EECS 127/227AT Optimization Models in Engineering
Spring 2020 Homework 9

This homework is due Friday, April 3, 2020 at 23:00 (11pm).
Self grades are due Friday, April 10, 2020 at 23:00 (11pm).

This version was compiled on 2020-04-04 15:46.

Submission Format: Your homework submission should consist of a single PDF file that contains all of your answers (any handwritten answers should be scanned) as well as your IPython notebook with solutions saved as a PDF.

1. Does strong duality hold?
Consider

$$\min_{(x,y)\in\mathcal{D}} e^{-x}$$
$$\text{s.t. } x^2/y \le 0$$

where $\mathcal{D} = \{(x,y) \mid y > 0\}$.

(a) Prove the problem is convex. Find the optimal value.
Hint: To prove the constraint function is convex, you will have to prove it is convex with respect to the vector $\begin{bmatrix} x & y \end{bmatrix}^\top$. *Consider computing the Hessian of the constraint function, its determinant and trace, and show that it is PSD by analyzing signs of its eigenvalues.*
Solution:
The second derivative of the objective function is e^{-x} which is non-negative thus the objective is a convex function. Furthermore, the constraint is jointly convex as can be verified by showing the Hessian is PSD (additionally one may notice this is the perspective function of x^2 which will be convex). The Hessian for $g(x,y) = \frac{x^2}{y}$ is given by,

$$\nabla^2 g(x,y) = \begin{bmatrix} \frac{2}{y} & -\frac{2x}{y^2} \\ -\frac{2x}{y^2} & \frac{2x^2}{y^3} \end{bmatrix}.$$

Suppose λ_1, λ_2 are the eigenvalues of $\nabla^2 g(x,y)$. The determinant of the Hessian is 0 which gives us $\lambda_1\lambda_2 = 0$. Further trace of the Hessian is $\frac{2}{y} + \frac{2x^2}{y^3} > 0$ (since $y > 0$) which gives $\lambda_1 + \lambda_2 > 0$. Thus one eigenvalue must be positive and the other must be 0 which shows that the Hessian is positive semidefinite.
Hence the problem is convex. Furthermore, the only feasible value of x is $x = 0$. Hence the optimal value is $e^{-0} = 1$.

(b) Next, we will proceed to find an optimal solution and an optimal value for the dual problem. The Lagrangian dual function $g(\lambda)$, can be written as:

$$g(\lambda) = \inf_{(x,y)\in\mathcal{D}} e^{-x} + \lambda\frac{x^2}{y}.$$

Explain why $g(\lambda)$ is lower bounded by 0 for $\lambda \geq 0$. *Note: Here we are not dualizing the constraint $y > 0$ that is in the definition of $\mathcal{D}$ — this is only dualizing the other constraint.*

Solution: This is true since both terms in the sum are non-negative because $y > 0$.

(c) Show that $g(\lambda) = 0$ for $\lambda \geq 0$.
Hint 1: To show that the infimum in ((b)) is 0, we want to show there exist (x, y) such that both e^{-x} and $\lambda \frac{x^2}{y}$ can get arbitrarily close to 0.
Hint 2: Consider a sequence $\{x_k\}$ going to $+\infty$ and a sequence $\{y_k\}$ also going to $+\infty$ such that $\lim_{k\to\infty} \frac{x_k^2}{y_k} = 0$. Simply put, we want to drive x to infinity in order to drive e^{-x} to 0, while having y grow faster than x^2, so that the second term also goes to 0.

Solution: To show that the infimum is 0, we pick any sequence $\{x_k\}$ going to $+\infty$ and pick a sequence $\{y_k\}$ also going to $+\infty$ such that $\lim_{k\to\infty} \frac{x_k^2}{y_k} = 0$. One example of such sequence pair is $y_k = x_k^4$ and $x_k = 2k$.
This gives $\lim_{k\to\infty} e^{-x} + \frac{x_k^2}{y_k} = 0$, so $g(\lambda) = \inf_{(x,y)\in\mathcal{D}} e^{-x} + \lambda\frac{x^2}{y} = 0$. Note that the (x_k, y_k) pair as $k \to \infty$ does not satisfy the constraint of the primal problem $x^2/y \leq 0$ but allows x^2/y to get arbitrarily close to 0.

(d) Now, write the dual problem and find an optimal solution λ^* and an optimal value d^* for the dual problem using the results above. What is the duality gap?

Solution: The Lagrange dual problem is

$$d^* = \sup_{\lambda\geq 0} g(\lambda)$$

where $g(\lambda) = \inf_{(x,y)\in\mathcal{D}} \; e^{-x} + \lambda x^2/y$. Note $g(\lambda) = 0$ from previous part. It follows from the previous problems that $d^* = 0$ and any $\lambda \geq 0$ is optimal. The duality gap is 1.

(e) Does Slater's Condition hold for this problem? Does Strong Duality hold?

Solution:
While the primal problem is convex, we cannot find a point that is strictly in the interior of the domain and satisfies the constraint as needed for Slater's condition. Specifically, for Slater's condition to hold we need the existence of an (x, y) pair such that $x^2/y < 0$. Note there is no such pair (x, y) since $y > 0$ and $x^2 \geq 0$. Hence Slater's condition does not hold for this problem.

Method 2:
From the previous parts we saw that $p^* \neq d^*$, and thus strong duality does not hold. Furthermore, the problem is convex. For convex problems we know that if Slater's condition holds then we must have strong duality (i.e Slater's is a sufficient condition). However since strong duality does not hold it implies that Slater's does not hold.

Note that this problem is an example illustrating that **convexity alone is not enough to guarantee strong duality for an optimization problem**.

2. Visualizing the dual problem

Download the Jupyter notebook `visualize_dual.ipynb`; complete the code where designated and answer the questions given in the space provided. (If you prefer, for questions that do not involve writing code, you can write solutions on separate paper or LaTeX PDF, just make sure to correctly mark the relevant pages when uploading to Gradescope.)

3. Sensitivity and dual variables

In this problem, we explore the interpretation of dual variables as sensitivity parameters of the primal problem. Recall the canonical **convex** primal problem

$$\begin{aligned} \min_{\vec{x}\in\mathbb{R}^n} \quad & f_0(\vec{x}) \\ \text{s.t.} \quad & f_i(\vec{x}) \leq 0, \quad i = 1, \ldots, m \\ & h_j(\vec{x}) = 0, \quad j = 1, \ldots, p \end{aligned}$$

where f_i are convex for all $i = 0, \ldots, m$ and h_j are affine for all $j = 1, \ldots, p$. **Assume strong duality holds.**

Here, we consider the *perturbed* problem

$$\begin{aligned} \min_{\vec{x}\in\mathbb{R}^n} \quad & f_0(\vec{x}) \\ \text{s.t.} \quad & f_i(\vec{x}) \leq u_i, \quad i = 1, \ldots, m \\ & h_j(\vec{x}) = v_j, \quad j = 1, \ldots, p \end{aligned}$$

and define

$$p^*(\vec{u}, \vec{v}) = \inf\{f_0(\vec{x}) \mid f_i(\vec{x}) \leq u_i \ \ \forall i, \ h_j(\vec{x}) = v_j \ \ \forall j\}$$

for perturbation vectors $\vec{u} = \begin{bmatrix} u_1 & \ldots & u_m \end{bmatrix}^\top$ and $\vec{v} = \begin{bmatrix} v_1 & \ldots & v_p \end{bmatrix}^\top$. In other words, $p^*(\vec{u}, \vec{v})$, **is a function of** $\vec{u}$ and $\vec{v}$ that gives the optimal value for the perturbed problem (if it is feasible). If the problem is infeasible (i.e. no points exist that satisfy the constraints), we say that $p^*(\vec{u}, \vec{v}) = +\infty$ otherwise. Note that $p^*(\vec{0}, \vec{0})$ is the original problem.

(a) Prove that $p^*(\vec{u}, \vec{v})$ is jointly convex [1] in $(\vec{u} \in \mathbb{R}^m, \ \vec{v} \in \mathbb{R}^p)$.
Hint: Let $\mathcal{D} = \{(\vec{x} \in \mathbb{R}^n, \vec{u} \in \mathbb{R}^m, \vec{v} \in \mathbb{R}^p) \mid f_i(\vec{x}) \leq u_i \ \ \forall i, \ h_j(\vec{x}) = v_j \ \ \forall j\}$ denote the feasible set for F. Now define $F(\vec{x}, \vec{u}, \vec{v})$ to be a function that is equal to $f_0(\vec{x})$ on $\mathcal{D}$ and $+\infty$ otherwise. Show $F(\vec{x}, \vec{u}, \vec{v})$ is convex and then observe that

$$p^*(\vec{u}, \vec{v}) = \min_{\vec{x}} F(\vec{x}, \vec{u}, \vec{v}).$$

Solution: The domain for $F(\vec{x}, \vec{u}, \vec{v})$ is convex because it is the intersection of epigraphs of convex functions $\{(\vec{x}, \vec{u}, \vec{v}) \mid f_i(\vec{x}) \leq u_i \ \ \forall i\}$ and hyperplanes $\{(\vec{x} \in \mathbb{R}^n, \vec{u} \in \mathbb{R}^m, \vec{v} \in \mathbb{R}^p) \mid \ h_j(\vec{x}) = v_j \ \ \forall j\}$. Furthermore, $f_0(\vec{x})$ is jointly convex in $(\vec{x}, \vec{u}, \vec{v})$ since it is convex in $\vec{x}$ and does not depend on $\vec{u}$ or $\vec{v}$. Hence F is convex.
Finally,

$$p^*(\vec{u}, \vec{v}) = \min_{\vec{x}} F(\vec{x}, \vec{u}, \vec{v})$$

is the pointwise minimization of a convex function — hence $p^*(\vec{u}, \vec{v})$ is convex in $(\vec{u}, \vec{v})$.

[1] Recall that a function $f : A \times B \to \mathbb{R}$ is jointly convex in $(\vec{a} \in A, \ \vec{b} \in B)$ if for all $\theta \in [0, 1]$, and for all $\vec{a}_1, \vec{a}_2 \in A, \ \vec{b}_1, \vec{b}_2 \in B$, we have that $f(\theta\vec{a}_1 + (1-\theta)\vec{a}_2, \ \theta\vec{b}_1 + (1-\theta)\vec{b}_2) \leq \theta f(\vec{a}_1, \vec{b}_1) + (1-\theta) f(\vec{a}_2, \vec{b}_2)$.

(b) Show that

$$p^*(\vec{u},\vec{v}) \geq p^*(\vec{0},\vec{0}) - \vec{\lambda}^{*\top}\vec{u} - \vec{\nu}^{*\top}\vec{v}$$

where $\vec{\lambda}^* \in \mathbb{R}^m$ and $\vec{\nu}^* \in \mathbb{R}^p$ are the optimal dual variables for the dual of the unperturbed primal problem (corresponding to inequality and equality constraints, respectively).
Hint: Consider the dual function $g(\vec{\lambda},\vec{\nu})$ for the perturbed problem at $\lambda = \lambda^*, \nu = \nu^*$ and upper-bound it by the Lagrangian at a feasible point $\vec{x}$. Here λ^* and ν^* are the optimal dual variables for the original unperturbed problem.

Solution:

Solution 1:

We let $g_{\vec{u},\vec{v}}$ and $\mathcal{L}_{\vec{u},\vec{v}}$ denote the dual function and Lagrangian for the perturbed problem respectively.

Now, we have:

$$p^*(\vec{u},\vec{v}) \geq g_{\vec{u},\vec{v}}(\vec{\lambda}^*,\vec{\nu}^*) = \inf_{\vec{x}} \mathcal{L}_{\vec{u},\vec{v}}(\vec{x},\vec{\lambda}^*,\vec{\nu}^*) = \inf_{\vec{x}} f_0(\vec{x}) + \sum_{i=1}^{m} \lambda_i^*(f_i(\vec{x}) - u_i) + \sum_{j=1}^{p} \nu_j(h_j(\vec{x}) - \nu_j)$$
$$= \inf_{\vec{x}} \mathcal{L}(\vec{x},\vec{\lambda}^*,\vec{\nu}^*) - \vec{\lambda}^{*\top}\vec{u} - \vec{\nu}^{*\top}\vec{v} = p^*(0,0) - \vec{\lambda}^{*\top}\vec{u} - \vec{\nu}^{*\top}\vec{v}$$

where the first inequality follows from weak duality and the final equality follows from strong duality for the unperturbed problem.

Solution 2:

If we let $g(\vec{\lambda},\vec{\nu})$ be the dual function for the unperturbed problem, we have the following:

For any $\vec{x}$ that is feasible for the perturbed problem, we have for dual function for the perturbed problem $g(\vec{\lambda},\vec{\nu})$ that

$$g(\vec{\lambda}^*,\vec{\nu}^*) \leq f_0(\vec{x}) + \sum_{i=1}^{m} \lambda_i^* f_i(\vec{x}) + \sum_{j=1}^{p} \nu_i^* h_i(\vec{x})$$
$$\leq f_0(\vec{x}) + \vec{\lambda}^{*\top}\vec{u} + \vec{\nu}^{*\top}\vec{v}.$$

By strong duality, we have that $g(\vec{\lambda}^*,\vec{\nu}^*) = p^*(\vec{0},\vec{0})$, and hence we have

$$f_0(\vec{x}) \geq p^*(\vec{0},\vec{0}) - \vec{\lambda}^{*\top}\vec{u} - \vec{\nu}^{*\top}\vec{v}.$$

Since this inequality holds for all $\vec{x}$ feasible for the perturbed problem, the result holds as desired.

(c) Suppose we only have one equality and one inequality constraint (i.e., $\vec{u} = u$ and $\vec{v} = v$ are scalars). For each of the following situations, argue whether:
(A) the value of $p^*(u,v)$ increases as compared with $p^*(0,0)$,
(B) the value of $p^*(u,v)$ decreases as compared with $p^*(0,0)$, or
(C) we can make no claims on the relationship between $p^*(u,v)$ and $p^*(0,0)$.
Hint: Use the bound you computed in (b).

i. λ^* is large (as compared with ν^*) and $u < 0$.
ii. λ^* is large (as compared with ν^*) and $u > 0$.
iii. ν^* is large (as compared with λ^*) and positive and $v < 0$.
iv. ν^* is large (as compared with λ^*) and negative and $v > 0$.

Note that we can think of u and v as variables we choose — by examining how the solution to our original primal problem changes, we can describe how "sensitive" our problem is to its different constraints!

Solution: For the scalar case, we can restate the relationship proved in (b) as

$$p^*(u, v) \geq p^*(0, 0) - \lambda^* u - \nu^* v.$$

For $u = v = 0$, this bound is (trivially) tight: it just says that $p^*(0,0) = p^*(0,0)$. This means that when we increase the value of u and v incrementally, we can use this bound to determine whether $p^*(u, v)$ increases or decreases as compared with $p^*(0, 0)$.

i. *(A)*. Since λ^* is large and u is negative, we know that $-\lambda^* u \geq 0$, so our lower bound on $p^*(u, v)$ increases; thus, $p^*(u, v)$ must increase accordingly.
ii. *(C)*. In this case, $-\lambda^* u \leq 0$, so the lower bound on $p^*(u, v)$ decreases; this doesn't allow us to say anything about whether $p^*(u, v)$ increases, decreases, or remains the same, since it could still obey this bound for any of the three cases.
iii. *(A)*. As in case (c)i., $-\nu^* v \geq 0$, so $p^*(u, v)$ increases.
iv. *(A)*. As in case (c)iii., $-\nu^* v \geq 0$, so $p^*(u, v)$ increases.

4. KKT with circles

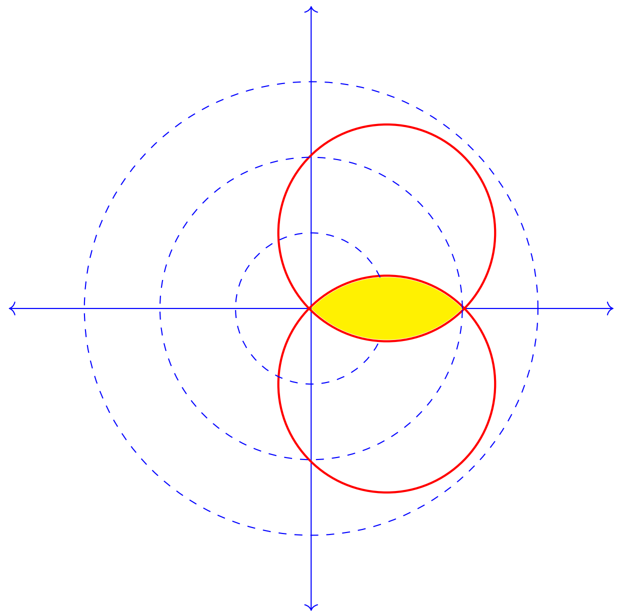
Consider the problem

$$\begin{aligned} \min_{\vec{x} \in \mathbb{R}^2} \quad & x_1^2 + x_2^2 \\ \text{s.t.} \quad & (x_1 - 1)^2 + (x_2 - 1)^2 \leq 2 \\ & (x_1 - 1)^2 + (x_2 + 1)^2 \leq 2 \end{aligned}$$

where $x = \begin{bmatrix} x_1 & x_2 \end{bmatrix}^\top \in \mathbb{R}^2$.

(a) Sketch the feasible region and the level sets of the objective function. Find the optimal point $\vec{x}^*$ and the optimal value p^*.

Solution:



The feasible region is given by the yellow area in the graphic above. The optimal solution is the closest point to the origin inside the feasible region. Since the origin is an element of this feasible region, we have $\vec{x}^* = \begin{bmatrix}0 & 0\end{bmatrix}^\top$, and $p^* = 0$.

(b) Does strong duality hold?

Solution: The problem is convex (i.e., the objective function and the feasible set are both convex). The feasible set contains interior points (e.g., $\vec{x} = \begin{bmatrix}1 & 0\end{bmatrix}^\top$), so Slater's condition is satisfied and thus strong duality holds.

(c) Write the KKT conditions for this optimization problem. Do there exist Lagrange multipliers λ_1^* and λ_2^* that prove the optimality of $\vec{x}^*$?

Solution: The Lagrangian is given by

$$\mathcal{L}(x, \lambda) = x_1^2 + x_2^2 + \lambda_1[(x_1 - 1)^2 + (x_2 - 1)^2 - 2] + \lambda_2[(x_1 - 1)^2 + (x_2 + 1)^2 - 2].$$

We can write the KKT conditions as follows:

i. Stationarity:

$$x_1^* + (\lambda_1^* + \lambda_2^*)(x_1^* - 1) = 0,$$
$$x_2^* + \lambda_1^*(x_2^* - 1) + \lambda_2^*(x_2^* + 1) = 0.$$

ii. Primal feasibility:

$$(x_1^* - 1)^2 + (x_2^* - 1)^2 - 2 \leq 0,$$
$$(x_1^* - 1)^2 + (x_2^* + 1)^2 - 2 \leq 0.$$

iii. Dual feasibility:

$$\lambda_1^* \geq 0,\ \lambda_2^* \geq 0.$$

iv. Complementary slackness:

$$\lambda_1^*[(x_1^* - 1)^2 + (x_2^* - 1)^2 - 2] = 0,$$
$$\lambda_2^*[(x_1^* - 1)^2 + (x_2^* + 1)^2 - 2] = 0.$$

From the stationarity conditions (along with dual feasibility), we can conclude that

$$x_1^* = 0,\ x_2^* = 0 \Rightarrow \lambda_1^* = \lambda_2^* = 0.$$

Since these values for $\vec{x}^*$ and $\vec{\lambda}^*$ satisfy the KKT conditions and strong duality holds, we can conclude that $\vec{x}^*$ is primal optimal (and, additionally, that $\vec{\lambda}^*$ is dual optimal).

5. Water filling

Consider the following problem:

$$\begin{aligned} \min_{\vec{x}} \ & -\sum_{i=1}^{n} \log(\alpha_i + x_i) \\ \text{s.t.} \ & x_i \geq 0,\ i = 1, \ldots, n, \\ & \vec{1}^\top \vec{x} = 1, \end{aligned}$$

where each scalar $\alpha_i > 0$ for $i = 1, \ldots, n$.

This problem arises in information theory when we wish to allocate power to a set of n communication channels. Each variable x_i represents the transmitter power allocated to the i^{th} channel, and $\log(\alpha_i + x_i)$ gives the capacity or communication rate of the channel, so the problem is to allocate a total power of one to the channels, in order to maximize the total communication rate.
Note: This is Example 5.2 on Page 245 of Boyd's book.

(a) Write the KKT conditions.

Solution: We can write the Lagrangian as:

$$\mathcal{L}(\vec{x}, \vec{\lambda}, \nu) = -\sum_{i=1}^{n} \log(\alpha_i + x_i) - \sum_{i=1}^{n} \lambda_i x_i + \nu(\vec{1}^\top \vec{x} - 1).$$

i. Stationarity:

$$-\frac{1}{\alpha_i + x_i^*} - \lambda_i^* + \nu^* = 0, \quad i = 1, \ldots, n$$

ii. Primal feasibility:

$$x_i^* \geq 0,\ i = 1, \ldots, n \qquad \vec{1}^\top \vec{x}^* = 1$$

iii. Dual feasibility:

$$\lambda_i^* \geq 0,\ i = 1, \ldots, n$$

iv. Complementary slackness:

$$\lambda_i^* x_i^* = 0,\ i = 1, \ldots, n$$

(b) Find the primal solution $\vec{x}^*$ using the KKT conditions.

Solution: The dual variable $\vec{\lambda}^*$ is a slack variable, and we can eliminate it by using the stationarity condition:

$$x_i^* \geq 0, \quad \vec{1}^\top \vec{x}^* = 1, \quad \nu^* \geq \frac{1}{\alpha_i + x_i^*}, \quad x_i^* \left(\nu^* - \frac{1}{\alpha_i + x_i^*}\right) = 0.$$

We can find the solution by considering two cases:

- If $\nu^* < 1/\alpha_i$, then we must have $x_i > 0$. Then complementary slackness gives $\nu^* = 1/(\alpha_i + x_i^*)$, or equivalently, $x_i^* = -\alpha_i + 1/\nu^*$.
- If $\nu^* > 1/\alpha_i$, by complementary slackness, we must have $x_i^* = 0$. To see this, assume $x_i^* > 0$:
$$x_i^* \implies \nu^* = \frac{1}{\alpha_i + x_i^*} \implies \nu^* \leq \frac{1}{\alpha_i},$$
which leads to a contradiction.

As a result, we can write the optimal solution x^* as:

$$\begin{aligned} x_i^* &= \begin{cases} -\alpha_i + 1/\nu^* & \text{if } \nu^* < 1/\alpha_i, \\ 0 & \text{if } \nu^* \geq 1/\alpha_i, \end{cases} \\ &= \max\{0, -\alpha_i + 1/\nu^*\} \end{aligned}$$

Lastly, we need to determine the value of ν^*, for which we can use the primal feasibility condition:

$$\sum_{i=1}^{n} \max\{0, -\alpha_i + 1/\nu^*\} = 1.$$

Note that

$$f(\nu) = \sum_{i=1}^{n} \max\{0, -\alpha_i + 1/\nu\}$$

is a strictly decreasing function over the region $(0, \infty)$; therefore, there exists a unique value for ν^* such that $f(\nu^*) = 1$.

This solution method is called water filling because $(\alpha_i + x_i)$ is increased to $1/\nu^*$, unless α_i already exceeds this threshold.

6. Homework process

Whom did you work with on this homework? List the names and SIDs of your group members.