Optimization 1 — Tutorial 12

January 14, 2021

Problem 1

Consider the problem

$$\begin{aligned}\min_{x,y,z} \quad & x^2 - y^4 + z^2 + \sqrt{x^2+z^2} + \left\| \begin{pmatrix} 2x \\ x+z \end{pmatrix} \right\|_1 \\ & x^2 + y^4 + z^2 + z \le 4.\end{aligned}$$

(a) Find a dual problem.

(b) Is the dual problem convex?

Solution

(a) Writing the Lagrangian

$$L(x,y,z,\lambda) = x^2 - y^4 + z^2 + \sqrt{x^2+z^2} + \left\| \begin{pmatrix} 2x \\ x+z \end{pmatrix} \right\|_1 + \lambda\left(x^2+y^4+z^2+z-4\right), \quad \lambda \ge 0,$$

- The problem is not separable and it is impossible to find an explicit minimizer. Therefore, we transform the problem into an equivalent one with a separable Lagrangian.
- Define $t_1 = x$, $t_2 = z$, $w_1 = 2x$ and $w_2 = x+z$ (other substitutions can also be used)

$$\begin{aligned}\min_{x,y,z,\mathbf{t},\mathbf{w}} \quad & x^2 - y^4 + z^2 + \sqrt{t_1^2+t_2^2} + \|\mathbf{w}\|_1 \\ & x^2+y^4+z^2+z \le 4, \\ & t_1 = x, \\ & t_2 = z, \\ & w_1 = 2x, \\ & w_2 = x+z.\end{aligned}$$

- Therefore

$$\begin{aligned}L(x,y,z,\mathbf{t},\mathbf{w},\lambda,\boldsymbol{\mu},\boldsymbol{\eta}) = {} & x^2 - y^4 + z^2 + \sqrt{t_1^2+t_2^2} + \|\mathbf{w}\|_1 \\ & + \lambda\left(x^2+y^4+z^2+z-4\right) \\ & + \mu_1(t_1-x) + \mu_2(t_2-z) \\ & + \eta_1(w_1-2x) + \eta_2(w_2-x-z), \quad \lambda \ge 0.\end{aligned}$$

- And now finding q is a separable problem

$$\begin{aligned}q(\lambda,\boldsymbol{\mu},\boldsymbol{\eta}) = {} & \min_{x,y,z,\mathbf{t},\mathbf{w}} L = \min_x \left\{(\lambda+1)x^2 - (\mu_1+2\eta_1+\eta_2)x\right\} \\ & + \min_y\left\{(\lambda-1)y^4\right\} + \min_z\left\{(\lambda+1)z^2 - (\eta_2+\mu_2)z\right\} \\ & + \min_{\mathbf{t}}\left\{\|\mathbf{t}\| + \boldsymbol{\mu}^T\mathbf{t}\right\} + \min_{\mathbf{w}}\left\{\|\mathbf{w}\|_1 + \boldsymbol{\eta}^T\mathbf{w}\right\}, \quad \lambda \ge 0.\end{aligned}$$

– $\min_x \left\{(\lambda+1)x^2 - (\mu_1+2\eta_1)x\right\} = -\frac{(\mu_1+2\eta_1+\eta_2)^2}{4(\lambda+1)}$ for $x = \frac{\mu_1+2\eta_1+\eta_2}{2(\lambda+1)}$.

- $\min\limits_{y}\left\{(\lambda-1)\,y^4\right\} = \begin{cases} 0, & \lambda \geq 1, \\ -\infty, & \lambda < 1. \end{cases}$
- $\min\limits_{z}\left\{(\lambda+1)\,z^2 - \eta_2 z\right\} = -\frac{(\eta_2+\mu_2)^2}{4(\lambda+1)}$ for $z = \frac{\eta_2+\mu_2}{2(\lambda+1)}$.
- To solve for $\mathbf{t} \in \mathbb{R}^2$ notice

$$\|\mathbf{t}\| + \boldsymbol{\mu}^T\mathbf{t} \geq \|\mathbf{t}\|_2 - \|\mathbf{t}\|\,\|\boldsymbol{\mu}\| = \|\mathbf{t}\|\,(1-\|\boldsymbol{\mu}\|)\,.$$

  * If $\|\boldsymbol{\mu}\| > 1$ choose $\mathbf{t} = -\alpha\boldsymbol{\mu}$ and

$$\|\mathbf{t}\| + \boldsymbol{\mu}^T\mathbf{t} = \alpha\,\|\boldsymbol{\mu}\| - \alpha\,\|\boldsymbol{\mu}\|^2 = \alpha\,\|\boldsymbol{\mu}\|\,(1-\|\boldsymbol{\mu}\|) \xrightarrow[\alpha\to\infty]{} -\infty.$$

  * If $\|\boldsymbol{\mu}\| \leq 1$ then $\|\mathbf{t}\| + \boldsymbol{\mu}^T\mathbf{t} \geq 0$ and this lower bound is attained for $\mathbf{t} = \mathbf{0}$.
  * So

$$\min_{\mathbf{t}}\left\{\|\mathbf{t}\| + \boldsymbol{\mu}^T\mathbf{t}\right\} = \begin{cases} 0, & \|\boldsymbol{\mu}\| \leq 1, \\ -\infty, & \|\boldsymbol{\mu}\| > 1. \end{cases}$$

- To solve for $\mathbf{w} \in \mathbb{R}^2$ notice

$$\|\mathbf{w}\|_1 + \boldsymbol{\eta}^T\mathbf{w} \geq \|\mathbf{w}\|_1 - \|\mathbf{w}\|_1\,\|\boldsymbol{\eta}\|_\infty = \|\mathbf{w}\|_1\,(1-\|\boldsymbol{\eta}\|_\infty)\,.$$

  * If $\|\boldsymbol{\eta}\|_\infty > 1$ choose $\mathbf{w} = -\alpha\,(0,\ldots,\operatorname{sign}(\eta_k),\ldots,0)^T$ for $k = \operatorname*{argmax}\limits_{i}|\eta_i|$, and

$$\|\mathbf{w}\|_1 + \boldsymbol{\eta}^T\mathbf{w} = \alpha - \alpha\,\|\boldsymbol{\eta}\|_\infty = \alpha\,(1-\|\boldsymbol{\eta}\|_\infty) \xrightarrow[\alpha\to\infty]{} -\infty.$$

  * If $\|\boldsymbol{\eta}\|_\infty \leq 1$ then $\|\mathbf{w}\|_1 + \boldsymbol{\mu}^T\mathbf{w} \geq 0$ and this lower bound is attained for $\mathbf{w} = \mathbf{0}$.
  * So

$$\min_{\mathbf{w}}\left\{\|\mathbf{w}\|_1 + \boldsymbol{\eta}^T\mathbf{w}\right\} = \begin{cases} 0, & \|\boldsymbol{\eta}\|_\infty \leq 1, \\ -\infty, & \|\boldsymbol{\eta}\|_\infty > 1. \end{cases}$$

- Therefore, the dual is

$$\begin{aligned} \max_{\lambda,\boldsymbol{\mu},\boldsymbol{\eta}} \quad & -\frac{(\mu_1+2\eta_1+\eta_2)^2}{4(\lambda+1)} - \frac{(\eta_2+\mu_2)^2}{4(\lambda+1)} - 4\lambda \\ & \lambda \geq 1, \\ & \|\boldsymbol{\mu}\| \leq 1, \\ & \|\boldsymbol{\eta}\|_\infty \leq 1. \end{aligned}$$

(b) The dual problem is always a convex problem.

Problem 2

Find a dual problem for

$$\begin{aligned} \min_{\mathbf{x}\in\mathbb{R}^n} \quad & \sum_{i=1}^{n}\left(\mathbf{a}_i\mathbf{x}_i^2 + 2\mathbf{b}_i\mathbf{x}_i + e^{\mathbf{c}_i\mathbf{x}_i}\right) \\ & \sum_{i=1}^{n}\mathbf{x}_i = 1, \end{aligned}$$

where $\mathbf{a} \in \mathbb{R}^n_{++}$ and $\mathbf{b},\mathbf{c} \in \mathbb{R}^n$. Verify that strong duality holds.

Solution

- Even though the Lagrangian is separable, we cannot solve explicitly

$$\min_{\mathbf{x}_i}\left\{\mathbf{a}_i\mathbf{x}_i^2 + 2\mathbf{b}_i\mathbf{x}_i + e^{\mathbf{c}_i\mathbf{x}_i} + \lambda\mathbf{x}_i\right\}.$$

- Define $\mathbf{y}_i = \mathbf{c}_i \mathbf{x}_i$

$$\begin{aligned} \min_{\mathbf{x}\in\mathbb{R}^n} \quad & \sum_{i=1}^{n} \left(\mathbf{a}_i \mathbf{x}_i^2 + 2\mathbf{b}_i \mathbf{x}_i + e^{\mathbf{y}_i}\right) \\ & \sum_{i=1}^{n} \mathbf{x}_i = 1, \\ & \mathbf{y}_i = \mathbf{c}_i \mathbf{x}_i, \quad \forall 1 \le i \le n. \end{aligned}$$

- The separable Lagrangian is

$$\begin{aligned} L(\mathbf{x}, \mathbf{y}, \lambda, \boldsymbol{\mu}) &= \sum_{i=1}^{n} \left(\mathbf{a}_i \mathbf{x}_i^2 + 2\mathbf{b}_i \mathbf{x}_i + e^{\mathbf{y}_i}\right) + \lambda \left(\sum_{i=1}^{n} \mathbf{x}_i - 1\right) + \sum_{i=1}^{n} \mu_i (\mathbf{y}_i - \mathbf{c}_i \mathbf{x}_i) \\ &= \sum_{i=1}^{n} \left(\mathbf{a}_i \mathbf{x}_i^2 + (2\mathbf{b}_i + \lambda - \mu_i \mathbf{c}_i) \mathbf{x}_i\right) + \sum_{i=1}^{n} \left(e^{\mathbf{y}_i} + \mu_i \mathbf{y}_i\right) - \lambda. \end{aligned}$$

  - $\min_{\mathbf{x}_i} \left\{\mathbf{a}_i \mathbf{x}_i^2 + (2\mathbf{b}_i + \lambda - \mu_i \mathbf{c}_i) \mathbf{x}_i\right\} = -\frac{(2\mathbf{b}_i + \lambda - \mu_i \mathbf{c}_i)^2}{4\mathbf{a}_i}$ for $\mathbf{x}_i = -\frac{2\mathbf{b}_i + \lambda - \mu_i \mathbf{c}_i}{2\mathbf{a}_i}$ (a convex problem so stationarity is sufficient for optimality).
  - $\min_{\mathbf{y}_i} \{e^{\mathbf{y}_i} + \mu_i \mathbf{y}_i\}$ is convex, so stationarity is sufficient for optimality. A stationary point satisfies $e^{\mathbf{y}_i} + \mu_i = 0$.
    - If $\mu_i < 0$ then the solution is $\mathbf{y}_i = \ln(-\mu_i)$ with optimal value $\mu_i (\ln(-\mu_i) - 1)$.
    - If $\mu_i > 0$ there are no stationary points (and there is no lower bound since $e^{\mathbf{y}_i} + \mu_i \mathbf{y}_i \to -\infty$ as $\mathbf{y}_i \to -\infty$).
    - for $\mu = 0$ the optimal value is not attained and 0 is a lower bound.
    - Therefore, under the convention $0 \ln(0) = 0$ we have

$$\min_{\mathbf{y}_i} \{e^{\mathbf{y}_i} + \mu_i \mathbf{y}_i\} = \begin{cases} \mu_i (\ln(-\mu_i) - 1), & \mu_i \le 0 \\ -\infty, & \mu_i > 0. \end{cases}$$

- The dual problem

$$\begin{aligned} \max_{\lambda, \boldsymbol{\mu}} \quad & \sum_{i=1}^{n} \left(\mu_i (\ln(-\mu_i) - 1) - \frac{(2\mathbf{b}_i + \lambda - \mu_i \mathbf{c}_i)^2}{4\mathbf{a}_i}\right) - \lambda \\ & \lambda \ge 0, \\ & \boldsymbol{\mu} \le \mathbf{0}. \end{aligned}$$

- Strong duality holds, since the equivalent primal is convex and coercive over a nonempty closed linearly constrained feasible set.

Problem 3

Consider the optimization problem

$$\begin{aligned} \min_{\mathbf{x}\in\mathbb{R}^n} \quad & \mathbf{a}^T \mathbf{x} + \sum_{i=1}^{n} \mathbf{x}_i \ln(\mathbf{x}_i) \\ & \mathbf{x} \in \triangle_n. \end{aligned}$$

(a) Show that the problem cannot have more than one optimal solution.

(b) Find a one-dimensional dual problem.

(c) Find the optimal solution of the dual and primal problems.

Solution

(a) The function $\mathbf{x}_i \ln(\mathbf{x}_i)$ is strictly convex over $\mathbb{R}_+$, so the objective is strictly convex over the domain. Therefore, it has one optimal solution or no solutions (moreover, since $\triangle_n$ is compact there is exactly one).

Alternatively, since this is a convex problem and the generalized Slater's condition is satisfied we have $\{\text{KKT}\} = \{\text{optimal}\}$ and also $\{\text{optimal}\} \neq \emptyset$.

- So we need to show that there is at most one feasible KKT point.

$$\begin{aligned} L(\mathbf{x}, \boldsymbol{\lambda}, \mu) &= \mathbf{a}^T \mathbf{x} + \sum_{i=1}^{n} \mathbf{x}_i \ln(\mathbf{x}_i) + \mu \left(\sum_{i=1}^{n} \mathbf{x}_i - 1 \right) - \boldsymbol{\lambda}^T \mathbf{x} \\ &= \sum_{i=1}^{n} (\mathbf{a}_i + \mu - \lambda_i) \mathbf{x}_i + \sum_{i=1}^{n} \mathbf{x}_i \ln(\mathbf{x}_i) - \mu, \quad \boldsymbol{\lambda} \in \mathbb{R}_+^n, \mu \in \mathbb{R} \end{aligned}$$

- The KKT conditions are

$$\begin{cases} \mathbf{a}_i + \mu - \lambda_i + \ln(\mathbf{x}_i) + 1 = 0, \\ \lambda_i \mathbf{x}_i = 0, \\ \sum_{i=1}^{n} \mathbf{x}_i = 1, \\ \mathbf{x}_i \geq 0 \end{cases}$$

- The system has exactly one solution $\mathbf{x}_i = e^{-\mathbf{a}_i - \mu - 1}$, $\mu = \ln(e^{-\mathbf{a}_i}) - 1$, $\lambda_i = 0$.

(b) We write a Lagrangian with only one dual variable

$$L(\mathbf{x}, \lambda) = \sum_{i=1}^{n} (\mathbf{a}_i \mathbf{x}_i + \mathbf{x}_i \ln(\mathbf{x}_i)) + \mu \sum_{i=1}^{n} (\mathbf{x}_i - 1), \quad 0 \leq \mathbf{x}_i, \mu \in \mathbb{R}$$

- This is a separable problem and we need to solve

$$\min_{0 \leq \mathbf{x}_i} \{(\mathbf{a}_i + \mu) \mathbf{x}_i + \mathbf{x}_i \ln(\mathbf{x}_i)\}.$$

- This is a convex objective, so stationarity is sufficient for optimality. If the stationary point is not feasible, then the optimal solution is on the boundary. A stationary point satisfies $\mathbf{a}_i + \mu + \ln(\mathbf{x}_i) + 1 = 0$, and it has a solution $\mathbf{x}_i = e^{-\mathbf{a}_i - \mu - 1} > 0$. This point is always feasible, and therefore optimal. The optimal value is

$$(\mathbf{a}_i + \mu) e^{-\mathbf{a}_i - \mu - 1} - e^{-\mathbf{a}_i - \mu - 1} (\mathbf{a}_i + \mu + 1) = -e^{-\mathbf{a}_i - \mu - 1}.$$

- The dual is

$$\max_{\mu \in \mathbb{R}} \left\{ -\sum_{i=1}^{n} e^{-\mathbf{a}_i - \mu - 1} - \mu \right\}.$$

(c) The dual is a non-constrained maximization of a concave function. Therefore, stationarity is sufficient for optimality. A stationary point satisfies

$$\sum_{i=1}^{n} e^{-\mathbf{a}_i - \mu - 1} = 1 \Longrightarrow \mu = \ln\left(\sum_{i=1}^{n} e^{-\mathbf{a}_i} \right) - 1.$$

and μ is the solution of the dual . For the primal problem, we have the identity $\mathbf{x}_i = e^{-\mathbf{a}_i - \mu - 1}$, so the primal solution is

$$\mathbf{x}_i = e^{-\mathbf{a}_i - \ln\left(\sum_{i=1}^{n} e^{-\mathbf{a}_i} \right) + 1 - 1} = \frac{e^{-\mathbf{a}_i}}{\sum_{i=1}^{n} e^{-\mathbf{a}_i}}.$$