Convex Optimization Exam, December 2020.

Please send your exam and code to `dm.daspremont@gmail.com` *by* **December 2, 2020 at 14:00.** *Late exams will not be graded.*

Exercise 1 (3 points) Write the following program as a LP and derive its dual,

$$\begin{array}{ll} \text{minimize} & \|x\|_\infty \\ \text{subject to} & Ax = b \end{array}$$

in the variable $x \in \mathbf{R}^n$.

Exercise 2 (3 points) Let D_{kl} be the Kullback-Leibler divergence. Prove the *information inequality*:

$$D_{\text{kl}}(u,v) \geq 0 \quad \text{for all } u,\ v \in \mathbf{R}^n_{++}.$$

Also show that $D_{\text{kl}}(u,v) = 0$ if and only if $u = v$.

Hint. The Kullback-Leibler divergence can be expressed as

$$D_{\text{kl}}(u,v) = f(u) - f(v) - \nabla f(v)^T (u - v),$$

where $f(v) = \sum_{i=1}^n v_i \log v_i$ is the negative entropy of v.

Exercise 3 (3 points) Show that the following program

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & x^T (A - bb^T) x \leq 0 \\ & b^T x \geq 0 \\ & Dx = g \end{array}$$

in the variable $x \in \mathbf{R}^n$, where $A \in \mathbf{S}_n$ and $A \succeq 0$, is convex. Compute a dual. *Hint: write it first as a Second Order Cone Program, with $A = V^T V$ for some $V \in \mathbf{S}_n$ because $A \succeq 0$.*

Exercise 4 (3 points) Compute a non trivial dual for the analytic centering problem

$$\text{minimize} - \sum_{i=1}^m \log(b_i - a_i^T x)$$

in the variable $x \in \mathbf{R}^n$, on the domain $\{x \in \mathbf{R}^n : a_i^T x < b_i\}$.

Exercise 5 (3 points) *A penalty method for equality constraints.* We consider the problem

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & Ax = b, \end{array} \tag{1}$$

where $f_0 : \mathbf{R}^n \to \mathbf{R}$ is convex and differentiable, and $A \in \mathbf{R}^{m \times n}$ with $\mathbf{Rank}\, A = m$.

In a *quadratic penalty method*, we form an auxiliary function

$$\phi(x) = f(x) + \alpha \|Ax - b\|_2^2,$$

where $\alpha > 0$ is a parameter. This auxiliary function consists of the objective plus the *penalty term* $\alpha \|Ax - b\|_2^2$. The idea is that a minimizer of the auxiliary function, $\tilde{x}$, should be an approximate solution of the original problem. Intuition suggests that the larger the penalty weight α, the better the approximation $\tilde{x}$ to a solution of the original problem.

Suppose $\tilde{x}$ is a minimizer of ϕ. Show how to find, from $\tilde{x}$, a dual feasible point for (1). Find the corresponding lower bound on the optimal value of (1).

Exercise 6 (5 points) Given m data points $x_i \in \mathbf{R}^n$ with labels $y_i \in \{-1, 1\}$. The goal of this exercise is to write a function to solve the classification problem

$$\begin{array}{lll} \text{minimize} & \frac{1}{2}\|w\|_2^2 + C\mathbf{1}^T z & \\ \text{subject to} & y_i(w^T x_i) \geq 1 - z_i, & i = 1, \ldots, m \\ & z \geq 0 & \end{array}$$

in the variables $w \in \mathbf{R}^n$, $z \in \mathbf{R}^m$.

- Derive the dual.
- Use the barrier method to solve both primal and dual problems up to arbitrary precision ϵ. Test your code on random clouds of points (e.g. generate two classes of data points by picking two bivariate Gaussian samples with different moments). Try various values of $C > 0$, e.g. $(10^{-1}, 1, 10, ...)$.
- Plot convergence of the objective function in Newton iterations for the barrier problem (i.e. inner iterations), with respect to the best value found, in semilog scale.

NOTE: Please use *graphics and tables* to illustrate your results as much as possible. Use general purpose languages such as PYTHON, JULIA or MATLAB. You can either return a compressed directory containing your code and graphics or a single jupyter notebook detailing your work.