Educational Administration: Theory and Practice
2024,30(2), 1500 - 1506
ISSN:2148-2403
https://kuey.net/

Research Article

Advanced Smoothing Techniques For Penalty-Based Optimization

Dharminder Singh[1*], Amanpreet Singh[2], Salil Kumar[3]

[1]Research Scholar,"Desh Bhagat University, Mandi Gobindgarh Punjab (India)
[2]Post Graduate Department of Mathematics, GSSDGS Khalsa College, Patiala, Punjab (India)"
[3] Department of Mathematics, Sri Guru Gobind Singh Khalsa College, Sector 26,Chandigarh

Citation: Dharminder Singh , et al (2024) Advanced Smoothing Techniques For Penalty-Based Optimization, Educational Administration: Theory and Practice, 30(2), 1500 - 1506
Doi: 10.53555/kuey.v30i2.7765

ARTICLE INFO	ABSTRACT
	This manuscript presents a novel approach for enhancing the effectiveness of objective penalty functions involving inequality constraints. The introduction of a penalty function that lacks smoothness is addressed by applying a novel smoothing technique to ensure its smoothness. The paper analyses the error estimates for both the original and smoothed problems. To verify practicality of the proposed approach, some numerical cases are solved.

INTRODUCTION

The classical optimization problem with inequality constraints can be expressed mathematically as:

$$minimize\ f(x)$$
$$(P)$$
$$\text{such that } g_i(x) \leq 0 \text{ where } i = 1,2, \dots m$$
$$(P)$$

In the given problem $(P), f(x)$ is our objective function with the inequality constraints written as $g_i(x)$. All these functions, namely f, g_i are real-valued, and $x \in R^n$. Let $H_0 = \{x \in R^n \text{ s.t. } g_i(x) \leq 0\}$
In contrast to various other general methods, the penalty function approach serves as an alternative technique for obtaining the solution of equation (P). This problem can be solved by converting it into a series of unconstrained optimization problems, thereby significantly easing the solving process. This methodology establishes a precedent by effectively breaking down the original problem into manageable components. A common penalty function to deal with (P) is:

$$\psi^{\square}(x,\delta) = f(x) + \sigma \sum_{i=1}^{m} max\{g_i(x),0\}^2$$
$$(1)$$

By incorporating the aforementioned penalty function, the original optimization problem (P) is reduced as:

$$Min\ \psi^{\square}(x,\delta) \text{ s.t. } x \in R^n$$
$$(2)$$

The penalty function described in equation (1) possesses smoothness but lacks exactness. By "exactness," we refer to the property where for some $\omega^{\square}$ is optimal solution for both (2) and (P) when $\omega \geq \omega^{\square}$.

HISTORICAL BACKGROUND

The initial groundwork in the formulation of the penalty function concept was established by Zangwill [1], who introduced the classical penalty function in following form:

$$\psi_\sigma(x) = f(x) + \sigma \sum_{i=1}^{m} max\{g_i(x),0\}$$
(3)

And the optimization problem for equation (3) reduces to:

$$Min\ \psi(x,\delta) \text{ s.t. } x \in R^n$$

The penalty function presented in (3) has the potential to be exact, but only under specific conditions. An exact penalty function was developed by Hans and Mangasarian (1979) [2] and introduced the concept of exact penalty functions. Furthermore, using precise penalty functions as inspiration, Rosenberg(1981) [3] suggested a globally convergent approach for convex programming.
During the exploration of algorithms for solving penalty problems, it was observed that gradually increasing the penalty parameter led to non-differentiability of penalty functions [1,3,4,5]. Consequently, smoothing becomes essential for penalty functions to facilitate the application of Newton methods or gradient-based methods. A smoothed exact penalty function was introduced by Pinar and Zenios (1995) [5] for use in the solution of convex constrained optimization problems that require both convex objective and constraint functions.
Smoothing nonlinear penalty functions for limited optimization problems was the subject of an article written by Yang et al. (2003) [6], which was published in a journal with a similar focus. In addition, Meng et al. (2004) [7] presented a technique for smoothing exact penalty functions within the framework of inequality constrained optimization issues.
The presence of smooth penalty functions is typically preferred in optimization problem solving due to the inherent lack of smoothness in exact penalty functions. Consequently, various innovative strategies have emerged in the field of exact penalty functions as discussed in [7,8,9,10,11]. The SPFM technique has been widely studied and introduced by Fiaccio and McCormick [8] as a general approach.
Building upon the concept of SPFM, Meng et al. [12] proposed and extensively investigated an objective function penalty method. Their research focused on utilizing penalty functions associated with the objective function to effectively tackle constrained optimization problems. Their approach offers a promising alternative to traditional penalty methods and presents potential advancements in the field of optimization. The formula for the penalty function represented by [12] is:

$$Q(x,L) = (f(x) - L)^c + \sum_{i=1}^{m} max\{g_i(x),0\}$$

In a previous study [12], it is prove that the penalty function exhibits favorable characteristics of smoothness. Additionally, the accuracy of the objective penalty function was mathematically established, affirming its effectiveness in optimization problems.

Moreover, researchers have explored various other forms penalty function that are constructed through a union of the objective function and constraint penalty. This amalgamation allows for the formulation of penalty functions that is helpful in providing a comprehensive approach to problem solving.

In the present article, the focus will be on discussing the process of smoothing the objective penalty function. By employing smoothing techniques, the objective penalty function can be transformed into a continuous and differentiable function, thereby enhancing its applicability in optimization algorithms. The article aims to explore various methods and approaches for achieving this smoothing effect.

A SECOND ORDERED SMOOTH PENALTY FUNCTION

Let us consider $p: R \to R$ given as:

$$p(t) = \begin{cases} 0 & t \le 0 \\ t^{\frac{2}{3}} & t \ge 0 \end{cases}$$
(4)

The function $p(t)$ is exact but not smooth. So to make it smooth write the optimization problem for it as:

$$\psi_\sigma(x) = f(x) + \sigma \sum_{i=1}^{m} p(g_i(x))$$
(5)

the associated smooth penalty optimization problem reduces as:

$$minimize\ \psi_{\sigma}(x) \text{ s.t. } x \in R^n \tag{6}$$

From the definition in (4), clearly, the function $p(t)$ on R^1 does not fall into the class of continuous functions. We propose the introduction of a new function that possesses the ideal characteristics of continuity and differentiability in order to avoid this limitation. Specifically, to find a function that possesses a continuous first-order derivative is our main objective. To fulfill these criteria, we define the smoothing function as follows:

$$p_{\varepsilon}(t) = \begin{cases} 0, & \text{if} \quad t \leq 0 \\ \dfrac{t^{4/3}}{2\varepsilon^{2/3}}, & \text{if} \quad t > 0 \text{ and } t \leq \varepsilon \\ t^{2/3} - \dfrac{\varepsilon^{2/3}}{2}, & \text{if} \quad t > \varepsilon \end{cases} \tag{7}$$

We prove that above $p_{\varepsilon}(t)$ is continuously differentiable and its derivative is given by:

$$p'_{\varepsilon}(t) = \begin{cases} 0 & t \leq 0 \\ \dfrac{2t^{1/3}}{3\varepsilon^{2/3}} & 0 \leq t \leq \varepsilon \\ \dfrac{1}{3t^{1/3}} & t \geq \varepsilon \end{cases} \tag{8}$$

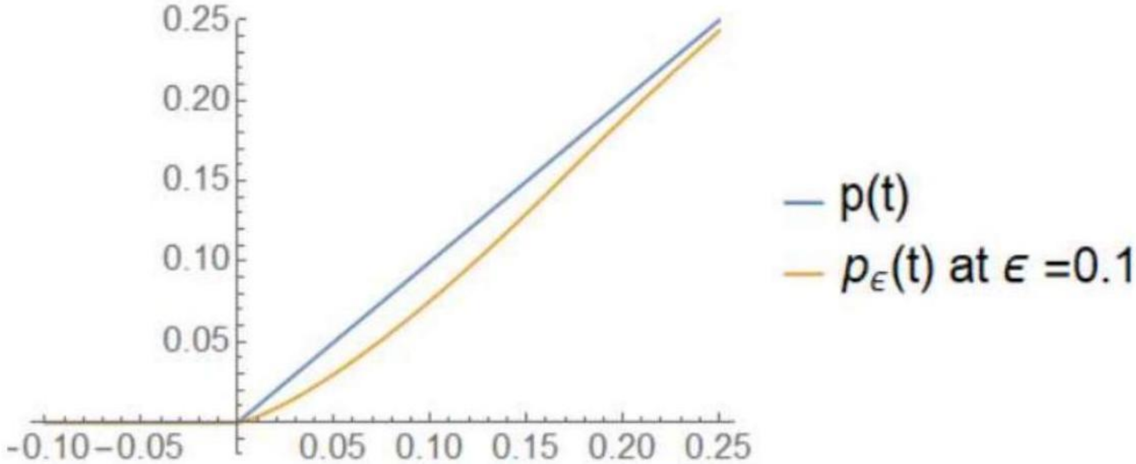

FIGURE 1. The behaviour of $p_{\varepsilon}(t)$ at $\varepsilon = 0.1$ and $p(t)$

The smoothing function introduced earlier exhibits remarkable properties of continuity and differentiability, as elucidated in the theorem presented below.

Proposition 1. For any $\varepsilon > 0$ we prove that:

1. $p_{\varepsilon}(t)$ belongs to C^1 on R
2. $\forall t \in R, p(t) \geq p_{\varepsilon}(t)$
3. $\lim_{\varepsilon \to 0} p_{\varepsilon}(t) = p(t)$

Proof. 1. The first part of the above theorem can be verified by showing the function to be continuous and differentiability at 0 and ε.

Continuity at 0

$LHL = Lt_{t\to 0^-} p_{\varepsilon}(t) = Lt_{t\to 0^-} 0 = 0$ $RHL = Lt_{t\to 0^+} p_{\varepsilon}(t) = Lt_{t\to 0^+} \frac{t^{4/3}}{2\varepsilon^{2/3}} = Lt_{h\to 0} \frac{1}{2}\frac{(0+h)^{4/3}}{\varepsilon^{2/3}} = 0$ Thus $LHL = RHL$

Also $p_{\varepsilon}(0) = 0$

Hence the function is continuous at 0

Continuity at ε :

$LHL = Lt_{t\to\varepsilon^-} p_\varepsilon(t) = Lt_{t\to\varepsilon^-} \frac{t^{4/3}}{2\varepsilon^{2/3}} = Lt_{h\to 0} \frac{1}{2}\frac{(\varepsilon-h)^{4/3}}{\varepsilon^{2/3}} = \frac{1}{2}\varepsilon^{2/3}$ $RHL =$

$Lt_{t\to\varepsilon^+} p_\varepsilon(t) = Lt_{t\to\varepsilon^+} t^{2/3} - \frac{\varepsilon^{2/3}}{2} = Lt_{h\to 0}(\varepsilon+h)^{2/3} - \frac{1}{2}\varepsilon^{2/3} = \varepsilon^{2/3} - \frac{1}{2}\varepsilon^{2/3} = \frac{1}{2}\varepsilon^{2/3}$ Thus LHL RHL, Also $p_\varepsilon(\varepsilon) = \frac{1}{2}\varepsilon^{2/3}$

Now the continuous first-order differentiability of the function $p_\varepsilon(t)$ will be demonstrated. Now from equation (11) we have

$$p'_\varepsilon(t) = \begin{cases} 0 & t \le 0 \\ \dfrac{2t^{1/3}}{3\varepsilon^{2/3}} & 0 \le t \le \varepsilon \\ \dfrac{2}{3t^{1/3}} & t \ge \varepsilon \end{cases}$$

Continuity at 0

$LHL = Lt_{t\to 0^-} p'_\varepsilon(t) = 0$ $RHL = Lt_{t\to 0^+} p'_\varepsilon(t) = Lt_{h\to 0} \frac{2(0+h)^{1/3}}{3\,\varepsilon^{2/3}} = 0$ Hence LHL = RHL, Also $p'_\varepsilon(0) = 0$

Hence the function is continuous at 0

Continuity at ε

$LHL = Lt_{t\to\varepsilon^-} p'_\varepsilon(t) = Lt_{t\to\varepsilon^-} \frac{2t^{1/3}}{3\varepsilon^{2/3}} = Lt_{h\to 0} \frac{2(\varepsilon-h)^{1/3}}{3\,\varepsilon^{2/3}} = \frac{2}{3}\varepsilon^{-1/3}$ $RHL = Lt_{t\to\varepsilon^+} p_\varepsilon'(t) = Lt_{t\to\varepsilon^+} \frac{2}{3t^{1/3}} = Lt_{h\to 0}\frac{2}{3}(\varepsilon+h)^{-1/3} = \frac{2}{3}\varepsilon^{-1/3}$ Thus LHL RHL.

Also $p'_\varepsilon(\varepsilon) = \frac{2}{3}\varepsilon^{-1/3}$

Hence the function $p'_\varepsilon(t)$ is continuous at 0 and ε.

Hence the function is first order differentiable.

2. Now consider

$$p(t) - p_\varepsilon(t) = \begin{cases} 0, & \text{if} \quad t \le 0 \\ t^{2/3} - \dfrac{t^{4/3}}{2\varepsilon^{2/3}}, & \text{if} \quad 0 \le t \le \varepsilon \\ \dfrac{\varepsilon^{2/3}}{2}, & \text{if} \quad t \ge \varepsilon \end{cases}$$

Now when $0 \le t \le \varepsilon$

Define $U(t) = t^{2/3} - \frac{t^{4/3}}{2\varepsilon^{2/3}}$ then

$U'(t) = \frac{2}{3t^{1/3}} - \frac{2t^{1/3}}{3\varepsilon^{2/3}}$ Thus $U'(t) > 0$.

Also $U(0) = 0$ and $U(\varepsilon) = \varepsilon^{2/3} - \frac{\varepsilon^{4/3}}{2\varepsilon^{2/3}} = \varepsilon^{2/3} - \frac{1}{2}\varepsilon^{2/3} = \frac{1}{2}\varepsilon^{2/3}$

Hence $0 \le p(t) - p_\varepsilon(t) \le \frac{1}{2}\varepsilon^{2/3}$.

3. From the proof of second part, it is quite obvious.

Let $\psi_{\sigma,\varepsilon}(x) = f(x) + \sigma \sum_{i=1}^{m} p_\varepsilon(g_i(x))$

This smooth penalty optimization problem is written as:

$$\min \psi_{\sigma,\varepsilon}(x) \text{ so that } x \in R^n \tag{9}$$

Subsequent theorem reveals the connection between (9) and (6) objective functions.

Proposition 2. Let $x \in H_0$, and $\varepsilon > 0$, in this case we prove that

$$0 \le \psi_\sigma(x) - \psi_{\sigma,\varepsilon}(x) \le \frac{1}{2} m\sigma\varepsilon^{\frac{2}{3}}$$

Proof. From Proposition 1, we have

$$0 \le p(g_i(x)) - p_\varepsilon(g_i(x)) \le \frac{1}{2}\varepsilon^{2/3}$$

Thus we see that

$$0 \le \psi_\sigma(x) - \psi_{\sigma,\varepsilon}(x) \le \frac{1}{2} m\sigma\varepsilon^{\frac{2}{3}}$$

Proposition 3. Consider the sequence positive numbers $\langle \varepsilon_j \rangle$ such that it converges to zero as j tends to infinity. Also suppose that for minimization problem $min_{x\in H_0}\psi_{\sigma,k}(x)$. Then $min_{x\in H_0}\psi_{\sigma}(x)$ has the optimal solution $\acute{x}$, where $\acute{x}$ is the limit point of sequence $\langle x_j \rangle$.
Proof. From given hypothesis, we observe that

$$\psi_{\sigma,\varepsilon_j}(x) \geq \psi_{\sigma,\varepsilon_j}(x_j), \forall x \in H_0$$

And Proposition 2 implies

$$\psi(x) \leq \psi_{\sigma}(x)$$

and

$$\psi_{\sigma}(x) \leq \psi_{\sigma,\varepsilon_j}(x) + \frac{1}{2}m\sigma\varepsilon_j^{\frac{2}{3}}$$

It follows that

$$\psi_{\sigma}(x_j) \leq \psi_{\sigma,\varepsilon_j}(x_j) + \frac{1}{2}m\sigma\varepsilon_j^{\frac{2}{3}} \leq \psi_{\sigma,\varepsilon_j}(x) + \frac{1}{2}m\sigma\varepsilon_j^{\frac{2}{3}} \leq \psi_{\sigma}(x) + \frac{1}{2}m\sigma\varepsilon_j^{\frac{2}{3}}$$

Let $j \to \infty$, we observe that

$$\psi_{\sigma}(\acute{x}) \leq \psi_{\sigma}(x)$$

This concludes the argument.
Proposition 4. Let $x_{\sigma}^{\square} \in H_0$ and $\acute{x}_{\sigma,\varepsilon} \in H_0$ in (6) and (9) respectively be the optimal solutions of problem with $\sigma > 0$ and $\varepsilon > 0$. Then $\psi_{\sigma}(x_{\sigma}^{\square}) - \psi_{\sigma,\varepsilon}(\acute{x}_{\sigma,\varepsilon})$ is bounded above by $\frac{1}{2}m\sigma\varepsilon^{\frac{2}{3}}$
When an error is sufficiently minor, the two theorems above imply that a solution to (9) is likewise a solution to (6).
After analyzing the previous discussion, it becomes evident that a solution that is approximately optimal for the problem (9) can also be considered approximately optimal for the problem (6), given that the solution for SP meets the feasibility criterion of ε.

ALGORITHM FOR SMOOTH PENALTY FUNCTION

The following algorithm is proposed as a solution for the given problem.
Step 1 Choose an initial point labeled as x_0. Set a stopping tolerance represented by $\varepsilon > 0$ which is a small positive value indicating the desired level of accuracy for the solution. Assign positive values for ε_0 and σ_0. Select two additional values: λ, which should be a decimal between 0 and 1 , and N which should be greater than 1 . Start the iteration with initial value $j = 0$ and follow the next step.
Step 2 Utilize the current point x_j (obtained from the previous step) as starting solution. Solve $min_{x\in R^{\square}}\psi_{\sigma_j,\varepsilon_j}(x)$ to get the next solution $x_j^{\square}$. The algorithm's subsequent steps or iterations can be performed after $x_j^{\square}$ has been achieved.
Step 3 If we get the desired ε-feasible solution as $x_j^{\square}$, in that case, the solution is close to optimal. Otherwise, write $x_{j+1} = x_j^{\square}$ with $\varepsilon_{j+1} = \lambda\varepsilon_j$ and $\sigma_{j+1} = N\sigma_j$, then follow the second step with $j = j + 1$.

NUMERICAL EXAMPLES

Now we solve a numerical example with the help of suggested algorithm using Mathematica software. Example 1. The Rosen-Suzen problem in [12] is

$$min f(x) = x_1^2 + x_2^2 + 2x_3^2 + x_4^2 - 5x_1 - 5x_2 - 21x_3 + 7x_4$$
$$\text{s.t. } g_1(x) = 2x_1^2 + x_2^2 + x_3^2 + 2x_1 + x_4 - 5 \leq 0$$
$$g_2(x) = x_1^2 + x_2^2 + x_3^2 + x_4^{+x_1-x_2+x_3-x_4-8\leq 0}$$
$$g_3(x) = x_1^2 + 2x_2^2 + x_3^2 + 2x_4^2 - x_1 - x_4 - 10 \leq 0$$

We solve this equation using Mathematica. Let $x^0 = (0,0,0,0)$ We take initial value of penalty parameter $\sigma_0 = 3, \varepsilon_0 = 0.2, \lambda = 0.1$ and $N = 3$
The values against distinct values of penalty parameters are presented in the table (I). From the table, we observe that above the proposed algorithm results in better value of objective function in comparison to l_1 and l_2 penalty function algorithm. In the fourth iteration of the proposed algorithm, the value of $f(x)$ was better than the value obtained in twenty-fifth iteration of the algorithm proposed in [13]

TABLE I. Numerical results using Mathematica

k	x^{k+1}	σ_k	ε_k	$f(x^{k+1})$
0	$(0.169255, 0.834042, 2.012210, -0.972317)$	3	0.2	-44.2534
1	$(0.169480, 0.835149, 2.00954, -0.966767)$	9	0.02	-44.2339
2	$(0.169559, 0.835532, 2.00863, -0.964877)$	27	0.002	-44.2338

Example 2. Consider the problem given in [13] is

$$\min f(x) = 1000 - x_1^2 - 2x_2^2 - x_3^2 - x_1x_2 - x_1x_3$$
$$\text{s.t. } g_1(x) = x_1^2 + x_2^2 + x_3^2 - 25 = 0$$
$$g_2(x) = (x_1 - 5)^2 + (x_2 - 5)^2 + (x_3 - 5)^2 - 25 \le 0$$

The algorithm begins with the value $x^0 = (0,0,0)$, gradually the value of penalty parameter is increased 10 times and tolerance value is decreased by multiplier 0.01 at each iteration step of the algorithm. The results using Mathematica are calculated below:

TABLE II. Numerical results using Mathematica

k	x^{k+1}	ρ_k	ε_k	$f(x^{k+1})$
0	(2.51017,4.22738,0.967762)	10	0.1	944.097939
1	(2.50102,4.22196,0.964757)	100	0.001	944.203874
2	(2.5001,4.22138,0.964624)	1000	0.00001	944.214474
3	(2.50001,4.22132,0.964611)	10000	0.0000001	944.215534

CONCLUSION

The article at hand introduces a smoothing objective penalty function that effectively assesses errors related to its usage. The article also presents supporting evidence and proofs regarding the error assessment. Additionally, the article for the constrained optimization problem presents an algorithmic design sequence that relies on the objective penalty function so as to obtain a solution. In addition, the essay provides further proofs that the algorithm converges globally.

REFERENCES

1. W. I. Zangwill, "Non-linear programming via penalty functions," Management science 13, 344-358 (1967).
2. S.-P. Han and O. L. Mangasarian, "Exact penalty functions in nonlinear programming," Mathematical programming 17, 251-269 (1979).
3. E. Rosenberg, "Globally convergent algorithms for convex programming," Mathematics of Operations Research 6, 437-444 (1981).
4. M. Ç. Pinar and S. A. Zenios, "On smoothing exact penalty functions for convex constrained optimization," SIAM Journal on Optimization 4, 486-511 (1994).
5. S. A. Zenios, M. C. Pinar, and R. S. Dembo, "A smooth penalty function algorithm for network-structured problems," European Journal of Operational Research 83, 220-236 (1995).
6. X. Yang, Z. Meng, X. Huang, and G. Pong, "Smoothing nonlinear penalty functions for constrained optimization problems," (2003).
7. Z. Meng, Q. Hu, C. Dang, and X. Yang, "An objective penalty function method for nonlinear programming," Applied mathematics letters 17, 683 – 689 (2004).
8. A. V. Fiacco and G. P. McCormick, Nonlinear programming: sequential unconstrained minimization techniques (SIAM, 1990).
9. R. Fletcher, "A model algorithm for composite nondifferentiable optimization problems," in Nondifferential and Variational Techniques in Optimization (Springer, 1982) pp. 67-76.
10. R. Fletcher, "An exact penalty function for nonlinear programming with inequalities," Mathematical Programming 5, 129-150 (1973).
11. D. Mauricio and N. Maculan, "A boolean penalty method for zero-one nonlinear programming," Journal of Global Optimization 16, 343-354 (2000).
12. Z. Meng, C. Dang, M. Jiang, and R. Shen, "A smoothing objective penalty function algorithm for inequality constrained optimization problems," Numerical functional analysis and optimization 32, 806-820 (2011).

13. J. Lasserre, "A globally convergent algorithm for exact penalty functions," European Journal of Operational Research 7, 389-395 (1981).
14. M. S. Bazaraa, H. D. Sherali, and C. M. Shetty, Nonlinear programming: theory and algorithms (John Wiley & Sons, 2013).
15. Z. Wu*, F. Bai, X. Yang, and L. Zhang, "An exact lower order penalty function and its smoothing in nonlinear programming," Optimization 53, 51-68 (2004).