{\nu})[\xi, \xi] > 0 \quad \forall \xi \in C_2(\bar{x}) \setminus \{0\}. \tag{3.5}$$

Piecewise SOSC is indeed sufficient for optimality, even though it is evidently weaker than MPCC-SOSC. Moreover, the condition obtained from (3.5) by replacing the strict inequality by non-strict is necessary for optimality [33].

Regarding dual behaviour of Newton-type methods when applied to MPCC, the following two possibilities should be considered:

- If the upper-level strict complementarity condition (ULSCC) holds, that is,

  $$(\bar{\nu}_1)_{I_0} > 0, \qquad (\bar{\nu}_2)_{I_0} > 0,$$

  then there is exactly one multiplier violating strict complementarity, namely the basic multiplier $\bar{\mu}$ corresponding to $\bar{\mu}_3 = \bar{\gamma}$.
- If ULSCC does not hold then there are no strictly complementary multipliers.

It is very tempting to extend our discussion above of general degenerate problems to MPCC. For example, suppose that ULSCC holds. Since MPCC-SOSC implies SOSC with any multiplier, it is tempting to deduce that under MPCC-SOSC, if the dual trajectory converges then most likely it converges to the (non-strictly complementary) basic multiplier. Similarly, under piecewise SOSC (3.5), there may exist $\hat{\gamma} \geq \bar{\gamma}$ such that SOSC holds only with those multipliers $\bar{\mu}$ that correspond to $\bar{\mu}_3 > \hat{\gamma}$. And then it is tempting to deduce that under piecewise SOSC, if the dual trajectory converges then most likely it converges either to the critical multiplier $\bar{\mu}$ corresponding to $\bar{\mu}_3 = \hat{\gamma}$, or to the (non-strictly complementary) basic multiplier. The cases of violated ULSCC can be also treated as above.

However, the discussion for general mathematical programming problems should be applied to MPCCs with some care, because of very special structure of the latter. Specifically, according to [12], a quite possible (and, in fact, favorable) scenario for SQP applied to MPCC reformulation with slack variables $y_1 = H_1(x)$ and $y_2 = H_2(x)$ is the following: the primal trajectory hits a point satisfying exact complementarity, that is, $y_1^k \geq 0$, $y_2^k \geq 0$ and $\langle y_1^k, y_2^k \rangle = 0$ for some k. In this case, the constraints of the QP approximation about the corresponding point (x^k, y_1^k, y_2^k) violate standard constraint qualifications and, hence, the corresponding QP multipliers would normally be not unique. Moreover, exact complementarity is preserved for all subsequent iterations. Therefore, specificity of the QP solver should be taken into account (otherwise, as already discussed above, if QP solver may pick any multiplier then dual behaviour should be considered essentially arbitrary). In [12], this specificity consists of saying that "a QP solver always chooses a linearly independent basis", and this (together with other things, including ULSCC) results in dual convergence to the basic multiplier. Note, however, that if one assumes piecewise SOSC instead of MPCC-SOSC then dual convergence to the basic multiplier is still a bad thing, since it may not satisfy SOSC (even though it is noncritical, in general).

At the same time, if $\langle y_1^k, y_2^k \rangle > 0$ for all k then our discussion for general problems still perfectly applies and the corresponding conclusions remain valid. In the next example, MPCC-LICQ and piecewise SOSC are satisfied but MPCC-SOSC and ULSCC are violated.

Example 3.5 (`ralph2` in MacMPEC [27]) The MPCC

$$\begin{array}{ll} \text{minimize} & x_1^2 + x_2^2 - 4x_1x_2 \\ \text{subject to} & x_1 \geq 0, \qquad x_2 \geq 0, \qquad x_1x_2 \leq 0, \end{array}$$

has the unique solution $\bar{x} = 0$, with $\mathcal{M}(\bar{x}) = \{\bar{\mu} \in \mathbf{R}^3 \mid \bar{\mu}_1 = \bar{\mu}_2 = 0, \bar{\mu}_3 \geq 0\}$. This solution satisfies MPCC-LICQ but not MPCC-SOSC.

The SQP subproblem (1.4) takes the form

$$\begin{array}{ll} \text{minimize} & 2x_1^k\xi_1 + 2x_2^k\xi_2 - 4x_2^k\xi_1 - 4x_1^k\xi_2 + \xi_1^2 + \xi_2^2 - (4 - \mu_3^k)\xi_1\xi_2 \\ \text{subject to} & x_1^k + \xi_1 \geq 0, \qquad x_2^k + \xi_2 \geq 0, \qquad x_1^kx_2^k + x_2^k\xi_1 + x_1^k\xi_2 \leq 0. \end{array}$$

Suppose that $0 \leq \mu_3^k < 6$, and let $x_1^k = x_2^k \neq 0$. It can be easily seen that the first two constraints cannot be active in this case, and that $\xi^k = -x^k/2$ is the unique stationary point of SQP subproblem, with the last constraint being active. Hence, the primal SQP step (with $\alpha_k = 1$) is given by $x^{k+1} = x^k/2$. In particular, only the last QP constraint remains active along the primal trajectory. Hence, $I = \{3\}$, while $I(\bar{x}) = \{1, 2, 3\}$. The multiplier of SQP subproblem is given by

$$\mu^{k+1} = (0, 0, 1 + \mu_3^k/2).$$

It follows that $\{\mu^k\} \to \bar{\mu} = (0, 0, 2)$. The unique constraint of (2.6) is degenerate at $\bar{x}$, and $\bar{\mu}_3 = 2$ is an associated critical multiplier. Hence, this $\bar{\mu}$ is a non-strictly complementary critical (with respect to the chosen set I) multiplier of (1.1), and $I \neq I(\bar{x})$. As predicted, convergence is only linear.

Note that in this example, (2.6) has another critical multiplier $\bar{\mu}_3 = 6$, despite of SOSC (but not SSOSC!) for (1.1) being valid with $\bar{\mu} = (0, 0, 6)$.

Figure 9 presents the projection of the set of multipliers onto (μ_1, μ_3)-plane (thick line) and the projections of some SQP dual trajectories for $x^0 = (1, 1)$, $\mu_2^0 = 0$. Note that for $\mu_3^0 \geq 6$, one-step termination happens at a noncritical multiplier.

In our numerical experiments with this example, SQP usually has finite termination, but sometimes converges slowly to the critical multiplier $\bar{\mu} = (0, 0, 2)$. SNM-NR demonstrates similar behaviour, though it also fails quite often. SNM-FB demonstrates all kinds of behaviour: sometimes it fails, sometimes converges superlinearly to a noncritical multiplier, and sometimes converges slowly to the critical multiplier $\bar{\mu} = (0, 0, 2)$.

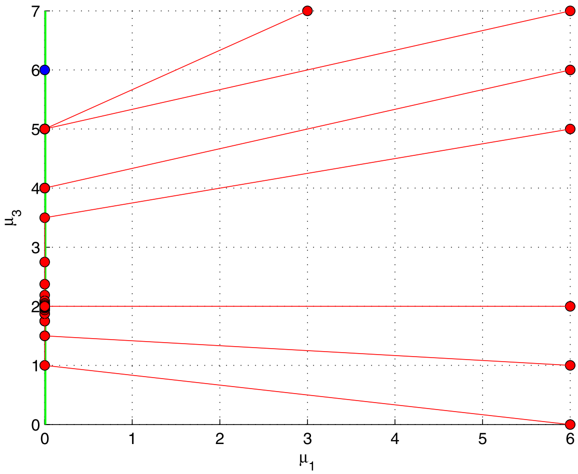


Fig. 9 Dual trajectories for Example 3.5, $x^0 = (1, 1)$, $\mu_2^0 = 0$

4 Further examples and numerical experiments

In this section, we put together some more examples of mathematical programming problems violating standard CQs, and briefly report on our numerical experience with them in Matlab environment.

We start with examples with no critical multipliers (Examples 4.1–4.15); they complement Example 3.1. These examples put in evidence that when there are no critical multipliers, SQP convergence is usually superlinear, despite degeneracy. Numerical results for these examples are reported in Table 1. For each problem, we used 10 random starting points (x^0, λ^0, μ^0) with components x_j^0 in $[\bar{x}_j - 10, \bar{x}_j + 10]$, $j = 1, \ldots, n$, and with components of (λ^0, μ^0) in $[-10, 10] \times [0, 10]$. Here, $\bar{x}$ is a solution specified in the corresponding example, or a convex combination (with equal coefficients) of solutions when there are many (like in Example 4.8). Example 4.4 is omitted from Table 1 because the algorithms converge to a nondegenerate local solution (see the discussion in Example 4.4). For each problem and each algorithm, we report on the number of successful runs (those for which the stopping criterion was satisfied after no more than 50 iterations), and on the number of times superlinear convergence was detected (in parentheses).

Example 4.1 ([38, p. 138]) The inequality-constrained problem

$$\begin{aligned} &\text{minimize} \quad \frac{1}{2}(x_1+1)^2 + \frac{1}{2}x_2^2 \\ &\text{subject to} \quad -x_1 \leq 0, \qquad (x_1-1)^2 + x_2^2 \leq 1, \end{aligned}$$

has the unique solution $\bar{x} = 0$, with $\mathcal{M}(\bar{x}) = \{\bar{\mu} \in \mathbf{R}^2 \mid \bar{\mu}_1 = 1 - 2\bar{\mu}_2, 0 \leq \bar{\mu}_2 \leq 1/2\}$. This solution satisfies MFCQ (but not LICQ), and SSOSC holds with all $\bar{\mu} \in \mathcal{M}(\bar{x})$.

SQP and SNM-FB converge superlinearly for this example (SQP actually has finite termination). For SNM-NR, partition (J_1^k, J_2^k) does not stabilize, the dual trajectory does not converge (has two distinct accumulation points), and the primal trajectory converges slowly, which sometimes results even in a failure because the iterations limit is reached.

Example 4.2 ([38, Example 1, "three-circle problem"]) The inequality-constrained problem

$$\begin{array}{ll} \text{minimize} & x_1 \\ \text{subject to} & (x_1-2)^2+x_2^2 \le 4, \qquad (x_1-4)^2+x_2^2 \le 16, \qquad x_1^2+(x_2-2)^2 \le 4, \end{array}$$

has the unique solution $\bar{x}=0$, with $\mathcal{M}(\bar{x})=\{\bar{\mu}\in\mathbf{R}^3 \mid \bar{\mu}_1=1/4-2\bar{\mu}_2, 0\le\bar{\mu}_2\le 1/8, \bar{\mu}_3=0\}$. This solution satisfies MFCQ (but not LICQ), and SSOSC holds with all $\bar{\mu}\in\mathcal{M}(\bar{x})$.

For this example, SQP has superlinear primal convergence ($I_k=\{2\}$ for large k). SNM-FB and SNM-NR either have superlinear primal convergence, or fail (for SNM-NR, Λ_k often becomes almost singular, and failures are typical).

Example 4.3 ([15, Problem 43], [9, Example 3], [38, Example 3]) The inequality-constrained problem

$$\begin{array}{ll} \text{minimize} & x_1^2+x_2^2+2x_3^2+x_4^2-5x_1-5x_2-21x_3+7x_4 \\ \text{subject to} & x_1^2+x_2^2+x_3^2+x_4^2+x_1-x_2+x_3-x_4\le 8, \\ & x_1^2+2x_2^2+x_3^2+2x_4^2-x_1-x_4\le 10, \\ & 2x_1^2+x_2^2+x_3^2+2x_1-x_2-x_4\le 5, \\ & -x_2^3-2x_1^2-x_4^2-x_1+3x_2+x_3-4x_4\le 7 \end{array}$$

has the unique solution $\bar{x}=(0,1,2,-1)$, with $\mathcal{M}(\bar{x})=\{\bar{\mu}\in\mathbf{R}^4 \mid \bar{\mu}_1=3-\bar{\mu}_3, \bar{\mu}_2=0, 2\le\bar{\mu}_3\le 3, \bar{\mu}_4=\bar{\mu}_3-2\}$. This solution satisfies MFCQ (but not LICQ), and SSOSC holds with all $\bar{\mu}\in\mathcal{M}(\bar{x})$.

SQP either converge superlinearly for this example (with $I_k=\{1,3\}$ for large k), or fails because of unbounded subproblems. SNM-FB usually fails, though sometimes converges superlinearly. SNM-NR fails.

Example 4.4 ([32], [2, p. 9]) The equality-constrained problem

$$\begin{array}{ll} \text{minimize} & x^2 \\ \text{subject to} & F(x)=0, \end{array}$$

where

$$F(x)=\begin{cases} x^6\sin(1/x) & \text{if } x\in\mathbf{R}\setminus\{0\}, \\ 0 & \text{if } x=0, \end{cases}$$

has the unique global solution $\bar{x}=0$ and infinitely many local solutions of the form $1/(\pi s)$, where s runs over integers. In particular, $\bar{x}$ is a nonisolated local solution, $\mathcal{M}(\bar{x})=\mathbf{R}$, and this solution violates the regularity condition (2.2) but satisfies SOSC with all $\bar{\lambda}\in\mathbf{R}$.

For this example, the Newton–Lagrange method converges superlinearly to local minimizers distinct from $\bar{x}$ (that is why this example is not reported in Table 1).

Example 4.5 ([13, (4)]) The inequality-constrained problem

$$\begin{aligned} &\text{minimize} && x^2 \\ &\text{subject to} && x^2 \leq 0, \end{aligned}$$

has the unique feasible point (hence, unique solution) $\bar{x} = 0$, and $\mathcal{M}(\bar{x}) = \mathbf{R}_+$. This solution violates MFCQ but satisfies SSOSC with all $\bar{\mu} \in \mathcal{M}(\bar{x})$.

SQP and SNM-FB converge superlinearly for this example (SQP actually has finite termination). For SNM-NR, partition (J_1^k, J_2^k) does not stabilize, the dual trajectory does not converge (has two distinct accumulation points $-1/2$ and 0), and the primal trajectory converges slowly.

Example 4.6 ([1]) The problem with mixed constraints

$$\begin{aligned} &\text{minimize} && x_1^2 \\ &\text{subject to} && x_2 = 0, \qquad x_2 \leq 0, \end{aligned}$$

has the unique solution $\bar{x} = 0$, with $\mathcal{M}(\bar{x}) = \{(\bar{\lambda}, \bar{\mu}) \in \mathbf{R} \times \mathbf{R}_+ \mid \bar{\lambda} + \bar{\mu} = 0\}$. This solution violates MFCQ but satisfies SSOSC for all $(\bar{\lambda}, \bar{\mu}) \in \mathcal{M}(\bar{x})$.

SQP and SNM-FB have finite (one-step) termination for this example. SNM-NR usually fails (because Λ_k becomes almost singular) but sometimes also shows one-step termination.

Example 4.7 (model MPCC) MPCC

$$\begin{aligned} &\text{minimize} && x_1 + 2x_2 \\ &\text{subject to} && x_1 \geq 0, \qquad x_2 \geq 0, \qquad x_1 x_2 \leq 0, \end{aligned}$$

has the unique solution $\bar{x} = 0$, with $\mathcal{M}(\bar{x}) = \{\bar{\mu} \in \mathbf{R}^3 \mid \bar{\mu}_1 = 1, \bar{\mu}_2 = 2, \bar{\mu}_3 \geq 0\}$. This solution satisfies MPCC-LICQ, and MPCC-SOSC holds (trivially, because $C(\bar{x}) = \{0\}$).

SQP has finite termination for this example. SNM-FB either converges superlinearly or fails. SNM-NR usually fails (because Λ_k becomes almost singular) but sometimes shows one-step termination.

Example 4.8 ([16]) MPCC

$$\begin{aligned} &\text{minimize} && \frac{1}{2}(x_1 - 1)^2 + \frac{1}{2}(x_2 - 2)^2 \\ &\text{subject to} && x_1 \geq 0, \qquad x_2 \geq 0, \qquad x_1 x_2 \leq 0, \end{aligned}$$

has two local solutions $\bar{x}^1 = (1, 0)$ and $\bar{x}^2 = (0, 2)$, with $\mathcal{M}(\bar{x}^1) = \{\bar{\mu} \in \mathbf{R}^3 \mid \bar{\mu}_1 = 0, \bar{\mu}_3 - 2 = \bar{\mu}_2 \geq 0\}$, $\mathcal{M}(\bar{x}^2) = \{\bar{\mu} \in \mathbf{R}^3 \mid 2\bar{\mu}_3 - 1 = \bar{\mu}_1 \geq 0, \bar{\mu}_2 = 0\}$, respectively. Both solutions satisfy MPCC-LICQ and MPCC-SOSC.

For this example, SQP has finite termination, while SNM-FB converges superlinearly. SNM-NR either has finite termination or fails (Λ_k becomes almost singular).

Example 4.9 ([4, Example 1]) MPCC

$$\begin{array}{ll} \text{minimize} & \frac{1}{2}(x_2-1)^2 \\ \text{subject to} & x_1 \geq 0, \qquad x_2 \geq 0, \qquad x_1x_2 \leq 0, \end{array}$$

has the unique solution $\bar{x} = (0, 1)$, with $\mathcal{M}(\bar{x}) = \{\bar{\mu} \in \mathbf{R}^3 \mid \bar{\mu}_1 = \bar{\mu}_3 \geq 0, \bar{\mu}_2 = 0\}$. This solution satisfies MPCC-LICQ and MPCC-SOSC.

For this example, SQP has finite termination. SNM-FB either converges superlinearly or fails. SNM-NR usually fails (Λ_k becomes almost singular), but sometimes has finite termination.

Example 4.10 ([4, Example 2]) MPCC

$$\begin{array}{ll} \text{minimize} & -x_1 + \int_0^{x_2} t^6 \sin(1/t)dt \\ \text{subject to} & x_1 \geq 0, \qquad x_2 \geq 0, \qquad x_1x_2 \leq 0, \end{array}$$

has the sequence of local solutions of the form $(0, 1/(\pi + 2\pi j))$, $j = 0, 1, \ldots$. All those solutions satisfy MPCC-LICQ and MPCC-SOSC. In our numerical experiments, we used $\bar{x} = 1/\pi$.

For this example, SQP and SNM-FB either converge superlinearly to $\bar{x}$ (for SQP, $I_k = \{1, 3\}$ for large k) or fail. Fails of SQP are quite frequent, due to unbounded subproblems. SNM-NR fails too (Λ_k becomes almost singular).

Example 4.11 ([26, Example 1 and p. 15], `jr1` in MacMPEC [27], [12, Sect. 2.1]) MPCC

$$\begin{array}{ll} \text{minimize} & (x_1-1)^2 + x_2^2 \\ \text{subject to} & x_2 \geq 0, \qquad x_2 - x_1 \geq 0, \qquad x_2(x_2 - x_1) \leq 0, \end{array}$$

has the unique solution $\bar{x} = (1/2, 1/2)$, with $\mathcal{M}(\bar{x}) = \{\bar{\mu} \in \mathbf{R}^3 \mid \bar{\mu}_1 = 0, \bar{\mu}_2 \geq 1, \bar{\mu}_3 = 2(\bar{\mu}_2 - 1)\}$. This solution satisfies MPCC-LICQ and MPCC-SOSC.

SQP and SNM-FB converge superlinearly for this example (SQP actually has finite termination, while SNM-FB sometimes fails). SNM-NR fails, apart from some rare cases of "accidental" finite termination.

Example 4.12 ([26, Example 1 and p. 15], `jr2` in MacMPEC [27], [12, Sect. 2.1]) MPCC

$$\begin{array}{ll} \text{minimize} & x_1^2 + (x_2-1)^2 \\ \text{subject to} & x_2 \geq 0, \qquad x_2 - x_1 \geq 0, \qquad x_2(x_2 - x_1) \leq 0, \end{array}$$

has the unique solution $\bar{x} = (1/2, 1/2)$, with $\mathcal{M}(\bar{x}) = \{\bar{\mu} \in \mathbf{R}^3 \mid \bar{\mu}_1 = 0, \bar{\mu}_2 \geq 0, \bar{\mu}_3 = 2(\bar{\mu}_2 + 1)\}$. This solution satisfies MPCC-LICQ and MPCC-SOSC.

SQP and SNM-FB converge superlinearly for this example (SQP actually has finite termination, while SNM-FB sometimes fails). SNM-NR fails, apart from some rare cases of "accidental" finite termination.

Example 4.13 ([39, (6.6)]) MPCC

$$\begin{aligned} &\text{minimize} \quad x_2 \\ &\text{subject to} \quad -x_2^2 \le -1, \qquad x_1 \ge 0, \qquad x_2 \ge 0, \qquad x_1x_2 \le 0, \end{aligned}$$

has the unique solution $\bar{x} = (0, 1)$, with $\mathcal{M}(\bar{x}) = \{\bar{\mu} \in \mathbf{R}^4 \mid \bar{\mu}_1 = 1/2, \bar{\mu}_2 = \bar{\mu}_4 \ge 0, \bar{\mu}_3 = 0\}$. This solution satisfies MPCC-LICQ and MPCC-SOSC (trivially, because $C(\bar{x}) = \{0\}$).

For this example, SQP either converges superlinearly (with $I_k = \{1, 4\}$ for large k) or fails because of inconsistent subproblems. SNM-FB either converges superlinearly or fails. In some cases, SNM-FB converges slowly by gradient steps. SNM-NR fails (Λ_k becomes almost singular).

Example 4.14 (`s14` in MacMPEC [27], [12, (2.3)]) MPCC

$$\begin{aligned} &\text{minimize} \quad x_1 + x_2 \\ &\text{subject to} \quad -x_2^2 \le -1, \qquad x_1 \ge 0, \qquad x_2 \ge 0, \qquad x_1x_2 \le 0, \end{aligned}$$

has the unique solution $\bar{x} = (0, 1)$, with $\mathcal{M}(\bar{x}) = \{\bar{\mu} \in \mathbf{R}^4 \mid \bar{\mu}_1 = 0, \bar{\mu}_2 - \bar{\mu}_4 = 1, \bar{\mu}_3 = 0, \bar{\mu}_4 \ge 0\}$. This solution satisfies MPCC-LICQ and MPCC-SOSC (trivially, because $C(\bar{x}) = \{0\}$).

For this example, SQP either converges superlinearly (with $I_k = \{1, 2, 4\}$ for large k) or fails because of inconsistent subproblems. SNM-FB either converges superlinearly or fails. SNM-NR fails (Λ_k becomes almost singular).

Example 4.15 (`s12` in MacMPEC [27], [12, Sect. 7.2)]) MPCC

$$\begin{aligned} &\text{minimize} \quad -x_1 - \frac{1}{2}x_2 \\ &\text{subject to} \quad x_1 + x_2 \le 2, \qquad x_1^2 - x_1 \ge 0, \qquad x_2 \ge 0, \qquad (x_1^2 - x_1)x_2 \le 0, \end{aligned}$$

has the unique solution $\bar{x} = (0, 2)$, with $\mathcal{M}(\bar{x}) = \{\bar{\mu} \in \mathbf{R}^4 \mid \bar{\mu}_1 = 1/2, 2\bar{\mu}_2 = 4\bar{\mu}_4 + 1, \bar{\mu}_3 = 0, \bar{\mu}_4 \ge 0\}$. This solution satisfies MPCC-LICQ and MPCC-SOSC (trivially, because $C(\bar{x}) = \{0\}$).

For this example, SQP either converges superlinearly (with $I_k = \{1, 2\}$ for large k) or fails because of inconsistent subproblems. SNM-FB sometimes converges superlinearly but usually fails. SNM-NR fails (Λ_k becomes almost singular).

We continue with examples (Examples 4.16–4.22) where critical multipliers do exist; they complement Examples 2.1–2.4, 3.2–3.5. These examples put in evidence that attraction of SQP iterates to critical multipliers is typical and, as a consequence, convergence is slow. Numerical results for these examples are reported in Table 2.

Table 1 Examples with no critical multipliers: the number of times convergence was declared for 10 different starting points, and the number of times it was superlinear

Examples	Algorithm		
	SQP	SNM-FB	SNM-NR
3.1	8(8)	8(8)	6(0)
4.1	10(10)	10(10)	4(0)
4.2	10(10)	6(6)	2(2)
4.3	5(5)	2(2)	0(0)
4.5	10(10)	10(10)	10(0)
4.6	10(10)	10(10)	3(3)
4.7	10(10)	4(4)	1(1)
4.8	10(10)	10(10)	4(4)
4.9	8(8)	6(6)	2(2)
4.10	3(3)	6(6)	0(0)
4.11	10(10)	8(8)	2(2)
4.12	10(10)	6(6)	2(2)
4.13	4(4)	4(3)	0(0)
4.14	7(7)	4(4)	0(0)
4.15	4(4)	2(2)	0(0)

Table 2 Examples where there exist critical multipliers: number of times convergence was declared for 10 different starting points, number of times it was superlinear, and number of times dual convergence to a critical multiplier was detected

Examples	Algorithm		
	SQP	SNM-FB	SNM-NR
2.1	10(0/10)	10(0/10)	10(0/10)
2.2	10(0/10)	7(0/7)	10(0/10)
2.3	10(0/10)	7(0/7)	10(0/10)
2.4	10(0/10)	7(0/7)	10(0/10)
3.2	10(0/10)	10(0/10)	10(0/10)
3.3	6(0/6)	9(2/7)	6(3/6)
3.4	9(2/9)	8(3/5)	5(0/5)
3.5	10(8/2)	5(2/3)	6(5/1)
4.16	0(0/0)	10(2/8)	1(1/0)
4.17	0(0/0)	8(0/8)	3(3/0)
4.18	10(0/10)	10(0/10)	10(0/10)
4.19	8(0/8)	10(0/10)	10(0/10)
4.20	10(0/10)	9(0/9)	10(0/10)
4.21	8(0/8)	7(0/7)	9(1/8)
4.22	7(5/2)	10(4/5)	3(0/1)

In addition to the information reported in Table 1, we now report on the number of runs for which convergence to a critical multiplier was detected (after slash in parentheses). Note that for equality-constrained problems (those in Examples 2.1–2.4), the three algorithms differ only in linesearch rules.

Example 4.16 ([5, p. 1350]) The inequality-constrained problem

$$\begin{array}{ll} \text{minimize} & x_3 \\ \text{subject to} & \langle Q_i(x_1, x_2), (x_1, x_2)\rangle - x_3 \leq 0, \quad i = 1, 2, 3, 4, \end{array}$$

where

$$Q_i = U_i^{\mathrm{T}} Q U_i, \quad U_i = \begin{pmatrix} \cos[\pi(i-1)/4] & \sin[\pi(i-1)/4] \\ -\sin[\pi(i-1)/4] & \cos[\pi(i-1)/4] \end{pmatrix}, \; i = 1, 2, 3, 4,$$

$$Q = \begin{pmatrix} 1 & 0 \\ 0 & -2 \end{pmatrix},$$

has the unique solution $\bar{x} = 0$, with $\mathcal{M}(\bar{x}) = \{\bar{\mu} \in \mathbf{R}^4_+ \mid \sum_{i=1}^4 \bar{\mu}_i = 1\}$. This solution satisfies MFCQ (but not LICQ), and violates SOSC for any $\bar{\mu} \in \mathcal{M}(\bar{x})$. Critical multipliers for $I = \{1, 2, 3, 4\}$ are those satisfying the equality $(2\bar{\mu}_1 - \bar{\mu}_2 - 4\bar{\mu}_3 - \bar{\mu}_4)(-4\bar{\mu}_1 - \bar{\mu}_2 + 2\bar{\mu}_3 - \bar{\mu}_4) - 9(\bar{\mu}_2 - \bar{\mu}_4)^2 = 0$.

For this example, SQP usually fails (because of unbounded subproblems), and SNM-NR also (Λ_k becomes almost singular), apart from some rare cases of finite termination. SNM-FB usually converges slowly, with dual trajectory converging to some critical multiplier (however, in some cases, superlinear convergence to a non-critical multiplier was observed).

Example 4.17 ([7, Example 6.1]) The inequality-constrained problem

$$\begin{array}{ll} \text{minimize} & x_3 \\ \text{subject to} & \langle Q_i(x_1, x_2), (x_1, x_2)\rangle - x_3 \leq 0, \quad i = 1, 2, 3, \end{array}$$

where

$$Q_1 = \begin{pmatrix} 0 & \sqrt{3} \\ \sqrt{3} & -2 \end{pmatrix}, \qquad Q_2 = \begin{pmatrix} 0 & -\sqrt{3} \\ -\sqrt{3} & -2 \end{pmatrix}, \qquad Q_3 = \begin{pmatrix} -3 & 0 \\ 0 & 1 \end{pmatrix},$$

has the unique solution $\bar{x} = 0$, with $\mathcal{M}(\bar{x}) = \{\bar{\mu} \in \mathbf{R}^3_+ \mid \sum_{i=1}^3 \bar{\mu}_i = 1\}$. This solution satisfies MFCQ (but not LICQ) and violates SOSC for any $\bar{\mu} \in \mathcal{M}(\bar{x})$. Critical multipliers for $I = \{1, 2, 3\}$ are those satisfying the equality $2(\bar{\mu}_1\bar{\mu}_2 + \bar{\mu}_1\bar{\mu}_3 + \bar{\mu}_2\bar{\mu}_3) - \bar{\mu}_1^2 - \bar{\mu}_2^2 - \bar{\mu}_3^2 = 0$.

The algorithms behave very similarly to what has been observed in Example 4.16.

Example 4.18 ([24, Example 2.1], [19, Example 2], [25, Example 1]) The equality-constrained problem

$$\begin{array}{ll} \text{minimize} & x^2 \\ \text{subject to} & x^2 = 0, \end{array}$$

has the unique feasible point (hence, unique solution) $\bar{x} = 0$, with $\mathcal{M}(\bar{x}) = \mathbf{R}$. This solution violates the regularity condition (2.2) but satisfies SOSC with all $\bar{\lambda} > -1$. The unique critical multiplier is $\bar{\lambda} = -1$.

For this example, all versions of the Newton–Lagrange method converge slowly; dual trajectory converges to the unique critical multiplier.

Example 4.19 The inequality-constrained problem

$$\begin{aligned} &\text{minimize} && x^3 \\ &\text{subject to} && x^2 \le 0, \end{aligned}$$

has the unique feasible point (hence, unique solution) $\bar{x} = 0$, with $\mathcal{M}(\bar{x}) = \mathbf{R}_+$. This solution violates MFCQ but satisfies SSOSC with all $\bar{\mu} \in \mathcal{M}(\bar{x})$ except for $\bar{\mu} = 0$, which is the unique critical multiplier.

For this example, SQP, SNM-NR and SNM-FB all converge slowly; dual trajectory converges to the unique critical multiplier. SQP sometimes fails because of unbounded subproblems.

Example 4.20 ([17], [19, Example 2], [25, Example 3]) The equality-constrained problem

$$\begin{aligned} &\text{minimize} && x_1^2 \\ &\text{subject to} && x_1^2 - x_2^2 = 0, \end{aligned}$$

has the unique solution $\bar{x} = 0$, with $\mathcal{M}(\bar{x}) = \mathbf{R}$. This solution violates the regularity condition (2.2) but satisfies SOSC with all $\bar{\lambda} \in (-1, 0)$. Critical multipliers are $\bar{\lambda} = -1$ and $\bar{\lambda} = 0$.

For this example, all versions of the Newton–Lagrange method converges slowly; dual trajectory converges to some critical multiplier.

Example 4.21 The equality-constrained problem

$$\begin{aligned} &\text{minimize} && x_1^2 + x_2^2 + x_3^2 \\ &\text{subject to} && \sin x_1 + \sin x_2 + \sin x_3 = 0, \qquad x_1 + x_2 + x_3 + x_1^2 + \sin x_1 x_3 = 0, \end{aligned}$$

has the unique solution $\bar{x} = 0$, with $\mathcal{M}(\bar{x}) = \{\bar{\lambda} \in \mathbf{R}^2 \mid \bar{\lambda}_1 + \bar{\lambda}_2 = 0\}$. This solution violates the regularity condition (2.2) but satisfies SOSC with all $\bar{\lambda} \in \mathcal{M}(\bar{x})$ such that $\bar{\lambda}_1 \in (-2, 6)$. Critical multipliers are $\bar{\lambda} = (-2, 2)$ and $\bar{\lambda} = (6, -6)$.

This is the only example where we had to reduce the region of distribution for random primal starting points, in order to avoid attraction to a different (nondegenerate) stationary point. Specifically, we used starting points x^0 with components x_j^0 in $[\bar{x}_j - 1, \bar{x}_j + 1]$, $j = 1, \ldots, n$. With this choice, all versions of the Newton–Lagrange method usually converge slowly to a critical multiplier.

Example 4.22 ([24, Example 4.2]) MPCC

$$\begin{aligned} &\text{minimize} && x_1 + x_2 + x_1^2 + x_2^2 + (x_3 - 1)^2 \\ &\text{subject to} && x_1 + x_2 \ge 0, \qquad x_1^2 - (x_3 - 1)^2 \ge 0, \qquad x_2 \ge 0, \qquad x_3 \ge 0, \\ &&& (x_1 + x_2)x_2 + (x_1^2 - (x_3 - 1)^2)x_3 \le 0, \end{aligned}$$

has solution $\bar{x} = (0, 0, 1)$, with $\mathcal{M}(\bar{x}) = \{\bar{\mu} \in \mathbf{R}^4 \mid \bar{\mu}_1 = 1, \bar{\mu}_2 \geq 0, \bar{\mu}_3 = \bar{\mu}_4 \geq 0, \bar{\mu}_5 \geq 0\}$. This solution violates MPCC-LICQ but satisfies MPCC-SOSC with some MPCC-multipliers. Critical multipliers for $I = \{1, 2, 3, 5\}$ are those satisfying the equality $\bar{\mu}_5 = 1 + \bar{\mu}_2$.

For this example, the methods demonstrate various kinds of behaviour depending on a starting point. SQP either has finite termination, or converges slowly to a critical multiplier, or fails (for various reasons). SNM-FB either slowly converges to a critical multiplier, or superlinearly converges to a noncritical one, or even sometimes linearly converges to a noncritical multiplier. SNM-NR either converges slowly to a critical multiplier, or converges slowly because partition (J_1^k, J_2^k) does not stabilize, but most times it fails.

We complete this collection with two examples where Lagrange multipliers do not exist.

Example 4.23 ([3, (2.15)]) MPCC

$$\begin{array}{ll} \text{minimize} & x_2 - x_1 \\ \text{subject to} & x_1 \leq 0, \qquad x_2 \leq 0, \qquad x_1 + x_2 \leq 0, \qquad x_2(x_1 + x_2) \leq 0, \end{array}$$

has solution $\bar{x} = 0$, which is not a strongly stationary point.

For this example, SQP either has finite termination or converges slowly, while $\mu_4^k \to +\infty$. SNM-FB and SNM-NR usually fail.

Example 4.24 (`scholtes4` in MacMPEC [27], [12, Sect. 7.1]) MPCC

$$\begin{array}{ll} \text{minimize} & -x_1 + x_2 + x_3 \\ \text{subject to} & x_1 - 4x_2 \leq 0, \qquad x_1 - 4x_3 \leq 0, \qquad x_2 \geq 0, \qquad x_3 \geq 0, \qquad x_1 x_3 \leq 0, \end{array}$$

has the unique solution $\bar{x} = 0$, which is not a strongly stationary point.

For this example, SQP usually has finite termination. SNM-FB fails. SNM-NR fails to make a step.

References

1. Andreani, R., Martínez, J.M.: On the solution of mathematical programming problems with equilibrium constraints. Math. Methods Oper. Res. **54**, 345–358 (2001)
2. Anitescu, M.: Nonlinear programs with unbounded Lagrange multiplier sets. Preprint ANL/MCS-P796-0200, Mathematics and Computer Science Division, Argonne National Laboratory, Argonne, IL (2000)
3. Anitescu, M.: On using the elastic mode in nonlinear programming approaches to mathematical programs with complementarity constraints. SIAM J. Optim. **15**, 1203–1236 (2005)
4. Anitescu, M., Tseng, P., Wright, S.J.: Elastic-mode algorithms for mathematical programs with equilibrium constraints: global convergence and stationarity properties. Math. Program. **110**, 337–371 (2007)
5. Arutyunov, A.V.: Perturbations of extremum problems with constraints and necessary optimality conditions. J. Sov. Math. **54**, 1342–1400 (1991)
6. Arutyunov, A.V.: Optimality Conditions: Abnormal and Degenerate Problems. Kluwer Academic, Dordrecht (2000)

7. Baccari, A., Trad, A.: On the classical necessary second-order optimality conditions in the presence of equality and inequality constraints. SIAM J. Optim. **15**, 394–408 (2004)
8. De Luca, T., Facchinei, F., Kanzow, C.: A theoretical and numerical comparison of some semismooth algorithms for complementarity problems. Comput. Optim. Appl. **16**, 173–205 (2000)
9. Facchinei, F., Fischer, A., Kanzow, C.: On the accurate identification of active constraints. SIAM J. Optim. **9**, 14–32 (1999)
10. Fischer, A.: Modified Wilson's method for nonlinear programs with nonunique multipliers. Math. Oper. Res. **24**, 699–727 (1999)
11. Fischer, A.: Local behaviour of an iterative framework for generalized equations with nonisolated solutions. Math. Program. **94**, 91–124 (2002)
12. Fletcher, R., Leyffer, S., Ralph, D., Scholtes, S.: Local convergence of SQP methods for mathematical programs with equilibrium constraints. SIAM J. Optim. **17**, 259–286 (2006)
13. Hager, W.W.: Stabilized sequential quadratic programming. Comput. Optim. Appl. **12**, 253–273 (1999)
14. Hager, W.W., Gowda, M.S.: Stability in the presence of degeneracy and error estimation. Math. Program. **85**, 181–192 (1999)
15. Hock, W., Schittkowski, K.: Test Examples for Nonlinear Programming Codes. Lect. Notes in Econom. and Math. Systems, vol. 187. Springer, Berlin (1981)
16. Hu, X.M., Ralph, D.: Convergence of a penalty method for mathematical problems with complementarity constraints. J. Optim. Theory Appl. **123**, 365–390 (2004)
17. Izmailov, A.F.: Lagrange methods for finding degenerate solutions of conditional extremum problems. Comput. Math. Math. Phys. **36**, 423–429 (1996)
18. Izmailov, A.F.: Mathematical programs with complementarity constraints: regularity, optimality conditions, and sensitivity. Comput. Math. Math. Phys. **44**, 1145–1164 (2004)
19. Izmailov, A.F.: On the analytical and numerical stability of critical Lagrange multipliers. Comput. Math. Math. Phys. **45**, 930–946 (2005)
20. Izmailov, A.F., Solodov, M.V.: Optimality conditions for irregular inequality-constrained problems. SIAM J. Control Optim. **40**, 1280–1295 (2001)
21. Izmailov, A.F., Solodov, M.V.: The theory of 2-regularity for mappings with Lipschitzian derivatives and its applications to optimality conditions. Math. Oper. Res. **27**, 614–635 (2002)
22. Izmailov, A.F., Solodov, M.V.: Complementarity constraint qualification via the theory of 2-regularity. SIAM J. Optim. **13**, 368–385 (2002)
23. Izmailov, A.F., Solodov, M.V.: Karush-Kuhn-Tucker systems: regularity conditions, error bounds and a class of Newton-type methods. Math. Program. **95**, 631–650 (2003)
24. Izmailov, A.F., Solodov, M.V.: Newton-type methods for optimization problems without constraint qualifications. SIAM J. Optim. **15**, 210–228 (2004)
25. Izmailov, A.F., Solodov, M.V.: On attraction of Newton-type iterates to multipliers violating second-order sufficiency conditions. Math. Program., DOI 10.1007/s10107-007-0158-9. Available at http://www.impa.br~optim/solodov.html
26. Jiang, H., Ralph, D.: QPECgen, a MATLAB generator for mathematical programs with quadratic objectives and affine variational inequality constraints. Techn. Rept., Univ. Melbourne, Dept. Math. (1997)
27. Leyffer, S.: MacMPEC: AMPL collection of MPECs. http://www-unix.mcs.anl.gov/leyffer/MacMPEC/
28. Li, D.-H., Qi, L.: A stabilized SQP method via linear equations. Applied Mathematics Technical Report AMR00/5, The University of New South Wales (2000)
29. Luo, Z.-Q., Pang, J.-S., Ralph, D.: Mathematical Programs with Equilibrium Constraints. Cambridge University Press, Cambridge (1996)
30. Qi, L.: Convergence analysis of some algorithms for solving nonsmooth equations. Math. Oper. Res. **18**, 227–244 (1993)
31. Qi, L., Sun, J.: A nonsmooth version of Newton's method. Math. Program. **58**, 353–3674 (1993)
32. Robinson, S.M.: Generalized equations and their solutions, Part II: applications to nonlinear programming. Math. Program. Study **19**, 200–221 (1982)
33. Scheel, H., Scholtes, S.: Mathematical programs with complementarity constraints: stationarity, optimality and sensitivity. Math. Oper. Res. **25**, 1–22 (2000)
34. Scholtes, S., Stöhr, M.: Exact penalization of mathematical programs with equilibrium constraints. SIAM J. Control Optim. **37**, 617–652 (1999)
35. Scholtes, S., Stöhr, M.: How stringent is the linear independence assumption for mathematical programs with complementarity constraints?. Math. Oper. Res. **26**, 851–863 (2001)

36. Wright, S.J.: Superlinear convergence of a stabilized SQP method to a degenerate solution. Comput. Optim. Appl. **11**, 253–275 (1998)
37. Wright, S.J.: Modifying SQP for degenerate problems. SIAM J. Optim. **13**, 470–497 (2002)
38. Wright, S.J.: Constraint identification and algorithm stabilization for degenerate nonlinear programs. Math. Program. **95**, 137–160 (2003)
39. Wright, S.J.: An algorithm for degenerate nonlinear programming with rapid local convergence. SIAM J. Optim. **15**, 673–696 (2005)