Optimization 1 — Tutorial 13

January 21, 2021

Problem 1

Let $\mathbf{A} \in \mathbb{R}^{m\times n}$. Prove that exactly one of the following two claims is satisfied.

(I) The system $\mathbf{Ax} = \mathbf{0}$, $\mathbf{x} > 0$ has a solution.

(II) There exists a vector $\mathbf{y} \in \mathbb{R}^n$ for which $\mathbf{A}^T\mathbf{y} \leq \mathbf{0}$ and $\mathbf{A}^T\mathbf{y} \neq \mathbf{0}$.

Solution

$\neg(I) \Rightarrow (II)$: Assume $\mathbf{Ax} = \mathbf{0}$, $\mathbf{x} > 0$ has no solution.

- Since $\mathbf{e} > 0$ then $\mathbf{A}(\mathbf{x} + \mathbf{e}) = \mathbf{0}, \mathbf{x} \geq \mathbf{0}$ also has no solution (otherwise $\mathbf{z} = \mathbf{x} + \mathbf{e} > 0$ is a solution of (I)).
- Define

$$
\begin{aligned}
(O) \quad & \mathbf{Ax} = -\mathbf{Ae},\ \mathbf{x} \geq \mathbf{0} \\
(P) \quad & \mathbf{A}^T\mathbf{y} \leq \mathbf{0},\ (-\mathbf{Ae})^T\mathbf{y} > \mathbf{0}
\end{aligned}
$$

- From Farkas' lemma exactly one of them has a solution.
- Since (O) has no solution, then (P) has a solution, which immediately implies (II).

$(II) \Rightarrow \neg(I)$: Assume that there exists a vector $\mathbf{y} \in \mathbb{R}^n$ for which $\mathbf{A}^T\mathbf{y} \leq \mathbf{0}$ and $\mathbf{A}^T\mathbf{y} \neq \mathbf{0}$.

- Assume on the contrary that $\mathbf{Ax} = \mathbf{0}$, $\mathbf{x} > 0$ has a solution ($(I), (II)$ hold together).
- Therefore $\mathbf{y}^T\mathbf{Ax} = \mathbf{x}^T\mathbf{A}^T\mathbf{y} = 0$, $\mathbf{x} > 0$ has a solution.
- Since $\mathbf{x} > 0$ and $\mathbf{A}^T\mathbf{y} \leq 0$, $\mathbf{A}^T\mathbf{y} \neq \mathbf{0}$ then $0 = \mathbf{x}^T\mathbf{A}^T\mathbf{y} < 0$ which is a contradiction.

Problem 2 (Winter 2012/2013)

Let $\mathbf{E} \in \mathbb{R}^{k\times n}$, $\mathbf{f} \in \mathbb{R}^n$, $\mathbf{a} \in \mathbb{R}^m$, $\mathbf{A} \in \mathbb{R}^{m\times n}$, $\mathbf{b} \in \mathbb{R}^m$ and $\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_n \in \mathbb{R}^n$. Consider the problem

$$
\begin{aligned}
\min_{\mathbf{x}\in\mathbb{R}^n, \mathbf{z}\in\mathbb{R}^m} \quad & \frac{1}{2}\|\mathbf{Ex}\|^2 + \frac{1}{2}\|\mathbf{z}\|^2 + \mathbf{f}^T\mathbf{x} + \mathbf{a}^T\mathbf{z} + \sum_{i=1}^{n} e^{\mathbf{c}_i^T\mathbf{x}} \\
\text{s.t.} \quad & \mathbf{Ax} + \mathbf{z} = \mathbf{b}, \\
& \mathbf{z} \geq 0.
\end{aligned}
$$

Assume that $\mathbf{E}$ has full column rank.

(a) Show that the objective function is coercive.

(b) Show that if the set $P = \{\mathbf{x} \in \mathbb{R}^n \colon \mathbf{Ax} \leq \mathbf{b}\}$ is non-empty, then strong duality holds.

(c) Find a dual problem.

Solution

(a) Since $\lambda_{\min}\left(\mathbf{E}^T\mathbf{E}\right) > 0$ we have

$$\begin{aligned} f(\mathbf{x},\mathbf{z}) &= \frac{1}{2}\|\mathbf{E}\mathbf{x}\|^2 + \frac{1}{2}\|\mathbf{z}\|^2 + \mathbf{f}^T\mathbf{x} + \mathbf{a}^T\mathbf{z} + \sum_{i=1}^{n} e^{\mathbf{c}_i^T\mathbf{x}} \\ &\geq \frac{1}{2}\lambda_{\min}\left(\mathbf{E}^T\mathbf{E}\right)\|\mathbf{x}\|^2 + \frac{1}{2}\|\mathbf{z}\|^2 - \begin{pmatrix}\mathbf{f}\\ \mathbf{a}\end{pmatrix}^T\begin{pmatrix}\mathbf{x}\\ \mathbf{z}\end{pmatrix} \\ &\geq \frac{1}{2}\min\left\{\lambda_{\min}\left(\mathbf{E}^T\mathbf{E}\right),1\right\}\left\|\begin{pmatrix}\mathbf{x}\\ \mathbf{z}\end{pmatrix}\right\|^2 - \left\|\begin{pmatrix}\mathbf{f}\\ \mathbf{a}\end{pmatrix}\right\|\left\|\begin{pmatrix}\mathbf{x}\\ \mathbf{z}\end{pmatrix}\right\| \xrightarrow[\left\|\begin{pmatrix}\mathbf{x}\\ \mathbf{z}\end{pmatrix}\right\|\to\infty]{} \infty. \end{aligned}$$

(b) We show the three required properties:

- For $\mathbf{x} \in P$ define $\mathbf{z} = \mathbf{b} - \mathbf{A}\mathbf{x} \geq 0$. So $(\mathbf{x},\mathbf{z})$ is feasible (generalized Slater's condition).
- The problem is of course convex (need to show).
- $f^* > -\infty$ from coerciveness of a continuous function over a closed feasible set.

(c) For separability we reformulate the problem as

$$\begin{aligned} \min_{\mathbf{x},\mathbf{z},\mathbf{y},w_i} \quad & \frac{1}{2}\|\mathbf{E}\mathbf{x}\|^2 + \frac{1}{2}\|\mathbf{z}\|^2 + \mathbf{f}^T\mathbf{x} + \mathbf{a}^T\mathbf{z} + \sum_{i=1}^{n} e^{\mathbf{w}_i} \\ \text{s.t.} \quad & \mathbf{A}\mathbf{x} + \mathbf{z} = \mathbf{b}, \\ & \mathbf{z} \geq 0, \\ & \mathbf{w}_i = \mathbf{c}_i^T\mathbf{x}, \quad \forall 1 \leq i \leq n. \end{aligned}$$

- The Lagrangian is

$$\begin{aligned} L(\mathbf{x},\mathbf{z},\mathbf{y},\mathbf{w},\boldsymbol{\theta},\boldsymbol{\varphi}) &= \frac{1}{2}\|\mathbf{E}\mathbf{x}\|^2 + \frac{1}{2}\|\mathbf{z}\|^2 + \mathbf{f}^T\mathbf{x} + \mathbf{a}^T\mathbf{z} + \sum_{i=1}^{n} e^{\mathbf{w}_i} \\ &\quad + \sum_{i=1}^{n}\theta_i\left(\mathbf{w}_i - \mathbf{c}_i^T\mathbf{x}\right) + \boldsymbol{\varphi}^T\left(\mathbf{A}\mathbf{x} + \mathbf{z} - \mathbf{b}\right), \quad \mathbf{z} \geq \mathbf{0}. \end{aligned}$$

- So

$$\begin{aligned} q(\boldsymbol{\theta},\boldsymbol{\varphi}) &= \min_{\mathbf{x}}\left\{\frac{1}{2}\|\mathbf{E}\mathbf{x}\|^2 + \left(\mathbf{f} - \sum_{i=1}^{n}\theta_i\mathbf{c}_i - \mathbf{A}^T\boldsymbol{\varphi}\right)^T\mathbf{x}\right\} \\ &\quad + \min_{\mathbf{z}\geq\mathbf{0}}\left\{\frac{1}{2}\|\mathbf{z}\|^2 + (\mathbf{a}+\boldsymbol{\varphi})^T\mathbf{z}\right\} + \min_{w_i}\sum_{i=1}^{n}\left(e^{\mathbf{w}_i} + \theta_i\mathbf{w}_i\right). \end{aligned}$$

  - The minimization w.r.t. $\mathbf{x}$ is convex and unconstrained, so stationarity is sufficient for optimality. Stationary points satisfy

$$\mathbf{E}^T\mathbf{E}\mathbf{x} + \mathbf{f} - \sum_{i=1}^{n}\theta_i\mathbf{c}_i - \mathbf{A}^T\boldsymbol{\varphi} = \mathbf{0} \Longrightarrow \mathbf{x} = -\left(\mathbf{E}^T\mathbf{E}\right)^{-1}\left(\mathbf{f} - \sum_{i=1}^{n}\theta_i\mathbf{c}_i - \mathbf{A}^T\boldsymbol{\varphi}\right),$$

  since $\mathbf{E}$ is of full column rank, and the minimal value is

$$-\frac{1}{2}\left(\mathbf{f} - \sum_{i=1}^{n}\theta_i\mathbf{c}_i - \mathbf{A}^T\boldsymbol{\varphi}\right)\left(\mathbf{E}^T\mathbf{E}\right)^{-1}\left(\mathbf{f} - \sum_{i=1}^{n}\theta_i\mathbf{c}_i - \mathbf{A}^T\boldsymbol{\varphi}\right) \equiv -\frac{1}{2}\mathbf{g}^T\left(\mathbf{E}^T\mathbf{E}\right)^{-1}\mathbf{g}.$$

  - $\min_{\mathbf{z}\geq\mathbf{0}}\left\{\frac{1}{2}\|\mathbf{z}\|^2 + (\mathbf{a}+\boldsymbol{\varphi})^T\mathbf{z}\right\}$ is separable with respect to the $\mathbf{z}_i$-s. From convexity, the optimum is a stationary point or on the boundary. Stationarity implies $\mathbf{z}_i + \mathbf{a}_i + \varphi_i = 0$, so from the constraint

$$\min_{\mathbf{z}_i\geq 0}\left\{\frac{1}{2}\mathbf{z}_i^2 + (\mathbf{a}_i+\varphi_i)\mathbf{z}_i\right\} = \begin{cases} -\frac{1}{2}(\mathbf{a}_i+\varphi_i)^2 \text{ for } \mathbf{z}_i = -\mathbf{a}_i - \varphi_i, & \mathbf{a}_i + \varphi_i \leq 0, \\ 0 \text{ for } \mathbf{z}_i = 0, & \mathbf{a}_i + \varphi_i > 0, \end{cases}$$

and compactly

$$\min_{\mathbf{z}\geq\mathbf{0}}\left\{\frac{1}{2}\|\mathbf{z}\|^2+(\mathbf{a}+\boldsymbol{\varphi})^T\mathbf{z}\right\}=-\frac{1}{2}\sum_{i=1}^{m}[\mathbf{a}_i+\varphi_i]_+^2 \text{ for } \mathbf{z}=[-\mathbf{a}-\boldsymbol{\varphi}]_+ .$$

$$-\min_{\mathbf{w}}\sum_{i=1}^{n}\left(e^{\mathbf{w}_i}+\theta_i\mathbf{w}_i\right)=\begin{cases}-\sum_{i=1}^{n}(\theta_i+\theta_i\ln(-\theta_i)), & \theta_i\leq 0,\\ -\infty, & \text{otherwise.}\end{cases}$$

- The dual is

$$\begin{aligned}\max_{\boldsymbol{\theta},\boldsymbol{\varphi}}\quad & -\frac{1}{2}\mathbf{g}^T\left(\mathbf{E}^T\mathbf{E}\right)^{-1}\mathbf{g}-\frac{1}{2}\sum_{i=1}^{m}[\mathbf{a}_i+\varphi_i]_+^2-\sum_{i=1}^{n}(\theta_i+\theta_i\ln(-\theta_i))\\ & \theta_i\leq 0,\quad \forall 1\leq i\leq n.\end{aligned}$$

Problem 3 (Winter 2012/2013)
Consider the optimization problem

$$\begin{aligned}\min_{x,y,z\in\mathbb{R}}\quad & xyz\\ \text{s.t.}\quad & x^2+2y^2+3z^2\leq 1.\end{aligned}$$

(a) Find all KKT points of the problem.

(b) Find all optimal solutions of the problem.

Solution

(a) The Lagrangian is

$$L(x,y,z,\lambda)=xyz+\lambda\left(x^2+2y^2+3z^2-1\right),\quad \lambda\geq 0.$$

- KKT conditions are

$$\begin{cases}yz+2\lambda x=0, & (1)\\ xz+4\lambda y=0, & (2)\\ xy+6\lambda z=0, & (3)\\ x^2+2y^2+3z^2\leq 1, & (4)\\ \lambda\left(x^2+2y^2+3z^2-1\right)=0, & (5)\\ \lambda\geq 0. & (6)\end{cases}$$

- If $\lambda=0$ then $yz=xz=xy=0$ (from $(1),(2),(3)$), so at least two of them are 0. From (4) we obtain the points (with value 0)
  - $(0,0,z)$ and $z^2\leq\frac{1}{3}\Longrightarrow -\sqrt{\frac{1}{3}}\leq z\leq\sqrt{\frac{1}{3}}$.
  - $(0,y,0)$ and $y^2\leq\frac{1}{2}\Longrightarrow -\sqrt{\frac{1}{2}}\leq y\leq\sqrt{\frac{1}{2}}$.
  - $(x,0,0)$ and $x^2\leq 1\Longrightarrow -1\leq x\leq 1$.
- If $\lambda>0$ then

$$\begin{cases}yz=-2\lambda x, & (i)\\ xz=-4\lambda y, & (ii)\\ xy=-6\lambda z, & (iii)\\ x^2+2y^2+3z^2-1=0, & (iv)\end{cases}$$

  - If either x,y,z is 0 then all three are 0 (from $(i),(ii),(iii)$), and (iv) is not met.
  - So $x,y,z\neq 0$ and therefore $xyz=-2\lambda x^2=-4\lambda y^2=-6\lambda z^2<0$.
  - Dividing by -2λ we have $x^2=2y^2=3z^2$.
  - From (iv) we have

* $3x^2 = 1 \Longrightarrow x = \pm\sqrt{\frac{1}{3}}$.
* $6y^2 = 1 \Longrightarrow y = \pm\sqrt{\frac{1}{6}}$.
* $9z^2 = 1 \Longrightarrow z = \pm\frac{1}{3}$.

– Since $xyz < 0$ we obtain $\left(\sqrt{\frac{1}{3}}, -\sqrt{\frac{1}{6}}, \frac{1}{3}\right)$, $\left(-\sqrt{\frac{1}{3}}, \sqrt{\frac{1}{6}}, \frac{1}{3}\right)$, $\left(\sqrt{\frac{1}{3}}, \sqrt{\frac{1}{6}}, -\frac{1}{3}\right)$ and $\left(-\sqrt{\frac{1}{3}}, -\sqrt{\frac{1}{6}}, -\frac{1}{3}\right)$.

– For all instances $\lambda = \frac{1}{6\sqrt{2}} > 0$, and therefore these are feasible KKT points.

(b) The constraint is convex and satisfies Slater's condition (for example when $x = y = z = 0$). The objective is continuous over a compact set (all coordinates are bounded), then $\emptyset \neq \{\text{optimal}\} \subseteq \{\text{locally optimal}\} \subseteq \{\text{KKT}\}$.

- All points for which at least one coordinate is 0 give a value of 0.
- All other four points give a value of $-\frac{1}{9\sqrt{2}} < 0$, so are all non-strict global minima.

Question: what if the constraint was an equality constraint? In this case the constraint is not convex, and therefore

$$\emptyset \neq \{\text{optimal}\} \subseteq \{\text{locally optimal}\} \subseteq \{\text{KKT}\} \cup \{\text{irregular}\}.$$

Problem 4 (Spring 2017/2018)

Show that the following problem is convex, in the sense that it is a minimization of a convex function over a convex feasible set.

$$\begin{aligned}
\min_{x,y,z\in\mathbb{R}} \quad & \sqrt{3x^2 + 4y^2 + 2xy + 4xz + 6z^2 + 4z + 8} + \frac{3x^2 + 4y^2 + 5}{10y - 3x} \\
\text{s.t.} \quad & \max\left\{|10x - 3y + 8z|, x^2 - \min\left\{\sqrt{z}, 3\right\}\right\} \leq 4 + x, \\
& \frac{e^{(3x+y)^2}}{z + y} \leq 1, \\
& -1 \leq x \leq 2 \\
& y, z \geq 1.
\end{aligned}$$

Solution

Note: give full detail when showing a problem is convex.

- Objective function:
  – First term:

$$\sqrt{(x+y)^2 + (x+2z)^2 + (z+2)^2 + x^2 + 3y^2 + z^2 + 4} = \left\| \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 2 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & \sqrt{3} & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} - \begin{pmatrix} 0 \\ 0 \\ -2 \\ 0 \\ 0 \\ 0 \\ 2 \end{pmatrix} \right\|^2,$$

  so it is convex under a linear transformation of the convex function $\|\cdot\|^2$.

  – Second term:

$$\frac{3x^2 + 4y^2 + 5}{10y - 3x} = \frac{\left\| \begin{pmatrix} 3 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ \sqrt{5} \end{pmatrix} \right\|^2}{\begin{pmatrix} -3 & 10 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}},$$

and since the denominator is positive (last constraints), this is quad_over_lin – thus convex.

- First constraint:
  - $|10x - 3y + 8z|$ is convex under a linear transformation of the convex function $|\cdot|$.
  - $x^2 - \min\{\sqrt{z}, 3\} = x^2 + \max\{-\sqrt{x}, 3\}$ thus convex since $x^2, -\sqrt{x}, 3$ are convex and maximum preserves convexity.
  - The RHS is linear, so the constraint is convex.
- Second constraint:
  - Equivalent to $e^{(3x+y)^2} \leq z + y$ since $z + y > 0$ (from last constraints).
  - $(3x + y)^2$ is convex (linear transformation).
  - e^t is convex and increasing function, thus $e^{(3x+y)^2}$ is convex.
  - The RHS is linear, so the constraint is convex.
- All other constraints are linear thus convex.