Invexity versus generalized convexity

Cristina Liliana Pripoae and Gabriel Teodor Pripoae

Abstract. In this paper, we make an attempt to answer the following general question: *when a smooth non-convex (from the classical viewpoint) function may be considered generalized convex, with respect to some properly chosen linear connection* ? We prove this is possible for a large family of real valued smooth functions on differentiable manifolds, provided they are regular or have all critical points strictly convex. Counterexamples are given to show that the result cannot be improved-in general-in order to cover the excepted cases. We extend Craven's characterization of invexity, from functions defined on $\mathbb{R}^n$ to functions on differentiable manifolds, and use this result to compare generalized strictly convexity with invexity.

M.S.C. 2010: 53A15,90C25,90C26.
Key words: invexity, generalized convexity, generalized strictly convexity, critical points of minimum type, local vs. global minimum points.

1 Introduction

The main problems in Optimization Theory are very hard to solve, despite their very simple statements: *Does a given function have minimum points*? *If the answer is affirmative, are these points local or global minimum points*? *Are the minimum points unique and if not, "how many" are they*?

For real valued smooth functions f on $\mathbb{R}^n$, there exist some well-known answers, based on the classical convexity theory ([11], [10]):

(i) if f is strictly convex at a critical point x_0, then x_0 is a local minimum point;

(ii) if f has a minimum point x_0, then f is convex around x_0;

(iii) if f is convex on $\mathbb{R}^n$ and x_0 is a critical point, then x_0 is a global minimum point;

(iv) if f is strictly convex on $\mathbb{R}^n$ and x_0 is a critical point, then x_0 is the unique global minimum point.

Classical convex optimization theory developed simple and powerful algorithms to reach a global minimum point (when there exists one)([11]). Unfortunately, the

BSG Proceedings 18. The Int. Conf. of Diff. Geom. and Dynamical Systems (DGDS-2010), October 8-11, 2010, Bucharest Romania, pp. 87-94.

strictly convexity is rare among smooth functions having a global minimum point and convexity alone gives not enough information.

This was one of the reasons which led to a generalization of the (strictly) convexity of functions, from the Euclidean to the Riemannian setting (see [11] for historical details). Despite the increase of the difficulty of the technical formalism, this new theory of convexity offered also pleasant surprises: many functions which are not (classically) convex become convex with respect to some properly chosen Riemannian metrics. One famous example is provided by the "Rosenbrock banana function" ([11]).

The price to pay for this more general setting is the search for an appropriate Riemannian metric, which sometimes may be tedious.

In [8], we extended the Riemannian convexity of functions, in the more general affine differential setting: the Hessian operator is constructed using an arbitrary linear connection (instead of the Levi-Civita one) and the geodesic links are replaced by auto-parallel curves links. Given a real valued smooth function on a differentiable manifold, one looks for a linear connection which "makes" it generalized (strictly) convex.

In this paper, for the sake of completeness, we expose the main affine differential tools needed to understand (affine differential) generalized (strictly) convexity (§2). Next, we recall the classical notion of smooth invex functions on Euclidean spaces (§3) and their generalization on differentiable manifolds ([8]); in the latter case, we extend their characterization, proven in the classical setting by Craven ([2]):

Theorem. *Let f be a real valued smooth function on a differentiable manifold. Then f is invex if and only if it is regular or every critical point is a global minimum one.*

This result shows that (affine differential) generalized strictly convexity remains a special case of invexity, when passing from functions defined on $\mathbb{R}^n$ to functions defined on differentiable manifolds.

In §4, we recall several examples of non-convex (from the classical theory viewpoint) functions which admit linear connections with respect to whom the respective functions become generalized convex ([8]). These examples were the starting point for the following result, proved by us in [9]:

Theorem. *Let f be a real valued smooth function on a differentiable manifold M. Suppose f is regular or has only one critical point, where f is strictly convex. Then there exists a linear connection with respect to which f is generalized convex.*

In §5, we generalize this result and prove that every real valued smooth function f on a differentiable manifold is generalized convex, with respect to some properly choosen linear connection, provided f is regular or f is strictly convex in all its critical points.

Counterexamples are given to show that the result cannot be improved in order to cover the excepted cases (§6): for a particular function *with critical non-extremum points* and for a particular function *with a maximum point*, we prove there do not exist linear connections which make them generalized convex.

But the most surprinsing examples arose from some (quite similar) functions having only one *degenerate* critical point of minimum type: one of them can be made generalized convex and the other does not.

2 Generalized convexity in affine differential geometry

Consider a differentiable manifold M and $\mathcal{F}(M)$ the algebra of the real valued smooth (i.e. C^∞-differentiable) functions on M. Denote by $\mathcal{X}(M)$ the $\mathcal{F}(M)$-module of vector fields on M and by $\mathcal{C}(M)$ the set of linear connections on M. We recall that a linear connection $\nabla \in \mathcal{C}(M)$ is an operator from $\mathcal{X}(M) \times \mathcal{X}(M)$ to $\mathcal{X}(M)$, $\mathcal{F}(M)$-linear in the first argument, $\mathbb{R}$-linear in the second argument and, for each function $f \in \mathcal{F}(M)$ and for each vector fields X, $Y \in \mathcal{X}(M)$, we have

$$\nabla_X fY = f\nabla_X Y + df(X)Y. \tag{2.1}$$

Each linear connection ∇ defines an *affine differentiable structure* on M. For $f \in \mathcal{F}(M)$, the Hessian operator with respect to ∇ is a (0,2)-tensor field, defined by

$$H_f(X,Y) = (\nabla_X df)(Y). \tag{2.2}$$

We say $f \in \mathcal{F}(M)$ is ∇-*generalized convex* (respectively ∇-*generalized strictly convex*) at $x_0 \in \mathbb{R}^n$ if its Hessian H_f is positive semi-definite (respectively positive definite) in x_0. In this case, we say x_0 is a convex (respectively strictly convex) point of f.
A function $f \in \mathcal{F}(M)$ is called *generalized convex* (respectively *generalized strictly convex*) if there exists a linear connection ∇ on M, with respect to which f becomes ∇-*generalized convex* (respectively ∇-*generalized strictly convex*).

(These definitions of convexity differ slightly from those in [11].)

Remark 2.1. (i) Any real valued smooth convex function on $\mathbb{R}^n$ is also generalized convex, with respect to the canonical (Euclidean) linear connection.

(ii) Any regular function on a differentiable manifold is generalized convex, with respect to some linear connection. Moreover, this linear connection may be chosen such that the Hessian of f identically vanish ([11]).

(iii) In local coordinates $(x^1,...,x^n)$, the components of a linear connection are smooth functions Γ^i_{jk}, $i,j,k \in \{1,...,n\}$, given by

$$\nabla_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j} = \Gamma^k_{ij} \frac{\partial}{\partial x^k}. \tag{2.3}$$

The Hessian of a smooth function f writes

$$H_{ij} = f_{ij} - \Gamma^k_{ij} f_k, \tag{2.4}$$

where f_k are the first order partial derivatives of f and f_{ij} are the second order partial derivatives of f.

Fix a function $f \in \mathcal{F}(M)$. We denote by $\mathcal{C}_f$ (resp. $\mathcal{C}^s_f$) the set of linear connections ∇ such that f is ∇-convex (respectively strictly convex). Fix $\nabla \in \mathcal{C}(M)$. We denote by $\mathcal{F}_\nabla$ (resp. $\mathcal{F}^s_\nabla$) the set of ∇-convex (resp. strictly convex) functions.

Remark 2.2. (i) All the previously defined sets are convex sets and differential (resp. affine differential) invariants.

(ii) Obviously, $\mathcal{F}_\nabla$ is non-void. If $\mathcal{F}^s_\nabla \neq \varnothing$, then the manifold M admits a Hessian structure (i.e. a Riemannian metric which is the Hessian of a smooth function).

(iii) The assertions (i)-(iv) from §1 remain true on differentiable manifolds, replacing the *(strictly) convexity* by *affine differential generalized (strictly) convexity.*

3 Invexity of smooth functions on differentiable manifolds

Consider an n-dimensional differentiable manifold M and $f : M \to \mathbb{R}$ a smooth function. Denote by T_pM the tangent space at $p \in M$ and $TM = \{(p, v) \mid p \in M, v \in T_pM\}$ the total space of the tangent bundle of M. Let η, $\theta : M \times M \to TM$ be two smooth functions (with respect with the product differentiable structure on $M \times M$), such that, for every p, $q \in M$ we have $\eta(p, q)$, $\theta(p, q) \in T_qM$. We may view η and θ as two families of vector fields, indexed by all the points of M (the first variable).

Definition 3.1. The function f is $(\eta, \theta) - invex$ if, for every p,$q \in M$,

$$f(p) - f(q) \geq df_q(\eta(p, q)) + df_p(\theta(q, p)). \tag{3.1}$$

If, moreover, for every two points p, $q \in M$ the previous relation holds with strict inequality, we say the function f is *strictly* $(\eta, \theta) - invex$.

Remark 3.1. Due originally to Hanson ([5]), for a vector fields family defined by the position vector in $\mathbb{R}^n$, the invexity was intensively studied by many authors, including Craven [2], Mond, Ben-Israel [1], Giorgi [4], Mishra [6], Udriste, Mititelu [7], [3]; we recover its (classical) definition, by taking $\theta = 0$ in (3.1).

Craven proved the following result, in the particular case of smooth functions in $\mathbb{R}^n$ ([2]; see also [1]).We extended it by the

Theorem 3.1. *Let f be a real valued smooth function on a differentiable manifold. Then f is invex if and only if it is regular or every critical point is a global minimum.*

Proof. Suppose f is η-invex, where $\eta : M \times M \to TM$ satisfies $\eta(p, q) \in T_qM$, for any points $p, q \in M$. Let q be a critical point of f. Then

$$f(p) - f(q) \geq df_q(\eta(p, q)) = 0,$$

hence q is a global minimum point for f.

Conversely, let $q \in M$.

Case I. Suppose $df_q = 0$; then $f(p) \geq f(q)$, for every $p \in M$. We define $\eta(p, q) = 0$, for every $p \in M$.

Case II. Suppose $df_q \neq 0$. Consider an arbitrary Riemanian metric g on M and define

$$\eta(p, q) = [g_q(gradf, gradf)]^{-1}[f(p) - f(q)]gradf_q.$$

Then $df_q(\eta(p, q)) = f(p) - f(q)$, so f is invex. □

Remark 3.2. Using the Remark 2.2.(iii), we see that any (affine differential) generalized strictly convex function is invex.

The example (iii) from §6 will show that the converse of the previous statement is false.

4 Case studies

In [9] we considered several functions admitting only critical points which are global minimum ones and proved that their lack of convexity was only apparent, because (even if they were not classically convex) they were generalized convex in an appropriate affine differential geometry. We recall briefly the respective examples (for details,including also some pictures, see [9]).

(i) The function

$$f_1 : \mathbb{R}^2 \to \mathbb{R}, \;\; f_1(x,y) = x^4 + y^4 - 6(x^3 + y^3) + 14(x^2 + y^2)$$

is convex (in classical sense) and has only one (global) minimum point (0,0).

(ii) We slighty modify f_1 to

$$f_2(x,y) = x^4 + y^4 - 6(x^3 + y^3) + 12(x^2 + y^2),$$

which is no more (classically) convex, as a short calculation shows. However, the minimum point property remains the same as for f_1. In global coordinates on $\mathbb{R}^2$, we found ([9]) a linear connection ∇, with respect to which f_2 is (affine differential) strictly convex.

(iii) The (classical) convexity loss becomes more evident for

$$f_3(x,y) = x^4 + y^4 - 6.3(x^3 + y^3) + 12(x^2 + y^2),$$

whose graph has a bigger "bump" on one side. Of course, the (global) minimum point property is stable with respect to these "bumps".

With respect to the same connection ∇ as in (ii), the function f_3 is also strictly convex. (Interestingly, when the "bump" grows, as for the value 6.54 instead of 6.3, the perturbed function cannot be made generalized convex, anyhow we would choose the linear connection; this situation occurs because the new function gets also local maximum points).

(iv) The uniqueness of the minimum point is necessary for the generalized strictly convexity of the function f. This does not hold for generalized convexity, as the following example shows.

Let the function $f_4 : \mathbb{R}^2 \to \mathbb{R}$, given by $f_4(x,y) = x^2 e^y$. This function has an infinity of minimum points $(0,y)$, with $y \in \mathbb{R}$, and no other critical point.It is easy to see that f_4 is not (classically) convex, because its (classical) Hessian is indefinite in the point (1,0).

In [9] we found a linear connection ∇ such that f_4 be ∇-convex.

5 The main result

As the previous examples suggest, smooth functions having all the critical points of minimum type are likely to be "made convex", by choosing an appropriate linear connection. In fact, we may prove the following result:

Theorem 5.1. *Let f be a real valued smooth function on a differentiable manifold M. Suppose f has only critical points which are strictly convex. Then the function f is generalized convex.*

Proof. Let $x_0 \in M$ be a minimum point of f and let $\nabla^1 \in \mathcal{C}(M)$ be an *arbitrary* linear connection. Since x_0 is a critical point, the Hessian H_f in x_0 does not depend on the choice of the linear connection in $\mathcal{C}(M)$; so the positiveness of H_f in x_0 implies f is ∇^1-strictly convex in a neighborhood U of x_0. If $U = M$, the theorem is proved.

If the union of all U's is strictly contained in M, denote V the complement in M of the topological closure of this reunion. The restriction of f to the open set V is regular; hence, by the Remark 2.1., (ii) there exists a linear connection $\nabla^2 \in \mathcal{C}(V)$ such that the restriction to V of f is ∇^2-convex.

Due to the (implicitly supposed) paracompactness of the manifold M, there exist an open sub-covering $\{W_i\}_i$ of the open covering of the U's (with, eventually, V) and a differentiable partition of unity $\{\phi_i\}_i$ associated to it. Consider a family $\{\nabla^i\}_i$ of linear connections, with $\nabla^i \in \mathcal{C}(W_i)$, such that ∇^i be the restriction of either a ∇^1 or, eventually, ∇^2.

The restriction of f to each set W_i is ∇^i-convex. We define $\nabla := \sum_i \phi_i \nabla^i$. First, we remark that ∇ is a linear connection on M. This is due to the fact that the set of linear connections behaves like an affine module over the ring of (germs of) smooth functions on M.

Secondly, the function f is ∇-convex. Indeed, let remark that if ∇' and ∇" are two linear connections and h a smooth function, then we may construct a new connection $\nabla''' = h\nabla' + (1-h)\nabla''$. Consider another function α; then the Hessian of α with respect to ∇''' writes $H'''_\alpha = hH'_\alpha + (1-h)H''_\alpha$. This proves that if, moreover, α is ∇'-convex and ∇''-convex, and if h takes values in the interval [0,1], then α is also ∇'''-convex. Since all the functions from the partition of unity have the previous property, it follows that the function f is convex with respect to the new constructed connection ∇. □

Combining the Theorem 5.1. with the Remark 2.1.(ii), we get the

Corollary 5.2. *Let f be a real valued smooth function on a differentiable manifold M. Suppose f is regular or has only critical points, all strictly convex. Then the function f is generalized convex.*

6 Counterexamples

The hypothesis of the Theorem 5.1. cannot be weakened, as the following counter-examples show.

(i) Consider the smooth function $f_5 : \mathbb{R}^2 \to \mathbb{R}$, $f_5(x,y) = x^3 + y^3$. This function is not (classically) convex and has an unique critical non-extremum point (0,0) ([9]).

Suppose ad absurdum that there exists a linear connection ∇ on M such that f_5 is ∇-convex. The Hessian of f_5 with respect to ∇ must be positive definite, so the following inequalities hold

$$2x - \Gamma^1_{11}x^2 - \Gamma^2_{11}y^2 \geq 0, \qquad 2y - \Gamma^1_{22}x^2 - \Gamma^2_{22}y^2 \geq 0,$$

for every real numbers x and y. In the first inequality, we make $y := 0$ and denote $b(x) := \frac{1}{2}\Gamma^1_{11}(x,0)$. Thus, there exists a smooth function $b \in \mathcal{F}(\mathbb{R})$ such that $x -$

$b(x)x^2 \geq 0$, for every real number x. We deduce $b(x) \leq \frac{1}{x}$ for $x < 0$ and $b(x) \leq \frac{1}{x}$ for $x > 0$, fact which contradicts the continuity of b.

We conclude the initial supposition is false, so f_5 cannot be made generalized convex in any affine differential geometry on $\mathbb{R}^2$.

(ii) Consider now the smooth function $f_6 : \mathbb{R}^2 \to \mathbb{R}$, given by $f_6(x,y) = sin(x^2 + y^2)$. This function has both minimum and maximum points ([9]).

Suppose ad absurdum that there exists a linear connection ∇ on M such that f_6 is ∇-convex. The Hessian H of f_6 with respect to ∇ has the following components:

$$H_{11}(x,y) = 2cos(x^2+y^2)(1-\Gamma^1_{11}x-\Gamma^2_{11}y) - 4x^2 sin(x^2+y^2)$$
$$H_{12}(x,y) = -2cos(x^2+y^2)(\Gamma^1_{12}x+\Gamma^2_{12}y) - 4xy sin(x^2+y^2)$$
$$H_{22}(x,y) = 2cos(x^2+y^2)(1-\Gamma^1_{22}x-\Gamma^2_{22}y) - 4y^2 sin(x^2+y^2),$$

for every real numbers x and y. We make $x := \sqrt{\frac{\pi}{2}}$ and $y := 0$. At this point, all the components of the Hessian vanish, excepted $H_{11} = -2\pi$, which shows that here the Hessian is negatively defined.

We deduce the initial supposition is false, so f_6 cannot be made generalized convex in any affine differential geometry on $\mathbb{R}^2$.

(iii) Let $f_7 : \mathbb{R}^2 \to \mathbb{R}$, given by $f_7(x,y) = x^4 + ax^2y^2 + y^4$, $a > -2$. The Hessian of f has positive trace and the determinant equals to $g(x,y) := (144 - 12a^2)x^2y^2 + + 24a(x^4+y^4)$. The point (0,0) is a global, unique, degenerated minimum point.

Case I. If $a = 1$, the function is (classically) convex.

Case II. Suppose $a \in (-2, 0)$. Then $g(0,y) < 0$ for any non-null y, so f is not (classically) convex around the origin.Moreover, there does not exist a linear connection ∇ on $\mathbb{R}^2$ such that f becomes ∇-convex.

Indeed, suppose ad absurdum such a linear connection ∇ exists. Then the Hessian of f_7, with respect to ∇, must be positive definite around the origin. In particular, its determinant

$$h(x,y) := [12x^2 + 2ay^2 - \Gamma^1_{11}(4x^3+2axy^2) - \Gamma^2_{11}(4y^3+2ax^2y)][12y^2+2ax^2-$$

$$-\Gamma^1_{22}(4x^3+2axy^2) - \Gamma^2_{22}(4y^3+2ax^2y)] - [4axy - \Gamma^1_{12}(4x^3+2axy^2) - \Gamma^2_{12}(4y^3+2ax^2y)]^2$$

must be non-negative there. But we remark that

$$h(0,y) = 8y^4[3a - (6\Gamma^2_{11} + a\Gamma^2_{22})y + 2(\Gamma^2_{11}\Gamma^2_{22} - (\Gamma^2_{12})^2)y^2]$$

is negative, for non-null (and close enough to zero) y. So, in this case, f_7 cannot be made generalized convex.

7 Conclusions

Let f be a real valued smooth function on a differentiable manifold M.

(i) The function f is invex if and only if it is regular or every critical point is a global minimum.

(ii) If f is generalized strictly convex (i.e. strictly convex with respect to some linear connection in M), then it is invex. The converse is false, in general, but is true for a special class of invex functions: the regular ones and those having exactly one generalized strictly convex critical point.

(iii) Regularity or the strictly convexity at all the critical points ensures f can be made (affine differential) generalized convex.

(iv) An open problem is to find an invariant who decides which non-strictly convex critical points allow generalized convexity and which forbid.

References

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

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

[3] M. Ferrara and S. Mititelu, *Mond-Weir duality in vector programming with generalized invex functions on differentiable manifolds*,Balkan J. Geom. Appl. 16 (2006),no.2, 80-87.

[4] G. Giorgi, *A note on the relationships between convexity and invexity*, J. Austral. Math. Soc.. Ser B., 32 (1990), 97-99.

[5] M. A. Hanson, *On sufficiency of the Kuhn-Tucker conditions*, J. Math. Anal. Appl. 80 (1981), 545-550.

[6] S.K. Mishra and G. Giorgi, *Invexity and optimization*, Nonconvex Optimization and Its Applications 88, Springer-Verlag, Berlin, 2008.

[7] S. Mititelu, *Generalized convexity*, (in Romanian), Balkan Geometry Press, 2009.

[8] C.L. Pripoae and G.T. Pripoae, *Generalized invexity on differentiable manifolds*, Balkan J. Geom. Appl. 10 (2005),no.1, 155-164.

[9] C.L. Pripoae and G.T. Pripoae, *Differential and affine differential invariants for convexity on differentiable manifolds*, Proc. Balkan Conference on Oper. Research, Constanta May 25-28 2005, Bucharest, 2007, 165-170.

[10] C. Udrişte and E. Tănăsescu, *Minima and maxima of real valued functions*, (in Romanian), Ed. Tehnica, Bucharest, 1980.

[11] C. Udrişte, *Convex functions and optimization methods on Riemannian manifolds*, Kluwer Acad. Publ., 1994.

Authors' addresses:

Cristina Liliana Pripoae
Department of Mathematics,
Faculty of Cybernetics, Statistics and Economic Informatics,
Academy of Economic Studies, Piaţa Romană no.6, Bucureşti, România.
E-mail: cristinapripoae@yahoo.com

Gabriel Teodor Pripoae
Department of Mathematics,
Faculty of Mathematics and Computer Science,
University of Bucharest, Str. Academiei no.14, Bucureşti, România.
E-mail: gpripoae@fmi.unibuc.ro