Characterizations of Strongly Quasiconvex Functions and its Applications in the Gradient Method

Felipe Lara
www.felipelara.cl

Instituto de Alta Investigación (IAI),
Universidad de Tarapacá,
Arica, Chile.

Thursday 30, May 2024,
VAO2024,
Catholic University of Milano,
Milano, Italy.

This talk is based on the papers:

(*La*) F. LARA, On strongly quasiconvex functions: existence results and proximal point algorithms, *JOTA*, **192**, 891–911, (2022).

(*MV*) F. LARA, R.T. MARCAVILLACA, P.T. VUONG, Characterizations, dinamycal systems and gradient methods for strongly quasiconvex functions, *Submitted*, (2024).

Generalized convex functions: weakly convex, Difference of Convex (DC), Invex, Quasiconvex functions

- Economic theory, especially in *consumer preference theory* (see (D)). *Quasiconcavity* is the mathematical formulation of *tendency to the diversification*.
- Fractional programming (applications in economics as min (cost/time), max (return/risk) among others).

(*D*) G. Debreu. "Theory of value: an axiomatic approach to economic equilibrium". John Wiley, New York, (1959).

Gérard Debreu

Nobel Prize in Economics 1983.

Strongly Convex and Quasiconvex [Polyak-1966]

A function h with a convex domain is said to be

(a) Strongly convex if there exists $\gamma \in]0, +\infty[$ such that for all $x, y \in \mathrm{dom}\, h$ and all $\lambda \in [0, 1]$, we have

$$h(\lambda y + (1-\lambda)x) \leq \lambda h(y) + (1-\lambda)h(x) - \lambda(1-\lambda)\frac{\gamma}{2}\|x - y\|^2, \quad (1)$$

(b) Strongly quasiconvex if there exists $\gamma \in]0, +\infty[$ such that for all $x, y \in \mathrm{dom}\, h$ and all $\lambda \in [0, 1]$, we have

$$h(\lambda y + (1-\lambda)x) \leq \max\{h(y), h(x)\} - \lambda(1-\lambda)\frac{\gamma}{2}\|x - y\|^2. \quad (2)$$

(P) B.T. Polyak, Existence theorems and convergence of minimizing sequences in extremum problems with restrictions, *Soviet Math.*, **7**, 72–75, (1966).

Summarizing (quasiconvex is denoted by qcx):

$$\begin{array}{ccccc} \text{strongly convex} & \Rightarrow & \text{strictly convex} & \Rightarrow & \text{convex} \\ \Downarrow & & \Downarrow & & \Downarrow \\ \text{strongly qcx} & \Rightarrow & \text{strictly qcx} & \Rightarrow & \text{qcx} \end{array}$$

Remark

There is no relationship between convex and strongly quasiconvex functions. Indeed, the function $h(x) = -x^2 - x$ is strongly quasiconvex on $[0, 1]$ without being convex, while $h(x) \equiv 1$ is convex without being strongly quasiconvex.

Geometric Interpretation

Proposition

Let $h : \mathbb{R}^n \to \overline{\mathbb{R}}$ be a proper function. Then

(a) h is convex if and only if

$$\operatorname{epi} h := \{(x, t) \in \mathbb{R}^n \times \mathbb{R} : \ h(x) \leq t\} \text{ is a convex set.} \tag{3}$$

(b) h is quasiconvex if and only if

$$S_\lambda(h) := \{x \in \mathbb{R}^n : \ h(x) \leq \lambda\} \text{ is a convex set for all } \lambda \in \mathbb{R}.$$

Example [J-1996]

Given $c > 0$, the function $h : \mathbb{R}^n \to \mathbb{R}$ given by $h(x) = \|x\|$ is strongly quasiconvex on $\mathbb{B}(0, c)$ without being strongly convex.

Example [Lara-2022]

Given $c > 0$, the function $h : \mathbb{R}^n \to \mathbb{R}$ given by $h(x) = \sqrt{\|x\|}$ is strongly quasiconvex on $\mathbb{B}(0, c)$ without being convex.

Example [LMV-2024]

The function $h : \mathbb{R} \to \mathbb{R}$ given by $h(x) = x^2 + 3\sin^2(x)$ is strongly quasiconvex on $\mathbb{R}$ without being convex.

(J) M. Jovanović, A note on strongly convex and quasiconvex functions, *Math. Notes*, **60**, 584–585, (1996).

New examples

Fractional programming

Given a subset K of $\mathbb{R}^n$, and functions $h : \mathbb{R}^n \to \mathbb{R}$ and $g : \mathbb{R}^n \to \mathbb{R}$, we define the fractional minimization problem by

$$\min_{x \in K} \varphi(x) \;=\; \min_{x \in K} \frac{h(x)}{g(x)}. \tag{FMP}$$

This problem is important in continuous optimization and mathematical programming due to its applications in several fields of the mathematical sciences, especially for economics purposes, for instance, in maximization of productivity as maximization of return/risk or profit/cost and minimization of cost/time among others.

Proposition [ILMY-2024]

Suppose that $\varphi(x) = \frac{h(x)}{g(x)}$ for all $x \in \mathrm{dom}\,\varphi$ with $\mathrm{dom}\,\varphi$ a convex set, h is strongly convex with modulus $\gamma > 0$, g is finite, positive and bounded from above by M on $\mathrm{dom}\,\varphi$. If any of the following conditions holds:

(a) g is affine,

(b) h is nonnegative on $\mathrm{dom}\,\varphi$ and g is concave,

(c) h is nonpositive on $\mathrm{dom}\,\varphi$ and g is convex,

then φ is strongly quasiconvex with modulus $\gamma' := \frac{\gamma}{M} > 0$.

Corollary [ILMY-2024]

Let $A, B \in \mathbb{R}^{n\times n}$, $a, b \in \mathbb{R}^n$, $\alpha, \beta \in \mathbb{R}$, and $\varphi : \mathbb{R}^n \to \mathbb{R}$ be the function given by:

$$\varphi(x) = \frac{h(x)}{g(x)} = \frac{\frac{1}{2}\langle Ax, x\rangle + \langle a, x\rangle + \alpha}{\frac{1}{2}\langle Bx, x\rangle + \langle b, x\rangle + \beta}. \tag{4}$$

Take $0 < m < M$ and define:

$$K = \{x \in \mathbb{R}^n : \; m \leq g(x) \leq M\}.$$

If A is a symmetric and positive definite matrix and at least one of the following conditions holds:

(a) $B = 0$ (the null matrix),

(b) h is nonnegative on K and B is negative semidefinite,

(c) h is nonpositive on K and B is positive semidefinite,

then h is strongly quasiconvex on K with modulus $\gamma = \frac{\lambda_{\min}(A)}{M}$, where $\lambda_{\min}(A)$ is the minimum eigenvalue of A.

Convexity vs Quasiconvexity

Proposition

If h is a proper function and $\beta > 0$, then

$$h \text{ is convex } \iff h + \frac{1}{2\beta}\|\cdot\|^2 \text{ is strongly convex.} \tag{5}$$

Remark

If h is a proper function and $\beta > 0$, then

$$h \text{ is quasiconvex } \not\Rightarrow h + \frac{1}{2\beta}\|\cdot\|^2 \text{ is strongly quasiconvex.} \tag{6}$$

Indeed, the continuous function $h : \mathbb{R} \to \mathbb{R}$ given by $h(x) = x^3$ is quasiconvex, but $x^3 + \frac{1}{2\beta}|x|^2$ is not strongly quasiconvex for any $\beta > 0$.

The Existence Result

Theorem [Polyak-1966]

Let K be a closed and convex set in $\mathbb{R}^n$ and $h : K \to \mathbb{R}$ be a lsc and strongly convex function with modulus $\gamma_h > 0$. Then, $\mathrm{argmin}_K\, h = \{\overline{x}\}$ and

$$h(\overline{x}) + \frac{\gamma_h}{2}\|y - \overline{x}\|^2 \leq h(y), \ \forall\, y \in \mathrm{dom}\, h. \tag{7}$$

Open Problem

Let K be a closed and convex set in $\mathbb{R}^n$ and $h : K \to \mathbb{R}$ be a lsc and **strongly quasiconvex** function with modulus $\gamma_h > 0$. Then, $\mathrm{argmin}_K\, h$ is a singleton ?.

(P) B.T. Polyak, Existence theorems and convergence of minimizing sequences in extremum problems with restrictions, *Soviet Math.*, **7**, 72–75, (1966).

Partial Advances

(*VN*) A.A. Vladimirov, Y.E. Nesterov, Y.N. Chekanov, O ravnomerno kvazivypuklyh funkcionalah [On uniformly quasiconvex functionals], *Vestn. Mosk. un-ta, vycis. mat. i kibern.*, **4**, 18–27, (1978).

(*V*) J.P. Vial, Strong convexity of sets and functions, *J. Math. Economics*, **9**, 187–205, (1982).

(*J*) M. Jovanović, On strong quasiconvex functions and boundedness of level sets, *Optimization*, **20**, 163–165, (1989).

(*CZ*) J.P. Crouzeix, J.A. Ferland, C. Zălinescu, α-convex sets and strong quasiconvexity, *MOR*, **22**, 998–1022, (1997).

Definition

A proper function $h : \mathbb{R}^n \to \overline{\mathbb{R}}$ is said to be:

(*i*) 2-supercoercive, if

$$\liminf_{\|x\|\to+\infty} \frac{h(x)}{\|x\|^2} > 0, \tag{8}$$

(*iv*) coercive, if

$$\lim_{\|x\|\to+\infty} h(x) = +\infty. \tag{9}$$

or equivalently, if $S_\lambda(h)$ is bounded for all $\lambda \in \mathbb{R}$.

Remark

Clearly, $(i) \Rightarrow (ii)$, but the reverse statements does not hold as the function $h(x) = \sqrt{x}$ shows.

Main Theorem

Theorem [Lara-2022]

Let K be a convex set in $\mathbb{R}^n$ and $h : K \to \mathbb{R}$ be a strongly quasiconvex function with modulus $\gamma_h > 0$. Then h is 2-supercoercive (in particular, supercoercive).

Corollary [Lara-2022; Kab-Lara-2022]

Let K be a closed and convex set in $\mathbb{R}^n$ and $h : K \to \mathbb{R}$ be a lsc and strongly quasiconvex function with modulus $\gamma_h > 0$. Then, $\mathrm{argmin}_K\, h = \{\overline{x}\}$ and

$$h(\overline{x}) + \frac{\gamma_h}{8}\|y - \overline{x}\|^2 \leq h(y), \ \forall\ y \in K. \tag{10}$$

Proximal Point Algorithm

Step 0. Take $x^1 \in K$, $k = 0$ and $\{c_k\}_{k\in\mathbb{N}}$ be a sequence of positive numbers bounded away from 0.

Step 1. Take $k = k + 1$ and

$$x^{k+1} \in \mathrm{Prox}_{c_k h}(K, x^k). \tag{11}$$

Step 2. If $x^{k+1} = x^k$, then STOP, $\{x^k\} = \mathrm{argmin}_K h$. Otherwise, go to Step 1.

Theorem [Lara-2022]

Let $K \subseteq \mathbb{R}^n$ be a closed and convex set, $h : \mathbb{R}^n \to \overline{\mathbb{R}}$ be a proper, lsc and strongly quasiconvex function with value $\gamma_h > 0$ such that $K \subseteq \mathrm{dom}\, h$, and $\{c_k\}_{k\in\mathbb{N}}$ be a sequence of positive numbers bounded away from 0. Then the sequence $\{x^k\}_{k\in\mathbb{N}}$, generated by relation (11), is a minimizing sequence of h, i.e., $h(x^k) \downarrow \min_{x\in K} h(x)$.

Algorithmic Consequences:

(*KL*) A. KABGANI, F. LARA, Strong subdifferentials: theory and applications in nonconvex optimization, *JoGO*, **84**, 349–368, (2022).

(*RM*) S.-M. GRAD, F. LARA, R.T. MARCAVILLACA, Relaxed-inertial proximal point type algorithms for quasiconvex minimization, *JoGO*, **85**, 615–635, (2023).

(*LM*) F. LARA, R.T. MARCAVILLACA, Bregman proximal point type algorithms for quasiconvex minimization, *Optimization*, DOI: 10.1080/02331934.2022.2112580, (2024).

Equilibrium Problems:

(*IL*) A. Iusem, F. Lara, Proximal point algorithms for quasiconvex pseudomonotone equilibrium problems, *JOTA*, **193**, 1–3, 443–461, (2022).

(*RM*) S.-M. Grad, F. Lara, R.T. Marcavillaca, Relaxed-inertial proximal point algorithms for nonconvex pseudomonotone equilibrium problems with applications, *JOTA*, DOI: 10.1007/s10957-023-02375-1, (2024).

(*RM*) F. Lara, R.T. Marcavillaca, L. H. Yen, An extragradient projection method for strongly quasiconvex equilibrium problems with applications. *COAM*, DOI: 10.1007/s40314-024-02626-5, (2024).

(*IY*) A. Iusem, F. Lara, R.T. Marcavillaca, L.H. Yen, A two-steps proximal point algorithm for nonconvex equilibrium problems with applications in fractional programming, *Submitted*, (2024).

Characterizations

Theorem

Let $K \subseteq \mathbb{R}^n$ be a convex set and $h : \mathbb{R}^n \to \mathbb{R}$ be a differentiable function. Then the following assertions hold:

(a) h is convex on K if and only if

$$\langle \nabla h(x) - \nabla h(y), x - y \rangle \geq 0, \ \forall \ x, y \in K. \tag{12}$$

(b) h is strongly convex on K with modulus $\gamma_h > 0$ if and only if

$$\langle \nabla h(x) - \nabla h(y), x - y \rangle \geq \gamma_h \|x - y\|^2, \ \forall \ x, y \in K. \tag{13}$$

Generalized Monotonicity

Given a nonempty set C in $\mathbb{R}^n$ and a set-valued operator $T : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$, T is said to be:

(a) strongly pseudomonotone on C with modulus $\gamma \geq 0$, if for all $x, y \in C$, we have

$$\langle v, x-y \rangle \geq 0 \Longrightarrow \langle u, y-x \rangle \leq -\gamma \|y-x\|^2, \ \forall\, v \in T(y), \ \forall\, u \in T(x). \tag{14}$$

(b) strongly quasimonotone on C with modulus $\gamma \geq 0$, if for all $x, y \in C$, we have

$$\langle v, x-y \rangle > 0 \Longrightarrow \langle u, y-x \rangle \leq -\gamma \|y-x\|^2, \ \forall\, v \in T(y), \ \forall\, u \in T(x). \tag{15}$$

Theorem [AE-1961]

Let $K \subseteq \mathbb{R}^n$ be a convex set and $h : \mathbb{R}^n \to \mathbb{R}$ be a differentiable function. Then h is quasiconvex on K if and only if for every $x, y \in K$, we have

$$h(x) \leq h(y) \implies \langle \nabla h(y), x - y \rangle \leq 0. \tag{16}$$

(*AE*) K.J. Arrow, A.C. Enthoven, Quasiconcave programming, *Econometrica*, **29**, 779–800, (1961).

Kenneth Joseph Arrow

Nobel Prize in Economics 1972.

Theorem [LMV-2024]

Let $K \subseteq \mathbb{R}^n$ be a convex set and $h : K \to \mathbb{R}$ be differentiable function. Then h is strongly quasiconvex on K with modulus $\gamma_h \geq 0$ if and only if for every $x, y \in K$, we have

$$h(x) \leq h(y) \implies \langle \nabla h(y), x - y \rangle \leq -\frac{\gamma_h}{2} \|y - x\|^2. \tag{17}$$

Corollary [AE-1961]

Let $K \subseteq \mathbb{R}^n$ be a convex set and $h : K \to \mathbb{R}$ be differentiable function. Then h is quasiconvex if and only if for every $x, y \in K$, we have

$$h(x) \leq h(y) \implies \langle \nabla h(y), x - y \rangle \leq 0. \tag{18}$$

Generalized Monotonicity

Proposition [LMV-2024]

Let $K \subseteq \mathbb{R}^n$ be a convex set and $h : K \to \mathbb{R}$ be differentiable function. Then h is strongly quasiconvex on K with modulus $\gamma_h \geq 0$ if and only if for each $x, y \in K$, we have

$$\langle \nabla h(x), y - x \rangle > -\frac{\gamma_h}{2}\|y - x\|^2 \implies \langle \nabla h(y), x - y \rangle \leq -\frac{\gamma_h}{2}\|y - x\|^2. \quad (19)$$

Corollary

If $\gamma_h = 0$, then

$$h \text{ is quasiconvex} \iff \nabla h \text{ is quasimonotone}. \quad (20)$$

Remark [LMV-2024]

$$\nabla h \text{ is } \gamma_h - \text{strongly monotone}$$
$$\Downarrow$$
$$\langle \nabla h(x), y - x \rangle > -\frac{\gamma_h}{2}\|y - x\|^2 \implies \langle \nabla h(y), x - y \rangle \leq -\frac{\gamma_h}{2}\|y - x\|^2$$
$$\Downarrow$$
$$\nabla h \text{ is } \frac{\gamma_h}{2} - \text{strongly pseudomonotone} \tag{21}$$
$$\Downarrow$$
$$\nabla h \text{ is } \frac{\gamma_h}{2} - \text{strongly quasimonotone.}$$

Proposition [LMV-2024]

Let $K \subseteq \mathbb{R}^n$ be a convex set and $h : K \to \mathbb{R}$ be differentiable function and $\gamma \geq 0$. Then

$$\begin{array}{ccc} h \,\mathrm{is}\, \gamma-\mathrm{strongly\,convex} & \Longleftrightarrow & \nabla h \,\mathrm{is}\, \frac{\gamma}{2}-\mathrm{strongly\,mon} \\ \Downarrow & & \Downarrow \\ h \,\mathrm{is}\, \gamma-\mathrm{strongly\,qcx} & \Longleftrightarrow & \nabla h \text{ satisfies relation } (19) \\ \Downarrow & & \Downarrow \\ h \,\mathrm{is}\, \gamma-\mathrm{sharply} & \Longleftrightarrow & \nabla h \,\mathrm{is}\, \frac{\gamma}{2}-\mathrm{strongly\,pseudomon.} \end{array}$$

Definition [KK-2000]

h is said to be sharply quasiconvex with modulus $\gamma \geq 0$ if for every $x, y \in K$, the following implication holds:

$$\langle \nabla h(y), x-y \rangle \geq 0 \Rightarrow h(\lambda y+(1-\lambda)x) \leq \max\{h(y), h(x)\}-\lambda(1-\lambda)\frac{\gamma}{2}\|y-x\|^2. \tag{22}$$

Generalized Convexity for Gradient Methods

Definition [NNG-2019]

A differentiable function $h : \mathbb{R}^n \to \mathbb{R}$ is said to be μ-quasi-strongly convex, $\mu > 0$, if

$$\langle \nabla h(x), x - x^* \rangle \geq h(x) - h(x^*) + \frac{\mu}{2}\|x - x^*\|^2, \ \forall \, x \in \mathbb{R}^n. \tag{23}$$

where x^* denotes the projection of x onto $\operatorname{argmin}_{\mathbb{R}^n} h$.

(*NN*) I. Necoara, Y. Nesterov, F. Glineur, Linear convergence of first-order methods for non-strongly convex optimization, *Math. Program.*, **175**, 69–107, (2019).

Definition

A differentiable function $h : \mathbb{R}^n \to \mathbb{R}$ satisfies the Polyak-Łojasiewicz (PL) property if there exists $\mu > 0$ such that

$$\|\nabla h(x)\|^2 \geq \mu(h(x) - h(\overline{x})), \ \forall \, x \in \mathbb{R}^n, \tag{24}$$

where $\overline{x} \in \mathrm{argmin}_{\mathbb{R}^n} h$.

(*L*) S. Łojasiewicz, A topological property of real analytic subsets. Coll. du CNRS, Les équations aux dérivées partielles, **117**, 87–89, (1963).

(*P*2) B.T. Polyak, Gradient methods for minimizing functionals, *Zh. Vychisl. Math. Mat. Fiz.*, **3**, 643–653, (1963).

Proposition [LMV-2024]

Let $h : \mathbb{R}^n \to \mathbb{R}$ be a differentiable function. If h is μ-quasi-strongly convex function with $\mu > 0$ and admits an unique minimizer, then h is strongly quasiconvex with modulus $\mu > 0$.

Proposition [LMV-2024]

Let $h : \mathbb{R}^n \to \mathbb{R}$ be a differentiable function. If h is strongly quasiconvex with modulus $\gamma > 0$ with L-Lipschitz continuous gradient ($L > 0$), then the PL property holds with modulus $\mu := \frac{\gamma^2}{2L} > 0$, that is,

$$\|\nabla h(x)\|^2 \geq \frac{\gamma^2}{2L}(h(x) - h(\overline{x})), \ \forall \ x \in \mathbb{R}^n, \qquad (25)$$

where $\overline{x} = \operatorname{argmin}_{\mathbb{R}^n} h$.

Remark [LMV-2024]

(*i*) The reverse statement does not hold in general as the Example 4.1.3 in the Book of Nesterov 2018 shows (satisfies the PL property but it is not strongly quasiconvex).

(*ii*) The function $h : \mathbb{R}^n \to \mathbb{R}$ given by $h(x) = x^2 + 3\sin^2 x$ is an example of a strongly quasiconvex function satisfying the PL property and without being convex.

(*N*) Y. Nesterov, "Lectures on convex optimization". Springer, Berlin, (2018).

The Problem

We are concerned with the study of the minimization problem

$$\min_{x \in \mathbb{R}^n} h(x). \tag{26}$$

where $h : \mathbb{R}^n \to \mathbb{R}$ is a continuously differentiable function. Our goal is the study of problem (26) via the first-order dynamical system:

$$\begin{cases} \dot{x}(t) + \nabla h(x(t))) = 0, \ t > 0, \\ x(t_0) = x_0. \end{cases} \tag{27}$$

Theorem [LMV-2024]

Let the function $h : \mathbb{R}^n \to \mathbb{R}$ be continuously differentiable and strongly quasiconvex with modulus $\gamma \geq 0$ and $\overline{x} = \mathrm{argmin}_{\mathbb{R}^n} h$. Then the following assertions hold:

(a) $t \mapsto h(x(t))$ is nonincreasing.

(b) Any trajectory $x(t)$ to the dynamical system (27) satisfy that

$$\|x(t) - \overline{x}\| \leq \|x_0 - \overline{x}\| e^{-\frac{\gamma}{2} t}, \tag{28}$$

i.e., $x(t)$ converges exponentially to the unique solution of (26);

(c) For any trajectory $x(t)$ to the dynamical system (27) there exists $T > 0$ and $L > 0$ such that

$$h(x(t)) - h(\overline{x}) \leq \min \left\{ \frac{L}{2} \|x_0 - \overline{x}\| e^{-\frac{\gamma}{2} t}, (h(x_0) - h(\overline{x}))\, e^{-\frac{\gamma^2}{2L} t} \right\}, \forall t \geq T, \tag{29}$$

as a consequence, $h(x(t))$ converges exponentially to $h^* = \min_{\mathbb{R}^n} h$.

The Gradient Method

Algorithm 1 The Gradient Method

Step 0. Take $x^0 \in \mathbb{R}^n$, $k = 0$ and a sequence $\{\beta_k\}_k \subseteq \mathbb{R}_{++}$.

Step 1. Compute $\nabla h(x^k)$ and

$$x^{k+1} = x^k - \beta_k \nabla h(x^k). \tag{30}$$

Step 2. If $x^{k+1} = x^k$, then STOP, $x^k \in \operatorname{argmin}_{\mathbb{R}^n} h$. Otherwise, take $k = k + 1$ and go to Step 1.

Theorem [LMV-2024]

Let $h : \mathbb{R}^n \to \mathbb{R}$ be a strongly quasiconvex function with modulus $\gamma > 0$ and differentiable with locally Lipschitz continuous gradient. Let $\overline{x} = \mathrm{argmin}_{\mathbb{R}^n} h$ and $\{\beta_k\}_k$ be a positive sequence satisfying

$$0 < \underline{\beta} \leq \beta_k \leq \overline{\beta} < \min\left\{\frac{\gamma}{L_0^2}, \frac{2}{L_0}\right\}. \tag{31}$$

Then,

$$\|x^{k+1} - \overline{x}\|^2 \leq \left(1 - \beta_k(\gamma - \beta_k L_0^2)\right) \|x^k - \overline{x}\|^2, \ \forall\ k \in \mathbb{N}. \tag{32}$$

and $\{x^k\}_k$ converges linearly to the unique solution $\overline{x}$.

Corollary [LMV-2024]

Assume the assumptions of previous Theorem holds, $\gamma < 2L_0$ and $\beta_k < \frac{\gamma}{L_0^2} \leq \frac{2}{L_0}$. Then we have an optimal convergence rate for the functional values:

$$h(x^k) - h(\overline{x}) \leq \left(1 - \frac{\gamma^2}{4L_0^2}\right)^{k-1} \|x^0 - \overline{x}\|^2. \qquad (33)$$

and

$$h(x^k) - h(\overline{x}) \leq \left(1 - \frac{\gamma^3}{4L_0^3}\left(1 - \frac{\gamma}{4L_0}\right)\right)^{k-1} (h(x^0) - h(\overline{x})). \qquad (34)$$

2nd-order Dynamical System

We consider the following gradient dynamical system:

$$\begin{cases} \ddot{x}(t) + \alpha \dot{x}(t) + \nabla h(x(t)) = 0, \ t > 0, \\ x(0) = x_0, \quad \dot{x}(0) = v_0. \end{cases} \tag{35}$$

Assumption

Tthere exists $\kappa \in]0, +\infty[$ such that for every trajectory $x(t)$ of the dynamical system (35) we have

$$\langle \nabla h(x(t)), x(t) - \overline{x} \rangle \geq \kappa (h(x(t)) - h(\overline{x})), \tag{36}$$

where $\overline{x} = \mathrm{argmin}_{\mathbb{R}^n} h$.

Remark [LMV-2024]

(*i*) Assumption (36) holds trivially with $\kappa = 1$ for convex functions.

(*ii*) If $h : \mathbb{R}^n \to \mathbb{R}$ is differentiable with L-Lipschitz gradient and strongly quasiconvex with modulus $\gamma > 0$, then by Main Theorem and Descent Lemma with $\nabla h(\overline{x}) = 0$, we have

$$\langle \nabla h(x), x - \overline{x} \rangle \geq \frac{\gamma}{2} \|x - \overline{x}\|^2 \geq \frac{\gamma}{L}(h(x) - h(\overline{x})),$$

which implies (36) with $\kappa = \frac{\gamma}{L}$.

Theorem [LMV-2024]

Let $h : \mathbb{R}^n \to \mathbb{R}$ be a differentiable and strongly quasiconvex function with modulus $\gamma > 0$ and $\overline{x} = \mathrm{argmin}_{\mathbb{R}^n} h$. Suppose that assumption (36) holds. Then any trajectory $x(t)$ generated by (35) converges exponentially to the unique solution $\overline{x}$ of problem (26), and the function values $h(x(t))$ converges exponentially to the optimal value $h(\overline{x})$.

Corollary [LMV-2024]

The following asymptotic exponential convergence rate holds:

$$h(x(t)) - h^* = \mathcal{O}\left(e^{-\lambda\kappa t}\right), \tag{37}$$

$$\|x(t) - \overline{x}\| = \mathcal{O}\left(e^{-\lambda\kappa t}\right), \tag{38}$$

with $\lambda = \min\{\sqrt{\frac{\gamma}{2\kappa}}, \frac{2\alpha}{\kappa+4}\}$.

Theorem [LMV-2024]

Let $h : \mathbb{R}^n \to \mathbb{R}$ be a differentiable with L-Lipschitz gradient and strongly quasiconvex function with modulus $\gamma > 0$, $\overline{x} = \mathrm{argmin}_{\mathbb{R}^n} h$, $\theta \in]0, 1[$ and $\beta \in]0, \frac{1-\theta^2}{L}]$. Then the sequence $\{x_k\}_k$, generated by the heavy ball method:

$$x_{k+1} = x_k + \theta(x_k - x_{k-1}) - \beta \nabla h(x_k). \tag{39}$$

converges linearly to $\overline{x}$ and the sequence $\{h(x_k)\}_k$ converges linearly to the optimal value $h(\overline{x})$.

Corollary [LMV-2024]

Under the assumptions of the previous Theorem, we have the following convergence rate

$$h(x_{k+1}) - h(\overline{x}) \leq \left(1 - \frac{\rho}{\sigma}\right)^{k} E_1, \tag{40}$$

$$\|x_{k+1} - x_k\|^2 \leq \frac{2\beta}{\theta^2}\left(1 - \frac{\rho}{\sigma}\right)^{k} E_1. \tag{41}$$

Moreover, for all $k \in \mathbb{N}$, we have

$$\|\nabla h(x_k)\| \leq \frac{(1+\theta)}{\theta}\left(1 - \frac{\rho}{\sigma}\right)^{\frac{k-1}{2}} \sqrt{\frac{2}{\beta}}\sqrt{E_1},$$

$$\|x_k - \overline{x}\| \leq \frac{2(1+\theta)}{\gamma\,\theta}\left(1 - \frac{\rho}{\sigma}\right)^{\frac{k-1}{2}} \sqrt{\frac{2}{\beta}}\sqrt{E_1}.$$

where $\theta \in]0, 1[$, $\beta \in]0, \frac{1-\theta^2}{L}]$, $\rho = \min\{\frac{\beta}{2}, \frac{1-\beta L-\theta^2}{2\beta}\}$, $\sigma = \max\{\frac{2L}{\gamma^2} + \beta, \frac{1}{\beta}\}$ and $E_1 := h(x_0) - h(\overline{x}) + \frac{\theta^2}{2\beta}\|x_1 - x_0\|^2$.

ADR) J.-F. Aujol, Ch. Dossal, A. Rondepierre, Convergence rate of the heavy ball method for quasi-strongly convex optimization, *SIOPT*, **32**, 1817–1842, (2022).

(BC) H.H. Bauschke, P.L. Combettes, "Convex Analysis and Monotone Operators Theory in Hilbert Spaces". *CMS Books in Mathematics*. Springer-Verlag, second edition, (2017).

(R) R.T. Rockafellar, Monotone operators and proximal point algorithms, *SICON.*, **14**, 877–898, (1976).

(V) P.T. Voung, A second-order dynamical system and its discretization for strongly pseudo-monotone variational inequality, *SICON*, **59**, 2875–2897, (2021).