November 18, 2014

Shalom Sergei: As I told you Shoham and I have discussed yesterday the directions for your research and below you will find guidelines to start your work toward your M.Sc. thesis. Enjoy!

Marc and Shoham

Research Topic: First Order Methods for Structured Optimization Problems with Applications

After, you will be completing the three first tasks below, by then your knowledge from the optimization courses will also have expanded, then we will present you more specific questions/problems to be tackled that will be the "meat" of your thesis. Please note that the 3 papers are completely independent, so that you can study them in parallel...Nevertheless, I would recommend that you first read [3] (which is very general and also light in math contents), then [1], and keep [2] for the "dessert"!

1. **Some Motivation.** A huge variety of fundamental scientific problems can be modeled via nonsmooth, convex and often nonconvex models ranging from machine learning to image/signal recovery problems etc...Our intent is not to develop and analyze *general* algorithms (this is hopeless!). Rather, we will focus on problems which special structures, in the hope that these can be beneficially exploited to devise smart and tractable methods. Here to give you a first contact with such kind of problems in the nonconvex area, we will focus on clustering problems.

   The paper [1] will give you an overview of the field within the optimization framework. As you will see this problem is very hard! You dont need to enter into all the details in that paper, (we intent to tackle the problem within different methods than the one described in the paper). The main objective is to give you an idea on this a rich area and its may important applications, and which leads to many interesting optimization questions.

2. **First Order Methods.** You are already familiar with first order methods for convex problems. Here we want to focus on similar methods but for *nonconvex problems*. The paper [2] will be give you the necessary mathematical background and tools that will be useful to tackle other class of problems. In fact, the last section of the paper is

on matrix factorization problems, which are also directly connected to the clustering problem.

3. **Dual and Lagrangian Decomposition Methods.** This is a wide subject which I did not covered in the algorithm course the year you took it. So, your task here will be to become familiar with these algorithms. It is important that you familiarize with these methods, both in convex and nonconvex situations. Indeed, for instance we would like to consider this approach for possibly tackling the clustering problem, an approach which to the best of our knowledge has never been studied.

   There is a huge body of literature on the subject....and the survey paper [3] is describing nicely both the ideas and applications of these methods. You don't need to read all the chapters...!.. but mainly the ones describing the structure of the main ideas and algorithms. The chapter 3 of the book [4] also gives nice description and more mathematical details on all these methods. It can be downloaded for free from:

   `https://dspace.mit.edu/handle/1721.1/3719`

We will start discussing more specific questions, (there are many!!–Both Theoretical and Computational) after you gain familiarity with the above material.

It is a good habit from the beginning of a research project to summarize and keep tract of what you are reading. This way, you not only gain insights in your work, but also save much time toward the end, and when ready, it becomes then much easier to summarize the output of your own research when writing your thesis.

Finally, if you are not yet familiar with the ultimate word processor for mathematicians: **Latex**...(as this document was produced)... This is the best time to start...!....

Bibliography

[1] M. Teboulle. "A unified continuous optimization framework for center-based clustering " methods. *Jounal of Machine Learning Research*, 8 (2007), 65–102.

[2] J. Bolte, S. Sabach, and M. Teboulle. " Proximal alternating linearized minimization for nonconvex and nonsmooth problems." *Mathematical Programmng, Ser. A*, 146 (2014), 459–494.

[3] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein "Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers." *Foundations and Trends in Machine Learning*, 3(1), (2011), 1122.

[4] D. P. Bertsekas and J. N. Tsitsiklis, *Parallel and Distributed Computation: Numerical Methods* (Prentice-Hall, New Jersey, 1989).