Differential and Affine Differential Invariants for Convexity on Differentiable Manifolds

Cristina Liliana Pripoae
Department of Mathematics, Academy of Economic Studies, Bucharest
Gabriel Teodor Pripoae
Faculty of Mathematics and Informatics, University of Bucharest

May 25, 2005

Abstract

We prove that some special functions which are non-convex (from the classical viewpoint) may be considered generalized convex, with respect to some properly choosed linear connections. This behaviour provides support to conjecture that this is true for any differentiable function, as far as none of its critical points is a maximum one.

Keywords: generalized convexity, differential invariants, affine differential invariants
MSC 2000: 53A15,90C25, 90C26

1 Introduction

In order to decide which local extremum points are global ones, classical optimization theory usually uses convexity criteria. In the Euclidean setting (i.e. for real valued functions defined on subsets of $\mathbf{R}^n$), the true essence of convexity (which is an affine property) is hidden by and mixed with metric properties. Considering generalized convexity on Riemannian manifolds (as in [5]) increases the degree of generality; many functions which are not (classically) convex become convex with respect to some properly choosed Riemannian metrics (the characteristic property along segment lines is replaced by the same property along geodesic paths).

In [1], [2] we extended the generalized convexity of functions, in the Affine differential setting: the Hessian operator is constructed by using an arbitrary linear connection (instead of the Levi-Civita one) and the geodesic links are replaced by auto-parallel curves links. Two natural questions arrise:

1. *given a linear connection, do there exist smooth functions which are generalized convex with respect to it* ?

2. *given a smooth function, do there exist linear connections which "make" it generalized convex*?

Partial answers to both problems were provided in [2]; in particular, for the Rosenbrock banana function, we found an infinite family of linear connections with respect to whom the respective function is (generalized) convex (see §2).

In this paper, we expose the main affine differential tools for generalized convexity (§2). Next, we consider several examples of non-convex (from the classical theory viewpoint) functions and find linear connections with respect to whom the respective functions become generalized convex (§3).

We prove that every differentiable function on a differentiable manifold, with one global minimum point, is generalized convex, with respect to some properly choosed linear connection (§4).

Finally, we give an example to support the conjecture that the previous result is true also for functions with an infinity of critical points, provided none of them is a maximum one (§5).

2 Generalized convexity in Affine differential geometry

Consider a differentiable manifold M and $\mathcal{F}(M)$ the algebra of the real valued differentiable (i.e. smooth) 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 $\mathcal{C}(M) \times \mathcal{C}(M)$ to $\mathcal{C}(M)$, $\mathcal{F}(M)$-linear in the first argument, **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

$$\nabla_X fY = f\nabla_X Y + df(X)Y \tag{1}$$

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

$$H_f(X,Y) = (\nabla_X df)(Y) \tag{2}$$

We say $f \in \mathcal{F}(M)$ is ∇-*convex* (respectively ∇-*strictly convex*) if its Hessian H_f is semi-positively defined (respectively positively defined).

Remarks 1. If ∇ is a linear connection on M, we denote by ∇^t and ∇^s the transposed connection and the symmetric connection, respectively, associated to ∇ by the formulas $\nabla^t_X Y = \nabla_Y X + [X,Y]$ and $2\nabla^s = \nabla + \nabla^t$.

Let f be a differentiable function on M. Then

(i) The Hessian of f with respect to ∇ is the transposed of the Hessian of f with respect to ∇^t. The Hessian of f with respect to ∇^s is the mean of the Hessians with respect to ∇ and ∇^t.

(ii) The folowing assertions are equivalent:
a) f is ∇-convex (resp. strictly convex);
b) f is ∇^t-convex (resp. strictly convex);
c) f is ∇^s-convex (resp. strictly convex).

(iii) H_f (with respect to ∇) is symmetric if and only if the torsion of ∇ belongs to the kernel of df; that is, $df(\nabla_X Y - \nabla_Y X - [X,Y]) = 0$, where $X, Y \in \mathcal{X}(M)$. (In particular, we recover the known fact that, with respect to a symmetric linear connection, all Hessian operators are symmetric).

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

$$\nabla_{\frac{\partial}{\partial x^j}} \frac{\partial}{\partial x^k} = \Gamma^i_{jk} \frac{\partial}{\partial x^i}$$

The Hessian of a differentiable function f writes

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

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

Notations 2. Fix a function $f \in \mathcal{F}(M)$. We denote by $\mathcal{C}_f$ (resp. $\mathcal{C}_f^s$) 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. $\mathcal{F}_\nabla^s$) the set of ∇-convex (resp. strictly convex) functions.

Remarks 3. (i) All the previously defined sets are convex sets and differential (resp. affine differential) invariants.

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

Examples 4. (i) With respect to the canonical connection ∇ of the Euclidean space $\mathbf{R}^n$, $\mathcal{F}_\nabla$ contains all the C^2-differentiable functions on $\mathbf{R}^n$, convex from the classical view point.

(ii) The Rosenbrock banana function $f : \mathbf{R}^2 \to \mathbf{R}$ is given by

$$f(x_1, x_2) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2$$

It is not a (classically) convex function; it becomes (generalized) convex, for example, with respect to the (Levi Civita connection associated to the) Riemannian metric $g = (4x_1^2 + 1)(dx_1)^2 - 4x_1 dx_1 dx_2 + (dx_2)^2$ ([5]).

Moreover, it is easy to show there exists an infinite family of linear connections, whose components are solutions of the linear equations system

$$a\Gamma_{11}^1 + b\Gamma_{11}^2 = 600x_1^2 - 200x_2 + 1 - \alpha^2$$

$$a\Gamma_{22}^1 + b\Gamma_{22}^2 = 100 - \beta^2$$

$$a\Gamma_{12}^1 + b\Gamma_{12}^2 = -200x_1 \pm \sqrt{\alpha^2\beta^2 - \gamma^2}$$

where

$$a = 200x_1^3 - 200x_1x_2 + x_1 - 1 \ , \ b = 100x_2 - 100x_1^2$$

and $\alpha^2, \beta^2, \gamma^2$ are arbitrary positive (parameter) functions on M, such that the squared root has sense.

Remark 5. (i) Let ∇ be a symmetric linear connection on M. Denote by R its curvature (1,3)-tensor field, given by

$$R(X, Y)Z = \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X,Y]}$$

for every vector fields $X, Y, Z \in \mathcal{X}(M)$.

Then ([1]) $\mathcal{F}_\nabla$ contain nonconstant functions if and only if there exist a nonconstant function f and a semi-positively defined (0,2)- tensor field α, such that

$$df(R(X, Y)Z) = (\nabla_Y \alpha)(X, Z) - (\nabla_X \alpha)(Y, Z) \tag{1}$$

for every vector fields $X, Y, Z \in \mathcal{X}(M)$. In particular, this implies

$$df(\nabla_X R(Y, Z)V + \nabla_Z R(X, Y)V + \nabla_Y R(Z, X)V) = 0 \tag{2}$$

(ii) Let f be a regular function on a differentiable manifold M, endowed with a symmetric linear connection ∇. Suppose the curvature of ∇ vanishes on a (Chevalley) distribution complementary to $Kerdf$. Then, there exist ∇-convex functions on M.

In particular, this situation occurs on every locally Euclidean manifold.

(iii) Let f be a regular function on M. There exists ([5]) a linear connection ∇^o such that f is ∇^o-linear affine (i.e. $H^o_f = 0$). We choose ξ a (local) generator of the complementary distribution of $Kerdf$, such that $df(\xi < 0$. Then $\mathcal{C}_f$ (resp. $\mathcal{C}^s_f$) is the set of all connections of the form

$$\nabla_X Y = \nabla^o_X Y + A(X,Y) + B(X,Y)\xi \tag{3}$$

where $A \in \mathcal{T}^1_2(M)$, $A(X,Y) \in Kerdf$; and $B \in \mathcal{T}^0_2(M)$ is semi-positively (resp. positively) defined ([1]).

In particular, consider (M,g) a Riemannian manifold. We may obtain a large family of linear connections in $\mathcal{C}_f$, by taking $B(X,Y) = b^2 g(X,Y)$, where b is an arbitrary function on M. (With slighty modified details we may obtain examples for $\mathcal{C}^s_f$).

3 Case studies

We consider several functions admitting only critical points which are global minimum ones. We prove that their (possible) lack of convexity is only apparent, because they may be considered convex in an appropriate differential affine geometry.

(i) First, let $f_1 : \mathbf{R}^2 \to \mathbf{R}$ be the function given by

$$f_1(x,y) = x^4 + y^4 - 6(x^3 + y^3) + 14(x^2 + y^2)$$

As its graph shows below, f_1 has a paraboloid-like shape (Fig.1). A direct computation proves f_1

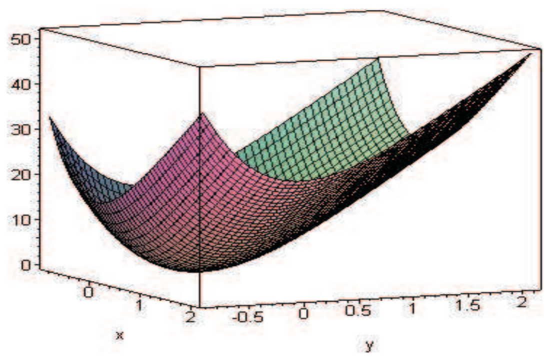

Figure 1: Graph of f_1

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

(ii) We slighty modify the previous function to

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

This new function is no more (classically) convex, as a short calculation shows and as its graph hints (Fig.2). However, the minimum point property remains the same as for f_1. In global

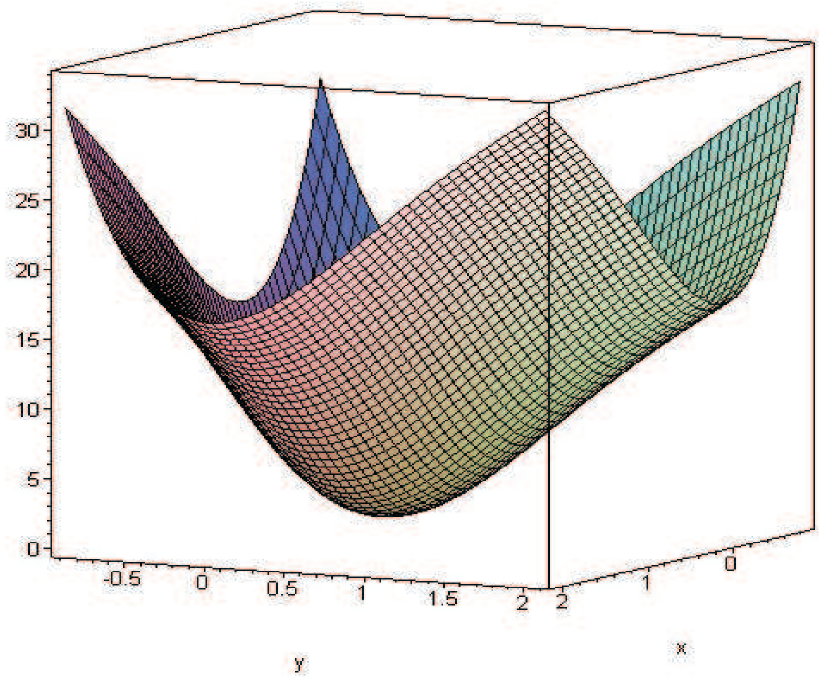

Figure 2: Graph of f_2

coordinates on $\mathbf{R}^2$, we find a linear connection ∇, with respect to which f_2 is (affine differential) strictly convex; the (only non-vanishing) connection components are:

$$\Gamma^1_{11} = -x \quad , \quad \Gamma^2_{22} = -y$$

The (only non-vanishing) components of the Hessian matrix of f_2, with respect to ∇, are

$$H_{11} = 4x^4 - 18x^3 + 36x^2 - 36x + 24$$

$$H_{22} = 4y^4 - 18y^3 + 36y^2 - 36y + 24$$

(the positivity of H follows from the positivity of these previous two functions).

(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" at the right side (Fig.3). Of course, the (global) minimum point property is stable with respect to these "bumps". With respect to the same connection ∇ as in (ii),

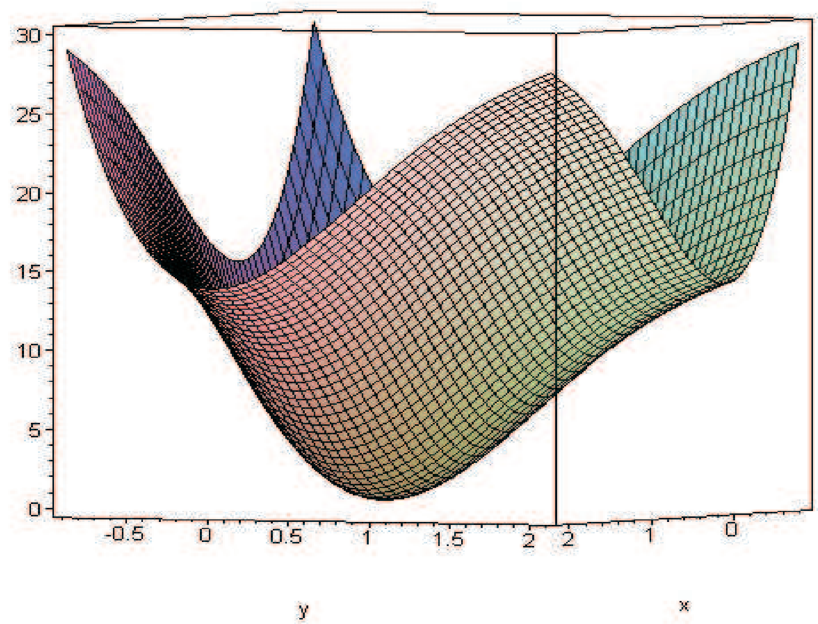

Figure 3: Graph of f_3

the function f_3 is also strictly convex. (Interestingly, when the "bump" grows, as for 6.45 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).

4 The main result

As the previous examples suggest, differentiable functions with only one minimum point are likely to be "made convex", by choosing an appropriate linear connection. In fact, we may prove the following result, which completes what was already known from the Remark 5,(iii) (i.e. in the case of regular differentiable functions).

Theorem 6. *Let f be a real valued differentiable function on a differentiable manifold M.Suppose f has only a critical point, which is of minimum type. Then there exists a linear connection in $\mathcal{C}_f$.*

Proof. Let $x_0 \in M$ be the 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-convex in a neighborhood U of x_0. If $U = M$, the theorem is proved.

If U is strictly contained in M, denote V the complement in M of the topological closure of U. The restriction of f to the open set V is regular; hence, by the Remark 5, (iii) 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 (implicitely supposed) paracompactness of the manifold M, there exist an open sub-covering $\{W_i\}_i$ of the open covering $\{U, V\}$ 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 ∇^1 or ∇^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 conection 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) functions.

Secondly, the function f is ∇-convex. Indeed, let remark that if ∇' and ∇" are two linear connections and h a differentiable function, then we may construct a new connection $\overline{\nabla} = h\nabla' + (1-h)\nabla$". Consider another function α; then the Hessian of α with respect to $\overline{\nabla}$ 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.

As all the functions from the partition of unity have the previous property, it follows that the function f is convex with respect with the new constructed connection ∇.

Combining the theorem with the Remark 5, (iii), we get the

Corollary. *Let f be a real valued differentiable function on a differentiable manifold M.Suppose f is regular or has only a critical point, which is of minimum type. Then there exists a linear connection in $\mathcal{C}_f$.*

5 Other example

We may wonder if, in the hypothesis of Theorem 6, the uniqueness of the critical point is necessary for the generalized convexity of the function f. As the following example shows, the respective result seems true in a more general context, where several (eventualy an infinity of) critical points exist (but none of them may be a maximum one !).

Consider the differentiable function $f_4 : \mathbf{R}^2 \to \mathbf{R}$, given by $f_4(x, y) = x^2 e^y$. This function is not (classically) convex, because its (classical) Hessian is indefinite in the point (1,0).

Let ∇ be a linear connection, whose components (except some vanishing ones) are:

$$\Gamma^1_{21} = \Gamma^1_{12} = 1 \quad , \quad \Gamma^2_{11} \leq 0 \quad , \quad \Gamma^2_{22} \leq 0$$

A direct computation of the Hessian (with respect to ∇) shows that f_4 is ∇-convex. In this case, the choosen function has an infinity of minimum points $(0, y)$, with $y \in \mathbf{R}$, but no other critical point. The shape of f_4 may be seen below. (Even it is not visually obvious, the lack of classical convexity follows as previously asserted).

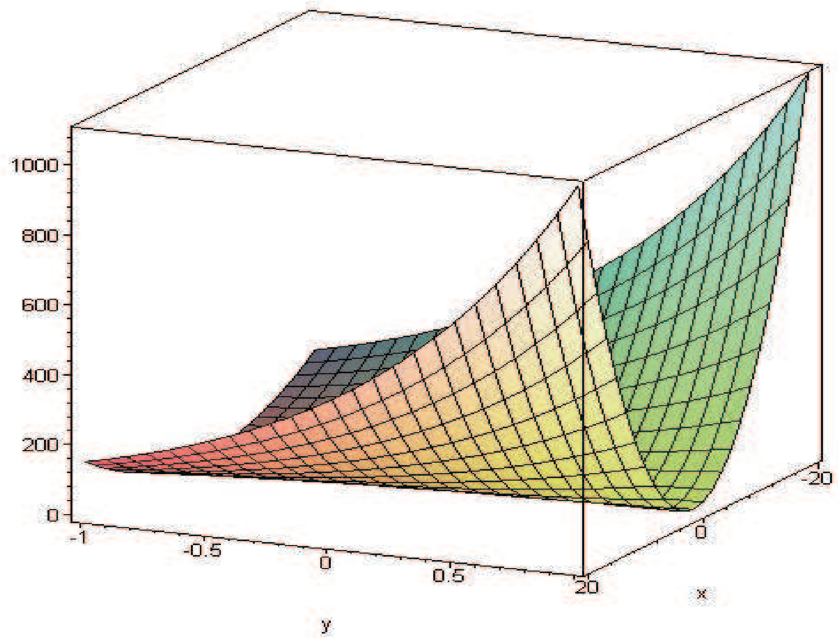

Figure 4: Graph of f_4

References

[1] C.L. Pripoae, Affine differential invariants associated to convex functions, Proc. Conf. S.S.M.R, Cluj-Napoca (May 1998), Digital Data Publ. House (Cluj-Napoca, 1999), 247-252

[2] C.L. Pripoae, G.T. Pripoae, General descent algorithm on affine manifolds, in Gr. Tsagas ed., Proc. Workshop in Global Analysis, Diff. Geom. and Lie Algebras, (Salonic, 1998), Balkan Geometry Press, 1999, 175-179

[3] C.L. Pripoae, New conditions for affine differential convexity, An. Univ. Bucuresti, s. Mat.-Info., LI (2002), no.1, 127-132

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

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