Nonlinear Optimization – Sheet 09

Exercise 1

We have

$$g_1(x) = x_1 + 4x_2 - 3; \qquad \nabla g_1(x) = \begin{pmatrix} 1 \\ 4 \end{pmatrix}$$

and

$$g_2(x) = x_2 - x_1; \qquad \nabla g_2(x) = \begin{pmatrix} -1 \\ 1 \end{pmatrix}.$$

As $\nabla g_1(x)$ and $\nabla g_2(x)$ are linearly independent for every x, the LICQ is satisfied everywhere, regardless of how many of the constraints are active. Therefore, the KKT-conditions for feasible x,

$$\nabla f(x) + g'(x)^\top \mu = 0 \qquad (I)$$

$$\begin{pmatrix} 2(x_1 - 2) \\ 4(x_2 - 1) \end{pmatrix} + \begin{pmatrix} 1 & -1 \\ 4 & 1 \end{pmatrix} \cdot \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} = 0 \qquad (I)$$

$$\mu^\top g(x) = 0 \qquad (II),$$

are necessary optimality conditions. The second condition can be extended to yield two conditions (due to the complementarity condition),

$$\mu_1(x_1 + 4x_2 - 3) = 0$$
$$\mu_2(x_2 - x_1) = 0.$$

We make a case distinction

$\mu_1 = 0 \wedge \mu_2 = 0$ In this case we find that $x = (2, -1)$ with $\mu = (0, 0)$ is the unique point satisfying the KKT-conditions.

$\mu_1 = 0 \wedge \mu_2 > 0$ In this case we find that the only solution of the system of equations is $x = (4/3, 4/3)$ with $\mu = (0, -4/3)$. However, μ_2 is not positive. Therefore, there are no KKT-points in this case.

$\mu_1 > 0 \wedge \mu_2 = 0$ In this case we find that $x = (5/3, 1/3)$ with $\mu = (2/3, 0)$ is the unique point satisfying the KKT-conditions.

$\mu_1 > 0 \wedge \mu_2 > 0$ In this case we find $x = (3/5, 3/5)$ with $\mu = (22/25, -48/25)$. However, μ_2 is not positive. Therefore, there are no KKT-points in this case.

Under the LICQ, KKT-conditions are necessary. Therefore, the only two candidates for local minima are $x = (2, -1)$ with $\mu = (0, 0)$ and $x = (5/3, 1/3)$ with $\mu = (2/3, 0)$.

For $x^* = (2, -1)$ with $\mu^* = (0, 0)$, we compute

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

At x^*, there are no active inequality constraints. Therefore, the critical cone is $\mathbb{R}^2$, in particular $d = (1, 0)$ is contained in the critical cone and we obtain

$$d^\top \cdot \begin{pmatrix} 2 & 0 \\ 0 & 4 \end{pmatrix} \cdot d \geq 2 \cdot \|d\|^2 \geq 0 \;\forall d \in \mathbb{R}^2.$$

As a result, the necessary second order optimality conditions are satisfied.

With Theorem 9.5 we can even conclude that x ist a strict local maximizer with $f(x) = f(2, -1) = -8$.

For $x^* = (5/3, 1/3)$ with $\mu^* = (2/3, 0)$, we compute

$$\mathcal{T}_{\text{NLP}}^{\text{critical}}(x) = \{d \in \mathbb{R}^2 |\; d_1 + 4d_2 = 0,\; d_1 - d_2 \leq 0\} = \mathbb{R}_{\geq 0} \cdot (-4, 1).$$

Therefore,

$$d^\top \cdot \mathcal{L}(x^*, \mu^*) \cdot d = \begin{pmatrix} 2 & 0 \\ 0 & 4 \end{pmatrix} \geq 2 \cdot \|d\|^2 \geq 0 \;\forall d \in \mathbb{R}^2.$$

In particular, the conditions for Theorem 9.5 are satisfied and $(5/3, 1/3)$ is a local maximizer with $f(5/3, 1/3) = -1$. Therefore, it is a global maximizer.

Exercise 2

1. We need to show that for (P1) there exists a point that satisfies GCQ but not ACQ. The feasible set is
$$F = \mathbb{R}_{\leq 0} \times 0 \cup 0 \times \mathbb{R}_{\leq 0}.$$
It is clear from the geometry of the problem that
$$\mathcal{T}_F(0) = F.$$
Furthermore,
$$\mathcal{T}_F^{\text{lin}}(0) = \{d \in \mathbb{R}^2 | d_1 \leq 0,\ d_2 \leq 0\}.$$
Therefore, the ACQ is not satisfied. However,
$$\mathcal{T}_F(0)^\circ = \{s \in \mathbb{R}^2 : s^\top x \leq 0\ \forall x \in \mathcal{T}_F(0) = F\} = \{s_1 \geq 0, s_2 \geq 0\}.$$
and
$$\mathcal{T}_F^{\text{lin}}(0)^\circ = \{s \in \mathbb{R}^2 : s^\top x \leq 0\ \forall x \in \mathcal{T}_F^{\text{lin}}(0)\} = \{s_1 \geq 0, s_2 \geq 0\},$$
i.e. the GCQ is satisfied at 0.

2. We need to show that for (P2) there exists a point that satisfies ACQ but not MFCQ. By adding the two constraints, we obtain $q(x_1) \leq 0$. That is equivalent to
$$-1 \leq x_1 \leq 1.$$
On this interval, $q(x_1) = 0$. Therefore, the constraints reduce to $-x_2 \leq 0$ and $x_2 \leq 0 \implies x_2 = 0$. Consequently, the feasible set is
$$F = [-1, 1] \times \{0\}.$$
For $x = (-1, 0)$ we have
$$\mathcal{T}_F(x) = \mathbb{R}_{\geq 0} \cdot (1, 0) \cup \mathbb{R}_{\geq 0} \cdot (-1, 0) = \mathbb{R} \times \{0\}.$$
We have
$$g_1'(x) = \begin{pmatrix} 0 \\ -1 \end{pmatrix}, \qquad g_2'(x) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$
As all inequality constraints are active at x, we obtain
$$\mathcal{T}_F^{\text{lin}}(x) = \{d \in \mathbb{R}^2 |\ d_2 \geq 0;\ d_2 \leq 0\} = \mathbb{R} \times \{0\}.$$
Therefore, the ACQ is satisfied here. If there is a vector d such that
$$g_1(x) \cdot d < 0 \Longleftrightarrow d_2 > 0$$
, then
$$g_2(x) \cdot d = d_2 > 0.$$
Therefore, the MFCQ is not satisfied at x.

3. We need to show that for P3 there exists a point that satisfies MFCQ but not LICQ. We compute
$$\nabla g_1(x) = \begin{pmatrix} -3x_1^2 \\ -1 \end{pmatrix}; \qquad \nabla g_2(x) = \begin{pmatrix} 0 \\ -1 \end{pmatrix}$$
Consider $x = (0, 1)$. x doesn't satisfy the LICQ as $\nabla g_1(x) = \nabla g_2(x)$. However, none of the inequality conditions is active. Therefore, MFCQ is trivially satisfied.

Exercise 3

We interpret (7.1s) to have $(x, s) \in \mathbb{R}^{n+n_{\text{ineq}}}$ and $g_i(x) + s_i = 0$ instead of $(x, s) \in \mathbb{R}^{n\times n_{\text{ineq}}}$ and $g_i(x) + s = 0$.

(i) The Lagrange-function $\mathcal{L}$ for (7.1) is given by

$$\mathcal{L}(x, \mu, \lambda) = f(x) + \sum_{i=0}^{n_{\text{eq}}} \mu_i g_i + \sum_{j=0}^{n_{\text{ineq}}} \lambda_j h_j(x)\,,$$

whereas the Lagrange-function $\tilde{\mathcal{L}}$ for (7.1s) is given by

$$\tilde{\mathcal{L}}(x, s, \mu, \lambda, \tilde{\lambda}) = f(x) - \sum_{i=0}^{n_{\text{eq}}} \mu_i s_i + \sum_{i=0}^{n_{\text{eq}}} \tilde{\lambda}_i (g_i(x) + s_i) + \sum_{j=0}^{n_{\text{ineq}}} \lambda_j h_j(x)\,.$$

We have

$$\begin{aligned}\nabla_{(x,s)} \tilde{\mathcal{L}}(x, s, \mu, \lambda, \tilde{\lambda}) &= \begin{bmatrix} \nabla_x f(x) \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ -\mu \end{bmatrix} + \begin{bmatrix} g'(x)^t \\ \mathrm{Id} \end{bmatrix} \tilde{\lambda} + \begin{bmatrix} h'(x)^t \lambda \\ 0 \end{bmatrix} \\ &= \begin{bmatrix} \nabla_x f(x) + g'(x)^t \tilde{\lambda} + h'(x)^t \lambda \\ \tilde{\lambda} - \mu \end{bmatrix}\end{aligned}$$

The KKT conditions for (7.1s) impose

$$\nabla_{(x,s)} \tilde{\mathcal{L}}(x, s, \mu, \lambda, \tilde{\lambda}) = 0 \tag{1}$$
$$g(x) + s = 0 \tag{2}$$
$$h(x) = 0 \tag{3}$$
$$\mu \geq 0\,, \quad -s \leq 0\,, \quad \mu^t s = 0\,. \tag{4}$$

Clearly, eq. (1) is equivalent to $\nabla_x \mathcal{L}(x, \mu, \lambda) = 0$. The condition pair eq. (2) and eq. (4) is equivalent to

$$\mu \geq 0\,, \quad g(x) \leq 0\,, \quad \mu^t g(x) = 0\,. \tag{5}$$

The implication eq. (2),eq. (4) $\implies$ eq. (5) is clear from multiplying eq. (2) with μ^t. The implication eq. (5) $\implies$ eq. (2),eq. (4) follows by setting $s = -g(x)$. Thus, the KKT systems are in one-to-one correspondence.

(ii) LICQ for (7.1s) imposes that the collection of vectors

$$\begin{aligned}\nabla_{(x,s)} h_j(x) &= \begin{bmatrix} \nabla_x h_j(x) \\ 0 \end{bmatrix} \\ \nabla_{(x,s)} (g_i(x) + s_i) &= \begin{bmatrix} \nabla_x g_i(x) \\ e_i \end{bmatrix} \\ \nabla_{(x,s)} - s_i &= \begin{bmatrix} 0 \\ -e_i \end{bmatrix}\end{aligned}$$

are linearly independent. Assume that LICQ is violated for (7.1), ie that there is a linear dependence

$$\sum_j a_j \nabla_x h_j(x) + \sum_i b_i \nabla_x g_i(x) = 0$$

for some $a_j, b_i \in \mathbb{R}$ not all zero. Then we get a linear dependence

$$\sum_j a_j \begin{bmatrix} \nabla_x h_j(x) \\ 0 \end{bmatrix} + \sum_i b_i \begin{bmatrix} \nabla_x g_i(x) \\ e_i \end{bmatrix} + \sum_i b_i \begin{bmatrix} 0 \\ -e_i \end{bmatrix} = 0\,.$$

Therefore LICQ for (7.1s) implies LICQ for (7.1). With similar explicit calculation, one can see that the conditions are equivalent.

MFCQ for (7.1s) imposes that the collection of vectors

$$\nabla_{(x,s)} h_j(x) = \begin{bmatrix} \nabla_x h_j(x) \\ 0 \end{bmatrix}$$
$$\nabla_{(x,s)} (g_i(x) + s_i) = \begin{bmatrix} \nabla_x g_i(x) \\ -e_i \end{bmatrix}$$

are linearly independent and that there is a vector $d \in \mathbb{R}^{n+n_{\text{ineq}}}$ such that

$$\begin{aligned} \begin{bmatrix} 0 \\ -e_i \end{bmatrix}^t d &< 0 && \forall i \in \mathscr{A}(x)\,, \\ \begin{bmatrix} \nabla_x h_j(x) \\ 0 \end{bmatrix}^t d &= 0 && \forall j\,, \\ \begin{bmatrix} \nabla_x g_i(x) \\ e_i \end{bmatrix}^t d &= 0 && \forall i\,. \end{aligned}$$

Assume that MFCQ is violated for (7.1), ie that there is a linear dependence

$$\sum_j a_j \nabla_x h_j(x) = 0$$

for some $a_j \in \mathbb{R}$ not all zero or that there is no MFCQ vector $d_0 \in \mathbb{R}^n$. The linear dependence transfers to (7.1s) as above, so assume that there is no MFCQ vector. Quite clearly, any MFCQ vector for (7.1s) provides one for (7.1) just by chopping the second half off, hence there cannot ba any MFCQ vector for (7.1s) Therefore MFCQ for (7.1s) implies MFCQ for (7.1).

For ACQ and GCQ, we observe that the feasible sets F for (7.1) and (7.1s) coincide (after projection to $\mathbb{R}^n$), and so do their tangent cones.

Exercise 4

i): We write q instead of n_{eq}. Let $A \in \mathbb{R}^{p\times n}, B \in \mathbb{R}^{q\times n}$ with $rk(B) = q$ and $p \leq n_{ineq}$. We set $C := \begin{bmatrix} -B^T & B^T \end{bmatrix} \in \mathbb{R}^{n\times 2q}$ and apply the Farkas Lemma to the pair (C^T, b), where $b = A^T\mu$ for fixed $\mu \geq 0$ and $\mu \neq 0$. This gives us, that either

1: There exists an $x = \begin{pmatrix} \lambda^+ \\ \lambda^- \end{pmatrix} \in \mathbb{R}^{2q}$ s.t. $Cx = b$ and $x \geq 0$
or
2: There exists a $d \in \mathbb{R}^n$ s.t. $d^TC \geq 0$ and $d^Tb < 0$.

Writing $\lambda = \lambda^+ - \lambda^-$ in the first case, we see that $Cx = b \iff -B^\top\lambda^+ + B^\top\lambda^- = A^T\mu$, so $A^T\mu + B^T\lambda = 0$.

In the second case, we have $d^T \begin{bmatrix} -B^T & B^T \end{bmatrix} = d^TC \geq 0$, so $-d^TB^T \geq 0$ and $d^TB^T \geq 0$, which implies $Bd = 0$. Also $\mu Ad = (Ad)^T\mu = d^TA^T\mu = d^Tb < 0$.

This concludes the proof.

ii): Let (x^*, λ^*, μ^*) be a KKT point for (7.1)

Then, using the obvious matrices in i), we get:
x^* satisfies MFCQ $\iff$ (0.1) has a solution $\iff$ every nontrivial solution (μ, λ) for (0.2) satisfies μ has component < 0. Now let (μ, λ) be different Lagrange multipliers, that solve the KKT-System at x^*. Then $(\mu - \mu^*, \lambda - \lambda^*)$ is a solution to (0.2), which implies that $\mu - \mu^*$ has a component < 0. This shows, that since $\mu \geq 0$, the components are bounded by the maximal component of μ^*. Therefore the set of λ s.t. $B^T\lambda = -A^T\mu - \nabla f(x^*)$ is also bounded. Obviously, the set of Lagrange multipliers is closed, so it is compact.