ACSIJ Advances in Computer Science: an International Journal, Vol. 5, Issue 3, No.21 , May 2016
ISSN : 2322-5157
www.ACSIJ.org

An SQP Algorithm for Recourse-based Stochastic Nonlinear Programming

Cunzhe Liu[1] and Xinshun Ma[2]

[1,2] Department of Mathematics and Physics, North China Electric Power University, Baoding , 071003, Hebei Province, China
[1]liuzhemath@163.com,[2]maxs@sina.com

Abstract

The stochastic nonlinear programming problem with completed recourse and nonlinear constraints is studied in this paper. We present a sequential quadratic programming method for solving the problem based on the certainty extended nonlinear model. This algorithm is obtained by combing the active set method and filter method. The convergence of the method is established under some standard assumptions. Moreover, a practical design is presented and numerical results are provided.

Keywords: *stochastic programming; nonlinear constraints; SQP*

1. Introduction

Stochastic programming is the basic method to solve the optimization problem in uncertain environment. The two-stage stochastic programming problem is based on random variables to be observed before and after the specific values, and meanwhile the decision variables and decision-making process are divided into two stages of mathematical programming to make decisions. Stochastic programming, especially the general theory and method of the recourse-based stochastic linear programming, has been studied in the literature [1-3]. As a direct extension of the linear case, the theory of stochastic convex programming has been studied rapidly [4-14] since the groundbreaking work of Rockafellar and Wet [12].

So far, two basic method frameworks have been developed to solve the recourse-based stochastic convex programming (including stochastic linear programming).The first method is based on the finite sample space, and the stochastic programming is transformed into a deterministic extensive mathematical programming problem. After that, we use the deterministic mathematical programming method to solve the problems, and it's suitable to solve the problems of small and medium scale. The second method based on Benders decomposition and the development of L-shaped algorithm, the large-scale problem is transformed into a small scale problem to be solved iteratively and it's suitable for solving large-scale problems in principle. However, in view of the general stochastic nonlinear problems, because of the complexity of the model structure and design algorithm, the effective solving algorithm is relatively less. Recently, the literature [16] gave the second solving algorithm framework for general nonlinear programming problems. In the decomposition algorithm of this paper, the Benders decomposition technique is essentially used to solve the sequential quadratic programming subproblem, as a result, the algorithm is limited to a conceptual framework and the computational efficiency in practical also need to be further tested.

With the new theory and new algorithm of nonlinear programming have been proposed, we can use the latest achievements of deterministic nonlinear programming algorithm directly. We use the first method to solve the stochastic nonlinear programming, and it is still a more effective and feasible method to solve the problem of small and medium size. In this paper, we present a sequential quadratic programming (SQP) method to solve the recourse-based stochastic nonlinear programming problems which are based on the equivalent nonlinear extensive model. The search direction is obtained by the ε-active set method which is used to solve the quadratic programming subproblems, and the step length is obtained by the filter method to avoid the difficulty of choosing penalty factor in standard method. The convergence of the method is established under some standard assumptions, and the numerical results show that the algorithm is effective.

2. Stochastic Nonlinear Programming and Its Equivalent Problem

Let us suppose that the triplet (Ω,A,μ) as a probability space, $\zeta=\zeta(\omega)$ is a random vector defined in the space, we consider recourse-based stochastic nonlinear programming

$$\begin{aligned} & min \quad \left\{ f\left(x\right)+E_{\zeta}\left[Q\left(x,\zeta(\omega)\right)\right]\right\} \\ & s.t. \quad \varphi(x)\geq 0 \end{aligned} \tag{1}$$

with

$$Q\left(x,\zeta\left(\omega\right)\right)=min\left\{g(y)\middle|a(x,\omega)+b(y,\omega)\geq 0\right\}. \tag{2}$$

Where, $x\in R^{n_1}$ and $y\in R^{n_2}$ are the first stage decision variable and the second stage decision variable respectively, for a given realization $\omega\in\Omega$, $a:R^{m_1\times n_1}\to R^{m_1}$

and $b: R^{m_2\times n_2} \to R^{m_2}$, $f(x)$ and $g(y)$ are real nonlinear function, $\varphi(x)$ is real linear or nonlinear constraint. $E_\zeta(\cdot)$ denotes the expected value operator about random vector ζ , $Q(x,\zeta(\omega))$ is the second-stage value function, and $Q(x)=E_\zeta\left[Q(x,\zeta(\omega))\right]$ is called the recourse function.

We assume that the random vector ζ has finite support, let $i=1,\cdots,N$ denote its possible realizations, and p_i is their probabilities. We may write the recourse-based stochastic programming (1)-(2) as follows:

$$\begin{cases} min \quad f(x)+\sum_{i=1}^{N} p_i g(y_i) \\ s.t. \quad \varphi(x)\ge 0, \\ \qquad a_i(x,\omega)+b_i(y_i,\omega)\ge 0, i=1,\cdots,N. \end{cases} \tag{3}$$

First, we give the following notations:

$$z=(x^T,y_1^T,\cdots,y_N^T)^T, \tag{4}$$

$$\nabla F(z)=\begin{pmatrix} \nabla f(x) \\ p_1\nabla g(y_1) \\ \vdots \\ p_N\nabla g(y_N) \end{pmatrix}, \tag{5}$$

$$u(z)=\begin{pmatrix} \varphi(x) \\ a_1(x)+b_1(y_1) \\ \vdots \\ a_N(x)+b_N(y_N) \end{pmatrix}. \tag{6}$$

As a result, the recourse-based stochastic nonlinear programming problem can be formulated as the following equivalent extensive form:

$$\begin{cases} \min\limits_{z\in R^{n_1+n_2*N}} \quad F(z) \\ s.t. \quad u_i(z)\ge 0, i=1,\cdots .N+1. \end{cases} \tag{7}$$

3. A filter Active Set SQP Algorithm

For stochastic programming model (7), we give the sequential quadratic programming methods to solve it, and it needs to solve a QP subproblem at kth iteration:

$$\begin{cases} \min\limits_{d\in R^{n_1+n_2*N}} \quad g_k^T d+\frac{1}{2}d^T B_k d \\ s.t. \qquad \nabla u_i(z_k)d+u_i(z_k)\ge 0, i=1,\cdots,N+1. \end{cases} \tag{8}$$

Where $g_k=g(z_k)=\nabla F(z_k)$, $B_k\in R^{(n_1+n_2*N)\times(n_1+n_2*N)}$ is an approximate Hessian of the Lagrangian function of problem (7).

In order to solve (8), we define the sets

$$I(z,\varepsilon)=\{\, i \mid u_i(z)\le\varepsilon \,\}, i\in\{1,\cdots,N+1\}, \tag{9}$$

and it needs to solve a subproblem at kth iteration:

$$QP_k \begin{cases} \min\limits_{d\in R^{n_1+n_2*N}} \quad g_k^T d+\frac{1}{2}d^T B_k d \\ s.t. \qquad \nabla u_i(z_k)d+u_i(z_k)\ge 0, i\in I(z_k,\varepsilon_k). \end{cases} \tag{10}$$

We assume that d_k is an optimal solution of (10), and it's also the search direction of the current iteration step.

The classical SQP algorithm uses the penalty function [17] as the value function to determine the linear search step length, but in practice, the selection of penalty factor is relatively difficult. In this paper, the filter method based on the literature [20] is used to determine the step length to replace the classical method.

For each iteration point z_k , we define the following constraint violation function:

$$G(z_k)=\sum_{i=1}^{N+1} min\{u_i(z),0\}, \tag{11}$$

Together with the function value $F(z_k)$, we can get a two-dimensional array

$$(F_k,G_k)=\left(F(z_k),G(z_k)\right). \tag{12}$$

A two-dimensional array (F_k,G_k) is said to dominate (F_l,G_l) if and only if doth $F_k\le F_l$ and $G_k\le G_l$. A filter is several pairs and no pair dominates any other. If a pair is not dominated by any other in the filter, the array is said to be accepted for inclusion in the filter. Obviously, the acceptance condition of the filter is not sufficient to guarantee the convergence of the algorithm, because it allows the later point gathered at the neighborhood of some point $(F(z_j),G(z_j))$ of the current filter, and at the same time $G(z_j)>0$, $G(c(z_k))\to 0(k\to\infty)$ can not be guaranteed. So we can not find a KKT point of the nonlinear programming problem. To avoid this, we need to improve the acceptable conditions of the filter, and do an envelope below the filter that prevents an arbitrary point close to the filter from being accepted.

We use P_k to denote the set of the entire iteration index $j(j<k)$, where (F_j,G_j) belongs to an array of the current filter. The iterative point z_k is said to be acceptable to the filter, if

$$G_k\le\alpha G_j \quad \text{or} \quad F_k\le F_j-\beta G_j,$$

holds for all $j\in P_k$. Here α,β are positive constants, and $0<\beta<\alpha<1$.

Algorithm

Step1 Given $z_1\in R^{n_1+n_2*N}$, B_1 is a positive definite matrix, $\varepsilon\ge 0$ k:=1.

Step2 Solve QP_k . Suppose that d_k is the solution and the search direction. If $\|d_k\|\le\varepsilon$, then stop, and output z_k as an approximate minimum point.

Step3 Let $l=1$, $\gamma_{k,l}=1$.

Step4 Set $\bar{z}_k=z_k+\gamma_{k,l}d_k$, if $\bar{z}_k$ is said to be acceptable for the filter, then let $\gamma_k=\gamma_{k,l}$, $z_{k+1}=\bar{z}_k$, and go to Step6.

Step5 Let $\gamma_{k,l+1}=\gamma_{k,l}/2$, $l=l+1$, and go to Step4.

Step6 If $g_k^T d_k > -\frac{1}{2} d_k^T B_k d_k$, then the point z_{k+1} is accepted by the filter.

Step7 Update B_k to B_{k+1} , $k := k+1$, and go to Step2.

Note: In the algorithm above, we compute B_k by some quasi-Newton [21-22] formula.

In order to prove the convergence of the algorithm, we need the following assumptions:

(A1) $F(z)$ is convex and twice continuously differentiable function.

(A2) B_k is symmetric positive definite and there exists two positive constants m and M such that

$$m\|d\|_2^2 \le d^T B_k d \le M\|d\|_2^2, \tag{13}$$

holds for all $k \ge 1$ and all $d \in R^{n_1+n_2*N}$.

(A3)All iterative points $\{z_k\}$ generated by the algorithm are located in a bounded closed convex set Ω .

(A4)For all $k \in K$, there is $z_k \to z^*$ (z^* is not necessarily a KKT point), and there exists $v \in R^{n_1+n_2*N}$ such that

$$C_i^T v > 0 \quad i \in I'(x^*,\varepsilon), \tag{14}$$

where $I'(x^*,\varepsilon) = \{i \mid i \in I(z_k,\varepsilon_k), k \in K\}$, $\varepsilon_k \ge 0$.

In order to prove the convergence of the algorithm, we give the following lemma.

Lemma 1 Let us assume that an infinite sequence (F_k, G_k) is acceptable by the filter, where $G_k > 0$, and $\{F_k\}$ is bounded below, then there is $\lim_{k\to\infty} G_k = 0$.

Proof: See the literature [20].

Theorem 2 Under the assumptions (A1)-(A4), the algorithm generates a sequence such that at least exiting an accumulation is the KKT point of the problem.

Proof: From the assumption (A3), there is an accumulation point z^* such that

$$\lim_{k\in K, k\to\infty} z_k = z^*,$$

where K is an infinite index set. From the process of the algorithm, we know that there are infinite or finite points into the filter. We will prove the two cases respectively in the following.

Case 1: there are infinite many points into the filter

There are infinite many points into the filter, then $\lim_{k\to\infty} G_k = 0$.

If $G_k = 0$, z_k is a feasible point. From the algorithm, we can know

$$g_k^T d_k + \frac{1}{2} d_k^T B_k d_k \le 0,$$

then z_k can not be accept by the filter, so $G_k > 0$. From lemma 1 and F_k is bounded below, we know $\lim_{k\to\infty} G_k = 0$, so z^* is a feasible point. We prove that z^* must be a KKT point in the following.

If z^* is not a KKT point, we suppose

$$K_1 = \left\{ k \in K \mid g_k^T d_k > -\frac{1}{2} d_k^T B_k d_k \right\}.$$

Then for $\forall k \in K_1$, there must exist $\varepsilon > 0$ such that $\|d_k\| > \varepsilon$ (otherwise, there exists $K_2 \subset K_1$ such that $\lim_{k\in K_2, k\to\infty} \|d_k\| = 0$, and they are contradictory).

From the KKT point of the subproblem QP_k , we know

$$\begin{aligned} g_k^T d_k &= d_k^T \nabla u(z_k) \upsilon_k - d_k^T B_k d_k \\ &= -\upsilon_k^T u(z_k) - d_k^T B_k d_k \end{aligned}.$$

If $u_i(z_k) \ge 0$, we have

$$g_k^T d_k \le -d_k^T B_k d_k \le -\frac{1}{2} d_k^T B_k d_k . \tag{15}$$

If $u_i(z_k) < 0$, we have

$$g_k^T d_k \le \upsilon_k^T G_k - d_k^T B_k d_k . \tag{16}$$

From $\lim_{k\to\infty} G_k = 0$, we know $\exists k_0$, if $k \ge k_0$, then

$$G_k \le \frac{1}{2M} d_k^T B_k d_k , \tag{17}$$

$$g_k^T d_k \le \upsilon_k^T - d_k^T B_k d_k \le -\frac{1}{2} d_k^T B_k d_k , \tag{18}$$

So for $\forall k \in K_1$, we have

$$g_k^T d_k \le -\frac{1}{2} d_k^T B_k d_k ,$$

Which contradicts the definition of K_1 , so z^* is a KKT point.

Case 2: there are finite many points into the filter

From the assumption, we know F_k is monotonous and bounded below. We have

$$\lim_{k\in K_1, k\to\infty} G_k = 0 ,$$

So z^* is a feasible point. From the algorithm, we know $\exists k_0$, if $k \ge k_0$, then

$$g_k^T d_k \le -\frac{1}{2} d_k^T B_k d_k ,$$

K_1 is a finite set, and there must exist $\eta > 0$ such that

$$\begin{aligned} F_k - F_{k+1} &\ge -\eta g_k^T d_k + o(\eta) \\ &\ge \frac{\eta}{2} d_k^T B_k d_k \ge \eta G_k = \eta o\left(\|d_k\|^2\right) \end{aligned}, \tag{19}$$

Let us sum the both sides of the above equation at the same time, and we have

$$+\infty > \sum_{k=k_0}^{\infty} F_k - F_{k+1} \ge \sum_{k=k_0}^{\infty} \eta o\left(\|d_k\|^2\right), \tag{20}$$

$$\sum_{k=k_0}^{\infty} \eta o\left(\|d_k\|^2\right) \le +\infty ,$$

So

$$\lim_{k\to\infty} \|d_k\| = 0 ,$$

From the above equation, we know z^* is a KKT point.

4. Numerical Experiment

There is no standard stochastic nonlinear programming test problem set available. In literature [23], Chen and Womersley have developed test problems for stochastic quadratic problems. We can also use this idea to generate the problem.

In this section, we consider the following stochastic nonlinear programming problem. The first stage objective function is from [24] and the second stage objective is a convex function with twice continuously differentiable. The constraints are from the literature [25-26], which guarantee the problem with completed recourse.

$$\begin{cases} min & (x_1-1)^a+(x_1-x_2)^b+ \\ & E_\xi\, min\left\{ce^{-y_1-y_2}+2(y_1+y_2)^2+2y_1+6y_2\right\} \\ s.t. & 3x_1+2x_2+me^{x_1}\le 15, y_1\le x_1, \\ & x_1+2x_2+ne^{x_2}\le 5, y_2\le x_2, \\ & x_1+x_2\ge 0, y_1\le \xi_1, y_2\le \xi_2, \\ & x_1,x_2,y_1,y_2\ge 0. \end{cases} \tag{21}$$

Where $\xi=(\xi_1,\xi_2)^T$, ξ_1 take 4, 5, 6 at the probability of 1/3 respectively, ξ_2 take 1, 3, 5 at the probability of 1/3 respectively, ξ_1 and ξ_2 are independent of each other. So ξ take each element of the set $\left\{(k_1,k_2)^T \mid k_1=4,5,6; k_2=1,3,5\right\}$ at the probability of 1/9.

We use the algorithm presented in this paper to solve the problem (21), so we have

$$f(x)=(x_1-1)^a+(x_1-x_2)^b,$$
$$g(y)=ce^{-y_1-y_2}+2(y_1+y_2)^2+2y_1+6y_2.$$

The algorithm runs in MATLAB7.1 programming environment. The initial iteration point $z_0=[0,0,0,0,0,0,0,0]'$, and let $a=2$, $b=3$, $c=1$, $d=1$, $\alpha=0.99$, $\beta=0.0001$. Table 1 gives the iterative output of the algorithm, in which the second (k=2) iteration results are as follows:

$$z_2=[0.9559,0.8609,\text{-}0.0000,\text{-}0.0000,\text{-}0.0000,\text{-}0.0000,\text{-}0.0000,\text{-}0.0000]',$$
$$g_2=[0.0085,\text{-}0.0118,0.0000,0.0000,0.0000,0.0000,0.0000,0.0000]',$$
$$d_2=[\text{-}0.0611,\text{-}0.0271,0.7778,0.7778,0.7778,0.7778,0.7778,0.7778]',$$
$$(F_2,G_2)=(1.0028,0.0431).$$

Table 1 Iterative process

Iterations	$(x_k)'$	Objective function
1	$(0.5000,0.0000)$	1.3750
2	$(0.9559,0.8609)$	1.0028
3	$(0.9644,0.8491)$	1.0028
4	$(0.9723,0.8473)$	1.0027
5	$(0.9715,0.8474)$	1.0027
6	$(0.9716,0.8474)$	1.0027

From table 1, we can know that the number of iterations is 6, the optimal solution is $(0.9716,0.8474)'$, and the optimal value is 1.0027.

If $a=b=1$, $c=m=n=0$, we use the algorithm presented in this paper to solve it, and the optimal solution is $(0,0)'$, and the optimal value is -1. In literature [3], this kind of problem is solved by the L-shaped method, and the result is identical. Experimental results show that the proposed algorithm in this paper is superior to the latter in both the number of the iterations and the computation time. Because of the global convergence of the nonlinear programming, the method of this paper is an efficient algorithm to solve the stochastic nonlinear programming problem of small and medium scale, which is more efficient than the classical L-shaped method.

5. Conclusion

We give an algorithm to solve the stochastic nonlinear programming problem with completed recourse and nonlinear constraints in this paper. The algorithm bases on SQP method and combines with the ε-active set and the filter method. The convergence of the algorithm is proved under certain conditions, and the numerical results show that the algorithm is effective.

References

[1] R.M. Van Slyke and R.J.-B. Wets, “L-shaped linear programs with applications to optimal control and stochastic programming”, SIAM Journal of Applied Mathematics, 1969, 17 (4): 638–663.

[2] P. Kall and S.W. Wallace, “Stochastic Programming”, John Wiley & Sons, New York, 1994.

[3] J.R. Birge and F. Louveaux, “Introduction to stochastic programming”, Berlin:Springer, 1997.

[4] R.T. Rockafellar and R.J-B. Wets, “A dual solution procedure for quadratic stochastic programs with simple recourse”, Berlin:Springer, 1983, 252-265.

[5] R.T. Rockafellar and R.J-B. Wets, "A Lagrangian finite-generation technique for solving linear-quadratic problems in stochastic programming", Math. Prog. Study, 1986, 28:63-93.

[6] R.T. Rockafellar and R.J-B. Wets, "Linear-quadratic problems with stochastic penalties: the finite generation algorithm", Berlin:Springer, 1987, 545-560.

[7] R.T. Rockafellar, "Linear-quadratic programming and optimal control", SIAM J. Contr. Optim, 1987, 25:781-814.

[8] R.T. Rockafellar and R.J-B. Wets, "Generalized linear-quadratic problems of deterministic and stochastic optimal control in discrete time", SIAM J. Contr. Optim, 1990, 28:810-822.

[9] X. Chen, L. Qi and S. Womersley, "Newton's method for quadratic stochastic programs with recourse", Journal of Computational and Applied Mathematics, 1995, 60:29-46.

[10] L. Qi and S. Womersley, "An SQP algorithm for extended linear-quadratic problems in stochastic programming", Annals of Operations Research, 1995, 56:251-285.

[11] Z. Wei, L. Qi and X. Chen, "An SQP-Type method and its application in stochastic programs", Journal of Optimization theory and applications, 2003, 205-228.

[12] U.V. Shanbhag, "Decomposition and sampling methods for stochastic equilibrium problems", phD thesis, Department of Management Science and Engineering (Operations Research), Stanford University, 2006.

[13] P. Li Yong and H. Huang Guo, "Interval-parameter Two-stage stochastic nonlinear programming for water resources management under uncertainty", Water Resour Manage, 2008, 22:681-698.

[14] Heinz W. Engl, "Existence of measurable optima in stochastic nonlinear programming and control", Appl. Math. Optim, 1979, 5:271-281.

[15] R.T. Rockafellar and R.J.-B. Wets, "Stochastic convex programming: Kuhn-Tucker conditions", Journal of Mathematical Economics, 1975,2: 349-370.

[16] A.A. Kulkarni and U.V. Shanbhag, "Recourse-based stochastic nonlinear programming properties and Benders-SQP algorithms", Computational Optimization and Applications , 2012,51: 77-123.

[17] D.Q. Mayne and E. Polak, "A superlinearly convergent algorithm for constrained optimization problem", Math. Prog. Study, 1982, 16:45-61.

[18] M.J.D. Powell and Y. Yuan, "A recursive quadratic programming algorithm that use differentiable exact penalty function", Math. Prog, 1986, 35:265-278.

[19] N. Maratos, "Exact penalty function algorithms for finite dimensional and control optimization problems", phD thesis, Imperial College Sci. Tech. University of London, 1978.

[20] R. Fletcher and S. Leyffer, "Nonlinear programming without a penalty function", Mathematical Programming, 2002, 91:239-269.

[21] G.C. Broyden, "The convergence of a class of double rand minimization algorithms : 2.the new algorithm", J. Inst. Math. Appl., 1970, 6:222-231.

[22] R. Fletcher, "A new approach to variable metric algorithms", Computer J., 1970, 13:317-322.

[23] X. Chen and S. Womersley, "Random test problems and parallel methods for quadratic programs and quadratic stochastic programs", Optim, Methods Softw. 2000,13:275-306.

[24] W. Hock and K. Schittkowski, "Test examples for nonlinear programming codes", Springer, New York, 1981.

[25] L. Qi, C.Ling, X. Tong and G. Zhou, "A smoothing projected Newton-type algorithm for semi-infinite programming", Comput Optim Appl, 2009, 42:1-30.

[26] Diwekar Urmila and David Amy, "Bonue algorithm for large scale stochastic nonlinear programming problems", Spring Berlin, 2015.

Cunzhe Liu (1990-), Master graduate student, Department of mathematics and Physics, North China Electric Power University, his research direction is the algorithm of stochastic optimization.

Xinshun Ma (1964-), Professor, Ph.D., Department of mathematics and Physics, North China Electric Power University, his current research interests concern the optimization theory and method, stochastic programming and electricity market.