Constraint qualifications for nonlinear programming

Consider the standard nonlinear program

$$\begin{array}{lll} \min & f(x) & \\ \text{s.t.} & g_i(x) \geq 0 & \forall i = 1, \ldots, m, \\ & h_j(x) = 0 & \forall 1 = 1, \ldots, p, \end{array} \tag{NLP}$$

with continuously differentiable functions $f, g_i, h_j : \mathbb{R}^n \to \mathbb{R}$. The feasible set of (NLP) will be denoted by Ω, i.e.,

$$\Omega := \{x \in \mathbb{R}^n \mid g_i(x) \leq 0\,(i = 1, \ldots, m), \;\; h_j(x) = 0\,(j = 1, \ldots, p)\}.$$

For a feasible point $\bar{x} \in \Omega$, the so-called *tangent cone* for (NLP) at $\bar{x}$ is just the tangent cone of Ω at $\bar{x}$, i.e.,

$$\begin{aligned} T(\bar{x}) = T_\Omega(\bar{x}) \;\; &:= \;\; \left\{d \;\middle|\; \exists\,\{x^k\} \subset \Omega, \{t_k\} \downarrow 0 : x^k \to \bar{x} \text{ and } \frac{x^k - \bar{x}}{t_k} \to d\right\} \\ &= \;\; \{d \mid \exists\{d^k\} \to d, \{t_k\} \downarrow 0 : \bar{x} + t_k d^k \in \Omega \;\forall k\}, \end{aligned}$$

and the *linearized cone* for (NLP) at $\bar{x}$ is

$$L(\bar{x}) := \left\{d \;\middle|\; \begin{array}{ll} \nabla g_i(\bar{x})^T d \leq 0 & (i : g_i(\bar{x}) = 0), \\ \nabla h_j(\bar{x})^T d = 0 & (j = 1, \ldots, p) \end{array} \right\}.$$

Note that both the tangent and the linearized cone are in fact closed cones and the latter is obviously polyhedral convex, where this is not necessarily true for the tangent cone. In addition to that, it is easy to see that the inclusion

$$T(\bar{x}) \subset L(\bar{x}) \tag{1}$$

holds for all $\bar{x} \in \Omega$.

The condition that equality holds in (1) is known as the *Abadie constraint qualification (ACQ)*, which we formally state in the following definition:

Definition 1 (ACQ) *Let $\bar{x}$ be feasible for (NLP). We say that the* Abadie constraint qualification *holds at $\bar{x}$ (and write* ACQ($\bar{x}$)*) if*

$$T(\bar{x}) = L(\bar{x}).$$

ACQ plays a key role for establishing necessary optimality conditions for (NLP), due to the well-known Karush-Kuhn-Tucker theorem, see [1, Theorem 5.1.3].

Theorem 2 (KKT conditions) *Let $\bar{x}$ be a local minimizer of (NLP) such that ACQ($\bar{x}$). Then there exist multipliers $\lambda \in \mathbb{R}^m$ and $\mu \in \mathbb{R}^p$ such that*

$$0 = \nabla f(\bar{x}) + \sum_{i=1}^{m} \lambda_i \nabla g_i(\bar{x}) + \sum_{j=1}^{p} \mu_j \nabla h_j(\bar{x}) \tag{2}$$

and

$$\lambda_i \geq 0, \quad \lambda_i g_i(\bar{x}) = 0 \quad (i = 1, \ldots, m). \tag{3}$$

In fact, the same result could be established under the following weaker condition:

Definition 3 (GCQ) *Let $\bar{x}$ be feasible for (NLP). We say that the* Guignard constraint qualification (GCQ) *holds at $\bar{x}$ (and write* GCQ($\bar{x}$)*) if*

$$T(\bar{x})^\circ = L(\bar{x})^\circ,$$

i.e., if the polar[1] *of the tangent equals the polar of the linearized cone.*

In general, we call a property of the feasible set a *constraint qualification* if it guarantees the KKT conditions to hold at a local minimizer.

GCQ is, in a sense, see [2], the weakest constraint qualification, and, as the following example shows, may be strictly weaker than ACQ.

Example 4 Consider

$$\min x_1^2 + x_2^2 \quad \text{s.t.} \quad x_1, x_2 \geq 0, \; x_1 x_2 = 0. \tag{4}$$

The global minimizer is $\bar{x} = (0, 0)^T$. We compute that

$$T(\bar{x}) = \{d \mid d_1, d_2 \geq 0, \; d_1 d_2 = 0\} \subsetneq \{d \mid d_1, d_2 \geq 0\} = L(\bar{x}).$$

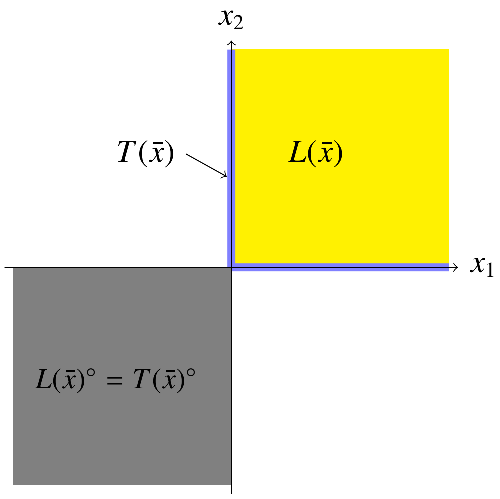

Figure 1: Tangent cone, linearized cone and their polars for (4) at $\bar{x}$

Hence, ACQ is violated at $\bar{x}$. On the other hand we have

$$T(\bar{x})^\circ = \{v \mid v_1, v_2 \leq 0\} = L(\bar{x})^\circ,$$

and hence GCQ is satsified at $\bar{x}$. Note, however, that ACQ (and hence GCQ) is satisfied at any other feasible point of (4).

[1]For a cone $C \subseteq \mathbb{R}^n$ its polar is given by $C^\circ := \{v \mid v^T d \leq 0 \; \forall d \in C\}$

ACQ and thus GCQ are considered to be pretty mild assumptions, which have a good chance to be statisfied. On the other hand they are hard to check, since the tangent cone may be hard to compute. Furthermore, they are, in general, not strong enough to guarantee convergence of possible algorithms for solving (NLP).

For these purposes the following two constraint qualifications are most useful.

Definition 5 *Let $\bar{x}$ be feasible for (NLP) and put $I(\bar{x}) := \{i \mid g_i(\bar{x}) = 0\}$. We say that*

a) the linear independence constraint qualification (LICQ) *holds at $\bar{x}$ (and write* LICQ($\bar{x}$)*) if the gradients*

$$\nabla g_i(\bar{x}) \ (i \in I(\bar{x})), \quad \nabla h_j(\bar{x}) \ (j = 1, \ldots, p)$$

are linearly independent.

b) the Mangasarian-Fromovitz constraint qualification (MFCQ) *holds at $\bar{x}$ (and write* MFCQ($\bar{x}$)*) if the gradients*

$$\nabla h_j(\bar{x}) \ (j = 1, \ldots, p)$$

are linearly independent and there exists a vector $d \in \mathbb{R}^n$ such that

$$\nabla g_i(\bar{x})^T d < 0 \ (i \in I(\bar{x})), \quad \nabla h_j(\bar{x})^T d = 0 \ (j = 1, \ldots, p).$$

Note that LICQ and MFCQ are violated at any feasible point of the nonlinear program from Example 4.

In order to establish our main theorem on the relation of the constraint qualifications introduced above, we need the following auxiliary result.

Lemma 6 *Let $\bar{x} \in \Omega$ such that MFCQ is satsified at $\bar{x}$. Then there exists $\varepsilon > 0$ and a C^1-curve $x : (-\varepsilon, \varepsilon) \to \mathbb{R}^n$ such that $x(t) \in \Omega$ for all $t \in [0, \varepsilon)$, $x(0) = \bar{x}$ and $x'(0) = d$.*

Proof. Define $H : \mathbb{R}^{p+1} \to \mathbb{R}^p$ by

$$H_i(y, t) = h_i(\bar{x} + td + h'(\bar{x})^T y) \quad \forall j = 1, \ldots, p,$$

where $h'(\bar{x})$ denotes the Jacobian of h at $\bar{x}$. The nonlinear equation $H(y, t) = 0$ has the solution$(\bar{y}, \bar{t}) = (0, 0)$ with

$$H'_y(0, 0) = h'(\bar{x}) h'(\bar{x})^T$$

and the latter matrix is non-singular (even positive definite) due to the linear independence of the vectors $\nabla h_j(\bar{x})\,(j = 1, \ldots, p)$. The *implicit function theorem* yields a C^1-function $y : (-\varepsilon, \varepsilon) \to \mathbb{R}^p$ such that $y(0) = 0$, $H(y(t), t) = 0$ and

$$y'(t) = -H'_y(y(t), t)^{-1} H'_t(y(t), t)$$

for all $t \in (-\varepsilon, \varepsilon)$. Hence, we have

$$y'(0) = -H'_y(0, 0)^{-1} H'_t(0, 0) = -H'_y(0, 0)^{-1} h'(\bar{x}) d = 0.$$

Now, put $x(t) = \bar{x} + td + h'(\bar{x})^T y(t)$ for all $t \in (-\varepsilon, \varepsilon)$. Reducing ε if necessary, $x : (-\varepsilon, \varepsilon) \to \mathbb{R}^n$ has all desired properties: Obviously, $x \in C^1$, $x(0) = \bar{x}$, $x'(0) = d$ and $h_i(x(t)) = 0$ for all $t \in (-\varepsilon, \varepsilon)$. Moreover, by continuity we have $g_i(x(t)) < 0$ for all $i \notin I(\bar{x})$ and $|t|$ sufficiently small. For $i \in I(\bar{x})$ we have $g_i(x(0)) = g_i(\bar{x}) = 0$ and

$$\frac{d}{dt} g_i(x(0)) = \nabla g_i(\bar{x})^T d < 0$$

and hence $g_i(x(t)) < 0$ for all $t > 0$ sufficiently small. □

Theorem 7 *Let $\bar{x}$ be feasible for (NLP). Then the following implications hold:*

$$\text{LICQ}(\bar{x}) \implies \text{MFCQ}(\bar{x}) \implies \text{ACQ}(\bar{x}) \implies \text{GCQ}(\bar{x}). \tag{5}$$

Proof. $\text{LICQ}(\bar{x}) \implies \text{MFCQ}(\bar{x})$: Obviously, the vectors $\nabla h_j(\bar{x})$ $(j = 1, \ldots, p)$ are linear independent. It remains to find a suitable vector d: For these purposes, consider the matrix

$$\begin{pmatrix} \nabla g_i(\bar{x})^T & (i \in I(\bar{x})) \\ \nabla h_j(\bar{x})^T & (j = 1, \ldots, p) \end{pmatrix} \in \mathbb{R}^{(|I(\bar{x})|+p)\times n},$$

which has full rank by $\text{LICQ}(\bar{x})$. Hence we can add rows to obtain non-singular matrix $A(\bar{x}) \in \mathbb{R}^{n\times n}$, and hence the linear equation

$$A(\bar{x})d = \begin{pmatrix} -e \\ 0 \end{pmatrix}$$

with e being the vector (in $\mathbb{R}^{|I(\bar{x})|}$) of all ones has a solution $\hat{d}$, which fulfills the requirements for $\text{MFCQ}(\bar{x})$.

$\text{MFCQ}(\bar{x}) \implies \text{ACQ}(\bar{x})$: In view of (1), it suffices to show that $L(\bar{x}) \subseteq T(\bar{x})$ holds. Hence, let $d \in L(\bar{x})$ and $\hat{d}$ given by $\text{MFCQ}(\bar{x})$ such that

$$\nabla g_i(\bar{x})^T \hat{d} < 0 \; \forall i \in I(\bar{x}), \quad \nabla h_j(\bar{x})^T \hat{d} = 0 \; \forall j.$$

Put $d(\delta) := d + \delta\hat{d}$ for $\delta > 0$. Then for all $\delta > 0$ we have

$$\nabla g_i(\bar{x})^T d(\delta) < 0 \; \forall i \in I(\bar{x}), \quad \nabla h_i(\bar{x})^T d(\delta) = 0 \; \forall i.$$

We claim that this implies $d(\delta) \in T(\bar{x})$ for all $\delta > 0$: By Lemma 6 there exists a C^1-curve $x : (-\varepsilon, \varepsilon) \to \mathbb{R}^n$ such that $x(t) \in \Omega$ for all $t \in [0, \varepsilon)$, $x(0) = \bar{x}$ and $x'(0) = d(\delta)$. For an arbitrary sequence $\{t_k\} \downarrow 0$ and $x^k := x(t_k)$ we hence infer that $x^k \to_X \bar{x}$ and thus

$$d(\delta) = x'(0) = \lim_{k\to\infty} \frac{x(t_k) - \bar{x}}{t_k - 0} = \lim_{k\to\infty} \frac{x^k - \bar{x}}{t_k} \in T(\bar{x}).$$

And since $T(\bar{x})$ is closed, this implies $d = \lim_{\delta\downarrow 0} d(\delta) \in T(\bar{x})$.

$\text{ACQ}(\bar{x}) \implies \text{GCQ}(\bar{x})$: Follows immediately from the definitions.

□

We would like to point out that there is a variety of other constraint qualifications for (NLP) occuring in the literature. A very comprehensive suryey on that topic is given by [3].

References

[1] J.V. Burke: *Numerical Optimization.* Course Notes, AMath/Math 516, University of Washington, Spring Term 2012.

[2] F.J. Gould and J.W. Tolle: *A necessary and sufficient qualification for constrained optimization.* SIAM J. Appl. Math. 20, 1971, pp. 164-172.

[3] D. W. Peterson: *A review of constraint qualifications in finite-dimensional spaces.* SIAM Review 15, 1973, pp. 639-654.