A tutorial on properties of the epigraph reformulation

Oliver Stein*

October 29, 2024

Abstract

This paper systematically surveys useful properties of the epigraph reformulation for optimization problems, and complements them by some new results. We focus on the complete compatibility of the original formulation and the epigraph reformulation with respect to solvability and unsolvability, the compatibility with respect to some, but not all, basic constraint qualifications, the formulation of first order optimality conditions for problems with max-type objective function, and the interpretation of feasibility and optimality cuts along epigraphs in the framework of cutting plane methods. Finally we introduce a generalized epigraph reformulation which is particularly useful for treating summands of objective and constraint functions independently in the reformulation.

1 Introduction

For a set $X \subseteq \mathbb{R}^n$ and a function $f : X \to \mathbb{R}$ the set

$$\text{epi}(f, X) \;=\; \{(x, \alpha) \in X \times \mathbb{R} \mid f(x) \leq \alpha\}$$

is called the epigraph of f on X (Fig. 1). Basic properties of epigraphs characterize important properties of f on X, e.g. convexity and lower semi-continuity.

*Institute for Operations Research (IOR), Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany, stein@kit.edu

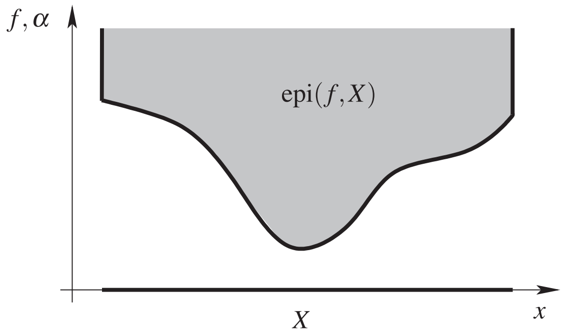

Figure 1: Epigraph of f on X

Indeed, the set X and the function $f : X \to \mathbb{R}$ are convex if and only if $\text{epi}(f, X)$ is convex, and in particular the existence of subgradients of convex functions (at points from the interior of their domain) can be shown by proving the existence of outer normal vectors to their epigraph [4, 9]. Furthermore, a function f is lower semi-continuous on X if and only if $\text{epi}(f, X)$ is closed relative to $X \times \mathbb{R}$ [10]. Together with a projection argument, the latter can, e.g., be employed to derive semi-continuity properties of optimal value functions in parametric optimization [3, 10]. For concepts like epi-convergence, epi-derivatives, epi-addition, and epi-multiplication of functions, etc., we refer to [3].

This paper collects properties of epigraphs which are useful in optimization models. Most of these properties have previously been stated elsewhere, or are simply known as 'folklore', but we are not aware of a systematic compilation of these results. The present paper aims to close this gap. It does not, however, intend to provide original references of the single results. Rather, for details it often refers to textbooks by the present author.

With any minimization problem

$$P: \quad \min_{x \in \mathbb{R}^n} f(x) \quad \text{s.t.} \quad x \in X$$

one can relate its epigraph reformulation (also called epigraphical reformulation)

$$P_{\text{epi}}: \quad \min_{(x,\alpha) \in \mathbb{R}^n \times \mathbb{R}} \alpha \quad \text{s.t.} \quad (x, \alpha) \in \text{epi}(f, X)$$

which in more explicit terms reads

$$P_{\text{epi}}: \quad \min_{(x,\alpha) \in \mathbb{R}^n \times \mathbb{R}} \alpha \quad \text{s.t.} \quad f(x) \le \alpha,\ x \in X.$$

Two major applications of the epigraph reformulation are the generation of a linear objective function, and the algorithmic treatment of objective functions f of max-type.

In fact, since the objective function $f_{\text{epi}}(x, \alpha) = \alpha$ of P_{epi} is linear in the vector of decision variables (x, α), P_{epi} may be treated by algorithms which require a linear objective function, like Kelley's cutting plane method for convex problems. Also, while the minimal value of a solvable nonconvex problem cannot always be computed as the minimal value of the convex hull problem of P, it is always identical to the minimal value of the convex hull problem of P_{epi}, thanks to the linear objective function of the latter [4].

If, for some finite index set K, the objective function of P is of the form $f = \max_{k \in K} f_k$ with functions f_k, $k \in K$, then the epigraph reformulation

$$\min_{(x,\alpha)\in\mathbb{R}^n\times\mathbb{R}} \alpha \quad \text{s.t.} \quad \max_{k\in K} f_k(x) \leq \alpha,\ x \in X$$

may be rewritten as

$$P_{\text{epi}}: \quad \min_{(x,\alpha)\in\mathbb{R}^n\times\mathbb{R}} \alpha \quad \text{s.t.} \quad f_K(x) \leq \alpha e,\ x \in X,$$

where f_K stands for the vector-valued function with entries f_k, $k \in K$, e denotes the all-ones vector of appropriate dimension, and the inequality is understood component-wise.

While P is in general a nonsmooth optimization problem, for smooth functions f_k and a smooth functional description of X the latter reformulation is a smooth problem. Also, if X is a polyhedral set and the functions f_k are affine, then the reformulation is an LP. If X and all functions $f_k : X \to \mathbb{R}$ are convex, then the convexity of the epigraphs $\text{epi}(f_k, X)$, the identity $\text{epi}(\max_{k\in K} f_k, X) = \bigcap_{k\in K} \text{epi}(f_k, X)$, and the fact that intersections of convex sets are again convex yield the convexity of $\max_{k\in K} f_k$. Therefore, both P and P_{epi} are convex optimization problems.

Example 1.1. *For $X \subseteq \mathbb{R}^n$ and $z \in \mathbb{R}^n$ the projection problem with respect to the Chebyshev norm is*

$$P: \quad \min_{x\in\mathbb{R}^n} \|x - z\|_\infty \quad s.t. \quad x \in X.$$

Its minimal value is the ℓ_∞-distance of z from X, and every minimal point is called ℓ_∞-projection of z to X. In view of the max-structures in the objective function,

$$\|x - z\|_\infty = \max_{k=1,\dots,n} |x_k - z_k| = \max_{k=1,\dots,n} \max\{\pm(x_k - z_k)\},$$

the epigraph reformulation can be written as

$$P_{\mathrm{epi}}: \quad \min_{x,\alpha} \alpha \quad s.t. \quad \pm(x-z) \leq \alpha e,\ x \in X.$$

If, for example, the set X is polyhedral, then the nonsmooth problem P is reformulated into the linear optimization problem P_{epi}.

This paper is structured as follows. In Section 2 we show that there is a one-to-one correspondence between the solvability of P and P_{epi} as well as one-to-one correspondences between the cases of unsolvability. We shall also prove a one-to-one correspondence between local minimal points. Section 3 treats one-to-one correspondences between the validity of three major constraint qualifications in P and P_{epi}, namely the Slater, Mangasarian-Fromovitz and Abadie constraint qualifications. It also discusses why such a correspondence does in general not hold for the linear independence constraint qualification. Section 4 uses these results to state necessary and sufficient first order optimality conditions for problems P with max-type objective functions. In Section 5 we briefly discuss how feasibility and optimality cuts in cutting plane methods are related to the epigraph reformulation, before Section 6 provides a useful generalization of the epigraph reformulation. Section 7 closes this paper with some final remarks.

2 Compatibility of solvability properties

The results in this section hold without any assumptions on X and $f : X \to \mathbb{R}$.

In preparation for the main result of this section, we first show that local minimality in the problem P_{epi} is necessarily 'global with respect to the epigraph variable α'.

Lemma 2.1. *A point $(\bar{x}, \bar{\alpha})$ is a local minimal point of P_{epi} if and only if the following three conditions hold:*

$$(\bar{x}, \bar{\alpha}) \in \mathrm{epi}(f, X), \tag{1}$$

$$\bar{\alpha} = f(\bar{x}), \tag{2}$$

there exists a neighborhood U of $\bar{x}$ with

$$\forall\, (x, \alpha) \in \mathrm{epi}(f, X) \cap (U \times \mathbb{R}) : \ \alpha \geq f(\bar{x}). \tag{3}$$

Proof. Let $(\bar{x}, \bar{\alpha})$ be a local minimal point of P_{epi}, that is, (1) holds, and there exist a neighborhood U of $\bar{x}$ as well as some $\varepsilon > 0$ with

$$\forall\, (x, \alpha) \in \text{epi}(f, X) \cap (U \times (\bar{\alpha} - \varepsilon, \bar{\alpha} + \varepsilon)) :\ \alpha \geq \bar{\alpha}. \tag{4}$$

From (1) we know $\bar{\alpha} \geq f(\bar{x})$. Assuming the violation of (2) hence means $\bar{\alpha} > f(\bar{x})$. Then, after possibly reducing ε, we may assume $\bar{\alpha} - \varepsilon > f(\bar{x})$. Defining $\widetilde{x} := \bar{x}$ and $\widetilde{\alpha} := \bar{\alpha} - \varepsilon/2$ thus implies

$$(\widetilde{x}, \widetilde{\alpha}) \in \text{epi}(f, X) \cap (U \times (\bar{\alpha} - \varepsilon, \bar{\alpha} + \varepsilon)) \ \text{ and } \ \widetilde{\alpha} < \bar{\alpha}.$$

This contradicts the local minimality of $(\bar{x}, \bar{\alpha})$ and proves (2). In particular, every local minimal point $(\bar{x}, \bar{\alpha})$ of P_{epi} is of the form $(\bar{x}, f(\bar{x}))$, and (4) can be rewritten as

$$\forall\, (x, \alpha) \in \text{epi}(f, X) \cap (U \times (f(\bar{x}) - \varepsilon, f(\bar{x}) + \varepsilon)) :\ \alpha \geq f(\bar{x}). \tag{5}$$

To finally prove (3), consider any $(x, \alpha) \in \text{epi}(f, X) \cap (U \times \mathbb{R})$. In view of (5) it remains to consider the cases $\alpha \geq f(\bar{x}) + \varepsilon$ and $\alpha \leq f(\bar{x}) - \varepsilon$.

The case $\alpha \geq f(\bar{x}) + \varepsilon$ yields the desired inequality $\alpha \geq f(\bar{x})$. On the other hand, $\alpha \leq f(\bar{x}) - \varepsilon$ implies $f(x) \leq \alpha \leq f(\bar{x}) - \varepsilon \leq f(\bar{x}) - \varepsilon/2$, so that $(x, f(\bar{x}) - \varepsilon/2)$ lies in $\text{epi}(f, X) \cap (U \times (f(\bar{x}) - \varepsilon, f(\bar{x}) + \varepsilon))$. Due to (5) this yields the contradiction $f(\bar{x}) - \varepsilon/2 \geq f(\bar{x})$. Therefore the case $\alpha \leq f(\bar{x}) - \varepsilon$ cannot occur, and (3) is shown.

The sufficiency part of the assertion is clear, since the global statement with respect to α in (3) implies the corresponding local statement. ☐

Theorem 2.2. *For every $X \subseteq \mathbb{R}^n$ and $f : X \to \mathbb{R}$ the problems P and P_{epi} are equivalent in the following sense:*

a) For every local or global minimal point $\bar{x}$ of P, $(\bar{x}, f(\bar{x}))$ is a local or global minimal point of P_{epi}, respectively.

b) For every local or global minimal point $(\bar{x}, \bar{\alpha})$ of P_{epi}, $\bar{x}$ is a local or global minimal point of P, respectively.

c) P is unbounded if and only if P_{epi} is unbounded.

d) P is infeasible if and only if P_{epi} is infeasible.

e) The infimum of P is finite and not attained if and only if the infimum of P_{epi} is finite and not attained.

f) The infima of P and P_{epi} coincide.

Proof. For the proof of part a, let $\bar{x}$ be a local minimal point of P, that is, $\bar{x} \in X$ holds, and there is some neighborhood U of $\bar{x}$ with $f(x) \geq f(\bar{x})$ for all $x \in X \cap U$.

In view of $(\bar{x}, f(\bar{x})) \in \text{epi}(f, X)$ and Lemma 2.1, it is sufficient to show (3) for the given neighborhood U. Indeed, all $(x, \alpha) \in \text{epi}(f, X) \cap (U \times \mathbb{R})$ fulfill $\alpha \geq f(x) \geq f(\bar{x})$.

The proof of the corresponding assertion for global minimal points follows from the choice $U = \mathbb{R}^n$ in the above arguments (where, alternatively, a direct proof would not need to rely on Lemma 2.1).

To see part b, let $(\bar{x}, \bar{\alpha})$ be a local minimal point of P_{epi} with corresponding neighborhood U. Then Lemma 2.1 implies $(\bar{x}, \bar{\alpha}) \in \text{epi}(f, X)$, from which $\bar{x} \in X$ follows, as well as (3). In view of $(x, f(x)) \in \text{epi}(f, X) \cap (U \times \mathbb{R})$ for all $x \in U$, the latter yields $f(x) \geq f(\bar{x})$ for these x, and thus the local minimality of $\bar{x}$ for P.

Again, the result about global minimal points can be shown by the choice $U = \mathbb{R}^n$.

In part c, the unboundedness of P is equivalent to the existence of some sequence $(x^\ell) \subseteq X$ with $f(x^\ell) \leq -\ell$ for all $\ell \in \mathbb{N}$. The definition $\alpha^\ell := f(x^\ell)$ thus yields a sequence $(x^\ell, \alpha^\ell) \in \text{epi}(f, X)$ with $\alpha^\ell \leq -\ell$, implying the unboundedness of P_{epi}. For the reverse direction, any sequence $(x^\ell, \alpha^\ell) \in \text{epi}(f, X)$ with $\alpha^\ell \leq -\ell$ enforces $f(x^\ell) \leq \alpha^\ell \leq -\ell$ and, thus, the unboundedness of P.

Part d is seen from the facts that $X = \emptyset$ implies $\text{epi}(f, X) = \emptyset$, and that $X \neq \emptyset$, with the choice $\bar{\alpha} := f(\bar{x})$ for some $\bar{x} \in X$, implies $\text{epi}(f, X) \neq \emptyset$.

Since apart from solvability (parts a and b), unboundedness (part c), and infeasibility (part d), only the nonattainment of a finite infimum can occur in any optimization problem [4], and since for the occurrence of the first three cases we have shown one-to-one correspondences between P and P_{epi}, also the statement of part e must hold.

To finally prove part f, we show that the sets of lower bounds of the problems P and P_{epi} coincide. This will include the formal cases of the infima $\pm\infty$, corresponding to infeasibility and unboundedness, respectively.

Indeed, let ω be a lower bound of P, that is, $f(x) \geq \omega$ holds for all $x \in X$. Then all $(x, \alpha) \in \text{epi}(f, X)$ satisfy $\alpha \geq f(x) \geq \omega$, so that ω is also a lower bound for P_{epi}. Vice versa, if ω is a lower bound for P_{epi}, then $\alpha \geq \omega$ holds for all $(x, \alpha) \in \text{epi}(f, X)$. In particular, with the choices $\alpha = f(x)$ we obtain $f(x) \geq \omega$ for all $x \in X$, which shows that ω is a lower bound for P. □

As employed in the proof of the above theorem, there exist exactly the three cases of unsolvability of optimization problems from Theorem 2.2c,d,e. In [4] also this result is shown by an epigraph argument, namely by studying the possible outcomes of the parallel projection of $\mathrm{epi}(f, X)$ to the 'α-space'. In case of solvability one obtains a set $[v, +\infty)$ with the minimal value $v \in \mathbb{R}$ of P, and else a set $(v, +\infty)$ with $v \in \{\pm\infty\}$ (for infeasible or unbounded problems, resp.) or $v \in \mathbb{R}$ (for nonattained infima). In all four cases the 'lower boundary point' v of the interval is the infimum of P.

In multicriteria optimization, that is, problems P with a vector-valued objective function $f : \mathbb{R}^n \to \mathbb{R}^m$, minimal points generalize to efficient points, and minimal values to nondominated points. One can likewise define the epigraph $\mathrm{epi}(f, X) = \{(x, \alpha) \in \mathbb{R}^n \times \mathbb{R}^m \mid f(x) \leq \alpha\}$, where the vector inequality is understood component-wise. Boundary points of its parallel projection to the 'α-space' $\mathbb{R}^m$ (the so-called upper image set) determine the set of nondominated points, in analogy to the above single-criterion case.

With respect to Theorem 2.2c we remark that for $X \neq \emptyset$ the feasible set of P_{epi} is always unbounded, but still the objective function of P_{epi} is bounded from below on this set, unless the problem P is unbounded.

The intrinsic unboundedness of the feasible set of P_{epi} can impose a formal problem for algorithms which require bounded feasible sets, like branch-and-bound methods. If such an algorithm cannot be appropriately modified, an artificial upper bound $\bar{\alpha}$ on the epigraph variable must be introduced. A first possibility for this is to compute an upper bound $\bar{\alpha}$ for f on X which, at least for a factorable function f and a box X, may be achieved by interval arithmetic [4]. A usually more tractable approach is to choose some $\bar{\alpha}$ such that the lower level set $\{x \in X \mid f(x) \leq \bar{\alpha}\}$ is nonempty, e.g., $\bar{\alpha} := f(\bar{x})$ with some feasible point $\bar{x} \in X$. Then the minimal point sets of P and

$$P_{\bar{\alpha}} : \quad \min_x \ f(x) \quad \text{s.t.} \quad f(x) \leq \bar{\alpha}, \ x \in X$$

coincide. Since the maximal value of f on the feasible set of $P_{\bar{\alpha}}$ cannot exceed $\bar{\alpha}$, an appropriate epigraph reformulation of $P_{\bar{\alpha}}$ is

$$\min_{x,\alpha} \ \alpha \quad \text{s.t.} \quad f(x) \leq \alpha \leq \bar{\alpha}, \ x \in X.$$

3 Compatibility of constraint qualifications

For some constraint qualifications, their validity in P can be characterized by their corresponding validity in P_{epi}.

The Slater constraint qualification

We start with the discussion of Slater's constraint qualification, which is often formulated for convexly described problems of the form

$$P: \quad \min_x f(x) \quad \text{s.t.} \quad g_I(x) \leq 0,\ Ax = b$$

with a finite index set I, convex functions f, g_i, $i \in I$, g_I denoting the vector-valued function with entries g_i, $i \in I$, as well as a matrix A and a vector b of appropriate dimensions. The set

$$X \;=\; \{x \in \mathbb{R}^n \mid g_I(x) \leq 0,\ Ax = b\} \tag{6}$$

is said to satisfy the Slater constraint qualification if there exists some point $x^\star \in \mathbb{R}^n$ with $g_I(x^\star) < 0$ and $Ax^\star = b$. Every such point $x^\star$ is called a Slater point of P.

Under the above assumptions, also the set

$$\text{epi}(f, X) \;=\; \{(x, \alpha) \in \mathbb{R}^n \times \mathbb{R} \mid f(x) \leq \alpha,\ g_I(x) \leq 0,\ Ax = b\} \tag{7}$$

is convexly described, since the function $f(x) - \alpha$ is convex in (x, α).

Theorem 3.1. *For a finite index set I, convex functions $f, g_i : \mathbb{R}^n \to \mathbb{R}$, $i \in I$, and a matrix A and a vector b of appropriate dimensions, the set X from* (6) *satisfies the Slater constraint qualification if and only if* epi(f, X) *from* (7) *satisfies the Slater constraint qualification.*

Proof. The 'if' part of the assertion is clear, since for every Slater point $(x^\star, \alpha^\star)$ of epi(f, X) from (7), the point $x^\star$ is a Slater point of X from (6). For the proof of the 'only if' part let $x^\star$ be a Slater point of X from (6). Then the point $(x^\star, \alpha^\star)$ with $\alpha^\star := f(x^\star) + 1$ satisfies $f(x^\star) < \alpha^\star$, $g_I(x^\star) < 0$, and $Ax^\star = b$. It is, thus, a Slater point of epi(f, X) from (7). □

Note that the arguments of the above proof do not rely on the convexity assumption for the involved functions, so that they also cover the Slater constraint qualification for nonconvex problems, if required.

In view of the subsequent developments, we remark that the Slater constraint qualification is a global constraint qualification on the entire set X, and that it does not require any differentiability properties on the functions f and g_i, $i \in I$. In particular, Theorem 3.1 covers the case $f = \max_{k \in K} f_k$ with convex functions f_k.

On the other hand, for nonconvex problems P, usually local constraint qualifications at given points $\bar{x} \in X$ are formulated, and they require some differentiability assumptions. While for the constraint functions we will indeed assume differentiability, the subsequent results will also hold for certain nondifferentiable objective functions, in particular $f = \max_{k \in K} f_k$ with differentiable functions f_k. More generally, the following results on the Mangasarian-Fromovitz and Abadie constraint qualifications can also be shown for objective and constraint functions which are only one-sided directionally differentiable in the sense of Hadamard. We will not pursue these results in the framework of this tutorial, but rather refer to [7] for appropriate notions of constraint qualifications in nonsmooth optimization.

Indeed, in the remainder of this section we consider the problem

$$P: \quad \min_x \max_{k \in K} f_k(x) \quad \text{s.t.} \quad g_I(x) \leq 0, \; h_J(x) = 0$$

with finite index sets I, J, K and differentiable functions f_k, $k \in K$, g_i, $i \in I$, h_j, $j \in J$, and h_J denoting the vector-valued function with entries h_j, $j \in J$. The feasible set of P is

$$X \;=\; \{x \in \mathbb{R}^n \mid g_I(x) \leq 0, \; h_J(x) = 0\}, \tag{8}$$

and we write the epigraph of $f = \max_{k \in K} f_k$ on X as

$$\text{epi}(f, X) \;=\; \{(x, \alpha) \in \mathbb{R}^n \times \mathbb{R} \,| f_K(x) \leq \alpha e, \; g_I(x) \leq 0, \; h_J(x) = 0\}. \tag{9}$$

Since constraint qualifications are mainly applied at local or global minimal points, in view of the results from Section 2 subsequently we shall focus on correspondences between constraint qualifications at points $\bar{x} \in X$ and $(\bar{x}, f(\bar{x})) \in \text{epi}(f, X)$.

The Mangasarian-Fromovitz constraint qualification

In the following let $I_0(\bar{x}) = \{i \in I \mid g_i(\bar{x}) = 0\}$ denote the active index set at $\bar{x} \in X$, let $\nabla g_{I_0(\bar{x})}(x)$ be the matrix with columns $\nabla g_i(x)$, $i \in I_0(\bar{x})$, as well as $\nabla h_J(x)$ the matrix with columns $\nabla h_j(x)$, $j \in J$.

The Mangasarian-Fromovitz constraint qualification (MFCQ) is said to hold at $\bar{x}$ in X from (8) if $\nabla h_J(\bar{x})$ possesses full column rank, and if a direction $d \in \mathbb{R}^n$ with

$$\nabla g_{I_0(\bar{x})}(\bar{x})^\intercal d < 0, \; \nabla h_J(\bar{x})^\intercal d = 0 \tag{10}$$

exists. Moreover, with the active index set

$$K_0(\bar{x}, f(\bar{x})) = \{k \in K \mid f_k(\bar{x}) = f(\bar{x})\} \left(= \{k \in K \mid f_k(\bar{x}) = \max_{\ell \in K} f_\ell(x)\} \right) \tag{11}$$

the MFCQ holds at $(\bar{x}, f(\bar{x}))$ in $\text{epi}(f, X)$ from (9) if and only if the vectors $(\nabla h_j(\bar{x}), 0)$, $j \in J$, are linearly independent, and if there exists a direction (d, δ) with (10) and

$$\nabla f_{K_0(\bar{x}, f(\bar{x}))}(\bar{x})^\intercal d \ < \ \delta e. \tag{12}$$

Theorem 3.2. *For finite index sets I, J, K, let the functions f_k, $k \in K$, g_i, $i \in I$, h_j, $j \in J$, be differentiable at $\bar{x} \in X$. Then the MFCQ holds at $\bar{x}$ in X from* (8) *if and only if the MFCQ holds at $(\bar{x}, f(\bar{x}))$ in* $\text{epi}(f, X)$ *from* (9).

Proof. Let the MFCQ hold at $(\bar{x}, f(\bar{x}))$ in $\text{epi}(f, X)$ from (9). Since the vectors $(\nabla h_j(\bar{x}), 0)$, $j \in J$, are linearly independent if and only if the vectors $\nabla h_j(\bar{x})$, $j \in J$, are, the MFCQ holds at $\bar{x}$ in X from (8). For the reverse direction, let the MFCQ hold at $\bar{x}$ in X from (8) with a corresponding MF vector d. Then the vectors $(\nabla h_j(\bar{x}), 0)$, $j \in J$, are linearly independent and (10) holds. To show the existence of some $\delta \in \mathbb{R}$ with (12) it suffices to put $\delta := \max_{k \in K_0(\bar{x}, f(\bar{x}))} \langle \nabla f_k(\bar{x}), d \rangle + 1$. □

For $|K| = 1$ Theorem 3.2 particularly covers the case of a differentiable objective function f. In this case $K_0(\bar{x}, f(\bar{x})) = K$ holds.

For the case that the convexly described set X from (6) is nonempty and the matrix A possesses full column rank, we remark that the following three properties are equivalent [5]: (i) X satisfies the Slater constraint qualification, (ii) the MFCQ holds somewhere in X, (iii) the MFCQ holds everywhere in X.

The Abadie constraint qualification

For the formulation of the Abadie constraint qualification at a point $\bar{x}$ in X from (8) let

$$L(\bar{x}, X) \ = \ \{d \in \mathbb{R}^n \mid \nabla g_{I_0(\bar{x})}(\bar{x})^\intercal d \leq 0, \ \nabla h_J(\bar{x})^\intercal d = 0\}$$

denote the (outer) linearization cone to X at $\bar{x}$, and

$$T(\bar{x}, X) \ = \ \{d \in \mathbb{R}^n \mid \exists\, (t^\ell) \searrow 0, \, (d^\ell) \to d \ \forall\, \ell \in \mathbb{N} : \ \bar{x} + t^\ell d^\ell \in X\}$$

the (outer) tangent cone to X at $\bar{x}$ (aka Bouligand tangent cone or contingent cone). While the inclusion $T(\bar{x}, X) \subseteq L(\bar{x}, X)$ is always true [5], the reverse inclusion may fail. The Abadie constraint qualification (ACQ) is said to hold at $\bar{x} \in X$ if also $L(\bar{x}, X) \subseteq T(\bar{x}, X)$ is true. The MFCQ implies the ACQ at any $\bar{x}$ in X from (8) ([5], and [2] for the case including equality constraints), but the ACQ may also hold when the MFCQ is violated.

With the active index set $K_0(\bar{x}, f(\bar{x}))$ from (11) we obtain the linearization cone of $\mathrm{epi}(f, X)$ from (9) at $(\bar{x}, f(\bar{x}))$

$$\begin{aligned} L(\,(\bar{x}, f(\bar{x})), \mathrm{epi}(f, X)\,) = \{(d, \delta) \in \mathbb{R}^n \times \mathbb{R} \mid \nabla f_{K_0(\bar{x}, f(\bar{x}))}(\bar{x})^\intercal d \le \delta e, \\ \nabla g_{I_0(\bar{x})}(\bar{x})^\intercal d \le 0,\ \nabla h_J(\bar{x})^\intercal d = 0\}, \end{aligned}$$

and (d, δ) lies in the tangent cone $T(\,(\bar{x}, f(\bar{x})), \mathrm{epi}(f, X)\,)$ if and only if there are sequences $(t^\ell) \searrow 0$ and $(d^\ell, \delta^\ell) \to (d, \delta)$ with $(\bar{x}, f(\bar{x})) + t^\ell(d^\ell, \delta^\ell) \in \mathrm{epi}(f, X)$ for all $\ell \in \mathbb{N}$. More explicitly, the latter means $\bar{x} + t^\ell d^\ell \in X$ and $f_k(\bar{x} + t^\ell d^\ell) \le f(\bar{x}) + t^\ell \delta^\ell$ for all $k \in K$ and $\ell \in \mathbb{N}$.

Theorem 3.3. *For finite index sets I, J, K, let the functions f_k, $k \in K$, g_i, $i \in I$, h_j, $j \in J$, be differentiable at $\bar{x} \in X$. Then the ACQ holds at $\bar{x}$ in X from* (8) *if and only if the ACQ holds at $(\bar{x}, f(\bar{x}))$ in* $\mathrm{epi}(f, X)$ *from* (9).

Proof. Let the ACQ hold at $(\bar{x}, f(\bar{x}))$ in $\mathrm{epi}(f, X)$ from (9). To show the ACQ at $\bar{x}$ in X from (8), choose some $d \in L(\bar{x}, X)$ and define

$$\delta := \max_{k \in K_0(\bar{x}, f(\bar{x}))} \langle \nabla f_k(\bar{x}), d \rangle. \tag{13}$$

Then (d, δ) lies in $L(\,(\bar{x}, f(\bar{x})), \mathrm{epi}(f, X)\,)$. By the assumption of ACQ at $(\bar{x}, f(\bar{x}))$ in $\mathrm{epi}(f, X)$ there are sequences $(t^\ell) \searrow 0$ and $(d^\ell, \delta^\ell) \to (d, \delta)$ with $(\bar{x}, f(\bar{x})) + t^\ell(d^\ell, \delta^\ell) \in \mathrm{epi}(f, X)$, $\ell \in \mathbb{N}$, which in particular implies $\bar{x} + t^\ell d^\ell \in X$, $\ell \in \mathbb{N}$. This shows $d \in T(\bar{x}, X)$ and, thus $L(\bar{x}, X) \subseteq T(\bar{x}, X)$.

For the proof of the reverse direction, let the ACQ hold at $\bar{x}$ in X from (8) and choose some $(d, \delta) \in L(\,(\bar{x}, f(\bar{x})), \mathrm{epi}(f, X)\,)$. Then we have $d \in L(\bar{x}, X)$ and, by the ACQ at $\bar{x}$ in X, the existence of sequences $(t^\ell) \searrow 0$ and $(d^\ell) \to d$ with $\bar{x} + t^\ell d^\ell \in X$, $\ell \in \mathbb{N}$. To show the validity of the ACQ at $(\bar{x}, f(\bar{x}))$ in $\mathrm{epi}(f, X)$ from (9), it thus suffices to construct a sequence $(\delta^\ell) \to \delta$ with

$$f_k(\bar{x} + t^\ell d^\ell) \le f(\bar{x}) + t^\ell \delta^\ell \text{ for all } k \in K,\ \ell \in \mathbb{N}. \tag{14}$$

For each $k \in K_0(\bar{x}, f(\bar{x}))$ the latter is equivalent to

$$\frac{f_k(\bar{x} + t^\ell d^\ell) - f_k(\bar{x})}{t^\ell} \le \delta^\ell \text{ for all } \ell \in \mathbb{N},$$

where the left hand side converges to $\langle \nabla f_k(\bar{x}), d \rangle$, since the differentiability of the function f_k at $\bar{x}$ implies its one-sided directional differentiability in the sense of Hadamard. This motivates to put

$$\delta^\ell \ := \ \delta + \max_{k \in K_0(\bar{x}, f(\bar{x}))} \left(\frac{f_k(\bar{x} + t^\ell d^\ell) - f_k(\bar{x})}{t^\ell} - \langle \nabla f_k(\bar{x}), d \rangle \right).$$

Indeed, we obtain $(\delta^\ell) \to \delta$, and each $k \in K_0(\bar{x}, f(\bar{x}))$ satisfies

$$\delta^\ell \ \geq \ \delta + \frac{f_k(\bar{x} + t^\ell d^\ell) - f_k(\bar{x})}{t^\ell} - \langle \nabla f_k(\bar{x}), d \rangle \ \geq \ \frac{f_k(\bar{x} + t^\ell d^\ell) - f(\bar{x})}{t^\ell},$$

where the last inequality follows from the choice $(d, \delta) \in L(\,(\bar{x}, f(\bar{x})), \mathrm{epi}(f, X)\,)$ and from $f_k(\bar{x}) = f(\bar{x})$ for $k \in K_0(\bar{x}, f(\bar{x}))$.

It remains to show (14) for all $k \notin K_0(\bar{x}, f(\bar{x}))$. In this case we have $f_k(\bar{x}) < f(\bar{x})$, so that the continuity of f_k at $\bar{x}$, together with $(t^\ell) \searrow 0$ and the boundedness of (d^ℓ, δ^ℓ), guarantees (14) for all sufficiently large ℓ. □

Again, for $|K| = 1$ Theorem 3.3 covers the case of a smooth objective function f.

We remark that also the proof of the inclusion $T(\bar{x}, X) \subseteq L(\bar{x}, X)$ relies on the one-sided directional differentiability in the sense of Hadamard of the involved differentiable functions, g_i, $i \in I_0(\bar{x})$, h_j, $j \in J$ [5].

Corollary 3.4. *Let the set X from* (8) *be described by affine functions g_i, $i \in I$, h_j, $j \in J$, and for a finite index set K let the functions f_k, $k \in K$, be differentiable at $\bar{x} \in X$. Then the ACQ holds at $(\bar{x}, f(\bar{x}))$ in* $\mathrm{epi}(f, X)$ *from* (9).

Proof. Under the polyhedrality assumption on X, the ACQ holds at every point $\bar{x}$ in X from (8) [5], so that the assertion follows from Theorem 3.3. □

The linear independence constraint qualification

As seen above, in simple words, the epigraph reformulation of a minimization problem with max-type objection function does not interfere with the validity of the Slater, Mangasarian-Fromovitz and Abadie constraint qualifications. In general, however, this is not true for the linear independence constraint qualification (LICQ), which is said to hold at $\bar{x}$ in X from (8) if the gradients $\nabla g_i(\bar{x})$, $i \in I_0(\bar{x})$, $\nabla h_j(\bar{x})$, $j \in J$, are linearly independent. The LICQ implies the MFCQ at any $\bar{x}$ in X from (8), but the MFCQ may also hold when the LICQ is violated.

The LICQ holds at $(\bar{x}, f(\bar{x}))$ in $\text{epi}(f, X)$ from (9) if the vectors

$$\begin{pmatrix} \nabla f_k(\bar{x}) \\ -1 \end{pmatrix}, \; k \in K_0(\bar{x}, f(\bar{x})), \; \begin{pmatrix} \nabla g_i(\bar{x}) \\ 0 \end{pmatrix}, \; i \in I_0(\bar{x}), \; \begin{pmatrix} \nabla h_j(\bar{x}) \\ 0 \end{pmatrix}, \; j \in J$$

are linearly independent. By elementary column transformations for the matrix formed by these columns, the latter condition is equivalent to the linear independence of the vectors

$$\nabla f_k(\bar{x}) - \nabla f_\ell(\bar{x}), \; k \in K_0(\bar{x}, f(\bar{x})) \setminus \{\ell\}, \; \nabla g_i(\bar{x}), \; i \in I_0(\bar{x}), \; \nabla h_j(\bar{x}), \; j \in J, \tag{15}$$

where ℓ is chosen arbitrarily from $K_0(\bar{x}, f(\bar{x}))$.

Hence, while the LICQ at $(\bar{x}, f(\bar{x}))$ in $\text{epi}(f, X)$ from (9) implies the LICQ at $\bar{x}$ in X from (8), vice versa this is not necessarily the case. In particular, from (15) one sees that the LICQ in the epigraph reformulation requires the vectors $\nabla f_k(\bar{x})$, $k \in K_0(\bar{x}, f(\bar{x}))$, to be affinely independent, and also that the relation

$$|K_0(\bar{x}, f(\bar{x}))| \; \leq \; n - |I_0(\bar{x})| - |J| + 1$$

must hold. For example, in the case $|I_0(\bar{x})| + |J| = n$ this is only possible for $|K_0(\bar{x}, f(\bar{x}))| = 1$. However, at least in the smooth case $|K| = 1$ the LICQ at $\bar{x}$ in X from (8) implies the LICQ at $(\bar{x}, f(\bar{x}))$ in $\text{epi}(f, X)$ from (9).

Recall the hierarchy LICQ $\Rightarrow$ MFCQ $\Rightarrow$ ACQ at any feasible point in X and in $\text{epi}(f, X)$, respectively, where the reverse directions are in general wrong, and where for convexly described problems the MFCQ is equivalent to the Slater constraint qualification in the sense mentioned above. In view of this hierarchy, the LICQ can be viewed as too strong to be compatible with the epigraph reformulation, while all discussed weaker constraint qualifications are compatible.

4 First order optimality conditions

The previous results allow us to formulate necessary and sufficient first order optimality conditions for optimization problems with max-type objective functions. Indeed, we again consider the problem

$$P: \quad \min_x \max_{k \in K} f_k(x) \quad \text{s.t.} \quad g_I(x) \leq 0, \; h_J(x) = 0$$

with finite index sets I, J, K and differentiable functions f_k, $k \in K$, g_i, $i \in I$, h_j, $j \in J$. Its feasible set X is described as in (8). We consider the epigraph reformulation of P in the form

$$P_{\text{epi}}: \quad \min_{x,\alpha} \alpha \quad \text{s.t.} \quad f_K(x) - \alpha e \le 0,\ g_I(x) \le 0,\ h_J(x) = 0,$$

that is, with a feasible set epi(f, X) described as in (9). The Karush-Kuhn-Tucker (KKT) conditions for this formulation of P_{epi} at a point $(\bar{x}, f(\bar{x})) \in \text{epi}(f, X)$ state the existence of multipliers $\kappa_k \ge 0$, $k \in K_0(\bar{x}, f(\bar{x}))$, $\lambda_i \ge 0$, $i \in I_0(\bar{x})$, $\mu_j \in \mathbb{R}$, $j \in J$, with

$$\begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \sum_{k \in K_0(\bar{x}, f(\bar{x}))} \kappa_k \begin{pmatrix} \nabla f_k(\bar{x}) \\ -1 \end{pmatrix} + \sum_{i \in I_0(\bar{x})} \lambda_i \begin{pmatrix} \nabla g_i(\bar{x}) \\ 0 \end{pmatrix} + \sum_{j \in J} \mu_j \begin{pmatrix} \nabla h_j(\bar{x}) \\ 0 \end{pmatrix}.$$

With the convex hull $\text{conv}(\{\nabla f_k(\bar{x}),\ k \in K_0(\bar{x}, f(\bar{x}))\})$, the convex conical hull $\text{cone}(\{\nabla g_i(\bar{x}),\ i \in I_0(\bar{x})\})$ and the 'linear hull' $\text{span}(\{\nabla h_j(\bar{x}),\ j \in J\})$ of the corresponding vectors, the KKT conditions at $(\bar{x}, f(\bar{x})) \in \text{epi}(f, X)$ can be written more briefly as

$$\begin{aligned} 0 \ \in\ & \text{conv}(\{\nabla f_k(\bar{x}),\ k \in K_0(\bar{x}, f(\bar{x}))\}) \\ & + \text{cone}(\{\nabla g_i(\bar{x}),\ i \in I_0(\bar{x})\}) + \text{span}(\{\nabla h_j(\bar{x}),\ j \in J\}). \end{aligned} \tag{16}$$

The sum of these sets is understood in the Minkowski sense.

Theorem 4.1. *Let $\bar{x}$ be a local minimal point of P at which the functions f_k, $k \in K$, g_i, $i \in I$, h_j, $j \in J$, are differentiable, and let the ACQ hold at $\bar{x}$ in X from* (8). *Then the KKT condition* (16) *is satisfied.*

Proof. By Theorem 2.2a the point $(\bar{x}, f(\bar{x}))$ is locally minimal for P_{epi}. Theorem 3.3 guarantees that the ACQ holds at $(\bar{x}, f(\bar{x}))$ in epi(f, X) from (9). The KKT theorem under ACQ [2, 5] thus yields (16). □

Analogous arguments and Corollary 3.4 show the next result.

Corollary 4.2. *Let X be polyhedral and let $\bar{x}$ be a local minimal point of P at which the functions f_k, $k \in K$, are differentiable. Then the corresponding KKT condition* (16) *is satisfied.*

For convexly described problems of the form

$$P: \quad \min_x \max_{k \in K} f_k(x) \quad \text{s.t.} \quad g_I(x) \le 0,\ Ax = b$$

we can state a necessary as well as a sufficient first order optimality condition. The feasible set X of P is described as in (6).

Theorem 4.3. *For convex functions f_k, $k \in K$, g_i, $i \in I$, with finite index sets I, K, as well as a matrix A and a vector b of appropriate dimensions, let the Slater constraint qualification hold in X from* (6)*, and let $\bar{x}$ be a minimal point of P at which the f_k, $k \in K$, g_i, $i \in I$ are differentiable. Then the KKT condition* (16) *holds, where the vectors $\nabla h_j(\bar{x})$, $j \in J$, are the columns of the transposed matrix $A^\intercal$.*

Proof. By Theorem 2.2a the point $(\bar{x}, f(\bar{x}))$ is minimal for P_{epi}. Theorem 3.1 guarantees that $\text{epi}(f, X)$ from (7) satisfies the Slater constraint qualification. Then also the differentiable description of the epigraph

$$\text{epi}(f, X) \;=\; \{(x, \alpha) \in \mathbb{R}^n \times \mathbb{R} \mid f_K(x) \leq \alpha e,\; g_I(x) \leq 0,\; Ax = b\} \qquad (17)$$

satisfies the Slater constraint qualification and, therefore, the KKT theorem for differentiably and convexly described problems [2, 5] yields the assertion. □

As common in convex optimization, the following first order sufficient optimality condition holds without the assumption of a constraint qualification.

Theorem 4.4. *For convex functions f_k, $k \in K$, g_i, $i \in I$, as well as a matrix A and a vector b of appropriate dimensions, let the KKT condition* (16) *hold for P_{epi} at a point $(\bar{x}, f(\bar{x}))$ in $\text{epi}(f, X)$ from* (17)*, so that the f_k, $k \in K$, g_i, $i \in I$ are differentiable at $\bar{x}$. Then $\bar{x}$ is a minimal point of P.*

Proof. By the sufficiency for minimality of the KKT condition in convexly described problems [4], $(\bar{x}, f(\bar{x}))$ is a minimal point of P_{epi}. Theorem 2.2b thus yields the assertion. □

For convex differentiable functions f_k, $k \in K$, in convex analysis it is shown that the set $\text{conv}(\{\nabla f_k(\bar{x}),\, k \in K_0(\bar{x}, f(\bar{x}))\})$ coincides with the convex subdifferential $\partial f(\bar{x})$ of $f = \max_{k \in K} f_k$ at $\bar{x}$ [9]. For the considered convexly described problems the KKT condition (16) may therefore be rewritten as

$$0 \;\in\; \partial f(\bar{x}) + \text{cone}(\{\nabla g_i(\bar{x}),\, i \in I_0(\bar{x})\}) + \text{range}(A^\intercal).$$

Moreover, for not necessarily convex, but continuously differentiable f_k, $k \in K$, the function f is locally Lipschitz continuous and subdifferentially regular, so that also its Clarke subdifferential coincides with $\text{conv}(\{\nabla f_k(\bar{x}),\, k \in K_0(\bar{x}, f(\bar{x}))\})$ [1].

5 Cutting plane methods

Cutting plane methods approximate algorithmically difficult by easier to handle optimization problems. In comparison to an LP the difficulties can arise, for example, from the presence of nonlinear functions or of discrete variables. Indeed, consider an optimization problem

$$P: \quad \min_x f(x) \quad \text{s.t.} \quad x \in M \cap X$$

with a polyhedral set X and some set M 'hosting the difficulties' like descriptions by nonlinear functions or the discreteness of variables. We assume $f : M \to \mathbb{R}$ to be defined on M, but not necessarily on the entire set X.

The epigraph reformulation

$$P_{\text{epi}}: \qquad \min_{x,\alpha} \alpha \quad \text{s.t.} \quad f(x) \le \alpha,\ x \in M \cap X$$

of P possesses three types of constraints on the vector (x, α) of decision variables, namely the inequality $f(x) \le \alpha$, the constraint $x \in M$ and the condition $x \in X$. The algorithmically difficult part of the problem is therefore modeled by the first two constraints, while the third constraint and the objective function are polyhedral and linear, respectively. The two difficult constraints describe the epigraph

$$\text{epi}(f, M) \;=\; \{(x, \alpha) \in M \times \mathbb{R} |\ f(x) \le \alpha\}$$

of f on M. A possible representation of the epigraph reformulation thus is

$$P_{\text{epi}}: \quad \min_{x,\alpha} \alpha \quad \text{s.t.} \quad (x, \alpha) \;\in\; \text{epi}(f, M) \,\cap\, (X \times \mathbb{R}).$$

A cutting plane method iteratively generates polyhedral relaxations $\widehat{\text{epi}(f, M)}$ of $\text{epi}(f, M)$ and solves the relaxed optimization problems

$$\widehat{P}_{\text{epi}}: \qquad \min_{x,\alpha} \alpha \quad \text{s.t.} \quad (x, \alpha) \in \widehat{\text{epi}(f, M)} \,\cap\, (X \times \mathbb{R}).$$

It terminates if a computed optimal point $(\widehat{x}^\star, \widehat{\alpha}^\star)$ (approximately) lies in $\text{epi}(f, M)$. Else, i.e. for $(\widehat{x}^\star, \widehat{\alpha}^\star) \notin \text{epi}(f, M)$, it adds a cut to the description of $\widehat{\text{epi}(f, M)}$. Such a cut is defined by a linear inequality $d^\intercal x + \delta\alpha \le b$ with $d \in \mathbb{R}^n$ and $\delta, b \in \mathbb{R}$, which firstly satisfies

$$d^\intercal \widehat{x}^\star + \delta \widehat{\alpha}^\star > b$$

i.e., it is violated by $(\widehat{x}^\star, \widehat{\alpha}^\star)$, and secondly is valid for $\text{epi}(f, M)$. The latter means

$$\forall\, (x, \alpha) \in \text{epi}(f, M): \quad d^\intercal x + \delta\alpha \leq b.$$

The fact that the set $\text{epi}(f, M)$ is described by two separate conditions induces the following effects. In the case $(\widehat{x}^\star, \widehat{\alpha}^\star) \notin \text{epi}(f, M)$ at least one of the two conditions $f(\widehat{x}^\star) \leq \widehat{\alpha}^\star$ and $\widehat{x}^\star \in M$ must be violated. The way in which a cut is constructed therefore depends on which of the two conditions is violated.

In the case $f(\widehat{x}^\star) > \widehat{\alpha}^\star$ it suggests itself to exploit a property of the function f at the point $\widehat{x}^\star$ to generate the cut. However, as long as it is not clear whether $\widehat{x}^\star \in M$ holds, f may not be defined at $\widehat{x}^\star$ and it may, thus, not make sense to look for such a property.

For this reason, cutting plane methods first check the condition $\widehat{x}^\star \in M$. In the case $\widehat{x}^\star \notin M$ they construct a cut with respect to this feasible set, called feasibility cut. It has the form of an inequality $d^\intercal x \leq b$ valid for M with $d^\intercal \widehat{x}^\star > b$, which does naturally not depend on the epigraph variable α (i.e., $\delta = 0$ holds). Because of $\text{epi}(f, M) \subseteq M \times \mathbb{R}$ this inequality is also valid for $\text{epi}(f, M)$. A feasibility cut can therefore be constructed independently of the validity of the second condition $f(x) \leq \alpha$ in the description of the epigraph. For example, if $M = \{x \in \mathbb{R}^n \mid g_i(x) \leq 0,\, i \in I\}$ is described by convex functions g_i, $i \in I$, one chooses some $k \in I$ with $g_k(\widehat{x}^\star) > 0$, some subgradient s of g_k at $\widehat{x}^\star$, and defines the Kelley cut

$$g_k(\widehat{x}^\star) + \langle s, x - \widehat{x}^\star \rangle \;\leq\; 0.$$

If no feasibility cut is required, $\widehat{x}^\star$ lies in the domain M of f, and in the case $f(\widehat{x}^\star) > \widehat{\alpha}^\star$ a cut can be constructed with information about f at the point $\widehat{x}^\star$. The validity of the corresponding inequality $d^\intercal x + \delta\alpha \leq b$ is not required for all $(x, \alpha) \in \mathbb{R}^n \times \mathbb{R}$ with $f(x) \leq \alpha$, since f is only defined on M, but only for all $(x, \alpha) \in M \times \mathbb{R}$ with $f(x) \leq \alpha$. This results in a valid inequality for $\text{epi}(f, M)$, called optimality cut.

If, for example, M and $f: M \to \mathbb{R}$ are convex, then also the function $f(x) - \alpha$ is convex, and with any subgradient s of f at $\widehat{x}^\star$ the corresponding Kelley cut is

$$0 \geq f(\widehat{x}^\star) - \widehat{\alpha}^\star + \left\langle \begin{pmatrix} s \\ -1 \end{pmatrix}, \begin{pmatrix} x \\ \alpha \end{pmatrix} - \begin{pmatrix} \widehat{x}^\star \\ \widehat{\alpha}^\star \end{pmatrix} \right\rangle = f(\widehat{x}^\star) + \langle s, x - \widehat{x}^\star \rangle - \alpha.$$

The point $(\widehat{x}^\star, \widehat{\alpha}^\star)$ violates this inequality due to $f(\widehat{x}^\star) > \widehat{\alpha}^\star$. Remarkably, the point $(\widehat{x}^\star, f(\widehat{x}^\star)) \in \text{epi}(f, M)$ satisfies the inequality with equality so that,

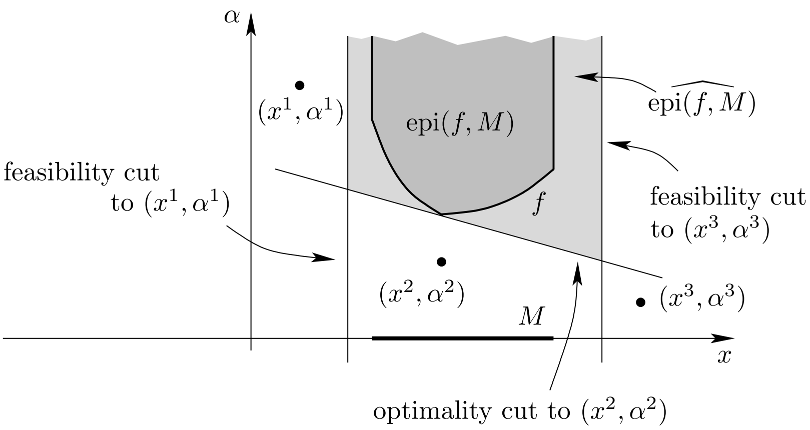

Figure 2: Feasibility and optimality cuts

in contrast to Kelley feasibility cuts, the Kelley optimality cuts are deepest possible in the sense that they define supporting hyperplanes to epi(f, M).

Figure 2 illustrates the concepts of feasibility and optimality cuts and the resulting relaxation of an epigraph.

6 Generalized epigraph reformulation

This section generalizes the epigraph reformulation in two aspects which are useful for establishing tractable optimization models. For example, the generalization allows to split the objective function into summands, which enter the reformulation separately. At the same time, we introduce this technique for the treatment of inequality constraints.

To this end, a functional $F : \mathbb{R}^m \to \mathbb{R}$ is called monotone (on $\mathbb{R}^m$), if

$$\forall\, x, y \in \mathbb{R}^m \ \text{ with } \ x \le y : \quad F(x) \le F(y)$$

holds, where the vector inequalities are understood component-wise.

For $X \subseteq \mathbb{R}^n$, functions $f : X \to \mathbb{R}^m$ and $g : X \to \mathbb{R}^p$ as well as monotone functionals $F : \mathbb{R}^m \to \mathbb{R}$ and $G : \mathbb{R}^p \to \mathbb{R}$, consider the problem

$$P : \quad \min_{x \in \mathbb{R}^n} F(f(x)) \quad \text{s.t.} \quad G(g(x)) \le 0,\ x \in X.$$

With the generalized epigraph

$$\text{gepi}(f, g, X) \ := \ \{(x, \alpha, \beta) \in X \times \mathbb{R}^m \times \mathbb{R}^p\} \mid f(x) \le \alpha,\ g(x) \le \beta\}.$$

we may introduce the generalized epigraph reformulation

$$P_{\text{gepi}}: \quad \min_{(x,\alpha,\beta)\in\mathbb{R}^n\times\mathbb{R}^m\times\mathbb{R}^p} F(\alpha) \quad \text{s.t.} \quad G(\beta)\le 0,\ (x,\alpha,\beta)\in \text{gepi}(f,g,X).$$

More explicitly, it reads

$$\begin{aligned} P_{\text{gepi}}: \quad \min_{(x,\alpha,\beta)\in\mathbb{R}^n\times\mathbb{R}^m\times\mathbb{R}^p} F(\alpha) \quad \text{s.t.} \quad & G(\beta)\le 0, \\ & f(x)\le\alpha,\ g(x)\le\beta,\ x\in X. \end{aligned}$$

Rather than explicitly generalizing all previous results from the standard epigraph reformulation, here we only illustrate how to work with this technique along the extension of a rudimentary version of Theorem 2.2.

Theorem 6.1. *For every $X\subseteq\mathbb{R}^n$, $f: X\to\mathbb{R}^m$, $g: X\to\mathbb{R}^p$ and monotone functionals $F:\mathbb{R}^m\to\mathbb{R}$, $G:\mathbb{R}^p\to\mathbb{R}$, the problems P and P_{gepi} are equivalent in the following sense:*

a) For every global minimal point $\bar{x}$ of P, $(\bar{x}, f(\bar{x}), g(\bar{x}))$ is a global minimal point of P_{gepi} with the same minimal value.

b) For every global minimal point $(\bar{x},\bar{\alpha},\bar{\beta})$ of P_{gepi}, $\bar{x}$ is a global minimal point of P with the same minimal value.

Proof. For the proof of part a, let $\bar{x}$ be a global minimal point of P. Define $\bar{\alpha} := f(\bar{x})$ and $\bar{\beta} = g(\bar{x})$. Thus the constraints $f(x)\le\alpha$ and $g(x)\le\beta$ of P_{gepi} are satisfied at $(\bar{x},\bar{\alpha},\bar{\beta})$, and the feasibility of $\bar{x}$ in P also implies $G(\bar{\beta})\le 0$ and $\bar{x}\in X$, so that $(\bar{x},\bar{\alpha},\bar{\beta})$ is feasible for P_{gepi}.

For every other feasible point (x,α,β) of P_{gepi} we have $x\in X$, and the monotonicity of G yields $G(g(x))\le G(\beta)\le 0$, so that x is feasible for P. The minimality of $\bar{x}$ for P and the monotonicity of F imply $F(\alpha)\ge F(f(x))\ge F(f(\bar{x})) = F(\bar{\alpha})$, so that the minimality of $(\bar{x},\bar{\alpha},\bar{\beta})$ is shown, along with the identity of the minimal values.

To show part b, let $(\bar{x},\bar{\alpha},\bar{\beta})$ be a global minimal point of P_{gepi}. Its feasibility yields $\bar{x}\in X$ and, by the monotonicity of G, $G(g(\bar{x}))\le G(\bar{\beta})\le 0$. Therefore $\bar{x}$ is feasible for P. Together with the monotonicity of F the feasibility of $(\bar{x},\bar{\alpha},\bar{\beta})$ also implies $F(f(\bar{x}))\le F(\bar{\alpha})$.

Let x be any other feasible point of P and define $\alpha := f(x)$ and $\beta := g(x)$. Then (x,α,β) is feasible for P_{gepi}. The optimality of $(\bar{x},\bar{\alpha},\bar{\beta})$ for P_{gepi} thus implies $F(f(x)) = F(\alpha)\ge F(\bar{\alpha})\ge F(f(\bar{x}))$, which shows the global minimality of $\bar{x}$ for P.

It remains to show equality in the last of the above inequalities. Indeed, for $F(\bar{\alpha}) > F(f(\bar{x}))$ the feasible point $(\bar{x}, f(\bar{x}), \bar{\beta})$ would possess a better objective function value than $(\bar{x}, \bar{\alpha}, \bar{\beta})$, which contradicts the minimality of the latter. Therefore $F(\bar{\alpha}) = F(f(\bar{x}))$ holds. $\square$

Example 6.2. *For $X \subseteq \mathbb{R}^n$ and $z \in \mathbb{R}^n$ the projection problem with respect to the ℓ_1-norm is*

$$P: \quad \min_{x \in \mathbb{R}^n} \|x - z\|_1 \quad s.t. \quad x \in X.$$

Its minimal value is the ℓ_1-distance of z from X, and every minimal point is called ℓ_1-projection of z to X. The structure

$$\|x - z\|_1 = \sum_{k=1,\ldots,n} |x_k - z_k|$$

of the objective function allows us to use the generalized epigraph reformulation with the monotone functional $F(\alpha) := \sum_{k=1}^n \alpha_k = \langle e, \alpha \rangle$. The generalized epigraph reformulation can be written as

$$P_{\text{gepi}}: \quad \min_{(x,\alpha) \in \mathbb{R}^n \times \mathbb{R}^n} \langle e, \alpha \rangle \quad s.t. \quad \pm(x - z) \le \alpha, \; x \in X.$$

As in Example 1.1, if the set X is polyhedral, then the nonsmooth problem P has been reformulated into an LP. We remark that the generalized epigraph reformulation of the ℓ_1-projection problem P doubles the number of variables, while a different popular approach for treating the ℓ_1-norm, by splitting the vector $x - z$ in its component-wise positive and negative parts, would even triple the number of variables.

7 Final remarks

The epigraph reformulation is also useful for minimization problems with objective function of sup-type, where the supremum is taken over possibly infinitely many functions. Indeed, for $X \subseteq \mathbb{R}^n$ and a set-valued mapping $Y : X \rightrightarrows \mathbb{R}^m$ let $f(x) = \sup_{y \in Y(x)} f(x, y)$. If for some $x \in X$ the maximization problem of $f(x, \cdot)$ over $Y(x)$ is not solvable, the usual conventions for the supremum apply, i.e. $f(x) = -\infty$ for $Y(x) = \emptyset$, etc. The epigraph reformulation of the problem

$$P: \quad \min_x \sup_{y \in Y(x)} f(x, y) \quad \text{s.t.} \quad x \in X$$

can then be written as

$$P_{\text{epi}}: \quad \min_{x,\alpha} \alpha \quad \text{s.t.} \quad x \in X,\ f(x,y) \leq \alpha \quad \forall\, y \in Y(x).$$

The inequality constraints of P_{epi} are of generalized semi-infinite type. Theory and methods for such problems may be found in [6].

Acknowledgements

The author wishes to thank Immanuel Bomze for posing a question which sparked the compilation of this tutorial, and for helpful comments on an earlier version of this manuscript.

References

[1] A. Bagirov, N. Karmitsa, M.M. Mäkelä, *Introduction to Nonsmooth Optimization*, Springer, Cham, 2014.

[2] C. Geiger, C. Kanzow, *Theorie und Numerik restringierter Optimierungsaufgaben,* Springer, Berlin, 2002.

[3] R.T. Rockafellar, R.J. Wets, *Variational Analysis,* Springer Science & Business Media, 2009.

[4] O. Stein, *Basic Concepts of Global Optimization,* Mathematics Study Ressources 5, Springer, 2024.

[5] O. Stein, *Basic Concepts of Nonlinear Optimization,* Mathematics Study Ressources 8, Springer, 2024.

[6] O. Stein, *Bi-level Strategies in Semi-infinite Programming,* Kluwer Academic Publishers, Boston, 2003.

[7] O. Stein, *On constraint qualifications in non-smooth optimization,* Journal of Optimization Theory and Applications, Vol. 121 (2004), 647-671.

[8] O. Stein, *Grundzüge der Gemischt-Ganzzahligen Optimierung,* SpringerSpektrum, 2024.

[9] O. Stein, *Grundzüge der Konvexen Analysis,* SpringerSpektrum, 2021.

[10] O. Stein, *Grundzüge der Parametrischen Optimierung,* SpringerSpektrum, 2021.