An Integral Optimality Condition for Global Optimization*

Zhang Lili[1,2] Li Jianyu[1,3†] Li Xingsi[1,2]

Abstract An integral optimality condition for global optimization problem is investigated by using a level set auxiliary function. The auxiliary function has one variant that represents an estimated optimal value of the objective function in primal optimization problem and one controlling parameter for accuracy. Necessary and sufficient condition for global optimality in terms of the behavior of the auxiliary function is derived. The integral global optimality condition is obtained via a limiting process of this auxiliary function. Furthermore, if the measure is the Lebesgue measure and the integral region takes a finite subset of the Natural Number set, then this integral global optimality condition divergences to the approximation scheme that used aggregate function to approximate the max-function in the finite minimax problem. So the integral global optimality condition is an extension of this approximation scheme in continuous maximum problem.

Keywords Operations research, integral optimality condition, level set, Lebesque measure, global optimization, aggregate function

Subject Classification (GB/T13745-92) 110.74

全局优化问题的一类积分型最优性条件

张丽丽 [1,2] 李建宇 [1,3†] 李兴斯 [1,2]

摘要 本文通过构造水平集辅助函数对一类积分全局最优性条件进行研究．所构造的辅助函数仅含有一个参数变量与一个控制变量，该参数变量用以表征对原问题目标函数最优值的估计，而控制变量用以控制积分型全局最优性条件的精度．对参数变量做极限运算即可得到积分型全局最优性条件．继而给出了用该辅助函数所刻画的全局最优性的充要条件，从而将原全局优化问题的求解转化为寻找一个非线性方程根的问题．更进一步地，若所取测

收稿日期： 2008 年 5 月 7 日．

* This research is supported by the National Natural Science Foundation of China (No. 10902077)

1. State Key Laboratory of Structural Analysis for Industrial Equipment, Department of Engineering Mechanics, Faculty of Vehicle Engineering and Mechanics, Dalian University of Technology, Dalian 116024, China; 大连理工大学工业装备结构分析国家重点实验室，运载工程与力学学部，工程力学系，大连， 116024

2. Department of Applied Mathematics, Dalian University of Technology, Dalian 116024, China; 大连理工大学应用数学系，大连， 116024

3. School of Mechanical Engineering, Tianjin University of Science & Technology, Tianjin 300222, China; 天津科技大学机械工程学院，天津， 300222

† Corresponding author 通讯作者

度为勒贝格测度且积分区域为自然数集合的一个有限子集，则该积分最优性条件便化为有限极大极小问题中利用凝聚函数对极大值函数进行逼近的近似系统．从而积分型全局最优性条件可以看作是该近似系统从离散到连续的一种推广．

关键词 运筹学，积分型最优性条件，水平集，勒贝格测度，全局优化，凝聚函数

学科分类号(GB/T13745-92) 110.74

1 Introduction

Let X be a topological space, f a real valued function on X and S a closed subset of X, (X,Ω,μ) be a measure space. The problem is to find the supremum of f over S

$$f^* = \sup_{x\in S}\{f(x)\} \tag{1.1}$$

We first give some assumptions:

Assumption 1 *f is continuous on S.*

Assumption 2 *There is a real number such that the intersection of the level set $L_c = \{x|f(x) \geqslant c\}$ and S is nonempty and compact.*

In this case, the set of global maxima is nonempty. i.e.

$$S^* = S\cap L_{f^*} = \{x|f(x) = f^*\} \neq \emptyset \tag{1.2}$$

Consequently, the problem is reduced to this problem

$$(P) \qquad f^* = \max_{x\in S\cap L_c}\{f(x)\} \tag{1.3}$$

In this paper, we shall maintain assumptions 1 and 2.

Optimality conditions, i.e. the conditions by which one can determine whether a point is a candidate of a minimum/maximum, play an important role in the theory of optimization. The optimality conditions for problem (P) have been investigated in the past decades, numerous studies of how to characterize the optimal value of a nonlinear function on a set had been proposed. However, almost all of them are of a local nature. Even then, they have tended to require several levels of differentiability unless some convexity hypotheses are imposed. The search for necessary and sufficient conditions for global optimality without requiring any convexity is an important and needed endeavor. Chew et.al.[1] gave an approach differs from the traditional, derivative-based ones by appealing to the theory of measure and integration. Theoretical algorithms based on the various characterizations of that global optimality were also developed. This integral global optimization theory and algorithms were then applied to vector optimization of upper robust mappings[2]. Hiriart-Urruty[3] gave the well-known approximation that is called the integral global optimality condition of the global maximum of f over the set X.

$$\max_{x\in X}\{f(x)\}=\lim_{p\to\infty}\left(\frac{1}{p}\ln\int_X e^{pf(x)}\mathrm{d}x\right) \tag{1.4}$$

Then, for a convex program, Lasserre[4] applied the approximation scheme (1.4) to the Lagrangian function yielded the logarithmic barrier function(LBF). Especially, for linear programming problem, by using the approximation (1.4) in Fenchel duality, the dual LBF could be retrieved from the primal LBF and vice versa.

The main motivation of this research is to extend the research in [3]. Applying the level set, we define an auxiliary function which has one variant that represents an estimated optimal value of the objective function and one controlling parameter for accuracy. Some properties of this auxiliary function are given. Based on these properties, we investigate the integral global optimality condition which is obtained via a limit process of the auxiliary function. Finally, finding the solution of primal global optimization problem is transformed to finding the root of a nonlinear equation, what is a consequence of the necessary and sufficient condition for global optimality that is characterized by the auxiliary function. Furthermore, an interesting relationship between global integral optimality condition and the approximate scheme that used aggregate function to approximate the max-function in finite minimax problem is revealed. The integral global optimality condition is an extension of this approximate scheme in continuous maximum problem. i.e. this approximate scheme is a special case such that the measure is the Lebesgue measure and the integral region takes a finite subset of the Natural Number set.

The rest of this paper is arranged as follows. Section 2 extends the work in [3], gives the integral global optimality condition and defines an auxiliary function, which can be used to obtain an equivalent equation of the integral global optimality condition. Some relative properties of this auxiliary function are also given. Section 3 proposes a special case, reveals an interesting relationship between integral global optimality condition and an approximate scheme in finite minimax problem. Section 4 concludes the paper.

2 Integral global optimality condition

For our purposes, it will be useful to define the following auxiliary function $F_p(f,c)$.

Definition 2.1 *Suppose $c < f^*$, we define an auxiliary function:*

$$F_p(f,c) = \frac{1}{p}\ln\int_{L_c} e^{pf(x)}\mathrm{d}\mu \tag{2.1}$$

It should be noticed that, although we denote the auxiliary function as $F_p(f,c)$, f is not the variable or parameter, p is a controlling parameter for accuracy, it can be treated as a constant in the deduction process. So c is the only one variant in this auxiliary function.

In our notations, for a set A, $m(A)$ denotes the measure value of this set, i.e. if $\mathrm{d}\mu$ is Lebesgue measure, $m(A)$ is the Lebesgue measure value of set A. The following are the properties of function $F_p(f,c)$.

Proposition 2.1 *For $c < f^*$, we have $\lim\limits_{p\to\infty} F_p(f,c) \geqslant c$.*

Proof By definition, $f(x) \geqslant c$ for $x \in L_c$, so that

$$F_p(f,c) = \frac{1}{p}\ln\int_{L_c} e^{pf(x)}d\mu \geqslant \frac{1}{p}\ln\int_{L_c} e^{pc}\mathrm{d}\mu = c + \frac{1}{p}\ln(m(L_c)) = c \quad (as\, p\to\infty)$$

Proposition 2.2 *If $f^* > c_1 > c_2$, then $F_p(f,c_2) \geqslant F_p(f,c_1)$.*

Proof Let $\underline{c} = \min\limits_{x\in S}\{f(x)\}$, we have

$$\begin{aligned}
F_p(f,c_2) &= \frac{1}{p}\ln\int_{L_{c_2}} e^{pf(x)}\mathrm{d}\mu \\
&= \frac{1}{p}\ln\left(\int_{L_{c_1}} e^{pf(x)}\mathrm{d}\mu + \int_{L_{c_2}/L_{c_1}} e^{pf(x)}\mathrm{d}\mu\right) \\
&\geqslant \frac{1}{p}\ln\left(\int_{L_{c_1}} e^{pf(x)}\mathrm{d}\mu + \left[m\left(L_{c_2}\right) - m\left(L_{c_1}\right)\right] e^{p\underline{c}}\right) \\
&\geqslant \frac{1}{p}\ln\left(\int_{L_{c_1}} e^{pf(x)}\mathrm{d}\mu\right) = F(f,c_1)
\end{aligned}$$

As $f^* > c_1 > c_2$,$L_{c_1} \subseteq L_{c_2}$ and $m\left(L_{c_2}\right) - m\left(L_{c_1}\right) \geqslant 0$, $e^{p\underline{c}} \geqslant 0$.

Lemma 2.1 *Suppose* $\{c_k\}$ *is a increasing sequence tends to* $c \leqslant f^*$ *as* $k \to \infty$. *Then*

$$L_c = \bigcap_{k=1}^{\infty} L_{c_k} = \lim_{k\to\infty} L_{c_k} \tag{2.2}$$

And

$$\lim_{k\to\infty} m\left(L_{c_k}\right) = m\left(L_c\right) \tag{2.3}$$

Proof According to the definition of level sets, we have $L_c \subset L_{c_k} \subset L_{c_{k-1}} \subset ....$ So $\lim\limits_{k\to\infty} L_{c_k} = \bigcap\limits_{k=1}^{\infty} L_{c_k}$. If $x \in \bigcap\limits_{k=1}^{\infty} L_{c_k}$, then $f(x) \geqslant c_k$, for all $k = 1, 2, ...$ Hence $f(x) \geqslant c$, i.e. $x \in L_c$. The conclusion follows immediately from the continuity of measure.

Proposition 2.3 *Suppose* $\{c_k\}$ *is a increasing sequence whose limit is* $c \leqslant f^*$. *Then*

$$F_p(f, c) = \lim_{c_k \uparrow c} F_p(f, c_k) \tag{2.4}$$

Proof According to proposition 2.2, the sequence $\{F_p(f, c_k)\}$ is decreasing and $F_p(f, c_k) \geqslant F_p(f, c)$ for $k = 1, 2, ...$ so that the limit $\lim\limits_{c_k \uparrow c} F_p(f, c_k)$ exists. Moreover,

$$\begin{aligned}
0 &\leqslant \frac{1}{p}\ln\int_{L_{c_k}} e^{pf(x)}\mathrm{d}\mu - \frac{1}{p}\ln\int_{L_c} e^{pf(x)}\mathrm{d}\mu \\
&= \frac{1}{p}\ln\left(\frac{\int_{L_{c_k}} e^{pf(x)}\mathrm{d}\mu}{\int_{L_c} e^{pf(x)}\mathrm{d}\mu}\right) \\
&= \frac{1}{p}\ln\left(\frac{\int_{L_c} e^{pf(x)}\mathrm{d}\mu + \int_{L_{c_k}/L_c} e^{pf(x)}\mathrm{d}\mu}{\int_{L_c} e^{pf(x)}d\mu}\right) \\
&\leqslant \frac{1}{p}\ln\left(1 + \frac{e^{pf^*}(m(L_{c_k}) - m(L_c))}{\int_{L_c} e^{pf(x)}\mathrm{d}\mu}\right)
\end{aligned}$$

As $c_k \to c$, $m(L_{c_k}) - m(L_c) \to 0$. So

$$0 \leqslant \frac{1}{p}\ln\int_{L_{c_k}} e^{pf(x)}\mathrm{d}\mu - \frac{1}{p}\ln\int_{L_c} e^{pf(x)}\mathrm{d}\mu \leqslant 0$$

the Eq. (2.4) is obtained.

We give the following definition of auxiliary function $F_p(f, c)$ at the value of f^*

.

Definition 2.2 *Let $c \leqslant f^*$ and $\{c_k\}$ be a increasing sequence whose limit is f^*. The auxiliary function $F_p(f, f^*)$ is defined to be:*

$$F_p(f, f^*) = \lim_{c_k \uparrow f^*} \left\{ \frac{1}{p} \ln \int_{L_{c_k}} e^{pf(x)} \mathrm{d}\mu \right\} \tag{2.5}$$

The above definition is well defined since $F_p(f, c_k)$ is a decreasing bounded sequence. Moreover, this limit does not depend on the choice of the increasing sequence.

By proposition 2.3, it is clear that definition 2.2 extends definition 2.1 to the case of $c \leqslant f^*$. Similarly, propositions 2.1, 2.2 remain valid for $c \leqslant f^*$.

Now we give the main result about this integral global optimality condition.

Theorem 2.1 *For problem (P), the following are equivalent:*
i) a point x^* is a global maximum with $f^* = f(x^*)$ as the corresponding global maximum value;
ii) $\lim\limits_{p\to\infty} F_p(f, c) \leqslant f^*$ for $c \leqslant f^*$;
iii) $\lim\limits_{p\to\infty} F_p(f, f^*) = f^*$

Proof *i)* $\Rightarrow$ *ii)* f^* is the global maximum value of f, so $f(x) \leqslant f^*$. For $c \leqslant f^*$, we have

$$F_p(f, c) = \frac{1}{p} \ln \int_{L_c} e^{pf(x)} d\mu \leqslant \frac{1}{p} \ln \int_{L_c} e^{pf^*} \mathrm{d}\mu = f^* + \frac{1}{p} \ln(m(L_c)) = f^* (as\, p \to \infty)$$

ii) $\Rightarrow$ *iii)* Combining with proposition 2.1, we obtain the result in *iii)*;
iii) $\Rightarrow$ *i)* Assume f^* is not the global maximum value and $\bar{f}$ is. Then $\bar{f} > f^*$, and $L_{\bar{f}} \subseteq L_{f^*}$, $m\left(L_{\bar{f}}\right) \leqslant m(L_{f^*})$, we have

$$\begin{aligned} F_p(f, f^*) &= \frac{1}{p} \ln \int_{L_{f^*}} e^{pf(x)} \mathrm{d}\mu \\ &\geqslant \frac{1}{p} \ln \int_{L_{\bar{f}}} e^{pf(x)} \mathrm{d}\mu \\ &\geqslant \frac{1}{p} \ln \int_{L_{\bar{f}}} e^{p\bar{f}} \mathrm{d}\mu \\ &= \bar{f} + \frac{1}{p} \ln\left(m\left(L_{\bar{f}}\right)\right) \end{aligned}$$

So

$$f^* = \lim_{p\to\infty} F_p(f, f^*) \geqslant \lim_{p\to\infty} \left\{ \bar{f} + \frac{1}{p} \ln\left(m\left(L_{\bar{f}}\right)\right) \right\} = \bar{f},$$

which is contract with the assumption.

According to theorem 2.1, solving problem (P) is transformed to find the root of a nonlinear equation:

$$g(c) = F_p(f, c) - c = \frac{1}{p} \ln \left(\int_{L_c \cap S} e^{pf(x)} \mathrm{d}\mu \right) - c \tag{2.6}$$

We will give an example to illustrate this result.

Example 2.1 [1] Consider the problem of finding the maximum of $f(x) = x$ over the robust set $S = [1, 4]$. For any c, the level set $L_c = \{x | x \geqslant c\} = [c, +\infty)$, so that

$$F_p(f, c) = \frac{1}{p} \ln \int_{L_c \cap S} e^{pf(x)} \mathrm{d}x = \frac{1}{p} \ln \int_c^4 e^{px} \mathrm{d}x = \frac{1}{p} \ln \left(\frac{1}{p} \left(e^{4p} - e^{pc} \right) \right)$$

Applying the theorem 2.1, we have

$$F(f, f^*) = f^* = \frac{1}{p} \ln \left(\frac{1}{p} \left(e^{4p} - e^{pf^*} \right) \right)$$

Hence, $f^* = \frac{1}{p} \ln \left(\frac{e^{4p}}{p+1} \right) \approx 4$, as $p = 177.4457$ and $S^* = \{4\}$.

3 A special case: The approximation scheme of aggregate function to max-function in finite min-max problem

In order to state the result as broadly as possible, we will work in the general context of $p - norm$ in functional analysis[5−7]. Let (Ω, F, μ) be a measure space, F a Sigma Algebra of subsets of Ω, and μ a measure on Ω.

Definition 3.1 [7] *A norm on a linear space Ω is a non-negative function $\|\bullet\| : \Omega \to R$ with the properties that:*

i) $\|x\| = 0$ if and only if $x = 0$ (faithfulness);

ii) $\|x + y\| \leqslant \|x\| + \|y\|$ for all $x, y \in \Omega$ (triangle inequality);

iii) $\|\alpha x\| = |\alpha| \, \|x\|$ for all $x \in \Omega$ and $\alpha \in K$ (homogeneity)

In Definition 3.1, we are assuming that K is R(real numbers) or C (complex numbers). For a number, $|\bullet|$ denotes the usual absolute value, and for a set, $|\bullet|$ denotes the cardinal number of this set.

Then we have the $p-norm$. For $1 \leqslant p < \infty$, the Linear (vector) space of the equivalence classes of F-measurable functions on Ω that satisfy $\int_{\Omega} |\phi(\omega)|^p \, \mathrm{d}\mu(\omega) < \infty$ are denoted by $L_p(\Omega, F, \mu)$, with vector-space operations defined pointwise and

$$\|\phi\|^p = \left(\int_{\Omega} |\phi(\omega)|^p \, \mathrm{d}\mu(\omega) \right)^{1/p} \tag{3.1}$$

The equivalence classes are the functions which can be represented by μ-measure whose values are uniquely determined except on a set of measure zero. In particular, for $\Omega \subseteq R^n$, μ is the Lebesgue measure, so we have $L_p(\Omega, F, \mu) = L_p(\Omega)$.

In this paper, we use the $p-norm$ on $L_p(\Omega)$. Let $\phi(x) = e^{f(x)}$, the $p-$norm of vector ϕ is defined as

$$\|\phi\|_p = \left(\int_{\Omega} |\phi|^p \, \mathrm{d}\mu \right)^{1/p} = \left(\int_{\Omega} e^{pf(x)} \mathrm{d}\mu \right)^{1/p} \tag{3.2}$$

If $\Omega = N$, μ is the equidistribution, i.e.$\mu(n) = 1\,(\forall n \in N)$, then $L_p(\Omega, F, \mu)$ is the space of sequences $\phi = \{\phi_n\}_{n=1}^{\infty}$ that satisfy $\sum\limits_{n=1}^{\infty} |\phi_n|^p < \infty$, denote this space as l^p, and the norm of this space is $\|\phi\|^p = \left(\sum\limits_{n=1}^{\infty} |\phi_n|^p \right)^{1/p}$. Taking a finite subset N_1 of N, and $|N_1| = k$, then we have $\|\phi\|^p = \left(\sum\limits_{n=1}^{k} |\phi_n|^p \right)^{1/p}$.

Consider a vector-valued function $f(x) = \{f_1(x), ..., f_k(x)\}$, let $\phi_i(x) = e^{f_i(x)}$, $1 \leqslant i \leqslant k$. Obviously the components of vector $\phi(x) = \{\phi_1(x), ..., \phi_k(x)\}$ are always positive, the $p-$norm of vector ϕ is defined as

$$\|\phi(x)\|_p = \left(\sum_{i=1}^{m} |\phi_i(x)|^p \right)^{1/p} = \left(\sum_{i=1}^{m} e^{pf_i(x)} \right)^{1/p} \tag{3.3}$$

Especially, the $\infty-$norm is defined as

$$\max_{1 \leqslant i \leqslant k} \{\phi_i(x)\} = \lim_{p \to \infty} \|\phi(x)\|_p \tag{3.4}$$

By taking a logarithmic operation on both sides of (3.3) and (3.4), respectively, we immediately have

$$\max_{1 \leqslant i \leqslant k} \{f_i(x)\} = \lim_{p \to \infty} \left(\frac{1}{p} \ln \sum_{i=1}^{m} e^{pf_i(x)} \right) \tag{3.5}$$

Which is just the approximate scheme that Li gave in [8], and the function on the right hand was called the aggregate function. This approximation had been investigated broadly and had obtained encouraging results, either from theoretically or algorithmically[8−12].

In [8], Li proved the uniformly approximation of aggregate function to the finite max-function. The max-function can either be the maximum value of several functions (As Li pointed in [9]) or be the maximum of finite real numbers.

Example 3.1 Finding the minimization of Rosenbrock function

$$f(x) = 100\left(x_2 - x_1^2\right)^2 + (x_1 - 1)^2$$

on feasible region $X := \{x_i \in \{0,1\}, i = 1,2\}$, the optimal value is $f^* = 0$. Using the aggregate function

$$\max_{x_1,x_2\in\{0,1\}} (-f(x)) \approx \frac{1}{p} \ln\left(e^{-p} + e^{-100p} + e^{-101p} + 1\right)$$

When $p = 10$,

$$\max_{x_1,x_2\in\{0,1\}} (-f(x)) \approx \frac{1}{p} \ln\left(e^{-p} + e^{-100p} + e^{-101p} + 1\right) = 4.539889921687054\mathrm{e} - 006$$

Considering an optimization problem with finite points in the feasible region:

$$\left(\bar{P}\right) \qquad \max_{x\in X} f(x) \tag{3.6}$$

the aggregate function can be used to solve this problem. As the finiteness, the derivation and proof of this are similar to [8]. However, it should be noticed that, this maximum problem is different from the generally finite maximum problem. The max-function of finite minimax problem is to find the maximum component of a vector, while the max-function of (P) and $\left(\bar{P}\right)$, may be called continuous max-function, is to find a point over a continuous set X which makes the f maximum.

4 Conclusion

A novel integral optimality condition has been investigated. Some properties of the integral global optimality condition are given by using the auxiliary level set function. Necessary and sufficient conditions for global optimality in terms of the behavior of the auxiliary function are derived, so solving the primal optimization problem is transformed to find the root of a nonlinear equation. Some algorithms for nonlinear equation, such as the bisection method, can be used directly. Furthermore, from the measure theory point of view, the approximation scheme of aggregate function to max-function in finite min-max problem is a special case of this integral global optimality condition. This paper is just a preliminary study, further research (including the implementations, algorithms, applications and so on) of this integral global optimality condition remains as an important research subject.

References

[1] Chew Soohong, Zheng Quan. Integral global optimization theory, implementation and application, Lecture notes in economics and mathematical systems[M]. Berlin: Springer-Verlag, 1988.

[2] Shi Shuzhong, Zheng Quan. Vector minimization of upper Robust mappings[J]. *OR Transactions*, 1997, **1**(2): 20-33.

[3] Hiriart-Urruty J. B. Conditions for global optimality[M]// Horst R., Pardalos P. M., Pardalos M. Panos. Handbook of global optimization. Dordrecht: Kluwer, 1994.

[4] Lasserre J. B. Why the logarithmic barrier function in convex and linear programming?[J]. *Operations Research Letters*, 2000, 27: 149-152.

[5] Ward T. B. Functional analysis lecture notes[OL]. [2008-4-20]. http://www.mth.uea.ac.uk/ h720/teaching/functionalanalysis/materials/FAnotes.pdf.

[6] Bonnans F. J., Shapiro A. Perturbation analysis of optimization problems[M]. New York: Springer-Verlag, 2000.

[7] Yosida K. *Functional analysis*[M]. Berlin: Springer-Verlag, 1965.

[8] Li Xingsi. An entropy-based aggregate method for minmax optimization[J]. *Engineering Optimization*, 1992, 18: 277-285.

[9] Li Xingsi. An aggregate function method for nonlinear programming[J]. *Science in China, Ser. A*, 1991, 12: 1283-1288.

[10] Li Xingsi. An Efficient approach to nonlinear Minimax problems[J]. *Chinese Science Bulletin*, 1992, **37**(10): 802-805.

[11] Li Xingsi. A smooth quasi-exact penalty function for nonlinear programming[J]. *Chinese Science Bulletin*, 1992, **37**(10): 806-809.

[12] Pan Shaohua, Li Xingsi. The asymptotic analysis of quasi-exact penalty function method[J]. *Numerical Mathematics, A Journal of Chinese Universities*(in Chinese), 2007, **29**(1): 47-56.