Magic Gradient Descent*

Albert Author[1] and Bernard D. Researcher[2]

***Abstract*— Describe in a few sentences what the paper is about and why it is interesting to read it.**

I. INTRODUCTION

Some general introducing sentences about the topic, motivation and relevance of problem/algorithm.

In this paper we give an introduction to the results presented in paper(s) [**?**].

We present the problem statement (optimization problem) the main results/algorithms, discuss the underlying ideas and illustrate the results by numerical simulations.

Notation. Define notation.

II. PROBLEM STATEMENT AND BACKGROUND

Provide a mathematical problem description. If necessary, some background material.

$$\begin{aligned} \underset{x}{\text{minimize}} \quad & f_0(x) \\ \text{subject to} \quad & f_i(x) \le 0,\ i = 1, \dots, m. \\ & A_{\text{eq}} x = b_{\text{eq}}. \end{aligned} \tag{1}$$

-convex quadratic opt prblem with inequality constraints -¿ how to handle ineq. constraints? -technical relevance: optimal control, mpc

III. MAIN RESULTS

A. Concept of Barrier Methods

Convex optimization Problems with no inequality constraints can be solved efficiently by using Newton's method. If inequality constraints are involved, Newton's method can not guarantee feasibiliy of a solution. It is hence desirable, to transform an inequality-constrained optimization problem into a only equality-constrained one. Therefore, we move the inequality constraints implicitley to the objective function. TODO!A simple and also precise way to do this, evaluate an indicator function

$$I_-(x) := \begin{cases} 0 & \text{for } u \neq 0 \\ \infty & \text{for } u > 0 \end{cases} \tag{2}$$

on the values of the inequality constraints $f_i, i = 1, \dots, m$. Then, the optimization Problem has the shape

$$\begin{aligned} \underset{x}{\text{minimize}} \quad & f_0(x) + \sum_{i=1}^{m} I_-(f_i(x)) \\ \text{subject to} \quad & A_{\text{eq}} x - b_{\text{eq}} = 0,\ i = 1, \dots, p. \end{aligned} \tag{3}$$

This problem is an equivalent to (1) and has no inequality constraints. However, it is clearly neither convex nor continuous (and hence not differentiable). Since we need these properties to solve the optimization problem computationally, we approximate the indicator function I_- by the function

$$\hat{I}_-(u) = \begin{cases} \frac{1}{t}\log(-u) & \text{for } u < 0, \\ \infty & \text{for } u \ge 0, \end{cases} \tag{4}$$

The parameter $t > 0$ sets the approximation's accuracy. The higher t is, the better the indicator function is approximated. By replacing the Indicator functions by $\hat{I}_-$, we obtain an approximation

$$\begin{aligned} \underset{x}{\text{minimize}} \quad & f_0(x) - \sum_{i=1}^{m} \frac{1}{t} \log(-f_i(x)) \\ \text{subject to} \quad & A_{\text{eq}} x - b_{\text{eq}} = 0 \end{aligned} \tag{5}$$

of problem (1).

Note, that $\frac{1}{t}\log(-u)$ is convex, increasing in u, and differentiable on the feasible set. Hence the entire function $\sum_{i=1}^{m} \hat{I}_-(f_i(x))$ is convex and (5) is a convex Problem with differentiable objective function. These properties allow us to solve (5) computationally. We call an optimal point $x^*(t)$ of (5) with parameter t a central point and a solution to its dual problem $(\lambda^*(t), \nu^*(t))$ a dual central point. The set of (dual) solutions of (5) for all $t > 0$ we call the (dual) central path. One can show, that solutions $(x^*(t), \lambda^*(t), \nu^*(t))$ of (5) converge to the solution (x^*, λ^*, ν^*) of (1) for t going towards zero. The proof is shown in [2].

B. Measure for the Approximation's quality

An immediately arising question is, what conclusions about the solution (x^*, λ^*, ν^*) of (1) can be drawn from a knowing a solution of (5) for a certain $t > 0$.

C. Newton's Method

Newton's method is an iterative process to solve nonlinear equality systems

$$F(x) = 0 \tag{6}$$

for a differentiable map $F : \mathbb{R}^n \longrightarrow \mathbb{R}^m$. The idea of this algorithm is as follows: At a given point x_k, the zero of the linear approximation of F around x_k is computed. This point is chosen as the next iterate x_{k+1}. In particular, a linear approximation of F in x_k is defined as

$$L(x) := F(x_k) + JF(x_k)(x - x_k) \text{ for } x \in \mathbb{R}^n, \tag{7}$$

where $JF(x_k)$ is the Jacobian of F at the point x_k. If $JF(x_k)$ invertible, the point $\tilde{x}$ with $L(\tilde{x}) = 0$ is exactly the solution of the linear equality $JF(x_k)x = -F(x)$. Technical

*Project within the course Convex Optimization, University of Stuttgart, July 2, 2020.

[1]Albert Author is a student of the Bachelor study program Mechatronics, University of Stuttgart, `albert.author@papercept.net`

[2]Bernard D. Researcher is a student of the Master study program Engineering ... ., University of Stuttgart, `b.d.researcher@ieee.org`

conditions and proofs about convergence rates of Newton's method can be found in [1].

For the purpose of optimizing a convex, twice differentiable objective function f_0 we want to find a zero of the gradient ∇f_0. Therefore we can apply the Newton Method to solve the non-linear equation

$$F(x) := \begin{pmatrix} \nabla f_0(x) \\ g(x) \end{pmatrix} = 0 \qquad \text{with } g(x) = \begin{pmatrix} g_1(x) \\ \vdots \\ g_p(x) \end{pmatrix}$$

. By convexity, satisfying $\nabla f_0(x^*) = 0$ is not only neccessary, but also sufficient for x^* to be a global minimum of f_0.

Present main theorems/algorithm. Explain idea, explain algorithm, provide a convergence proof, discuss main properties (advantages and disadvantages) Use algorithm environment in Latex to present algorithm (pseudo-code)

IV. EXAMPLES

Show and discuss simulation examples etc....

V. CONCLUSIONS

Summarize the main points (with more details than in the preceding introduction). The paper should not be between 4 and 8 pages.

APPENDIX

Add for example your Matlab code here. (Code should be nicely formated and documented).

Appendixes should appear before the acknowledgment.

ACKNOWLEDGMENT

REFERENCES

[1] Carsten Scherer *Vorlesungsskript Einführung in die Optimierung* 2019: Lehrstuhl für Mathematische Systemtheorie, Universität Stuttgart.

[2] Stephen Boyd, Lieven Vandenberghe *Convex Optimization* 2004: Cambridge University Press.