Problem Sheet 9

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

Part A

(9.1) 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.

(9.2) A quadratically constraint quadratic problem (QCQP) has the form

$$\begin{aligned} \text{minimize} \quad & \frac{1}{2}\boldsymbol{x}^\top \boldsymbol{P}\boldsymbol{x} + \boldsymbol{q}^\top \boldsymbol{x} + r \\ \text{subject to} \quad & \frac{1}{2}\boldsymbol{x}^\top \boldsymbol{P}_i\boldsymbol{x} + \boldsymbol{q}_i^\top \boldsymbol{x} + r_i \leq 0, \ 1 \leq i \leq m, \end{aligned}$$

with $\boldsymbol{P}$ symmetric positive definite and $\boldsymbol{P}_1, \dots, \boldsymbol{P}_m$ symmetric positive semidefinite. Derive the Lagrange dual of this problem.

(9.3) 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 (9.2) can be formulated as a semidefinite programming problem.

Part B

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

(9.5) 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.