Exercise Sheet 6 (due date for Exercises 6.1 - 6.3: Feb. 04)

Exercise 6.1 (Homework)

The goal of this exercise is to study a primal-dual path-following method for Linear Programming.

Remark: *This method can also be adapted to conic programs over a symetric cone $K \subset \mathbb{R}^n$ (such as SOCPs and SDPs), but this requires to introduce the notion of* Jordan algebra*, in which $\mathbb{R}^n$ is equipped with a particuliar symmetric vector product $\circ : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ such that $K = \{\boldsymbol{x} \circ \boldsymbol{x} : \boldsymbol{x} \in \mathbb{R}^n\}$. For the case of LPs, the cone is $K = \mathbb{R}^n_+$, and the product is simply the elementwise product over $\mathbb{R}^n$: $\boldsymbol{x} \circ \boldsymbol{y} := \boldsymbol{x} \odot \boldsymbol{y}$, where*

$$(\boldsymbol{x} \odot \boldsymbol{y})_i := x_i y_i, \quad \forall i \in [n].$$

We consider a pair of primal and dual LPs:

$$p^* = \inf\{\boldsymbol{c}^T\boldsymbol{x} |\ A\boldsymbol{x} = \boldsymbol{b}, \boldsymbol{x} \geq \boldsymbol{0}\}, \qquad d^* = \sup\{\boldsymbol{b}^T\boldsymbol{y} |\ A^T\boldsymbol{y} + \boldsymbol{s} = \boldsymbol{c}, \boldsymbol{s} \geq \boldsymbol{0}\}.$$

For $t > 0$, denote by $\boldsymbol{x}^*(t), \boldsymbol{y}^*(t)$ and $\boldsymbol{s}^*(t)$ the central path variables at t:

$$\boldsymbol{x}^*(t) = \operatorname*{argmin}_{\boldsymbol{x}}\{t\boldsymbol{c}^T\boldsymbol{x} - \sum_i \log(x_i) |\ A\boldsymbol{x} = \boldsymbol{b}\}, \qquad \boldsymbol{y}^*(t), \boldsymbol{s}^*(t) = \operatorname*{argmax}_{\boldsymbol{y},\boldsymbol{s}}\{t\boldsymbol{b}^T\boldsymbol{y} + \sum_i \log(s_i) |\ A^T\boldsymbol{y} + \boldsymbol{s} = \boldsymbol{c}\}.$$

It can be shown that $\boldsymbol{x}^*(t), \boldsymbol{y}^*(t), \boldsymbol{s}^*(t)$ are the solutions of the KKT system $\begin{cases} A\boldsymbol{x} = \boldsymbol{b} \\ A^T\boldsymbol{y} + \boldsymbol{s} = \boldsymbol{c} \\ \boldsymbol{x} \odot \boldsymbol{s} = \frac{1}{t}\boldsymbol{1} \\ \boldsymbol{x}, \boldsymbol{s} > \boldsymbol{0}. \end{cases}$

Primal-Dual path following methods work by maintaining primal and dual strictly feasible variables $\boldsymbol{x}_k, \boldsymbol{y}_k, \boldsymbol{s}_k$ that are *not too far from* the central path variables at t_k. The condition of complementary slackness holds for $t_k \to \infty$. Formally, for a parameter $\theta > 0$, the notion of neighborhood of the central path is defined by

$$(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{s}) \in \mathcal{N}_\theta := \Big\{(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{s}) : \boldsymbol{x} > \boldsymbol{0}, \boldsymbol{s} > \boldsymbol{0}, A\boldsymbol{x} = \boldsymbol{b}, A^T\boldsymbol{y} + \boldsymbol{s} = \boldsymbol{c}, \|\boldsymbol{x} \odot \boldsymbol{s} - \mu(\boldsymbol{x}, \boldsymbol{s})\boldsymbol{1}\| \leq \boldsymbol{\theta}\mu(\boldsymbol{x}, \boldsymbol{s})\Big\},$$

where $\mu(\boldsymbol{x}, \boldsymbol{s}) := \frac{\boldsymbol{x}^T\boldsymbol{s}}{n}$.

1. Show that $\boldsymbol{x} \odot \boldsymbol{s} = \frac{1}{t}\boldsymbol{1} \implies \frac{1}{t} = \mu(\boldsymbol{x}, \boldsymbol{s})$.
2. Let $(\boldsymbol{x}_k, \boldsymbol{y}_k, \boldsymbol{s}_k) \in \mathcal{N}_\theta$, and define $\epsilon = n\mu(\boldsymbol{x}_k, \boldsymbol{s}_k)$. Show that $\boldsymbol{x}_k$ and $\boldsymbol{y}_k$ are ϵ-suboptimal for the primal and dual LPs, respectively.
3. At iteration k, assume $(\boldsymbol{x}_k, \boldsymbol{y}_k, \boldsymbol{s}_k) \in \mathcal{N}_\theta$, and the iterates are updated as $\boldsymbol{x}_{k+1} := \boldsymbol{x}_k + \Delta\boldsymbol{x},\ \boldsymbol{y}_{k+1} := \boldsymbol{y}_k + \Delta\boldsymbol{y},\ \boldsymbol{s}_{k+1} := \boldsymbol{s}_k + \Delta\boldsymbol{s}$, where the directions $\Delta\boldsymbol{x}, \Delta\boldsymbol{y}, \Delta\boldsymbol{s}$ are computed by solving the linear system

$$\begin{cases} A\Delta\boldsymbol{x} &= \boldsymbol{0} \\ A^T\Delta\boldsymbol{y} + \Delta\boldsymbol{s} &= \boldsymbol{0} \\ \boldsymbol{s}_k \odot \Delta\boldsymbol{x} + \boldsymbol{x}_k \odot \Delta\boldsymbol{s} &= \tau_{k+1}\boldsymbol{1} - \boldsymbol{x}_k \odot \boldsymbol{s}_k. \end{cases}$$

   where $\tau_{k+1} := (1 - \frac{\delta}{\sqrt{n}})\mu(\boldsymbol{x}_k, \boldsymbol{s}_k)$ for some parameter $\delta \in (0, 1)$. Show that

$$\mu(\boldsymbol{x}_{k+1}, \boldsymbol{s}_{k+1}) = (1 - \frac{\delta}{\sqrt{n}})\mu(\boldsymbol{x}_k, \boldsymbol{s}_k).$$

For the last question, you can use the following result, without proving it: *For well chosen values of δ and θ (e.g. $\delta = \theta = 0.4$), we have*

$$\boldsymbol{s}_{k+1} > \boldsymbol{0},\ \boldsymbol{x}_{k+1} > \boldsymbol{0},\ \|\boldsymbol{x}_{k+1} \odot \boldsymbol{s}_{k+1} - \mu(\boldsymbol{x}_{k+1}, \boldsymbol{s}_{k+1})\boldsymbol{1}\| \leq \boldsymbol{\theta}\mu(\boldsymbol{x}_{k+1}, \boldsymbol{s}_{k+1}).$$

4. Assume the algorithm is initialized with $(\boldsymbol{x}_0, \boldsymbol{y}_0, \boldsymbol{s}_0) \in \mathcal{N}_\theta$, and define $t_0 := 1/\mu(\boldsymbol{x}_0, \boldsymbol{s}_0)$. Show that, if the parameters θ and δ are *well-chosen* (according to the above result), then $(\boldsymbol{x}_k, \boldsymbol{y}_k, \boldsymbol{s}_k) \in \mathcal{N}_\theta, \forall k \in \mathbb{N}$, and give an asymptotic bound (depending on t_0, $n \to \infty$ and $\epsilon \to 0$) for the number of iterations required to obtain ϵ-suboptimal primal and dual solutions.

Exercise 6.2 (Homework)

The S-lemma – Part 1

The goal of this exercise –and the following one– is to prove the *S-lemma*:
Let A, B be symmetric matrices, and assume that $\exists \boldsymbol{x}_0 : \boldsymbol{x}_0^T A \boldsymbol{x}_0 > 0$. Then,

$$\Big(\boldsymbol{x}^T A \boldsymbol{x} \geq 0 \implies \boldsymbol{x}^T B \boldsymbol{x} \geq 0\Big) \iff \Big(\exists \lambda \geq 0 : B - \lambda A \succeq 0\Big). \tag{1}$$

Remark: The condition on the right hand side can be equivalently written as $\inf_{\{\boldsymbol{x}:\boldsymbol{x}^T A\boldsymbol{x}\geq 0\}} \{\boldsymbol{x}^T A \boldsymbol{x}\} \geq 0.$

This result is useful to derive the robust counterpart of quadratic optimization problems, and can be interpreted as a *constraint qualification* for **nonconvex** quadratic problems with a single quadratic constraint.

1. One of the two directions of the equivalence (1) is easy to show. Which one ?

2. To prove the other implication, we consider the following SDP:

$$\begin{aligned} p^* = \inf \quad & \langle B, X \rangle \\ s.t. \quad & \langle A, X \rangle \geq 0 \\ & \langle I, X \rangle = 1 \\ & X \succeq 0. \end{aligned} \tag{P}$$

   Show that if Problem (P) is feasible, then it must have an optimal solution X^*.

3. Show that if $p^* = \langle B, X^* \rangle \geq 0$ and $\exists \boldsymbol{x}_0 : \boldsymbol{x}_0^T A \boldsymbol{x}_0 > 0$, then the right hand side of (1) holds.

Exercise 6.3 (Homework)

Proof of the S-lemma – Part 2

1. Let X^* be the optimal solution of Problem (P) (cf. Exercise 6.2). Define $\bar{A} = (X^*)^{1/2} A (X^*)^{1/2}$ and $\bar{B} = (X^*)^{1/2} B (X^*)^{1/2}$, and consider an eigenvalue decomposition $\bar{A} = U \Lambda U^T$. Show that if $\boldsymbol{z} \in \{-1, 1\}^n$, then $(U\boldsymbol{z})^T \bar{A} (U\boldsymbol{z}) \geq 0$.

2. Compute the value of the matrix $Z = \frac{1}{2^n} \sum_{\boldsymbol{z} \in V} \boldsymbol{z}\boldsymbol{z}^T$, where the sum goes over the 2^n vertices ($V = \{\boldsymbol{z}_1, \ldots, \boldsymbol{z}_{2^n}\}$) of the hypercube $\{-1, 1\}^n$. *Hint: You can use a statistical reasonning.*

3. Compute the value of the quantity
$$\frac{1}{2^n} \sum_{\boldsymbol{z} \in V} (U\boldsymbol{z})^T \bar{B} (U\boldsymbol{z}),$$
   and conclude that if the condition at the LHS of (1) holds, then we have $p^* \geq 0$.