Logarithmic convexifying of polynomials

Abdulljabar Naji Ahmed Abdullah

SUMMARY OF THE DOCTORAL DISSERTATION

One of the fundamental problems of analysis, technology, economics and other branches of science is the search for minima and critical points of functions. One of the methods leading to this goal is the deformation of a given function to a convex function, searching for critical points of this deformation and iterating this process. Reducing a function to a convex or strongly convex function leads to easy determination of critical points and minima of this deformation. These are the unique points where the gradient is zero.The classic approach to convexifying of a function $f : \mathbb{R}^n \to \mathbb{R}$ on bounded and convex sets is to add a strongly convex function $b : \mathbb{R}^n \to \mathbb{R}$ such that $f + b$ is a strongly convex function on this set (see for instance papers of A.N.Tikhonov, W.B.Liu, C.A.Flouudas and S.Zlobec for quadratic function $b(x) = \gamma|x|^2$, $\gamma > 0$). We describe this more precisely.

Let $b : \mathbb{R}^n \to \mathbb{R}$ be a $\mathscr{C}^k$ class μ-strongly convex function, $k \geq 2$, $\mu > 0$. Let $X \subset \mathbb{R}^n$ be a compact and convex set, let $f : \mathbb{R}^n \to \mathbb{R}$ be a function of class $\mathscr{C}^k$ and let $D \in \mathbb{R}$ be a positive number such that

$$|\partial^2_\beta f(x)| \leq D \quad \text{for } \ x \in X \text{ and } \beta \in S^{n-1},$$

where S^{n-1} the unit sphere in $\mathbb{R}^n$, and $\partial^2_\beta f(x)$ is the second order derivative of f in the direction β at x. One can directly check that (see. uwaga2.2.3 and fakt 3.1.1):

For any $\xi \in \mathbb{R}^n$ and $N > D/\mu$, the function $\phi_{N,\xi} : \mathbb{R}^n \to \mathbb{R}$ defined by

$$\phi_{N,\xi}(x) = Nb(x - \xi) + f(x), \quad x \in \mathbb{R}^n,$$

is strongly convex on X (more precisely $N\mu - D$-strongly convex).

In this paper, we will compare the above approach to convexifying of a function that takes only positive values with another approach of multiplying it by a power of a strongly convex function. The latter approach was proposed in 2015 by K. Kurdyka, S. Spodzieja and continued by K. Kurdyka, K. Rudnicka, S. Spodzieja. More precisely, by K. Kurdyka, S. Spodzieja, a positive function f of class $\mathscr{C}^2$ is convex on a compact and convex set $X \subset \mathbb{R}^n$ by multiplying the function f by $(1 + |x|^2)^N$ for some N, and by K. Kurdyka, K. Rudnicka, S. Spodzieja – by multiplying the function f by $\exp(N|x|^2)$.

In Chapter 2 we generalize these results and show that: If $X \subset \mathbb{R}^n$ is a compact and convex set and $f : X \to \mathbb{R}$ is a function of the class $\mathscr{C}^2$ that takes only positive values, then for any strongly convex function $b : \mathbb{R}^n \to \mathbb{R}$ there is $N_0 > 0$ such that for each $N \geq N_0$ and $\xi \in X$, the function

$$\varphi_{N,\xi}(x) = b^N(x - \xi)f(x), \quad x \in \mathbb{R}^n, \tag{1}$$

is strongly convex on X.

In the case where the function f is a polynomial, the exponent N can be estimated effectively in terms of the radius of the set X (i.e., $\sup\{|x| : x \in X\}$) of the modules of the polynomial coefficients and $m = \inf\{f(x) : x \in X\}$ (see wniosek 2.2.2). Therefore, in the case of positive functions on compact and convex sets, both adding a multiple of a strongly convex function to the function and multiplying it by the power of such a function have a similar effect, but the first method uses a smaller coefficient N. If we additionally assume that b is a logarithmically strongly convex function (i,e., $\ln b$ is a strongly convex function), then $\varphi_{N,\xi}$ is also a logarithmically strongly convex function (see wniosek 2.2.9). In the case when X is a semialgebraic, compact and convex set, the coefficients of the polynomials describing X and the coefficients of the polynomial f are integers (or rational numbers), the exponent N can be determined fully efficiently (see Twierdzenia 2.2.10 and 2.2.12).These theorems are obtained using the result of G. Jeronimo, D. Perrucci, E. Tsigaridas (2013).

In Chapter 3, we will compare the classical approach to problem convexifying of a function with the above for any strongly convex function $b : \mathbb{R}^n \to \mathbb{R}$ and a positive function f on a closed and convex (not necessarily bounded) set. In this chapter, we convexiying of a function f by multiplying it by $b(N(x-\xi))$ instead of $b^N(x-\xi)$. This approach simplifies some calculations.

Uwaga 2.2.3 and fakt 3.1.1 are difficult to apply to unbounded sets. Namely, we have

Fact 3.1.2. Let $b : \mathbb{R}^n \to (0,+\infty)$ be a convex function of class $\mathscr{C}^2$, let $X \subset \mathbb{R}^n$ be a convex and closed set, let $f : \mathbb{R}^n \to \mathbb{R}$ be a function of class $\mathscr{C}^2$ and let $N > 0$. If for any $\xi \in X$, the function $\phi_{N,\xi}$ defined by (1) is convex on X, then

$$\partial^2_\beta \phi_{N,\xi}(\xi) = N\partial^2_\beta b(0) + \partial^2_\beta f(\xi) \geq 0 \quad \text{for any } \beta \in S^{n-1}.$$

In particular, $\partial^2_\beta f$, $\beta \in S^{n-1}$, are bounded together from below on X.

Therefore, fakt 3.1.1 and uwaga 2.2.3 can be extended to the case of unbounded sets only if the second-order directional derivatives $\partial^2_\beta f$, $\beta \in S^{n-1}$, together are bounded from below on X. In the general case, instead of the constant N, we need to choose a function that depends on $|\xi|$. Namely, assume that f has a polynomial second-order growth, i.e.,

$$|\partial^2_\beta f(x)| \leq D(1+|x|)^\alpha \quad \text{for } \ x \in X \text{ i } \beta \in S^{n-1},$$

for some $D > 0$, and $\alpha \in \mathbb{N}$ and $b : \mathbb{R}^n \to \mathbb{R}$ is a function of class $\mathscr{C}^k$, $k \geq 2$, and logarithmically μ-strongly convex, $\mu > 0$, such that $0 = \operatorname{argmin}_{\mathbb{R}^n} b$ i $b(0) = 1$, where $\operatorname{argmin}_X b$ is a point in X where b takes the smallest value in X.

Let

$$N(|\xi|) = \frac{D}{\mu}\left(|\xi| + 1 + \sqrt{\frac{\alpha}{\mu}}\right)^\alpha + 1$$

Then, for any $\xi \in \mathbb{R}^n$ a function $\phi_\xi : \mathbb{R}^n \to \mathbb{R}$ defined by

$$\phi_\xi(x) = N(|\xi|)b(x-\xi) + f(x), \quad x \in \mathbb{R}^n,$$

is strongly convex on X (more precisely μ-strongly convex), see Lemma 3.2.1.

In particular, the assertion of the above lemma can be obtained for the function $\psi_\xi(x) = Nb(\xi)b(x-\xi) + f(x)$, for a sufficiently large constant N (see lemat 3.2.5).

In the case when we obtain the convexifying of a function by multiplying it by $x \mapsto b(N(x-\xi))$, where b is a strongly convex function or logarithmically strongly convex, we must of course assume that the function only takes positive values on X. Then we have

Fact 3.1.5. Let $b : \mathbb{R}^n \to (0, +\infty)$ be a μ-strongly convex function of class $\mathscr{C}^2$ such that $0 = \operatorname{argmin}_{\mathbb{R}^n} b$, let $X \subset \mathbb{R}^n$ be a convex and closed set and let $f : \mathbb{R}^n \to (0, +\infty)$ be a function of class $\mathscr{C}^2$. Let

$$\varphi_{N,\xi}(x) = b(N(x-\xi))f(x),$$

where $N > 0$.

(i) If for any $\xi \in X$, $\varphi_{N,\xi}$ is a strictly convex function on X, then $C\partial^2_\beta f(x) \geq -f(x)$ for any $x \in X$ and $\beta \in S^{n-1}$ and some constant $C > 0$.

(ii) If b is the logarithmically μ-strongly convex function and $C\partial^2_\beta f(x) \geq -f(x)$ and $Cf(x) \geq |\partial_\beta f(x)|$ for any $x \in X$, $\beta \in S^{n-1}$ and some constant $C > 0$, then for $N > \frac{2C}{\sqrt{mu}} + \frac{1}{C\mu}$, the function $\varphi_{N,\xi}$ is strictly convex on X.

The main difficulty in applying the above fact is the estimation of the constant C. This difficulty can be overcome when we convexifying of the polynomial. More specifically, let $f \in \mathbb{R}[x]$, where $x = (x_1, \dots, x_n)$ is a system of a variable, be a polynomial of degree d, and let $f = f_0 + \cdots + f_d$, where f_j is a homogeneous polynomial of degree j or zero. Let $f_{d*} = \min_{|x|=1} f_d(x)$. Obviously, $f_{d*} > 0$ if and only if the leading form f_d of the polynomial f takes only positive values $\mathbb{R}^n \setminus \{0\}$. Then we can get a convexifying of polynomial f by multiplying it by the function $b(N(x-\xi))$, $\xi \in \mathbb{R}^n$. Namely, we have

Theorem 3.3.1. *Assume that $f_{d*} > 0$ and there exists $m > 0$ such that*

$$f(x) \geq m \quad \text{for } x \in \mathbb{R}^m.$$

Then there is an effectively compatible N_0 such that for any $N > N_0$ and for any $\xi \in \mathbb{R}^n$ the function $\varphi_{N,\xi} : \mathbb{R}^n \to \mathbb{R}$ defined by

$$\varphi_{N,\xi}(x) = b(N(x-\xi))f(x)$$

is μ-strongly convex in $\mathbb{R}^n$.

Chapters 4, 5 and 6 we deal with iterations of a mapping that assigns to each point the only critical point of the convexifying of function f.

In Chapter 4, we assume that $X \subset \mathbb{R}^n$ is a convex and compact set, and that the function $b : \mathbb{R}^n \to \mathbb{R}$ is strongly convex of class $\mathscr{C}^k$, $k \geq 2$, such that $0 = \operatorname{argmin}_{\mathbb{R}^n}$. Let $f : \mathbb{R}^n \to \mathbb{R}$ be a function of the class $\mathscr{C}^k$. Then there is a number $N \geq 1$ such that for any $\xi \in \mathbb{R}^n$, the function

$$\phi_{N,\xi}(x) = Nb(x-\xi) + f(x)$$

is strongly convex on X. We define a mapping

$$\kappa_N : X \ni \xi \mapsto \operatorname{argmin}_X \phi_{N,\xi} \in \mathbb{R}^n.$$

If

$$X_{f\le r} := \{x \in \mathbb{R}^n : f(x) \le r\} \subset X,$$

we show that (see lemat 4.1.3 and 4.1.4) the following properties hold:

(i) The mapping κ_N is a diffeomorphism of class $\mathscr{C}^{k-1}$ from $X_{f\le r}$ to $Y = \kappa_N(X_{f\le r}) \subset X_{f\le r}$.

(ii)The set of fixed points of $\kappa_N|_{X_{f\le r}}$ is equal to $\Sigma_f \cap X_{f\le r}$, where Σ_f is the set of critical points of f.

In theorem 4.2.1(c)we show that: If $f : \mathbb{R}^n \to \mathbb{R}$ is a semialgebraic function of class $\mathscr{C}^2$ then for any $\xi \in X_{f\le r}$, the limit point $\lim_{\nu\to\infty} \kappa_N^\nu(\xi)$ exists and belongs to Σ_f.

The proof of this theorem, is based on showing the monotonicity of the sequence $f(\xi_\nu)$ (see wniosek 4.1.5) and applying the comparison principle from K. Kurdyka and S. Spodzieja to show that the series $\sum_{\nu=0}^{\infty} \operatorname{dist}(\kappa_N^\nu(\xi), f^{-1}(f(\kappa_N^{\nu+1}(\xi))))$ is convergent. The idea of this proof is based on the proof of Theorem 7.5 from the paper by K. Kurdyka and S. Spodzieja. The proof of this theorem is not a direct transfer of the proof Theorem 7.5 from the paper by K. Kurdyka and S. Spodzieja, because in this paper is considered convexification of f by multiplying it by $(1+|x|^2)^N$, and we consider convexification this function by adding Nb to it.

Assuming that the function f is semialgebraic, the above theorem allows us to define the mapping

$$\kappa_{N,*} : X_{f\le r} \to \Sigma_f \cap X_{f\le r}\,,$$

given by $\kappa_{N,*}(\xi) = \lim_{\nu\to\infty} \kappa_N^\nu(\xi)$.

Assuming that the function f has only one critical value on $X_{f\le r}$, we will show that the mapping $\kappa_{N,*}$ is continuous. Namely, we have

Theorem 4.3.1. *Let $0 \in \operatorname{Int} X_{f\le r}$ and let $f(0)$ be the minimal value of f. Then there exists $f(0) < \delta < r$ such that the sequence κ_N^ν uniformly convergents to $\kappa_{N,*}$ in the set $U = X_{f\le\delta}$. In particular the mapping*

$$\kappa_{N,*}|_U : U \to U \cap \Sigma_f$$

is continuous and $\kappa_{N,*}(\xi) = \xi$ for $\xi \in U \cap \Sigma_f$. Consequently $\kappa_{N,*}|_U$ is a deformation retraction and the set $U \cap \Sigma_f$ is a retract of U.

In Chapter 5, we transfer some properties of the mapping κ_N to the case of unbounded sets. Among other things, in the case where the convexification of the function f is of the form

$$\psi_\xi(x) = N(\xi)b(x-\xi) + f(x), \qquad (\xi, x) \in \mathbb{R}^n \times \mathbb{R}^n$$

and for a convex and closed set X,

$$\kappa(\xi) = \operatorname{argmin}_X \psi_\xi \in X,$$

we show that: For any $\xi \in X$, the point $\kappa(\xi)$ is the anique lower critical point of ψ_ξ on X. A similar fact holds for the function

$$\kappa_N : X \ni \xi \mapsto \operatorname{argmin}_X b(N(x-\xi))f(x).$$

In Chapter 6, we trasver the results from chapter four for the mapping $X_{f\leq r} \ni \xi \mapsto \operatorname{argmin}_{|x|\leq R} b^N(x-\xi)f(x) \in X$, assuming that the set $X_{f\leq r}$ is compact and convex. In this case, all of the above properties of κ_N are true. If we additionally assume that $b(x) = \exp(|x|^2)$ and f is a polynomial, then the mapping κ_N has some additional properties, among others it is an analytic and semialgebraic mapping, i.e. it is a Nash maping. The mapping $\kappa_N : X_{f,r} \to \kappa_N(X_{f,r})$ is the inverse of

$$\kappa_N(X_{f,r}) \ni x \mapsto x + \frac{1}{2Nf(x)}\nabla f(x) \in X_{f,r}\,,$$

so it is an analytic and semialgebraic mapping, i.e., it is a Nash mapping.

At the end of Chapter 6, we deal with the convergence problem of the sequence $\frac{\xi_\nu}{|\xi_\nu|}$, where $\xi_\nu = \kappa_N^\nu(\xi)$, $\nu \in \mathbb{N}$, and $\xi \in \mathbb{R}^n$, i.e. the problem of convergence of a sequence of spherical parts of the sequence ξ_ν. This is a transfer of Rene Thom's problem for the gradient field trajectory (solved by K. Kurdyka, T. Mostowski, A, Parusiński) to the discrete case. We consider this problem under assumption that $\xi_\nu \to 0$, when $\nu \to \infty$, and with some additional quite restrictive assumptions.

Part of the results of this work has already been published in the work of A.N. Abdullah, K. Rosiak, S. Spodzieja. This applies to point 2.2 and Chapter 6.