MTH 377/577 CONVEX OPTIMIZATION

Winter Semester 2023

Indraprastha Institute of Information Technology Delhi

END-SEMESTER EXAM

(Time: 1 hour 15 min, Total Points: 30)

Q1. (10 points). Write a detailed solution to the following optimization problem

$$\begin{array}{lll} \max & \log(u_1 - 2) + \log(u_2 - 1) & \\ \text{subject to} & u_1 + 2u_2 \leq 12 & (\lambda_1) \\ & 2u_1 + u_2 \leq 12 & (\lambda_2) \end{array}$$

A1. Bring the problem to the standard form:

$$\begin{array}{llll} \min & f_0(u_1, u_2) := -\log(u_1 - 2) - \log(u_2 - 1) & & \\ \text{subject to} & f_1(u_1, u_2) := u_1 + 2u_2 - 12 \leq 0 & (\lambda_1) & (PF1) \\ & f_2(u_1, u_2) := 2u_1 + u_2 - 12 \leq 0 & (\lambda_2) & (PF2) \end{array}$$

Since $\log(u_i - d_i)$ is a concave transformation of an affine function, it is concave. Therefore, f_0 is a convex function as the sum of two convex functions. f_1 and f_2 are both affine functions of (u_1, u_2), so they are convex as well. All the functions involves are differentiable. Therefore this is a differentiable convex optimization problem. The point $(3, 2)$ is strictly feasible, so Slater's condition holds and strong duality obtains. This means KKT conditions are necessary an sufficient for optimality and are given by

$$\begin{array}{rl} (PF1),(PF2) & \\ \lambda_1 \geq 0 & (DF1) \\ \lambda_2 \geq 0 & (DF2) \\ \lambda_1(u_1 + 2u_2 - 12) = 0 & (CS1) \\ \lambda_2(2u_1 + u_2 - 12) = 0 & (CS2) \\ -\dfrac{1}{u_1 - 2} + \lambda_1 + 2\lambda_2 = 0 & (LO1) \\ -\dfrac{1}{u_2 - 1} + 2\lambda_1 + \lambda_2 = 0 & (LO2) \end{array}$$

From (LO1) and (LO2), we have that

$$\lambda_1 + 2\lambda_2 = \frac{1}{u_1 - 2} > 0 \qquad (LO1)$$

$$2\lambda_1 + \lambda_2 = \frac{1}{u_2 - 1} > 0 \qquad (LO2)$$

Therefore both λ_1 and λ_2 cannot be 0. Let us look for a primal-dual solution where $\lambda_1 > 0$ and $\lambda_2 > 0$. This implies, by (CS1) and (CS2), that (PF1) and (PF2) bind. That is

$$u_1 + 2u_2 = 12$$
$$2u_1 + u_2 = 12$$

which gives the primal point $(u_1, u_2) = (4, 4)$. (LO1) and (LO2) then give the dual point as $(\lambda_1, \lambda_2) = (1/18, 2/9)$. Since this primal-dual pair satisfies the KKT conditions, it is an optimal primal-dual pair.

Grading Comment: 2 points for reasoning that this is a convex optimization problem, 2 points for the conclusion of strong duality, 3 points for writing the KKT conditions, 3 points for computing the primal-dual solution pair.

Q2. (12 points). For the standard form of the optimization problem in Q1,

(a) (4 points). Write down the Lagrangian function $L(u_1, u_2, \lambda_1, \lambda_2)$. Given (u_1, u_2), is L is an affine function of (λ_1, λ_2)? Given (λ_1, λ_2), is L a convex function of (u_1, u_2)? Support your answer by adequate reasoning.

(b) (4 points). Find the dual function $g(\lambda_1, \lambda_2)$, and compute the value of $g(1, 1)$.

(c) (2 points). Which of the two constraints in problem of Q1 is the more active constraint? In what sense is it more active?

(d) (2 points). Suppose we perturb the constraints in problem of Q1 by the perturbation vector $(z_1, z_2) = (1, 1)$. This means the constraints now look like

$$u_1 + 2u_2 \leq 12 + 1 = 13$$
$$2u_1 + u_2 \leq 12 + 1 = 13$$

Give an upper bound on the magnitude by which the primal optimal value $p^\star$ can change.

A2. (a) The Lagrangian is given by

$$L(u_1, u_2, \lambda_1, \lambda_2) = -\log(u_1 - 2) - \log(u_2 - 1) + \lambda_1(u_1 + 2u_2 - 12) + \lambda_2(2u_1 + u_2 - 12)$$

Given (u_1, u_2), L is an affine function of (λ_1, λ_2) because it is of the form $a_1\lambda_1 + a_2\lambda_2 + b$, where a_1, a_2 and b are fixed constants.

Given (λ_1, λ_2), L is a convex function of (u_1, u_2). This is because,

$$L(u_1, u_2, \lambda_1, \lambda_2) = \underbrace{-\log(u_1 - 2) - \log(u_2 - 1)}_{\text{convex function of } (u_1,u_2)} + \underbrace{(\lambda_1 + 2\lambda_2)u_1 + (2\lambda_1 + \lambda_2)u_2 - 12(\lambda_1 + \lambda_2)}_{\text{affine function of } (u_1,u_2)}$$

Affine functions are convex as well. L is, therefore, convex as a sum of two convex functions.

Grading Comment: 1 point for writing down Lagrangian correctly, 1 point for the conclusion of affinity, 2 points for conclusion of convexity.

(b) The dual function $g(\lambda_1, \lambda_2)$ is the optimal value function of the unconstrained optimization problem

$$\min_{u_1,u_2} \qquad L(u_1, u_2, \lambda_1, \lambda_2)$$

The necessary and sufficient first order conditions of this problem are the Lagrangian optimality conditions (LO1) and (LO2) in Q1, which give the unconstrained minimum as

$$u_1 = \frac{1}{\lambda_1 + 2\lambda_2} + 2, \quad u_2 = \frac{1}{2\lambda_1 + \lambda_2} + 1$$

Substituting this in the Lagrangian, we get the dual function

$$\begin{aligned} g(\lambda_1, \lambda_2) &= \log(\lambda_1 + 2\lambda_2) + \log(2\lambda_1 + \lambda_2) \\ &\quad + \lambda_1\Big(\frac{1}{\lambda_1 + 2\lambda_2} + 2 + \frac{2}{2\lambda_1 + \lambda_2} + 2 - 12\Big) + \lambda_2\Big(\frac{2}{\lambda_1 + 2\lambda_2} + 4 + \frac{1}{2\lambda_1 + \lambda_2} + 1 - 12\Big) \end{aligned}$$

We can compute $g(1, 1) = 2\log(3) - 13$.

Grading Comment: 1 point if the student understands what a dual function is, 1 point for solving for u_1 and u_2, 1 point for writing the dual function and 1 point for computing $g(1, 1)$.

(c) The more active constraint is the one with the large value of dual variable. From the solution of Q1, we know $\lambda_2 > \lambda_1$, so the second constraint is the more active one. Using the marginal value interpretation of dual variables, this means that if we increase the right side of the second constraint by a little bit, the primal optimal value increases by more than if we were to increase the right side of the first constraint the same way.

Grading Comment: 1 point for identifying the second constraint and 1 point for the explanation.

(d) The magnitude by which the primal optimal value $p^\star$ can change is bounded above by $\lambda_1 z_1 + \lambda_2 z_2 = 1/18 + 2/9 = 5/18$.

Q3. (4 points). For the standard form of optimization problem in Q1,

(a) (2 points). Write down the parameterized log barrier function for the constraints and prove that it is a convex function.

(b) (2 points). Using the parameterized log barrier function, write down an approximate unconstrained optimization problem associated with the constrained problem of Q1. Despite being an approximation, why is this unconstrained problem a good basis for a descent based algorithm?

A3. (a) The parameterized log barrier function is given by

$$\frac{1}{t}B(u_1, u_2) := \frac{1}{t}\Big(-\log(12 - u_1 - 2u_2) - \log(12 - 2u_1 - u_2)\Big) \qquad \text{where } t > 0$$

Now $(u_1, u_2) \mapsto 12 - u_1 - 2u_2$ is an affine function; $-\log(12 - u_1 - 2u_2)$ is convex as a convex transformation of an affine function; $B(u_1, u_2)$ is convex as a sum of two convex functions; and finally, $\frac{1}{t}B(u_1, u_2)$ is convex as a positive scaling of a convex function.

Grading Comment: 1 point for writing down parameterized log barrier function and 1 point for the convexity explanation.

(b) An approximate unconstrained problem is

$$\min_{u_1,u_2} \quad -\log(u_1 - 2) - \log(u_2 - 1) + \frac{1}{t}\Big(-\log(12 - u_1 - 2u_2) - \log(12 - 2u_1 - u_2)\Big)$$

This is a good basis for an unconstrained descent based optimization algorithm because the objective function is differentiable and convex. So descent based methods can be expected to work well.

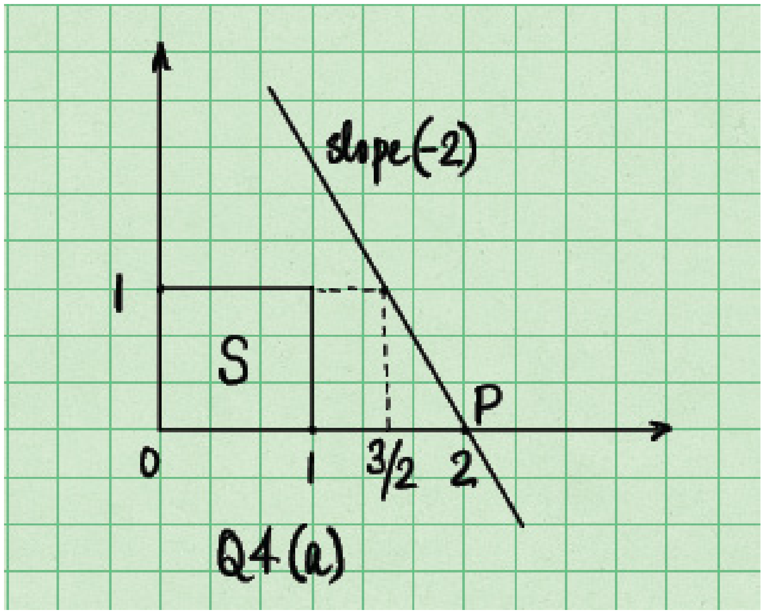
Grading Comment: 1 point for writing down the problem and 1 point for the explanation.

Q4. (a) (2 points) Let S be the square with vertices $(0, 0), (1, 0), (0, 1)$ and $(1, 1)$ and let P be the point $(2, 0)$. Is the line passing through the point $(3/2, 1)$ and slope -2 a separating hyperplane for the sets S and P ? If yes, identify this hyperplane with its (normal vector, scalar). If not, argue why not.

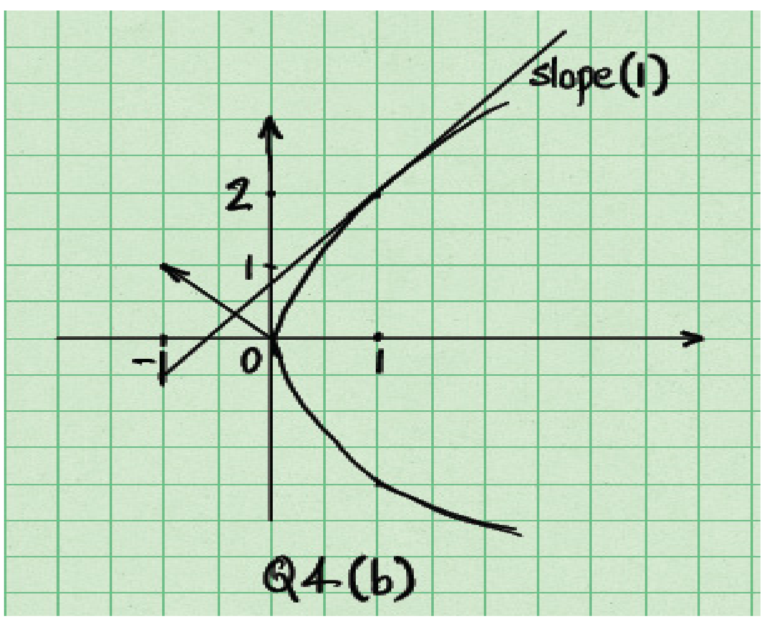
(c) (2 points) Does there exist a supporting hyperplane to the set $P = \{(x, y) \in \mathbb{R}^2 : y^2 \leq 4x\}$ at the point $(1, 2)$? If not, argue why not. If yes, argue why, identify the hyperplane with its (normal vector, scalar), and identify whether P lies in the positive or negative half space associated with the hyperplane.

A4. (a) The given line passes through P and a figure helps to see that it is indeed a separating hyperplane for S and P. Its equation is $2x + y = 4$. So its normal vector is $(2, 1)$ and its scalar is 4.

Grading: 1 point each for both questions.



(b) Yes, the supporting hyperplane exists because P is a convex set and $(1,2)$ is a point on the boundary of P. So the Supporting Hyperplane Theorem applies. In this case, it is tangent to P at this point. The slope of the tangent line to P at the point $(1,2)$ is 1. So the equation of this line is $y-2=1.(x-1)$ or $y-x=1$. This supporting hyperplane is then identified with the normal vector $a=(-1,1)$ and scalar $b=1$. Of course, as is clear from the figure, P lies in the negative halfspace associated with this hyperplane.



Grading: 1 point each for both the questions. It is possible someone describes the normal vector as $a=(1,-1)$ and scalar as $b=-1$. This is correct but in this case, P will lie in the positive halfspace.