TMA4180 Optimization I
Project 2: Constrained optimization

Anton Evgrafov

April 2, 2018

1 Introduction

In this project we will work with the same least-squares type objective function as in Project 1. For simplicity we will only focus on the convex model 2 from project 1. Additionally, we will restrict the set of separating curves we are interested in. Namely we will require, that the set

$$\hat{S}_{A,b} := \{\, z \in \mathbb{R}^n \mid z^{\mathrm{T}} A z + b^{\mathrm{T}} z \leq 1 \,\}. \tag{1}$$

is an ellipsoid, which is neither too large nor too small. This can be expressed by introducing two positive bounds $0 < \underline{\lambda} < \overline{\lambda} < +\infty$ on the eigenvalues of the matrix A. Therefore, ideally we would like to consider the problem:

$$\begin{aligned} &\underset{x=(A,b)}{\text{minimize}}\, \hat{f}(x) = \sum_{i=1}^{m} \hat{r}_i(A,b)^2, \\ &\text{s.t.}\, \lambda_{\min}(A) \geq \underline{\lambda}, \\ &\quad\;\; \lambda_{\max}(A) \leq \overline{\lambda}, \end{aligned} \tag{2}$$

where $x \in \mathbb{R}^{n(n+1)/2+n}$, $\lambda_{\min}(\cdot)$ and $\lambda_{\max}(\cdot)$ denote the smallest and the largest eigenvalue of a given real symmetric matrix, and as before

$$\hat{r}_i(A,b) := \begin{cases} \max\{z_i^{\mathrm{T}} A z_i + b^{\mathrm{T}} z_i - 1, 0\}, & \text{if } w_i > 0, \\ \max\{1 - z_i^{\mathrm{T}} A z_i - b^{\mathrm{T}} z_i, 0\}, & \text{otherwise.} \end{cases} \tag{3}$$

Question 1

According to Courant–Fischer–Weyl theorem the smallest and the largest eigenvalues of a symmetric matrix A admit a variational characterization

$$\lambda_{\min}(A) = \min_{\|x\|_2=1} x^{\mathrm{T}} A x, \qquad \lambda_{\max}(A) = \max_{\|x\|_2=1} x^{\mathrm{T}} A x. \tag{4}$$

Using this characterization and the results from project 1, show that the problem (2) is convex.

2 Simplified model

Generally speaking, the functions $\lambda_{\min}(\cdot)$ and $\lambda_{\max}(\cdot)$ are non-smooth, and the problem (2) is somewhat beyond the scope of this course (cf.: Semidefinite programming). Instead, we will focus on a 2D case (that is, A is a 2×2 symmetric matrix) and further simplify the problem. From linear algebra we know that $\mathrm{tr}(A) = A_{11} + A_{22} = \lambda_1 + \lambda_2$, and $\det(A) = A_{11}A_{22} - A_{12}^2 = \lambda_1\lambda_2$, where $\lambda_1, \lambda_2 \in \mathbb{R}$ are the eigenvalues of A. Thus we can ensure that the eigenvalues will not be too small or too large by considering the following constraints:

$$\begin{aligned} &\underset{x=(A,b)}{\text{minimize}}\, \hat{f}(x), \\ &\text{s.t.} \begin{cases} \underline{\lambda} \le A_{11} \le \overline{\lambda}, \\ \underline{\lambda} \le A_{22} \le \overline{\lambda}, \\ (A_{11}A_{22})^{1/2} \ge (\underline{\lambda}^2 + A_{12}^2)^{1/2}, \end{cases} \end{aligned} \tag{5}$$

where the last constraints results from $\det(A) \ge \underline{\lambda}^2$.

Question 2

Show that the function $(A_{11}, A_{22}, A_{12}) \mapsto (A_{11}A_{22})^{1/2} - (\underline{\lambda}^2 + A_{12}^2)^{1/2}$ is twice continuously differentiable and concave on the set $\{\, (A_{11}, A_{22}, A_{12}) \in \mathbb{R}^3 \mid A_{11} > 0, A_{22} > 0 \,\}$.

Question 3

Rewrite the problem (5) in standard form. Show that the constraints satisfy Slater's constraint qualification. Conclude that KKT conditions are both necessary and sufficient for global optimality for this problem.

Question 4

Write down the KKT conditions for the problem (5) in standard form.

Question 5

Solve the problem (5) numerically. Illustrate the behaviour/verify correctness of your numerical strategy and its implementation on a range of test examples. (Generally speaking, the problem should be more difficult to solve if $\underline{\lambda} \to 0$.) Discuss the results in your report.

Note: you should aim at numerical methods that preserve the differentiability and concavity of the constraints, i.e., the diagonal coefficients of the matrix A should remain positive during the computation. Barrier and SQP (when started from the point satisfying the positive diagonal requirement) methods will satisfy this requirement.

Another possibility is to keep the linear constraints as is and apply the barrier for the non-linear constraint only. The resulting bound constrained subproblem can be solved e.g. by utilizing the projected gradient method. A similar approach can also be taken within the augmented Lagrangian framework.

3 Test problems

The guidelines given in project 1 apply. Try to come up with a test case, where the optimal solution to the unconstrained problem is not an ellipsoid — such a test case should illustrate the differences between the constrained and the unconstrained cases.