Differential conditions for constrained nonlinear programming via Pareto optimization

by
Adriano Pascoletti and Paolo Serafini
University of Udine, Department of Mathematics and Computer Science, and CISM, International Center for Mechanical Sciences, Udine, Italy

Abstract: We deal with differential conditions for local optimality. The conditions we derive for inequality constrained problems do not require constraint qualifications and are the broadest conditions based only on first and second order derivatives. A similar result is proved for equality constrained problems, although necessary conditions require regularity of the equality constraints.

Keywords: differential conditions, constraint qualifications, local optimality, Pareto optimality.

1. INTRODUCTION

In this paper we deal with the following problem:

$$\begin{array}{ll} \max & f_0(x) \\ & f_i(x) \geq 0 \quad i = 1, \ldots, m \end{array} \tag{1}$$

where $f_i : R^n \to R$, $i = 0, \ldots, m$, are twice differentiable maps and the maximization is local. In (1) only inequality constraints are considered. We prefer to deal separately with the case in which equality constraints are also included, and postpone this case to Section 4.

The characterization of the local maxima of (1) by differential conditions has been object of extensive research starting from the pioneering paper by Kuhn and Tucker [11] and the earlier result by Karush [10]. All results related to necessary and sufficient differential conditions require additional assumptions on the constraint behavior near the optima. These assumptions have been generally referred to as *constraint qualifications* (CQ). Although investigation on several types of constraint qualifications has started long ago (we only mention the results by [2], [12], [1]) the subject is still focus of intense research as recent papers show. Among many others, we mention in particular [4], [9], [13], [14]. Recently the attention has been also devoted to extending the constraint qualifications to more complex problems. See [17], [8], [6]. We mention in particular [14] in which the gap between necessary and sufficient conditions for optimality is closed in the case of polynomial problems by resorting to an elaborate mathematical formulation.

If the optimality conditions are limited to the inspection of the first and second order derivatives there is no way to prevent the gap, but there is a possibility to free the conditions from constraint qualifications. The key is an old result by Wan [16] which was proved in the context of multi objective optimization. Instead of building a Lagrangean function $f_0(x) + \sum_i \lambda_i f_i(x)$, we reformulate (1) as an unconstrained Pareto optimization problem with the $(m+1)$ criteria $(f_0, \ldots, f_m)$, where all functions play the same role and there is no dichotomy between objective and constraints. We then apply the differential conditions

Corresponding author: Paolo Serafini, email: serafini@dimi.uniud.it, postal address:
Dipartimento di Matematica e Informatica, Via delle Scienze 206, 33100 Udine, Italy

for Pareto optimality stated in [16] and derive necessary and sufficient conditions which embed the known results as particular cases.

We remark that also Ben Tal [3] obtained conditions without constraint qualifications but at the expense of having multipliers which are functions of feasible directions and are not fixed. This makes the conditions not easily manageable.

An important feature of the conditions derived in this paper is that they are the broadest conditions one can state by using only first and second order derivatives. By this we mean that if a point is a local maximum whose optimality can be detected by means of first and second order conditions, then optimality is detected also from the conditions stated in this paper. Conversely if a point is not a local maximum and its nonoptimality can be detected by means of first and second order conditions, then nonoptimality is detected also from our conditions.

2. PARETO OPTIMALITY CONDITIONS

Let $F(x) = (f_0(x), \ldots, f_m(x))^\top$. In the sequel we use the following notation for two vectors a and b: $a > b$ means $a_i > b_i$ for each i; $a \geq b$ means $a_i \geq b_i$ for each i and $a \neq b$; $a \geqq b$ means $a_i \geq b_i$ for each i. Let us consider the following alternative definitions of Pareto optima:

Definition 1: *A point $\hat{x} \in R^n$ is* weakly Pareto optimal (WP) *if there exists a neighborhood N of $\hat{x}$ such that $F(x) > F(\hat{x})$ for no $x \in N$.* ∎

Definition 2: *A point $\hat{x} \in R^n$ is* Pareto optimal (P) *if there exists a neighborhood N of $\hat{x}$ such that $F(x) \geq F(\hat{x})$ for no $x \in N$.* ∎

Definition 3: *A point $\hat{x} \in R^n$ is* strongly Pareto optimal (SP) *if there exists a neighborhood N of $\hat{x}$ such that $F(x) \geqq F(\hat{x})$ for no $x \in N$, $x \neq \hat{x}$.* ∎

The following implications hold trivially

$$\hat{x} \text{ is SP} \quad \Longrightarrow \quad \hat{x} \text{ is P} \quad \Longrightarrow \quad \hat{x} \text{ is WP} \tag{2}$$

Let us denote by $DF(x)$ and $D^2F(x)$ respectively the first and second order derivatives of F at x. Let $p := m + 1 - \operatorname{rank} DF(x) = \dim \ker DF(x)^\top$, $Q(x) : R^{m+1} \to R^p$ be any linear surjective operator such that $Q(x)\, DF(x) = 0$, and

$$K(x) := \left\{ h \in R^n : DF(x)\, h \geqq 0 \right\}$$

In the sequel we may occasionally drop in the notation the dependence on x of $DF(x)$, $D^2F(x)$, $Q(x)$ and $K(x)$ if the context makes it clear. Note that, by definition, the rows of Q are linearly independent and span $\ker DF^\top$. We may now state the following differential conditions proved in [16]:

Condition P1 *(first order necessity)*: *If x is WP, then there exists $\pi \in R^{m+1}$, $\pi \geq 0$, such that $\pi\, DF(x) = 0$.* ∎

Condition P2 *(first order sufficiency)*: *If $\ker DF(x) = \{0\}$ and there exists $\pi > 0$ such that $\pi\, DF(x) = 0$, then x is SP.* ∎

Condition P3 *(second order necessity)*: *If x is WP, then $Q(x)\, D^2F(x)(h,h) \neq Q(x)\, c$ for any $h \in K(x)$, $c > 0$.* ∎

Condition P4 *(second order sufficiency)*: *If there exists $\pi \geq 0$ such that $\pi\, DF(x) = 0$ and*

$$Q(x)\, D^2F(x)(h,h) \neq Q(x)\, c \quad \text{for any } h \in K(x),\ h \neq 0,\ c \geqq 0$$

then x is SP. ∎

We note that, whenever $p = 1$ ($\operatorname{rank} DF(x) = (m+1) - 1$), then π may be identified with $Q(x)$ and Conditions P3 and P4 require the quadratic form $\pi\, D^2F(x)$ to be negative semidefinite and definite respectively on $K(x)$. Moreover, if $\pi > 0$, then $K(x) = \ker DF(x)$. Hence in these cases the stated conditions become simpler. However, if $p > 1$ ($\operatorname{rank} DF(x) < (m+1) - 1$) then there are alternative linearly independent multipliers and, by using the operator Q, it is possible to take care of all of them simultaneously. This provides a stronger condition than working with just the vector π (or any vector obtained as linear combination of the rows of Q). Example 3 in Section 5 shows that working with just the vector π fails in detecting the optimum while by using Q the optimum is identified.

If $p = 0$ then $\ker DF(x)^\top = \{0\}$ (obviously fulfilled only if $n \geq m+1$) and clearly x cannot be optimal by P1. Note the existence of a first order sufficient condition (obviously fulfilled only if $n \leq m$). We also note that a stronger first order sufficiency condition is provided by $K(x) = \{0\}$. We omit the proof which can be found in [15].

Corollary P2 *(first order sufficiency)*: *If $K(x) = \{0\}$ then x is SP.*

3. CONDITIONS FOR LOCAL MAXIMA

Now we link Conditions P1-P4 with problem (1). We assume without loss of generality that all constraints are active at $\hat{x}$, i.e. $f_i(\hat{x}) = 0$, $i = 1\ldots, m$ (we may always take neighborhoods feasible for the nonactive constraints). Then a local maximum is defined as:

Definition 4: *A point $\hat{x}$ is a* maximum (M) *if there exists a neighborhood N of $\hat{x}$ such that $f_0(x) > f_0(\hat{x})$ and $f_i(x) \geq f_i(\hat{x})$, $i = 1, \ldots, m$, for no $x \in N$.* ∎

Definition 5: *A point $\hat{x}$ is a* strict maximum (SM) *if there exists a neighborhood N of $\hat{x}$ such that $f_i(x) \geq f_i(\hat{x})$, $i = 0, \ldots, m$, for no $x \in N$, $x \neq \hat{x}$.* ∎

Clearly the following implications hold:

$$\hat{x} \text{ is SP} \iff \hat{x} \text{ is SM} \implies \hat{x} \text{ is M} \implies \hat{x} \text{ is WP}$$

and therefore we may simply derive from Conditions P1–P4 the following conditions:

Condition M1 *(first order necessity)*: *If x is M, then there exists $\pi \in R^{m+1}$, $\pi \geq 0$, such that $\pi\, DF(x) = 0$.* ∎

Condition M2 *(first order sufficiency)*: *If* $\ker DF(x) = \{0\}$ *and there exists* $\pi > 0$ *such that* $\pi\, DF(x) = 0$, *then* x *is SM.* ∎

Corollary M2 *(first order sufficiency)*: *If* $K(x) = \{0\}$ *then* x *is SM.* ∎

Condition M3 *(second order necessity)*: *If* x *is M, then* $Q(x)\, D^2F(x)(h,h) \neq Q(x)\, c$ *for any* $h \in K(x)$, $c > 0$. ∎

Condition M4 *(second order sufficiency)*: *If there exists* $\pi \geq 0$ *such that* $\pi\, DF(x) = 0$ *and*

$$Q(x)\, D^2F(x)(h,h) \neq Q(x)\, c \quad \textit{for any } h \in K(x),\ h \neq 0,\ c \geqq 0$$

then x *is SM.* ∎

We note that Condition M1 is the well known Fritz John optimality condition. We recall that if $\pi_0 = 0$ this necessary condition gives little information because the point is candidate for optimality for any objective function f_0. Indeed CQ have been introduced in order to avoid this degeneracy. On the contrary our approach needs no CQ since the degenerate cases can be handled by second order conditions. It is worth pointing out that even if, in degenerate cases, the objective function disappears in $Q(x)\, D^2F(x)$, the dependence on f_0 can still be found in $K(x)$.

Condition M2 is a sufficient condition which does not rely on convexity properties of the functions f_i. It is essentially a linear programming complementarity condition with the additional requirement of $\ker DF(x) = \{0\}$ to rule out second order terms in the solution neighborhood.

Since Conditions P3-P4 are the "broadest" ones based only on first and second order derivatives (as stated in [16] Theorem 4), also Conditions M3-M4 are the broadest ones in the sense specified in the Introduction. This also means that the gap between necessary and sufficient conditions cannot be filled if we limit ourselves to consider first and second order derivatives.

We shall see from some examples that Conditions M1-M4 can determine optimality or nonoptimality even in cases where usual CQ do not hold. We briefly recall the most important CQ we shall refer to in the examples. Let us assume that all constraints are active at $\hat{x}$. Let $\tilde{F}(x) := (f_1(x), \ldots, f_m(x))^\top$ and $\tilde{K}(x) := \left\{h : D\tilde{F}(x)\, h \geqq 0\right\}$. Then we consider the following CQ:

– constraint regularity (RCQ): $\operatorname{rank} D\tilde{F}(\hat{x}) = m$; in other words $\pi\, DF(\hat{x}) = 0$ only if $\pi_0 \neq 0$;

– Kuhn-Tucker (KTCQ): for every $h \in \tilde{K}(\hat{x})$ there exists a function $\tau \in [0,1] \to x(\tau) \in R^n$ such that $x(\tau)$ is feasible, $\hat{x} = x(0)$, $dx(\tau)/d\tau|_{\tau=0} = \alpha\, h$, with $\alpha > 0$;

– Mangasarian-Fromovitz (MFCQ): there exist $h \in R^n$ such that $D\tilde{F}(\bar{x})\, h > 0$. In other words $\tilde{K}(\hat{x})$ has nonempty interior and no row of $D\tilde{F}(\hat{x})$ is null;

– Abadie (ACQ): $\tilde{K}(\hat{x})$ equals the Bouligand tangent cone $T_{\hat{x}}$ to the feasible set at $\hat{x}$. We recall that $T_{\hat{x}}$ is the set of $v \in R^n$ for which there exists a sequence of feasible points $x^{(n)} \to \hat{x}$ and a sequence of positive scalars λ_n such that $\lambda_n(x^{(n)} - \hat{x}) \to v$.

If the RCQ is fulfilled, then the Conditions M1, M3 and M4 become the well known nonlinear programming conditions with Lagrange multipliers λ such that $\pi = (1, \lambda)$. From M2 we may state the following side result:

Proposition 1: *If $m > n$, $\ker DF(\hat{x}) = \{0\}$, and there exist $\lambda_i > 0$ such that $Df_0(\hat{x}) + \sum_i \lambda_i \, Df_i(\hat{x}) = 0$, then $\hat{x}$ is SM.* ∎

4. EQUALITY CONSTRAINTS

The main result of this paper consists in Conditions M1-M4. Now the obvious question is the extension of the result to equality constraints. Therefore in this section we deal with the following problem:

$$\begin{aligned} \max \quad & f_0(x) \\ & f_i(x) \geq 0 \quad i = 1, \ldots, m \\ & h_j(x) = 0 \quad j = 1, \ldots, q \end{aligned} \tag{3}$$

where the functions f_i are as above and $h_j : R^n \to R$, $j = 1, \ldots, q$, are twice differentiable maps and the maximization is local. Let $F(x)$ as above and $H(x) = (h_1(x), \ldots, h_q(x))^\top$.

The conditions we may state in this case are not as neat as the ones with inequalities. Here regularity of the equality constraints does play a role in the necessary conditions, and cannot be avoided, unless maybe a thorough investigation with other tools is carried out, but this is out of the scope of this paper.

We consider the case of equality constraints in two different ways. In the first approach we embed equalities into (1) by converting them in a standard way into pairs of opposite inequalities, so that we face the inequality constrained problem:

$$\begin{aligned} \max \quad & f_0(x) \\ & f_i(x) \geq 0 \quad i = 1, \ldots, m \\ & h_j(x) \geq 0 \quad j = 1, \ldots, q \\ & -h_j(x) \geq 0 \quad j = 1, \ldots, q \end{aligned} \tag{4}$$

However, this approach has the drawback that all feasible points are trivially WP. So the necessary conditions M1 and M3 are automatically satisfied by all points and no information on candidates to optimality is provided. Indeed it is immediate to see that condition P1 is always verified. We will return to the necessary conditions in the second part of this section.

Let us now examine the sufficient conditions. It is useful to keep split in the notation the functions F and H. By applying condition M2 we ask for

$$\ker \begin{pmatrix} DF \\ DH \\ -DH \end{pmatrix} = \ker \begin{pmatrix} DF \\ DH \end{pmatrix} = \{0\}$$

and multipliers $\pi > 0$, $\mu^+ > 0$ and $\mu^- > 0$ such that $\pi \, DF + (\mu^+ - \mu^-) \, DH = 0$. This is equivalent to ask for multipliers $\pi > 0$ and μ (unrestricted in sign) such that $\pi \, DF + \mu \, DH = 0$.

By applying condition M4 we first verify the existence of multipliers $\pi \geqq 0$, $\mu^+ \geqq 0$ and $\mu^- \geqq 0$ (not all zero) such that $\pi \, DF + (\mu^+ - \mu^-) \, DH = 0$ is satisfied. However, this condition is always verified by the multipliers $\pi = 0$, $\mu^+ = \mu^- > 0$. Let us call *trivial* the multipliers $\pi \geqq 0$, $\mu = \mu^+ - \mu^-$, such that $(\pi, \mu) = 0$. We shall see shortly that if the condition is satisfied only by the trivial multipliers the second order sufficient condition gives no information. Let p such that

$$\operatorname{rank} \begin{pmatrix} DF \\ DH \\ -DH \end{pmatrix} = \operatorname{rank} \begin{pmatrix} DF \\ DH \end{pmatrix} = (m + 1 + 2\,q) - (q + p)$$

i.e. $p = \dim\ker(DF^\top, DH^\top)$. According to Condition M4 we define a linear surjective operator $\mathbf{Q} : R^{m+1+2\,q} \to R^{q+p}$ such that

$$\mathbf{Q}\begin{pmatrix} DF \\ DH \\ -DH \end{pmatrix} = 0 \qquad \text{as} \qquad \mathbf{Q} = \begin{pmatrix} Q & R^+ & R^- \\ O & I_q & I_q \end{pmatrix}$$

and this is equivalent to have $Q\,DF + R\,DH = 0$, with $R = R^+ - R^-$.

We briefly discuss the case when only the trivial multipliers exist. In this case the rows of DF and DH are together linearly independent (moreover, $p = 0$ and the matrix $\mathbf{Q}$ consists only of the lower part), and by applying the inverse function theorem the point is not optimal, so Condition M4 cannot hold.

The cone K is defined by $K = \{h : DF\,h \geqq 0,\ DH\,h = 0\}$. Then M4 is verified if, for no $h \neq 0$, $h \in K$, and for no triple of vectors $c = (c^{(1)}, c^{(2)}, c^{(3)}) \geqq 0$ of suitable dimensions, we have

$$\mathbf{Q}\begin{pmatrix} D^2F \\ D^2H \\ -D^2H \end{pmatrix}(h,h) = \mathbf{Q}\begin{pmatrix} c^{(1)} \\ c^{(2)} \\ c^{(3)} \end{pmatrix} \tag{5}$$

i.e.

$$(Q\,D^2F + R\,D^2H)(h,h) = Q\,c^{(1)} + R\,(c^{(2)} - c^{(3)}) \tag{6}$$

and

$$0 = c^{(2)} + c^{(3)} \tag{7}$$

Conditions (7) and $c^{(2)}, c^{(3)} \geqq 0$ imply $c^{(2)} = c^{(3)} = 0$, so we ask for

$$(Q\,D^2F + R\,D^2H)(h,h) \neq Q\,c \qquad \text{for any } c \geqq 0 \text{ and } h \neq 0,\ h \in K \tag{8}$$

In the second approach, in order to get more meaningful necessary conditions, we use the result that in a neighborhood of a feasible point x, the set $\{x : H(x) = 0\}$ is diffeomorphic to an open set in R^{n-q} if $\operatorname{rank} DH = q$. Note that we are just assuming the RCQ for the equalities. We implicitly take care of the equality constraints by considering a diffeomorphism $C : R^{n-q} \to R^n$ such that $H \circ C$ is identically zero. Therefore we have $D(H \circ C) = DH\,DC = 0$ and also $D^2(H \circ C) = 0$, i.e.

$$D^2(H \circ C)(h', h') = D^2H\,(DC\,h', DC\,h') + DH\,D^2C(h', h') = 0 \quad \forall h' \in R^{n-q} \tag{9}$$

Note that $DH\,DC = 0$ together with $\operatorname{rank} DH = q$ imply that the rows of DH span $\ker DC^\top$.

Then the maximization problem with inequality constraints only is reformulated for the functions

$$\mathcal{F} : R^{n-q} \to R^n \to R^{m+1}, \quad \mathcal{F} = F \circ C$$

for which $D\mathcal{F} = DF\,DC$. The optimality conditions are now defined for $\mathcal{F}$. The existence of $\pi \geq 0$ such that $\pi\,D\mathcal{F} = 0$ is equivalent to $\pi\,DF\,DC = 0$, i.e. $(\pi\,DF)^\top \in \ker DC^\top$, which, by the previous observation, is in turn equivalent to $\pi\,DF + \mu\,DH = 0$, for some μ (unrestricted in sign).

As for the second order conditions, let $p := (m+1) - \operatorname{rank} D\mathcal{F} = \dim\ker D\mathcal{F}^\top$ (clearly p is independent of the diffeomorphism), and $Q : R^{m+1} \to R^p$ surjective such that $Q\,D\mathcal{F} = 0$, i.e. $Q\,DF\,DC = 0$. Again this is equivalent to $(Q, R) : R^{m+1+q} \to R^p$ surjective such that

$$Q\,DF + R\,DH = 0 \tag{10}$$

The cone K has to be redefined on R^{n-q} and becomes $\mathcal{K} = \{h' \in R^{n-q} : DF\,DC\,h' \geqq 0\}$. Note that $DC\,\mathcal{K} = \{h \in R^n : DH\,h = 0,\ DF\,h \geq 0\}$. From (9) and (10) we derive that the evaluation of $Q\,D^2\mathcal{F}(h', h')$ for $h' \in \mathcal{K}$, is equivalent to the evaluation of $(Q\,D^2F + R\,D^2H)\,(h,h)$ for h such that $DF\,h \geq 0,\ DH\,h = 0$. Summarizing we have the following conditions if equality constraints are present:

Condition M1' *(first order necessity with equality constraints)*: *If x is M, and the equality constraints satisfy the RCQ, then there exists $\pi \in R^{m+1}$, $\pi \geq 0$, and $\mu \in R^q$, such that $\pi\, DF(x) + \mu\, DH(x) = 0$.* ∎

Condition M2' *(first order sufficiency with equality constraints)*: *If*

$$\ker \begin{pmatrix} DF(x) \\ DH(x) \end{pmatrix} = \{0\}$$

and there exists $\pi > 0$ and μ such that $\pi\, DF(x) + \mu\, DH(x) = 0$, then x is SM. ∎

Let $K(x) = \{h \in R^n : DF(x)\, h \geqq 0,\ DH(x)\, h = 0\}$ and $(Q(x), R(x)) : R^{m+1+q} \to R^p$ linear surjective such that $Q(x)\, DF(x) + R(x)\, DH(x) = 0$. An argument similar to the one used to prove Corollary P2 leads to

Corollary M2' *(first order sufficiency with equality constraints)*: *If $K(x) = \{0\}$ then x is SM.* ∎

Condition M3' *(second order necessity with equality constraints)*: *If x is M and the equality constraints satisfy the RCQ, then $(Q(x)\, D^2F(x) + R(x)\, D^2H(x))(h,h) \neq Q(x)\, c$ for any $h \in K(x)$, $c > 0$.* ∎

Condition M4' *(second order sufficiency with equality constraints))*: *If there exists $\pi \geq 0$ and μ such that $\pi\, DF(x) + \mu\, DH(x) = 0$ and*

$$(Q(x)\, D^2F(x) + R(x)\, D^2H(x))(h,h) \neq Q(x)\, c \quad \textit{for any } h \in K(x),\ h \neq 0,\ c \geqq 0$$

then x is SM. ∎

5. EXAMPLES

The following examples may give a good idea of the scope of the stated conditions. In all examples the point to which the conditions are applied is the origin and we drop in the notation the dependence on the point.

In the first example the optimality is detected by first order sufficiency conditions. Examples 2–4 show the application of second order conditions when none of the quoted constraint qualifications holds: in Examples 2 and 3 optimality is detected by Condition M4, and in Example 4 nonoptimality follows from Condition M3. Example 5 discusses the unavoidable gap between second order necessary and sufficient conditions, and Example 6 deals with the maximization of a generic function over a nonconvex set (the union of two cones). The last two examples consider equality constraints (nonregular in Example 7).

Example 1: Let $f_1(x_1, x_2) = -x_1$, $f_2(x_1, x_2) = x_2$, $f_3(x_1, x_2) = x_1^2 - x_2$ and $f_0(x_1, x_2)$ a generic function with $\partial f_0 / \partial x_1 > 0$. Then $K = \{0\}$ and optimality is verified by Corollary M2. Note that RCQ and MFCQ are not satisfied; KTCQ and ACQ are satisfied. ∎

Example 2: Let $f_1(x_1,x_2) = -x_1^3 - x_2$, $f_2(x_1,x_2) = x_1^4 + x_2$ and $f_0(x_1,x_2)$ a generic function with $\partial f_0/\partial x_1 = 0$ and $\partial^2 f_0/\partial x_1^2 < 0$. Then $K = \{h : h_2 = 0\}$ and

$$Q = \begin{pmatrix} 0 & 1 & 1 \\ 1 & \partial f_0/\partial x_2 & 0 \end{pmatrix}$$

so that

$$Q\, D^2F(h,h) = \begin{pmatrix} 0 \\ \partial^2 f_0/\partial x_1^2 \cdot h_1^2 \end{pmatrix} \neq Q\,c = \begin{pmatrix} c_2 + c_3 \\ c_1 + c_2\, \partial f_0/\partial x_2 \end{pmatrix}$$

and Condition M4 is satisfied and the origin is optimal. Note that RCQ, KTCQ, MFCQ and ACQ are not satisfied. ∎

Example 3: this example can be found in [3] and reported also in [5]. In [3] necessary and sufficient conditions are successfully applied to this example. Their application is more complex than the derivation below. The functions are the quadratic forms

$$f_0(x_1,x_2,x_3) = -2\,x_2\,x_3 - \frac{1}{2}\,x_1^2, \quad f_1(x_1,x_2,x_3) = -2\,x_1\,x_3 - \frac{1}{2}\,x_2^2, \quad f_2(x_1,x_2,x_3) = -2\,x_1\,x_2 - \frac{1}{2}\,x_3^2$$

Then $DF = O_{3\times 3}$, $K = R^3$, $Q = I_3$. Hence

$$Q\, D^2F(h,h) = \begin{pmatrix} -4\,h_2\,h_3 - h_1^2 \\ -4\,h_1\,h_3 - h_2^2 \\ -4\,h_1\,h_2 - h_3^2 \end{pmatrix}$$

If $Q\,D^2F(h,h) \geq 0$ for some $h \neq 0$, and wlog $h_1 \neq 0$, then $h_2\,h_3 < 0$, i.e. $h_2 \neq 0$ and $h_3 \neq 0$. Hence also $h_1\,h_2 < 0$ and $h_1\,h_3 < 0$. But the three strict inequalities are incompatible. Therefore Condition M4 is satisfied and the origin is optimal. None of the quoted CQ holds because the first derivatives vanish at the origin.

We want to show the role played by Q with respect to the vector π. Given any vector $\pi = \{\pi_1, \pi_2, \pi_3\} \geq 0$ (“any” because $DF = O_{3\times 3}$), the usual second order sufficiency condition calls for negative definiteness of the matrix

$$\pi\, D^2F = -\begin{pmatrix} \pi_1 & 2\,\pi_3 & 2\,\pi_2 \\ 2\,\pi_3 & \pi_2 & 2\,\pi_1 \\ 2\,\pi_2 & 2\pi_1 & \pi_3 \end{pmatrix}$$

An analysis of $\pi\,D^2F$ shows that there is always a positive eigenvalue $\lambda \geq \min \pi_i$ and so the usual second order sufficiency condition does not detect optimality. Apparently, requiring all components of $Q\,D^2F(h,h)$ to be nonpositive is stronger than requiring a positive combination of the components to be nonpositive. Note that the usual second order necessary condition (requiring negative semidefiniteness of $\pi\,D^2F$) cannot be applied due to the missing CQ. ∎

Example 4: Let $f_0(x_1,x_2) = x_1$, $f_1(x_1,x_2) = -x_1^4 + x_1^2 - x_2^2$. Then $K = \{h : h_1 \geq 0\}$, $\pi = Q = (0 \quad 1)$ and there exists $h \in K$, $h \neq 0$, such that $\pi\,D^2F(h,h) = 2\,(h_1^2 - h_2^2) > 0$. In this case nonoptimality is detected by condition M3. Note that $\pi_0 = 0$ and RCQ, KTCQ, MFCQ, ACQ are not satisfied. ∎

Example 5: this example (derived from [5] Example 4.2) supports the statement that these are the broadest conditions using first and second order derivatives. Let $f_0(x_1, x_2) = x_1 - x_2^2$ and $f_1(x_1, x_2) = -x_2^2$. Then $K = \{h : h_1 \geq 0\}$, $Q = \pi = (0, 1)$, $\pi\, D^2F(h, h) = -2\, h_2^2$. Condition M3 is verified but the point is not optimal. Note that if we change f_1 into $-x_1^4 - x_2^2$, the point is optimal (it is the only feasible point) and DF and D^2F are unchanged. Hence it is impossible to get a certificate of nonoptimality for the example based only on first and second order derivatives. None of the quoted CQ holds. ∎

Example 6: this example (derived from a similar example in [4]) considers the nonconvex feasible set defined by two linearly independent row vectors a_1 and a_2, $f_1(x_1, x_2) = (a_1\, x)\,(a_2\, x) \geq 0$, and a generic objective function $f_0(x_1, x_2)$. Being $K = \{h : Df_0\, h \geqq 0\}$ the condition $Df_0 \neq 0$ implies $Q = \pi = (\,0 \quad 1\,)$ and for $h \in K$, $c_1, c_2 > 0$ we have to check

$$\pi\, D^2F(h, h) = 2\,(a_1\, h)\,(a_2\, h) = \pi\, c = c_2$$

Since there exists $h \in K$ such that $(a_1\, h)\,(a_2\, h) > 0$ the necessary condition M3 cannot be satisfied and the origin is not optimal. If $Df_0 = 0$ then $K = R^2$, $Q = I_2$ and we have to check

$$Q\, D^2F(h, h) = \begin{pmatrix} D^2 f_0(h, h) \\ 2\,(a_1\, h)\,(a_2\, h) \end{pmatrix} = Q\, c = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}$$

If the Hessian $D^2 f_0$ is negative definite we have a proof of optimality by M4. However, the origin can be optimal even if $D^2 f_0$ is not negative definite. We have to consider the two cones $C^1 := \left\{h : D^2 f_0(h, h) > 0\right\}$ and $C^2 := \{h : (a_1\, h)\,(a_2\, h) > 0\}$. If $C^1 \cap C^2 \neq \emptyset$, then the origin cannot be optimal by M3 and if $\bar{C}^1 \cap \bar{C}^2 = \{0\}$ then it is optimal by M4. ∎

Example 7: this is Example 7.5 in [4] with objective function $f_0(x_1, x_2) = -x_1^2 - x_2^2 + a\, x_1\, x_2$, with a any real number. The inequality constraints are given by $f_1(x_1, x_2) = x_1$, $f_2(x_1, x_2) = x_2$, and the nonregular equality constraint by $h(x_1, x_2) = x_1\, x_2$. Then K is the nonnegative orthant,

$$DF = \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad DH = (\,0 \quad 0\,), \quad (\,Q \quad R\,) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

If $(Q\, D^2F + R\, D^2H)(h, h) = Q\, c$, i.e. $-h_1^2 - h_2^2 + a\, h_1\, h_2 = c_1$ and $h_1\, h_2 = 0$, then the first expression becomes $-h_1^2 - h_2^2 < 0$. Therefore Condition M4' is satisfied and the point is optimal. ∎

Example 8: this is Example 4.4 of [14] after a translation of $\hat{x} = (48, 36, 72)^\top$ to the origin.
$f_0(x_1, x_2, x_3) = -(x_1 + 48)^2 - (x_3 + 72)^2, \quad f_1(x_1, x_2, x_3) = x_1 + x_2 + x_3,$
$f_2(x_1, x_2, x_3) = (x_1 + 48)^2 + (x_2 - 64)^2 - 6400, \quad f_3(x_1, x_2, x_3) = (x_1 - 27)^2 + (x_2 + 36)^2 - 2025,$
and the (regular) equality constraint $h(x_1, x_2, x_3) = (x_1 + 48)^2 - 32(x_3 + 72)$. The RCQ, MFCQ are not satisfied (but KTCQ, ACQ, and a CQ introduced in [14] are). Optimality can be simply established by Corollary M2' since $K = \{0\}$. ∎

6. REFERENCES

[1] Abadie, J.M., "On the Kuhn-Tucker theorem", in *Nonlinear Programming*, J. Abadie ed., Wiley, New York, 1967, 21–67.

[2] Arrow, K.J., Hurwicz, L. and Uzawa, H., "Constraint qualifications in maximization problems", *Naval Research Logistics Quarterly*, **8** (1961), 175–191.

[3] Ben-Tal, A., "Second-order and related extremality conditions in nonlinear programming", *J. Opt. Theory and Appl.*, **31** (1980), 143–165.

[4] Bertsekas, D.P. and Ozdaglar, A.E., "Pseudonormality and a Lagrange multiplier theory for constrained optimization", *J. Opt. Theory and Appl.*, **114** (2002) 287–343.

[5] Cambini, R., "Second order-optimality conditions in multiobjective programming", *Optimization*, **44** (1998) 139–160.

[6] Guerra Vazquez, F. and Rückmann J.J.R. , "Extensions of the Kuhn -Tucker constraint qualification to generalized semi-infinite programming", *SIAM J. Optim.*, **15** (2005) 926–937.

[7] He, Y. and Sun, J., "A weaker constraint qualification and a second order necessary optimality condition for mathematical programming", http://www.bschool.nus.edu.sg/depart/ds/sunjiehomepage /downloads/ NewPapers/soc-s.pdf, (2005).

[8] Izmailov, A.F. and Solodov, M.V., "Complementarity constraint qualification via the theory of 2-regularity", *SIAM J. Optim.*, **13** (2002) 368–385.

[9] Jeyakumar, V., Lee, G.M. and Dinh, N., "New Sequential Lagrange multiplier conditions characterizing optimality without constraint qualification for convex programs", *SIAM J. Optim.*, **14** (2003) 534–547.

[10] Karush, W., *Minima of functions of several variables with inequalities as side constraints*, M.Sc. dissertation, Department of Mathematics, University of Chicago, Chicago, IL, USA, (1939).

[11] Kuhn, H.W. and Tucker, A.W., "Nonlinear programming", in *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, J. Neyman ed., University of California Press, Berkeley, CA, USA, (1951).

[12] Mangasarian, O.L. and Fromovitz, S., "The Fritz John necessary optimality conditions in the presence of equality and inequality constraints", *J. Math. Anal. Appl.*, **17** (1967) 37–47.

[13] Martinez, J.M. and Svaiter, B.F., "A practical optimality condition without constraint qualifications for nonlinear programming", *J. Opt. Theory and Appl.*, **118** (2003) 117–133.

[14] Schichl, H. and Neumaier, A., "Transposition theorems and qualification-free optimality conditions", http://www.mat.univie.ac.at/~neum/ms/trans.pdf, (2005).

[15] Staib, T., "On necessary and sufficient optimality conditions for multicriterial optimization problems", *Mathematical Methods of Operations Research (ZOR)*, **35** (1991) 231–248.

[16] Wan, Y.H., "On local Pareto optima", *J. Math. Economics*, **2** (1975) 35–42.

[17] Ye, J.J., "Constraint qualifications and necessary optimality conditions for optimization problems with variational inequality constraints", *SIAM J. Optim.*, **10** (2000) 943–962.