Monotonic Analysis with Applications to Optimization

A. M. Rubinov

School of Information Technology and Mathematical Sciences and

Centre for Informatics and Applied Optimization,

University of Ballarat

Examination of many problems arising in optimisation and related topics can be successfully accomplished if these problems enjoy a convex structure. Convexity is, however, a restrictive hypothesis so the question arises: is it possible to substitute convexity for another structure at least for some broad classes of non-convex problems. Monotonicity is one of these structures.

Monotonic analysis is a theory of monotonicity that is based on such notions of convex (and abstract convex) analysis as subdifferentials, Fenchel-Moreau conjugacy, polarity. *Many Lipschitz problems can be converted to problems with monotonic data*, so convex problems are much simpler than monotonic problems. However some theoretical schemes of convex duality and some numerical schemes of convex programming can be successfully applied also in monotonic setting.

Abstract convexity

Let H be a set of functions defined on a set X. For a function $f : X \to \mathbb{R}_{+\infty} \equiv \mathbb{R} \cup \{+\infty\}$ consider the set of all H-minorants of f (the support set of f w.r.t H):

$$\text{supp}\,(f, H) = \{h \in H : h \leq f\},$$

where $h \leq f$ means that $h(x) \leq f(x)$ for all $x \in X$.

1. A function $f : X \to \mathbb{R}_{+\infty}$ is called H-convex (abstract convex w.r.t. H) if $f(x) = \sup\{h(x) : h \in \text{supp}\,(f, H)\}, \quad \forall\, x \in X$.

2. The H-subdifferential of a H-convex f at a point y is

$$\partial_H f(y) = \{l \in L : f(x) \geq l(x) + (f(y) - l(y)) \;\; \forall x \in X\}.$$

3. A nonempty set $U \subset X$ is called abstract convex w.r.t. H (or H-convex) if for each $x \in X \setminus U$ there exists $h \in H$ such that $h(x) > \sup_{u \in U} h(u)$.

Monotonic analysis is a branch of abstract convexity, where special sets of elementary functions H are used that lead to some classes of monotone functions and corresponding sets. I consider only the simplest situation ($X = \mathbb{R}^n_+$) and only one of the classes of elementary functions (multiplicative min-type functions).

Let $X = \mathbb{R}^n_+$. For each $l \in \mathbb{R}^n$ consider the function

$$l(x) = \langle l, x \rangle := \min_{i : l_i > 0} l_i x_i.$$

Let L be the set of all such functions.

The function $p : \mathbb{R}^n_+ \to \mathbb{R}_{+\infty}$ is L-convex $\iff$ p is IPH (increasing: $x \geq y \implies p(x) \geq p(y)$ and positively homogeneous: $p(\lambda x) = \lambda p(x)$ for $\lambda > 0$.

IPH functions are monotone analogues of sublinear (convex PH) functions

Examples of IPH functions

$$p(x) = \sum_i l_i x_i, \quad p(x) = \max_i l_i x_i, \quad p(x) = \min_i l_i x_i, \quad (l \geq 0).$$

$$p(x) = (\sum_i x_i^k)^{1/k}, \quad k > 0$$

$$p(x) = \prod_i x_i^{\alpha_i}, \qquad \alpha_i \geq 0, \;\; \sum_i \alpha_i = 1$$

Let p be IPH. Polar function

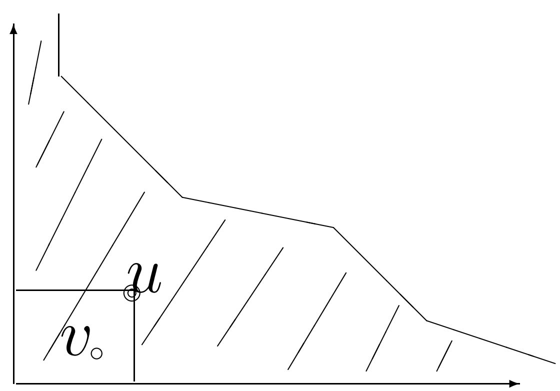
$p^\circ(l) = \max_{x:p(x)>0} \langle l, x\rangle / p(x) = \max_{x:p(x)=1} \langle l, x\rangle.$

Theorem: $p^\circ(l) = \dfrac{1}{p(1/l)}$ where

$$(1/l)_i = \frac{1}{l_i} \quad \text{if} \quad l_i > 0, \qquad (1/l)_i = 0 \quad \text{if} \quad l_i = 0.$$

Let $p(y) > 0$. The subdifferential $\partial_L p(y)$ of an IPH function p at y contains the vector $p(y)/y$.

A set $U \subset \mathbb{R}^n_+$ is called normal if $u \in U$, $0 \le v \le u \implies v \in U$.



Theorem A set U is L-convex if and only if U is closed and normal.

The support set

$$\text{supp}\,(p, L) = \{l : \langle l, x\rangle \le p(x)\ \ \forall x\}$$

of an IPH function p is normal. The level set

$$S(p) = \{x : p(x) \le 1\}$$

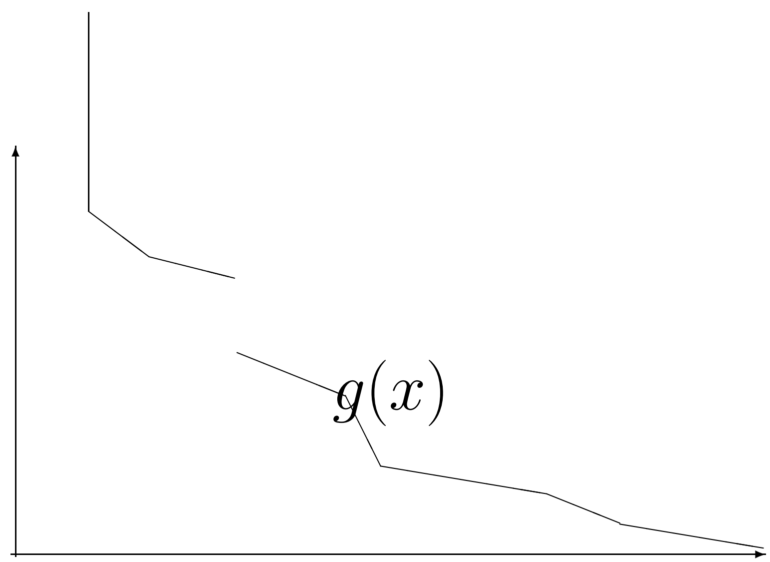
is also normal. We have

$$S(p) = \text{supp}\,(p^\circ, L).$$

Let $g : \mathbb{R}^n_+ \to \mathbb{R}_{+\infty}$ be an usc decreasing function. Then

$$\mathrm{hyp}^+ g = \{(x, \lambda) : p(x) \geq \lambda \geq 0\}$$

is a closed normal subset of $\mathbb{R}^{n+1}$.



Let p be an IPH function such that $\operatorname{supp}(p, L) = \operatorname{hyp}^+ g$. We can use both p and p° in the study of g.

Applications

Monotonic analysis has many applications to optimization, stability of inequality systems, best approximation, inequalities of Hadamard type etc. We consider only applications to optimization.

1. Optimization theory Let $f : \mathbb{R}^n \to \bar{\mathbb{R}}$ and $g : \mathbb{R}^n \to \mathbb{R}^m$. Consider problem (P):

$$\text{minimize} \quad f(x) \quad \text{subject to} \quad g(x) \leq 0.$$

The perturbation (value) function of this problem is:

$$\beta(y) = \inf\{f(x) : g(x) \leq y\}.$$

β is decreasing and $\beta(0)$ is the optimal value of (P). The behavior of β near zero is important for the theoretical examination of (P).

The main properties of Lagrange and penalty functions and its generalizations can be expressed in terms of β.

Assume that $f(x) \geq 0$ and β as usc. Then $\text{hyp}^+\beta$ is a closed normal set which is the level set of an IPH function. We can use the theory of IPH functions in the study of perturbation functions.

Let p be an IPH function defined on $\mathbb{R}_+ \times \mathbb{R}^m_+$.

The function

$$L^+(x,\lambda) = p(f(x), \lambda_1 g_1^+(x), \ldots, \lambda_m g_m^+(x)), \quad x \in \mathbb{R}^n, \lambda \in \mathbb{R}^m_+$$

is called a penalty type function generated by p. (Here $g_i^+(x) = \max(g_i(x), 0)$.)

Consider function

$$L_c^+(x,\lambda) = (f(x)+c)^k + \lambda \sum_{i=1}^{m} g_i^+(x).$$

Then $\lambda \to 0$ as $c \to +\infty$. We can find not very large values of of both c and λ that are good for unconstrained minimization of L_c^+. Functions L_c^+ can be used for the minimization of concave functions subject to linear constraints.

The examination of penalty-type functions generated by IPH functions can be accomplished by means of monotonic analysis.

(See: A. M. Rubinov, X.Q, Yang,

Lagrange-type functions in constrained nonconvex optimization,

Kluwer, 2003)

2. Optimization: Numerical methods

The Cutting Angle Method (CAM) has been developed for the minimization of IPH function on the simplex. This is a *monotonic analogue* of the cutting plane method in convex optimization.

The main idea behind the CAM is the following:

Let p be an IPH function. Let $x^1, \ldots, x^k \in \mathbb{R}^n_{++}$ and $l^k \in \partial_L p(x^k)$.

Let

$p_k(x) = \max_k \langle l^k, x \rangle = \max_k \min_i l_i^k x_i$.

For each compact set $U \subset \mathbb{R}^n_{++}$ it is possible to find points x_k such that $p_k \to p$ uniformly on U. Then minimizers of p_k converge to a minimizer of p as $k \to \infty$.

Let f be a Lipschitz function defined on the unit simplex $S_n = \{x \in \mathbb{R}^n_+ : \sum_i x_i = 1\}$ and let

$$g_c(x) = \|x\|_1 \left(f\left(\frac{x}{\|x\|_1}\right) + c \right), \quad x \in \mathbb{R}^n_+ \setminus \{0\}.$$

Then $g_c(x) = f(x) + c$ for $x \in S_n$.

Theorem There exists $c > 0$ such that g_c is IPH.

The following Lipschitz programming problems are equivalent:

$$\min f(x) \quad \text{s.t.} \quad x \in S_n, \qquad \min g_c(x) \quad \text{s.t.} \quad x \in S_n.$$

A Lipschitz programming problem:

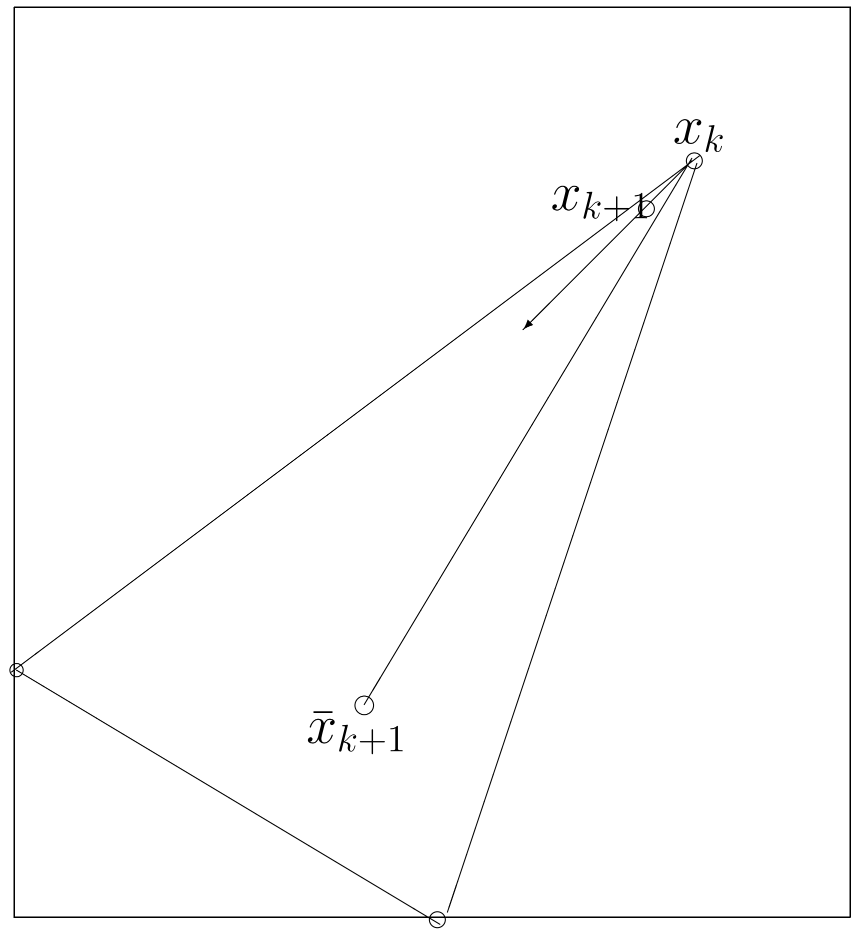
$$\min f(x) \quad \text{s.t.} \quad x \in S$$

can be solved by CAM.

Observation If $f(x)$ is far enough from the global minimum then the value of f can be reduced by CAM very quickly.

A combination of the CAM and the local Discrete Gradient method allows one to find global minimizers or at least deep local minimizers for nonlinear (in particular, nonsmooth) Lipschitz problems.

The main idea of this combination is as follows:


x_k
x_{k+1}
$\bar{x}_{k+1}$