

# The Reduced Gradient Method Application

Prof. Parvaneh Basaligheh

University of Tech and Management Malaysia

[pbasaligeh@utmcc.my](mailto:pbasaligeh@utmcc.my)

## *Article Info*

*Page Number: 128 – 132*

*Publication Issue:*

*Vol 71 No. 1 (2022)*

## *Article History*

*Article Received: 18 November 2021*

*Revised: 03 December 2021*

*Accepted: 19 December 2021*

*Publication: 28 January 2022*

## *Abstract*

This study proposes a decomposition method through which they can be a successful solution to multi-stage stochastic nonlinear programs. The proposed method entails the scenario analysis method. The proposed method also performs its role via search direction generation in such a way that sets of quadratic programming sub-issues are solved in a parallel way, especially when the size is significant lower, compared to the case involving original problems at the respective iterations. Relative to the dual multiplier derivation, which focuses on non-anticipativity constraints, the proposed system advocates for the introduction of generalized reduced gradient approaches.

**Keywords:** - Gradient Method, SNLP etc.

---

## **1 Introduction**

For optimization problems, one of the aspects representing or depicting a critical class entails stochastic nonlinear programming (SNLP) problems. These problems are ever-existing in the real-world scenarios. Also, most of the existing systems remain nonlinear inherently, pointing to the criticality of employing nonlinear frameworks towards representing the systems, as well as the optimization of such systems by using nonlinear programming approaches. Another critical attribute to acknowledge is the uncertainty concept. In most cases, details of the systems under implementation are not known with precision and accuracy. Rather, variables and parameters are discerned relative to issues such as the ranges in which they occur or represent, as well as states of probability distributions. Given such scenarios, it is important to note that optimization can be realized via the incorporation of stochastic programming frameworks. Indeed, the emergence of the problems of multi-stage stochastic programming is felt and well-documented in practical scenarios, including portfolio selections and manpower and production planning. The following program is worth considering:

$$\min_{x \in X} \hat{c}_0(x) + E_{\xi_1} Q_1(x, \xi_1) \tag{1}$$

where  $X = \{x | c_0(x) = 0\} \subseteq \mathfrak{R}^{n_0}$ , the recourse function

$$Q_1(x, \xi_1) = \min_{y_1} q_1(x, y_1, \xi_1) + E_{\xi_2} Q_2(x, y_1, \xi_1, \xi_2) \quad \text{subject to } c_1(x, y_1, \xi_1) = 0 \tag{2}$$

and for  $t=2, \dots, T-1$ , recursively we have

$$Q_t(x, y_1, \dots, y_{t-1}, \xi_1, \dots, \xi_t) = \min_{y_t} q_t(x, y_1, \dots, y_{t-1}, y_t, \xi_1, \dots, \xi_t) + E_{\xi_{t+1}} Q_{t+1}(x, y_1, \dots, y_t, \xi_1, \dots, \xi_t, \xi_{t+1}) \tag{3}$$

$$\text{subject to } c_t(x, y_1, \dots, y_t, \xi_1, \dots, \xi_t) = 0 \tag{4}$$

In this case, the deterministic vector is represented by  $Q_T = 0$ .  $x \in \mathfrak{R}^{n_0}$ , the random vector realization  $\hat{\xi}_i$ , and the  $i$ -th stage decision vector  $\xi_i$ .  $y_i \in \mathfrak{R}^{n_i}$ , whose generations occur recursively through  $x, y_1, \dots, y_{i-1}$  and  $\hat{\xi}_1, \dots, \hat{\xi}_i$ . Given  $\mathfrak{R}^{n_0}$  and its associated real-valued functions in the form of  $\hat{c}_0$  and  $c_0$ , we have  $y_i(x, y_1, \dots, y_{i-1}, \hat{\xi}_1, \dots, \hat{\xi}_i)$ . With a correlation between  $c_t$  and  $\hat{\xi}_1, \dots, \hat{\xi}_t$  documented, the form remains random.

## 2 Methodology

Given  $\xi = (\xi_1, \dots, \xi_{T-1})$  as a discrete random vector, should  $c_{it} (i=1, \dots, S_t)$  be the infinite realizations of  $c_t$ , then on stage  $t$ , constraint functions become  $c_{ti}$ .

In [1], a comprehensive description regarding multi-stage stochastic program establishment is provided.

The introduction of the scenario analysis approach has been informed by the need to ensure that multi-stage stochastic programs are handled, an example of such a study being [2]. This approach focuses on a scenario's finite number with time.

Given  $\xi = (\xi_1, \dots, \xi_{T-1})$  and the central assumption being that the associated space of probability is  $(\Omega, \theta, P)$ , if  $\xi^{(s)} = (\xi_1^{(s)}, \xi_2^{(s)}, \dots, \xi_{T-1}^{(s)})$  is the scenario at hand, with the probability distribution known and fixed, then it is possible to reformulate (1)-(4) in such a way that:

$$\begin{aligned} &\min \sum_{s=1}^S f_s(z^{(s)}) \\ &s.t. \quad h_s(z^{(s)}) = 0, s = 1, 2, \dots, S \\ &\quad \sum_{s=1}^S A_s(z^{(s)}) = 0, \end{aligned} \tag{5}$$

In this case,  $z^{(s)} = (x^{(s)}, y_1^{(s)}, \dots, y_{T-1}^{(s)}) \in \mathfrak{R}^n, n = \sum_{i=0}^{T-1} n_i$

$$f_s(z^{(s)}) = p_s(\hat{c}_0(x^{(s)}) + \sum_{t=1}^{T-1} q_t(x^{(s)}, y_1^{(s)}, \dots, y_t^{(s)}, \xi_1^{(s)}, \dots, \xi_t^{(s)})) \tag{6}$$

Deriving further, we let  $h_s(z^{(s)}) = (c_0(x^{(s)}); c_1(x^{(s)}, y_1^{(s)}, \xi_1^{(s)}); \dots; c_{T-1}(x^{(s)}, y_1^{(s)}, \dots, y_{T-1}^{(s)}, \xi_1^{(s)}, \dots, \xi_{T-1}^{(s)}))$

At this point, the non-anticipativity constraints are constraints (7), suggesting that when some situations share some history to a moment selected, their decision all the way to the given moment must have been similar, with additional details in relation to this reformulation given in [2]. In the current investigation,

(1.5)-(1.7) are solved as the focal points. The study's assumption is that should  $f_s : \mathfrak{R}^n \rightarrow \mathfrak{R}$  and  $h_s : \mathfrak{R}^n \rightarrow \mathfrak{R}^m$  be continuously differentiable functions, twice in particular, then  $z^{(s)} \in \mathfrak{R}^n, h_{si} : \mathfrak{R}^n \rightarrow \mathfrak{R} (i=1, \dots, m)$  and  $h_s(z^{(s)}) = (h_{s1}(z^{(s)}), \dots, h_{sm}(z^{(s)}))^T$ . Also, it can be seen that a special structure is obtained in the full row rank matrix representing  $A_i \in \mathfrak{R}^{m_i \times n}, A = (A_1, A_2, \dots, A_s) \in \mathfrak{R}^{m_0 \times nS}$ .

### 3 Results and Discussion

In situations where the scenario S is large, it is worth indicating that (5)-(7) also tend to be large, proving difficult to solve directly through techniques of general-purpose optimization. The eventuality is that decomposition methods have been established in a quest to address this gap, seeking to solve stochastic programming issues [3]. Additionally, computer parallelization has proved insightful and critical relative to feasible implementations of decomposition approaches. It is also notable that many studies have examined this area [4-7], which concerns nonlinear and linear programming. For the majority, they have focused on renowned decomposition guidelines [8], as well as the augmented Lagrangian function and Lagrangian-based duality theory. One of the efficient approaches involves the L-shaped decomposition technique, aiding in solving multi-stage stochastic linear issues. In the respective cycles, there is a recursive generation of optimal and feasible cut sets, causing a reduction in feasible regions. Other studies have documented other approaches [9, 10], including the logarithmic barrier approach, through which the programs could be solved and, in turn, steer system optimization. Given that all the methods rely on stochastic linear programs' special properties and structures, it is imperative to highlight that the generalization of the methods towards solving stochastic nonlinear programs is difficult.

With the scenario analysis approach employed [2], an iterative algorithm has been proposed in the form of the PHA (progressive hedging method). Also, this method has been applied towards stochastic generalized network solutions [10], with findings demonstrating satisfactory numerical observations. Other studies that have implemented the PHA technique [4, 5] contend that its challenge involves how to select a proper penalty variable. Hence, it has been affirmed not to be the best approach for application in loosely-coupled scenario analysis issues. In response, studies have proposed the bundle-based decomposition technique [6], which has been observed to be superior, compared to PHA. More recently, a newer approach has been proposed [7], capable of addressing non-anticipativity constraints using a Lagrangian dual method, as well as combining logarithmic barrier techniques. Indeed, this method's implementation is ongoing.

This study's focus is on the efforts seeking to establish a nonlinear programming model targeting problems with much nonlinearity, as well as linear constraints existing in large sparse sets; with the objective function on the focus. From experience, most of the linear programming issues are large (inordinately) because of their attempt towards approximation via piecewise linearization, translating into nonlinear issues. Also, most of the problems in the real-world have fewer variables linked to the objective function's nonlinearity. The eventuality is that problems that are considered in this study exist in the forms:

$$\text{minimize } F(x) = f(x^N) + c^T x, \quad (7)$$

$$\text{subject to } Ax = b, \quad (8)$$

$$l < x < u \quad (9)$$

where  $A$  is  $m \times n, m \leq n$ .

In this case,  $x^N$  and  $x^L$ , which are linear portions, are achieved via the partitioning of  $x$

Nonlinear variables, in this case, are represented by  $x^N$ , with  $A$  and  $c$  affecting all variables. In some situations,  $x^N$  in  $c^T x$  could be introduced in  $f(x^N)$ . In other scenarios,  $c$  becomes 0. The central assumption in

this case is that  $f(x^N)$ , given a feasible zone with gradient, could be differentiated continuously. This trend gives:

$$\nabla f(x^N) = g(x)$$

Another assumption at this stage is that at any feasible point, represented by  $x^N$ , indeed,  $f$  and  $g$  can be computed. The eventual model is similar to the reduced-gradient approach [8], as well as the variable-reduction technique, observations drawn from the assertions by Gill and Murray [9]. This model could also be seen to extend the work of [2], who focused on the revised simplex approach.

#### 4 Conclusion

When  $x$  and  $F(x)$  are partitioned to achieve nonlinear and linear terms, descriptive procedures can be realized. However, through  $f(x)$  and  $g(x)$ ,  $F(x)$  and  $\nabla F(x)$  can be denoted easily.

In this case, matrices are expected to involve upper case letters while vectors are represented using lower case letters. For scalars presented in Greek, they are presented via lower-case letters. Thus,  $\epsilon > 0$  is the floating-point arithmetic's precision. When one subscript characterize a vector, we have:

$$z = (z^{(1)T}, z^{(2)T}, \dots, z^{(S)T})^T$$

And

$$\nabla_{z^{(s)}} = \frac{\partial \cdot}{\partial z^{(s)}}, \nabla \cdot (z) = \frac{\partial \cdot}{\partial z}, \nabla \cdot (\mu) = \frac{\partial \cdot}{\partial \mu} \quad (10)$$

The implication is that  $\nabla \cdot (z) = (\nabla_1^T \nabla_2^T \dots \nabla_s^T)^T$ .

To simplify the statement, certain functions are employed. In such cases, corresponding variables are omitted. Examples include  $\nabla_s h_s(z^{(s)})$  representing  $\nabla_s h_s, h_{k_s}$  and  $h_s(z_k^{(s)})$ . If specificity is lacking, then vectors in the entirety are in the form of column vectors  $\|\cdot\|$

#### References

- [1] P. Kall and S.W. Wallace. 1994. *Stochastic Programming*. John Wiley & Sons, NY.
- [2] R.T. Rockafellar and R.J-B. Wets. 1991. *Scenarios and Policy Aggregation in Optimization Under Uncertainty*. Math. Oper. Res., 16, 119-147.
- [3] A. Ruszczyński. 1997. *Decomposition Methods in Stochastic Programming*. Math. Prog., 79, 333-353.
- [4] L.S. Lasdon. 1970. *Optimization Theory for Large Systems*, Macmillan, New York.
- [5] B. Feinberg. 1989. *Coercion Functions and Decentralized Linear Programming*, Math. Oper. Res., 14, 177-187.
- [6] S.P. Han. 1989. *A Decomposition Method and Its Application To Convex Programming*, Math. Oper. Res., 14, 237-248.
- [7] A. Ruszczyński. 1995. *On Convergence Of An Augmented Lagrangian Decomposition Method For Sparse Convex Optimization*. Math. Oper. Res., 20, 634-656.
- [8] G.B. Dantzig and P. Wolfe. 1961. *Decomposition Principle for Linear Programs*. Oper. Res., 8, 101-111.
- [9] J.R. Birge. 1995. *Current Trends In Stochastic Programming Computation And Applications, Technique Report 95-15*. Department of Industrial and Operations Engineering, University of Michigan.
- [10] J.R. Birge and F. Louveaux. 1999. *Introduction to Stochastic Programming*, Springer-Verlag, 1997.

- [11] Bin Saion, M. P. . (2022). Simulating Leakage Impact on Steel Industrial System Functionality. *International Journal of New Practices in Management and Engineering*, 10(03), 12–15. <https://doi.org/10.17762/ijnpme.v10i03.129>
- [12] Yadav, S. ., S. . Ali, and A. . Arora. “Design Simulation of Improved Power Quality Conditioner System for Enhancement of Power Quality”. *International Journal on Recent and Innovation Trends in Computing and Communication*, vol. 9, no. 12, Dec. 2021, pp. 19-26, doi:10.17762/ijritcc.v9i12.5496.