# Homework 2 - Analysis of High Dimensional Data

Let $f_i : \mathbb{R}^n \to \mathbb{R}$ and $P_i$ be bounded polyhedral subsets of $\mathbb{R}^n$ with non-empty intersection. We want to solve the following problem:

$$\textbf{PRIMAL:} \min \sum_{i=1}^{m} f_i(x), \quad x \in \mathbb{R}^n$$
$$\text{subject to } x \in P_i, i = 1, ..., m$$

Note that the constraint $x \in P_i \iff A_i x \leq b_i, A_i \in \mathbb{R}^{n_i \times n}, b_i \in \mathbb{R}^{n_i}$ and that these are all convex. Thus we can reformulate the above problem as:

$$\textbf{PRIMAL:} \min \sum_{i=1}^{m} f_i(x)$$
$$\text{subject to } A_i x \leq b_i, \quad i = 1, ..., m$$

Note that so long as $f_i(x) > -\infty \quad \forall x$ that are feasible, we can always find a maximum, say $p^*$. In addition, if the $f_i$ are convex, then $\sum_{i=1}^{m} f_i(x)$ is convex and this is a convex problem, which means that local minima are also global minima. This assumption is also need to apply KKT 1-4 and Slater's condition to show that strong duality holds as we will later see.

The Lagrangian is equal to

$$L(x, \lambda_1, ..., \lambda_m) = \sum_{i=1}^{m} f_i(x) + \sum_{i=1}^{m} \lambda_i'(A_i x - b_i), \quad \lambda_i \in \mathbb{R}^n$$

Thus the dual is

$$\begin{aligned} g(\lambda_1, ..., \lambda_m) &= \inf_x L(x, \lambda_1, ..., \lambda_m) \\ &= \inf_x \sum_{i=1}^{m} f_i(x) + \sum_{i=1}^{m} \lambda_i'(A_i x - b_i) \end{aligned}$$

Once way to minimize $L(x, \lambda_1, ..., \lambda_m)$ is to take derivative with respect to $x$. If $f_i$ is differentiable,

$$\begin{aligned} \nabla_x L(x, \lambda_1, ..., \lambda_m) &= \nabla_x \sum_{i=1}^{m} f_i(x) + \sum_{i=1}^{m} \lambda_i'(A_i x - b_i) \\ &= \sum_{i=1}^{m} \nabla_x f_i(x) + \sum_{i=1}^{m} A_i' \lambda_i \end{aligned}$$

Thus let $x^*$ be such that $\sum_{i=1}^m \nabla_x f_i(x^*) + \sum_{i=1}^m A_i'\lambda_i = 0$. This also implies that KKT-4 holds. If $f$ is not differentiable then we have to repeat the same argument using subdifferentials and show that 0 is in the set of subdifferentials. It is clear that KKT-1 holds as $A_i x - b_i \leq 0 \quad \forall i = 1, ..., m$.

The dual problem can be formulated as:

$$\textbf{DUAL:} \max g(\lambda_1, ..., \lambda_m)$$
$$\text{subject to } \lambda_i \succeq 0, \quad i = 1, ..., m$$

Note that the dual problem is always concave and hence a local maximum is a global maximum, although there may be more than one global maxima. Let's call this $d^*$. From the formulation, it is clear that KKT-2 holds as well. Finally, we need to show that KKT-3, or complementary slackness, holds. For this we consider 2 situations - one where $x \in (\cap_{i=1}^m P_i)^o$ and the other where $x \in \partial(\cap_{i=1}^m P_i)$. Here $A^o$ denotes the interior of A and $\partial(A)$ denotes the boundary of A. If $x \in (\cap_{i=1}^m P_i)^o$ and $f_i$ is convex for $i = 1, ..., m$ we can apply Slater's condition directly. Thus strong duality holds, i.e. the duality gap is 0 and there is no need to apply KKT conditions. In the other case when $x \in \partial(\cap_{i=1}^m P_i)$, suppose we have that $A_i x = b_i, \ i = 1, ..., k$ and $A_i x \leq b_i, \ i = k+1, ..., m$. One way to ensure complementary slackness holds is to set $\lambda_i = 0 \in \mathbb{R}^n, \ i = k+1, ..., m$, which makes these into non-binding constraints. Recall that to show that there is no duality gap in this case we need KKT 1-4 in addition to $f_i$ being convex and differentiable. These are all satisfied if we pick feasible $(x, \{\lambda_i\}_{i=1}^m)$ according to the primal and dual problems (KKT 1-2) such that the following additional conditions hold:

1. $\sum_{i=1}^m \nabla_x f_i(x) + \sum_{i=1}^m A_i'\lambda_i$=0. (KKT-4)
2. $\lambda_i = 0 \in \mathbb{R}^n$ whenever $A_i x \leq b_i$. (KKT-3)