

Referee Report

OUTER APPROXIMATION SCHEME FOR WEAKLY CONVEX CONSTRAINED OPTIMIZATION PROBLEMS

September 10, 2024

The Ph.D. thesis focuses on the solution of weakly convex constrained optimization problems using outer approximation schemes.

This is a class of optimization problems strictly containing the convex ones due to the fact that any weakly convex function is indeed convex.

In contrast to more classical approaches, the proposed one is based on a series of outer approximations of the feasible region based on spheres rather than planes. This is indeed a more natural approach given certain technical properties of the weakly convex functions.

The cutting sphere algorithm and the related convergence theory introduced and presented in this Ph.D. work, rely on a series of interesting and technically challenging results there disseminated.

To the best of my knowledge, the main contributions of the work include:

1. the reformulation of a general weakly convex problem as an Euclidean projection problem into a suitable set, see Section 3.2;
2. the proof of the existence of convergent subsequence of the Cutting Sphere Algorithm to the optimal solution of the reformulated optimization problem, see Theorem 30;
3. the introduction of computationally realizable versions of the Cutting Sphere Algorithm and related theoretical analysis, see Chapter 6.

Overall, I am positively impressed by the quality and the extensiveness of the presented results and I am confident about author's potential to successfully move along the difficult road of seeking efficient algorithms for the solution of weakly convex optimization problems.

Moreover, due to the generality of the considered problem, the work might have a far reaching impact as it would provide practitioners with a solid theoretical backbone to model and solve real world problems where simple convexity is lacking. Among other weakly convex problems, we mention explicitly the *Robust Matrix Sensing* and the *Robust Phase Retrieval* problems as meaningful representative of weakly convex optimization problems.

For all the above mentioned reasons, I am sure that the dissertation is mature and ready to be accepted, despite some minor remarks that I still have and put in the next part of this report. I also attach an annotated version of the thesis that might help the candidate in locating the pointed out parts of the work.

1. P.ii: “and” repeated 2 times; QCQP and S-QCQP not defined. Moreover: “normed squared” → “squared norm” ?
2. P1: “ $\beta \rightarrow \beta_1$ ”;
3. P2: the meaning of the sentence around (1.2) is not clear. “Problem” → “Problems”.
4. P3: remove “the”;
5. P4: add def of Uniformly Convex Banach Space;

6. P5: x_0 is fixed, clash of notation;
7. P6: the sentence toward the end of the page is not clear. What does it mean for a sequence of operators to converge to a common fixed point? Moreover Tk should be T_k ;
8. P7: end of page. How your approach differs from the papers mentioned above?
9. P10: in the statement of Proposition 4 there is no need to repeat the the definition of paraconvexity. The statement should something like “ f is ρ -weakly convex iff f is ρ -paraconvex”;
10. P12: $\gamma_\rho \rightarrow \Gamma_\rho$;
11. P15: “appear be” $\rightarrow$ “appear to be”;
12. P16: above equ. (2.14) h should be f ?
13. P20: from equ. (2.3) the notation $\partial_{(\cdot,\cdot)}^\epsilon$ is introduced but seems to be not defined. Clarifications on such notation should be introduced;
14. P31: A clarification of the equivalence of (1.1) with (3.4) would be beneficial for readability; The proof of Theorem 20 should be properly expanded and the better explained; this is somehow a central result and many of the crucial passages seem to be hidden. E.g., it is stated that for a feasible solution (x, x_n) of (3.4), it holds $x \in D$. Assuming the equivalence of (1.1) and (3.4) I can only conclude that $x \in \arg \min_{z \in S} F(z)$, but not $x \in D$ as stated. Indeed, it seems that D might be strictly contained in the set of the minimisers of F . Otherwise, I believe, it should be proved that there exist (ρ, η) s.t. (3.1) is true for all the minimisers minimisers of F on S .
15. P32: where is the constraints $g_0(y) \leq 0$ defined?
16. P33: replace 0 wit $\emptyset$;
17. P34: + should be $-$. Why it is assumed that k is finite?
18. P39: the condition $z \in E$ is not stated in Algorithm 1 but used in the subsequent proof;
19. P41: Add “then” after the underlined text;
20. P49: Would you please expand the reasons for the existence of the solution of (OP_k) after eq. (4.1)?
21. P50: $A_J, A_k - - > A_0, A_i$? Same for c_J, c_k ?
22. P53: add “of” after “complement”;
23. P54: x^T is missing after x ; the notation $A_i(x) \succeq 0$ is not defined and (4.7) seems to be a simple rewriting of (4.6); the feasible sets of (SDP) and (QCQP) are sup[posed to be convex or are convex under the current assumptions (i.e., convex QCQP)?
24. P66: $A_J - - > A_0$?
25. P69: Ω_0 not defined;
26. P670: $A_J - - > A_0$?
27. P72: Section 2 should be Section 4.6.1; Eq. (4.41), J should be 0?
28. P75: J should be 0?
29. P76: J should be 0?
30. P77: J should be 0?
31. P78: J should be 0?
32. P79: Example 55 could be a Lemma, as it proves the convexity og GIS for (QCQP); J should be 0?
33. P81: “has” $- - >$ “is”;
34. P89: “has” $- - >$ “is”; J should be 0? In (KKT) γ_k is missing;
35. P90: J should be 0?

36. P92: q_i should be f_i ? γ_i should multiply also b_i ?
37. P94: it is no clear to me how you proved that x belongs also to the second part of $\mathcal{O}_k^\alpha$ (see the highlighted part);
38. P106: the number of constraints m_k is initialized in Algorithm 2. it would be useful to remind the reader how m_k is updated during the algorithm;
39. P127: “instances” $\rightarrow$ “instance”;
40. P128: I think that the computational results presented at the beginning of this page might be misleading. The efficiency/effectiveness of Algorithm 3 might appear limited given the dimension of the considered problem and the presented computational times. Rather than focusing on time I suggest to enrich the section with graphs regarding, e.g., the behaviour of the objective function w.r.t. the iterations. Actually, it would be interesting also in having a computational study regarding the Algorithmic performance in relation to the algorithm hyper-parameters. It would be interesting, for example, understanding how the algorithm behave w.r.t. the parameter $\bar{m}$ or the initial point (as we know the exact solution) or the inaccuracy ε . Moreover, the values of these parameter used for the presented experiments is missing.

Edinburgh, 10th September 2024

Stefano Epalle