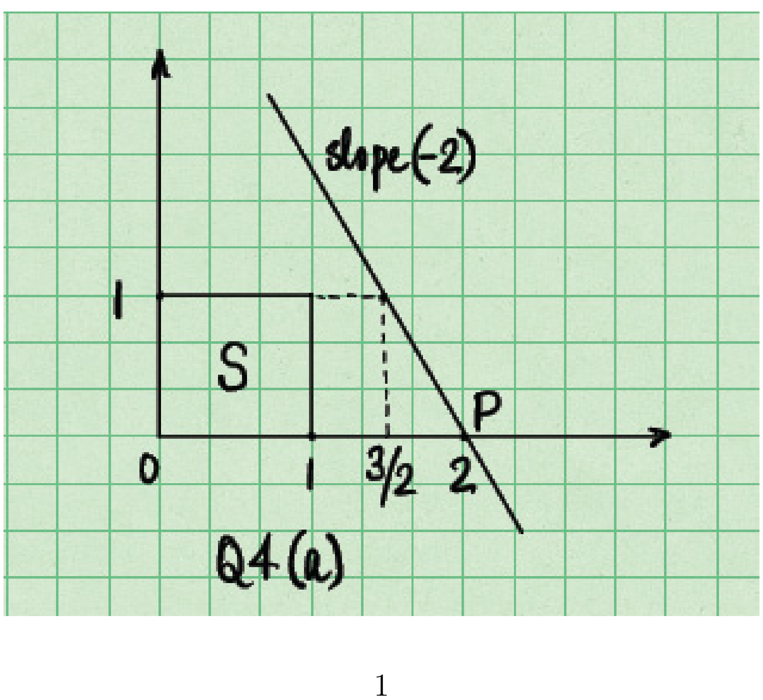
MTH 377/577 CONVEX OPTIMIZATION
Winter Semester 2022
Indraprastha Institute of Information Technology Delhi
END-SEMESTER EXAM
(Time: 2 hours, Total Points: 30)

Q1. (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.

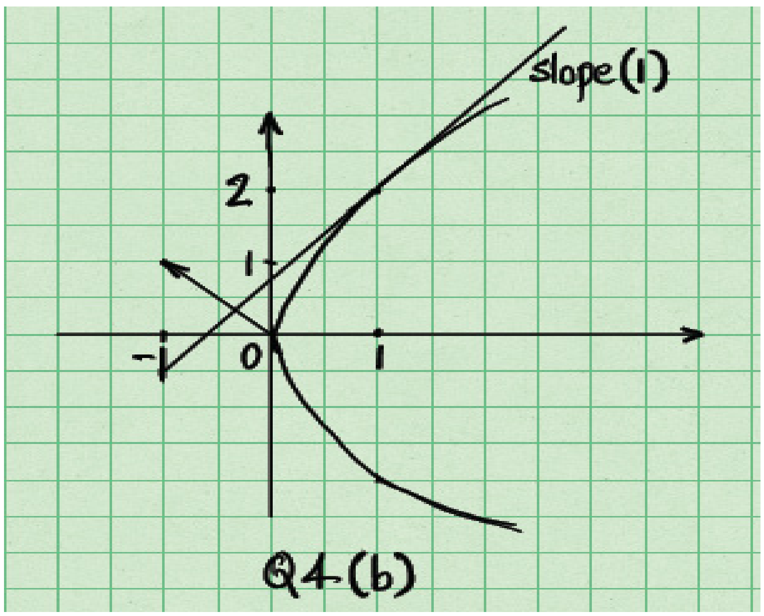
(c) (3 points) Does there exist a supporting hyperplane to the set $P=\{(x,y)\in\mathbb{R}^2 : y^2\leq 4x\}$ at the point $(1,2)$? If not, argue why not. If yes, argue why, identify the hyperplane with its (normal vector, scalar), and identify whether P lies in the positive or negative half space associated with the hyperplane.

A1. (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 all the three 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.

Q2. (5 points) Write the dual problem for the following linear program

$$\begin{aligned} \min \quad & 12x_1 + 3x_2 + 4x_3 \\ \text{subject to} \quad & 4x_1 + 2x_2 + 3x_3 \geq 2 \\ & 8x_1 + x_2 + 2x_3 \geq 3 \\ & x_1 \geq 0, \quad x_2 \geq 0, \quad x_3 \geq 0 \end{aligned}$$

A2. Rewrite the problem in the standard form

$$
\begin{array}{rll}
 & \min \quad 12x_1 + 3x_2 + 4x_3 & \\
\text{subject to} & 2 - 4x_1 - 2x_2 - 3x_3 \leq 0 & (\lambda_1) \\
 & 3 - 8x_1 - x_2 - 2x_3 \leq 0 & (\lambda_2) \\
 & -x_1 \leq 0 & (\lambda_3) \\
 & -x_2 \leq 0 & (\lambda_4) \\
 & -x_3 \leq 0 & (\lambda_5)
\end{array}
$$

The Lagrangian function is written as

$$
\begin{aligned}
L(x, \lambda) &= 12x_1 + 3x_2 + 4x_3 + \lambda_1(2 - 4x_1 - 2x_2 - 3x_3) + \lambda_2(3 - 8x_1 - x_2 - 2x_3) - \lambda_3 x_1 - \lambda_4 x_2 - \lambda_5 x_3 \\
&= (12 - 4\lambda_1 - 8\lambda_2 - \lambda_3)x_1 + (3 - 2\lambda_1 - \lambda_2 - \lambda_4)x_2 + (4 - 3\lambda_1 - 2\lambda_2 - \lambda_5)x_3 + 2\lambda_1 + 3\lambda_2
\end{aligned}
$$

The dual function is found as the solution to the unconstrained Lagrangian optimization problem and is given by

$$
\begin{aligned}
g(\lambda) &= \min_{x_1, x_2, x_3} L(x, \lambda) \\
&= \begin{cases} 2\lambda_1 + 3\lambda_2 \\ \text{if} \quad 12 - 4\lambda_1 - 8\lambda_2 - \lambda_3 = 0; 3 - 2\lambda_1 - \lambda_2 - \lambda_4 = 0; 4 - 3\lambda_1 - 2\lambda_2 - \lambda_5 = 0 \\ -\infty \\ \text{otherwise} \end{cases}
\end{aligned}
$$

The dual problem is $\max_{\lambda \geq 0} g(\lambda)$ which in this context is elaborated by making the domain constraints explicit, is given by

$$
\begin{array}{rl}
 & \max \quad 2\lambda_1 + 3\lambda_2 \\
\text{subject to} & 4\lambda_1 + 8\lambda_2 + \lambda_3 = 12 \\
 & 2\lambda_1 + \lambda_2 + \lambda_4 = 3 \\
 & 3\lambda_1 + 2\lambda_2 + \lambda_5 = 4 \\
 & \lambda_1 \geq 0 \\
 & \lambda_2 \geq 0 \\
 & \lambda_3 \geq 0 \\
 & \lambda_4 \geq 0 \\
 & \lambda_5 \geq 0
\end{array}
$$

An equivalent formulation in which λ_3, λ_4 and λ_5 are eliminated is given by

$$\begin{array}{rl} \max & 2\lambda_1 + 3\lambda_2 \\ \text{subject to} & 4\lambda_1 + 8\lambda_2 \leq 12 \\ & 2\lambda_1 + \lambda_2 \leq 3 \\ & 3\lambda_1 + 2\lambda_2 \leq 4 \\ & \lambda_1 \geq 0 \\ & \lambda_2 \geq 0 \end{array}$$

Grading. 1 point for writing the Lagrangian correctly, 2 points for writing the dual function correctly, 2 points for writing the dual problem correctly. If a student has not done the last equivalent reformulation, that may be excused.

Q3. (5 points) Write an optimality criterion for a differentiable concave maximization problem in standard form (irrespective of whether strong duality obtains). Use the criterion to derive a supporting hyperplane interpretation of optimality, taking care to identify the hyperplane. Draw an accompanying picture illustrating this interpretation.

A3. The geometric optimality criterion for differentiable concave maximization problem says that: $\mathbf{x} \in F$ is optimal if and only if the directional derivative in any feasible direction $\mathbf{y} - \mathbf{x}$ from $\mathbf{x}$ is nonpositive. Formally,

$$\forall \mathbf{y} \in F, \quad D_{\mathbf{y}-\mathbf{x}} f_0(\mathbf{x}) = \langle \nabla f_0(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle \leq 0 \tag{1}$$

(1) can be equivalently written as

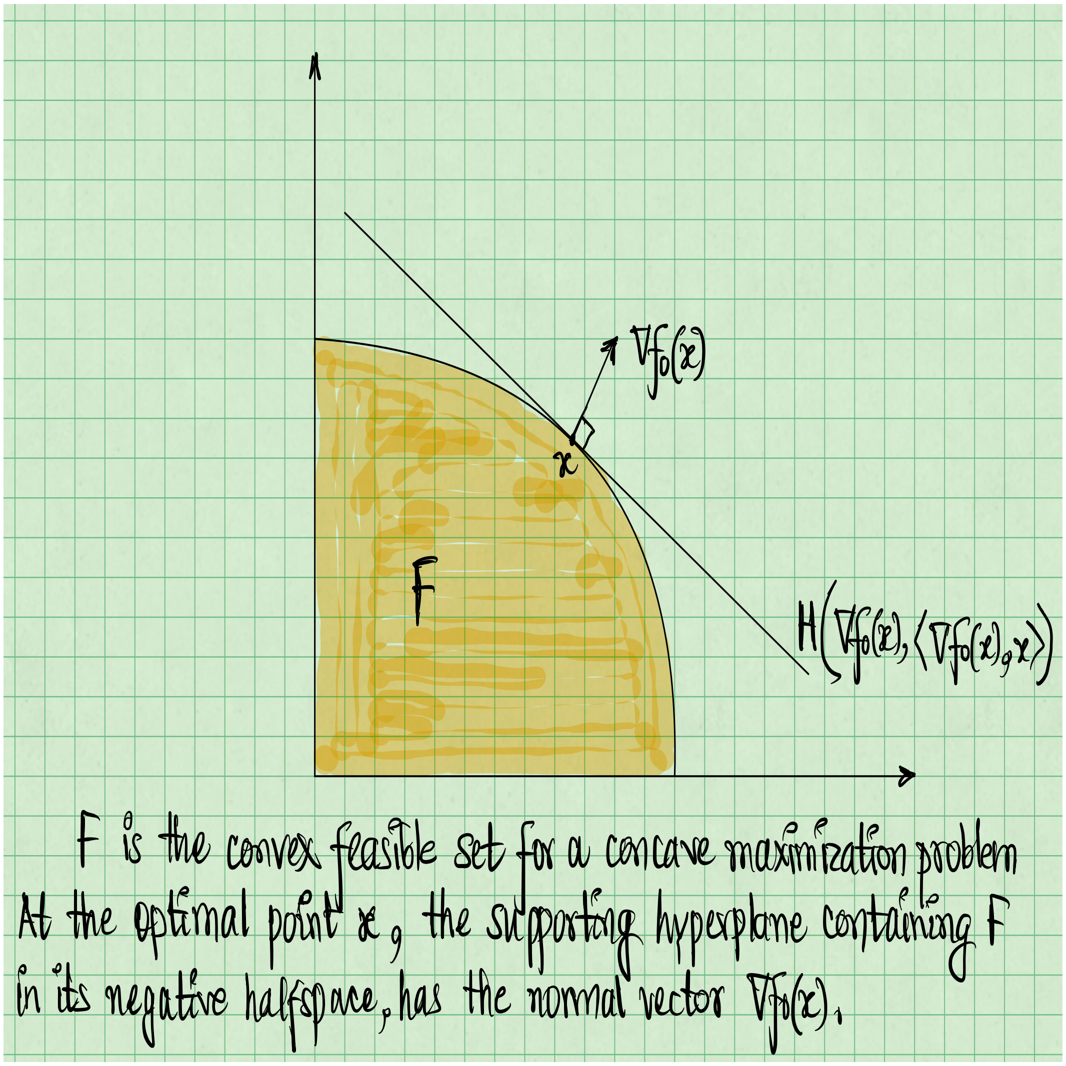
$$\forall \mathbf{y} \in F, \quad \langle \nabla f_0(\mathbf{x}), \mathbf{y} \rangle \leq \langle \nabla f_0(\mathbf{x}), \mathbf{x} \rangle \tag{2}$$

(2) is the basis for a supporting hyperplane interpretation of the optimality criterion (1) of nonpositive directional derivative. Writing (2) explicitly in terms of the associated hyperplane, we have

$$F \subseteq H^{-}\big(\nabla f_0(\mathbf{x}), \langle \nabla f_0(\mathbf{x}), \mathbf{x} \rangle\big) \tag{3}$$

$$\mathbf{x} \in F \cap H\big(\nabla f_0(\mathbf{x}), \langle \nabla f_0(\mathbf{x}), \mathbf{x} \rangle\big) \tag{4}$$

(3) and (4) remind us that an optimal point of a concave maximization problem must be a boundary point of the feasible set F. Moreover, it is that boundary point of F at which the supporting hyperplane which contains F in its negative halfspace has the normal vector equal to the gradient vector of the objective function. The accompanying figure illustrates this point.



Grading Comment: 1 point for optimality criterion (1), 3 points for deriving the interpretation and identifying the hyperplane. If a student shows $F \subset H^-$, then the normal vector should be $\nabla f_0(\mathbf{x})$ and the constant term should be $\langle \nabla f_0(\mathbf{x}), \mathbf{x} \rangle$; if a student shows $F \subset H^+$, then the normal vector should be $-\nabla f_0(\mathbf{x})$ and the constant term should be $\langle -\nabla f_0(\mathbf{x}), \mathbf{x} \rangle$. Finally, 1 point for the figure.

Q4. (5 points) Prove that $(x, y) = (2/\sqrt{3}, 3/2)$ is an optimal solution to

$$\begin{array}{rl} \max & 4x + 6y - x^3 - 2y^2 \\ \text{subject to} & x + 3y \leq 8 \\ & 5x + 2y \leq 14 \\ & x \geq 0, \quad y \geq 0 \end{array}$$

A4. Bring the problem to the standard form:

$$
\begin{array}{rlll}
\min & f_0(x,y) := -4x - 6y + x^3 + 2y^2 & & \\
\text{subject to} & f_1(x,y) := x + 3y - 8 \leq 0 & (\lambda_1) & (PF1) \\
& f_2(x,y) := 5x + 2y - 14 \leq 0 & (\lambda_2) & (PF2) \\
& f_3(x,y) := -x \leq 0 & (\lambda_3) & (PF3) \\
& f_4(x,y) := -y \leq 0 & (\lambda_4) & (PF4)
\end{array}
$$

Now, the Hessian matrix of the objective function is given by

$$D^2 f_0(x,y) = \begin{bmatrix} 6x & 0 \\ 0 & 4 \end{bmatrix}$$

The eigenvalues of the Hessian matrix $D^2 f_0(x,y)$ are $6x$ and 4. Both of these are nonnegative at any point (x,y) where $x \geq 0$. So f_0 is convex on $\mathbb{R}_+ \times \mathbb{R}$. The constraints are all affine. So this is a convex optimization problem for which Strong Duality obtains. This means KKT conditions are necessary an sufficient for optimality and are given by

$$
\begin{array}{rl}
-4 + 3x^2 + \lambda_1 + 5\lambda_2 - \lambda_3 = 0 & (LO1) \\
-6 + 4y + 3\lambda_1 + 2\lambda_2 - \lambda_4 = 0 & (LO2) \\
\lambda_1(x + 3y - 8) = 0 & (CS1) \\
\lambda_2(5x + 2y - 14) = 0 & (CS2) \\
\lambda_3 x = 0 & (CS3) \\
\lambda_4 y = 0 & (CS4) \\
\lambda = (\lambda_1, \lambda_2, \lambda_3, \lambda_4) \geq 0 & (DF) \\
\text{(PF) constraints} &
\end{array}
$$

The given point $(x,y) = (2/\sqrt{3}, 3/2)$ is an optimal solution to the problem iff it satisfies KKT conditions. Let's verify. (CS3) and (CS4) imply $\lambda_3 = \lambda_4 = 0$. Also at the given point, (PF1) and (PF2) are both slack; using (CS1) and $(CS2)$, this implies $\lambda_1 = \lambda_2 = 0$. We have deduced that $\lambda_1 = \lambda_2 = \lambda_3 = \lambda_4 = 0$. This gives a complete description of the primal-dual point, which satisfies (LO1) and (LO2), and thereby all the KKT conditions. Therefore, it is an optimum.

Grading Comment: 2 points for reasoning that this is a convex optimization problem with strong duality, 2 for writing the KKT conditions, 1 point for computing the associated dual point and KKT verification.

Q5. Consider a convex optimization problem in standard form as given by

$$
\begin{array}{rll}
\min & f_0(x) & \\
\text{subject to} & f_i(x) \leq 0 & \text{for every } i = 1, \ldots, m \\
& Ax - b = 0 &
\end{array}
$$

where f_0 as well as every f_i is a convex function on $\mathbb{R}^n$.

(a) (3 points) Consider the sets

$$\mathcal{A} = \{(u, v, t) \in \mathbb{R}^m \times \mathbb{R}^p \times \mathbb{R} : \forall i = 1, \ldots, m, f_i(x) \leq u_i; Ax - b = v; f_0(x) \leq t\}$$
$$\mathcal{B} = \{(0, 0, s) \in \mathbb{R}^m \times \mathbb{R}^p \times \mathbb{R} : s < p^\star\}$$

Show that $\mathcal{A}$ and $\mathcal{B}$ are convex sets.

(b) (4 points) Let $p^\star$ be the optimal value of the convex optimization problem described above. Use the separating hyperplane theorem to show that there exists $(\tilde{\lambda}, \tilde{\nu}, \mu) \neq 0$ such that

$$\forall x \in \mathcal{D}, \quad \sum_{i=1}^{m} \tilde{\lambda}_i f_i(x) + \tilde{\nu}^T(Ax - b) + \mu f_0(x) \geq \mu p^\star$$

(c) (3 points) Assuming $\mu > 0$, show that strong duality obtains for the convex optimization problem.

A5. (a) Let $f(x) = (f_1(x), \ldots, f_m(x))$. Fix (u, v, t) and (u', v', t') in $\mathcal{A}$ and θ in $[0, 1]$. Then

$$\exists x \in \mathcal{D} \quad f(x) \leq u \quad Ax - b = v \quad f_0(x) \leq t \tag{5}$$
$$\exists x' \in \mathcal{D} \quad f(x') \leq u' \quad Ax' - b = v' \quad f_0(x') \leq t' \tag{6}$$

Multiplying the first set of equations and inequalities by θ, the second set by θ', and adding up, we have

$$\forall i = 1, \ldots, m \quad \theta f_i(x) + (1 - \theta) f_i(x') \leq \theta u_i + (1 - \theta) u_i' \tag{7}$$
$$\theta Ax + (1 - \theta) Ax' - b = \theta v + (1 - \theta) v' \tag{8}$$
$$\theta f_0(x) + (1 - \theta) f_0(x') \leq \theta t + (1 - \theta) t' \tag{9}$$

Since f_0 as well as each f_i is convex, we have

$$\forall i = 1, \ldots, m \quad f_i(\theta x + (1 - \theta) x') \leq \theta f_i(x) + (1 - \theta) f_i(x') \tag{10}$$
$$A(\theta x + (1 - \theta) x') = \theta Ax + (1 - \theta) Ax' \tag{11}$$
$$f_0(\theta x + (1 - \theta) x') \leq \theta f_0(x) + (1 - \theta) f_0(x') \tag{12}$$

(10) - (11) - (12) and (7) - (8) - (9) imply

$$\forall i = 1, \ldots, m \quad f_i(\theta x + (1 - \theta) x') \leq \theta u_i + (1 - \theta) u_i' \tag{13}$$
$$A(\theta x + (1 - \theta) x') = \theta v + (1 - \theta) v' \tag{14}$$
$$f_0(\theta x + (1 - \theta) x') \leq \theta t + (1 - \theta) t' \tag{15}$$

Since $\theta x + (1 - \theta) x'$ is a point in $\mathcal{D}$ for which (13) - (14) - (15) hold, we have that $\theta(u, v, t) + (1 - \theta)(u', v', t') \in \mathcal{A}$. So $\mathcal{A}$ is convex.

Fix $(0,0,s)$ and $(0,0,s')$ in $\mathcal{B}$ and θ in $[0,1]$. Then $s < p^\star$ and $s' < p^\star$ implying $\theta s + (1-\theta)s' < p^\star$. So $(0,0,\theta s + (1-\theta)s') \in \mathcal{B}$. So $\mathcal{B}$ is convex.

Grading Comment: 2 points for convexity of $\mathcal{A}$, 1 point for convexity of $\mathcal{B}$.

(b) *Step 1.* In the first step, we show $\mathcal{A} \cap \mathcal{B} = \emptyset$. Suppose, by way of contradiction, $(u,v,t) \in \mathcal{A} \cap \mathcal{B}$. Then since $(u,v,t) \in \mathcal{A}$, there exists $x \in \mathcal{D}$ such that $f_0(x) \leq t$. Moreover, since $(u,v,t) \in \mathcal{B}$, $t < p^\star$. The conjunction of the conclusion in the preceding two statements imply that there exists $x \in \mathcal{D}$ such that $f_0(x) \leq p^\star$, which is a contradiction to the definition of $p^\star$ as the primal optimal value.

Step 2. From the first step, $\mathcal{A}$ and $\mathcal{B}$ are disjoint convex sets. By separating hyperplane theorem, there exists $(\tilde{\lambda}, \tilde{\nu}, \mu) \neq 0$ and $(\tilde{\lambda}, \tilde{\nu}, \mu) \in \mathbb{R}^m \times \mathbb{R}^p \times \mathbb{R}$ and $\alpha \in \mathbb{R}$ such that $A \subseteq H^+((\tilde{\lambda}, \tilde{\nu}, \mu), \alpha)$ and $B \subseteq H^-((\tilde{\lambda}, \tilde{\nu}, \mu), \alpha)$. This means

$$\forall (u,v,t) \in \mathcal{A}, \quad \tilde{\lambda}^T u + \tilde{\nu}^T v + \mu t \geq \alpha \tag{16}$$

$$\forall (u,v,t) \in \mathcal{B}, \quad \tilde{\lambda}^T u + \tilde{\nu}^T v + \mu t \leq \alpha \tag{17}$$

(17) is equivalently written as: $\forall t < p^\star$, $\mu t \leq \alpha$, implying that $\mu p^\star \leq \alpha$. Applying (16) for all $(u,v,t) \in \mathcal{A}$ for which $\exists x \in \mathcal{D}$ such that $(u,v,t) = (f(x), Ax - b, f_0(x))$, we have

$$\forall x \in \mathcal{D}, \quad \sum_{i=1}^m \tilde{\lambda}_i f_i(x) + \tilde{\nu}^T (Ax - b) + \mu f_0(x) \geq \alpha \geq \mu p^\star \tag{18}$$

Grading Comment: 1 points for Step 1 showing disjointness, 2 point for applying separating hyperplane theorem to conclude (16) - (17), 1 point for finally deducing (18) from (16) - (17).

(c) Divide (18) by μ to get

$$\forall x \in \mathcal{D}, \quad L\Big(x, \frac{\tilde{\lambda}}{\mu}, \frac{\tilde{\nu}}{\mu}\Big) \geq p^\star \tag{19}$$

Let $\lambda = \frac{\tilde{\lambda}}{\mu}$ and $\nu = \frac{\tilde{\nu}}{\mu}$, we have

$$g(\lambda, \nu) = \min_{x \in \mathcal{D}} L\Big(x, \frac{\tilde{\lambda}}{\mu}, \frac{\tilde{\nu}}{\mu}\Big) \geq p^\star \tag{20}$$

The conjunction of (19) and weak duality give us the strong duality conclusion.

Grading Comment: 1 points for (19), 1 point for (20), and 1 point for the final statement deducing strong duality.