RAIRO-Oper. Res. 49 (2015) 451–472
DOI: 10.1051/ro/2014046

RAIRO Operations Research
www.rairo-ro.org

FRITZ JOHN TYPE OPTIMALITY AND DUALITY IN NONLINEAR PROGRAMMING UNDER WEAK PSEUDO-INVEXITY

Hachem Slimani[1] and Mohammed Said Radjef[2]

Abstract. In this paper, we use a generalized Fritz John condition to derive optimality conditions and duality results for a nonlinear programming with inequality constraints, under weak invexity with respect to different $(\eta_i)_i$ assumption. The equivalence between saddle points and optima, and a characterization of optimal solutions are established under suitable generalized invexity requirements. Moreover, we prove weak, strong, converse and strict duality results for a Mond-Weir type dual. It is shown in this study, with examples, that the introduced generalized Fritz John condition combining with the invexity with respect to different $(\eta_i)_i$ are especially easy in application and useful in the sense of sufficient optimality conditions and of characterization of solutions.

Keywords. Nonlinear programming, weak (FJ)-pseudo-invexity, generalized Fritz John condition, generalized Fritz John stationary point, optimality, duality, saddle point.

Mathematics Subject Classification. 26A51, 90C26, 90C30, 90C46.

Received February 24, 2012. Accepted August 1, 2014.

[1] LaMOS Research Unit, Computer Science Department, University of Bejaia, 06000 Bejaia, Algeria. haslimani@gmail.com

[2] LaMOS Research Unit, Operational Research Department, University of Bejaia, 06000 Bejaia, Algeria. radjefms@gmail.com

Article published by EDP Sciences

1. Introduction

In optimization theory, optimality conditions and duality results for differentiable nonlinear constrained problems are important theoretically as well as computationally. In the literature, most of the studies of such problems are established involving the classical Karush–Kuhn–Tucker and/or Fritz John conditions. The John criterion [17], known in the literature under the complete name Fritz John criterion, is in a sense more general than the Karush–Kuhn–Tucker one which is due to Karush [18] and Kuhn–Tucker [21]. From the Fritz John criterion, the Karush–Kuhn–Tucker one is obtained by adding an assumption which is to impose a suitable constraint-qualification [22] on the constraints of the problem. Thus, in case of necessary optimality criteria, the only restriction on a constrained program is that the constraints should satisfy certain qualification but for sufficient optimality criteria and duality results to hold, the objective and constraints functions are required to satisfy some convexity or generalized convexity requirements, see, for example, Bazaraa *et al.* [5] and Mangasarian [22].

Among the different classes of functions introduced as generalizations of the convexity, the concept of invexity has received more attention from the researchers, especially the specialists in optimization. The concept of invexity is introduced by Hanson [13] for the differentiable functions by generalizing the difference $(x - x_0)$ in the definition of convex function to any function $\eta(x, x_0)$. He proved that if, in a mathematical programming problem, instead of the convexity assumption, the objective and constraint functions are invex with respect to the same vector function η, then both the sufficiency of Karush-Kuhn-Tucker conditions and weak and strong Wolfe duality still hold. Further, Ben Israel and Mond [7] considered a class of functions called pre-invex and also showed that the class of invex functions is equivalent to the class of functions whose stationary points are global minima, see also Craven and Glover [9]. Hanson and Mond [14] introduced two other classes of functions called type I and type II functions for the scalar optimization problem, which were further generalized to pseudo-type I and quasi-type I by Rueda and Hanson [30] and sufficient optimality conditions are obtained involving these functions. Kaul and Kaur [19] showed that the Karush-Kuhn-Tucker (Fritz John) necessary conditions are sufficient for optimality under the hypotheses of the pseudo-invexity and the quasi-invexity (the invexity and the strict invexity) with respect to the same function η for the objective and constraint functions respectively. Martin [23] introduced a weaker invexity called Kuhn-Tucker invexity, or KT-invexity, which is necessary and sufficient for every Kuhn-Tucker stationary point to be a global minimizer in the classical mathematical programming problem. Further properties and applications of invexity for some more general problems were studied by Antczak [1–4], Bector *et al.* [6], Craven [8], Fulga and Preda [12], Jeyakumar and Mond [16], Kaul *et al.* [20], Mishra *et al.* [24,25], Osuna-Gomez *et al.* [28], Pini and Singh [29], Soleimani-damaneh and Sarabi [35], and others.

However, one major difficulty in this extension of convexity is that invex problems require the same function $\eta(x, x_0)$ for the objective and constraint functions. This requirement turns out to be a major restriction in applications. In reference [31], a constrained nonlinear programming is considered and KT-invex, weakly KT-pseudo-invex and type I problems with respect to different $(\eta_i)_i$ are defined (each function occurring in the studied problem is considered with respect to its own function η_i instead of the same function η). A new Kuhn-Tucker type necessary condition is introduced for nonlinear programming problems and duality results are obtained, for Wolfe and Mond-Weir type dual programs, under generalized invexity assumptions. In [34], the invexity with respect to different $(\eta_i)_i$ is used in the nondifferentiable case. Fritz John type necessary, Karush-Kuhn-Tucker type necessary and sufficient optimality conditions and duality results are obtained for nondifferentiable multiobjective programming (see also [32, 33]).

In parallel to all these developments and advances of the invexity and its extensions in theory, some applications in practice begin to take place. Recently, Dinuzzo *et al.* [10] have obtained some kernel function in Machine Learning which is not quasi-convex (and hence also neither convex nor pseudoconvex) but it is invex. Nickisch and Seeger [27] have studied a multiple kernel learning problem and have used the invexity to deal with the optimization which is nonconvex.

In this paper, we study Fritz John type optimality and duality for constrained nonlinear programming with inequality constraints. We introduce a generalized Fritz John condition which is necessary and sufficient for a feasible point to be an optimal solution under weak invexity with respect to different $(\eta_i)_i$. In particular, we obtain optimality conditions of Fritz John type that extend previous conditions of Kuhn-Tucker type presented in references [31–33], and generalize results obtained in the literature on this topic. Moreover, we establish the equivalence between saddle points and optima, and a characterization of solutions under suitable generalized invexity assumptions. The result characterizing the optimal solutions is proved under weaker hypotheses than the one given in [32, 33], and it allows to characterize optimal solutions which are not characterized by previously known results using the concept of Kuhn-Tucker stationary point [3, 23]. Furthermore, by using the introduced generalized Fritz John condition, we formulate a Mond-Weir type dual for which we prove several duality results. These latter results are more general than those obtained in references [31–33] because here we use the generalized Fritz John condition instead of the generalized Kuhn-Tucher condition introduced in reference [31]. By way of illustration, several examples are provided.

2. Preliminaries and definitions

Invex functions were introduced to optimization theory by Hanson [13], and called by Craven [8], as a very broad generalization of convex functions.

Definition 2.1. [13] Let D be a nonempty open set of $\mathbb{R}^n$ and $\eta : D \times D \to \mathbb{R}^n$ be a vector function. A function $f : D \to \mathbb{R}$ is said to be (def) at $x_0 \in D$ on D with respect to η, if the function f is differentiable at x_0 and for each $x \in D$, (cond) holds.

(i) def: invex,
cond:

$$f(x) - f(x_0) \geq [\nabla f(x_0)]^t \eta(x, x_0). \tag{2.1}$$

(ii) def: pseudo-invex,
cond:

$$[\nabla f(x_0)]^t \eta(x, x_0) \geq 0 \Rightarrow f(x) - f(x_0) \geq 0. \tag{2.2}$$

(iii) def: quasi-invex,
cond:

$$f(x) - f(x_0) \leq 0 \Rightarrow [\nabla f(x_0)]^t \eta(x, x_0) \leq 0. \tag{2.3}$$

If the inequality in (2.1) (resp. second (implied) inequality in (2.3)) is strict ($x \neq x_0$), we say that f is strictly invex (resp. strictly quasi-invex) at x_0 on D with respect to η. f is said to be (strictly) invex (resp. pseudo-invex or (strictly) quasi-invex) on D with respect to η, if f is (strictly) invex (resp. pseudo-invex or (strictly) quasi-invex) at each $x_0 \in D$ on D with respect to the same η.

Remark 2.2. When the function $\eta(x, x_0) = x - x_0$, the definition of (strict) invexity (resp. pseudo-invexity and quasi-invexity) reduces to the definition of (strict) convexity (resp. pseudo-convexity and quasi-convexity).

Craven and Glover [9] and Ben-Israel and Mond [7] stated that the class of invex functions are all those functions whose stationary points are global minima. Moreover, Ben-Israel and Mond [7] proved that the class of invex functions coincides with the one of pseudoinvex functions, and every function f, with $\nabla f \not\equiv 0$, is invex.

Proposition 2.3. [7] *Any differentiable function $f : D \subseteq \mathbb{R}^n \to \mathbb{R}$ at a point $x_0 \in D$, with $\nabla f(x_0) \neq 0$, is invex at x_0 on D with respect to $\eta(x, x_0) = [f(x) - f(x_0)]\frac{[\nabla f(x_0)]}{[\nabla f(x_0)]^t[\nabla f(x_0)]}$, $\forall\, x \in D$.*

Now, we give others η for which a given scalar function is pseudo-invex.

Proposition 2.4. *Any differentiable function $f : D \subseteq \mathbb{R}^n \to \mathbb{R}$ at a point $x_0 \in D$, with $\nabla f(x_0) \neq 0$, is pseudo-invex at x_0 on D with respect to $\eta(x, x_0) = [f(x) - f(x_0)][\nabla f(x_0)]$, $\forall\, x \in D$ or $\eta(x, x_0) = [f(x) - f(x_0)]t(x_0)$, $\forall\, x \in D$ where $t(x_0) \in \mathbb{R}^n$ with $t_i(x_0) = \begin{cases} 1, & \text{if } \frac{\partial f}{\partial x_i}(x_0) \geq 0, \\ -1, & \text{otherwise,} \end{cases}$ for all $i = 1, \ldots, n$.*

Example 2.5. The function $f : \mathbb{R}^2 \to \mathbb{R}$ defined by $f(x) = -x_1^2 - 2x_2$ is pseudo-invex on $\mathbb{R}^2$ with respect to $\eta(x, \tilde{x}) = (-x_1^2 - 2x_2 + \tilde{x}_1^2 + 2\tilde{x}_2)(-2\tilde{x}_1, -2)^t$. Furthermore, f is pseudo-invex at each $x_0 \in [\mathbb{R}_+ \times \mathbb{R}]$ on $\mathbb{R}^2$ with respect to $\eta_1(x, \tilde{x}) = (-x_1^2 - 2x_2 + \tilde{x}_1^2 + 2\tilde{x}_2)(-1, -1)^t$ and it is pseudo invex at each $x_0 \in [(\mathbb{R}_- \backslash \{0\}) \times \mathbb{R}]$ on $\mathbb{R}^2$ with respect to $\eta_2(x, \tilde{x}) = (-x_1^2 - 2x_2 + \tilde{x}_1^2 + 2\tilde{x}_2)(1, -1)^t$.

In Slimani and Radjef [31], a new concept of weak KT-pseudo-invexity is introduced and duality results have been obtained for a constrained nonlinear programming. Now, we define a concept of (semi strictly) weak pseudo-invexity for scalar functions given as follows.

Definition 2.6. Let D be a nonempty open set of $\mathbb{R}^n$ and $\eta : D \times D \to \mathbb{R}^n$ be a vector function. A function $f : D \to \mathbb{R}$ is said to be *weakly pseudo-invex* at $x_0 \in D$ on D with respect to η, if the function f is differentiable at x_0 and for each $x \in D$:

$$f(x) - f(x_0) < 0 \Rightarrow \exists\ \bar{x} \in D,\ [\nabla f(x_0)]^t \eta(\bar{x}, x_0) < 0. \tag{2.4}$$

f is said to be weakly pseudo-invex on D with respect to η, if f is weakly pseudo-invex at each $x_0 \in D$ on D with respect to the same η.
In the relation (2.4), if we have $f(x)-f(x_0) \leq 0$ $(x \neq x_0)$, instead of $f(x)-f(x_0) < 0$, we say that f is semi strictly weakly pseudo-invex at x_0 on D with respect to η.

Remark 2.7. Note that, in Definition 2.6, $\bar{x}$ depends on x and x_0, *i.e.* $\bar{x} = \bar{x}(x, x_0)$. As particular case, if $\bar{x} = x$, we obtain the pseudo-invexity (resp. the strict quasi-invexity for the second case) of a scalar function f, given in Definition 2.1.

If a function f is pseudo-invex at x_0 with respect to η, then it is weakly pseudo-invex at x_0 with respect to the same η (take $\bar{x} = x$). However, if f is weakly pseudo-invex at x_0 with respect to η, then f may not be pseudo-invex at x_0 with respect to the same η but it will be pseudo-invex at x_0 with respect to $\tilde{\eta}$ with $\tilde{\eta}(x, x_0) = \eta(\bar{x}(x, x_0), x_0)$, $\forall\ x \in D$. Note that also, we can use Proposition 2.4 to obtain a function $\hat{\eta}$ for which f to be pseudo-invex (and then weakly pseudo-invex) at x_0. Thus the classes of pseudo-invex functions and weakly pseudo-invex functions coincide.

Example 2.8. The function $f : \mathbb{R}^2 \to \mathbb{R}$ defined by $f(x) = -2x_1^3 - x_2$ is weakly pseudo-invex at $x_0 = (0, 0)$ on $\mathbb{R}^2$ with respect to $\eta(x, x_0) = (x - x_0) \in \mathbb{R}^2$ (take $\bar{x} = [(0, f(x_0) - f(x))^t + x_0] \in \mathbb{R}^2$). But f is not pseudo-invex at x_0 on $\mathbb{R}^2$ with respect to the same η because for $x = (1, -1)$, $f(x) - f(x_0) < 0$ and $[\nabla f(x_0)]^t \eta(x, x_0) > 0$. However, f is pseudo-invex at x_0 on $\mathbb{R}^2$ with respect to $\tilde{\eta}(x, x_0) = (0, f(x_0) - f(x))^t$. Note that if we use Proposition 2.4, we obtain $\hat{\eta}(x, x_0) = (f(x) - f(x_0))(1, -1)^t$ for which f is pseudo-invex (and then weakly pseudo-invex) at x_0 on $\mathbb{R}^2$.

Definition 2.9. [11] A function $f : D \to \mathbb{R}^N$ is a convexlike function if for any $x, y \in D$ and $0 \leq \lambda \leq 1$, there exists $z \in D$ such that

$$f(z) \leq \lambda f(x) + (1 - \lambda) f(y).$$

Consider the following constrained nonlinear programming problem (P):

$$(P) \qquad \begin{array}{c} \text{Minimize } f(x), \\ \text{subject to } \ g_j(x) \leq 0,\ j \in K = \{1, \ldots, k\}, \end{array}$$

where $f, g_j : D \to \mathbb{R},\ j \in K$, D is an open set of $\mathbb{R}^n$; $X = \{x \in D : g_j(x) \leq 0,\ j \in K\}$ is the set of all feasible solutions for (P). For $x_0 \in D$, we denote by $J(x_0)$ the set $\{j \in K : g_j(x_0) = 0\}$ and by $\tilde{J}(x_0)$ (resp. $\bar{J}(x_0)$) the set $\{j \in K : g_j(x_0) < 0$ (resp. $g_j(x_0) > 0)\}$. We have $J(x_0) \cup \tilde{J}(x_0) \cup \bar{J}(x_0) = K$ and if $x_0 \in X, \bar{J}(x_0) = \varnothing$. $J_0 = |J(x_0)|$ is the cardinal of the set $J(x_0)$, g_J is the semi-vector of g composed of the active constraints at the point x_0.

The concept of Kuhn-Tucker (Fritz John) stationary point for (P) is very used in the literature for establishing optimality conditions for the problem (P). It is defined as follows.

Definition 2.10. [22] A feasible point $x_0 \in X$ is said to be a Kuhn-Tucker (resp. Fritz John) stationary point for (P), if the functions f and g are differentiable at x_0 and there exists $\lambda \in \mathbb{R}_+^{J_0}$ (resp. $(\mu, \lambda) \in \mathbb{R}_+^{1+J_0},\ (\mu, \lambda) \neq 0$) such that:

$$\nabla f(x_0) + \sum_{j \in J(x_0)} \lambda_j \nabla g_j(x_0) = 0, \tag{2.5}$$

$$(\text{resp.} \quad \mu \nabla f(x_0) + \sum_{j \in J(x_0)} \lambda_j \nabla g_j(x_0) = 0). \tag{2.6}$$

The problem (P) is said to be HC-invex at $x_0 \in X$ if f and $g_j,\ j \in K$ are invex at x_0 (with respect to the same function η). Thus, if the problem (P) is HC-invex, then every Kuhn-Tucker stationary point is a minimizer of (P) [13]. Martin [23] remarked that the converse is not true in general, and he proposed a weaker notion, called KT-invexity, which assures that every Kuhn-Tucker stationary point is a minimizer of problem (P) if and only if problem (P) is KT-invex.

Definition 2.11. [23] Let $\eta : X \times X \to \mathbb{R}^n$ be a vector function. The problem (P) is said to be *KT-invex* on the feasible set X with respect to η, if the functions f and g are differentiable on X and for each $x, x_0 \in X$:

$$f(x) - f(x_0) \geq [\nabla f(x_0)]^t \eta(x, x_0), \tag{2.7}$$

$$-[\nabla g_j(x_0)]^t \eta(x, x_0) \geq 0,\ \forall\ j \in J(x_0). \tag{2.8}$$

The following result established by Martin [23] is considered as an optimality criterion for problem (P) and as a characterization of the KT-invexity notion with respect to η.

Theorem 2.12. [23] *Every Kuhn-Tucker stationary point of problem (P) is a global minimizer if and only if (P) is KT-invex on X with respect to η.*

Antczak [3] showed that the result in Theorem 2.12 remains true under a generalized KT-invexity called KT-$(0, r)$-invexity with respect to η.

Hayashi and Komiya [15] have proved the following alternative lemma which will be used to prove Fritz John type necessary optimality conditions and to establish a characterization of optimal solutions for (P).

Lemma 2.13. *Let S be a nonempty set in $\mathbb{R}^n$ and let $\psi : S \to \mathbb{R}^m$ be a convexlike function. Then either*

$$\psi(x) < 0 \textit{ has a solution } x \in S,$$

or

$$p^T \psi(x) \geq 0 \textit{ for all } x \in S, \textit{ for some } p \in \mathbb{R}^m_+, \ p \neq 0,$$

but both alternatives are never true (Here the symbol T denotes the transpose of matrix).

Now, before establishing optimality conditions for (P), we give the following simple propositions that we will use.

Proposition 2.14. *Let S be a nonempty subset of $\mathbb{R}^n$. If a function $\varphi : S \to]-\infty, 0]$ is strictly quasi-invex at $x_0 \in S$ on S with respect to $\theta : S \times S \to \mathbb{R}^n$ and $\varphi(x_0) = 0$, then $[\nabla \varphi(x_0)]^t \theta(x, x_0) < 0$, $\forall\ x \in S$.*

Proof. For $x \in S$, we have $\varphi(x) \leq 0 = \varphi(x_0)$, which by strict quasi-invexity of φ at x_0 on S with respect to θ implies $[\nabla \varphi(x_0)]^t \theta(x, x_0) < 0$. $\square$

Corollary 2.15. *Let $x_0 \in X$ be a feasible solution of (P). For each $j \in J(x_0)$, if g_j is strictly quasi-invex at x_0 on X with respect to $\theta_j : X \times X \to \mathbb{R}^n$, then $[\nabla g_j(x_0)]^t \theta_j(x, x_0) < 0$, $\forall\ x \in X$.*

Proposition 2.16. *Let x_0 be a feasible solution of (P). For each $j \in K$, if $\nabla g_j(x_0) \neq 0$ and the components of $\theta_j : X \times X \to \mathbb{R}^n$ are defined by $\theta^l_j(x, x_0) = \begin{cases} g_j(x) - \delta, & \textit{if } \frac{\partial g_j}{\partial x_l}(x_0) \geq 0, \\ -g_j(x) + \delta, & \textit{otherwise,} \end{cases}$ for all $l = 1, \ldots, n$ with $\delta \in \mathbb{R}$, $\delta > 0$, then $[\nabla g_j(x_0)]^t \theta_j(x, x_0) < 0$, $\forall\ x \in X$.*

Proof. We have $[\nabla g_j(x_0)]^t \theta_j(x, x_0) = \displaystyle\sum_{l=1}^{n} \frac{\partial g_j}{\partial x_l}(x_0) s^j_l(x_0)[g_j(x) - \delta] < 0$, $\forall\ x \in X$, $\forall\ j \in K$ with $s^j_l(x_0) = \begin{cases} 1, & \text{if } \frac{\partial g_j}{\partial x_l}(x_0) \geq 0, \\ -1, & \text{otherwise,} \end{cases}$ for all $l = 1, \ldots, n$. $\square$

3. Optimality conditions

In this section, we give Fritz John type necessary and sufficient optimality conditions for a feasible point to be an optimal solution of (P). For the sufficiency conditions, we use the weak invexity with respect to different functions η and $(\theta_j)_j$. Moreover, we prove that the equivalence between saddle points and optima is held under semi strictly weak pseudo-invexity.

To prove necessary conditions for the problem (P), we need to prove the following lemma.

Lemma 3.1. *Suppose that*

(i) x_0 *is a (local) optimal solution for (P);*
(ii) g_j *is continuous at* x_0 *for* $j \in \tilde{J}(x_0)$, *the functions* $f,\ g_j,\ j \in J(x_0)$ *are differentiable at* x_0 *and there exist vector functions* $\eta : X \times D \to \mathbb{R}^n$ *and* $\theta_j : X \times D \to \mathbb{R}^n,\ j \in J(x_0)$ *which satisfy at* x_0 *the following inequalities,*

$$[\nabla g_j(x_0)]^t \eta(x, x_0) \leq [\nabla g_j(x_0)]^t \theta_j(x, x_0),\ \forall\ x \in X,\ \forall\ j \in J(x_0), \tag{3.1}$$

Then the system of inequalities

$$[\nabla f(x_0)]^t \eta(x, x_0) < 0, \tag{3.2}$$

$$[\nabla g_j(x_0)]^t \theta_j(x, x_0) < 0,\ j \in J(x_0), \tag{3.3}$$

has no solution $x \in X$.

Proof. Let $x_0 \in X$ be a local optimal solution for (P) and suppose there exists $\tilde{x} \in X$ such that the inequalities (3.2)-(3.3) are true. Let $\varphi_f(x_0, \tilde{x}, \tau) = f(x_0 + \tau\eta(\tilde{x}, x_0)) - f(x_0)$. We observe that this function vanishes at $\tau = 0$ and $\lim_{\tau \to 0^+} \tau^{-1}[\varphi_f(x_0, \tilde{x}, \tau) - \varphi_f(x_0, \tilde{x}, 0)] = \lim_{\tau \to 0^+} \tau^{-1}[f(x_0 + \tau\eta(\tilde{x}, x_0)) - f(x_0)] = [\nabla f(x_0)]^t \eta(\tilde{x}, x_0) < 0$ using (3.2).
It follows that, $\varphi_f(x_0, \tilde{x}, \tau) < 0$ if τ is in some open interval $(0, \delta_f),\ \delta_f > 0$. Thus,

$$f(x_0 + \tau\eta(\tilde{x}, x_0)) < f(x_0),\ \tau \in (0, \delta_f).$$

Similarly, by using (3.1) with (3.3), we get

$$g_j(x_0 + \tau\eta(\tilde{x}, x_0)) < g_j(x_0) = 0,\ \tau \in (0, \delta_{g_j}),\ \forall\ j \in J(x_0),$$

where for all $j \in J(x_0),\ \delta_{g_j} > 0$.
Now, since for $j \in \tilde{J}(x_0),\ g_j(x_0) < 0$ and g_j is continuous at x_0, therefore, there exists $\delta_j > 0$ such that

$$g_j(x_0 + \tau\eta(\tilde{x}, x_0)) < 0,\ \tau \in (0, \delta_j),\ \forall\ j \in \tilde{J}(x_0).$$

Let $\delta_0 = \min\{\delta_f,\ \delta_{g_j},\ j \in J(x_0),\ \delta_j,\ j \in \tilde{J}(x_0)\}$. Then

$$(x_0 + \tau\eta(\tilde{x}, x_0)) \in N_{\delta_0}(x_0),\ \tau \in (0, \delta_0), \tag{3.4}$$

where $N_{\delta_0}(x_0)$ is a neighborhood of x_0. Now, for all $\tau \in (0, \delta_0)$ we have

$$f(x_0 + \tau\eta(\tilde{x}, x_0)) < f(x_0), \tag{3.5}$$

$$g_j(x_0 + \tau\eta(\tilde{x}, x_0)) < 0,\ j \in K. \tag{3.6}$$

By (3.4) and (3.6), we get $(x_0 + \tau\eta(\tilde{x}, x_0)) \in N_{\delta_0}(x_0) \cap X$, for all $\tau \in (0, \delta_0)$. Hence (3.5) is a contradiction to the assumption that x_0 is a (local) optimal solution for (P). Thus, there exists no $x \in X$ satisfying the system (3.2)-(3.3), and the lemma is proved. $\square$

In the next theorem, we obtain Fritz John type necessary optimality conditions with different functions η and $(\theta_j)_j$ associated to the objective and constraints functions of (P).

Theorem 3.2. *(Fritz John type necessary optimality conditions) Suppose that*

(i) x_0 *is a (local) optimal solution for (P);*
(ii) g_j *is continuous at* x_0 *for* $j \in \tilde{J}(x_0)$*, the functions* $f,\ g_j,\ j \in J(x_0)$ *are differentiable at* x_0 *and there exist vector functions* $\eta : X \times D \to \mathbb{R}^n$ *and* $\theta_j : X \times D \to \mathbb{R}^n,\ j \in J(x_0)$ *which satisfy at* x_0 *the inequalities (3.1);*
(iii) $H(x) = ([\nabla f(x_0)]^t \eta(x, x_0),\ [\nabla g_j(x_0)]^t \theta_j(x, x_0),\ j \in J(x_0)) \in \mathbb{R}^{1+J_0}$ *is a convexlike function of* x *on* X*.*

Then there exists $(\mu, \lambda) \in \mathbb{R}_+^{1+J_0},\ (\mu, \lambda) \neq 0$ *such that* (x_0, μ, λ) *satisfies the following generalized Fritz John condition*

$$\mu[\nabla f(x_0)]^t \eta(x, x_0) + \sum_{j \in J(x_0)} \lambda_j [\nabla g_j(x_0)]^t \theta_j(x, x_0) \geq 0,\ \forall\ x \in X. \tag{3.7}$$

Furthermore, if $g_j,\ j \in \tilde{J}(x_0)$ *are also differentiable at* x_0*, the condition (3.7) can be written in the following equivalent form, where* $\lambda = (\lambda_1, \ldots, \lambda_k) \in \mathbb{R}_+^k$*:*

$$\mu[\nabla f(x_0)]^t \eta(x, x_0) + \sum_{j=1}^{k} \lambda_j [\nabla g_j(x_0)]^t \theta_j(x, x_0) \geq 0,\ \forall\ x \in X, \tag{3.8}$$

$$\lambda^t g(x_0) = 0. \tag{3.9}$$

Proof. If the conditions (i) and (ii) are satisfied, then, by Lemma 3.1 the system (3.2)-(3.3) has no solution for $x \in X$. Since, by hypothesis (iii), $H(x) = ([\nabla f(x_0)]^t \eta(x, x_0),\ [\nabla g_j(x_0)]^t \theta_j(x, x_0),\ j \in J(x_0))$ is a convexlike function of x on X, therefore, by Lemma 2.13, there exists $(\mu, \lambda) \in \mathbb{R}_+^{1+J_0},\ (\mu, \lambda) \neq 0$ such that the relation (3.7) is satisfied.
The equivalent form of the necessary conditions is readily obtained by setting for all $j \in \tilde{J}(x_0) = K - J(x_0),\ \lambda_j = 0$ and θ_j any function. □

Now, using the generalized Fritz John condition (3.7), we establish sufficient conditions for a feasible point to be an optimal solution of (P) under weak invexity with respect to different η and $(\theta_j)_j$.

Theorem 3.3. *Let* $x_0 \in X$ *and suppose that:*

(i) f *is weakly pseudo-invex at* x_0 *on* X *with respect to* $\eta : X \times X \to \mathbb{R}^n$*;*
(ii) g_J *is differentiable at* x_0 *and for all* $j \in J(x_0)$*, there exists a function* $\theta_j : X \times X \to \mathbb{R}^n$ *such that* $[\nabla g_j(x_0)]^t \theta_j(x, x_0) < 0,\ \forall\ x \in X$*.*

If there exists a vector $(\mu, \lambda) \in \mathbb{R}_+^{1+J_0},\ (\mu, \lambda) \neq 0$ *such that the generalized Fritz John condition (3.7) is satisfied, then the point* x_0 *is an optimal solution of (P).*

Proof. Let us suppose that x_0 is not an optimal solution of (P). Then there exists a feasible point x such that $f(x) - f(x_0) < 0$.
Since f is weakly pseudo-invex at x_0 on X with respect to η, it follows that

$$\exists\ \bar{x} \in X,\ [\nabla f(x_0)]^t \eta(\bar{x}, x_0) < 0. \tag{3.10}$$

By hypothesis, we have

$$[\nabla g_j(x_0)]^t \theta_j(\bar{x}, x_0) < 0,\ \forall\ j \in J(x_0). \tag{3.11}$$

As $(\mu, \lambda) \geq 0,\ (\mu, \lambda) \neq 0$ and from (3.10) and (3.11), it follows that

$$\mu[\nabla f(x_0)]^t \eta(\bar{x}, x_0) + \sum_{j \in J(x_0)} \lambda_j [\nabla g_j(x_0)]^t \theta_j(\bar{x}, x_0) < 0,$$

which contradicts (3.7), and therefore, x_0 is an optimal solution of (P). □

Remark 3.4. According to Corollary 2.15, we can replace in the above theorem the hypothesis (ii) by $\forall\ j \in J(x_0)$, g_j is strictly quasi-invex at x_0 on X with respect to $\theta_j : X \times X \to \mathbb{R}^n$.

Although the classical Fritz John necessary optimality condition is more reduced than the generalized Fritz John necessary optimality condition, this latter, combining with the invexity with respect to different $(\eta_i)_i$, has its usefulness in the sufficient optimality conditions. For illustration, in the following example, we consider a feasible point x_0 which is not a Kuhn-Tucker stationary point of problem and hence all the sufficient optimality conditions using this concept are not applicable to conclude on its optimality. Furthermore, we show that there exists no function η for which the objective and constraint functions are both (generalized) invex. Thus, we can not also use the sufficient optimality conditions using the (generalized) invexity with respect to the same η (in particular for $\eta(x, x_0) = x - x_0$) and the classical Fritz John conditions. Therefore, we appeal to the invexity with respect to different $(\eta_i)_i$, the generalized Fritz John condition and by using Theorem 3.3, we conclude on optimality of the point x_0.

Example 3.5. We consider the following nonlinear programming problem

$$\begin{aligned} &Minimize\ f(x) = -x_1, \\ &subject\ to\ g_1(x) = x_1^3 - x_2 \leq 0, \\ &\qquad\qquad g_2(x) = x_2 \leq 0, \end{aligned} \tag{3.12}$$

where $f : \mathbb{R}^2 \to \mathbb{R}$ and $g = (g_1, g_2) : \mathbb{R}^2 \to \mathbb{R}^2$. The set of all feasible solutions of problem is $X = \{x = (x_1, x_2) \in \mathbb{R}^2 : x_1^3 - x_2 \leq 0 \text{ and } x_2 \leq 0\}$.
For this problem, we have $x_0 = (0, 0) \in X$ is not a Kuhn-Tucker stationary point of problem (3.12), because the condition (2.5) of Kuhn-Tucker at x_0 takes a form $\nabla f(x_0) + \lambda_1 \nabla g_1(x_0) + \lambda_2 \nabla g_2(x_0) = (-1, -\lambda_1 + \lambda_2) \neq (0, 0),\ \forall\ (\lambda_1, \lambda_2) \geq 0$. Thus, all the sufficient optimality conditions using this concept, for example from [1, 3–5, 13, 14, 19, 22, 23], are not applicable.

Furthermore, we have x_0 is a Fritz John stationary point for (P) with respect to $(\mu^0, \lambda^0) \in \mathbb{R}_+^{1+J_0}$, but it is not difficult to prove that there exists no a function $\eta : X \times X \to \mathbb{R}^2$ for which the functions g_1 and g_2 are both (strictly) (pseudo)-invex at x_0 (take $x = (-2, -1) \in X$). Also, the Lagrangian $L(., \mu^0, \lambda^0)$ is not strictly $B - (p, r)$-invex at x_0 with respect to η and b on X (where $b(x, x_0) > 0,\ x \neq x_0$) [1] because $\mu^0 = 0,\ L(x, \mu^0, \lambda^0) \not> L(x_0, \mu^0, \lambda^0),\ \forall\ x \in X$ and "$L(x, \mu^0, \lambda^0) \geq L(x_0, \mu^0, \lambda^0),\ \forall\ x \in X$" is not true. Hence, the sufficient optimality conditions using the (generalized) invexity with respect to the same η (in particular for $\eta(x, x_0) = x - x_0$) with the classical Fritz John conditions are not applicable, for example Theorem 25 of Antczak [1], Theorem 4.2.12 of Bazaraa *et al.* [5], page 187, Theorem 3.2 of Kaul and Kaur [19] and Theorem 7.2.3 of Mangasarian [22], page 96.
However, by using the invexity with respect to different $(\eta_i)_i$ and the generalized Fritz John condition (3.7), we obtain

- f is pseudo-invex at x_0 on D (and then on X) with respect to $\eta(x, x_0) = (x_1, -x_1)$ using Proposition 2.4;
- g is differentiable at x_0, g_1 and g_2 are active constraints at x_0 and by using Proposition 2.16, for $\theta_1(x, x_0) = (x_1^3 - x_2 - 1, -x_1^3 + x_2 + 1)$ and $\theta_2(x, x_0) = (x_2 - 1, x_2 - 1)$, we obtain that $[\nabla g_j(x_0)]^t \theta_j(x, x_0) < 0,\ \forall\ x \in X,\ \forall\ j \in J(x_0) = \{1, 2\}$.

The generalized Fritz John condition (3.7) at x_0 for $\mu = 1$ and $\lambda_1 = \lambda_2 = 0$ takes the form $\mu[\nabla f(x_0)]^t \eta(x, x_0) = -x_1 \geq 0,\ \forall\ x \in X$. It follows that, by Theorem 3.3, x_0 is an optimal solution for the given nonlinear programming problem.

As particular case of Theorem 3.3, if the functions θ_j are equal to η, $\forall\ j \in J(x_0)$ and by using the classical Fritz John condition, we obtain the following theorem.

Theorem 3.6. *Let $x_0 \in X$ and suppose that:*

(i) f is weakly pseudo-invex at x_0 on X with respect to $\eta : X \times X \to \mathbb{R}^n$;
(ii) g_J is differentiable at x_0 and $\forall\ j \in J(x_0)$, $[\nabla g_j(x_0)]^t \eta(x, x_0) < 0,\ \forall\ x \in X$.

If there exists a vector $(\mu, \lambda) \in \mathbb{R}_+^{1+J_0},\ (\mu, \lambda) \neq 0$ such that the Fritz John condition (2.6) is satisfied, then the point x_0 is an optimal solution of (P).

Proof. It suffices to multiply the relation (2.6) by $\eta(x, x_0)$ and use Theorem 3.3. □

Remark 3.7. Kaul and Kaur [19], Theorem 3.2, proved that the Fritz John condition (2.6) is sufficient for x_0 to be an optimal solution of (P), if the objective and active constraint functions, f and $g_j,\ j \in J(x_0)$, are invex and strictly invex, respectively, at x_0 on X with respect to the same η. It is shown, in Theorem 3.6 (see Remark 3.4), that the result is also true under weak hypothesis, when f is weakly pseudo-invex and $\forall\ j \in J(x_0),\ g_j$ is strictly quasi-invex at x_0 on X with respect to η.

Using Proposition 2.14 and proceeding in the same manner as in the proof of Theorem 3.3, we can prove the following result.

Theorem 3.8. *Let $x_0 \in X$ and suppose that f is weakly pseudo-invex at x_0 on X with respect to $\eta : X \times X \to \mathbb{R}^n$. If there exists a vector $(\mu, \lambda) \in \mathbb{R}_+^{1+J_0}$, $(\mu, \lambda) \neq 0$ such that the scalar function $\Psi(x) = \sum_{j \in J(x_0)} \lambda_j g_j(x)$ is strictly quasi-invex at x_0 on X with respect to $\theta : X \times X \to \mathbb{R}^n$ and*

$$\mu[\nabla f(x_0)]^t \eta(x, x_0) + \sum_{j \in J(x_0)} \lambda_j [\nabla g_j(x_0)]^t \theta(x, x_0) \geq 0, \ \forall \ x \in X, \tag{3.13}$$

then the point x_0 is an optimal solution of (P).

Remark 3.9. In the above theorem, the strict quasi-invexity assumption of the scalar function $\Psi(x) = \sum_{j \in J(x_0)} \lambda_j g_j(x)$ at x_0 on X with respect to θ can be replaced by the relation $\sum_{j \in J(x_0)} \lambda_j [\nabla g_j(x_0)]^t \theta(x, x_0) < 0, \ \forall \ x \in X$.

Remark 3.10. Note that we have not used any alternative theorem to prove the Fritz John type sufficient optimality conditions (Thms. 3.3, 3.6 and 3.8), unlike to the usual procedure used in the literature where alternative theorems (Gordan, Motzkin, etc.) are used to prove Fritz John sufficient optimality conditions for non-linear scalar and multiobjective programming problems, see for example Bazaraa *et al.* [5], Kaul and Kaur [19] and Mangasarian [22].

Now, we give sufficient optimality conditions for existence of optima and saddle points by using the generalized Fritz John conditions (3.8) and (3.9) and the semi strict weakly pseudo-invexity of the Lagrangian.

Definition 3.11. [22] The Lagrange function, or Lagrangian, associated with the constrained minimization problem (P) is the function $L : D \times \mathbb{R} \times \mathbb{R}^k \to \mathbb{R}$ defined by $L(x, \mu, \lambda) = \mu f(x) + \lambda^t g(x)$.

Theorem 3.12. *(Generalized Fritz John saddle point conditions) Let $x_0 \in X$. Moreover, we assume that x_0 satisfies the generalized Fritz John conditions (3.8) and (3.9) with respect to μ^0, λ^0 and $\eta : X \times X \to \mathbb{R}^n$, i.e.*

$$\mu^0[\nabla f(x_0)]^t \eta(x, x_0) + \sum_{j=1}^{k} \lambda_j^0 [\nabla g_j(x_0)]^t \eta(x, x_0) \geq 0, \ \forall \ x \in X, \tag{3.14}$$

$${\lambda^0}^t g(x_0) = 0, \tag{3.15}$$

and the Lagrangian $L(., \mu^0, \lambda^0)$ is semi strictly weakly pseudo-invex at x_0 on X with respect to the same η. Then, x_0 is an optimal solution of (P) and (x_0, μ^0, λ^0) is a saddle point of the Lagrangian; thus

$$L(x_0, \mu^0, \lambda) \leq L(x_0, \mu^0, \lambda^0) \leq L(x, \mu^0, \lambda^0), \ \forall \ x \in X, \ \forall \ \lambda \in \mathbb{R}_+^k. \tag{3.16}$$

Proof. Let us suppose that x_0 is not an optimal solution of (P). Then there exists a feasible point x such that $f(x) - f(x_0) < 0$.
It follows from here, by $(\mu^0, \lambda^0) \geq 0$, the definition of $J(x_0)$ and X, that

$$\mu^0 f(x) + {\lambda_J^0}^t g_J(x) \leq \mu^0 f(x_0) + {\lambda_J^0}^t g_J(x_0). \tag{3.17}$$

From (3.15), we get $\lambda_j^0 = 0,\ \forall\, j \in K - J(x_0)$, and the relation (3.17) becomes

$$\mu^0 f(x) + {\lambda^0}^t g(x) \leq \mu^0 f(x_0) + {\lambda^0}^t g(x_0). \tag{3.18}$$

From (3.18) and the semi strictly weak pseudo-invexity of $L(., \mu^0, \lambda^0)$ at x_0 on X with respect to η, we obtain

$$\exists\ \bar{x} \in X,\ [\nabla(\mu^0 f + {\lambda^0}^t g)(x_0)]^t \eta(\bar{x}, x_0) < 0,$$

i.e.

$$\exists\ \bar{x} \in X,\ \mu^0[\nabla f(x_0)]^t \eta(\bar{x}, x_0) + \sum_{j=1}^{k} \lambda_j^0 [\nabla g_j(x_0)]^t \eta(\bar{x}, x_0) < 0,$$

which contradicts (3.14), and therefore, x_0 is an optimal solution of (P).
For the saddle point of the Lagrangian, we proceed by contradiction. Suppose that there exists $x \in X$ $(x \neq x_0)$ such that $\mu^0 f(x_0) + {\lambda^0}^t g(x_0) > \mu^0 f(x) + {\lambda^0}^t g(x)$. By the semi strictly weak pseudo-invexity of $L(., \mu^0, \lambda^0)$ at x_0 on X with respect to η, we obtain

$$\exists\ \bar{x} \in X,\ \mu^0[\nabla f(x_0)]^t \eta(\bar{x}, x_0) + \sum_{j=1}^{k} \lambda_j^0 [\nabla g_j(x_0)]^t \eta(\bar{x}, x_0) < 0,$$

which contradicts (3.14), and therefore

$$\mu^0 f(x_0) + {\lambda^0}^t g(x_0) \leq \mu^0 f(x) + {\lambda^0}^t g(x),\ \forall\ x \in X. \tag{3.19}$$

From the feasibility of x_0 and the relation (3.15), we obtain

$$\mu^0 f(x_0) + \lambda^t g(x_0) \leq \mu^0 f(x_0) + {\lambda^0}^t g(x_0),\ \forall\ \lambda \in \mathbb{R}_+^k. \tag{3.20}$$

Using (3.19) and (3.20), we obtain, by the definition of Lagrangian, that the relation (3.16) is satisfied. $\square$

It is known that, for the constrained mathematical programming problem (P), the equivalence of saddle points of the Lagrangian and minima is held under the convexity assumption (and constraint qualification). Antczak [1] showed that the result is still held for optimization problems with $B-(p,r)$-invex functions. In the following theorem, we prove that such equivalence between saddle points and minima remains true under semi strictly weak pseudo-invexity.

Theorem 3.13. *(Equivalence of saddle points and minima) For problem (P), we assume that the Lagrangian is semi strictly weakly pseudo-invex at x_0 on X with respect to $\eta : X \times X \to \mathbb{R}^n$ and for all $j \in K$, $[\nabla g_j(x_0)]^t \eta(x, x_0) \geq 0$, $\forall\ x \in X$. Then, x_0 is an optimal solution of (P) if and only if there exists $(\mu^0, \lambda^0) \in \mathbb{R}_+^{1+k}$ such that (x_0, μ^0, λ^0) satisfies the saddle point condition (3.16).*

Proof. (Necessary condition) Suppose that x_0 is an optimal solution of (P). Then x_0 is a Fritz John stationary point, *i.e.* there exists $(\mu^0, \lambda^0) \in \mathbb{R}_+^{1+J_0}$, $(\mu^0, \lambda^0) \neq 0$ such that

$$\mu^0 \nabla f(x_0) + \sum_{j=1}^{k} \lambda_j^0 \nabla g_j(x_0) = 0, \tag{3.21}$$

$${\lambda^0}^t g(x_0) = 0. \tag{3.22}$$

Multiplying the relation (3.21) by $\eta(x, x_0)$, we obtain that x_0 satisfies the generalized Fritz John conditions (3.8) and (3.9) with respect to μ^0, λ^0 and η. From Theorem 3.12, it follows that the saddle point condition (3.16) is satisfied at (x_0, μ^0, λ^0).
(Sufficient condition) Suppose that there exists $(\mu^0, \lambda^0) \in \mathbb{R}_+^{1+k}$ such that (x_0, μ^0, λ^0) is a saddle point of the Lagrangian of (P). From the inequality $L(x_0, \mu^0, \lambda) \leq L(x_0, \mu^0, \lambda^0)$, which holds for any $\lambda \in \mathbb{R}_+^k$, we have

$$\lambda^t g(x_0) \leq {\lambda^0}^t g(x_0). \tag{3.23}$$

If we put $\lambda = 0$ in (3.23), we obtain ${\lambda^0}^t g(x_0) \geq 0$, and since also $x_0 \in X$, hence ${\lambda^0}^t g(x_0) = 0$. Let x be any feasible point for (P), then ${\lambda^0}^t g(x) \leq 0$. Now, by using $L(x_0, \mu^0, \lambda^0) \leq L(x, \mu^0, \lambda^0)$, we conclude that the inequality

$$\mu^0 f(x_0) \leq \mu^0 f(x_0) + {\lambda^0}^t g(x_0) \leq \mu^0 f(x) + {\lambda^0}^t g(x) \leq \mu^0 f(x),$$

holds for all $x \in X$, *i.e.*

$$\mu^0 f(x_0) \leq \mu^0 f(x),\ \forall\ x \in X. \tag{3.24}$$

Now, suppose that $\mu^0 = 0$. Then we have $\mu^0 f(x) + {\lambda^0}^t g(x) \leq \mu^0 f(x_0) + {\lambda^0}^t g(x_0)$, $\forall\ x \in X$ $(x \neq x_0)$, and by semi strictly weak pseudo-invexity of Lagrangian at x_0 on X with respect to η, we obtain

$$\exists\ \bar{x} \in X,\ [\nabla(\mu^0 f + {\lambda^0}^t g)(x_0)]^t \eta(\bar{x}, x_0) < 0,$$

i.e.

$$\exists\ \bar{x} \in X,\ \sum_{j=1}^{k} \lambda_j^0 [\nabla g_j(x_0)]^t \eta(\bar{x}, x_0) < 0,$$

which contradicts the hypotheses of theorem. Thus, $\mu^0 > 0$ and hence, by (3.24), we obtain $f(x_0) \leq f(x)$, $\forall\ x \in X$. It follows that x_0 is an optimal solution of (P). □

4. Characterization of solutions

In Example 3.5, we have seen that there exist optimal solutions which, on the one hand, are not Kuhn-Tucker stationary points and, on the other hand, even if they are Fritz John stationary points some generalized convexities do not allow to conclude on their optimality. In order to characterize such optimal solutions, we need to define a new wide class of stationary points with an appropriate generalized invexity. Thus, using the generalized Fritz John condition (3.7), we define a new class of Fritz John type stationary points for (P) and we establish a new characterization of solutions under suitable generalized invexity requirement.

Definition 4.1. Let x_0 be a feasible point of (P) and $\eta : X \times X \to \mathbb{R}^n,\ \theta_j : X \times X \to \mathbb{R}^n,\ j \in J(x_0)$ be vector functions. x_0 is said to be a *generalized Fritz John stationary point* with respect to η and $(\theta_j)_{j\in J(x_0)}$, if the functions f and g are differentiable at x_0 and there exists a vector $(\mu,\lambda) \in \mathbb{R}_+^{1+J_0},\ (\mu,\lambda) \neq 0$ such that $(x_0,\mu,\lambda,\eta,(\theta_j)_{j\in J(x_0)})$ satisfies the generalized Fritz John condition (3.7) of Theorem 3.3.

Remark 4.2. The concept of generalized Kuhn-Tucker stationary point can be defined by setting $\mu = 1$ and $\lambda \in \mathbb{R}_+^{J_0}$ in Definition 4.1.

Martin [23] has characterized the optimal solutions for (P) associated with the Kuhn-Tucker stationary points by using the concept of KT-invexity with respect to η. Antczak [3] showed that this characterization is still true under KT-$(0,r)$-invexity with respect to η. Now we characterize the optimal solutions for (P) by using the concept of generalized Fritz John stationary point and new kind of invex functions which we define as follows.

Definition 4.3. Let $x_0 \in D$ and $\eta : X \times D \to \mathbb{R}^n,\ \theta_j : X \times D \to \mathbb{R}^n,\ j \in\ J(x_0)$ be vector functions. The problem (P) is said to be *weakly Fritz John-pseudo-invex (or weakly FJ-pseudo-invex)* at x_0 on X with respect to η and $(\theta_j)_{j\in J(x_0)}$, if the functions f and g are differentiable at x_0 and for each $x \in X$:

$$f(x) - f(x_0) < 0 \Rightarrow \exists\ \bar{x} \in X, \begin{cases} [\nabla f(x_0)]^t \eta(\bar{x},x_0) < 0, \\ [\nabla g_j(x_0)]^t\, \theta_j(\bar{x},x_0) < 0,\ \forall\ j \in J(x_0). \end{cases} \tag{4.1}$$

If $\bar{x} = x$, in the relation (4.1), we say that (P) is FJ-pseudo-invex at x_0 on X with respect to η and $(\theta_j)_{j\in J(x_0)}$. The problem (P) is said to be (weakly) FJ-pseudo-invex on D with respect to η and $(\theta_j)_j$, if it is (weakly) FJ-pseudo-invex at each $x_0 \in D$ on X with respect to the same η and $(\theta_j)_{j\in J(x_0)}$. In the relation (4.1), if we have $f(x) - f(x_0) \leq 0$ $(x \neq x_0)$, instead of $f(x) - f(x_0) < 0$, we say that (P) is semi strictly (weakly) FJ-pseudo-invex at x_0 on X with respect to η and $(\theta_j)_{j\in J(x_0)}$.

The following result is proved under weaker hypotheses than Theorem 2.4.6 given in [32,33]. The result remains true by using the concept of convexlikeness instead of the concepts of invexity and preinvexity. Thus, to prove the result we use the alternative Lemma 2.13 of Hayashi and Komiya [15] instead of the one given by Weir and Mond [36], Theorem 2.1.

Theorem 4.4. *Suppose that the functions f and g are differentiable on X and let $\eta : X \times X \to \mathbb{R}^n$ and $\theta_j : X \times X \to \mathbb{R}^n,\ j \in K$ be functions such that for all $x_0 \in X,\ H(x, x_0) = ([\nabla f(x_0)]^t \eta(x, x_0), [\nabla g_j(x_0)]^t \theta_j(x, x_0),\ j \in J(x_0)) \in \mathbb{R}^{1+J_0}$ is a convexlike function of x on X. Then, every generalized Fritz John stationary point with respect to η and $(\theta_j)_j$ of problem (P) is a global minimizer if and only if (P) is weakly FJ-pseudo-invex on X with respect to η and $(\theta_j)_j$.*

Proof. **(1)** (Sufficient condition) Let $x_0 \in X$ be a generalized Fritz John stationary point with respect to η and $(\theta_j)_{j\in J(x_0)}$ for (P). If (P) is weakly FJ-pseudo-invex at x_0 on X with respect to η and $(\theta_j)_{j\in J(x_0)}$, then, in the same manner as in Theorem 3.3, we obtain that x_0 is a global minimizer of (P).
(2) (Necessary condition) For the converse, suppose that every generalized Fritz John stationary point with respect to η and $(\theta_j)_j$ of problem (P) is a global minimizer.
Let us suppose that there exist two feasible points $\tilde{x}$ and x_0 such that

$$f(\tilde{x}) - f(x_0) < 0. \tag{4.2}$$

This means that x_0 is not a global minimizer of (P), and by using the initial hypothesis, x_0 is not a generalized Fritz John stationary point with respect to η and $(\theta_j)_{j\in J(x_0)}$ for (P), *i.e.*

$$\left(\mu[\nabla f(x_0)]^t \eta(x, x_0) + \sum_{j\in J(x_0)} \lambda_j [\nabla g_j(x_0)]^t \theta_j(x, x_0) \geq 0,\ \forall\, x \in X.\right)$$

is not satisfied for all $(\mu, \lambda) \in \mathbb{R}_+^{1+J_0},\ (\mu, \lambda) \neq 0$. Therefore, by Lemma 2.13, the system

$$\begin{cases} [\nabla f(x_0)]^t \eta(x, x_0) < 0, \\ [\nabla g_j(x_0)]^t\, \theta_j(x, x_0) < 0, \quad \forall\, j \in J(x_0). \end{cases}$$

has a solution $x = \bar{x} \in X$. In consequence, (P) is weakly FJ-pseudo-invex on X with respect to η and $(\theta_j)_j$. $\square$

Remark 4.5. Note that the hypothesis "for all $x_0 \in X,\ H(x, x_0) = ([\nabla f(x_0)]^t \eta(x, x_0), [\nabla g_j(x_0)]^t \theta_j(x, x_0),\ j \in J(x_0)) \in \mathbb{R}^{1+J_0}$ is a convexlike function of x on X" is needed just to prove the necessary optimality condition of Theorem 4.4.

In the following examples, we show that there exist optimal solutions of (P) which are not characterized neither by Theorem 2.12, established by Martin [23], nor by Theorem 5 of Antczak [3], but they are characterized by Theorem 4.4.

Example 4.6. We reconsider the problem given in Example 3.5.

$$\begin{aligned} &\text{Minimize } f(x) = -x_1, \\ &\text{subject to } \; g_1(x) = x_1^3 - x_2 \leq 0, \\ &\qquad\qquad\; g_2(x) = x_2 \leq 0, \end{aligned} \tag{4.3}$$

where $f : \mathbb{R}^2 \to \mathbb{R}$ and $g = (g_1, g_2) : \mathbb{R}^2 \to \mathbb{R}^2$. The set of all feasible solutions of problem is $X = \{x = (x_1, x_2) \in \mathbb{R}^2 : x_1^3 - x_2 \leq 0 \text{ and } x_2 \leq 0\}$.
We have $x_0 = (0, 0) \in X$ is not a Kuhn-Tucker stationary point of problem (4.3). Thus, the point x_0 does not belong to the set of optimal solutions characterized by Theorem 2.12 (Thm. 2.1 in Martin [23]) or by Theorem 5 of Antczak [3], even if the problem (4.3) is KT-invex and KT-$(0, r)$-invex on X with respect to $\hat{\eta}(x, \tilde{x}) = (x_1 - \tilde{x}_1, 0)$.
However, the problem (4.3) is weakly FJ-pseudo-invex on X with respect to $\eta(x, \tilde{x}) = (\eta'(x, \tilde{x}), 0)$, $\theta_1(x, \tilde{x}) = (-1, x_2\tilde{x}_2)$ and $\theta_2(x, \tilde{x}) = (0, \theta_2'(x, \tilde{x}))$ such that η' (resp. θ_2') can be any positive (resp. negative) function on $X \times X$ (take $\bar{x}(x, \tilde{x}) = (\sqrt[3]{a}, a) \in X$, with $a \in \,]-\infty, 0[$). Furthermore, x_0 is a generalized Fritz John stationary point with respect to η and $(\theta_j)_{j=1,2}$ (take $\mu = 0$, $\lambda_1 = 1$ and $\lambda_2 = 0$), it follows that, by using the sufficient condition of Theorem 4.4, x_0 is an optimal solution of problem (4.3).

Example 4.7. We consider the following nonlinear programming problem

$$\begin{aligned} &\text{Minimize } f(x) = -x^4, \\ &\text{subject to } \; g(x) = (x-2)^3 \leq 0, \end{aligned} \tag{4.4}$$

where $f, g :]0, +\infty[\to \mathbb{R}$. The set of feasible solutions of problem is $X = \,]0, 2]$.
We have $x_0 = 2 \in X$ is not a Kuhn-Tucker stationary point of problem (4.4), because the condition of Kuhn-Tucker at x_0 takes a form $\nabla f(x_0) + \lambda \nabla g(x_0) = -32 < 0, \forall\ \lambda \geq 0$. Thus, the point x_0 does not belong to the set of optimal solutions characterized by Theorem 2.12 (Thm. 2.1 in Martin [23]) or by Theorem 5 of Antczak [3], even if the problem (4.4) is KT-invex and KT-$(0, r)$-invex at x_0 on X with respect to $\hat{\eta}(x, \tilde{x}) = \frac{1}{32}(x^4 - 16)$.
However, the problem (4.4) is weakly FJ-pseudo-invex on X with respect to $\eta(x, \tilde{x})$ and $\theta(x, \tilde{x})$ such that η (resp. θ) can be any positive (resp. negative) function on $X \times X$ (take $\bar{x}(x, \tilde{x}) = \frac{x+\tilde{x}}{2} \in X$). Furthermore, x_0 is a generalized Fritz John stationary point with respect to η and θ (take $\mu = 0$ and $\lambda = 1$), it follows that, by using the sufficient condition of Theorem 4.4, x_0 is an optimal solution of problem (4.4).

5. Mond-Weir type duality

In relation to (P) and using the generalized Fritz John condition (3.7), we formulate the following dual problem which is in the format of Mond-Weir [26].

$$(MWD) \quad \text{Maximize } f(y),$$

subject to

$$\mu[\nabla f(y)]^t\eta(x,y) + \sum_{j\in J(y)} \lambda_j[\nabla g_j(y)]^t\theta_j(x,y) \geq 0,\ \forall\ x\in X,$$

$$y\in D,\quad (\mu,\lambda)\in \mathbb{R}_+^{1+|J(y)|},\quad (\mu,\lambda)\neq 0,$$
$$\eta : X\times D\to\mathbb{R}^n,\quad \theta_j : X\times D\to\mathbb{R}^n,\ \forall\ j\in J(y).$$

Let $Y = \{(y,\mu,\lambda,\eta,(\theta_j)_{j\in J(y)}) : \mu[\nabla f(y)]^t\eta(x,y) + \sum_{j\in J(y)}\lambda_j[\nabla g_j(y)]^t\theta_j(x,y) \geq 0, \forall\, x\in X; y\in D,\ (\mu,\lambda)\in\mathbb{R}_+^{1+|J(y)|}, (\mu,\lambda)\neq 0; \eta : X\times D\to\mathbb{R}^n, \theta_j : X\times D\to\mathbb{R}^n, \forall\ j\in J(y)\}$ be the set of all feasible solutions of problem (MWD). We denote by Pr_DY the projection of the set Y on D, that is, by definition $Pr_DY = \{y\in D : (y,\mu,\lambda,\eta,(\theta_j)_{j\in J(y)})\in Y\}$.

In what follows, we establish some duality results between (P) and (MWD) by using the concept of weak FJ-pseudo-invexity with respect to different η and $(\theta_j)_j$. Note that, we proceed in the same way as in [31] to prove the following results.

Theorem 5.1. *(Weak duality) Suppose that*

(i) $x\in X$*;*
(ii) $(y,\mu,\lambda,\eta,(\theta_j)_{j\in J(y)})\in Y$*;*
(iii)] the problem (P) is weakly FJ-pseudo-invex at y on X with respect to η and $(\theta_j)_{j\in J(y)}$.

Then $f(x)\not< f(y)$*.*

Proof. We proceed by contradiction. Suppose that $f(x) < f(y)$.
Since (P) is weakly FJ-pseudo-invex at y on X with respect to η and $(\theta_j)_{j\in J(y)}$, it follows that

$$\exists\ \bar{x}\in X,\ \begin{cases} [\nabla f(y)]^t\,\eta(\bar{x},y) < 0, \\ [\nabla g_j(y)]^t\,\theta_j(\bar{x},y) < 0,\ \forall\ j\in J(y). \end{cases} \tag{5.1}$$

As $(\mu,\lambda)\geq 0,\ (\mu,\lambda)\neq 0$ and from (5.1), we obtain

$$\mu[\nabla f(y)]^t\eta(\bar{x},y) + \sum_{j\in J(y)}\lambda_j[\nabla g_j(y)]^t\theta_j(\bar{x},y) < 0,$$

which contradicts (ii), and the conclusion follows. □

Now, we establish the following strong duality result between (P) and (MWD) without using any constraint qualification.

Theorem 5.2. *(Strong duality) Suppose that*

(i) x_0 *is a (local) optimal solution for (P);*
(ii) g_j *is continuous at* x_0 *for* $j\in\tilde{J}(x_0)$*, the functions* $f,\ g_j,\ j\in J(x_0)$ *are differentiable at* x_0 *and there exist vector functions* $\eta : X\times D\to\mathbb{R}^n$ *and* $\theta_j : X\times D\to\mathbb{R}^n,\ j\in J(x_0)$ *which satisfy at* x_0 *the inequalities (3.1);*

(iii) $H(x) = ([\nabla f(x_0)]^t \eta(x, x_0),\ [\nabla g_j(x_0)]^t \theta_j(x, x_0),\ j \in J(x_0)) \in \mathbb{R}^{1+J_0}$ *is a convexlike function of* x *on* X*.*

Then there exists $(\mu, \lambda) \in \mathbb{R}_+^{1+J_0}$, $(\mu, \lambda) \neq 0$ *such that* $(x_0, \mu, \lambda, \eta, (\theta_j)_{j\in J(x_0)}) \in Y$ *and the objective functions of (P) and (MWD) have the same values at* x_0 *and* $(x_0, \mu, \lambda, \eta, (\theta_j)_{j\in J(x_0)})$*, respectively. If, further, the problem (P) is weakly FJ-pseudo-invex at any* $\bar{y} \in Pr_D Y$ *on* X *with respect to* $\bar{\eta}$ *and* $(\bar{\theta}_j)_{j\in J(\bar{y})}$ *(with* $(\bar{y}, \bar{\mu}, \bar{\lambda}, \bar{\eta}, (\bar{\theta}_j)_{j\in J(\bar{y})}) \in Y$*), then* $(x_0, \mu, \lambda, \eta, (\theta_j)_{j\in J(x_0)}) \in Y$ *is an optimal solution of (MWD).*

Proof. By Theorem 3.2, there exist a vector $(\mu, \lambda) \in \mathbb{R}_+^{1+J_0}$ with $(\mu, \lambda) \neq 0$ such that

$$\mu[\nabla f(x_0)]^t \eta(x, x_0) + \sum_{j\in J(x_0)} \lambda_j [\nabla g_j(x_0)]^t \theta_j(x, x_0) \geq 0,\ \forall\ x \in X.$$

It follows that $(x_0, \mu, \lambda, \eta, (\theta_j)_{j\in J(x_0)}) \in Y$.
Trivially, the objective function values of (P) and (MWD) are equal.
Next, suppose that $(x_0, \mu, \lambda, \eta, (\theta_j)_{j\in J(x_0)}) \in Y$ is not an optimal solution of (MWD). Then there exists $(y^*, \mu^*, \lambda^*, \eta^*, (\theta_j^*)_{j\in J(y^*)}) \in Y$ such that $f(x_0) < f(y^*)$, which violates the weak duality Theorem 5.1. Hence $(x_0, \mu, \lambda, \eta, (\theta_j)_{j\in J(x_0)}) \in Y$ is indeed an optimal solution of (MWD). □

Following the same lines as in Theorem 3.3, we prove the converse duality theorem to the problems (P) and (MWD) under weak FJ-pseudo-invexity with respect to η and $(\theta_j)_j$.

Theorem 5.3. *(Converse duality) Suppose that*

(i) $(y, \mu, \lambda, \eta, (\theta_j)_{j\in J(y)}) \in Y$*;*
(ii) $y \in X$*;*
(iii) the problem (P) is weakly FJ-pseudo-invex at y *on* X *with respect to* η *and* $(\theta_j)_{j\in J(y)}$*.*

Then y *is an optimal solution for (P).*

For a strict duality to hold between (P) and (MWD), we need to use the concept of semi strictly weakly FJ-pseudo-invex with respect to η and $(\theta_j)_j$.

Theorem 5.4. *(Strict duality) Suppose that*

(i) $x \in X$ *and* $(y, \mu, \lambda, \eta, (\theta_j)_{j\in J(y)}) \in Y$ *such that* $f(x) = f(y)$*;*
(ii) the problem (P) is semi strictly weakly FJ-pseudo-invex at y *on* X *with respect to* η *and* $(\theta_j)_{j\in J(y)}$*.*

Then $x = y$*.*

Proof. We proceed by contradiction. Suppose that $x \neq y$.
By hypothesis we have $f(x) = f(y)$ and since (P) is semi strictly weakly FJ-pseudo-invex at y on X with respect to η and $(\theta_j)_{j\in J(y)}$, it follows that

$$\exists\ \bar{x} \in X,\ \begin{cases} [\nabla f(y)]^t\, \eta(\bar{x}, y) < 0, \\ [\nabla g_j(y)]^t\, \theta_j(\bar{x}, y) < 0,\ \forall\ j \in J(y). \end{cases} \tag{5.2}$$

As $(\mu, \lambda) \geq 0,\ (\mu, \lambda) \neq 0$ and from (5.2), we obtain

$$\mu[\nabla f(y)]^t \eta(\bar{x}, y) + \sum_{j \in J(y)} \lambda_j [\nabla g_j(y)]^t \theta_j(\bar{x}, y) < 0,$$

which contradicts the fact that $(y, \mu, \lambda, \eta, (\theta_j)_{j \in J(y)}) \in Y$. Hence $x = y$. □

6. Conclusion

In this paper, we have considered new concepts of (semi strictly) weak pseudo-invexity and weak FJ-pseudo-invexity to study Fritz John type optimality and duality for nonlinear programming with inequality constraints. To realize this study, we have introduced a generalized Fritz John condition which is both necessary and sufficient for a feasible point to be an optimal solution under weak invexity with respect to different η and $(\theta_j)_j$. The sufficient optimality conditions are obtained without using of any alternative theorem unlike the usual procedure. We have established simple propositions which helped us to construct easily these different functions (η and $(\theta_j)_j$) to verify the optimality of a feasible point (Ex. 3.5). Besides, we have shown that the equivalence between saddle points and minima remains true under semi strictly weak pseudo-invexity. Moreover, a new concept of Fritz John type stationary point is used to establish a characterization of solutions under suitable generalized invexity assumption. This allows to characterize optimal solutions which are not characterized by previously known results using the concept of Kuhn-Tucker stationary point. Furthermore, a Mond-Weir type dual is formulated and weak, strong, converse and strict duality results are proved.

In this contribution, it is shown by examples that the introduced generalized Fritz John condition combining with the invexity with respect to different $(\eta_i)_i$ are especially easy in application and useful in the sense of sufficient optimality conditions and of characterization of solutions. They give good results for nonconvex programming when many results in the literature, including the Kuhn-Tucker sufficient optimality conditions and the Fritz John ones combining with the (generalized) invexity with respect to the same η, are not applicable. In this way, previously known results in this area are generalized and extended.

Acknowledgements. The authors are grateful to the anonymous referee for his/her valuable suggestions and comments which have helped to improve the final version of the paper.

References

[1] T. Antczak, A class of B-(p,r)-invex functions and mathematical programming. *J. Math. Anal. Appl.* **286** (2003) 187–206.

[2] T. Antczak, New optimality conditions and duality results of G type in differentiable mathematical programming. *Nonlinear Anal.* **66** (2007) 1617–1632.

[3] T. Antczak, On (p, r)-invexity-type nonlinear programming problems. *J. Math. Anal. Appl.* **264** (2001) 382–397.

[4] T. Antczak, r-preinvexity and r-invexity in mathemetical programming. *Comp. Math. Appl.* **50** (2005) 551–566.

[5] M.S. Bazaraa, H.D. Sherali and C.M. Shetty, Nonlinear Programming: Theory and Algorithms, 3rd edn. Wiley, New York (2006).

[6] C.R. Bector, S.K. Suneja and C.S. Lalitha, Generalized B-vex functions and generalized B-vex programming. *J. Optim. Theor. Appl.* **76** (1993) 561–576.

[7] A. Ben-Israel and B. Mond, What is invexity ? *J. Austral. Math. Soc. Ser. B* **28** (1986) 1–9.

[8] B.D. Craven, Invex functions and constrained local minima. *Bull. Austral. Math. Soc.* **24** (1981) 357–366.

[9] B.D. Craven and B.M. Glover, Invex functions and duality. *J. Austral. Math. Soc. Ser. A* **39** (1985) 1–20.

[10] F. Dinuzzo, C.S. Ong, P. Gehler and G. Pillonetto, *Learning output kernels with block coordinate descent, in Proc. of the 28th International Conference on Machine Learning*, Bellevue, WA, USA, 2011.

[11] K.H. Elster, R. Nehse, *Optimality conditions for some nonconvex problems.* Springer-Verlag, New York 1980.

[12] C. Fulga and V. Preda, Nonlinear programming with E-preinvex and local E-preinvex functions. *Eur. J. Oper. Res.* **192** (2009) 737–743.

[13] M.A. Hanson, On sufficiency of the Kuhn-Tuker conditions. *J. Math. Anal. Appl.* **80** (1981) 445–550.

[14] M.A. Hanson and B. Mond, Necessary and sufficient conditions in constrained optimization. *Math. Program.* **37** (1987) 51–58.

[15] M. Hayashi and H. Komiya, Perfect duality for convexlike programs. *J. Optim. Theor. Appl.* **38** (1982) 179–189.

[16] V. Jeyakumar and B. Mond, On generalized convex mathematical programming. *J. Austral. Math. Soc. Ser. B* **34** (1992) 43–53.

[17] F. John, Extremum problems with inequalities as side conditions, in K.O. Friedrichs, O.E. Neugebauer and J.J. Stoker Eds., *Studies and Essays, Courant Anniversary Volume.* Wiley (Interscience), New York (1948) 187–204.

[18] W. Karush, *Minima of functions of several variables with inequalities as side conditions*, Master's Thesis. Department of Mathematics, University of Chicago, 1939.

[19] R.N. Kaul and S. Kaur, Optimality criteria in nonlinear programming involving nonconvex functions. *J. Math. Anal. Appl.* **105** (1985) 104–112.

[20] R.N. Kaul, S.K. Suneja and M.K. Srivastava, Optimality criteria and duality in multiple-objective optimization involving generalized invexity. *J. Optim. Theor. Appl.* **80** (1994) 465–482.

[21] H.W. Kuhn and A.W. Tucker, *Nonlinear programming, in Proc. of the Second Berkeley Symposium on Mathematical Statistics and Probability*, edited by J. Neyman, University of California Press, Berkeley, California (1951) 481–492.

[22] O.L. Mangasarian, *Nonlinear Programming.* McGrawHill, New York (1969).

[23] D.H. Martin, The essence of invexity. *J. Optim. Theor. Appl.* **47** (1985) 65–76.

[24] S.K. Mishra, R.P. Pant and J.S. Rautela, Duality in vector optimization under type I α-invexity in banach spaces. *Numer. Funct. Anal. Optim.* **29** (2008) 1128–1139.

[25] S.K. Mishra, S.Y. Wang and K.K. Lai, On non-smooth α-invex functions and vector variational-like inequality. *Optim Lett.* **02** (2008) 91–98.

[26] B. Mond and T. Weir, Generalized concavity and duality, *in Generalized concavity in optimization and economics*, edited by S. Schaible and W.T. Ziemba, Academic Press, New York (1981) 263–276.

[27] H. Nickisch and M. Seeger, Multiple kernel learning: a unifying probabilistic viewpoint. `arXiv:1103.0897v2 [stat.ML]` 30 March (2011).

[28] R. Osuna-Gomez, A. Beato-Morero and A. Rufian-Lizana, Generalized convexity in multiobjective programming. *J. Math. Anal. Appl.* **233** (1999) 205–220.

[29] R. Pini and C. Singh, A survey of recent [1985-1995] advances in generalized convexity with applications to duality theory and optimality conditions. *Optim.* **39** (1997) 311–360.

[30] N.G. Rueda and M.A. Hanson, Optimality criteria in mathematical programming involving generalized invexity. *J. Math. Anal. Appl.* **130** (1988) 375–385.

[31] H. Slimani and M.S. Radjef, Duality for nonlinear programming under generalized Kuhn-Tucker condition. *Int. J. Optim. Theor. Methods Appl.* **1** (2009) 75–86.

[32] H. Slimani and M.S. Radjef, *Fonctions invexes généralisées et optimisation vectorielle: optimalité, caractérisations, dualité et applications*. Editions Universitaires Européennes, Saarbrücken, 2011.

[33] H. Slimani and M.S. Radjef, *Multiobjective programming under generalized invexity: optimality, duality, applications*. LAP Lambert Academic Publishing, Saarbrücken, 2010.

[34] H. Slimani and M.S. Radjef, Nondifferentiable multiobjective programming under generalized d_I-invexity. *Eur. J. Oper. Res.* **202** (2010) 32–41.

[35] M. Soleimani-damaneh and M.E. Sarabi, Sufficient conditions for nonsmooth r-invexity. *Numer. Funct. Anal. Optim.* **29** (2008) 674–686.

[36] T. Weir and B. Mond, Pre-invex functions in multiple objective optimization. *J. Math. Anal. Appl.* **136** (1988) 29–38.