We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900
Open access books available

185,000
International authors and editors

200M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com



Chapter 7

Bilevel Disjunctive Optimization on Affine Manifolds

Constantin Udriste, Henri Bonnel, Ionel Tevy and
Ali Sapeeh Rasheed

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/intechopen.75643

Abstract

Bilevel optimization is a special kind of optimization where one problem is embedded within another. The outer optimization task is commonly referred to as the upper-level optimization task, and the inner optimization task is commonly referred to as the lower-level optimization task. These problems involve two kinds of variables: upper-level variables and lower-level variables. Bilevel optimization was first realized in the field of game theory by a German economist von Stackelberg who published a book (1934) that described this hierarchical problem. Now the bilevel optimization problems are commonly found in a number of real-world problems: transportation, economics, decision science, business, engineering, and so on. In this chapter, we provide a general formulation for bilevel disjunctive optimization problem on affine manifolds. These problems contain two levels of optimization tasks where one optimization task is nested within the other. The outer optimization problem is commonly referred to as the leaders (upper level) optimization problem and the inner optimization problem is known as the followers (or lower level) optimization problem. The two levels have their own objectives and constraints. Topics affine convex functions, optimizations with auto-parallel restrictions, affine convexity of posynomial functions, bilevel disjunctive problem and algorithm, models of bilevel disjunctive programming problems, and properties of minimum functions.

Keywords: convex programming, affine manifolds, optimization along curves, bilevel disjunctive optimization, minimum functions

Mathematics Subject Classification 2010: 90C25 90C29, 90C30

IntechOpen

© 2018 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/3.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. CC BY

1. Affine convex functions

In optimization problems [16, 17, 19, 23–27], one can use an *affine manifold* as a pair (M,Γ), where M is a smooth real n-dimensional manifold, and Γ is an affine symmetric connection on M. The connection Γ produces auto-parallel curves $x(t)$ via ODE system

$$\ddot{x}^h(t)+\Gamma^h_{ij}(x(t))\dot{x}^i(t)\dot{x}^j(t)=0.$$

They are used for defining the convexity of subsets in M and convexity of functions $f: D\subset M\to\mathbb{R}$ (see also [3, 6]).

Definition 1.1 *An affine manifold (M,Γ) is called autoparallely complete if any auto-parallel $x(t)$ starting at $p\in M$ is defined for all values of the parameter $t\in\mathbb{R}$.*

Theorem 1.1 [1] *Let M be a (Hausdorff, connected, smooth) compact n-manifold endowed with an affine connection Γ and let $p\in M$. If the holonomy group $Hol_p(\Gamma)$ (regarded as a subgroup of the group $Gl(T_pM)$ of all the linear automorphisms of the tangent space T_pM) has compact closure, then (M,Γ) is autoparallely complete.*

Let (M,Γ) be an auto-parallely complete affine manifold. For a C^2 function $f: M\to\mathbb{R}$, we define the tensor $Hess_\Gamma f$ of components

$$(Hess_\Gamma f)_{ij}=\frac{\partial^2 f}{\partial x^i\partial x^j}-\Gamma^h_{ij}\frac{\partial f}{\partial x^h}.$$

Definition 1.2 *A C^2 function $f: M\to\mathbb{R}$ is called:*

(1) linear affine with respect to Γ if $Hess_\Gamma f=0$, throughout;

(2) affine convex (convex with respect to Γ) if $Hess_\Gamma f\geqslant 0$ (positive semidefinite), throughout.

The function f is: (1) linear affine if its restriction $f(x(t))$ on each autoparallel $x(t)$ satisfies $f(x(t))=at+b$, for some numbers a, b that may depend on $x(t)$; (2) affine convex if its restriction $f(x(t))$ is convex on each auto-parallel $x(t)$.

Theorem 1.2 *If there exists a linear affine nonconstant function f on (M,Γ), then the curvature tensor field $R^h{}_{ikj}$ is in* Ker df.

Proof. For given Γ, if we consider

$$\frac{\partial^2 f}{\partial x^i\partial x^j}=\Gamma^h_{ij}\frac{\partial f}{\partial x^h}$$

as a PDEs system (a particular case of a Frobenius-Mayer system of PDEs) with $\frac{1}{2}n(n+1)$ equations and the unknown function f, then we need the complete integrability conditions

$$\frac{\partial^3 f}{\partial x^i \partial x^j \partial x^k} = \frac{\partial^3 f}{\partial x^k \partial x^i \partial x^j}.$$

Since,

$$\frac{\partial^3 f}{\partial x^i \partial x^j \partial x^k} = \left(\frac{\partial \Gamma^h_{ij}}{\partial x^k} + \Gamma^l_{ij} \Gamma^h_{kl} \right) \frac{\partial f}{\partial x^h},$$

it follows

$$\frac{\partial f}{\partial x^h} R^h{}_{ikj} = 0, R^h{}_{ikj} = \frac{\partial \Gamma^h_{ij}}{\partial x^k} - \frac{\partial \Gamma^h_{ki}}{\partial x^j} + \Gamma^l_{ij} \Gamma^h_{kl} - \Gamma^l_{ki} \Gamma^h_{jl}.$$

Corollary 1.1 *If there exists n linear affine functions $f_l, l = 1, \ldots, n$ on (M, Γ), whose df_l are linearly independent, then Γ is flat, that is, $R^h{}_{ikj} = 0$.*

Of course this only means the curvature tensor is zero on the topologically trivial region we used to set up our co-vector fields $df_l(x)$. But we can always cover any manifold by an atlas of topologically trivial regions, so this allows us to deduce that the curvature tensor vanishes throughout the manifold.

Remark 1.1 *There is actually no need to extend $df_l(x)$ to the entire manifold. If this could be done, then $df_l(x)$ would now be everywhere nonzero co-vector fields; but there are topologies, for example, S^2, for which we know such things do not exist. Therefore, there are topological manifolds for which we are forced to work on topologically trivial regions.*

The following theorem is well-known [16, 17, 19, 23]. Due to its importance, now we offer new proofs (based on catastrophe theory, decomposing a tensor into a specific product, and using slackness variables).

Theorem 1.3 *Let $f : M \to \mathbb{R}$ be a C^2 function.*

(1) If f is regular or has only one minimum point, then there exists a connection Γ such that f is affine convex.

(2) If f has a maximum point x_0, then there is no connection Γ making f affine convex throughout.

Proof. For the Hessian $(Hess_\Gamma f)_{ij}$ be positive semidefinite, we need n conditions like inequalities and equalities. The number of unknowns Γ^h_{ij} is $\frac{1}{2}n^2(n+1)$. The inequalities can be replaced by equalities using slackness variables.

The first central idea for the proof is to use the catastrophe theory, since almost all families $f(x,c)$, $x=\left(x^1,\ldots,,x^n\right)\in\mathbb{R}^n$, $c=(c_1,\ldots,c_m)\in\mathbb{R}^m$, of real differentiable functions, with $m\leq 4$ parameters, are structurally stable and are equivalent, in the vicinity of any point, with one of the following forms [15]:

We eliminate the case with maximum point, that is., Morse 0-saddle and the saddle point. Around each critical point (in a chart), the canonical form $f(x,c)$ is affine convex, with respect to appropriate locally defined linear connections that can be found easily. Using change of coordinates and the partition of unity, we glue all these connections to a global one, making $f(x,c)$ affine convex on M.

At any critical point x_0, the affine Hessian $Hess_\Gamma f$ is reduced to Euclidean Hessian, $\frac{\partial^2 f}{\partial x^i \partial x^j}(x_0)$. Then the maximum point condition or the saddle condition is contradictory to affine convexity condition.

A direct proof based on decomposition of a tensor: Let (M,Γ) be an affine manifold and $f:M\to\mathbb{R}$ be a C^2 function.

Suppose f has no critical points (is regular). If the function f is not convex with respect to Γ, we look to find a new connection $\overline{\Gamma}_{ij}^h=\Gamma_{ij}^h+T_{ij}^h$, with the unknown a tensor field T_{ij}^h, such that

$$\frac{\partial^2 f}{\partial x^i \partial x^j}(x)-\overline{\Gamma}_{ij}^h(x)\frac{\partial f}{\partial x^h}(x)=\sigma_{ij}(x), x\in M,$$

where $\sigma_{ij}(x)$ is a positive semi-definite tensor. A very particular solution is the decomposition $T_{ij}^h(x)=a^h(x)b_{ij}(x)$, where the vector field a has the property

$$D_a f=a^h(x)\frac{\partial f}{\partial x^h}(x)\neq 0, x\in M$$

and the tensor b_{ij} is

$$b_{ij}(x)=\frac{1}{D_a f}\left(\frac{\partial^2 f}{\partial x^i \partial x^j}(x)-\Gamma_{ij}^h(x)\frac{\partial f}{\partial x^h}(x)-\sigma_{ij}(x)\right), x\in M.$$

Remark 1.2 *The connection $\overline{\Gamma}_{ij}^h$ is strongly dependent on both the function f and the tensor field σ_{ij}.*

Suppose f has a minimum point x_0. In this case, observe that we must have the condition $\sigma_{ij}(x_0)=\frac{\partial^2 f}{\partial x^i \partial x^j}(x_0)$. Can we make the previous reason for $x\neq x_0$ and then extend the obtained connection by continuity? The answer is generally negative. Indeed, let us compute

$$b_{ij}(x_0)=\lim_{x\to x_0}\frac{1}{D_a f}\left(\frac{\partial^2 f}{\partial x^i \partial x^j}(x)-\Gamma_{ij}^h(x)\frac{\partial f}{\partial x^h}(x)-\sigma_{ij}(x)\right).$$

Here we cannot plug in the point x_0 because we get $\frac{0}{0}$, an indeterminate form.

To contradict, we fix an auto-parallel $\gamma(t), t \in [0, \epsilon)$, starting from minimum point $x_0 = \gamma(0)$, tangent to $\dot{\gamma}(0) = v$ and we compute (via l'Hôpital rule)

$$b_{ij}(x_0; v) = \lim_{t \to 0} b_{ij}(\gamma(t)) = \frac{\left(\frac{\partial^3 f}{\partial x^i \partial x^j \partial x^k}(x_0) - \Gamma_{ij}^h(x_0) \frac{\partial^2 f}{\partial x^h \partial x^k}(x_0) - \frac{\partial \sigma_{ij}}{\partial x^k}(x_0)\right) v^k}{a^h(x_0) \frac{\partial^2 f}{\partial x^h \partial x^k}(x_0) v^k}.$$

But this result depends on the direction v (different values along two different auto-parallels).

In some particular cases, we can eliminate the dependence on the vector v. For example, the conditions

$$\frac{\partial^3 f}{\partial x^i \partial x^j \partial x^l}(x_0) - \Gamma_{ij}^h(x_0) \frac{\partial^2 f}{\partial x^h \partial x^l}(x_0) - \frac{\partial \sigma_{ij}}{\partial x^l}(x_0)$$

$$= \rho \left(\frac{\partial^3 f}{\partial x^i \partial x^j \partial x^k}(x_0) - \Gamma_{ij}^h(x_0) \frac{\partial^2 f}{\partial x^h \partial x^k}(x_0) - \frac{\partial \sigma_{ij}}{\partial x^k}(x_0) \right),$$

$$a^h(x_0) \frac{\partial^2 f}{\partial x^h \partial x^l}(x_0) = \rho \left(a^h(x_0) \frac{\partial^2 f}{\partial x^h \partial x^k}(x_0) \right)$$

are sufficient to do this.

A particular condition for independence on v is

$$\frac{\partial^3 f}{\partial x^i \partial x^j \partial x^k}(x_0) - \Gamma_{ij}^h(x_0) \frac{\partial^2 f}{\partial x^h \partial x^k}(x_0) - \frac{\partial \sigma_{ij}}{\partial x^k}(x_0) = 0.$$

In this particular condition, we can show that we can build connections of previous type good everywhere.

1.1. Lightning through examples

Let us lightning our previous statements by the following examples.

Example 1.1 *(for the first part of the theorem) Let us consider the function* $f : \mathbf{R}^2 \to \mathbf{R}, f(x, y) = x^3 + y^3 + 3x + 3y$ *and* $\Gamma_{ij}^h = 0, i, j, h = 1, 2$. *Then* $\frac{\partial f}{\partial x} = 3x^2 + 3$, $\frac{\partial f}{\partial y} = 3y^2 + 3$ *and* f *has no critical point. Moreover, the Euclidean Hessian of* f *is not positive semi-definite overall. Let us make the above construction for* $\sigma_{ij}(x, y) = \delta_{ij}$. *Taking* $a^1 = a^2 = 1$, *we obtain the connection*

$$\overline{\Gamma}_{11}^h = \frac{6x - 1}{3x^2 + 3y^2 + 2}, \overline{\Gamma}_{22}^h = \frac{6y - 1}{3x^2 + 3y^2 + 2}, \overline{\Gamma}_{12}^h = \overline{\Gamma}_{21}^h = 0, \ h = 1, 2,$$

that is not unique.

Example 1.2 *(for one minimum point) Let us consider the function* $f : \mathbf{R}^2 \to \mathbf{R}, f(x, y) = 1 - e^{-}$ $\left(x^2 + y^2\right)$ *and* $\Gamma_{ij}^h = 0, i, j, h = 1, 2$. *Then* $\frac{\partial f}{\partial x} = 2xe^{-\left(x^2+y^2\right)}$, $\frac{\partial f}{\partial y} = 2ye^{-\left(x^2+y^2\right)}$ *and* f *has a unique critical*

minimum point $(0,0)$. *However, the Euclidean Hessian of* f *is not positive semi-definite overall. We make previous reason for* $\sigma_{ij} = 2e^{-\left(x^2+y^2\right)}\delta_{ij}$, $a^1 = \frac{\partial f}{\partial x}$, $a^2 = \frac{\partial f}{\partial y}$ *Hence we obtain* $\overline{\Gamma}^h_{ij} = T^h_{ij}$,

$$\overline{\Gamma}^1_{11} = -\frac{2x^3}{x^2+y^2}, \overline{\Gamma}^1_{12} = \overline{\Gamma}^1_{21} = -\frac{2x^2y}{x^2+y^2}, \overline{\Gamma}^1_{22} = -\frac{2xy^2}{x^2+y^2},$$

$$\overline{\Gamma}^2_{11} = -\frac{2x^2y}{x^2+y^2}, \overline{\Gamma}^2_{12} = \overline{\Gamma}^2_{21} = -\frac{2xy^2}{x^2+y^2}, \overline{\Gamma}^2_{22} = -\frac{2y^3}{x^2+y^2}.$$

Observe that $\lim_{(x,y)\to(0,0)} T^h_{ij}(x,y) = 0$. *Hence take* $\overline{\Gamma}^h_{ij}(0,0) = 0$.

The next example shows what happens if we come out of the conditions of the previous theorem.

Example 1.3 *Let us take the function* $f : \mathbb{R} \to \mathbb{R}, f(x) = x^3$, *where the critical point* $x = 0$ *is an inflection point. We take* $\Gamma(x) = -1 - \frac{2}{x^2}$, *which is not defined at the critical point* $x = 0$, *but the relation of convexity is realized by prolongation,*

$$\sigma(x) = f''(x) - \Gamma(x)f'(x) = 3\left(x^2 + 2x + 2\right) > 0, \quad \forall x \in \mathbb{R}.$$

Let us consider the ODE of auto-parallels

$$x''(t) - \left(1 + \frac{2}{t^2}\right)x'(t)^2 = 0, t \neq 0.$$

The solutions

$$x(t) = -\frac{1}{2}\ln|-2 + t^2 - ct| + \frac{c}{\sqrt{8+c^2}}\ \text{arctanh}\ \frac{2t-c}{\sqrt{8+c^2}} + c_1$$

are auto-parallels on $(\mathbb{R}\setminus\{0, t_1, t_2\}, \Gamma)$, *where* t_1, t_2 *are real solutions of* $-2 + t^2 - ct = 0$. *These curves are extended at* $t = 0$ *by continuity. The manifold* $(\mathbb{R}, \Gamma)$ *is not auto-parallely complete. Since the image* $x(\mathbb{R})$ *is not a "segment", the function* $f : \mathbb{R} \to \mathbb{R}, f(x) = x^3$ *is not globally convex.*

Remark 1.3 *For* $n \geq 2$, *there exists* C^1 *functions* $\varphi : \mathbb{R}^n \to \mathbb{R}$ *which have two minimum points without having another extremum point. As example,*

$$\varphi\left(x^1, x^2\right) = \left(x^{1^2} - 1\right)^2 + \left(x^{1^2}x^2 - x^1 - 1\right)^2$$

has two (global) minimum points $p = (-1, 0)$, $q = (1, 2)$.

The restriction

$$\varphi\left(x^1, x^2\right) = \left(x^{1^4} + x^{1^4}x^{2^2} + 2x^1 + 2\right) - \left(x^{1^2} + 2x^{1^3}x^2 + 2x^{1^2}x^2\right),\ x^1 > 0, x^2 > 0$$

is difference of two affine convex functions (see Section 2).

Our chapter is based also on some ideas in: [3] (convex mappings between Riemannian manifolds), [7] (geometric modeling in probability and statistics), [13] (arc length in metric and Finsler manifolds), [14] (applications of Hahn-Banach principle to moment and optimization problems), [21] (geodesic connectedness of semi-Riemannian manifolds), and [28] (tangent and cotangent bundles). For algorithms, we recommend the paper [20] (sequential and parallel algorithms).

2. Optimizations with autoparallel restrictions

2.1. Direct theory

The auto-parallel curves $x(t)$ on the affine manifold (M, Γ) are solutions of the second order ODE system

$$\ddot{x}^h(t) + \Gamma^h_{ij}(x(t))\dot{x}^i(t)\dot{x}^j(t) = 0, x(t_0) = x_0, \dot{x}(t_0) = \xi_0.$$

Obviously, the complete notation is $x(t, x_0, \xi_0)$, with

$$x(t_0, x_0, \xi_0) = x_0, \dot{x}(t_0, x_0, \xi_0) = \xi_0.$$

Definition 2.1 *Let $D \subset M$ be open and connected and $f : D \to \mathbb{R}$ a C^2 function. The point $x_0 \in D$ is called minimum (maximum) point of f conditioned by the auto-parallel system, together with initial conditions, if for the maximal solution $x(t, x_0, \xi_0) : I \to D$, there exists a neighborhood I_{t_0} of t_0 such that*

$$f(x(t, x_0, \xi_0)) \geq \ (\leq)\ f(x_0), \ \ \forall t \in I_{t_0} \subset I.$$

Theorem 2.1 *If $x_0 \in D$ is an extremum point of f conditioned by the previous second order system, then $df(x_0)(\xi_0) = 0$.*

Definition 2.2 *The points $x \in D$ which are solutions of the equation $df(x)(\xi) = 0$ are called critical points of f conditioned by the previous spray.*

Theorem 2.2 *If $x_0 \in D$ is a conditioned critical point of the function $f : D \to \mathbb{R}$ of class C^2 constrained by the previous auto-parallel system and if the number*

$$(Hess f)_{ij}\, \xi^i_0 \xi^j_0 = \left(\frac{\partial^2 f}{\partial x^i \partial x^j} - \frac{\partial f}{\partial x^h}\, \Gamma^h_{ij}\right)(x_0)\, \xi^i_0 \xi^j_0$$

is strictly positive (negative), then x_0 is a minimum (maximum) point of f constrained by the auto-parallel system.

Example 2.1 *We compute the Christoffel symbols on the unit sphere S^2, using spherical coordinates (θ, φ) and the Riemannian metric*

$$g_{\theta\theta} = 1, g_{\theta\varphi} = g_{\varphi\theta} = 0, g_{\varphi\varphi} = \sin^2\theta.$$

When $\theta \neq 0, \pi$, we find

$$\Gamma^{\theta}_{\varphi\varphi} = -\frac{1}{2}\sin 2\theta, \Gamma^{\varphi}_{\varphi\theta} = \Gamma^{\varphi}_{\theta\varphi} = \cot\theta,$$

and all the other Γs are equal to zero. We can show that the apparent singularity at $\theta = 0, \pi$ can be removed by a better choice of coordinates at the poles of the sphere. Thus, the above affine connection extends to the whole sphere.

The second order system defining auto-parallel curves (geodesics) on S^2 are

$$\ddot{\theta}(t) - \frac{1}{2}\sin 2\theta(t)\,\dot{\varphi}(t)\dot{\varphi}(t) = 0, \ddot{\varphi}(t) - 2\cot\theta(t)\,\dot{\varphi}(t)\dot{\theta}(t) = 0.$$

The solutions are great circles on the sphere. For example, $\theta = \alpha\, t + \beta$ and φ = const.

We compute the curvature tensor R of the unit sphere S^2. Since there are only two independent coordinates, all the non-zero components of curvature tensor R are given by $R^i_j = R^i_{j\theta\varphi} = -R^i_{j\varphi\theta}$, where $i, j = \theta, \varphi$. We get $R^{\theta}_{\varphi} = \sin^2\theta, R^{\varphi}_{\theta} = -1$ and the other components are 0.

Let $\left(\theta\left(t, \left(\theta_0, \varphi_0\right), \xi\right), \varphi\left(t, \left(\theta_0, \varphi_0\right), \xi\right), t \in R$ be the maximal auto-parallel which satisfies $\theta\left(t_0, \left(\theta_0, \varphi_0\right), \xi\right) = \theta_0$, $\dot{\theta}\left(t_0, \left(\theta_0, \varphi_0\right), \xi\right) = \xi^1$; $\varphi\left(t_0, \left(\theta_0, \varphi_0\right), \xi\right) = \varphi_0$, $\dot{\varphi}\left(t_0, \left(\theta_0, \varphi_0\right), \xi\right) = \xi^2$. We wish to compute $\min f(\theta, \varphi) = R^{\theta}_{\varphi} = \sin^2\theta$ with the restriction $\left(\theta\left(t, \left(\theta_0, \varphi_0\right), \xi\right), \varphi\left(t, \left(\theta_0, \varphi_0\right), \xi\right)\right), t \in R$.

Since $df = (2\sin\theta\cos\theta, 0)$, the critical point condition $df(\theta, \varphi)(\xi) = 0$ becomes $\sin\theta\cos\theta\,\xi^1 = 0$. Consequently, the critical points are either $(\theta_0 = k\pi, k \in \mathbb{Z}; \varphi)$, $\left(\xi^1, \xi^2\right) \neq (0, 0)$, or $\left(\theta_1 = (2k+1)\frac{\pi}{2}, k \in \mathbb{Z}; \varphi\right)$, $\left(\xi^1, \xi^2\right) \neq (0, 0)$, or $(\theta; \varphi)$, $\left(\xi^1 = 0; \xi^2 \neq 0\right)$.

The components of the Hessian of f are

$$(Hessf)_{\theta\theta} = \frac{\partial^2 f}{\partial\theta\partial\theta} = 2\cos 2\theta, (Hessf)_{\theta\varphi} = 0, (Hessf)_{\varphi\varphi} = \frac{1}{2}\sin^2 2\theta.$$

At the critical points (θ_0, φ) or (θ_1, φ), the Hessian of f is positive or negative semi-definite. On the other hand, along $\left(\xi^1 = 0, \xi^2 \neq 0\right)$, we find $(Hessf)_{ij}\,\xi^i\xi^j = \frac{1}{2}\sin^2 2\theta\left(\xi^2\right)^2 > 0, \xi^2 \neq 0$. Consequently, each point $\left(\theta \neq \frac{k\pi}{2}, \varphi\right)$, is a minimum point of f along each auto-parallel, starting from given point and tangent to $\left(\xi^1 = 0, \xi^2 \neq 0\right)$.

2.2. Theory via the associated spray

This point of view regarding extrema comes from paper [22].

The second order system of auto-parallels induces a spray (special vector field) $Y(x,y) = \left(y^h, \Gamma^h_{ij}(x)y^i y^j\right)$ on the tangent bundle TM, that is,

$$\dot{x}^h(t) = y^h(t), \dot{y}^h(t) + \Gamma^h_{ij}(x(t))y^i(t)y^j(t) = 0.$$

The solutions $\gamma(t) = (x(t), y(t)) : I \to D$ of class C^2 are called field lines of Y. They depend on the initial condition $\gamma(t)|_{t=t_0} = (x_0, y_0)$, and therefore the notation $\gamma(t, x_0, y_0)$ is more suggestive.

Definition 2.3 *Let $D \subset TM$ be open and connected and $f : D \to \mathbb{R}$ a C^2 function. The point $(x_0, y_0) \in D$ is called minimum (maximum) point of f conditioned by the previous spray, if for the maximal field line $\gamma(t, x_0, y_0)$, $t \in I$, there exists a neighborhood I_{t_0} of t_0 such that*

$$f(\gamma(t, x_0, y_0)) \geq \ (\leq)\ f(x_0, y_0), \ \ \forall t \in I_{t_0} \subset I.$$

Theorem 2.3 *If $(x_0, y_0) \in D$ is an extremum point of f conditioned by the previous spray, then (x_0, y_0) is a point where Y is in* $\operatorname{Ker} df$.

Definition 2.4 *The points $(x, y) \in D$ which are solutions of the equation*

$$D_Y f(x,y) = df(Y)(x,y) = 0$$

are called critical points of f conditioned by the previous spray.

Theorem 2.4 *If $(x_0, y_0) \in D$ is a conditioned critical point of the function $f : D \to \mathbb{R}$ of class C^2 constrained by the previous spray and if the number*

$$\left(d^2 f(Y,Y) + df(D_Y Y)\right)(x_0, y_0)$$

is strictly positive (negative), then (x_0, y_0) is a minimum (maximum) point of f constrained by the spray.

Example 2.2 *We consider the Volterra-Hamilton ODE system* [2].

$$\frac{dx^1}{dt}(t) = y^1(t), \frac{dx^2}{dt}(t) = y^2(t),$$

$$\frac{dy^1}{dt}(t) = \lambda y^1(t) - \alpha_1 y^{1^2}(t) - 2\alpha_2 y^1(t)y^2(t),$$

$$\frac{dy^2}{dt}(t) = \lambda y^1(t) - \beta_1 y^{2^2}(t) - 2\beta_2 y^1(t)y^2(t),$$

which models production in a Gause-Witt 2-species evolving in $\mathbb{R}^4$*: (1) competition if* $\alpha_1 > 0$, $\alpha_2 > 0$, $\beta_1 > 0$, $\beta_2 > 0$ *and (2) parasitism if* $\alpha_1 > 0$, $\alpha_2 < 0$, $\beta_1 > 0$, $\beta_2 > 0$.

Changing the real parameter t into an affine parameter s, we find the connection with constant coefficients

$$\Gamma^1_{11} = \frac{1}{3}(\alpha_1 - 2\beta_2), \Gamma^2_{22} = \frac{1}{3}(\beta_1 - 2\alpha_2),$$

$$\Gamma^1_{12} = \frac{1}{3}(2\alpha_2 - \beta_1), \Gamma^2_{12} = \frac{1}{3}(2\beta_2 - \alpha_1).$$

Let $x(t, x_0, y_0), t \in I$ *be the maximal field line which satisfies* $x(t_0, x_0, y_0) = (x_0, y_0)$. *We wish to compute* $\max f(x^1, x^2, y^1, y^2) = y^2$ *with the restriction* $x = x(t, x_0, y_0)$.

We apply the previous theory. Introduce the vector field

$$Y = \left(y^1, y^2, \lambda y^1 - \alpha_1 y^{1^2} - 2\alpha_2 y^1 y^2, \lambda y^1 - \beta_1 y^{2^2} - 2\beta_2 y^1 y^2\right).$$

We set the critical point condition $df(Y) = 0$. *Since* $df = (0, 0, 0, 1)$, *it follows the relation* $\lambda y^1 - \beta_1 y^{2^2} - 2\beta_2 y^1 y^2 = 0$, *that is, the critical point set is a conic in* $y^1 O y^2$.

Since $d^2 f = 0$, *the sufficiency condition is reduced to* $df(D_Y Y)(x_0, y_0) < 0$, *that is,*

$$\left(\lambda - \frac{\alpha_1 \beta_1 y^{22}}{\lambda - 2\beta_2 y^2} - 2\alpha_2 y^2\right)(y_0) < 0.$$

This last relation is equivalent either to

$$(\lambda - 2\alpha_2 y_0^2)(\lambda - 2\beta_2 y_0^2) - \alpha_1 \beta_1 y_0^2 2 < 0, \lambda - 2\beta_2 y_0^2 > 0$$

or to

$$(\lambda - 2\alpha_2 y_0^2)(\lambda - 2\beta_2 y_0^2) - \alpha_1 \beta_1 y_0^2 2 > 0, \lambda - 2\beta_2 y_0^2 < 0.$$

Each critical point satisfying one of the last two conditions is a maximum point.

3. Affine convexity of posynomial functions

For the general theory regarding geometric programming (based on posynomial, signomial functions, etc.), see [11].

Theorem 3.1 *Each posynomial function is affine convex, with respect to some affine connection.*

Proof. A posynomial function has the form

$$f : \mathbb{R}^n_{++} \to \mathbb{R},\ f(x) = \sum_{k=1}^{K} c_k \prod_{i=1}^{n} \left(x^i\right)^{a_{ik}},$$

where all the coefficients c_k are positive real numbers, and the exponents a_{ik} are real numbers. Let us consider the auto-parallel curves of the form

$$\gamma(t) = \left(\left(a^1\right)^{1-t}\left(b^1\right)^t, \left(a^2\right)^{1-t}\left(b^2\right)^t, \ldots, \left(a^n\right)^{1-t}\left(b^n\right)^t\right), t \in [0,1],$$

joining the points $a = \left(a^1, \ldots, a^n\right)$ and $b = \left(b^1, \ldots, b^n\right)$, which fix, as example, the affine connection

$$\Gamma^h_{hj} = \Gamma^h_{jh} = -\frac{1}{2}\frac{\mu^h}{\mu^j x^j}, \text{and otherwise } \Gamma^h_{ij} = 0.$$

It follows

$$f(\gamma(t)) = \sum_{k=1}^{K} c_k \prod_{i=1}^{n} \left(\left(a^i\right)^{a_{ik}}\right)^{1-t} \left(\left(b^i\right)^{a_{ik}}\right)^t$$

$$= \sum_{k=1}^{K} c_k \left(\prod_{i=1}^{n} \left(a^i\right)^{a_{ik}}\right)^{1-t} \left(\prod_{i=1}^{n} \left(b^i\right)^{a_{ik}}\right)^t.$$

One term in this sum is of the form $\psi_k(t) = A_k^{1-t}B_k^t$, and hence $\ddot{\psi}_k(t) = A_k^{1-t}B_k^t(\ln A_k - \ln B_k)^2 > 0$.

Remark 3.1 *Posynomial functions belong to the class of functions satisfying the statement "product of two convex function is convex".*

Corollary 3.1 *Each signomial function is difference of two affine convex posynomials, with respect to some affine connection.*

Proof. A signomial function has the form

$$f : \mathbb{R}^n_{++} \to \mathbb{R},\ f(x) = \sum_{k=1}^{K} c_k \prod_{i=1}^{n} \left(x^i\right)^{a_{ik}},$$

where all the exponents a_{ik} are real numbers and the coefficients c_k are either positive or negative. Without loss of generality, suppose that for $k = 1, \ldots, k_0$ we have $c_k > 0$ and for $k = k_0 + 1, \ldots, K$ we have $c_k < 0$. We use the decomposition

$$f(x) = \sum_{k=1}^{k_0} c_k \prod_{i=1}^{n} \left(x^i\right)^{a_{ik}} - \sum_{k=k_0+1}^{K} |c_k| \prod_{i=1}^{n} \left(x^i\right)^{a_{ik}},$$

we apply the Theorem and the implication $u''(t) \geq v''(t) \Rightarrow u - v$ convex. □

Corollary 3.2 *(1) The polynomial functions with positive coefficients, restricted to* $\mathbb{R}^n_{++}$*, are affine convex functions.*

(2) The polynomial functions with positive and negative terms, restricted to $\mathbb{R}^n_{++}$*, are differences of two affine convex functions.*

Proudnikov [18] gives the necessary and sufficient conditions for representing Lipschitz multivariable function as a difference of two convex functions. An algorithm and a geometric interpretation of this representation are also given. The outcome of this algorithm is a sequence of pairs of convex functions that converge uniformly to a pair of convex functions if the conditions of the formulated theorems are satisfied.

4. Bilevel disjunctive problem

Let $\left(M_1, {}^1\Gamma\right)$, the *leader decision affine manifold*, and $\left(M_2, {}^2\Gamma\right)$, the *follower decision affine manifold*, be two connected affine manifolds of dimension n_1 and n_2, respectively. Moreover, $\left(M_2, {}^2\Gamma\right)$ is supposed to be complete. Let also $f : M_1 \times M_2 \to \mathbb{R}$ be the *leader objective function*, and let $F = (F_1, \ldots, F_r) : M_1 \times M_2 \to \mathbb{R}^r$ be the *follower multiobjective function*.

The components $F_i : M_1 \times M_2 \to \mathbb{R}$ are (possibly) conflicting objective functions.

A *bilevel optimization problem* means a decision of leader with regard to a multi-objective optimum of the follower (in fact, a constrained optimization problem whose constraints are obtained from optimization problems). For details, see [5, 10, 12].

Let $x \in M_1$, $y \in M_2$ be the generic points. In this chapter, the *disjunctive solution set of a follower multiobjective optimization problem* is defined by

(1) the set-valued function

$$\psi : M_1 \rightrightarrows M_2, \psi(x) = \operatorname{Argmin}_{y \in M_2} F(x, y),$$

where

$$\operatorname{Argmin}_{y \in M_2} F(x, y) := \cup_{i=1}^{r} \operatorname{Argmin}_{y \in M_2} F_i(x, y)$$

or

(2) the set-valued function

$$\psi : M_1 \rightrightarrows M_2,\ \psi(x) = \operatorname{Argmax}_{y \in M_2} F(x,y),$$

where

$$\operatorname{Argmax}_{y \in M_2} F(x,y) := \cup_{i=1}^{r} \operatorname{Argmax}_{y \in M_2} F_i(x,y).$$

We deal with two bilevel problems:

(1) The *optimistic bilevel disjunctive problem*

$$(OBDP) \quad \min_{x \in M_1} \min_{y \in \psi(x)} f(x,y).$$

In this case, the follower cooperates with the leader; that is, for each $x \in M_1$, the follower chooses among all its disjunctive solutions (his best responses) one which is the best for the leader (assuming that such a solution exists).

(2) The *pessimistic bilevel disjunctive problem*

$$(PBDP) \quad \min_{x \in M_1} \max_{y \in \psi(x)} f(x,y).$$

In this case, there is no cooperation between the leader and the follower, and the leader expects the worst scenario; that is, for each $x \in M_1$, the follower may choose among all its disjunctive solutions (his best responses) one which is unfavorable for the leader.

So, a general optimization problem becomes a pessimistic bilevel problem.

Theorem 4.1 *The value*

$$\min_x [f(x,y) : y \in \psi(x)]$$

exists if and only if, for an index i, the minimum $\min_x \left[f(x,y) : y \in \psi_i(x)\right]$ *exists and, for each* $j \neq i$, *either* $\min_x \left[f(x,y) : y \in \psi_j(x)\right]$ *exists or* $\psi_j = \emptyset$. *In this case,*

$$\min_x [f(x,y) : y \in \psi(x)]$$

coincides to the minimum of minima that exist.

Proof. Let us consider the multi-functions $\phi_i(x) = f\left(x, \psi_i(x)\right)$ and $\phi(x) = f(x, \psi(x))$. Then $\phi(x) = \cup_{i=1}^{k} \phi_i(x)$. It follows that $\min_x \phi(x)$ exists if and only if either $\min_x \phi_i(x)$ exists or $\psi_i = \emptyset$, and at least one minimum exists.

Taking minimum of minima that exist, we find

$$\min_x [f(x,y) : y \in \psi(x)]. \qquad \square$$

Theorem 4.2 *Suppose M_1 is a compact manifold. If for each $x \in M_1$, at least one partial function $y \to F_i(x,y)$ is affine convex and has a critical point, then the problem (OBDP) has a solution.*

Proof. In our hypothesis, the set $\psi(x)$ is nonvoid, for any x, and the compacity assures the existence of $\min_x f(x, \psi(x))$.

In the next Theorem, we shall use the Value Function Method or Utility Function Method. □

Theorem 4.3 *If a C^1 increasing scalarization partial function*

$$y \to L(x,y) = u(F_1(x,y), \ldots, F_k(x,y))$$

has a minimum, then there exists an index i such that $\psi_i(x) \neq \emptyset$. Moreover, if $f(x,y)$ is bounded, then the bilevel problem

$$\min_x [f(x,y) : y \in \psi(x)]$$

has solution.

Proof. Let $\min_y L(x,y) = L(x,y^*)$. Suppose that for each $i = 1, \ldots, k$, $\min_y F_i(x,y) < F_i(x,y^*)$. Then y^* would not be minimum point for the partial function $y \to L(x,y)$. Hence, there exists an index i such that $y^* \in \psi_i(x)$. □

Boundedness of f implies that the bilevel problem has solution once it is well-posed, but the fact that the problem is well-posed is shown in the first part of the proof.

4.1. Bilevel disjunctive programming algorithm

An important concept for making wise tradeoffs among competing objectives is bilevel disjunctive programming optimality, on affine manifolds, introduced in this chapter.

We present an exact algorithm for obtaining the bilevel disjunctive solutions to the multi-objective optimization in the following section.

Step 1: Solve

$$\psi_i(x) = \text{Argmin}_{y \in M_2} F_i(x,y), i = 1, \ldots, m.$$

Let $\psi(x) = \cup_{i=1}^{r} \psi_i(x)$ be a subset in M_2 representing the mapping of optimal solutions for the follower multi-objective function.

Step 2: Build the mapping $f(x, \psi(x))$.

Step 3: Solve the leader's following program

$$\min_x [f(x,y), y \in \psi(x)].$$

From numerical point of view, we can use the Newton algorithm for optimization on affine manifolds, which is given in [19].

http://dx.doi.org/10.5772/intechopen.75643

5. Models of bilevel disjunctive programming problems

The manifold M is understood from the context. The connection Γ_{ij}^{h} can be realized in each case, imposing convexity conditions.

Example 5.1 *Let us solve the problem (cite* [7]*, p. 7;* [9]*):*

$$\min_{(x_1,x_2)} F(x_1,x_2,y) = (x_1 - y, x_2)$$

subject to

$$(x_1,x_2) \in Argmin_{(x_1,x_2)} \{(x_1,x_2)\,|\, y^2 - x_1^2 - x_2^2 \geq 0\},$$

$$1 + x_1 + x_2 \geq 0,\ -1 \leq x_1, x_2 \leq 1, 0 \leq y \leq 1.$$

Both the lower and the upper level optimization tasks have two objectives each. For a fixed y value, the feasible region of the lower-level problem is the area inside a circle with center at origin $(x_1 = x_2 = 0)$ *and radius equal to y. The Pareto-optimal set for the lower-level optimization task, preserving a fixed y, is the bottom-left quarter of the circle,*

$$\{(x_1,x_2) \in \mathbb{R}^2\,|\, x_1^2 + x_2^2 = y^2, x_1 \leq 0, x_2 \leq 0\}.$$

The linear constraint in the upper level optimization task does not allow the entire quarter circle to be feasible for some y. Thus, at most a couple of points from the quarter circle belongs to the Pareto-optimal set of the overall problem. Eichfelder [8] reported the following Pareto-optimal set of solutions

$$A = \left\{(x_1,x_2,y) \in \mathbb{R}^3\,|\, x_1 = -1 - x_2, x_2 = -\frac{1}{2} \pm \frac{1}{2}\sqrt{2y^2 - 1}, y \in \left[\frac{1}{\sqrt{2}}, 1\right]\right\}.$$

The Pareto-optimal front in $F_1 - F_2$ *space can be written in parametric form*

$$\left\{(F_1,F_2) \in \mathbb{R}^2\,|\, F_1 = -1 - F_2 - t, F_2 = -\frac{1}{2} \pm \frac{1}{2}\sqrt{2t^2 - 1}, t \in \left[\frac{1}{\sqrt{2}}, 1\right]\right\}.$$

Example 5.2 *Consider the bilevel programming problem*

$$\min_x \left[(x-y)^2 + x^2 : -20 \leq x \leq 20, y \in \psi(x)\right],$$

where the set-valued function is

$$\psi(x) = Argmin_y [xy : -x - 1 \leqslant y \leqslant -x + 1].$$

Explicitly,

$$\psi(x) = \begin{cases} [-1,1] & if\ \ x = 0 \\ -x - 1 & if\ \ x > 0 \\ -x + 1 & if\ \ x < 0. \end{cases}$$

Since $F(x,y)=(x-y)^2+x^2$, *we get*

$$F(x,\psi(x))=\begin{cases}[0,1] & \text{if } x=0\\ (-2x-1)^2+x^2 & \text{if } x>0\\ (-2x+1)^2+x^2 & \text{if } x<0.\end{cases}$$

on the regions where the functions are defined.

Taking into account $(-2x-1)^2+x^2>0$ *and* $(-2x+1)^2+x^2>0$, *it follows that* $(x^\circ,y^\circ)=(0,0)$ *is the unique optimistic optimal solution of the problem. Now, if the leader is not exactly enough in choosing his solution, then the real outcome of the problem has an objective function value above* 1 *which is far away from the optimistic optimal value zero.*

Example 5.3 *Let* $F(x,y)=(F_1(x,y),F_2(x,y))$ *and a Pareto disjunctive problem*

$$\psi(x)=Argmin_y\,F(x,y)=Argmin_y\,F_1(x,y)\cup Argmin_y\,F_2(x,y).$$

Then it appears a bilevel disjunctive programming problem of the form

$$\min_x\ [f(x,y),y\in\psi(x)].$$

This problem is interesting excepting the case $\psi(x)=\emptyset,\forall x$. *If* $y\to F_1(x,y)$ *and* $y\to F_2(x,y)$ *are convex functions, then* $\psi(x)\neq\emptyset$.

To write an example, we use

$$F_1(x,y)=[xy:-x-1\leqslant y\leqslant -x+1],F_2(x,y)=\left[x^2+y^2:y\geqslant -x+1\right]$$

and we consider a bilevel disjunctive programming problem of the form

$$\min_x\ \left[(x-y)^2+x^2:-20\leq x\leq 20,y\in\psi(x)\right],$$

with

$$\psi(x)=\psi_1(x)\cup\psi_2(x),$$

where

$$\psi_1(x)=Argmin_y[xy:-x-1\leqslant y\leqslant -x+1]=\begin{cases}[-1,1] & \text{if } x=0\\ -x-1 & \text{if } x>0\\ -x+1 & \text{if } x<0,\end{cases}$$

$$\psi_2(x)=Argmin_y\left[x^2+y^2:y\geqslant -x+1\right]=\begin{cases}-x+1 & \text{if } x\leq 1\\ 0 & \text{if } x>1,\end{cases}$$

$$\psi(x)=\begin{cases}[-1,1] & \text{if } x=0\\ \{-x-1,-x+1\} & \text{if } 0<x\leq 1\\ \{-x-1,0\} & \text{if } x>1\\ -x+1 & \text{if } x<0.\end{cases}$$

The objective $f(x,y) = (x-y)^2 + x^2$ *and the multi-function* $\psi(x)$ *produce a multi-function*

$$f(x,\psi(x)) = \begin{cases} [0,1] & if \quad x = 0 \\ \left\{(2x+1)^2 + x^2, (2x-1)^2 + x^2\right\} & if \quad 0 < x \leq 1 \\ \left\{(2x+1)^2 + x^2, 2x^2\right\} & if \quad x > 1 \\ (2x-1)^2 + x^2 & if \quad x < 0. \end{cases}$$

In context, we find the inferior envelope

$$y(x) = \begin{cases} 0 & if \quad x = 0 \\ -x+1 & if \quad 0 < x \leq 1 \\ 0 & if \quad x > 1 \\ -x+1 & if \quad x < 0 \end{cases}$$

and then

$$f(x,y(x)) = \begin{cases} 0 & if \quad x = 0 \\ (2x-1)^2 + x^2 & if \quad x \in (-\infty, 0) \cup (0,1] \\ 2x^2 & if \quad x > 1. \end{cases}$$

Since $(2x-1)^2 + x^2 > 0$*, the unique optimal solution is* $(x^\circ, y^\circ) = (0,0)$.

If we consider only $\psi_1(x)$ *as active, then the unique optimal solution* $(0,0)$ *is maintained. If* $\psi_2(x)$ *is active, then the optimal solution is* $(0,1)$.

6. Properties of minimum functions

Let $\left(M_1, {}^1\Gamma\right)$, the *leader decision affine manifold*, and $\left(M_2, {}^2\Gamma\right)$, the *follower decision affine manifold*, be two connected affine manifolds of dimension n_1 and n_2, respectively. Starting from a function with two vector variables

$$\varphi : M_1 \times M_2 \to \mathbb{R}, (x,y) \to \varphi(x,y),$$

and taking the infimum after one variable, let say y, we build a function

$$f(x) = \inf_y \{\varphi(x,y) : y \in a(x)\},$$

which is called *minimum function*.

A *minimum function* is usually specified by a pointwise mapping a of the manifold M_1 in the subsets of a manifold M_2 and by a functional $\varphi(x,y)$ on $M_1 \times M_2$. In this context, some differential properties of such functions were previously examined in [4]. Now we add new properties related to increase and convexity ideas.

First we give a new proof to Brian White Theorem (see Mean Curvature Flow, p. 7, Internet 2017).

Theorem 6.1 *Suppose that M_1 is compact, $M_2 = [0,T]$ and $f : M_1 \times [0,T] \to \mathbb{R}$. Let $\phi(t) = \min_x f(x,t)$. If, for each x with $\phi(t) = f(x,t)$, we have $\frac{\partial f}{\partial t}(x,t) \geq 0$, then ϕ is an increasing function.*

Proof. We shall prove the statement in three steps.

(1) *If f is continuous, then ϕ is (uniformly) continuous.*

Indeed, f is continuous on the compact $M_1 \times [0,1]$, hence uniformly continuous. So, for $\varepsilon > 0$ it exists $\delta > 0$ such that if $|t_1 - t_2| < \delta$, then $|f(x,t_1) - f(x,t_2)| < \varepsilon$, for any $x \in M_1$, or

$$-\varepsilon < f(x,t_1) - f(x,t_2) < \varepsilon$$

On one hand, if we put $\phi(t_1) = f(x_1,t_1)$ and $\phi(t_2) = f(x_2,t_2)$, then we have

$$f(x,t_1) > f(x,t_2) - \varepsilon \geq f(x_2,t_2) - \varepsilon.$$

Hence $\min_x f(x,t_1) \geq f(x_2,t_2) - \varepsilon$, so is $\phi(t_1) - \phi(t_2) \geq -\varepsilon$.

On the other hand,

$$f(x,t_2) + \varepsilon > f(x,t_1) \geq f(x_1,t_1).$$

Hence $\min_x f(x,t_2) + \varepsilon \geq f(x_1,t_1)$, so is $\phi(t_1) - \phi(t_2) \leq \varepsilon$.

Finally, $|\phi(t_1) - \phi(t_2)| \leq \varepsilon$, for $|t_1 - t_2| < \delta$, that is, ϕ is (uniformly) continuous.

(2) *Let us fix $t_0 \in (0;T]$. If $\phi(t_0) = f(x_0,t_0)$ and $\frac{\partial f}{\partial t}(x_0,t_0) \geq 0$, then it exists $\delta > 0$ such that $\phi(t) \leq \phi(t_0)$, for any $t \in (t_0 - \delta; t_0)$.*

Suppose $\frac{\partial f}{\partial t}(x_0,t_0) > 0$, it exists $\delta > 0$ such that $f(x_0,t) \leq f(x_0,t_0)$, for each $t \in (t_0 - \delta; t_0)$. It follows $\min_x f(x,t) \leq f(x_0,t) \leq f(x_0,t_0)$, and so is $\phi(t) \leq \phi(t_0)$.

If $\frac{\partial f}{\partial t}(x_0,t_0) = 0$, then we use $\overline{f}(x,t) = f(x,t) + \varepsilon t,\ \varepsilon > 0$. For $\overline{f}$, the above proof holds, and we take $\varepsilon \to 0$.

(3) *ϕ is an increasing function.*

Let $0 \leq a < b \leq T$ and note $A = \{t \in [a,b] \,|\, \phi(t) \leq \phi(b)\}$. A is not empty. If $\alpha = \inf A$, then, by the step (2), $\alpha < b$ and, by the step (1), $\alpha \in A$. If $\alpha > a$, we can use the step (2) for $t_0 = \alpha$ and it would result that α was not the lower bound of A. Hence $\alpha = a$ and $\phi(a) \leq \phi(b)$.

Remark The third step shows that a function having the properties (1) and (2) is increasing. For this the continuity is essential. Only property (2) is not enough. For example, the function defined by $\phi(t) = t$ on $[0,1]$ and $\phi(t) = 1 - t$ on $(1,2]$ has only the property (2), but it is not increasing on $[0,2]$.

Remark Suppose that f is a C^2 function and $\min_x f(x,t) = f(x_0(t),t)$, where $x_0(t)$ is an interior point of M. Since $x_0(t)$ is a critical point, we have

$$\phi^{'}(t) = \frac{\partial f}{\partial t}(x_0(t), t) + < \frac{\partial f}{\partial x}(x_0(t), t), x_0^{'}(t) >= \frac{\partial f}{\partial t}(x_0(t), t) \geq 0.$$

Consequently, $\phi(t)$ is an increasing function. If M_1 has a nonvoid boundary, then the monotony extends by continuity (see also the evolution of an extremum problem).

□

Example 6.1 *The single-time perspective of a function* $f : \mathbb{R}^n \to \mathbb{R}$ *is the function* $g : \mathbb{R}^n \times \mathbb{R}_+ \to \mathbb{R}$, $g(x,t) = tf(x/t)$, $\operatorname{dom} g = \{(x,t) | x/t \in \operatorname{dom} f, t > 0\}$. *The single-time perspective* g *is convex if* f *is convex.*

The single-time perspective is an example verifying Theorem 7.1. Indeed, the critical point condition for g, *in* x, $\frac{\partial g}{\partial x} = 0$, *gives* $x = tx_0$, *where* x_0 *is a critical point of* f. *Consequently,* $\phi(t) = \min_x g(x,t) = tf(x_0)$. *On the other hand, in the minimum point, we have* $\frac{\partial g}{\partial t}(x,t) = f(x_0)$. *Then* $\phi(t)$ *is increasing if* $f(x_0) \geq 0$, *as in Theorem 4.1.*

Theorem 6.2 *Suppose that* M_1 *is compact and* $f : M_1 \times M_2 \to \mathbb{R}$. *Let* $\phi(y) = \min_x f(x,y)$. *If, for each* x *with* $\phi(y) = f(x,y)$, *we have* $\frac{\partial f}{\partial y^\alpha}(x,y) \geq 0$, *then* $\phi(y)$ *is a partially increasing function.*

Proof. Suppose that f is a C^2 function and $\min_x f(x,y) = f(x_0(y), y)$, where $x_0(y)$ is an interior point of M_1. Since $x_0(y)$ is a critical point, we have

$$\frac{\partial \phi}{\partial y^\alpha} = \frac{\partial f}{\partial y^\alpha}(x_0(y), y) + < \frac{\partial f}{\partial x}(x_0(y), y), \frac{\partial x_0}{\partial y^\alpha} >= \frac{\partial f}{\partial y^\alpha}(x_0(y), y) \geq 0.$$

Consequently, $\phi(y)$ is a partially increasing function. If M has a non-void boundary, then the monotony extends by continuity. □

Theorem 6.3 *Suppose that* M_1 *is compact and* $f : M_1 \times M_2 \to \mathbb{R}$. *Let* $\phi(y) = \min_x f(x,y)$. *If, for each* x *with* $\phi(y) = f(x,y)$, *we have* $d_y^2 f(x,y) \leq 0$, *then* $\phi(y)$ *is an affine concave function.*

Proof. Without loss of generality, we work on Euclidean case. Suppose that f is a C^2 function and $\min_x f(x,y) = f(x(y), y)$, where $x(y)$ is an interior point of M_1. Since $x(y)$ is a critical point, we must have

$$\frac{\partial f}{\partial x^i}(x(y), y) = 0.$$

Taking the partial derivative with respect to y^α and the scalar product with $\frac{\partial x^i}{\partial y^\beta}$ it follows

$$\frac{\partial^2 f}{\partial x^i \partial x^j}\frac{\partial x^j}{\partial y^\alpha}\frac{\partial x^i}{\partial y^\beta} + \frac{\partial^2 f}{\partial y^\alpha \partial x^i}\frac{\partial x^i}{\partial y^\beta} = 0.$$

On the other hand

$$d_y\phi(y) = d_y f(x(y), y) = \left(\frac{\partial f}{\partial x^i}\frac{\partial x^i}{\partial y^\alpha} + \frac{\partial f}{\partial y^\alpha}\right) dy^\alpha = \frac{\partial f}{\partial y^\alpha} dy^\alpha$$

$$d_y^2\phi(y) = \left(\frac{\partial^2 f}{\partial y^\alpha \partial x^i}\frac{\partial x^i}{\partial y^\beta} + \frac{\partial^2 f}{\partial y^\alpha \partial y^\beta}\right) dy^\alpha dy^\beta$$

$$= \left(-\frac{\partial^2 f}{\partial x^j \partial x^i}\frac{\partial x^i}{\partial y^\beta}\frac{\partial x^j}{\partial y^\alpha} + \frac{\partial^2 f}{\partial y^\alpha \partial y^\beta}\right) dy^\alpha dy^\beta \leq 0.$$

□

Theorem 6.4 *Let $f : M_1 \times M_2 \to \mathbb{R}$ be a C^2 function and*

$$\phi(y) = \min_x f(x, y) = f(x(y), y).$$

If the set $A = \{(x(y), y) : y \in M_2\}$ is affine convex and $f|_A$ is affine convex, then $\phi(y)$ is affine convex.

Proof. Suppose f is a C^2 function. At points $(x(y), y)$, we have

$$0 \leq d^2 f(x(y), y) = \left(\frac{\partial^2 f}{\partial x^i \partial x^j}\frac{\partial x^i}{\partial y^\alpha}\frac{\partial x^j}{\partial y^\beta} + 2\frac{\partial^2 f}{\partial x^i \partial y^\alpha}\frac{\partial x^i}{\partial y^\beta} + \frac{\partial^2 f}{\partial y^\alpha \partial y^\beta}\right) dy^\alpha dy^\beta$$

$$= \left(\frac{\partial^2 f}{\partial x^i \partial y^\alpha}\frac{\partial x^i}{\partial y^\beta} + \frac{\partial^2 f}{\partial y^\alpha \partial y^\beta}\right) dy^\alpha dy^\beta = d^2\phi(y).$$

Author details

Constantin Udriste[1]*, Henri Bonnel[2], Ionel Tevy[1] and Ali Sapeeh Rasheed[1]

*Address all correspondence to: udriste@mathem.pub.ro

1 Department of Mathematics and Informatics, Faculty of Applied Sciences, University Politehnica of Bucharest, Bucharest, Romania

2 ERIM, Université de la Nouvelle-Calédonie, Nouma Cédex, New Caledonia, France

References

[1] Aké LA, Sánchez M. Compact affine manifolds with precompact holonomy are geodesically complete. Journal of Mathematical Analysis and Applications. 2016;**436**(2):1369-1371. DOI: 10.1016/j.jmaa.2015.12.037

[2] Antonelli PL, Ingarden RS, Matsumoto M. The Theory of Sparays and Finsler Spaces with Applications in Physics and Biology. Netherlands: Kluwer Academic Publishers, Springer; 1993. pp. 97-132. DOI: 10.1007/978-94-015-8194-3-4

[3] Arsinte V, Bejenaru A. Convex mappings between Riemannian manifolds. Balkan Journal of Geometry and Its Applications. 2016;**21**(1):1-14

[4] Beresnev VV, Pshenichnyi BN. The differential properties of minimum functions. USSR Computational Mathematics and Mathematical Physics. 1974;**14**(3):101-113. DOI: 10.1016/0041-5553(74)90105-0

[5] Bonnel H, Todjihounde L, Udrişte C. Semivectorial bilevel optimization on Riemannian manifolds. Journal of Optimization Theory and Applications. 2015;**167**(2):464-486. DOI: 10.1007/s10957-015-0789-6

[6] Boyd S, Vandenberghe L. Convex Optimization. United Kingdom: Cambridge University Press; 2009. 716p

[7] Calin O, Udrişte C. Geometric Modeling in Probability and Statistics. New York: Springer; 2014. DOI: 10.1007/978-3-319-07779-6

[8] Combettes PL. Perspective functions: Properties, constructions, and examples. Set-Valued and Variational Analysis. Dordrecht: Springer Science+Business Media; 2017:1-18. DOI: 10.1007/s11228-017-0407-x

[9] Das P, Roy TK. Multi-objective geometric programming and its application in gravel box problem. Journal of Global Research in Computer Science. 2014;**5**(7):6-11

[10] Deb K, Sinha A. Solving bilevel multi-objective optimization problems using evolutionary algorithms: KanGAL Report Number 2008005; 2017

[11] Duffin RJ, Peterson EL, Zener C. Geometric Programming. New Jersey: John Wiley and Sons; 1967

[12] Eichfelder G. Solving nonlinear multiobjective bilevel optimization problems with coupled upper level constraints: Technical Report Preprint No. 320, Preprint-Series of the Institut für Angewandte Mathematik, University of Erlangen-Nuremberg, Germany; 2007

[13] Myers B. Arc length in metric and Finsler manifolds. Annals of Mathematics. 1938;**39**(2): 463-471. DOI: 10.2307/1968797

[14] Olteanu O, Udrişte C. Applications of Hahn-Banach principle to moment and optimization problems. Nonlinear Functional Analysis and its Applications. 2005;**10**(5):725-742

[15] Poston T, Stewart I. Catastrophe Theory and Its Applications. Pitman; 1977. 491p

[16] Pripoae CL. Affine differential invariants associated to convex functions. In: Proceedings of Conference of SSM, Cluj, 1998; Digital Data Publishing House, Cluj-Napoca; 1999. pp. 247-252

[17] Pripoae CL, Pripoae GT. Generalized convexity in the affine differential setting. In: The International Conference "Differential Geometry—Dynamical Systems" (DGDS); 10–13 October 2013; Bucharest-Romania: BSG Proceedings 21; 2013. pp. 156-166

[18] Proudnikov I. The necessary and sufficient conditions for representing Lipschitz multivariable function as a difference of two convex functions. arXiv preprint arXiv:1409.5081v4. 2014

[19] Rasheed AS, Udrişte C, Ţevy I. Barrier algorithm for optimization on affine manifolds. BSG Proceedings. 2017;**24**:47-63

[20] Strongin DRG, Sergeyev YD. Global Optimization with Non-Convex Constraints: Sequential and Parallel Algorithms. London: Springer; 2000. 704p. DOI: 10.1007/978-1-4615-4677-1

[21] Sánchez M. Geodesic connectedness of semi-Riemannian manifolds. arXiv preprint math/0005039. 2001;**47**:3085-3102. DOI: 10.1016/s0362-546x(01)00427-8

[22] Udrişte C, Dogaru O. Extrema conditioned by orbits. (in Romanian). Buletinul Institutului Politehnic Bucureşti, Seria Mecanică. 1989;**51**:3-9

[23] Udrişte C. Convex Functions and Optimization Methods on Riemannian Manifolds. Netherlands: Kluwer Academic Publishers; 1994 . 365p. DOI: 10.1007/978-94-015-8390-9

[24] Udrişte C. Sufficient decreasing on Riemannian manifolds. Balkan Journal of Geometry and Its Applications. 1996;**1**:111-123

[25] Udrişte C. Riemannian convexity in programming (II). Balkan Journal of Geometry and Its Applications. 1996;**1**(1):99-109

[26] Udrişte C. Optimization methods on Riemannian manifolds. In: Proceedings of (IRB) International Workshop; 8–12 August 1995; Monteroduni. Italy: Algebras, Groups and Geometries; 1997. pp. 339-359

[27] Udrişte C, Bercu G. Riemannian Hessian metrics. Analele Universitatii din Bucuresti. 2005;**55**:189-204

[28] Yano K, Ishihara S. Tangent and Cotangent Bundles: Differential Geometry. New York: Dekker; 1973. 423p