R. Herzog, G. Müller, E. Herberg, M. Hashemi
Heidelberg University

Exercise 9 - Solution

Date issued: 12th June 2023
Date due: 20th June 2023

Homework Problem 9.1 (Finding Solutions using First and Second Order Information) 6 Points

Consider the problem

$$\left.\begin{aligned} \text{Maximize} \quad & -(x_1-2)^2 - 2(x_2-1)^2 \quad \text{where } x \in \mathbb{R}^2 \\ \text{subject to} \quad & x_1 + 4x_2 \le 3 \\ \text{and} \quad & x_1 \ge x_2 \end{aligned}\right\}$$

Determine, which admissible points satisfy a constraint qualification (ACQ/GCQ/MFCQ/LICQ) and use first and second order information to compute all stationary points and solve the problem, i. e., find all optima and explain why they are local and/or global solutions.

Solution.

We rewrite the problem as a minimization problem by considering the negative of the cost function and consequently obtain

$$\begin{aligned} f(x) &= (x_1-2)^2 + 2(x_2-1)^2 & g_1(x) &= x_1 + 4x_2 - 3 & g_2(x) &= x_2 - x_1, \\ f'(x) &= (2x_1 - 4, 4x_2 - 4) & g_1'(x) &= (1, 4) & g_2'(x) &= (-1, 1). \end{aligned}$$

Since $g_i'(x) \neq 0$ for all x and $i = 1, 2$ and both are obviously always linear independent, the LICQ is satisfied at all feasible x. (1 Point)

The corresponding KKT system for a multiplier $\mu \ge 0 \in \mathbb{R}^2$ and feasible $x = (x_1, x_2)^\intercal$ is

$$\begin{aligned} 2x_1 - 4 + \mu_1 - \mu_2 &= 0 \\ 4x_2 - 4 + 4\mu_1 + \mu_2 &= 0 \\ \mu_1(x_1 + 4x_2 - 3) &= 0 \\ \mu_2(x_2 - x_1) &= 0. \end{aligned}$$

Due to the complementarity conditions, there are four cases:

(i) Both constraints are inactive ($\mu_1 = 0,\ \mu_2 = 0$): The KKT system will then reduce to

$$\begin{aligned} 2x_1 - 4 &= 0 \\ 4x_2 - 4 &= 0 \end{aligned}$$

with the solution $x^* = (x_1^*, x_2^*) = (2, 1)$, which violates the first constraint so there is no stationary point, and hence no solution, for this case. (1 Point)

(ii) The first constraint is inactive and the second is active ($\mu_1 = 0$ and $x_1 = x_2$): Here the KKT system is

$$\begin{aligned} 2x_1 - 4 - \mu_2 &= 0, \\ 4x_2 - 4 + \mu_2 &= 0, \\ x_1 &= x_2. \end{aligned}$$

Adding the first two equations and using the active second constraint, we obtain the solution $\left(\frac{4}{3}, \frac{4}{3}\right)$ with $\mu_2 = -\frac{4}{3}$, which violates the nonnegativity condition so there is no stationary point, and hence no solution, for this case. (1 Point)

(iii) The second constraint is inactive and the first constraint is active ($\mu_2 = 0$ and $x_1 + 4x_2 = 3$): Here, the KKT system is

$$\begin{aligned} 2x_1 - 4 + \mu_1 &= 0 \\ 4x_2 - 4 + 4\mu_1 &= 0 \\ x_1 + 4x_2 &= 3. \end{aligned}$$

Adding the second equation to -4 times the first equation, we get

$$\begin{aligned} -8x_1 + 4x_2 &= -12 \\ x_1 + 4x_2 &= 3 \end{aligned}$$

with the unique solution $\left(\frac{5}{3}, \frac{1}{3}\right)$ and $\mu_1 = \frac{2}{3}$. This is a stationary point as x is feasible an μ is nonnegative. (1 Point)

(iv) Both constraints are active ($x_1 + 4x_2 = 3$ and $x_1 = x_2$): Here the KKT system is

$$\begin{aligned} 2x_1 - 4 + \mu_1 - \mu_2 &= 0 \\ 4x_2 - 4 + 4\mu_1 + \mu_2 &= 0 \\ x_1 + 4x_2 - 3 &= 0 \\ x_2 - x_1 &= 0 \end{aligned}$$

with the solution $\left(\frac{3}{5}, \frac{3}{5}\right)$ and $\mu_1 = \frac{22}{25}$, $\mu_2 = -\frac{48}{25}$, which violates the nonnegativity constaint for μ, so there is no stationary point, and hence no solution, for this case. (1 Point)

The feasible set is not compact, so we need second order information to check the optimality of the stationary point. We compute

$$\mathcal{L}_{xx}(x, \mu) = \begin{pmatrix} 2 & 0 \\ 0 & 4 \end{pmatrix}$$

which is positive definite. Hence, the solution to the problem is a local maximizer of the original (maximization) problem at $(\frac{5}{3}, \frac{1}{3})$, with functional value $-J(\frac{5}{3}, \frac{1}{3}) = 1$. Due to convexity of the cost functional and the feasible set, the local maximizer is the only maximizer and is a global one. (1 Point)

Homework Problem 9.2 (Comparing the Strength of CQs) 6 Points

From the lecture notes, we know that

$$\boxed{\text{LICQ}} \xRightarrow{\text{Lemma 8.17}} \boxed{\text{MFCQ}} \xRightarrow{\text{Corollary 8.14}} \boxed{\text{ACQ}} \xRightarrow{\text{Definition 8.6}} \boxed{\text{GCQ}}. \tag{8.15}$$

Show that generally

$$\boxed{\text{LICQ}} \overset{\text{(P3)}}{\not\Leftarrow} \boxed{\text{MFCQ}} \overset{\text{(P2)}}{\not\Leftarrow} \boxed{\text{ACQ}} \overset{\text{(P1)}}{\not\Leftarrow} \boxed{\text{GCQ}}$$

by investigating the following problems P1 to P3 at $x^* = (0, 0)^\intercal$:

$$\left.\begin{aligned} \text{Minimize} \quad & f(x) \qquad \text{where } x \in \mathbb{R}^2 \\ \text{subject to} \quad & x_1 \le 0 \\ & x_2 \le 0 \\ & x_1 x_2 = 0 \end{aligned}\right\} \tag{P1}$$

$$\left.\begin{aligned} \text{Minimize} \quad & f(x) \qquad \text{where } x \in \mathbb{R}^2 \\ \text{subject to} \quad & q(x_1) - x_2 \le 0 \\ & q(x_1) + x_2 \le 0 \end{aligned}\right\} \quad \text{for} \quad q(x_1) := \begin{cases} (x_1 + 1)^2, & x_1 < -1, \\ 0, & -1 \le x_1 \le 1, \\ (x_1 - 1)^2, & x_1 > 1, \end{cases} \tag{P2}$$

$$\left.\begin{aligned} \text{Minimize} \quad & f(x) \qquad \text{where } x \in \mathbb{R}^2 \\ \text{subject to} \quad & -x_1^3 - x_2 \le 0 \\ & -x_2 \le 0 \end{aligned}\right\} \tag{P3}$$

Solution.

(i) We start out showing that LICQ is not satisfied at the origin in problem P3, but MFCQ is. **Note:** LICQ requires that the active constraints do not carry the same tangent information while MFCQ only requires that there is a direction pointing into the interior of all active inequality constraints' feasibility regions. The example was constructed using this information.

At the point in question, both inequality constraints of problem P3 are active and we have

$$\begin{aligned} g_1'(x) &= \left(-3x_1^2, -1\right) \quad &\implies g_1'(x^*) = (0, -1) \\ g_2'(x) &= (0, -1) \quad &\implies g_2'(x^*) = (0, -1) \end{aligned}$$

so LICQ in fact does not hold. However, as there are no equality constraints, the direction $d = (0, 1)^\intercal$ shows that MFCQ in fact holds.

(ii) Next up, we show that MFCQ is violated at the origin for problem P2, but ACQ holds. **Note:** MFCQ requires that there is a direction pointing into the interior of all active inequality constraints' feasibility regions, so this prehibits that two active, countering inequality constraints collaps to yield an equality constraint. ACQ only requires that the tangent information and the linearized information match up. The example was constructed using this information. (2 Points)

At the point in question, both inequality constraints of problem P2 are active. Note that q is continuously differentiable with

$$q'(x_1) = \begin{cases} 2(x_1 + 1), & x_1 < -1, \\ 0, & -1 \le x_1 \le 1, \\ 2(x_1 - 1), & x_1 > 1, \end{cases} \tag{0.1}$$

and we have

$$\begin{aligned} g_1'(x) &= (q'(x_1), -1) \quad &\implies g_1'(x^*) = (0, -1) \\ g_2'(x) &= (q'(x_1), 1) \quad &\implies g_2'(x^*) = (0, 1) \end{aligned}$$

so $g_1'(x^*)d = d_2 = -g_2'(x^*)d$ for any $d \in \mathbb{R}^2$ cannot be simultaneously greater and less than 0, so MFCQ is violated. However, we have that

$$\mathcal{T}_F(x^*) = \mathbb{R} \times \{0\} = \mathcal{T}_F^{\text{lin}}(x^*),$$

so ACQ is satisfied. (2 Points)

(*iii*) Finally, we show that ACQ is violated in problem P1, but GCQ is satisfied. **Note:** ACQ needs the tangent and linearizing information/sets to match up while GCQ only requires that their polar cones coincide, i. e., that they look "the same on the outside" to linear functionals applied to them. This is a property of complementarity constrained problems such as the one used in this example.

For the feasible set $F = \{x \in \mathbb{R}^2 \mid x_1, x_2 \ge 0,\ x_1 x_2 = 0\}$, with $x^* = (0,0) \in F$, we have that

$$\begin{aligned} g_1'(x) &= (1,0) &&\Longrightarrow g_1'(x^*) = (1,0) \\ g_2'(x) &= (0,1) &&\Longrightarrow g_2'(x^*) = (0,1) \\ h'(x) &= (x_2, x_1) &&\Longrightarrow h'(x^*) = (0,0) \end{aligned}$$

we easily obtain the cones

$$\begin{aligned} \mathcal{T}_F(x^*) &= F \\ \mathcal{T}_F^{\text{lin}}(x^*) &= \mathbb{R}^2_{\le} = \operatorname{conv}(F). \end{aligned}$$

Accordingly, the cones don't match because the linearizing cone fills in the entire lower left quadrant. The set's boundaries still coincide though, and we obtain

$$\mathcal{T}_F(x^*)^\circ = \mathbb{R}^2_{\ge} = \mathcal{T}_F^{\text{lin}}(x^*)^\circ,$$

so GCQ is satisfied. (2 Points)

Homework Problem 9.3 (CQs are invariant under Slack Transformation) 10 Points

We can reformulate the original nonlinear problem

$$\left.\begin{aligned} \text{Minimize} \quad & f(x) && \text{where } x \in \mathbb{R}^n \\ \text{subject to} \quad & g_i(x) \le 0 && \text{for } i = 1, \ldots, n_{\text{ineq}} \\ \text{and} \quad & h_j(x) = 0 && \text{for } j = 1, \ldots, n_{\text{eq}} \end{aligned}\right\} \tag{7.1}$$

by introducing a so called **slack variable** $s \in \mathbb{R}^{n_{\text{ineq}}}$ to obtain the simple one-sided box-constrained problem

$$\left.\begin{aligned} \text{Minimize} \quad & f(x) && \text{where } (x, s) \in \mathbb{R}^{n \times n_{\text{ineq}}} \\ \text{subject to} \quad & g_i(x) + s = 0 && \text{for } i = 1, \ldots, n_{\text{ineq}} \\ \text{and} \quad & -s \le 0 && \\ \text{and} \quad & h_j(x) = 0 && \text{for } j = 1, \ldots, n_{\text{eq}} \end{aligned}\right\}. \tag{7.1$_\text{s}$}$$

(i) Derive the KKT-system of (7.1$_s$) and show that there is a one-to-one connection between the solutions of the KKT systems corresponding to (7.1) and (7.1$_s$).

(ii) Show that GCQ/ACQ/MFCQ/LICQ is satisfied at a feasible (x, s) for (7.1$_s$) if the respective condition is satisfied at x for (7.1).

For which CQs can you show equivalence?

Solution.

(i) All quantities corresponding to the slacked system (KKT$_s$) will be denoted with a tilde, e. g., for (7.1$_s$), we set the constraints $\widetilde{g}\colon \mathbb{R}^{n \times n_{\text{ineq}}} \longrightarrow \mathbb{R}^{n_{\text{ineq}}}$ and $\widetilde{h}\colon \mathbb{R}^{n \times n_{\text{ineq}}} \longrightarrow \mathbb{R}^{n_{\text{eq}}+n_{\text{ineq}}}$ as

$$\begin{aligned} \widetilde{g}(x,s) &\coloneqq -s & \text{with} \quad \widetilde{g}'(x,s) &= \begin{bmatrix} 0 & -\mathrm{Id} \end{bmatrix} \\ \widetilde{h}(x,s) &\coloneqq \begin{bmatrix} h(x) \\ g(x)+s \end{bmatrix} & \text{with} \quad \widetilde{h}'(x,s) &= \begin{bmatrix} h'(x) & 0 \\ g'(x) & \mathrm{Id} \end{bmatrix} \end{aligned}$$

and denote the corresponding feasible set by $\widetilde{F} \subseteq \mathbb{R}^{n \times n_{\text{ineq}}}$.

The KKT conditions are

$$\begin{aligned} \nabla f(x) + g'(x)^\intercal \mu + h'(x)^\intercal \lambda &= 0, \\ \mu \ge 0, \quad g(x) \le 0, \quad \mu^\intercal g(x) &= 0 \\ h(x) &= 0, \end{aligned} \tag{KKT}$$

and

$$\begin{aligned} \nabla f(x) + \widetilde{g}'(x,s)^\intercal \widetilde{\mu} + \widetilde{h}'(x,s)^\intercal \widetilde{\lambda} &= 0, \\ \widetilde{\mu} \ge 0, \quad \widetilde{g}(x,s) \le 0, \quad \widetilde{\mu}^\intercal \widetilde{g}(x,s) &= 0 \\ \widetilde{h}(x,s) &= 0, \end{aligned} \tag{KKT$_s$}$$

for multipliers $\widetilde{\mu} \in \mathbb{R}^n_{\text{ineq}}$ and $\widetilde{\lambda} = (\widetilde{\lambda}_x, \widetilde{\lambda}_s) \in \mathbb{R}^{n_{\text{eq}}+n_{\text{ineq}}}$ respectively, where (KKT$_s$) expands to the system

$$\begin{aligned} \begin{bmatrix} \nabla_x f(x) \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ -\widetilde{\mu} \end{bmatrix} + \begin{bmatrix} h'(x)^\intercal \widetilde{\lambda}_x + g'(x) \widetilde{\lambda}_s \\ \widetilde{\lambda}_s \end{bmatrix} &= 0, \\ \widetilde{\mu} \ge 0, \quad -s \le 0, \quad \widetilde{\mu}^\intercal s &= 0 \\ g(x) = -s, \; h(x) &= 0, \end{aligned} \tag{KKT$_s^*$}$$

showing the one to one correspondence of the KKT solutions, since we can always replace $g(x)$ with s and identify the multipliers $\mu = \widetilde{\mu} = \widetilde{\lambda}_s$ and $\widetilde{\lambda}_x = \lambda$. (2 Points)

(ii) First off, note that if we define the map

$$\Phi\colon \mathbb{R}^n \longrightarrow \mathbb{R}^{n+n_{\text{ineq}}}, \quad \Phi(x) := \begin{pmatrix} x \\ -g(x) \end{pmatrix} \quad \text{with} \quad \Phi'(x) = \begin{bmatrix} \text{Id} \\ -g'(x) \end{bmatrix}$$

then clearly $\Phi(F) = \widetilde{F}$ and since Φ is injective, it is invertible on its image, i. e., we can define

$$\Phi^{-1}(x, s) = x$$

as the (right) inverse on the image (this is not a full inverse because Φ is clearly not surjective!). Additionally, we immediately see that $\mathcal{A}(x) = \widetilde{\mathcal{A}(\Phi(x))}$.

(a) For LICQ, we observe that

$$\begin{bmatrix} \widetilde{g}_i{}'(x, s)|_{i \in \widetilde{\mathcal{A}(x,s)}} \\ \widetilde{h}_j{}'(x, s)|_{j=1,\ldots,n_{\text{eq}}} \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} 0 & -\text{Id} \end{bmatrix}_{i \in \widetilde{\mathcal{A}(x,s)}} \\ \begin{bmatrix} h'(x) & 0 \\ g'(x) & \text{Id} \end{bmatrix} \end{bmatrix},$$

which has full row-rank if and only if

$$\begin{bmatrix} g_i'(x)|_{i \in \mathcal{A}(x)} \\ h_j'(x)|_{j=1,\ldots,n_{\text{eq}}} \end{bmatrix}$$

does, so LICQ holds at x for (7.1) if and only if it holds at $(x, s) = \Phi(x)$ for (7.1_{s}).

(b) For MFCQ, we observe that

$$\widetilde{h}'(x, s) = \begin{bmatrix} h'(x) & 0 \\ g'(x) & \text{Id} \end{bmatrix}$$

has full (row) rank if and only if $h'(x)$ does. Additionally, existence of a $d = (d_x, d_s) \in \mathbb{R}^{n+n_{\text{ineq}}}$ such that

$$\begin{aligned} \widetilde{g}_i'(x, s)d = -d_{s,i} \quad & < 0, \; i \in \widetilde{\mathcal{A}(x, s)} \\ \widetilde{h}'(x)d = \begin{bmatrix} h'(x)d_x \\ g'(x)d_x + d_s \end{bmatrix} \quad & = 0 \end{aligned}$$

is equivalent to the existence of a d_x such that

$$\begin{aligned} g_i'(x)d_x &< 0 \; i \in \mathcal{A}(x) \\ h'(x)d_x &= 0, \end{aligned}$$

so MFCQ holds at x for (7.1) if and only if it holds at $(x, s) = \Phi(x)$ for (7.1_{s}).

(c) For ACQ, we can find a fairly compact closed form of the tangent- and linearizing cones for the slacked problem. Starting with the tangent cone $\mathcal{T}_{(x,s)}(\widetilde{F})$, we know that by definition this contains all directions $(d_x, d_s) \in \mathbb{R}^{n+n_{\text{ineq}}}$ such that there exists a positive null sequence $t^{(k)}$ and a sequence $(x^{(k)}, s^{(k)}) \in \widetilde{F} = \Phi(F)$ such that

$$\frac{\Phi(x^{(k)}) - \Phi(x)}{t_k} = \frac{(x^{(k)}, s^{(k)}) - (x, s)}{t^{(k)}} \to (d_x, d_s)$$

so the d_x components are exactly the directions in $\mathcal{T}_F(x)$. For the s and $s^{(k)}$ we know that $s = -g(x)$ and $s^{(k)} = -g(x^{(k)})$, so that, using the mean value theorem for the continuous differentiable g, we obtain that

$$d_s \xleftarrow{k\to\infty} \frac{s^{(k)} - s}{t^{(k)}} = -\frac{g(x^{(k)}) - g(x)}{t^{(k)}} = -\frac{g'(\xi^{(k)})(x^{(k)} - x)}{t^{(k)}} \xrightarrow{k\to\infty} -g'(x)d_x,$$

which shows that the d_s components of $d \in \mathcal{T}_{(x,s)}(\widetilde{F})$ are exactly the corresponding $-g'(x)$-transformed tangent directions $d_x \in \mathcal{T}_F(x)$. Over all, we obtain that $\mathcal{T}_{(x,s)}(\widetilde{F}) = \Phi'(x)\mathcal{T}_F(x)$ for all x in F.

For the linearizing cone we know that for $(x, s) = \Phi(x)$ (by definition)

$$\begin{aligned}
\mathcal{T}^{\text{lin}}_{(x,s)}(\widetilde{F}) &:= \left\{ d = (d_x, d_s) \in \mathbb{R}^{n+n_{\text{ineq}}} \,\middle|\, \begin{array}{ll} \widetilde{g}_i{}'(x)\, d \le 0 & \text{for all } i \in \widetilde{\mathcal{A}(x,s)} \\ \widetilde{h}_j{}'(x)\, d = 0 & \text{for all } j = 1, \dots, n_{\text{eq}} + n_{\text{ineq}} \end{array} \right\} \\
&= \{(d_x, d_s) \in \mathbb{R}^{n+n_{\text{ineq}}} \mid d_{s,i} \ge 0,\ i \in \widetilde{\mathcal{A}(x,s)},\ g'(x)d_x + d_s = 0,\ h'(x)d_x = 0\} \\
&= \{(d_x, -g'(x)d_x) \mid \text{for } d \in \mathbb{R}^n \text{ with } g'(x)d_x \le 0,\ i \in \mathcal{A}(x),\ h'(x)d_x = 0\} \\
&= \Phi'(x)\mathcal{T}^{\text{lin}}_F(x).
\end{aligned}$$

So for $(x, s) = \Phi(x)$ we have the transformations

$$\begin{aligned}
\mathcal{T}_{(x,s)}(\widetilde{F}) &= \Phi'(x)\mathcal{T}_F(x) \\
\mathcal{T}^{\text{lin}}_{(x,s)}(\widetilde{F}) &= \Phi'(x)\mathcal{T}^{\text{lin}}_F(x).
\end{aligned} \tag{0.2}$$

This shows that if $\mathcal{T}_F(x) = \mathcal{T}^{\text{lin}}_F(x)$, i. e., ACQ is satisfied at x for (7.1), then ACQ holds at (x, s) for (7.1$_s$).

In fact, because

$$\Phi'(x) = \begin{bmatrix} \text{Id} \\ -g'(x) \end{bmatrix},$$

we know that $\Phi'(x)$ has the left inverse

$$\Phi^{-L} = \begin{bmatrix} \mathrm{Id} & 0 \end{bmatrix}$$

(the projection/restriction to the first n components) and therefore (0.2) implies

$$\begin{aligned} \Phi'(x)^{-L}\mathcal{T}_{(x,s)}(\widetilde{F}) &= \mathcal{T}_F(x) \\ \Phi'(x)^{-L}\mathcal{T}^{\text{lin}}_{(x,s)}(\widetilde{F}) &= \mathcal{T}^{\text{lin}}_F(x), \end{aligned} \tag{0.3}$$

which shows the reverse implication, so if ACQ is satisfied at (x, s) for (7.1_s), then ACQ is also satisfied at x for (7.1).

(d) For GCQ, we can reuse the transformation property (0.2) that we derived for the ACQ investigation. For any x with corresponding $(x, s) = \Phi(x)$ we have that

$$\begin{aligned} (\mathcal{T}_{(x,s)}(\widetilde{F}))^\circ &= (\Phi'(x)\mathcal{T}_F(x))^\circ \\ &= \{f = (f_x, f_s) \in \mathbb{R}^{n+n_\text{ineq}} \mid f^\mathsf{T}\Phi'(x)d \le 0 \text{ for all } d \in \mathcal{T}_F(x)\} \\ &= \{f = (f_x, f_s) \in \mathbb{R}^{n+n_\text{ineq}} \mid \left(\Phi'(x)^\mathsf{T}f\right)^\mathsf{T} d \le 0 \text{ for all } d \in \mathcal{T}_F(x)\} \\ &= \{f = (f_x, f_s) \in \mathbb{R}^{n+n_\text{ineq}} \mid \Phi'(x)^\mathsf{T}f \in \mathcal{T}_F(x)^\circ\} \end{aligned}$$

and analogously

$$(\mathcal{T}^{\text{lin}}_{(x,s)}(\widetilde{F}))^\circ = \{f = (f_x, f_s) \in \mathbb{R}^{n+n_\text{ineq}} \mid \Phi'(x)^\mathsf{T}f \in \mathcal{T}^{\text{lin}}_F(x)^\circ\}.$$

Therefore, when $\mathcal{T}_F(x)^\circ = \mathcal{T}^{\text{lin}}_F(x)^\circ$ (GCQ is satisfied at x for (7.1)), then of course $\mathcal{T}_{(x,s)}(\widetilde{F})^\circ = \mathcal{T}^{\text{lin}}_{(x,s)}(\widetilde{F})^\circ$ (GCQ holds at $\Phi(x)$ for (7.1_s)).

The reverse implication I don't expect to hold but I have yet to construct an example showing that.

(8 Points)

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

(i) Use Farkas' Lemma (Lemma 7.11 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 $p \le n_\text{ineq}$ either the system

$$Ad < 0, \quad Bd = 0 \tag{0.4}$$

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

$$A^\top \mu + B^\top \lambda = 0 \tag{0.5}$$

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

Hint: Start with the existence of a nontrivial solution to (0.5). 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 (7.1). 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 $\{d \in \mathbb{R}^n \,|\, \widetilde{B}\, d \geq 0\}$ satisfy $\widetilde{c}^\top d \geq 0$

(c) There is no element of $\{d \in \mathbb{R}^n \,|\, \widetilde{B}\, d \geq 0\}$ 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 (0.5) 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

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

then the system (0.4) simply states that there is no MFCQ direction and (0.5) 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 $(d\mu_{\mathcal{(x^*)}}, d\lambda)$ 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.