Optimization 1 — Tutorial 9

December 24, 2020

Separation Theorem

Let $\emptyset \neq C \subseteq \mathbb{R}^n$ be closed and convex. Let $\mathbf{y} \notin C$. Then there exist $\mathbf{p} \in \mathbb{R}^n \setminus \{\mathbf{0}_n\}$ and $\alpha \in \mathbb{R}$ such that

$$\mathbf{p}^T \mathbf{y} > \alpha \text{ and } \mathbf{p}^T \mathbf{x} \leq \alpha \text{ for all } \mathbf{x} \in C.$$

Farkas' Lemma

Let $\mathbf{c} \in \mathbb{R}^n$ and $\mathbf{A} \in \mathbb{R}^{m \times n}$. Then exactly one of the following systems has a solution:

(I) $\mathbf{A}\mathbf{x} \leq \mathbf{0}_n$, $\mathbf{c}^T \mathbf{x} > 0$.

(II) $\mathbf{A}^T \mathbf{y} = \mathbf{c}$, $\mathbf{y} \geq 0$.

KKT Conditions for Convex and Linearly Constrained Problems

Consider the optimization problem

$$\begin{aligned} \min_{\mathbf{x} \in \mathbb{R}^n} \quad & f(\mathbf{x}) \\ \text{s.t.} \quad & \mathbf{A}\mathbf{x} \leq \mathbf{a}, \\ & \mathbf{B}\mathbf{x} = \mathbf{b}. \end{aligned}$$

where $f\colon \mathbb{R}^n \to \mathbb{R}$ is a convex and continuously differentiable function, $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{m \times n}$ and $\mathbf{a}, \mathbf{b} \in \mathbb{R}^m$. Let $\mathbf{x}^*$ be a feasible solution of the problem. Then $\mathbf{x}^*$ is an optimal solution of the problem if and only if there exist $\lambda \in \mathbb{R}^m_+$ and $\mu \in \mathbb{R}^m$ such that

$$\begin{cases} \nabla f(\mathbf{x}^*) + \mathbf{A}^T \lambda + \mathbf{B}^T \mu = \mathbf{0}_n, \\ \lambda^T (\mathbf{A}\mathbf{x}^* - \mathbf{b}) = \mathbf{0}_n. \end{cases}$$

Problem 1

Prove the non-homogeneous Farkas' lemma: Let $\mathbf{A} \in \mathbb{R}^{m \times n}$, $\mathbf{c} \in \mathbb{R}^n$, $\mathbf{b} \in \mathbb{R}^m$ and $d \in \mathbb{R}$. Suppose that there exists $0 \leq \mathbf{y}_0 \in \mathbb{R}^m$ such that $\mathbf{A}^T \mathbf{y}_0 = \mathbf{c}$. Prove that exactly one of the following two systems is feasible:

(I) $\quad \mathbf{A}\mathbf{x} \leq \mathbf{b}$, $\mathbf{c}^T \mathbf{x} > d$.

(II) $\quad \mathbf{A}^T \mathbf{y} = \mathbf{c}$, $\mathbf{b}^T \mathbf{y} \leq d$, $\mathbf{y} \geq 0$.

Solution

$\Leftarrow$: If (II) is feasible then (I) is infeasible ($(II) \implies \neg(I)$):

Note: equivalently, we can show that

- If (I) is feasible then (II) is infeasible; or,
- (I) and (II) cannot both be feasible.

Now we prove the required:

- Assume on the contrary that (I) holds. From (I) we have $\mathbf{A}\mathbf{x} \leq \mathbf{b} \implies \mathbf{x}^T \mathbf{A}^T \leq \mathbf{b}^T$, and for any $\mathbf{y} \geq 0$ we have $\mathbf{x}^T \mathbf{A}^T \mathbf{y} \leq \mathbf{b}^T \mathbf{y}$.
- Since (II) holds, there exists $\mathbf{y} \geq 0$ such $\mathbf{A}^T \mathbf{y} = \mathbf{c}$. Plugging in the above yields $\mathbf{x}^T \mathbf{c} \leq \mathbf{b}^T \mathbf{y}$.
- From (I) and (II) we have $d < \mathbf{x}^T \mathbf{c} \leq \mathbf{b}^T \mathbf{y} \leq d$ which is a contradiction.

$\Rightarrow$: If (II) is infeasible then (I) is feasible ($\neg(II) \implies (I)$):

Note: equivalently, we can show that

- If (I) is infeasible then (II) is feasible; or,
- (I) and (II) cannot both be infeasible.

Proof using (homogeneous) Farkas' lemma:

- Define the following equivalent system to (II):

$$(II') \quad \mathbf{A}^T \mathbf{y} = \mathbf{c}, \mathbf{b}^T \mathbf{y} + t = d, t \geq 0, \mathbf{y} \geq 0.$$

  - (II) and (II') are equivalent in the sense that if (II) has a solution $\mathbf{y} \geq 0$, then $\left(\mathbf{y}, t = d - \mathbf{b}^T \mathbf{y}\right) \geq 0$ is a solution to (II').
  - If (II') has a solution $(\mathbf{y}, t) \geq 0$, then $\mathbf{y} \geq 0$ is a solution to (II). So (II) has a solution if and only if (II') has a solution.

- We notice that we can rewrite (II') as

$$\begin{pmatrix} \mathbf{A}^T & \mathbf{0}_n \\ -\mathbf{b}^T & -1 \end{pmatrix} \begin{pmatrix} \mathbf{y} \\ t \end{pmatrix} = \begin{pmatrix} \mathbf{c} \\ -d \end{pmatrix}, \quad \begin{pmatrix} \mathbf{y} \\ t \end{pmatrix} \geq 0.$$

- Since we assume that (II) is infeasible, then (II') is infeasible.
- Using (homogeneous) Farkas' lemma, the following system is feasible

$$\begin{pmatrix} \mathbf{A} & -\mathbf{b} \\ \mathbf{0}_n^T & -1 \end{pmatrix} \begin{pmatrix} \mathbf{x} \\ s \end{pmatrix} \leq \mathbf{0}_{n+1}, \quad \begin{pmatrix} \mathbf{c} \\ -d \end{pmatrix}^T \begin{pmatrix} \mathbf{x} \\ s \end{pmatrix} > 0, \quad \begin{pmatrix} \mathbf{x} \\ s \end{pmatrix} \in \mathbb{R}^{n+1}.$$

- Writing this system explicitly we have $\mathbf{A}\mathbf{x} \leq s\mathbf{b}$, $0 \leq s$, $\mathbf{c}^T \mathbf{x} > ds$. We will show that $s > 0$.
  - If $s > 0$ then $\mathbf{A}\left(\frac{\mathbf{x}}{s}\right) \leq \mathbf{b}$, $\mathbf{c}^T \left(\frac{\mathbf{x}}{s}\right) > d$ and (I) has a solution.

- Assume $s = 0$, meaning $\mathbf{Ax} \le s\mathbf{b} = \mathbf{0}_n$ and $\mathbf{c}^T\mathbf{x} > 0$.
  * We know that there exists $0 \le \mathbf{y}_0 \in \mathbb{R}^m$ such that $\mathbf{A}^T\mathbf{y}_0 = \mathbf{c}$.
  * Therefore, from the (homogeneous) Farkas' lemma we derive that there is no solution to the system $\mathbf{Ax} \le \mathbf{0}_n$ and $\mathbf{c}^T\mathbf{x} > 0$, which is a contradiction.
- Therefore, $s > 0$ and (I) has a solution.

Proof using the separation theorem:

- Consider the closed and convex set

$$S = \left\{(\mathbf{z}, w) \in \mathbb{R}^{n+1} \colon \exists \mathbf{y} \ge 0 \text{ such that } \mathbf{z} = \mathbf{A}^T\mathbf{y},\ \mathbf{b}^T\mathbf{y} \le w\right\}.$$

  (It closed and convex since it is the image of a closed and convex set in $(\mathbf{z}, w, \mathbf{y})$ under the linear projection onto $(\mathbf{z}, w)$).
- Since we assume that (II) is infeasible, then $(\mathbf{c}, d) \notin S$.
- By the separation theorem, there exist $(\mathbf{p}, q) \in \mathbb{R}^{n+1} \setminus \{\mathbf{0}_n\}$ and $\alpha \in \mathbb{R}$ such that $\mathbf{p}^T\mathbf{c} + qd < \alpha$ and $\mathbf{p}^T\mathbf{z} + qw \ge \alpha$ for any $(\mathbf{z}, w) \in S$.
- If $q = 0$ then $\mathbf{p} \ne \mathbf{0}$, and so $\mathbf{p}^T\mathbf{c} < \alpha \le \mathbf{p}^T\mathbf{z} = \mathbf{p}^T\mathbf{A}^T\mathbf{y}$.
  - Therefore $\mathbf{p}^T\left(\mathbf{c} - \mathbf{A}^T\mathbf{y}\right) < 0$ for any $\mathbf{y} \ge 0$, which contradicts the fact that $\mathbf{A}^T\mathbf{y}_0 = \mathbf{c}$ for $\mathbf{y}_0 \ge 0$.
- If $q < 0$ we can take $w \to \infty$, in contradiction to the fact that $\mathbf{p}^T\mathbf{z} + qw \ge \alpha$ (notice that we cannot take $w \to -\infty$ since $\mathbf{b}^T\mathbf{y} \le w$).
- So $q > 0$. We divide by q and get $\tilde{\mathbf{p}}^T\mathbf{c} + d < \tilde{\alpha}$ and $\tilde{\mathbf{p}}^T\mathbf{z} + w \ge \tilde{\alpha}$, where $\tilde{\mathbf{p}} = \frac{\mathbf{p}}{q}$ and $\tilde{\alpha} = \frac{\alpha}{q}$.
  - Choose $w = \mathbf{b}^T\mathbf{y}$ and then

$$\tilde{\alpha} \le \tilde{\mathbf{p}}^T\mathbf{z} + \mathbf{b}^T\mathbf{y} = \tilde{\mathbf{p}}^T\mathbf{A}^T\mathbf{y} + \mathbf{b}^T\mathbf{y} = \mathbf{y}^T\left(\mathbf{A}\tilde{\mathbf{p}} + \mathbf{b}\right).$$

  - Since this is true for any $\mathbf{y} \ge 0$ then $\mathbf{A}\tilde{\mathbf{p}} + \mathbf{b} \ge 0$ (otherwise $\mathbf{y}^T\left(\mathbf{A}\tilde{\mathbf{p}} + \mathbf{b}\right)$ is not necessarily bounded from below by $\tilde{\alpha}$).
  - Therefore $\mathbf{A}\left(-\tilde{\mathbf{p}}\right) \le \mathbf{b}$ and $\left(-\tilde{\mathbf{p}}\right)^T\mathbf{c} > d$. This means that (I) has a solution as required.

Problem 2

Let $\mathbf{Q} \in \mathbb{R}^{n\times n}$, $\mathbf{A} \in \mathbb{R}^{m\times n}$, $\mathbf{b} \in \mathbb{R}^n$ and $\mathbf{c} \in \mathbb{R}^m$. Suppose that $\mathbf{Q} \succ 0$ and that $\mathbf{A}$ has full row rank. Solve the following problem:

$$\begin{aligned} \min_{\mathbf{x}\in\mathbb{R}^n} \quad & \frac{1}{2}\mathbf{x}^T\mathbf{Q}\mathbf{x} - \mathbf{b}^T\mathbf{x} \\ \text{s.t.} \quad & \mathbf{Ax} = \mathbf{c}. \end{aligned}$$

Solution

- The problem is feasible since $\mathbf{A}$ has full row rank and so $\mathbf{x} = \mathbf{A}^T\left(\mathbf{A}\mathbf{A}^T\right)^{-1}\mathbf{c}$ is a solution.
- The problem admits a minimizer since $\mathbf{Q} \succ 0$ so the objective is coercive over a closed set.
- The Lagrangian is

$$L\left(\mathbf{x}, \mu\right) = \frac{1}{2}\mathbf{x}^T\mathbf{Q}\mathbf{x} - \mathbf{b}^T\mathbf{x} + \mu^T\left(\mathbf{Ax} - \mathbf{c}\right).$$

- Since the problem is convex with linear constraints, a solution is optimal if and only if it is a KKT point. The KKT conditions are:

$$\begin{cases} \nabla_{\mathbf{x}}L\left(\mathbf{x}, \mu\right) = \mathbf{Qx} - \mathbf{b} + \mathbf{A}^T\mu = \mathbf{0}_n, & (i) \\ \mathbf{Ax} = \mathbf{c}. & (ii) \end{cases}$$

- From (i) we have $\mathbf{x} = \mathbf{Q}^{-1}\left(\mathbf{b} - \mathbf{A}^T\mu\right)$. Plugging into (ii) (feasibility constraint) we have

$$\mathbf{A}\mathbf{Q}^{-1}\left(\mathbf{b} - \mathbf{A}^T\mu\right) = \mathbf{c} \Longleftrightarrow \mathbf{A}\mathbf{Q}^{-1}\mathbf{A}^T\mu = \mathbf{A}\mathbf{Q}^{-1}\mathbf{b} - \mathbf{c}.$$

- We will show that $\mathbf{A}\mathbf{Q}^{-1}\mathbf{A}^T \succ 0$: notice that for any $\mathbf{y} \in \mathbb{R}^m$ we have

$$\begin{aligned} \mathbf{y}^T\mathbf{A}\mathbf{Q}^{-1}\mathbf{A}^T\mathbf{y} = 0 \Longleftrightarrow \left(\mathbf{A}^T\mathbf{y}\right)\mathbf{Q}^{-1}\left(\mathbf{A}^T\mathbf{y}\right) = 0 &\underset{\mathbf{Q}^{-1}\succ 0}{\Longleftrightarrow} \mathbf{A}^T\mathbf{y} = \mathbf{0}_n \\ &\underset{\mathbf{A}\text{ full row rank}}{\Longleftrightarrow} \sum_{i=1}^{n}\mathbf{A}_i^T\mathbf{y}_i = \mathbf{0}_n \Longleftrightarrow \mathbf{y} = \mathbf{0}_n, \end{aligned}$$

  and therefore $\mathbf{A}\mathbf{Q}^{-1}\mathbf{A}^T \succ 0$.

- This means that $\mu = \left(\mathbf{A}\mathbf{Q}^{-1}\mathbf{A}^T\right)^{-1}\left(\mathbf{A}\mathbf{Q}^{-1}\mathbf{b} - \mathbf{c}\right) \in \mathbb{R}^m$. So

$$\mathbf{x} = \mathbf{Q}^{-1}\left(\mathbf{b} - \mathbf{A}^T\left(\mathbf{A}\mathbf{Q}^{-1}\mathbf{A}^T\right)^{-1}\left(\mathbf{A}\mathbf{Q}^{-1}\mathbf{b} - \mathbf{c}\right)\right)$$

  is a KKT point and thus an optimal solution.