

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

J. M. Martínez

Published online: 27 February 2015
© Sociedad de Estadística e Investigación Operativa 2015

Consider the minimization of x^2 subject to $x^2 = 0$. The solution of this trivial problem is $x_* = 0$. This solution satisfies the Lagrange (KKT) conditions and every $\lambda \in \mathbb{R}$ is an admissible Lagrange multiplier. One of these multipliers ($\lambda_* = -1$) has undesirable properties: the distance between (x_*, λ_*) and (x, λ) is not bounded by a multiple of the norm of the KKT system computed at (x, λ) . This means that the norm of the KKT system is not a safe estimator or the primal–dual distance to the solution. Roughly speaking, multipliers with this characteristic are said to be *critical*. The paper by Izmailov and Solodov surveys all the present knowledge about critical multipliers.

The authors present their subject as a humorous list of bad news. Restricting the analysis to the equality-constrained smooth optimization problem, a critical vector of multipliers associated with a stationary point x_* is defined by the property that the Hessian of the corresponding Lagrangian is singular on the tangent subspace to the constraints. This is obviously an undesirable property and one of its effects is reflected in the example mentioned in the first paragraph of this report. Only when the multiplier is noncritical an error-bound property can be guaranteed. Moreover, non-criticality is equivalent to the fulfillment of a primal–dual error bound estimation (Proposition 2.2 of the paper).

The second bad news is really very bad. The Newton–Lagrange method (Newton’s method applied to the KKT system) has the persistent tendency of converging to critical multipliers, when they exist. This means that if one is using Newton for solving a constrained optimization problem and one uses a stopping criterion based on the norm of the KKT residual, stopping may occur at a point that is far from the solution, even if

This comment refers to the invited paper available at doi:[10.1007/s11750-015-0372-1](https://doi.org/10.1007/s11750-015-0372-1).

J. M. Martínez (✉)
Campinas, Brazil
e-mail: martinez@ime.unicamp.br

convergence to the solution is deemed to take place. In cases as the minimization of x^2 subject to $x^2 = 0$ one has the feeling that Newton “fails” because of its “greediness”: it tries, at early iterations, to satisfy prematurely the dual feasibility. The paradox is that it succeeds, leading to a small KKT residual when the primal approximation is still poor.

The third bad news is that the tendency of convergence to critical multipliers is shared by other methods, as quasi-Newton SQP and linearly constrained Lagrangian methods.

In the first three sections, the paper presents several insightful examples and families of problems that illustrate the convergence to critical multipliers. The examples are accompanied by illustrative pictures that help the reader to understand the phenomenon. The authors warn that no enough general theory is available yet, a fact that, undoubtedly, will stimulate future research.

The fourth section is dedicated to the analysis of the stabilized SQP method (stabilized Newton–Lagrange, or SNLM, in this paper) and the augmented Lagrangian method from the local point of view. Under suitable assumptions, starting close enough to a primal–dual solution with noncritical multiplier, both methods converge superlinearly to a primal–dual pair at which the multiplier is noncritical. The stabilization function of the stabilized Lagrange method cannot tend to zero too fast, but no analogous restriction has been detected for the augmented Lagrangian method. In any case, starting for arbitrary initial points even the methods for which nice local results exist has been observed to converge to critical multipliers.

In the fifth (and final) section of the paper, the augmented Lagrangian method and the stabilized Newton–Lagrange method are discussed with more detail. In the first case, attempts to accelerate the final iterations with cheaper stabilized Newton–Lagrange iterations are commented, and in the case of the SNLM method, globalization procedures by means of augmented Lagrangian and primal–dual merit functions are briefly discussed. Concerning the main objective of this paper, a great amount of open questions, many of which have not been formulated yet, arise.

Reading this very insightful survey, two aspects related with my own research came to my mind. The first concerns the behavior of Newton and other methods in the case of convergence to points at which the KKT conditions do not hold at all, that is, there are no Lagrange multipliers at the solution. The question in this case is whether the primal–dual sequence generated by a method satisfies asymptotically approximate KKT conditions or not. Roughly speaking, the answer is negative for NLM but positive for the Augmented Lagrangian and other methods ([Andreani et al. 2014](#)), a curiously analogous answer to the one obtained by Izmailov and Solodov for the case of critical multipliers. Moreover, the behavior of SNLM in the absence of Lagrange multipliers remains to be an open problem. My second concern is purely practical: using Lagrange multipliers in the Newton iteration is obviously good in noncritical cases, since it leads to superlinear or quadratic convergence, but has a computational cost that could be significative. Moreover, the sparsity of the Hessian of an objective function could be totally destroyed in the Hessian of the Lagrangian due to different patterns of the constraints Hessians. In some practical situations, the overall efficiency of Newton and other methods is increased if one merely discards the Hessians of the constraints, which amounts to “pretend” that the constraints are linear

for the purpose of computing increments. We observed clearly this phenomenon in the family of hard-sphere problems (Krejić et al. 2000). The simplification normally sacrifices superlinear convergence, but may not have serious inconvenients from the global point of view. It could be argued that, in the cases that a method is condemned to slow convergence due to approximation to critical multipliers, a radical simplification in the Hessian of the Lagrangian could preserve linear convergence with some improvement in terms of computer time.

Summing up, this is a very motivating survey that will attract the attention of readers to the addressed questions and to connections with unsuspected related theoretical and practical problems.

References

- Andreani R, Martínez JM, Santos LT, Svaiter BF (2014) On the behavior of constrained optimization methods when Lagrange multipliers do not exist. *Optim Methods Softw* 29:646–657
- Krejić N, Martínez JM, Mello MP, Pilotta EA (2000) Validation of an augmented Lagrangian algorithm with a Gauss–Newton Hessian approximation using a set of hard-spheres problems. *Comput Optim Appl* 16:247–263