Exercises: Constrained Optimization

Exercise 1 (Penalization). *We consider the following problem*

$$(P) \quad \min_{x\in\mathbb{R}^n} \quad f(x)$$
$$s.t. \quad Ax = b, \quad x \leq 0$$

with value v and the following penalized versions

$$(P_t^{in}) \quad \min_{x\in\mathbb{R}^n} f(x) - t\sum_{i=1}^n \ln(-x_i)$$
$$s.t. Ax = b, \quad x < 0$$

and

$$(P_t^{out}) \quad \min_{x\in\mathbb{R}^n} f(x) + t\sum_{i=1}^n (x_i)^+$$
$$s.t. \quad Ax = b$$

with associated value v_t^{in} and v_t^{out}, and an optimal solution x_t^{in} and x_t^{out}.

1. *Intuitively, assuming that f is "well behaved", for t going to which value does (P_t^{in}) tend to the original problem (P)? In which sense?*
2. *What can you say about x_t^{in}?*
3. *Can you compare v_t^{in} and v?*
4. *Same questions for (P_t^{out}).*

Answers: For t going to 0 we have that (P_t^{in}) tends toward (P) : in the sense that $v_t^{(in)} \to v$ and x_t goes toward an optimal solution . For t small enough we have $v_t^{in} \geq v$. In any case x_t^{in} is admissible.
For t going to $+\infty$, we have that (P_t^{out}) tends toward (P) in the sense that $v_t^{(out)} \to v$ and x_t^{out} goes toward an optimal solution. For t large enough, x_t^{out} is optimal for (P). We always have $v_t^{(out)} \leq v$.

Exercise 2 (Decomposition by prices). *We consider the following energy problem:*

- *you are an energy producer with N production units*
- *you have to satisfy a given demand planning for the next 24h (i.e. the total output at time t should be equal to d_t)*
- *the time step is the hour, and each unit have a production cost for each planning given as a convex quadratic function of the planning*
- *For each unit i, the production planning $u^i = (u_t^i)_{t\in[24]}$ has to satisfy polyhedral constraints $u^i \in U^i$.*

1. *Model this problem as an optimization problem. In which class does it belongs ? How many variables ?*
2. *Apply Uzawa's algorithm to this problem. Why could this be an interesting idea ?*
3. *Give an economic interpretation to this method.*
4. *What would happen if each unit had production constraints ?*

Exercise 3 (Kelley's convergence). *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

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

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*

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

5. *Show that $\varepsilon + f(x^{(k_1)}) + \left\langle g^{(k_1)}, x^{(k_2)} - x^{(k_1)} \right\rangle < f(x^{(k_2)})$*

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).*

Answers:

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 Kincluded in the domain of a convex function f, f is L-Lipschitz. Here $\mathrm{dom}(f) = \mathbb{R}^n$, 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 $K_\varepsilon \subset [N_\varepsilon]$. By the subgradient inequality, for any $\kappa \leq k$ and any $x \in X$,

$$f(x) \geq f(x^{(\kappa)}) + \langle g^{(\kappa)}, x - x^{(\kappa)} \rangle.$$

   Hence the cutting-plane model underestimates f on X, and therefore $\underline{v}^{(k)} \leq v$ for all k. Moreover, since $(x^{(k_2)}, \underline{v}^{(k_2)})$ is feasible for $\mathcal{P}^{(k_2-1)}$, it satisfies in particular the cut $\kappa = k_1$, giving

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

4. 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.

5. 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.

6. 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.