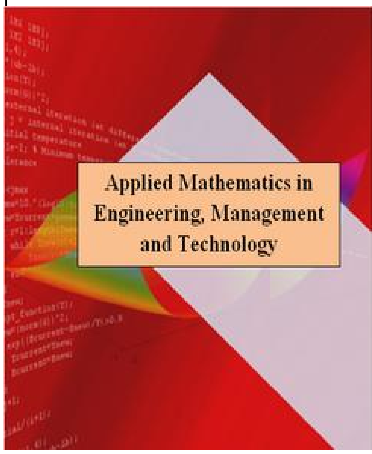
Applied mathematics in Engineering, Management and Technology 1 (3) 2013:210-216
www.amiemt-journal.com

A Penalty Line Search Method with Steering Rules

Maryam Dehghan Nayeri
Department of Mathematics, Faculty of mathematics, Sharif University Of Technology, Tehran, Iran



Abstract

In this paper an exact penalty line search is introduced for solving constrained nonlinear programing. By this method it is possible to solve some problems with combining penalty method and SQP method, and in any iteration for producing new stepone can uses linear search and solves unconstrained optimization by exact penalty function. In this method a linear programming subproblem with trust area constraint uses to update penalty parameter. This method ran and tested over various problems. Theresults show effectiveness of the method.

Keywords. Steering rules, SQP method, Optimizing with nonlinear constraint

1. Introduction

The exact penalty methods have introduced effective techniques for solving nonlinear programming problems. By these methods, many difficulties, such as incompatibility of linear constraints, have been overcome [9], and are successful in certain categories of problems in which linearized constraints are not satisfied [2,4,7,11,13]. The penalty methods solve vast range of problems, however, they have some matters in choosing and updating proper penalty parameter. Various ways are introduced for updating new parameter which they converge to improper points in most cases [12]. This is one of ineffective of this method.
Nearly a strategy is introduced for updating penalty parameter which uses trust region methods [3,5]. In this way, penalty parameter has to be updated such that efficient progress to the optimizing has guaranteed and for this matter, in some especialcases we need to solve an associated subproblems. In some problems with many constraints, this method have been combined with impressive set methods and that made it effective.
The mentioned ways are from the kind of trust region. In this situation, at first, one of the linear search penalty methods for nonlinear programming will be described. The linear search methods, however, as the trust region methods control step lengths. They can also produce very large and without result search direction in neighborhood of points which do not satisfy standard regular conditions. This situation may fails the linear search method, while these difficulties have been overcome by the trust region methods. Thus in this paper a method has introduced such that one can converges toward the solution continuously. The method, also, as like the trust region methods, possess convergence attributes. An attractive ability in the new algorithm is that for various problems has analogue operation.

2. Classical penalty function algorithm

Suppose the linear programming problem bellow:

$$\min f(x) \tag{1}$$

$$s.t. \quad g(x) \le 0 \tag{2}$$

$$h(x) = 0 \tag{3}$$

Function f is the target function and h and g are equality (p numbers) and inequality constrains (m numbers) respectively. The methods in which penalty functions are used convert a constrained problem to one or a sequence of unconstrained problems [4,15]. The constrains would input to the target function with a penalty parameter such that violation in them would results to penalty. Above problems would relates to the below with a classes of penalty functions:

$$\Phi(x,p) = f(x) + p\,Q\left(\left\| g(x)_+, h(x) \right\|\right) \qquad (4)$$

in which p is a real number and $\| \, . \, \|$ is any vector norm in R^{m+k} and $\left(g(x)_+\right)_j = \max\{0, g_j(x)\}(j=1,\dots,m)$ and Q is a function of non-negative real number R_+ into itself with specification:

$$Q(0) = 0, \qquad (5)$$

$$Q(e) > 0 \quad \text{for } e > 0, \qquad (6)$$

$$\infty > Q'(0^+) = \lim_{e \to 0^+} \frac{Q(e) - Q(0)}{e} > 0. \qquad (7)$$

Obviously (7) is equivalent to $Q(0)$ has a positive and finite value when Q is differentiable at zero. One of penalty function classes, is classical exact penalty function which is written as below:

$$\Phi_1(x,p) = f(x) + p\left(\sum_{j=1}^{m} \left\| g_j(x)_+ \right\| + \sum_{j=1}^{k} \left| h_j(x) \right|\right) \qquad (8)$$

Above equation would achieve by letting $Q(e) = e$ in equation (4) and using norm 1.

2.1. A sample choosing for penalty parameter

In the $\mathbf{l}_1$ classical penalty methods, choosing and updating penalty parameter p_k is very difficult. If the initial choosing, i.e. p_0 be small, it is possible that the penalty function algorithm which is introduced at the end of this section, found no x_{k+1} which runs the step 2 and then fall in to a trap at the step 1 and we could not find any proper penalty parameter. In addition, it is possible that in the initial iteration we stay far from the solution what in these situation a new step lays over the prior, because the minimization of $\Phi_{p_1}(x)$ should be terminated. On the other hand, if p_k be very large, it is possible that minimizing get too difficulties, for verifiability of constraints may result in a small step which this cause to so many iterations.

3. Algorithm; Linear search penalty method with steering rules

3.1. A linear search penalty method

The exact penalty method solves the nonlinear programming as below:

$$\begin{aligned} &\min f(x) \\ &s.t. \quad h_i(x) = 0 \qquad i \in e \\ &\qquad g_i(x) \geq 0 \qquad i \in I \end{aligned} \qquad (9)$$

by minimizing the exact penalty function

$$j_p(x) = f(x) + p u(x) \qquad (10)$$

where $u(x)$ is verification for violation of constraints and p is called penalty parameter.

3.2. Algorithm (3-1); Linear search method with steering rules

Data $x_0 \in R$ parameters:

$p_1 > 0, r > 0, e_1 \in (0,1), e_2 \in (0,e_1), t \in (0,1), h, h_1, h_2 \in (0,1), 0 < \Delta_{\min} \leq \Delta \leq \Delta_{\max}$ (for example, $p_1 = 1, r = 10, e_1 = 0.1, e_2 = 0.1, t = 0.5, h_1 = 0.25, h_2 = 0.75$)

Initial input $w_1 = I$
Let $k = 1$.

Step I: Solve the subproblem below:

$$\min_d q_k^p(d). \qquad (11)$$

Let $d_k(p_k)$ as a piece of its solution $p = p_k$. If $\|d_k(p_k)\|_\infty < 10^{-6}$ and $\|v(x_k)\| \le 10^{-6}$, stop and print x_k is KKT point.
Step II: If $m_k(d_k(p_k)) = 0$ then let $p_+ = p_k$ and go to step VI.
Step III: Solve the subproblem

$$\min m_k(d)$$
$$s.t. \quad \|d\|_\infty \le \Delta_k \qquad (12)$$

where Δ_k is given and then name a piece of its solution as $d_k^{LP}(p_k)$. If

$$m_k(0) - m_k(d_k^{Lp}) \le 10^{-16} \quad \text{and} \quad m_k(0) > 0 \; (13)$$

then stop and name x_k an infeasible stationary point for the penalty function of (10).
Step IV (updating penalty parametere):
IV-a: If $m_k(d_k^{LP}) = 0$, then find penalty parameter $p_+ > p_k + r$ in according to the vector $d(p_+)$ such that solves the problem (11) in such a way that:

$$m_k(d(p_+)) = 0. \qquad (14)$$

IV-b: Otherwise let $p_+ = p_k$. If the inequality

$$m_k(0) - m_k(d(p_+)) \ge e_1[m_k(0) - m_k(d_{LP}^k)] \quad (15)$$

does not hold, then evaluate $p_+ > p_k + r$ such that the vector $d(p_+)$ solves problem (14) and condition (15) holds.
Step V: and if the inequality

$$q_k^{p^+}(0) - q_k^{p^+}(d(p_+)) \ge e_2 p_+[m_k(0) - m_k(d_k^{LP})] \qquad (16)$$

does not hold, then evaluate $p_+ = p_k + r$ such that the vector $d(p_+)$ according to that holds (16).

Step VI (evaluating step length): Suppose the sequence $\{1, t, t^2, \ldots\}$ and calculate $0 < a_k < 1$ such that the relation below holds:

$$\Phi_{p_+}(x_k) - \Phi_{p_+}(x_k + a_k d_k) \ge h a_k [q_{p_+}^k(0) - q_{p_+}^k(d_k)] \qquad (17)$$

Step VII: Let $\Delta_{k+1} \in [\Delta_{\min}, \Delta_{\max}]$.
Step VIII: Let $p_{k+1} = p_+$ and $x_{k+1} = x_k + a_k d_k$.
The results of general convergence are presented in the next section, based on the linear search methods together with affection of penalty approximation, because the constraint of trust region in problem (12) only roles in choosing penalty parameter, and have an absolutely indirect role in the next step. Remember that for the linearized model below:

$$q_k^p = f(x_k) + \nabla^T f(x_t) d + \frac{1}{2} d^T w_k d + p\, m_k(d),$$

Hessian matrix w_k is positive definite and q_k^p is convex, so d_k^p abtained from the solution of problem (14) is unique.

4. Convergence analyzing

In this section the general convergence of Algorithm 3-1 will be describe. First of all suppose the assumptions below about iterative sequence $\{x_k\}$ and matrices w_k produced by Algorithm 3-1.

Assumptions:

A-I: Functions $h_i(i \in e)$, $g_i(i \in I)$ and f over a bounded convex set including the sequence x_k are differentiable with continuous derivatives.

A-II: Matrices w_k over a bounded convex set are positive definite uniformly and also there exists values $0 < m_{\min} < m_{\max}$ such that

$$m_{\min} \|p\|^2 \leq p^T w_k p \leq m_{\max} \|p\|^2,$$

for $p \in R^n$.

Directional derivative of function f in point a and in direction of p is shown by $Df(x,p)$. The point x is called stationary of the penalty function whenever for any direction p in R^n,

$$Dj_p(x,p) \geq 0.$$

In addition, point x is called an infeasible stationary point for problem (9) if $u(x) > 0$ then $Du(x,p) \geq 0$, for all $p \in R^n$. if problem (9) has an infeasible stationary point, then it is called local infeasible.

4.1 Algorithm well-definition

Lemma 1. Suppose x_k neither is a KKT point for nonlinear programming of (3-1) nor an stationary point of $u(x)$. Then we have

1. In Step IV of Algorithm 3-1, always there exists value p_+ according to vector $d_k(p_+)$ such that if $m_k(d^{LP}) = 0$, then condition (14) holds, and if $m_k(d^{LP}) > 0$ then condition (15) holds.
2. In Step V of Algorithm 3-1, always there exists value p_+ according to the vector $d_k(p_+)$ such that condition (16) holds.
3. In Step VI of Algorithm 3-1, d_k is a decreasing direction of $j_{p_+}(x)$, and so there exists a k such that condition (17) holds.

Theorem 2. Suppose that Algorithm 3-1 produce an infinite sequence of iterations in which assumption A-I and A-II hold. Let $\{p_k\}$ is bounded such that for all k sufficiently large, $p_k = \bar{p}$. Then any accumulation point x_* of $\{x_k\}$ is a stationary point of penalty function $j_p(x)$.

5. Calculative remarks for running

One of the most effective methods for solving constrained optimization problems is producing step by solving a second order subproblem. In spite of linear constrained iterative methods, when are effective just as most of the constrains are linear, the Second Order Periodic (SQP) method, is effective for solving nonlinear problems. The SQP method is described in this section.

Suppose the problem below:

$$\min f(x)$$
$$s.t. \quad c_i(x) = 0 \qquad i = 1,2,\ldots,m \qquad (18)$$

where $f: R^n \to R$, $c_i: R \to R$ ($i = 1,\ldots,m$) are functions with continuous derivatives up to order two. Indeed, problems in which there are only equality constrains, are not so important, but they help to understand the problems with general constrains.

Basic idea of SQP, is modeling (18) in iteration x_k with a second order subproblem programming as below

$$\min \nabla^T f(x_k) d + \frac{1}{2} d^T B_k d$$

$$s.t. \quad \nabla^T c_i(x_k)\,d + c_i(x_k) = 0 \qquad i = 1,2,\ldots,m, \tag{19}$$

and using the minimum of subproblem for obtaining the new point x_{k+1}.
Maybe the easiest definition of SQP, which is introduced below, is an application of Newton methods for holding optimal conditions KKT.
Assumptions:
B-I: Matrix $\nabla^T c_k$, has perfect row order.
B-II: Matrix B_k is positive definite in tangent space constrains, i.e.

$$d^T B_k d > 0, \qquad \forall d \in \{z : \nabla^T c(x_k) z = 0\}, \quad d \neq 0.$$

The assumption B-I is the linearly independence of gradients of constrains which in follow assumed to be hold. The assumption B-II guarantee d_k satisfies in second order sufficient conditions of KKT.
Now, by subtracting $\nabla c_k m_k$ from both side of Lagrange-Newton system, we can solve system below for $d = d_k$ and $m = m_{k+1}$ equivalently:

$$\begin{pmatrix} B_k & \nabla c_k \\ (\nabla c_k)^T & 0 \end{pmatrix} \begin{pmatrix} d \\ -m \end{pmatrix} = -\begin{pmatrix} \nabla f_k \\ c_k \end{pmatrix} \tag{20}$$

and let $x_{k+1} = x_k + d_k$.
Note that instead of systems (20) one can solves a sequence of second order functions minimizing subproblems as below:

$$\min \nabla^T f(x_k)\,d + \frac{1}{2} d^T B_k\, d$$
$$s.t. \quad \nabla^T c(x_k)\,d + c(x_k) = 0. \tag{21}$$

If assumptions B-I and B-II hold, then this problem has a unique solution (d_k, m_k) which satisfies in system (20). Thus the SQP in a simple form can be presented as below form:
Algorithm (5-1): SQP Algorithm
(0) Choose values m_1, x_1 and let $k = 1$.
(1) Solve the problem (19) for d_k and let m_{k+1} as the linear constrains Lagrange coefficient vector.
(2) Let $x_{k+1} = x_k + d_k$.
(3) If the convergence test holds, then stop (with approximately solution (x_{k+1}, m_{k+1})).
(4) Let $k = k+1$ and go to the step (1).

5.1. Consequences

Algorithm (3-1) is tested over various complex problems which in some of them incompatible linearized constrains exists or the MFCQ conditions does not hold or the problem is infeasible. The obtained results show that Algorithm, has excellent operation. Problems are selected from [8] and [9] and the results are shown in the table 5-1 to 5-3. Algorithm 3-1 is comprised with methods of kind interior point with linear search, SQP with filter, SQP with impressive set methods, penalty function methods, classical SQP and etc. From the obtained results in tables 5-2 and 5-3 it is obvious that the operation of Algorithm is much better than above methods.

Table 1. Numerical results of Algorithm 3-1

Name	Iter	p_k	$KKT(x)$	$feas(x)$	$f(x)$	$\Phi_p(x)$
Hs5	5	10	0	0	3.6543	3.6543
Hs15	10	1000	0.001	0	1	1
Hs18	21	100000	6.14E-7	12	27.49797	1200028
Hs23	4	1	4.38E-8	2.000016	1.75E-6	2.000018
Hs30	10	1	7.91E-7	0	1.00001-	1.00001
Hs36	2	1000	2.84E-9	0	-3300	-3300
Hs39	5	100	0	0	1	1
Hs40	10	1	0	0	-1	7
Hs42	2	1	0	0	0	0
Hs43	10	100	0	0	2	10

Example 1. The problem

$$\begin{aligned} &\min x_1 \\ s.t.\quad & x_1^2+1-x_2=0 \\ & x_1-1-x_3=0 \\ & x_2\geq 0,\quad x_3\geq 0 \end{aligned}$$

introduced by Wächter and Biegler in [14]. In this problem constraints are incompatible linearized. It has been shown that there exists a class of linear search interior point that may converge to a infeasible stationary point. In this paper, by using Algorithm 3-1 we converge from initial point $(-3,1,1)$ to the stationary point $x^*=(1,2,0)$. In the table 5-2 are shown the information resulted after running program.

Table 2. Numerical results of Example 5-1

it	p_k	x_1	x_2	$KKT(x)$	$feas(x)$	$f(x)$	$j_{p_+}(x)$
0	1	-3	1	.00E+006	1.40E+01	-3.00E+00	1.10E+01
1	1	-1.5405	1.2432	.54E+003	4.67E+00	-1.54E+00	3.13E+00
2	1	-0.9428	1.5316	1.94E+00	2.30E+00	-9.43E-01	1.36E+00
3	1	0.8115	1.6414	1.81E+00	1.83E+00	-8.12E-01	1.02E+00
4	1	-0.8043	1.6468	1.80E+00	1.80E+00	-8.04E-01	1.00E+00
5	1	-0.2924	0.8234	6.10E+00	1.55E+00	-2.92E-01	1.53E+01
6	10	1	2	8.25E+00	1.19E+00	3.54E-01	1.23E+01
7	10	0.40109	2	2.13E-17	2.69E-01	1.00E+00	1

Example 2. Suppose problem bellow

$$\begin{aligned} &\min (x_2-1)^2 \\ s.t.\quad & x_1^2=0 \\ & x_1^3=0 \end{aligned}$$

This problem has been investigated by Chen and Goldfarb [8]. In this problem the conditions MFCQ are broken in solution $x^*=(0,1)^T$ and the linearized constraints are also incompatible in any infeasible point. Fletcher [10] remarks that he hasconverged by initiating from infeasible point $(0,1)^T$ to $(0,0)^T$ that is not the solution of problem. In addition, in SQP it runs by filter and proved that exactly converges to $(0,0)^T$. After running by Algorithm 3-1 illustrated that this Algorithmdoes not show this treat. By this method, the iteration sequence method moves to the solution from the initial steps and does not converge to origin. In table 5-3 it has been shown that the linearized constraints never satisfied. But Algorithm obtains the penalty value $p=1$ which produce sufficient progress.

Table 3. Numerical results of Example 5-2

it	p_k	x_1	x_2	$KKT(x)$	$feas(x_k)$	$f(x)$	$j_{p_+}(x)$
0	1	1	0	2.00E+00	2.00E+00	1.00E+00	3.00E+00
1	1	0.6667	0.6667	6.67E-01	7.41E-01	1.11E-01	8.52E-01
2	1	0.4444	0.8736	3.95E-01	2.85E-01	1.60E-02	3.01E-01
3	1	0.2222	0.952	2.59E-01	6.04E-02	2.30E-03	6.27E-02
4	1	0.1111	1	6.48E-02	1.37E-02	0.00E+00	1.37E-02
5	1	0.0556	1	1.62E-02	3.26E-03	4.93E-32	3.26E-03
6	1	0.0278	1	4.05E-03	7.93E-04	4.93E-32	7.93E-04
7	1	0.0139	1	1.01E-03	1.96E-04	1.23E-32	1.96E-04
8	1	0.0069	1	2.53E-04	4.86E-05	0.00E+00	4.86E-05
9	1	0.0035	1	6.33E-05	1.21E-05	4.93E-32	1.21E-05
10	1	0.0017	1	1.58E-05	3.02E-06	4.93E-32	3.02E-06
11	1	0.0009	1	3.96E-06	7.54E-07	1.23E-32	7.54E-07
12	1	0.0009	1	8.48E-07	7.54E-07	1.23E-32	7.54E-07

References

[1] R. H. Byrd, N. I. M. Gould, J. Nocedal, R. A. Waltz, An algorithm for nonlinear optimization using linear programming and equality constrained subproblems, Mathematical Programming, 23(4):(2004) 27-48.
[2] R. H. Byrd, N. I. M. Gould, J. Nocedal, R. A. Waltz, On the convergence of uccessivelinearquadratic programming algorithms. SIAM Journal on Optimization,45(5): (2006) 471-489.
[3] R. H. Byrd, J. Nocedal, R. A.Waltz, Steering exact penalty methods, Optimization Methods and Software,16(4): (2008) 266-290.
[4] S. P. Han , O. L. Mangasarian, Exact penalty functions in nonlinear programming, Mathematical Programming, 45(8): (1979) 251-269.
[5] J. Nocedal , S. J. Wright, Numerical Optimization, Springer Series in OperationsResearch, Springer, 1999.
[6] J.L. Zhang, X.S. Zhang, A robust SQP method for optimization with inequality constraints, J.Comput, 78(3): (2003) 247-256.
[7] D.GLuenberg, linear and Nonlinear Programming, 2nd edition, Addision-Wesely Publishing Co.,1984.
[8] L. Chen , D. Goldfarb, Interior-point ‘l2 penalty methods for nonlinear programming with strong global convergence properties, Mathematical Programming, 100(5):(2005) 1-36.
[9] O. L. Mangasarian , S. Fromovitz, The Fritz- John necessary optimality conditions in the presence of equality and inequality constraints, Journal of Mathematical Analysisand Applications,106(6): (1967) 37-47.
[10] R. Fletcher, S. Leyher, and Ph. L. Toint, On the global convergence of a fiter-SQP algorithm, SIAM Journal on Optimization, 34(8): (2002)44-59.
[11] E. O. Omojokun, Trust region algorithms for optimization with nonlinear equality and inequality constraints, P.h.D thesis, University of Colorado, Boulder,Colorado, USA, 1989.
[12] X.M. Hu, D. Ralph, Convergence of a penalty method for mathematical programming with complementarity constraints, Journal of Optimization Theory and Applications, 50(7): (2004) 365-390.
[13] R. H. Byrd, Robust trust region methods for constrained optimization, Third SIAM Conference on Optimization, Houston, Texas, May,74(5): (1987) 365-56
[14] Wächter, A. and L. T. Biegler, Failure of global convergence for a class of interior point methods for nonlinear programming, Mathematical Programming, 88(3):565574, 2000.
[15] H.Y. Benson, A. Sen, D.F. Shanno, R.J. Vanderbei, Nterior-point algorithms,penalty methods and equilibrium problems, Computational Optimization and Applications, 38(23):(2006)155-182.
[16] W.Rudin, Principles of Mathematical Analysis, 1976.