E. Herberg, M. Marić, V. Stein
Heidelberg University

Exercise 9 - Solution

Date issued: 10th June 2024
Date due: 18th June 2024

Homework Problem 9.1 (Lin. Cone and CQs Depend on Description of Feasible Set) 3 Points

The optimization problems

$$\begin{aligned} \text{Minimiere} \quad & f(x) \quad \text{über } x \in \mathbb{R} \\ \text{unter} \quad & x = 0 \end{aligned} \tag{P_1}$$

and

$$\begin{aligned} \text{Minimiere} \quad & f(x) \quad \text{über } x \in \mathbb{R} \\ \text{unter} \quad & x^2 = 0 \end{aligned} \tag{P_2}$$

for any $f \in C^1(\mathbb{R})$ have their obvious solution (because sole feasible point) at $x^* = 0$.

Show that the Abadie and Guignard constraint qualifications are satisfied at $x^* = 0$ for (P_1) but not (P_2).

Solution.

We are dealing with the two feasible sets

$$F^{(1)} = \{x \in \mathbb{R} \mid \underbrace{x}_{h^{(1)}(x)} = 0\}, \quad F^{(2)} = \{x \in \mathbb{R} \mid \underbrace{x^2}_{h^{(2)}(x)} = 0\}$$

Both coincide with the singleton $F^{(1)} = F^{(2)} = \{0\} =: F$.

Accordingly, the tangent cones coincide as well, as they don't depend on the description of the feasible set. Since the only sequence in the feasible sets is the constant zero sequence, we know that

$$\mathcal{T}_F(x) = \{0\} \text{ and therefore } \mathcal{T}_F(x)^\circ = \mathbb{R}.$$

For linearizing cones, we compute

$$\begin{aligned} h^{(1)\prime}(x) &= 1 \quad &\Rightarrow \quad h^{(1)\prime}(x^*) &= 1 \\ h^{(2)\prime}(x) &= 2x \quad &\Rightarrow \quad h^{(2)\prime}(x^*) &= 0 \end{aligned}$$

and therefore obtain that

$$\begin{aligned} \mathcal{T}^{\text{lin}}_{F^{(1)}}(x^*) &= \{0\} \quad \text{and} \quad \mathcal{T}^{\text{lin}}_{F^{(2)}}(x^*)^\circ &= \mathbb{R} \quad &\Leftrightarrow \quad \text{ACQ/GCQ} \\ \mathcal{T}^{\text{lin}}_{F^{(2)}}(x^*) &= \mathbb{R} \quad \text{and} \quad \mathcal{T}^{\text{lin}}_{F^{(2)}}(x^*)^\circ &= \{0\} \quad &\Leftrightarrow \text{no ACQ/GCQ}. \end{aligned}$$

Note: The message here is that unnecessarily increasing the order of a constraint can remove linearization information and kill CQs.

(3 Points)

Homework Problem 9.2 (ACQ for Problems with Affine Constraints) 6 Points

Consider

$$F := \left\{x \in \mathbb{R}^n \,\middle|\, g_i(x) \le 0 \text{ for all } i = 1, \ldots, n_{\text{ineq}},\; h_j(x) = 0 \text{ for all } j = 1, \ldots, n_{\text{eq}}\right\} \tag{5.2}$$

and

$$F^{\text{lin}}(x) = \left\{ y \in \mathbb{R}^n \,\middle|\, \begin{aligned} g_i(x) + g_i'(x)\,(y - x) &\le 0 \quad \text{for all } i = 1, \ldots, n_{\text{ineq}} \\ h_j(x) + h_j'(x)\,(y - x) &= 0 \quad \text{for all } j = 1, \ldots, n_{\text{eq}} \end{aligned} \right\}$$

for $x \in F$.

(i) Show that $\mathcal{T}^{\text{lin}}_F(x) = \mathcal{T}_{F^{\text{lin}}(x)}(x)$ for $x \in F$. (Remark 5.6 Statement (i))

(ii) Show that $\mathcal{T}^{\text{lin}}_F(x)$ is a closed convex cone. (Remark 5.6 Statement (ii))

(iii) Prove Theorem 6.9 by showing that the Abadie CQ holds at any feasible point of problems of the form

$$\left.\begin{aligned} \text{Minimize} \quad & f(x) \qquad \text{where } x \in \mathbb{R}^n \\ \text{subject to} \quad & A_{\text{ineq}}\, x \le b_{\text{ineq}} \\ \text{and} \quad & A_{\text{eq}}\, x = b_{\text{eq}} \end{aligned}\right\} \tag{6.10}$$

Solution.

(*i*) Let $x \in F$.

To show that $\mathcal{T}_F^{\text{lin}}(x) \subseteq \mathcal{T}_{F^{\text{lin}}(x)}(x)$, let $d \in \mathcal{T}_F^{\text{lin}}(x)$. By definition,

$$\begin{aligned} g_i'(x)\, d &\le 0 \qquad \text{for all } i \in \mathcal{A}(x) \\ h_j'(x)\, d &= 0 \qquad \text{for all } j = 1, \dots, n_{\text{eq}} \end{aligned}. \tag{0.1}$$

For all $t > 0$ with

$$t < \inf_{i \in \mathcal{I}(x), g_i'(x)d > 0} \frac{g_i(x)}{g_i'(x)d}$$

we have that

$$\begin{aligned} g_i(x) + g_i'(x)\,((x+td) - x) &= g_i(x) + g_i'(x)\, td && \le 0 && \text{for all } i = 1, \dots, n_{\text{ineq}} \\ h_j(x) + h_j'(x)\,((x+td) - x) &= h_j(x) + h_j'(x)\, td = h_j(x) && = 0 && \text{for all } j = 1, \dots, n_{\text{eq}} \end{aligned}$$

so $x + td \in F^{\text{lin}}(x)$ for all of those t, so choosing $y^{(k)} = x + t^{(k)}d$ for any sufficiently small positive nullsequence $t^{(k)}$ yields a sequence in $F^{\text{lin}}(x)$ such that

$$\frac{y^{(k)} - x}{t^{(k)}} = d \longrightarrow d$$

so $d \in \mathcal{T}_{F^{\text{lin}}(x)}(x)$.

To show that $\mathcal{T}_F^{\text{lin}}(x) \supseteq \mathcal{T}_{F^{\text{lin}}(x)}(x)$, let $d \in \mathcal{T}_{F^{\text{lin}}(x)}(x)$. Then there exists a sequence $y^{(k)}$ in $F^{\text{lin}}(x)$ and a positive nullsequence $t^{(k)}$ such that $d^{(k)} := \frac{y^{(k)} - x}{t^{(k)}} \longrightarrow d$. The definition of $F^{\text{lin}}(x)$ and the feasibility of x immediately implies (0.1) for $d^{(k)}$ instead of d, so the same holds for d due to continuity.

(*ii*) Closedness is an immediate consequence of the continuous differentiability of g and h and convexity is clear because both conditions defining the linearizing cone are – by design – linear.

(*iii*) The feasible set is of the form

$$\begin{aligned} g(x) &:= A_{\text{ineq}}\, x - b_{\text{ineq}} \le 0 \\ \text{and} \quad h(x) &:= \; A_{\text{eq}}\, x - b_{\text{eq}} = 0 \end{aligned}$$

with affine linear constraints g and h, where $g' \equiv A_{\text{ineq}}$ and $h' \equiv A_{\text{eq}}$.

Accordingly, for any feasible $x \in F$, we have that

$$\begin{aligned} g_i(x) + g_i'(x)\,(y - x) &= A_{\text{ineq}}\, x - b_{\text{ineq}} + A_{\text{ineq}}(y - x) && = A_{\text{ineq}}\, y - b_{\text{ineq}} && = g(y) && \text{for all } i = 1, \dots, n_{\text{ineq}} \\ h_j(x) + h_j'(x)\,(y - x) &= A_{\text{eq}}\, x - b_{\text{eq}} + A_{\text{ineq}}(y - x) && = A_{\text{eq}}\, y - b_{\text{eq}} && = h(y) && \text{for all } j = 1, \dots, n_{\text{eq}} \end{aligned},$$

which means that $F^{\text{lin}}(x) = F$ for every $x \in F$ and therefore

$$\mathcal{T}_F^{\text{lin}}(x) = \mathcal{T}_{F^{\text{lin}}(x)}(x) = \mathcal{T}_F(x)$$

which is exactly the ACQ.

(6 Points)

Homework Problem 9.3 (Alternative formulation of the second MFCQ condition) 6 Points

Prove Lemma 6.11, i.e. show that the following three statements are equivalent for $x \in F$.

(ii) There exists a vector $d \in \mathbb{R}^n$ such that

$$\begin{aligned} g_i'(x)\, d &< 0 \quad \text{for all } i \in \mathcal{A}(x), \\ h_j'(x)\, d &= 0 \quad \text{for all } j = 1, \dots, n_{\text{eq}}. \end{aligned}$$

(iii) There exists a vector $d \in \mathbb{R}^n$ such that

$$\begin{aligned} g(x) + g'(x)\, d &< 0, \\ h(x) + h'(x)\, d &= 0. \end{aligned}$$

(iv) There exists a vector $d \in \mathbb{R}^n$ such that

$$\begin{aligned} g_i'(x)\, d &\le -1 \qquad \text{for all } i \in \mathcal{A}(x), \\ h_j'(x)\, d &= 0 \qquad \text{for all } j = 1, \dots, n_{\text{eq}}. \end{aligned}$$

Solution.

$(ii) \Rightarrow (iii)$

Let $d \in \mathbb{R}^n$ fulfill (ii), then we get

$$g_i(x) + g_i'(x)\, d = \begin{cases} \underbrace{g_i(x)}_{=0} + \underbrace{g_i'(x)\, d}_{<0} < 0 & \text{for all } i \in \mathcal{A}(x), \\ \underbrace{g_i(x)}_{<0} + g_i'(x)\, d & \text{for all } i \in \mathcal{I}(x), \end{cases}$$

$$\underbrace{h_j(x)}_{=0\,(x \in F)} + \underbrace{h_j'(x)\, d}_{=0\,(ii)} = 0 \qquad \text{for all } j = 1, \dots, n_{\text{eq}}.$$

Now let $c > 0$, such that $\tilde{d} := c\, d \in \mathbb{R}^n$ fulfills

$$g_i'(x)\, \tilde{d} < -g_i(x) \qquad \text{for all } i \in \mathcal{I}(x)$$

This is always possible, since $-g_i(x) > 0$ for all $i \in \mathcal{I}(x)$. Furthermore, multiplication with a positive scalar does not disrupt our previous findings, so we have constructed a suitable $\tilde{d} \in \mathbb{R}^n$ and (iii) holds. (2 Points)

$(iii) \Rightarrow (ii)$

Let $d \in \mathbb{R}^n$ fulfill (iii), then it holds

$$g_i'(x)\, d = \underbrace{g_i(x)}_{=0} + g_i'(x)\, d < 0 \qquad \text{for all } i \in \mathcal{A}(x)$$

$$h_j'(x)\, d = \underbrace{h_j(x)}_{=0} + h_j'(x)\, d = 0 \qquad \text{for all } j = 1, \dots, n_{\text{eq}}$$

Consequently, (ii) is true. (1 Point)

$(ii) \Rightarrow (iv)$

Let $d \in \mathbb{R}^n$ fulfill (ii). Then for every $i \in \mathcal{A}(x)$ we can find $c_i > 0$, such that

$$g_i'(x)\, d = -c_i.$$

With $\tilde{d} := \min_{i \in \mathcal{A}(x)} \{\frac{1}{c_i}\}\, d$ we have

$$g_i'(x)\, \tilde{d} \le -1 \qquad \text{for all } i \in \mathcal{A}(x).$$

Additionally, we have

$$h_j'(x)\, \tilde{d} = \min_{i \in \mathcal{A}(x)} \{\frac{1}{c_i}\} \underbrace{h_j'(x)\, d}_{=0} = 0 \qquad {}^{\text{for all}}\ j = 1, \dots, n_{\text{eq}}.$$

Altogether, we found $\tilde{d} \in \mathbb{R}^n$ that fulfills the desired conditions. (2 Points)

$(iv) \Rightarrow (ii)$

For $d \in \mathbb{R}^n$ fulfilling (iv) we directly see that it also fulfills (ii), since $-1 < 0$. (1 Point)

Homework Problem 9.4 (Multiplier Compactness is Equivalent to MFCQ) 6 Points

(i) Use Farkas' Lemma (Lemma 5.11 in the lecture notes) to show that for $A \in \mathbb{R}^{p \times n}$ and $B \in \mathbb{R}^{n_{\text{eq}} \times n}$ with $\operatorname{rank}(B) = n_{\text{eq}}$ and $p \leq n_{\text{ineq}}$ either the system

$$Ad < 0, \quad Bd = 0 \tag{0.2}$$

has a solution $d \in \mathbb{R}^n$ or

$$A^\top \mu + B^\top \lambda = 0 \tag{0.3}$$

has a solution $(\mu, \lambda) \neq 0$ with $\mu \geq 0$.

Hint: Start with the existence of a nontrivial solution to (0.3). Focus the nontriviality on μ. Transform the conditions $\mu \neq 0, \mu \geq 0$ into a linear condition with a sign condition using a normalization step with respect to $\|\cdot\|_1$. Split λ into its positive and negative part. Apply Farkas' Lemma. Success.

(ii) Let (x^*, λ^*, μ^*) be a KKT-point of (5.1). Show that MFCQ is satisfied at x^* if and only if the set of Lagrange multipliers that solve the KKT system for x^* is compact.

Solution.

(i) We show that existence of a solution (μ, λ) with $\mu \geq 0$ for the system

$$A^\top \mu + B^\top \lambda = 0$$

is equivalent to there not existing a $d \in \mathbb{R}^n$ such that

$$Ad < 0, \quad Bd = 0.$$

Farkas' Lemma (extended by a simple logical negation) states that the following are equivalent for $\widetilde{B} \in \mathbb{R}^{m \times n}$ and $\widetilde{c} \in \mathbb{R}^n$. Then the following are equivalent.

(a) The linear system $\widetilde{B}^\top \xi = \widetilde{c}$ has a non-negative solution $\xi \geq 0$.

(b) All elements of $\{d \in \mathbb{R}^n \mid \widetilde{B} d \geq 0\}$ satisfy $\widetilde{c}^\top d \geq 0$

(c) There is no element of $\{d \in \mathbb{R}^n \mid \widetilde{B} d \geq 0\}$ that satisfies $\widetilde{c}^\top d < 0$. .

And hence we can observe the following equivalences, which show the claim:

$$
\begin{aligned}
& A^\mathsf{T}\mu + B^\mathsf{T}\lambda = 0 \text{ has a solution with } (\mu,\lambda) \neq 0, \mu \geq 0 \\
\Leftrightarrow \quad & A^\mathsf{T}\mu + B^\mathsf{T}\lambda = 0 \text{ has a solution with } \mu \neq 0, \mu \geq 0 && \text{(because } \operatorname{rank}(B) = n_{\text{eq}}) \\
\Leftrightarrow \quad & A^\mathsf{T}\mu + B^\mathsf{T}\lambda = 0 \text{ has a solution with } \mu \geq 0 \text{ and } \sum \mu_i = \|\mu\|_1 = 1 && \text{(divide by } \|\mu\|_1) \\
\Leftrightarrow \quad & \underbrace{\begin{pmatrix} A^\mathsf{T} & B^\mathsf{T} & -B^\mathsf{T} \\ \mathbf{1}^\mathsf{T} & 0 & 0 \end{pmatrix}}_{\widetilde{B}^\mathsf{T}} \underbrace{\begin{pmatrix} \mu \\ \lambda^+ \\ \lambda^- \end{pmatrix}}_{\xi} = \underbrace{\begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}}_{\widetilde{c}} \text{ has a solution with} \begin{pmatrix} \mu \\ \lambda^+ \\ \lambda^- \end{pmatrix} \geq 0 && \text{(split } \lambda) \\
\Leftrightarrow \quad & \nexists \begin{pmatrix} d \\ d_0 \end{pmatrix} \text{ such that } \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}^\mathsf{T} \begin{pmatrix} d \\ d_0 \end{pmatrix} < 0 \text{ and } \begin{pmatrix} A & \mathbf{1} \\ B & 0 \\ -B & 0 \end{pmatrix} \begin{pmatrix} d \\ d_0 \end{pmatrix} \geq 0 && \text{(Farkas' Lemma)} \\
\Leftrightarrow \quad & \nexists d : Ad < 0,\ Bd = 0 && \text{(replacing } d \text{ with } -d)
\end{aligned}
$$

Note that the first couple of equivalent transformations were simply en effort to rewrite the nontriviality conditions of (0.3) as an equality that we could include in the Farkas Lemma (for which we needed to split λ in its positive and negative part. (3 Points)

(*ii*) If we set

$$A = \big(g_i'(x^*)\big)_{i \in \mathcal{A}(x^*)} \quad \text{and} \quad B = h'(x^*),$$

then the system (0.2) simply states that there is no MFCQ direction and (0.3) gives us a nontrivial element of the kernel of the matrix that appears in the KKT stationarity condition (i. e., we can scale that solution and add it to the multiplier wihtout breaking the KKT condition).

So: Assume that MFCQ is not satisfied at x^*, i. e., there is no $d \in \mathbb{R}^n$ such that $h'(x^*)d = 0$ and $g_i'(x^*)d \leq 0$ for all $i \in \mathcal{A}(x^*)$ then we invoke Statement (*i*) to obtain a non-trivial solution $(\widetilde{d\mu}, d\lambda) \neq 0$, $\widetilde{d\mu} \geq 0$ such that $A^\mathsf{T}\widetilde{d\mu} + B^\mathsf{T} d\lambda = 0$ (the kernel element with additional structure). We extend $\widetilde{d\mu} \in \mathbb{R}^{|\mathcal{A}(x^*)|}$ to $\mu \in \mathbb{R}^n_{\text{ineq}}$ by setting $\widetilde{d\mu}_i = 0$ for $i \notin \mathcal{A}(x^*)$. Then the ray $(\mu^* + t d\mu, \lambda^* + t d\lambda)$ for $t \geq 0$ is unbounded and all points on are multipliers for the KKT system for x^*, i. e. the set of Lagrange multipliers for x^* is not compact. (1.5 Points)

Now, assume that the set of Lagrange multipliers at x^* is non-compact. Because the Lagrange multipliers for the KKT system and for x^* are the intersection of a linear subspace (the solution space of the linear system intersected with the hyperplane defined by $g(x)$) and the set $\{(\mu,\lambda) \mid \mu \geq 0\}$, both of which are closed, the only way that this set can be non-compact is by being unbounded, so it contains a ray $(\mu^* + td\mu, \lambda + td\lambda)$ for $(d\mu, d\lambda) \neq 0$ and $t \geq 0$. From $\mu^* + td\mu \geq 0$ for all $t \geq 0$, we can summize that $d\mu \geq 0$ and therefore $(d\mu_{(x^*)}, d\lambda)$ is a non-trivial solution of $A^\intercal \mu + B^\intercal \lambda = 0,\ \mu \geq 0$. Due to Statement (i), we know that MFCQ is not satisfied at x^*. (1.5 Points)

Please submit your solutions as a single pdf and an archive of programs via moodle.