DISCUSSION

Comments on: Critical Lagrange multipliers: what we currently know about them, how they spoil our lives, and what we can do about it

Boris S. Mordukhovich

Abstract In these comments on the excellent survey paper by Izmailov and Solodov, we briefly discuss the main issues of critical Lagrange multipliers discovered and analyzed by the authors and then formulate some topics of interest for the future research. The latter topics concern the study of a possible influence of critical Lagrange multipliers on the convergence of primal–dual numerical algorithms of finding not general global and local minimizers but those satisfying certain desired stability properties known as tilt and full stability, which are briefly surveyed in these comments.

1 Critical Lagrange multipliers for general minimizers

This paper is devoted to considering the classical problems of nonlinear programming (NLP) with twice continuously differentiable ($\mathcal{C}^2$-smooth) data. Although the authors confine their attention to problems with only equality constraints, the phenomena they discuss occur also in the case of NLP problems with inequalities. This fact mentioned in the Introduction is worth to be emphasized for my comments below.

The striking phenomena of numerical analysis, discovered and strongly investigated by the authors during recent years, concern NLP and other classes of variational problems admitting *critical multipliers* (dual variables), which constitute the main topic of the discussion paper. Simple examples reveal that such multipliers may appear even in

---

B. S. Mordukhovich (✉)
Department of Mathematics, Wayne State University, Detroit, MI 48202, USA
e-mail: boris@math.wayne.edu

the case of unique Lagrange multipliers. This happens, in particular, under the classical *linear independence constraint qualification* (LICQ) that reduces to the regularity condition (1.3) of the discussion paper for problems with only equality constraints. As a rule, critical multipliers, defined via the reduced Hessian of the corresponding Lagrangian, are nonempty for optimal solutions with nonunique Lagrange multipliers and are largely responsible, as shown by the authors and their collaborators, for slow (not superlinear) convergence of major primal–dual numerical algorithms including Newton and Newton-related methods, the augmented Lagrangian method, and the sequential quadratic programming method.

On the other hand, the set of critical multipliers is empty under the validity of LICQ and the so-called strong second-order sufficient condition (SSOSC) by Robinson (1980), which is condition (2.1) of the discussion paper for equality constraints. It is well known that the simultaneous fulfillment of LICQ and SSOSC at the given optimal solution is *necessary and sufficient* for Robinson (1980) *strong regularity* of the associated Karush–Kuhn–Tucker (KKT) system; see Dontchev and Rockafellar (1996). The latter property has been highly recognized in the study of both qualitative and numerical aspects of optimization including applications to primal–dual algorithms. It is worth mentioning that, in the case of nonunique Lagrange multipliers, SSOSC can hold for one multiplier while violate for another one, which may be critical.

The authors clearly and nicely demonstrate, by both theoretical results and numerous examples, that critical multipliers, which are defined by an algebraic equation and hence constitute just a "thin" subset of Lagrange multipliers, play a crucial *negative* role in numerical optimization yielding slow convergence of major primal–dual algorithms. This strong message is a surprising discovery (at least for me, and surely for many other experts in optimization), which calls for improvements of computational algorithms to avoid the convergence to critical multipliers. Some of such "dual stabilization techniques" are presented in the discussion paper, mainly from the viewpoint of local convergence to optimal solutions, while global aspects of these techniques are also discussed with formulating important open questions.

I am highly impressed by both "negative" and "positive" results of this fundamental direction in optimization theory and applications pioneered and largely developed by the authors. In my subsequent comments, I would like to discuss some further possible developments in this direction partly close to my own recent research.

2 Critical Lagrange multipliers for tilt-stable minimizers

My major suggestion is to explore the possibility of arising such an "attraction phenomenon" of convergence to critical minimizers not for the case of *general* global or local minimizers but for *special ones* having some desired *stability properties* with respect to parameter perturbations. To the best of my knowledge, the first notion of this type was introduced and investigated by Poliquin and Rockafellar (1998) under the name of "tilt-stable" local minimizers in a very general framework. The motivation of these authors was expressed as follows: "What is the best paradigm for modern purposes? Optimization no longer revolves around making lists of solution candidates

to be checked out one by one, if it ever did. The role of optimality conditions is seen rather in the justification of numerical algorithms, in particular their stopping criteria, convergence properties, and robustness. From this angle, the goal of theory could be different. Instead of focusing on the threshold between necessity and sufficiency, one might more profitably try to characterize the strongest manifestations of optimality that support computational work. Indeed, this idea has motivated much of the effort that has gone into parametric optimization—the study of how local solutions to a problem may react to shift in data."

Largely motivated by these purposes, Poliquin and Rockafellar (1998) defined the notion of tilt stability of local minimizers for extended real-valued functions $f : \mathbb{R}^n \to \bar{\mathbb{R}} := (-\infty, \infty]$ in the following way: $\bar{x}$ is is *tilt-stable* local minimizer of f if $f(\bar{x}) < \infty$ and there exists a constant $\gamma > 0$ such that the mapping

$$M_\gamma : v \mapsto \text{argmin}\big\{ f(x) - \langle v, x \rangle \big| \, \|x - \bar{x}\| \leq \gamma \big\}$$

is single-valued and Lipschitzian on some neighborhood of $v = 0$ with $M_\gamma(0) = \bar{x}$.

The meaning of this definition is that the small linear perturbation $\langle v, x \rangle$, which *tilts* the objective f in one direction or another, should not affect the local solution disproportionately to the size of v or threaten its uniqueness.

The first significant observation made by Poliquin and Rockafellar (1998) is that for $\mathcal{C}^2$-smooth functions $f : \mathbb{R}^n \to \mathbb{R}$ having $\nabla f(\bar{x}) = 0$, the point $\bar{x}$ is a tilt-stable local minimizer of f *if and only if* the Hessian matrix $\nabla^2 f(\bar{x})$ is positive definite. This illustrates the striking difference between tilt-stable and general local minimizers in the very classical setting. Indeed, it is well known that the positive definiteness of $\nabla^2 f(\bar{x})$ is a sufficient condition for general local minimizers of f while the Hessian positive semidefiniteness (not positive definiteness) is a necessary condition for them.

In the same paper, the authors generalize this result to a broad class of extended real-valued functions f, which are both *prox-regular* and *subdifferentially continuous* at $\bar{x}$ for $\bar{v} = 0 \in \partial f(\bar{x})$; see Rockafellar and Wets (1998) for more details about this remarkable class containing the vast majority of functions typically encountered in finite-dimensional optimization and second-order variational analysis. The main result by Poliquin and Rockafellar (1998) gives a complete characterization of tilt-stable minimizers for the aforementioned class of extended real-valued functions via the *generalized Hessian*, or the *second-order subdifferential*, $\partial^2 f(\bar{x}, \bar{v})$ of f at $\bar{x}$ with respect to $\bar{v} \in \partial f(\bar{x})$ introduced by Mordukhovich (1992). This characterization is formulated as follows: $\bar{x}$ is a tilt-stable minimizer of f with $0 \in \partial f(\bar{x})$ *if and only if* the set-valued mapping $\partial^2 f(\bar{x}, 0) : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is positive definite in the sense that

$$\langle z, w \rangle > 0 \quad \text{whenever} \quad z \in \partial^2 f(\bar{x}, 0)(w), \; w \neq 0.$$

While the extended real-valued framework of minimizing $f : \mathbb{R}^n \to \bar{\mathbb{R}}$ implicitly incorporates constraints via $x \in \text{dom} f$, the passage to characterizations of tilt-stable minimizers for specific classes of constrained optimization problems (e.g., for classical NLPs with $\mathcal{C}^2$-smooth data considered in the discussion paper by Izmailov and Solodov) is not an easy job. It requires at least *two major components*: to develop *second-order subdifferential calculus* with equalities (sum and chain rules, etc.) for

the generalized Hessians and *explicit calculations* of them for structural functions arising in specific settings of constrained optimization. Some results in these directions can be found in the books by Mordukhovich (2006a, b), but they are not sufficient for applications to tilt stability. Second-order chain rules for several subclasses of *amenable compositions*, more appropriate for these purposes, were derived by Mordukhovich and Rockafellar (2012) and applied therein to characterizing tilt-stable minimizers for NLPs with $\mathcal{C}^2$-smooth equality and inequality constraints in $I\!R^n$.

The main characterization obtained in the latter paper reads as follows: under the validity of LICQ at $\bar{x}$, this point is a tilt-stable local minimizer of the given nonlinear program *if and only if* it satisfies SSOSC along with the corresponding (unique) Lagrange multiplier. This characterization excludes the existence of critical multipliers (independently of tilt stability) due to the discussion above. However, it does not imply that the absence of critical multipliers is always the case for tilt-stable minimizers, since neither LICQ nor SSOSC is *not necessary* for tilt stability of local solutions to nonlinear programs with $\mathcal{C}^2$-smooth data in contrast to strong regularity of the associated KKT systems for which the validity of both LICQ and SSOSC is necessary.

A second-order subdifferential characterization of tilt-stable minimizers for NLPs with equality and inequality constraints described by $\mathcal{C}^2$-smooth functions in finite dimensions, more appropriate for applications to the discussion paper by Izmailov and Solodov, has been recently obtained by Mordukhovich and Nghia (2014a) via developing a new approach completely different from (and simpler than) those in Poliquin and Rockafellar (1998) and Mordukhovich and Rockafellar (2012). This characterization applies to nonregular/degenerate problems and replaces the LICQ assumption by the simultaneous validity of the *Mangasarian–Fromovitz* and *constant rank* constraint qualifications (MFCQ and CRCQ, respectively). The characterization is given in terms of the new *uniform second-order sufficient condition* (USOSC), which is weaker than SSOSC. The example of a linear-quadratic NLP presented in this paper detects a tilt-stable local minimizer via USOSC under the validity of MFCQ and CRCQ but the violation of LICQ and SSOSC, and thus of Robinson's strong regularity and also strong stability in the sense of Kojima (1980). Furthermore, in this example the set of the corresponding Lagrange multipliers is not a singleton while, as one can check, critical multiplies *do not exist*.

Meantime, other strategies to study tilt stability of local minimizers have been tested over the years, especially in the recent time. Bonnans and Shapiro (2000) characterized this notion via the so-called "uniform quadratic growth condition with respect to tilt perturbations" for conic programs with $\mathcal{C}^2$-smooth data under the Robinson constraint qualification, which reduces to MFCQ for nonlinear programs. This line of research was developed by Lewis and Zhang (2013) under a certain nondegeneracy condition and by Drusvyatskiy and Lewis (2013) without imposing such a condition in the abstract setting of extended real-valued prox-regular functions in finite dimensions. Further results in this direction were obtained by Mordukhovich and Nghia (2013) in the case of general Asplund (in particular, reflexive) spaces with establishing *quantitative* relationships between moduli of the *uniform second-order growth* and *prox-regularity* of the objective. The uniform second-order growth conditions used in these papers are closely related to *strong metric regularity* of subgradient mappings.

Subsequently Drusvyatskiy, Mordukhovich et al. (2014a) replaced the latter characterization of tilt-stable minimizers by much more convenient and largely investigated notion of *metric regularity*, in addition to the *positive semidefiniteness* of the generalized Hessian of prox-regular functions in finite-dimensional spaces. The tilt stability characterizations obtained in this direction can be applied to *degenerate* minimization problems in the extended real-valued framework, but its realization in the case of NLPs and other remarkable classes of problems in constrained optimization has not been investigated yet.

The given short (hope useful) survey on tilt stability of local minimizers in constrained optimization raises challenging questions in connection with the discussion paper by Izmailov and Solodov, which have not been posted before:

(i) Does tilt stability of local minimizers for NLP and other nice classes of optimization problems with C^2-smooth data ensure the *absence of critical multipliers*?
(ii) Derive verifiable characterizations of (or even sufficient conditions for) tilt stability of local minimizers entirely in terms of the problem data for NLPs and some other remarkable classes of variational problems *without nondegeneracy* (like LICQ) assumptions via conditions weaker than SSOSC.

So far we obtained some partial results on (ii) for NLPs with inequality constraints. In the underlying example from Mordukhovich and Nghia (2014a) illustrating (ii), we have a set (not singleton) of Lagrange multipliers at a tilt-stable minimizer with no critical multipliers. Thus, it does not give a counterexample to (i).

Next, we will discuss (very briefly) these issues for other types of local minimizers, which are *fully stable* in either the Lipschitzian or Hölderian sense.

3 Critical Lagrange multipliers for fully stable minimizers

The notion of Lipschitzian full stability was introduced by Levy et al. (2000) in the extended real-valued framework of finite-dimensional optimization as follows: given $f : \mathbb{R}^n \times \mathbb{R}^d \to \bar{\mathbb{R}}$ and a nominal parameter pair $(\bar{v}, \bar{p}) \in \mathbb{R}^n \times \mathbb{R}^d$, a point $\bar{x}$ is said to be a *Lipschitzian fully stable local minimizer* of f if there is a number $\gamma > 0$ such that the argminimum mapping

$$M_\gamma : (v, p) \mapsto \operatorname{argmin}\big\{ f(x, p) - \langle v, x \rangle \big| \; \|x - \bar{x}\| \leq \gamma \big\}$$

is single-valued and locally Lipschitz continuous around $(\bar{v}, \bar{p})$ with $M_\gamma(\bar{v}, \bar{p}) = \bar{x}$, and moreover the local optimal value function

$$m_\gamma : (v, p) \mapsto \inf \big\{ f(x, p) - \langle v, x \rangle \big| \; \|x - \bar{x}\| \leq \gamma \big\}$$

is also locally Lipschitz continuous around the reference parameter pair $(\bar{v}, \bar{p})$.

We can see that the difference between tilt and full stability is that the local single-valuedness and Lipschitz continuity of the solution map is required not only with respect to tilt perturbations v but also with respect to *basic* perturbations p, which enter the objective in the general way while *together* with the tilt component. The

main result of Levy, Poliquin, and Rockafellar (2000) provides a characterization of Lipschitzian full stability in terms of a *partial version* of the *generalized Hessian* for a broad class of extended real-valued functions $f(x, p)$, which are prox-regular and subdifferentially continuous in x with compatible parametrization by p at the reference point under the validity of a certain basic constraint qualification.

Based on the new second-order calculus rules specially developed for the later construction, precise calculations of it in the settings of interest for the corresponding optimization problems as well as other variational techniques, a number of effective characterizations of Lipschitzian full stability of local minimizers have been recently obtained for several remarkable classes in finite-dimensional constrained optimization, namely for NLP, *mathematical programs with polyhedral constraints*, and problems of the so-called *extended nonlinear programming* in Mordukhovich, Rockafellar, and Sarabi (2013); for *second-order cone programming* in Mordukhovich et al. (2014b); for *semidefinite programming* in Mordukhovich et al. (2014a); and for *minimax* optimization problems in Mordukhovich and Sarabi (2014). All these characterizations, given entirely in terms of the problem data, are established under the corresponding *nondegeneracy* condition, which are counterparts of LICQ for the aforementioned classes of constrained optimization.

The only exception from this stream is the paper by Mordukhovich and Nghia (2014b), which contains an explicit characterization of Lipschitzian full stability for local minimizers of NLPs via a *partial* version of USOSC (not of SSOSC) while imposing the corresponding partial versions of MFCQ and CRCQ. Roughly speaking, this development is a full stability counterpart of the tilt stability characterization for degenerating NLPs given in Mordukhovich and Nghia (2014a, b). The example presented in the (2014b) paper shows that Lipschitzian full stability may hold in degenerate problems without the partial LICQ and SSOSC conditions and with multivalued sets of Lagrange multipliers. However, critical multiplier *do not appear* in this example.

It is shown furthermore in the aforementioned (2014b) that Lipschitzian full stability and its characterizations are equally important in infinite dimensions, particularly in the framework of *polyhedric programming* in Hilbert spaces with subsequent applications to *optimal control* of semilinear elliptic partial differential equations.

Another possibility to avoid nondegeneracy conditions, at least for problems with inequality constraints and the like, is to consider local minimizers that are *Hölderian fully stable* in the sense introduced by Mordukhovich and Nghia (2014b) in the extended real-valued framework of optimization as follows:

$$\|M_\gamma(v_1, p_1) - M_\gamma(v_2, p_2)\| \le \ell \|v_1 - v_2\| + \kappa \, d(p_1, p_2)^{\frac{1}{2}}$$

for the above argminimum mapping near $(\bar{v}, \bar{p})$ with some positive constants ℓ, κ, where d stands for the distance in the space of basic parameters. Although some characterizations of this notion were derived in the latter paper via uniform second-order growth and generalized Hessian conditions in the abstract setting without imposing nondegeneracy, these characterizations have not been yet realized for particular classes of constrained optimization problems as NLP and those mentioned above.

Most recently in Mordukhovich and Nghia (2014c), these notions and ideas have been developed (in some surprising way) to general classes of parametric variational systems (PVS) given in the generalized equation form

$$v \in g(x, p) + \partial_x f(x, p)$$

via the (limiting) partial subdifferential of a prox-regular function $f : X \times P \to \bar{I\!R}$. The PVS form clearly includes variational and quasi-variational inequalities, complementarity systems, and variational conditions. Some of the characterizations of Lipschitzian and Hölderian full stability of solutions to PVS do not require any nondegeneracy, but they should be more elaborated for particular classes of variational systems in both finite and infinite dimensions.

Regarding relationships between Lipschitzian and Hölderian full stability and critical multipliers, those questions posted above in (i) and (ii) for the case of tilt stability can be just repeated here for the full stability notions under consideration.

4 Conclusion

The excellent discussion paper by Izmailov and Solodov is devoted to recently discovered phenomena in numerical optimization concerning the existence of critical multipliers and their negative influence on convergence rates of primal–dual algorithms. It occurs to be largely related to the recent research on proper stability concepts for local minimizers in constrained optimization and other variational problems. The *main message* of these comments is to develop and/or apply computational algorithms *not to general minimizers* but to their *special classes* satisfying the desired properties of tilt and full stability, which are discussed above. I am inclined to believe that seeking just such stable minimizers will help us to *avoid troubles* with appearing critical multipliers and slow convergence of major primal–dual algorithms. My *conjecture* is that the answer is *yes* to the question formulated in (i) for both cases of tilt-stable and fully stable minimizers under appropriate assumptions in the corresponding case without any nondegeneracy condition. Then the question in (ii) is how to recognize that we have tilt-stable or fully stable optimal solutions of the problem entirely via its initial data. As the comments above show, much more research should be done in this direction for degenerate problems. My general feeling is that tilt and full stability could play a *positive role* for problems with possible degeneracy similar to strong regularity in nondegenerate problems. As we now know, tilt stability of local minimizers is *equivalent* to strong regularity of the associated KKT system for major classes of $\mathcal{C}^2$-smooth problems in constrained optimization *under nondegeneracy*, which is *unavoidable* for strong regularity, but not for tilt and full stability.

References

Bonnans JF, Shapiro A (2000) Perturbation analysis of optimization problems. Springer, New York
Dontchev AL, Rockafellar RE (1996) Characterizations of strong regularity for variational inequalities over polyhedral convex sets. SIAM J Optim 6:1087–1105

Drusvyatskiy D, Lewis AS (2013) Tilt stability, uniform quadratic growth, and strong metric regularity of the subdifferential. SIAM J Optim 23:256–267

Drusvyatskiy D, Mordukhovich BS, Nghia TTA (2014) Second-order growth, tilt stability, and metric regularity of the subdifferential. J Convex Anal 21:4

Kojima M (1980) Strongly stable stationary solutions in nonlinear programming. In: Robinson SM (ed) Analysis and computation of fixed points. Academic Press, New York, pp 93–138

Levy AB, Poliquin RA, Rockafellar RT (2000) Stability of locally optimal solutions. SIAM J Optim 10:580–604

Lewis AS, Zhang S (2013) Partial smoothness, tilt stability, and generalized Hessians. SIAM J Optim 23:74–94

Mordukhovich BS (1992) Sensitivity analysis in nonsmooth optimization. In: Field DA, Komkov V (eds) Theoretical aspects of industrial design. SIAM Volum Appl Math 58:32–46

Mordukhovich BS (2006a) Variational analysis and generalized differentiation, I: basic theory. Springer, Berlin

Mordukhovich BS (2006b) Variational analysis and generalized differentiation, II: applications. Springer, Berlin

Mordukhovich BS, Nghia TTA (2013) Second-order variational analysis and characterizations of tilt-stable optimal solutions in infinite-dimensional spaces. Nonlinear Anal 86:159–180

Mordukhovich BS, Nghia TTA (2014a) Second-order characterizations of tilt stability with applications to nonlinear programming. Math Program. doi:10.1007/s10107-013-0739-8

Mordukhovich BS, Nghia TTA (2014b) Full Lipschitzian and Hölderian stability in optimization with applications to mathematical programming and optimal control. SIAM J Optim 24:1344–1381

Mordukhovich BS, Nghia TTA (2014c) Local strong maximal monotonicity and full stability for parametric variational systems. Preprint http://arxiv.org/abs/1409.2018

Mordukhovich BS, Rockafellar RT (2012) Second-order subdifferential calculus with applications to tilt stability in optimization. SIAM J Optim 22:953–986

Mordukhovich BS, Sarabi ME (2014) Variational analysis and full stability of optimal solutions to constrained and minimax problems. Nonlinear Anal. doi:10.1016/jna.2014.10.013

Mordukhovich BS, Rockafellar RT, Sarabi ME (2013) Characterizations of full stability in constrained optimization. SIAM J Optim 23:1810–1848

Mordukhovich BS, Nghia TTA, Rockafellar RT (2014a) Full stability in finite-dimensional optimization. Math Oper Res. doi:10.1287/moor.2014.0669

Mordukhovich BS, Outrata JV, Sarabi ME (2014b) Full stability of locally optimal solutions in second-order cone programming. SIAM J Optim 24:1581–1613

Poliquin RA, Rockafellar RT (1998) Tilt stability of a local minimum. SIAM J Optim 8:287–299

Robinson SM (1980) Strongly regular generalized equations. Math Oper Res 5:43–62

Rockafellar RT, Wets RJB (1998) Variational analysis. Springer, Berlin