Take home exam -1

April 4, 2022

Personal work, to be posted on Teams for April 16 23:59 in pdf or handwritten.

Exercise 1. *We consider the set of $n \times n$ symmetric real matrices $S_n(\mathbb{R})$.*

1. *Show that $\langle A, B\rangle = \operatorname{tr}(AB)$ is a scalar product on S^n.*
2. *Show that the set of semi-definite positive matrices $K = S_n^+(\mathbb{R})$ is a cone.*
3. *Show that $K = S_n^+(\mathbb{R})$ is self dual (i.e. $K = K^{++}$ for this scalar product).*

Solution.

1. *It is symetric and bilinear.* $\operatorname{tr}(AA) = \sum_{ij} a_{ij}^2 = 0$ *implie* $A = 0$.
2. *Let A and B be in S^n_{++}, and $t > 0$ and $t' \geq 0$. Then we have, for all $x \in \mathbb{R}^n$, $x \neq 0$,*
$$x^\top (tA + t'B)x = tx^\top Ax + t'x^\top Bx > 0.$$
3. *Let $Y \in S_n \setminus S_n^+$. Then there exists $v \in \mathbb{R}^n$, $v^\top Yv < 0$. Moreover, $v^\top Yv = tr(v^\top Xv) = tr(v^\top vX) < 0$. Hence we have $X = v^\top v \in S_n^+$ such that $\langle Y, X\rangle < 0$, i.e. $Y \notin (S_n^+)^\star$.*
On the other hand, consider $Y \in S_n^+$. We have the following decomposition $Y = \sum_{i=1}^n \lambda_i q_i^\top q_i$, where $\lambda_i \geq 0$ are the eigenvalues, and q_i the associated eigenvectors. Thus, for any $X \in S_n^+$, we have
$$\langle Y, X\rangle = \operatorname{tr}(X \sum_{i=1}^n \lambda_i q_i^\top q_i) = \operatorname{tr}(\sum_{i=1}^n \lambda_i q_i^\top X q_i) \geq 0$$
hence $Y \in (S_n^+)^\star$

Exercise 2. *We consider the following problem.*

$$\underset{x_1,x_2}{\text{Min}} \quad x_1^2 + x_2^2 \tag{1}$$
$$s.t. \quad (x_1 - 1)^2 + (x_2 - 1)^2 \leq 1 \tag{2}$$
$$(x_1 - 1)^2 + (x_2 + 1)^2 \leq 1 \tag{3}$$

1. *Classify this problem.*
2. *Find the optimal solution and value of this problem.*
3. *Write and solve the KKT equation for this problem.*
4. *Derive and solve the Lagrangian dual of this problem.*

5. *Do we have strong duality ? If yes, could we have known it from the start ? If not, can you comment on why ?*

Solution. 1. *This is a convex QCQP*

2. *The only admissible point, and hence the optimal solution is* $(1, 0)$*, with value* 1.

3. *The Lagrangian is*

$$\mathcal{L}(x, \lambda) = x_1^2 + x_2^2 + \lambda_1((x_1 - 1)^2 + (x_2 - 1)^2 - 1) + \lambda_2((x_1 - 1)^2 + (x_2 + 1)^2 - 1)$$

KKT condition are

- *Gradient of Lagrangian is null :*

$$\begin{aligned} 2x_1 + 2\lambda_1(x_1 - 1) + 2\lambda_2(x_1 - 1) = 0 \\ 2x_2 + 2\lambda_1(x_2 - 1) + 2\lambda_2(x_2 + 1) = 0 \end{aligned}$$

- x *is primal feasible :* $(x_1 - 1)^2 + (x_2 - 1)^2 \leq 1$ *and* $(x_1 - 1)^2 + (x_2 + 1)^2 \leq 1$
- λ *is dual feasible* $\lambda_1 \geq 0$, $\lambda_2 \geq 0$.
- *Complementary slackness:*

$$\begin{aligned} \lambda_1 = 0 \quad & \textit{or} \quad (x_1 - 1)^2 + (x_2 - 1)^2 = 1 \\ \lambda_2 = 0 \quad & \textit{or} \quad (x_1 - 1)^2 + (x_2 + 1)^2 = 1 \end{aligned}$$

x *feasible is* $x = (1, 0)$*, which imply* $2 = 0$ *which is impossible. Thus there is no pair* (x, λ) *satisfying the KKT equations. The KKT equations fails to give the optimal solution because the constraints are not qualified.*

4. *The Lagrange dual function is*

$$\begin{aligned} g(\lambda_1, \lambda_2) &= \inf_{x_1, x_2} \mathcal{L}(x, \lambda) \\ &= \inf_{x_1, x_2} (1 + \lambda_1 + \lambda_2)(x_1^2 + x_2)^2 - 2(\lambda_1 + \lambda_2)x_1 - 2(\lambda_1 - \lambda_2)x_2 + \lambda_1 + \lambda_2 \\ &= \lambda_1 + \lambda_2 - \frac{(\lambda_1 + \lambda_2)^2 + (\lambda_1 - \lambda_2)^2}{1 + \lambda_1 + \lambda_2} \qquad \textit{if } 1 + \lambda_1 + \lambda_2 > 0 \end{aligned}$$

The dual problem reads

$$\begin{aligned} \underset{\lambda}{\text{Max}} \quad & \frac{\lambda_1 + \lambda_2 - (\lambda_1 - \lambda_2)^2}{1 + \lambda_1 + \lambda_2} \\ \textit{s.t.} \quad & \lambda_1 \geq 0, \lambda_2 \geq 0 \end{aligned}$$

By symmetry the optimum is attained at $\lambda_1 = \lambda_2$*, thus the dual reads*

$$\underset{\lambda_1 \geq 0}{\text{Max}} \quad \frac{2\lambda_1}{2\lambda_1 + 1}$$

Which has value 1 *and no solution.*

5. *The dual problem have the same value as the primal problem, thus we have strong duality.*

   However there does not exist a dual multiplier, which is why there is no solution to the KKT equations.

   We could not guarantee the existence of a primal-dual optimal solution through KKT as the constraints were not qualified.

Exercise 3. *We are going to prove that, if $f : \mathbb{R}^n \to \mathbb{R}$ is convex, and X a non-empty polytope (bounded polyhedron) then Kelley's cutting plane algorithm is converging.*
Consider $x_1 \in X$. We consider a sequence of points $(x^{(k)})_{k\in\mathbb{N}}$ such that $x^{(k+1)}$ is an optimal solution to

$$(\mathcal{P}^{(k)}) \qquad \underline{v}^{(k+1)} = \underset{x\in X}{\text{Min}} \quad z$$
$$s.t. \quad f(x^{(\kappa)}) + \langle g^{(\kappa)}, x - x^{(\kappa)} \rangle \leq z \qquad \forall \kappa \in [k]$$

where $g^{(k)} \in \partial f(x^{(k)})$.
Denote $v = \min_{x\in X} f(x)$.

1. *Show that v exists and is finite, and that there exists a sequence $x^{(k)}$.*
2. *Show that there exists L such that, for all k_1 and k_2, we have $\|f(x^{(k_1)}) - f(x^{(k_2)})\| \leq L\|x^{(k_1)} - x^{(k_2)}\|$, and $\|g^{(k)}\| \leq L$.*
3. *Let $K_\varepsilon = \{k \in \mathbb{N} \mid f(x^{(k)}) > v + \varepsilon\}$ be the set of index such that $x^{(k)}$ is not an ε-optimal solution. Show that $f(x_k) \to v$ if and only if K_ε is finite for all $\varepsilon > 0$*
4. *Consider $k_1, k_2 \in K_\varepsilon$, such that $k_2 > k_1$. Show that*
   $$f(x^{(k_1)}) + \langle g^{(k_1)}, x^{(k_2)} - x^{(k_1)} \rangle \leq \underline{v}^{(k_2)} \leq v$$
5. *Show that $\varepsilon + f(x^{(k_1)}) + \langle g^{(k_1)}, x^{(k_2)} - x^{(k_1)} \rangle < f(x^{(k_2)})$*
6. *Show that $\varepsilon < 2L\|x^{(k_2)} - x^{(k_1)}\|$.*
7. *Prove that $f(x^{(k)}) \to v$.*
8. *(Optional - hard) Find a complexity bound for the method (that is a number of iteration N_ε after which you are sure to have obtained a ε-optimal solution).*

Solution.

1. *f is finite convex and thus continuous on X which is compact, yielding the existence and finiteness of v.*

   f is subdifferentiable, thus we have the existence of $g^{(k)}$, and an optimal solution to $\mathcal{P}^{(k)}$ exists as the solution of a bounded linear programm.

2. *We have seen that on any compact K included in the domain of a convex function f, f is L-Lipschitz. Here $\text{dom}(f) = \mathbb{R}^n$, so on the compact $K = X + B(0, \varepsilon)$ f is L-Lipschitz, and on X any subgradient g is of norm lower than L.*

3. *$f(x_k) \to v$ iff $\forall \varepsilon > 0, \exists N_\varepsilon \in \mathbb{N}, \quad k \geq K \implies f(x_k) \leq v + \varepsilon$. Hence $K_\varepsilon \subset [N_\varepsilon]$.*

4. *By subgradient inequality* $f(y) \geq f(x^{(k)}) + \left\langle g^{(k)}, y - x^{(k)} \right\rangle$. *Thus, for all* k, $v \geq v^{(k)}$. *Further, note that* $v^{(k)} = f(x^{(k)})$, *hence using* $k = k_1$, *and* $y = x^{(k_2)}$ *we get*

$$f(x^{(k_1)}) + \left\langle g^{(k_1)}, x^{(k_2)} - x^{(k_1)} \right\rangle \leq \underline{v}^{(k_2)} \leq v.$$

5. *As* $k_2 \in K_\varepsilon$, *we have* $f(x^{(k_2)}) = v^{(k_2)} > v + \varepsilon \geq f(x^{(k_1)}) + \left\langle g^{(k_1)}, x^{(k_2)} - x^{(k_1)} \right\rangle + \varepsilon$ *by the previous question.*

6. *We have*

$$\varepsilon < |f(x^{(k_2)}) - f(x^{(k_1)})| + |\left\langle g^{(k_1)}, x^{(k_2)} - x^{(k_1)} \right\rangle| \leq 2L\|x^{(k_2)} - x^{(k_1)}\|$$

by Cauchy-Schwartz and question 2.

7. *If* $f(x^{(k)}) \not\to v$, *then there exists* $\varepsilon > 0$ *such that* $(x^{(k)})_{k \in K_\varepsilon})$ *is not finite. As* X *is compact we can exctract a converging subsequence, that is* $x^{(\sigma(k)}$ *such that* $x^{(\sigma(k))} \to x^\star$ *and* $\sigma(k) \in K_\varepsilon$, *which is in contradiction with the result of 6.*