

REPORT

on the doctoral dissertation

“Abstract Convexity: Duality, Primal-Dual and Proximal based Algorithms”

by The Hung Tran

September-13-2024

The dissertation presents a comprehensive study of a broad range of topics within and around abstract and generalized convexity, a short excerpt of the full list of which is given in the title. The topics and obtained results nicely fit together producing a rather complete theory with not many gaps left. It is pleasing to see that the theory is complemented with algorithmic applications.

The research focuses on the so-called composite minimization problems, which are of importance from the application viewpoint and, at the same time, give rise to interesting theoretical questions related to abstract convexity. Such problems involve two types of functions and, if they are treated as abstractly convex, the functions are in general determined by two different classes of “elementary” functions on respective spaces. This gives researchers some freedom in choosing appropriate classes and, at the same time, causes the need for developing special approaches taking into account interactions between the classes. This adds a new dimension to the traditional research directions like (conjugate and Lagrange) duality, zero duality gap, and optimality conditions.

In the dissertation and related publications, the candidate has developed an interesting approach to addressing the mentioned issues for composite minimization problems. The dissertation is well structured. The first two-chapter part addresses theoretical abstract convexity questions for composite minimization problems, while the second two-chapter part is devoted to algorithms and their convergence analysis. All major results are illustrated by examples.

In the first part, the dissertation makes a good use of the class of $\Phi_{lsc}^{\mathbb{R}}$ -convex functions (on a Hilbert space) which includes, in particular, weakly convex functions. It studies the corresponding subdifferentials and conjugates, and establishes connections between $\Phi_{lsc}^{\mathbb{R}}$ -subdifferentials and proximal subdifferentials.

A general composite minimization problem is then considered. It involves two component functions defined in general on two different spaces (sets) and, hence, requires considering two different classes Φ and Ψ of elementary functions. The candidate illustrates the versatility of this problem, defines the conjugate dual problem in a natural way and proves sufficient and necessary conditions for zero duality gap and then examines existence of solutions to the dual problem, i.e., strong duality. These conditions are then discussed and compared with the known ones when appropriate classes Φ and Ψ are chosen. Several important special cases are considered including the generalized Moreau envelope.

A certain Lagrangian function is introduced employing arbitrary classes Φ and Ψ of elementary functions, and the Lagrange dual problem is defined. The relationship between the conjugate and Lagrange dual problem as well as conditions ensuring zero duality gap are discussed.

The content of the second part is strongly based on the findings presented in the first part. The candidate constructs the $\Phi_{lsc}^{\mathbb{R}}$ -proximal operator, shows its basic properties and representations related to minimizing $\Phi_{lsc}^{\mathbb{R}}$ -convex functions, and illustrates the computation of elements of the operator. He then gives a detailed convergence analysis of a series of

corresponding $\Phi_{lsc}^{\mathbb{R}}$ -algorithms: proximal point, forward-backward, projected subgradient as well as certain primal-dual algorithms. A primal-dual algorithm for weakly convex functions is also studied. The performance of the algorithms is illustrated by a series of numerical examples.

Overall, the dissertation demonstrates the candidate's very good general theoretical knowledge in the discipline and his ability to independently conduct scientific research.

I recommend the thesis for acceptance and admission to public defence.

A. Kruger

Professor Alexander Kruger

Certain aspects of the dissertation could be improved. I would like to see more rigour in some mathematical arguments. There are also issues related to presentation: many unnecessary repetitions, duplicated assumptions, and some cases of inconsistency in notation and terminology. The dissertation is not always easy to read because of numerous language issues. Below are a few specific remarks, although I do not expect the candidate to address all of them before the defence.

1. The candidate needs to pay more attention to the assumptions and conventions. For instance, in the survey Sect. 3.1, the function f is assumed to act between X and $(-\infty, +\infty]$, and this assumption is repeated in many definitions and statements, which could probably be avoided. (This also applies to many other conventions.) At the same time, some definitions (e.g., Definition 2) go without this assumption; can f be different there? How can it be “by convention, $f \equiv -\infty$ ” (page 10)? Should f be assumed proper? Otherwise, definition (3.2) would give $-\infty$. Should Φ be assumed nonempty? X is assumed in Definition 3 to be a nonempty set; can it be empty in the other definitions?
2. The meaning of the “converse inclusion” in the “moreover” part of Proposition 2 needs to be explained. Converse to what? See the claim at the bottom of page 14.
3. Remark 1. The reference to (4) in part (i) to be checked. I am not sure about the purpose of the constant term claimed in part (ii). What is the meaning of the claim “ ∂_{Φ} is not unique” in part (iv)?
4. Example 3. Inequality (3.15) is formulated for the case $a > -1/(2\gamma)$ but then used in the case $a < -1/(2\gamma)$.
5. I do not understand Proposition 7. The inclusion seems to hold by definition for ANY subdifferential and ANY function.
6. In Lemma 2, the same condition is involved in both the assumptions and the conclusions.
7. Page 41. How to understand $\sup_{x \in X} -\infty$? I am suspicious about the inequality $\sup_{x \in X} -\infty > -\infty$.
8. Is it necessary to write $\varepsilon/2$ in Theorems 4 and 5 if the respective inequalities are formulated “for every $\varepsilon > 0$ ”?

9. It is assumed at the beginning of Sect. 4.2.1 that $a \geq 0$, but then it changes to $a > 0$ in the proof of Proposition 8 and to $a \leq 0$ in Corollary 1. Furthermore, it is observed after the proof of the corollary that “the above corollary does not hold with $a < 0$ ”.
10. Page 69. Should the inequality in line 2 and the first inequality in line 4 be equalities?
11. Page 71. “For every $\alpha \in \mathbb{R}$, let $\alpha < \dots$ ”. I am unable to understand this assumption.
12. Page 87. The last expression in the representation of $J_\gamma(x)$ in Definition 9 is not mathematically correct. This also applies to the representations in Remarks 14 and 15.
13. Pages 96 and 120. “is bounded and converges”. Can ‘bounded’ be dropped?
14. Pages 96, 102 and 120. “is decreasing and converges”. Can ‘converges’ be dropped?
15. Can squares be dropped in conditions (v) and (vi) of Theorems 12 and 13, conditions (i) and (ii) of Proposition 16, and in Propositions 18 and 20?
16. Example 19. The expression for the dual problem given on page 144 is only valid when $|y| \leq 1$; hence the supremum should not be taken over the whole of $\mathbb{R}$ (compare with Example 20).

Strict inequality (6.22) is violated already when $x = y = 0$. I guess this is a typo, and one should consider the non-strict inequality.

The candidate produces level plots of the function in the left-hand side of (6.22) for three values of μ to show that the Lagrangian is not inf-sharp around $(0, 0)$. This approach to “proving” things does not look convincing. Considering just three values of the parameter is not enough to conclude that something never happens.

The claimed conclusion can be easily obtained analytically. It suffices to show that

$$\inf_{(x,y) \in \mathbb{R}^2 \setminus \{(0,0)\}} \frac{\mathcal{H}(x,y)}{\sqrt{x^2 + y^2}} = 0.$$

Just take $x = 0$ and let $|y| \downarrow 0$.

17. Example 20. The given three charts add nothing to the conclusion already made.
18. Example 21. It is not obvious to me that $(0, 0)$ is a saddle point. Can there be other saddle points? The answer can also affect Example 23.
 “This implies that Lagrangian is not inf-sharp with constant 1.” No, it does not!
 “In fact, it is locally inf-sharp with constant $\mu < 1$ ”. I am not sure if the given charts prove this claim.
19. When assigning “*argmin*” to a variable, it would be not bad to explain why it is a singleton (if it is).
20. In Propositions 18 and 20, Corollary 4 and Theorems 18 and 20, it should be noted that μ comes from Definition 12.
21. Page 156. What is convergence to a set?
22. Page 162. Check the third displayed formula.
23. The bibliography list does not seem to adhere to a particular style. I would prefer to see more consistency in presenting the names/initials of the authors. I suggest to avoid using ‘et al’ in the list.