

PENALTY APPROACH FOR SOLVING CONSTRAINED NONLINEAR PROBLEMS

A.S. Soomro, Muhammad Zakaullah Khan and Doulat Ram Kundnani

Institute of Mathematics & Computer Science, University of Sindh, Jamshoro

Abstract

In this paper, the penalty function method is used for solving constrained nonlinear problems. This method transforms a constrained problem into a sequence of unconstrained problems. Then Newton's search method is used to minimize it.

Keywords: Unconstrained Optimization, Nonlinear Programming, Penalty Function Method.

1. Introduction

We have used the penalty function method for solving constrained nonlinear programming problems. This method is used here to transform a constrained problem into a sequence of unconstrained problems and is then minimized by the Newton's search method.

The penalty method can be used to solve the problem to minimize $f(x)$ subject to $g(x) \leq 0$ and $h(x) = 0$, and $x \in X$. In a penalty function method a penalty term is added to the objective function for any violation of the constraints. This method generates a sequence of infeasible points whose limit is an optimal solution to the original problem. This method can be used for inequality and equality constrained problems.

In this paper, we have used the penalty method for solving a constrained problem with a single constraint $h(x) = 0$, and have obtained a very close solution to the primal problem.

Consider a primal problem with a single constraint $h(x) = 0$.

Minimize $f(x)$

Subject to $h(x) = 0$, and $x \in X$.

The penalty problem is

Minimize $f(x) + \mu\alpha(x)$, $\mu > 0$, where $\alpha(x) = h^2(x)$.

We refer to the function $f(x) + \mu\alpha(x)$ as the auxiliary function.

2. Penalty Function Method

The penalty approach, (see [2], [3] and [4]), is used to transform a constrained problem into a single unconstrained problem or into a sequence of unconstrained problems. The constraints are placed into the objective function via a penalty parameter in such a way that penalizes any violation of the constraints. Consider the following problem with the single constraint $h(x) = 0$.

Minimize $f(x)$

Subject to $h(x) = 0$

Above problem can be replaced by the following unconstrained problem:

Minimize $f(x) + \mu h^2(x)$

Subject to $x \in E_n$

We see that an optimal solution must have $h^2(x)$ close to zero, otherwise a large penalty $\mu h^2(x)$ will be incurred. Now consider the following problem with the single inequality constraint $g(x) \leq 0$.

Minimize $f(x)$
 Subject to $g(x) \leq 0$

It is clear that the form Minimize $f(x) + \mu g^2(x)$ is not appropriate, since a penalty will be incurred whether $g(x) < 0$ or $g(x) > 0$. A penalty is desired only if the point x is not feasible, that is $g(x) > 0$. The unconstrained problem is therefore given by

Minimize $f(x) + \mu \text{maximum}\{0, g(x)\}$
 Subject to $x \in E_n$

If $g(x) \leq 0$, then $\text{maximum}\{0, g(x)\} = 0$ and no penalty is incurred. On the other hand, if $g(x) > 0$, then $\text{maximum}\{0, g(x)\} > 0$ and the penalty term $\mu g(x)$ is realized.

In general, a suitable penalty function must incur a positive penalty for infeasible points and no penalty for feasible points. If the constraints are of the form $g_i(x) \leq 0, i = 1, \dots, m$ and $h_i(x) = 0, i = 1, \dots, k$ then a suitable penalty function $\alpha(x)$ is defined by

$$\alpha(x) = \sum_{i=1}^m [\text{maximum}\{0, g_i(x)\}]^p + \sum_{i=1}^k |h_i(x)|^p, p > 0 \quad (2.1)$$

We refer to $f(x) + \mu \alpha(x)$ as the auxiliary function.

2.1 Lemma

Suppose that $f, g_1, \dots, g_m$ and $h_1, \dots, h_k$ are continuous functions on E_n . Let X be a nonempty set in E_n . Let α be a continuous function on E_n satisfying (2.1) and suppose that for each μ , there is an $x_\mu \in X$ such that $\theta(\mu) = f(x_\mu) + \mu \alpha(x_\mu)$, then the following statements hold:

(i) $\inf\{f(x): x \in X, g(x) \leq 0, h(x) = 0\} \geq \sup_{\mu \geq 0} \theta(\mu)$, where $\theta(\mu) = \inf\{f(x) + \mu \alpha(x): x \in X\}$, g is a vector function whose components are $g_1, \dots, g_m$ and h is a vector function whose components are $h_1, \dots, h_k$.

(ii) $f(x_\mu)$ is a nondecreasing function of $\mu \geq 0$, $\theta(\mu)$ is a nondecreasing function of μ , and $\alpha(x_\mu)$ is a nonincreasing function of μ .

Proof: See [1].

2.2 Theorem

Consider the following problem:

Minimize $f(x)$
 Subject to $g_i(x) \leq 0, i = 1, \dots, m$
 $h_i(x) = 0, i = 1, \dots, k$
 $x \in X$

where $g_1, \dots, g_m, h_1, \dots, h_k$ are continuous functions on E_n , and let X be a nonempty set in E_n . Suppose that the problem has a feasible solution, and let α be a continuous function satisfying (2.1). Furthermore, suppose that for each μ , there exists a solution $x_\mu \in X$ to the problem to minimize $f(x) + \mu \alpha(x)$ subject to $x \in X$, and $\{x_\mu\}$ is contained in a compact subset of X . Then,

$$\inf\{f(x): x \in X, g(x) \leq 0, h(x) = 0\} \geq \sup_{\mu \geq 0} \theta(\mu) = \lim_{\mu \rightarrow \infty} \theta(\mu)$$

where $\theta(\mu) = \inf\{f(x) + \mu \alpha(x): x \in X\} = f(x_\mu) + \mu \alpha(x_\mu)$. Furthermore the limit x^* of any convergent subsequence of $\{x_\mu\}$ is an optimal solution to the original problem, and $\mu \alpha(x_\mu) \rightarrow 0$, as $\mu \rightarrow \infty$.

Proof: See [1].

Corollary

If $\alpha(x_\mu) = 0$ for some μ , then x_μ is an optimal solution to the problem.

For proof of the above corollary refer [1].

3. Algorithm

Step 1. Let $\epsilon > 0$ be a termination. Choose an initial point x_1 , a penalty parameter $\mu_1 > 0$, and a scalar $\beta > 1$. let $k = 1$ and go to step 2.

Step 2. Starting with x_k , solve the following problem:

$$\begin{aligned} &\text{Minimize } f(x) + \mu_k \alpha(x) \\ &\text{Subject to } x \in X \end{aligned}$$

Let x_{k+1} be an optimal solution and go to step 3.

Step 3. If $\mu_k \alpha(x_{k+1}) < \epsilon$ stop, otherwise, let $\mu_{k+1} = \beta \mu_k$, replace k by $k+1$ and go to step 2.

Example 1. Consider the following primal problem:

$$\begin{aligned} &\text{Minimize } (x_1 - 2)^4 + (x_1 - 2x_2)^2 \\ &\text{Subject to } x_1^2 - x_2 = 0 \\ &\quad x \in X \end{aligned}$$

The optimal solution to the above problem is $x^* = (1, 1)$ and $f(x^*) = 2$, where $x^* = (x_1, x_2)^t$.

The Penalty problem is

$$\text{Minimize } (x_1 - 2)^4 + (x_1 - 2x_2)^2 + \mu_k (x_1^2 - x_2)^2$$

The starting point is taken as $x_1 = (2.0, 1.0)$, the initial penalty parameter and the scalar β are taken as $\mu_1 = 0.1$ and 10.0 , where the objective function is 0.0 .

By applying the penalty method, we obtained the optimal solution which is reported in the following table.

K	μ_k	$x_{k+1} = x_{\mu k}$	$f(x_{k+1})$	$\alpha(x_{\mu k}) = h^2(x_{\mu k})$	$\theta(\mu_k)$	$\mu_k \alpha(x_{\mu k})$
1	0.1	1.4293488, 0.7391304	0.1084365	1.7001670	0.2784532	0.1700167
2	1.0	1.1924640, 0.7501566	0.5200250	0.4513329	0.9713579	0.4513329
3	10.0	1.0690540, 0.9581832	1.4690400	0.03411141	1.81011540	0.3411141
4	100.0	1.0267500, 1.0316930	1.9718330	0.0005072607	2.0225588	0.05072607
5	1000.0	1.0018350, 1.0010520	1.9932190	0.00000861177	2.0000900	0.00681177

The procedure could have been stopped at the fourth iteration when the point $x_5^t = (1.0267500, 1.0316930)$ is reached, where $\alpha(x_{\mu k}) = 0.0005072607$. However to show more clearly that $\mu_k \alpha(x_{\mu k})$ does converge to zero, one more iteration was carried out, where $x_6^t = (1.0018350, 1.006520)$ is an accurate point and the value of an objective function $f(x) = 1.9932190$ is very close to an exact value of the objective function. Noting that $f(x_{\mu k})$ and $\theta(\mu_k)$ are

nondecreasing functions of μ_k and $\alpha(x_{\mu k})$ is a nonincreasing function of μ_k .

4. Computational Difficulties Associated with Penalty Functions

The solution to the penalty problem can be made arbitrarily close to the optimal solution of the original problem by choosing μ sufficiently large. However, if we choose a very large μ for solving a

penalty problem, we may get into some computational difficulties of ill-conditioning. Most algorithms using penalty functions employ a sequence of increasing penalty parameters. With each new value of the penalty parameter, an optimization technique is employed, starting with the optimal solution corresponding to the previously chosen parametric value.

References

- [1] Bazaraa, M.S. and Shetty, C. M. "Nonlinear Programming: Theory and Algorithms", John Wiley & Sons, Inc., Canada, (1979).
- [2] Bazaraa, M.S. Goode, J.J. "Sufficient conditions for a globally exact penalty function without convexity", Math. Prog. Study, 19 (1982) 1-15.
- [3] Zheng Quan; Zhang Lian-Sheng, "Penalty functions and global optimization problems with inequality constraints", Mathematical Numerical Science, 222, (1980).
- [4] Zhangwill, W. I. "Nonlinear programming via penalty functions", Management Science, 13 (1967) 344-358.