Prof. Dr. Karsten Urban
M. Sc. Lewin Ernst

Winter term 2024/2025
Sheet 06, 32 Points

Numerical Optimization
Solution to exercise sheet

review on 27.11.2024 during the exercise class

1. *(Fun with cones and tangential cones)*

   a) Let $\mathcal{S}$[1] $\subset \mathbb{R}^n$ be a set and let $x \in \mathcal{S}$ be given. Show that the tangential cone $\mathcal{T}(\mathcal{S}, x)$ is actually a cone and is a closed set.

   b) The conical hull of a set $\mathcal{S} \subset \mathbb{R}^n$ at $x \in \mathcal{S}$ is given by $\operatorname{cone}(\mathcal{S}-x) := \{\alpha(y-x) : y \in \mathcal{S}, \alpha > 0\}$. Let $\mathcal{S} \subset \mathbb{R}^n$ be convex and let $x \in \mathcal{S}$ be given. Show that the tangential cone $\mathcal{T}(\mathcal{S}, x)$ is the closure of the conical hull of $\mathcal{S}$ in x, i.e. $\mathcal{T}(\mathcal{S}, x) = \overline{\operatorname{cone}(\mathcal{S} - x)}$.

(6 + 6 = 12 Points)

Solution:

a) First we show that $\mathcal{T}(\mathcal{S}, x)$ is a cone. Let $d \in \mathcal{T}(\mathcal{S}, x)$ be given. Therefore there is $(\eta^{(\ell)})_{\ell\in\mathbb{N}}$ and $(x^{(\ell)})_{\ell\in\mathbb{N}} \subset \mathcal{S}$ such that $\lim\limits_{\ell\to\infty} x^{(\ell)} = x$ and $\lim\limits_{\ell\to\infty} \eta^{(\ell)}(x^{(\ell)}-x) = d$. Now let $\alpha > 0$ be given. Then we take the sequence $(\alpha\eta^{(\ell)})_{\ell\in\mathbb{N}}$ for which $\lim\limits_{\ell\to\infty} \alpha\eta^{(\ell)}(x^{(\ell)}-x) = \lim\limits_{\ell\to\infty} \alpha \lim\limits_{\ell\to\infty} \eta^{(\ell)}(x^{(\ell)}-x) = \alpha d$ holds.

We shall now show that the set $\mathcal{T}(\mathcal{S}, x)$ is closed in $\mathbb{R}^n$ (the norm does not matter). We show that the limit points of all convergence sequences are contained in the set, i.e. for all $(v^{(k)})_{k\in\mathbb{N}} \subset \mathcal{T}(\mathcal{S}, x)$ with $\lim\limits_{k\to\infty} v^{(k)} = v$ it holds $v \in \mathcal{T}(\mathcal{S}, x)$. Due to $v^{(k)} \in \mathcal{T}(\mathcal{S}, x)$ we know there are sequences $(\eta_k^{(\ell)})_{\ell\in\mathbb{N}} \subset \mathbb{R}^+$ and $(x_k^{(\ell)})_{\ell\in\mathbb{N}} \subset \mathcal{S}$ with $\lim\limits_{\ell\to\infty} x_k^{(\ell)} = x$ and $\lim\limits_{\ell\to\infty} \eta_k^{(\ell)}(x_k^{(\ell)} - x) = v^{(k)}$. By definition of limits there are for all $\epsilon_k := \frac{1}{k}$, $k \in \mathbb{N}$ a $N(\epsilon_k) \in \mathbb{N}$ such that

$$\|x - x_k^{(\ell)}\| < \epsilon_k \quad \text{and} \quad \|v^{(k)} - \eta_k^{(\ell)}(x_k^{(\ell)} - x)\| < \epsilon_k, \quad \forall \ell \geq N(\epsilon_k) =: N_k$$

For $v \in \mathbb{R}^n$ we define the sequences as

$$(\eta^{(k)})_{k\in\mathbb{N}} := (\eta_k^{(N_k)})_{k\in\mathbb{N}} \quad \text{and} \quad (x^{(k)})_{k\in\mathbb{N}} := (x_k^{(N_k)})_{k\in\mathbb{N}}.$$

Then it is clear that $\lim\limits_{k\to\infty} x^{(k)} = x$, because

$$0 \leq \|x - x^{(k)}\| = \|x - x_k^{(N_k)}\| < \frac{1}{k} \overset{k\to\infty}{\longrightarrow} 0.$$

So by the squeeze theorem we have convergence. The same trick can be used to show $\lim\limits_{k\to\infty} \eta^{(k)}(x^{(k)} - x) = v$. In fact, we get

$$\begin{aligned}\lim_{k\to\infty} \|v - \eta^{(k)}(x^{(k)} - x)\| &= \lim_{k\to\infty} \|v - \eta_k^{(N_k)}(x_k^{(N_k)} - x)\| \\ &= \lim_{k\to\infty} \|v - v_k + v_k - \eta_k^{(N_k)}(x_k^{(N_k)} - x)\| \\ &\leq \lim_{k\to\infty} \|v - v_k\| + \lim_{k\to\infty} \frac{1}{k} = 0\end{aligned}$$

[1]We use this notation to emphasize that it does not have to be a cone.

b) We first show $\mathcal{T}(\mathcal{S},x) \subseteq \overline{\text{cone}(\mathcal{S}-x)}$:
Let $d \in \mathcal{T}(\mathcal{S},x)$ be given, then there is $(\eta^{(\ell)})_{\ell\in\mathbb{N}}$ and $(x^{(\ell)})_{\ell\in\mathbb{N}} \subset \mathcal{S}$ such that $\lim\limits_{\ell\to\infty} x^{(\ell)} = x$ and $\lim\limits_{\ell\to\infty} \eta^{(\ell)}(x^{(\ell)} - x) = d$. Due to the definition of the conical hull $\text{cone}(\mathcal{S}-x)$ for each $k \in \mathbb{N}$ it holds $\eta^{(\ell)}(x^{(\ell)} - x) \in \text{cone}(\mathcal{S}-x)$ and therefore $\left(\eta^{(\ell)}(x^{(\ell)} - x)\right)_{k\in\mathbb{N}} \subset \text{cone}(\mathcal{S}-x)$. Due to closedness of $\overline{\text{cone}(\mathcal{S}-x)}$ we have $d \in \overline{\text{cone}(\mathcal{S}-x)}$.
Now we show the other inclusion $\overline{\text{cone}(\mathcal{S}-x)} \subseteq \mathcal{T}(\mathcal{S},x)$:
Let $d \in \text{cone}(\mathcal{S}-x)$ be given, therefore we have the existence of $\alpha > 0$ and $y \in \mathcal{S}$ such that $d = \alpha(y-x)$. We define the sequence $(x^{(\ell)})_{\ell\in\mathbb{N}} \subset \mathcal{S}$ by

$$x^{(\ell)} := x + \frac{\alpha}{\ell+\alpha}(y-x) = \left(1 - \frac{\alpha}{\ell+\alpha}\right)x + \frac{\alpha}{\ell+\alpha}y$$

which is the convex combination of $x \in \mathcal{S}$ and $y \in \mathcal{S}$ for $\lambda = \frac{\alpha}{\ell+\alpha} \in (0,1)$, $\ell \in \mathbb{N}$ converging to x for $\ell \to \infty$. Further, we define $(\eta^{(\ell)})_{\ell\in\mathbb{N}} \subset \mathbb{R}^+$ by $\eta^{(\ell)} := (\ell+\alpha)$. Then we have

$$\lim_{\ell\to\infty} \eta^{(\ell)}(x^{(\ell)} - x) = \lim_{\ell\to\infty} (\ell+\alpha)\left(x + \frac{\alpha}{\ell+\alpha}(y-x) - x\right) = d.$$

Therefore we have $d \in \mathcal{T}(\mathcal{S},x)$ and due to the closedness of $\mathcal{T}(\mathcal{S},x)$ it follows that $\overline{\text{cone}(\mathcal{S}-x)} \subseteq \mathcal{T}(\mathcal{S},x)$.

2. *(More cone fun: tangential cone versus linearized tangential cone)*
Consider the following cosntrained optimization problems

$$(I)\begin{cases} \min\limits_{x\in\mathbb{R}^2} & f(x) \\ \text{s.t.} & g_1(x) \le 0, \\ & g_2(x) \le 0, \end{cases} \qquad (II)\begin{cases} \min\limits_{x\in\mathbb{R}^2} & f(x) \\ & g_1(x) \le 0, \\ \text{s.t.} & g_2(x) \le 0, \\ & g_3(x) \le 0, \end{cases}$$

with

$$f(x) := (x_1-2)^2 + x_2^2, \qquad g_1(x) := (x_1-1)^3 + x_2, \qquad g_2(x) := -x_2, \qquad g_3(x) := x_1 - 1.$$

Solve the following task for (I) and (II).

(a) Determine the unique minimum $x^* \in \mathbb{R}^2$. Note: A proof is not necessary.
(b) Draw the feasible region $\mathcal{F}$.
(c) Determine the tangential cone $\mathcal{T}(\mathcal{F},x^*)$ and the linearized tangential cone $\mathcal{T}_{\text{lin}}(\mathcal{F},x^*)$ (it is enough to derive $\mathcal{T}(\mathcal{F},x^*)$ for one of the problems and the other can be given without explanations).
(d) Draw the tangential cone $\mathcal{T}(\mathcal{F},x^*)$ and the linearized tangential cone $\mathcal{T}_{\text{lin}}(\mathcal{F},x^*)$.
(e) Check the necessary conditions for a minimum of Theorem 3.1.8, i.e.

$$x^* \in \mathcal{F} \text{ is a local solution} \qquad \Rightarrow \qquad \nabla f(x^*)^T d \ge 0 \qquad \forall d \in \mathcal{T}(\mathcal{F},x^*).$$

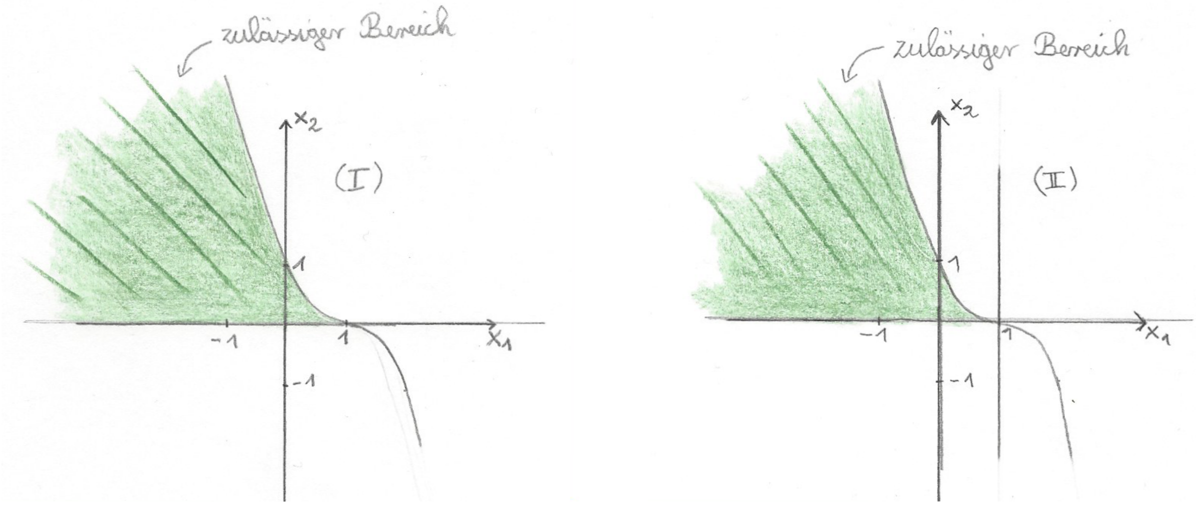
What do you observe if you replace $\mathcal{T}(\mathcal{F},x^*)$ with $\mathcal{T}_{\text{lin}}(\mathcal{F},x^*)$? Is this condition still necessary? How is this related to Abadie Constraint Qualification in Definition 3.1.11?

(2 + 2 + 6 + 2 + 2 = 14 Points)

Solution:

a) The function f is a elliptical paraboloid with origin $(2,0)$. Due to symmetry and monotonicity of f we have that the point $x \in S$ closest to $(2,0)$ is the minimum. With b) it is clear that $x^* = (1,0)^T$ is the minimum for (I) and (II).

b) Drawing the feasible sets $\mathcal{F}_I$ and $\mathcal{F}_{II}$ for problem (I) and (II) gives:



Obviously we have $\mathcal{F} := \mathcal{F}_I = \mathcal{F}_{II}$, i.e. the inequality $g_3(x) \le 0$ does not affect the position of the minimum at all.

c) For the linearized tangential cone we need some function evaluations for $x^* = (1,0)$:

$$g_1(x^*) = g_2(x^*) = g_3(x^*) = 0, \qquad \nabla g_1(x^*) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad \nabla g_2(x^*) = \begin{pmatrix} 0 \\ -1 \end{pmatrix}, \quad \nabla g_3(x^*) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

- Problem (I): The constraints for this problem are

$$g(x) := \begin{pmatrix} g_1(x) \\ g_2(x) \end{pmatrix}, \quad g'(x) := \begin{pmatrix} \nabla g_1(x)^T \\ \nabla g_2(x)^T \end{pmatrix} \qquad \text{d.h.} \qquad g(x^*) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad g'(x^*) = \begin{pmatrix} 0 & 1 \\ 0 & -1 \end{pmatrix}.$$

For the *linearized tangential cone* we get

$$\mathcal{T}_{\text{lin}}(\mathcal{F}, x^*) = \{d \in \mathbb{R}^2 \,:\, g'(x^*)d \le 0\} = \{d \in \mathbb{R}^2 \,:\, d_2 = 0\}.$$

For the tangential cone $\mathcal{T}(\mathcal{F}, x^*)$ we make the guess

$$\mathcal{T}(\mathcal{F}, x^*) = \{x \in \mathbb{R}^2 : x_1 \le 0,\, x_2 = 0\}.$$

We shall prove this now. First, let $d \in \mathcal{T}(\mathcal{F}, x^*)$, therefore we have sequences $(\eta^{(\ell)})_{\ell\in\mathbb{N}} \subset \mathbb{R}^+$ and $(x^{(\ell)})_{\ell\in\mathbb{N}} \subset \mathcal{F}$ with $\lim\limits_{\ell\to\infty} x^{(\ell)} = x^*$ and $\lim\limits_{\ell\to\infty} \eta^{(\ell)}(x^{(\ell)} - x^*) = d$. Now, due to the constraints we have for all $x \in \mathcal{F}$

$$x_2 \ge 0 \quad \Rightarrow \quad x_2 - x_2^* \ge 0, \tag{1}$$

and

$$(x_1 - 1)^3 \le -x_2 \le 0 \quad \Rightarrow \quad x_1 \le 1 \quad \Rightarrow \quad x_1 - x_1^* \le 0. \tag{2}$$

The latter inequality gives us for $(x^{(\ell)})_{\ell\in\mathbb{N}} \subset \mathcal{F}$

$$\eta^{(\ell)}(x_1^{(\ell)} - x_1^*) \le 0,\ \forall \ell \in \mathbb{N} \quad \Rightarrow \quad d_1 \le 0$$

and from (1) and (2) it follows

$$\begin{aligned}
0 &\overset{(1)}{\geq} -\eta^{(\ell)}(x_2^{(\ell)} - x_2^*) \overset{x_2^*=0}{=} -\eta^{(\ell)} x_2^{(\ell)} \overset{(2)}{\geq} \eta^{(\ell)}(x_1^{(\ell)} - 1)^3 \\
&= \underbrace{\eta^{(\ell)}(x_1^{(\ell)} - 1)}_{\overset{\ell\to\infty}{\longrightarrow} d_1} \underbrace{(x_1^{(\ell)} - 1)^2}_{\overset{\ell\to\infty}{\longrightarrow} 0} \overset{\ell\to\infty}{\longrightarrow} 0.
\end{aligned}$$

Hence $d_2 = 0$ and so $d \in \{x \in \mathbb{R}^2 : x_1 \leq 0,\, x_2 = 0\}$. Now, we shall prove the other inclusion, so let $d \in \{x \in \mathbb{R}^2 : x_1 \leq 0,\, x_2 = 0\}$ be given. We define the sequences $(\eta^{(\ell)})_{\ell\in\mathbb{N}} \subset \mathbb{R}^+$ by $\eta^{(\ell)} := \ell$ and $(x^{(\ell)})_{\ell\in\mathbb{N}} \subset \mathcal{F}$ by $x^{(\ell)} := (d_1/\ell + 1, 0)^T \in \mathcal{F}$. Hence we have

$$\lim_{\ell\to\infty} x^{(\ell)} = \lim_{\ell\to\infty} (d_1/\ell + 1, 0)^T = (1,0)^T = x^*$$

and

$$\lim_{\ell\to\infty} \eta^{(\ell)}(x^{(\ell)} - x^*) = \lim_{\ell\to\infty} \ell((1/\ell d_1 + 1, 0)^T - (1,0)^T) = (d_1, 0)^T$$

- Problem (II): The constraint evaluation become

$$g(x) := \begin{pmatrix} g_1(x) \\ g_2(x) \\ g_3(x) \end{pmatrix}, \quad g'(x) := \begin{pmatrix} \nabla g_1(x)^T \\ \nabla g_2(x)^T \\ \nabla g_3(x)^T \end{pmatrix} \qquad \text{d.h.} \qquad g(x^*) = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \quad g'(x^*) = \begin{pmatrix} 0 & 1 \\ 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

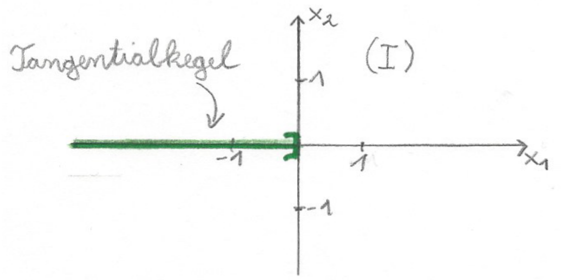
With that we get for the *linearized tangential cone*

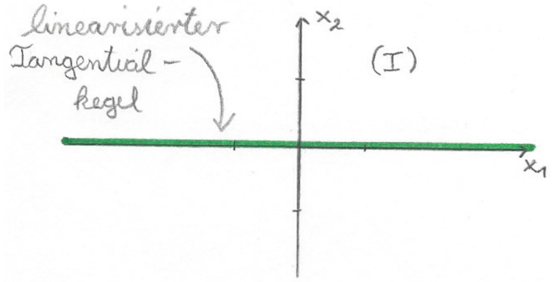
$$\begin{aligned}
\mathcal{T}(\mathcal{F}, x^*) &= \{d \in \mathbb{R}^2 \,:\, \begin{pmatrix} 0 & 1 \\ 0 & -1 \\ 1 & 0 \end{pmatrix} d \leq 0\} \\
&= \{d \in \mathbb{R}^2 \,:\, d_1 \leq 0,\, d_2 = 0\}.
\end{aligned}$$

The *tangential cone* is the same as before, i.e.

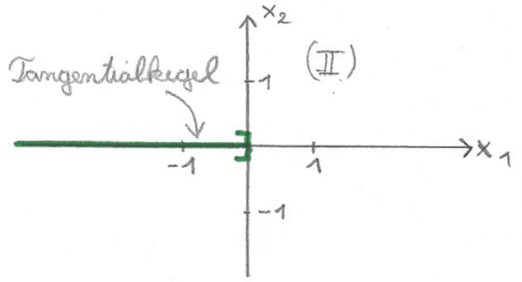
$$\mathcal{T}(\mathcal{F}, x^*) = \{x \in \mathbb{R}^2 : x_1 \leq 0,\, x_2 = 0\}.$$

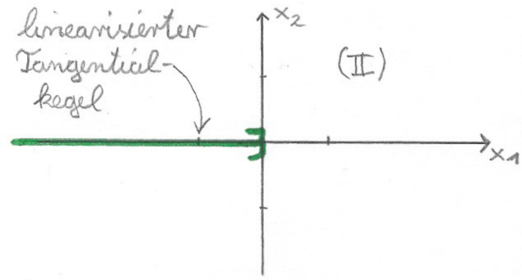
d) Drawing the tangential cone $\mathcal{T}(\mathcal{F}, x^*)$ and the linearized tangential cone $\mathcal{T}_{\text{lin}}(\mathcal{F}, x^*)$ for problem (I) and (II) results in:





$\mathcal{T}(\mathcal{F}, x^*) = \{d \in \mathbb{R}^2 \,:\, d_1 \leq 0,\, d_2 = 0\}$ $\qquad$ $\mathcal{T}_{\text{lin}}(\mathcal{F}, x^*) = \{d \in \mathbb{R}^2 \,:\, d_2 = 0\}$





$\mathcal{T}(\mathcal{F}, x^*) = \{d \in \mathbb{R}^2 \,:\, d_1 \leq 0,\, d_2 = 0\}$ $\qquad$ $\mathcal{T}_{\text{lin}}(\mathcal{F}, x^*) = \{d \in \mathbb{R}^2 \,:\, d_1 \leq 0,\, d_2 = 0\}$

e) The gradient of f in x^* is given by

$$\nabla f(x^*) = \begin{pmatrix} -2 \\ 0 \end{pmatrix}.$$

For the condition $\nabla f(x^*)^T d \geq 0$ to hold, we have to have $d_1 \leq 0$ and $d_2 \in \mathbb{R}$ arbitrary. Obviously we have $d_1 \leq 0$ for all $d \in \mathcal{T}(\mathcal{F}, x^*)$ for problem (I) and (II). For the linearized tangential cone $\mathcal{T}_{\text{lin}}(\mathcal{F}, x^*)$ we only have this for problem (II), whereas for problem (I) also directions with $d_1 > 0$ are contained in the set. Therefore, we do not have a necessary condition if we replace the tangential cone with the linearized tangential cone. But the linearized tangential cone is really what we want to work with, so that we need the Abadie Constraint Qualification to asure the necessity.

3. *(Constraint optimization)*
Consider the objective function $f : \mathbb{R}^2 \to \mathbb{R}$ with

$$f(x, y) = 2x^2 + xy + 3y^2 - 4x - y.$$

How should the parameter $\alpha, \beta \in \mathbb{R}$ be selected, such that the point $x^* = (3, -2)^T \in \mathbb{R}$ is a minimum subject to the constrain $h(x, y) = 0$, where

$$h(x, y) := 3x + \alpha y - \beta?$$

(6 Points)

Solution: The necessary condition for $(3, -2)^T \in \mathbb{R}^2$ to be a minimum is with respect to the constraint $h(x, y) = 0$ the existence of a Lagrange multiplier $\lambda^* \in \mathbb{R}$, s.t. $\nabla L(3, -2, \lambda^*) = 0$.

The optimization problem is convex with linear equality constraints, therefore this condition is even sufficient (see Theorem 3.1.20).

The Lagrange function is given by

$$\mathcal{L}(x, y, \lambda) = f(x, y) + \lambda g(x, y) = 2x^2 + xy + 3y^2 - 4x - y + \lambda(3x + \alpha y - \beta).$$

and for the gradient it holds:

$$\nabla \mathcal{L}(x, y, \lambda) = \begin{pmatrix} 4x + y - 4 + 3\lambda \\ x + 6y - 1 + \alpha\lambda \\ 3x + \alpha y - \beta \end{pmatrix}.$$

We plug in the point $(3, -2)^T \in \mathbb{R}^2$ for which we get:

$$0 \overset{!}{=} \nabla \mathcal{L}(3, -2, \lambda) = \begin{pmatrix} 6 + 3\lambda \\ -10 + \alpha\lambda \\ 9 + \alpha y - \beta \end{pmatrix},$$

This leads us to the solution $\lambda = -2$, $\alpha = -5$ and $\beta = 19$.