Problem Sheet 8

Problems in Part A will be discussed in class. Problems in Part B come with solutions and should be tried at home.

Part A

(8.1) Consider the general convex optimization problem

$$\text{minimize} \quad f(\boldsymbol{x}) \qquad \text{subject to} \quad \boldsymbol{f}(\boldsymbol{x}) \leq \boldsymbol{0}, \quad \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}.$$

The central path consists of the set of solutions $\boldsymbol{x}(t)$, $t > 0$, of the barrier problem

$$\text{minimize} \quad tf(\boldsymbol{x}) + \varphi(\boldsymbol{x}) \qquad \text{subject to} \quad \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b},$$

where $\varphi(\boldsymbol{x}) = -\sum_{i=1}^m \log(-f_i(\boldsymbol{x}))$ is the logarithmic barrier function. Show that a point $\boldsymbol{x}$ is equal to a point $\boldsymbol{x}^*(t)$ on the central path if and only if there exist dual multipliers $\boldsymbol{\lambda}^*$ and $\boldsymbol{\mu}^*$ such that the following conditions are satisfied:

$$\begin{aligned} \boldsymbol{f}(\boldsymbol{x}^*) &\leq \boldsymbol{0} \\ \boldsymbol{A}\boldsymbol{x}^* &= \boldsymbol{b} \\ \boldsymbol{\lambda}^* &\geq \boldsymbol{0} \\ -\lambda_i^* f_i(\boldsymbol{x}^*) &= \frac{1}{t}, \ 1 \leq i \leq m \\ \nabla_{\boldsymbol{x}} f(\boldsymbol{x}^*) + \sum_{i=1}^m \lambda_i^* \nabla_{\boldsymbol{x}} f_i(\boldsymbol{x}^*) + \boldsymbol{A}^\top \boldsymbol{\mu}^* &= \boldsymbol{0}, \end{aligned}$$

(8.2) Let $\{\boldsymbol{x}_1, \dots, \boldsymbol{x}_n\}$ be a series of data points with $\boldsymbol{x}_i \in \mathbb{R}^p$ for $1 \leq i \leq n$, and associated labels $\{y_1, \dots, y_n\}$ with $y_i \in \{-1, 1\}$. Consider the following version of the Support Vector Machine optimization problem that allows for few mistakes:

$$\begin{aligned} \text{minimize} \quad & \frac{1}{2}\|\boldsymbol{w}\|^2 + \mu \sum_{j=1}^n s_j \\ \text{subject to} \quad & y_i(\boldsymbol{w}^\top \boldsymbol{x}_i + b) - 1 + s_i \geq 0, \quad 1 \leq i \leq n \\ & s_i \geq 0, \quad 1 \leq i \leq n, \end{aligned}$$

Formulate the Lagrange dual and the KKT conditions for this problem. Show that the Lagrange dual does only depend on the inner products $\langle \boldsymbol{x}_i, \boldsymbol{x}_j \rangle$ of the data points.

(8.3) A matrix

$$\boldsymbol{A} = \begin{pmatrix} \boldsymbol{B} & \boldsymbol{v} \\ \boldsymbol{v}^\top & b \end{pmatrix},$$

is positive definite if and only if $b - \boldsymbol{v}^\top \boldsymbol{B}^{-1} \boldsymbol{v} \geq 0$. Use this, and the fact that a symmetric matrix factors as $\boldsymbol{A} = \boldsymbol{M}^\top \boldsymbol{M}$ for some $\boldsymbol{M}$, to show that the QCQP from Problem Sheet 7 can be formulated as a semidefinite programming problem.

Part B

(8.4) Given a symmetric matrix $\boldsymbol{A}$, formulate the problem of computing the largest eigenvalue $\lambda_{\max}(\boldsymbol{A})$ as a semidefinite programming problem.

(8.5) In many applications one is interested in finding a matrix of low rank that satisfies certain constraints. For example, one could have a covariance matrix, or a matrix containing user ratings of products, or a matrix whose entries are the squared distances between objects, but where only some entries are known. A common heuristic is to replace the rank of a symmetric matrix with the sum of the eiganvalues

(a) Show that for a symmetric matrix $\boldsymbol{A}$, the sum of the eigenvalues $\lambda_1 + \cdots + \lambda_n$ equals the trace $\mathrm{tr}(\boldsymbol{A})$. We can therefore write

$$\lambda_1 + \cdots + \lambda_n = \mathrm{tr}(\boldsymbol{A}) = \boldsymbol{I} \bullet \boldsymbol{A}.$$

(b) Formulate the problem of minimizing the trace of a symmetric positive semidefinite matrix $\boldsymbol{X}$ subject to constraints of the form

$$x_{ij} = a_{ij}$$

for some subset of indices $(i, j) \in \Omega \subseteq \{1, \ldots, n\}^2$. The problem is that of finding the matrix of smallest trace with some predetermined entries. Determine the dual of this problem.

(c) Using CVXPY in Python or CVX in MATLAB, perform the following experiment:

- Generate a random matrix $\boldsymbol{X}_0 \in \mathrm{SYM}_{100}$ of rank 10.
- For an increasing subcollection of "known" entries from $\boldsymbol{X}_0$, solve the trace minimization problem and determine if the solution of this optimization problem coincides with the matrix $\boldsymbol{X}_0$, thus effectively recovering it from only limited information.