Universitat de Barcelona

Numerical Linear Algebra

master in fundamental principles of data science

Project 1

Direct methods in optimization with constraints

DATA SCIENCE @ UNIVERSITAT DE BARCELONA

Author
Vladislav Nikolov Vasilev

Faculty of Mathematics and Computer Science

Academic year 2021-2022

Contents

1 Introduction

The main goal of this project is to study the basic numerical linear algebra behind optimization problems. In this case, we are going to consider a convex optimization problem which has equality and inequality constraints. The goal is to find a value of $x \in \mathbb{R}^n$ that solves the following problem:

$$\begin{aligned} &\text{minimize } f(x) = \frac{1}{2}x^T G x + g^T x \\ &\text{subject to } A^T x = b, \quad C^T x \geq d \end{aligned} \tag{1}$$

This constrained minimization problem can be solved using Lagrange multipliers. In order to do so, we have to transform the inequality constraints into equality constraints. Therefore, we introduce the slack variable $s = C^T x - d \in \mathbb{R}^m, s \geq 0$. The Lagrangian is given by the following expression:

$$L(x, \gamma, \lambda, s) = \frac{1}{2}x^T G x + g^T x - \gamma^T (A^T x - b) - \lambda^T (C^T x - d - s) \tag{2}$$

The previous expression can be rewritten as:

$$\begin{aligned} Gx + g - A\gamma - C\lambda &= 0 \\ b - A^T x &= 0 \\ s + d - C^T x &= 0 \\ s_i \lambda_i &= 0, \quad i = 1, \dots, m \end{aligned} \tag{3}$$

To solve this problem, we are going to use the Newton's method. Additionally, each step of the method is going to have two correction substeps that that will help us stay in the feasible region of the problem.

T1: In order to solve this problem, let us define $z = (x, \gamma, \lambda, s)$ and $F : \mathbb{R}^N \to \mathbb{R}^N$, where $N = n + p + 2m$. The function F can be defined as follows:

$$F(z) = F(x, \gamma, \lambda, s) = (Gx + g - A\gamma - C\lambda, b - A^T x, s + d - C^T x, s_i \lambda_i)$$

Our goal is to solve $F(z) = 0$ using Newton's method. This involves computing a Newton step δ_z so that for a given point z_k we have that $z_{k+1} = z_k + \delta_z, \forall k \in \mathbb{N}$. Knowing that $F(z_{k+1}) = F(z_k) + J_F \delta_z$, we can see that this is equivalent to solving the system $J_F \delta_z = -F(z_k)$, where J_F is the Jacobian matrix. This matrix is defined as follows:

$$J_F = \begin{pmatrix} \frac{\partial F_1}{\partial x} & \frac{\partial F_1}{\partial \gamma} & \frac{\partial F_1}{\partial \lambda} & \frac{\partial F_1}{\partial s} \\ \frac{\partial F_2}{\partial x} & \frac{\partial F_2}{\partial \gamma} & \frac{\partial F_2}{\partial \lambda} & \frac{\partial F_2}{\partial s} \\ \frac{\partial F_3}{\partial x} & \frac{\partial F_3}{\partial \gamma} & \frac{\partial F_3}{\partial \lambda} & \frac{\partial F_3}{\partial s} \\ \frac{\partial F_4}{\partial x} & \frac{\partial F_4}{\partial \gamma} & \frac{\partial F_4}{\partial \lambda} & \frac{\partial F_4}{\partial s} \end{pmatrix} = \begin{pmatrix} G & -A & -C & 0 \\ -A^T & 0 & 0 & 0 \\ -C^T & 0 & 0 & I \\ 0 & 0 & S & \Lambda \end{pmatrix} := M_{\text{KKT}}$$

where I is a $m \times m$ identity matrix and S and Λ are $m \times m$ diagonal matrices containing the values of s and λ, respectively.

Thus, in order to obtain δ_z at each step, we have to solve a linear system of equations defined by the matrix M_{KKT} and the right hand vector $-F(z_k)$.

2 Solving the KKT system without inequalities

The first case that we are going to study is the KKT system without inequalities. This allows us to represent the matrix M_{KKT} as a 3×3 block matrix:

$$M_{\text{KKT}} = \begin{pmatrix} G & -C & 0 \\ -C^T & 0 & I \\ 0 & S & \Lambda \end{pmatrix}$$

There are some strategies that all

2.1 Naive approach

2.2 Other strategies

2.3 Experimentation

3 Solving the general `KKT` system

3.1 LU factorization

3.2 LDL^T factorization

3.3 Experimentation

4 Conclusions

5 Additional comments

References

[1] Texto referencia
`https://url.referencia.com`