시뮬레이션을 통한 시스템 최적화 과정에서 공통 난수 활용의 이점 분석

A Benefit Analysis of Using Common Random Numbers When Optimizing a System by Simulation Experiments

박진원*
Jin-Won Park

Abstract

One of the primary goals of the simulation experiments is to understand the overall system behavior and to analyze the system, ultimately to optimize the system. Optimizing the system includes determining the optimum condition of the system parameters of interest. This paper is concerned with the simulation methodology for estimating the unknown objective function for the system of interest and optimizing the system with respect to the controllable factors. In the process of estimating the unknown objective function, which is assumed to be a second order spline function, we use common random numbers for different set of the controllable factors resulting in more accurate parameter estimation for the objective function. We will show some mathematical result for the benefit of using common random numbers.

Key Words: Simulation Optimization, Common Random Numbers, Correlated Experiments

* 홍익대학교 과학기술대학 전자전기컴퓨터공학부

I. Introduction

Computer simulation experiments are widely used in both the academic and the industrial fields. Often, we exploit simulation for the purpose of understanding the overall system behavior and analyzing the system, ultimately optimizing the system. Optimizing the system involves determining the optimum condition of the system parameters of interest.

There have been many different approaches for finding an optimum condition of a system when the system behavior is random, so they can not be dealt with the conventional mathematical programming methods. One of the typical approaches is using simulation experiments.

The general discussion on the optimization methods using the simulation experiments for finding an optimum condition of a system with random behavior was given in Park and Lee[2] and Yang and Lee[4]. Yang and Lee[4] classified the simulation optimization problems into 2 classes, single response problems and multiple response problems. Based on their classification, we are dealing with single response problems, of which the solution methods are available in the form of the Gradient Based Search Methods[1,5,12], the Stochastic Approximation Methods[10,17], the Heuristic Search Methods[9,11,13] and the Response Surface Methods[6,8,15,16,19]. Among them, we are focusing on the Response Surface Method which is also defined as the functional approach where we are approximating the system behavior of interest as a second order spline function with respect to the controllable factors. The controllable factors are often called as the decision variables. This approach was first initiated by Smith [19] and was later extended by Daugherty & Turnquist[8], and Park[3,16].

In the process of estimating the objective function, which is unknown but is assumed to be a second order spline function, we use common random numbers for different set of the decision variables, resulting in more accurate parameter estimation for the objective function. Obviously, more accurate parameter estimation will give us more accurate optimum solution to the system of interest.

Conceptually, using common random numbers in simulation optimization procedure is similar to setting the same environment in physical experiments so as to reduce the effect of nuisance variables. A few research have been devoted to use common random numbers in order to reduce the variance of the parameter estimator for the unknown objective function in RSM[7,16,18,20], since it is complicated to analyze the effect of using common random numbers.

The primary goal of this paper is to analyze the impact of using common random numbers and show some mathematical result for the benefit of using common random numbers in RSM type simulation optimization procedures.

The paper is organized as follows. In section II, after the introduction in this section, we will set up the simulation optimization problem using RSM, which was well defined in Park[3,16], Kwon[1] and Yang and Lee[4]. The major result of using common random numbers in estimating the parameters for the unknown objective function is given in section III. Finally, some comments and the possible future research topics will be included in section V. Also, the mathematical result and the proof for the theorem concerning the benefit of using common random numbers will be explained in detail in the Appendix.

II. The Simulation Optimization Problem

Let us consider a search method for finding the optimal operating condition over controllable factors of a system by performing sets of simulation experiments along with solving mathematical programming problems [4,10,16]. Under the assumption of the followings, we will perform a set of simulation experiments over a setting of controllable factors. First, the system performance measure of interest is not mathematically tractable. Second, the expected value of the system performance measure of interest is assumed to be unimodal in terms of the decision variables of the system, thus we have unique optimal solution even if we can not find it analytically. Third, we will approximate the unknown expected value of the system performance measure as a second order linear function. Fourth, we may have a linear constraint with respect to the decision variables.

Under the foregoing assumptions, we may have a linear model of

$$\begin{aligned} \textit{Optimize} \quad & Y(x) = \bar{y} \\ \textit{where} \quad & E(Y(x)) = E(f(x) + e(x)) \\ & = E(q(x) + h(x) + e(x)) \\ & = y_0(x) \\ s.t. \quad & \sum_{j=1}^{j=k} a_j x_j \leq M, M > 0 \\ & a_j > 0, x_j \geq 0 \end{aligned} \tag{1}$$

Here in equation (1), $q(x)$ is a quadratic function, $h(x)$ is the remainder, assuming $e(x)$ is distributed as Normal$(0, \eta^2)$. As we see in equation (1), the form of the objective function is not known, but its value at $\underline{x}$, a setting of experimental design with respect to the decision variables, will be estimated through a simulation run. Also, we have a single linear constraint imposed on the nonnegative decision variables.

III. Common Random Numbers

When we estimate the response function $f(x)$ in equation (1), we perform a set of simulation experiments at n design points with respect to the decision variables $\underline{x}$ of k elements. Let us denote the output of simulation run at design point l as Y_l. Then, we may have a second order linear model of

$$\begin{aligned} Y_l = \; & b_0 + \sum_{j=1}^{j=k} b_j x_{jl} \\ & + \sum_{i=1}^{i=k} \sum_{j=1}^{j=k}{}_{i \leq j} b_{ij} x_{il} x_{jl} + \varepsilon_l \\ & \text{for } l = 1, 2, .., n. \end{aligned} \tag{2}$$

Most linear models assume that

$$\begin{aligned} & Var(Y_l) = Var(\varepsilon_l) = \sigma^2 \\ & \quad \text{for } l = 1, 2, .., n. \\ & \textit{and} \\ & Cov(Y_l, Y_m) = Cov(\varepsilon_l, \varepsilon_m) = 0, \\ & \quad \text{for } l, m = 1, 2, .., n, \ l \neq m \end{aligned} \tag{3}$$

Under assumption (3), the best(minimum variance, uniform, unbiased) estimator for

$$b = [b_0, b_1, \ldots, b_k, b_{11}, \ldots, b_{k-1,k}] \quad (4)$$

is

$$\beta = (XX')^{-1}X'Y \quad (5)$$

and its covariance matrix is

$$Cov(\beta) = \sigma^2(XX')^{-1} \quad (6)$$

where

$$X' = \begin{bmatrix} 1 & x_{11} & x_{21} & \ldots & x_{p1} \\ 1 & x_{12} & x_{22} & \ldots & x_{p2} \\ . & . & . & . & \\ . & . & . & . & \\ 1 & x_{1n} & x_{2n} & \ldots & x_{pn} \end{bmatrix} \quad (7)$$

Here in equation (7), X' is an experimental design matrix augmented by the vector of 1, interactive terms $x_{ij}, i, j = 1, 2, \ldots, k$, $i < j$, and the square terms x_j^2, $j = 1, 2, \ldots, k$, where k is the number of the decision variables, n is the number of the design points. Thus, if we have 2 decision variables and 10 design points, then n is 10 and p is 5, which consists of $x_1, x_2, x_1x_2, x_1^2, x_2^2$. Also, σ^2 is the common variance of ε_l, $l = 1, 2, \ldots, n$.

Suppose that we can correlate Y_l and Y_m through simulation experiments at design points $\underline{x}^l$ and $\underline{x}^m$ using common random numbers for each design point. For the linear model of equation (2), the experimental design which is a part of equation (7), consists of the first order portion of $x_{1j}, x_{2j}, \ldots, x_{kj}$, where k is the number of the decision variables and j is the index for design point j.

If covariance matrix V of the error terms $\varepsilon = (\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n)$ is known and

$$V = \sigma^2(1-\rho)I + \sigma^2\rho J$$

where

ρ is the common correlation coefficient for ε_l and ε_m, $l \neq m$, $\sigma^2 = Var(\varepsilon_l)$, $l = 1, 2, \ldots, n$, I is an n x n identity matrix, and J is n x n matrix of all the elements being 1, (8)

then the best estimator for b is known as

$$\beta = (XX')^{-1}X'Y \quad (9)$$

and the covariance matrix as

$$Cov(\beta) = (XV^{-1}X')^{-1} \quad (10)$$

or

$$(XV^{-1}X')^{-1} \quad (11)$$

$$= (1-\rho)\sigma^2(XX')^{-1}$$

except element (1,1) which is for β_0.

Note that equations (9) and (10) are from Lewis and Odell[14]. Equation (11) was proved by Schruben and Margolin[18] when we have design matrix X' being orthogonally blockable. However, we proved equation (11) when design matrix X' is not orthogonally blockable but is of general form. We present the proof in the Appendix.

When we have design matrix X' being not orthogonally blockable but of general form, we have the freedom of selecting the design points arbitrarily over the feasible experimental region, which is extremely convenient when we have non cubic form of experimental region such as triangular type. The triangular type feasible region often occurs when we have optimization problems with a linear constraint inequality equation.

The result obtained in equation (11) says that if we have a particular type of correlated experiments, we can estimate parameters $\boldsymbol{b}$ in equations (2),(4) more accurately by the factor of $(1-\rho)$ except b_0 term. Note that the constant term, b_0 does not affect the shape and the location of the optimum solution of the problem. It only has an effect on the optimum value of the system. Thus, if we use common random numbers at different design points and consequently have positively correlated simula tion experiments, then we can find the optimum solution of the system of interest more precisely, which is the primary impact of using common random numbers.

Another observation for the result obtained in equation (11) reveals that if $\rho=1$, where the simulation outputs at different experimental points are perfectly correlated, then all the elements of the covariance matrix of V becomes 0. That is the ideal case where we can estimate the parameters of the model without errors. If that is the case, we can find the optimum solution correctly if the true objective function is the form of the model described in equation (2). However, if the true objective function of the system is not the form of equation (2), then we may have to devote to find an appropriate form of the objective function in equation (2).

Another extreme where $\rho = 0$, which means that there are no correlation between the simulation outputs at different design points, then we have to conclude that there is no advantage of using common random numbers in this process of simulation optimization via RSM. However, in many real situations, the common correlation coefficient ρ value turned out to be between 0 and 1[16,18].

IV. Conclusion and Further Research Topics

We have presented a benefit of using common random numbers for estimating the parameters of the objective function in RSM. The major result obtained in section III revealed that if we use common random numbers for estimating the parameters of the unknown objective function, it is helpful by the factor of ρ, where ρ is the common correlation coefficient between the system behavior measures at different experimental design points. The result is valid regardless of the orthogonality of the experimental design.

Apparently, it is critical to check if we have realized the covariance matrix for the error terms ε not being significantly departed from a special structure V as in equation (8). Thus, it is required to go through the covariance structure test procedure, which is left for the further research topic.

Appendix

Let X' be an experimental design augmented as in equation (7). Also, let Y_l $l=1,2,\ldots,n$ be n simulation outputs at n design points, and V be the covariance matrix of $\boldsymbol{Y}=(Y_1, Y_2, \ldots, Y_n)$.

Theorem:

The covariance matrix of the best estimator

$\boldsymbol{\beta}$ in equation (9) for $\boldsymbol{b}$ in equation (2) is given by

$$Cov(\beta) = (XV^{-1}X')^{-1} \quad \text{or} \tag{A1}$$

$$(XV^{-1}X')^{-1} = (1-\rho)\sigma^2(XX')^{-1} \tag{A2}$$

except element (1,1) regardless of the form of the design matrix, which means that the design matrix may not be orthogonal.

Proof:

step 1

Without loss of generality, let us assume that $\sigma^2 = 1$. We will show that

$$V^{-1} = \begin{bmatrix} a & b & b & \dots & b \\ b & a & b & \dots & b \\ \cdot & \cdot & \cdot & \cdot & \\ \cdot & \cdot & \cdot & \cdot & \\ b & b & b & \dots & a \end{bmatrix} \tag{A3}$$

where

$$a = (1+(n-2)\rho)/((1+(n-1)\rho)(1-\rho))$$
$$b = -\rho/((1+(n-1)\rho)/(1-\rho)), \quad \rho \neq 1. \tag{A4}$$

Thus,

$$V^{-1} = (a-b)I + bJ \tag{A5}$$

where I is n x n unit matrix and J is n x n matrix with all the elements being 1.

step 2

We will show that

$$\begin{aligned}(X\ V^{-1}X')^{-1} &= [X((a-b)I+bJ)X']^{-1} \\ &= [X(I/(1-\rho) - \rho J/((1+(n-1)\rho)(1-\rho)))X']^{-1} \\ &= (1-\rho)[XX' - (\rho/(1+(n-1)\rho))XJX']^{-1}.\end{aligned} \tag{A6}$$

Then, consider $D = (XX')^{-1}$ and $C = (XX' - hXJX')^{-1}$ where $h = \rho/(1+(n-1)\rho)$. We will show that $d_{ij} = c_{ij}$ except $i=j=1$ where $D = [d_{ij}]$ and $C = [c_{ij}]$.

Now, let us show steps 1 and 2 in detail.

step 1

1) V is a non-singular, symmetric matrix, so is V^{-1}, since $V = V'$, $V^{-1} = (V')^{-1} = (V^{-1})'$.
2) Let $V^{-1} = [v_{ij}]$, and V_{ij} be the submatrix of V formed by deleting the i-th row and the jth column of V, then

$$\begin{aligned} v_{ij} &= (-1)^{i+j}\,|V_{ji}|\,/det\ V \\ &= (-1)^{i+j}\,|V_{ij}|\,/det\ V \end{aligned} \tag{A7}$$

since V is symmetric and $V_{ij} = V_{ji}$.

Clearly, $V_{ii} = V_{jj}$, so $v_{ii} = v_{jj}$. (A8)

Consider v_{ij} and $v_{i+1,j}$ where $i > j$, then

$$\begin{aligned} v_{ij} &= (-1)^{i+j} \mid V_{ij} \mid / det\ V \\ &= (-1)^{i+j+1} \mid V_{i+1,j} \mid / det\ V \qquad (A9) \\ &= v_{i+1,j} \end{aligned}$$

since $V_{i+1,j}$ can be obtained by interchanging the *(i-1)*st column and the i-th column of V_{ij}, thus

$\mid V_{i+1,j} \mid = (-1) \mid V_{ij} \mid$.

Also, consider v_{ij} and $v_{i,j+1}$ where $i > j+1$,

$$\begin{aligned} v_{ij} &= (-1)^{i+j} \mid V_{ij} \mid / det\ V \\ &= (-1)^{i+j+1} \mid V_{i,j+1} \mid / det\ V \qquad (A10) \\ &= v_{i,j+1} \end{aligned}$$

since $V_{i,j+1}$ can be obtained by interchanging the i-th row and the *(j+1)*-st row of $V_{i,j}$, thus

$\mid V_{i,j+1} \mid = (-1) \mid V_{ij} \mid$.

From equations (A7) to (A10),

$v_{11} = v_{22} = \ldots = v_{nn}$

$v_{21} = v_{31} = v_{32} = v_{34} = \ldots = v_{n-1,n}$.

Also, the symmetricity of V^{-1} indicates that

$v_{21} = v_{31} = v_{32} = v_{43} = \ldots = v_{n,n-1}$

$v_{12} = v_{13} = v_{23} = v_{34} = \ldots = v_{n-1,n}$.

Finally, V^{-1} has the form of equation (A3).

Let us find a, b in V^{-1}. Since $VV^{-1} = I$,

$a + (n-1)b\rho = 1$, and

$a\rho + b + (n-2)b\rho = 0, n \geq 2$,

we have

$a = (1 + (n-2)\rho)/h$

$b = -\rho/h$,

where

$h = (1 + (n-1)\rho)(1-\rho)$ and $\rho \neq 1$.

<u>step 2</u>

First, for simplicity, let us define $A = XX'$ and $B = XX' - hXJX'$ in equation (A6), so that $D = A^{-1}$ and $C = B^{-1}$.

$$B = XX' - hXJX' = A - hXJX'$$

$$= \begin{bmatrix} n & \Sigma x_{1j} & \Sigma x_{2j} & .. & \Sigma x_{pj} \\ & \Sigma x_{1j}^2 & \Sigma x_{1j}\Sigma x_{2j} & .. & \Sigma x_{1j}\Sigma x_{pj} \\ & & . & . & \\ sym & & & . & \\ & & & & \Sigma x_{pj}^2 \end{bmatrix}$$

$$-h \begin{bmatrix} n^2 & n\Sigma x_{1j} & n\Sigma x_{2j} & .. & n\Sigma x_{pj} \\ & (\Sigma x_{1j})^2 & \Sigma x_{1j}\Sigma x_{2j} & .. & \Sigma x_{1j}\Sigma x_{pj} \\ & & . & . & \\ sym & & & . & \\ & & & & (\Sigma x_{pj})^2 \end{bmatrix}$$

$$= \begin{bmatrix} kn & k\Sigma x_{1j} & k\Sigma x_{2j} & .. & k\Sigma x_{pj} \\ & Q_{11} & Q_{12} & .. & Q_{1p} \\ & & . & . & \\ sym & & & . & \\ & & & & Q_{pp} \end{bmatrix} \qquad (A11)$$

where

$\Sigma = \Sigma_{j=1}^{p}$, $k = 1 - hn$,

$h = \rho/(1 + (n-1)\rho)$,

$$Q_{11}=\Sigma x_{1j}^2-h\Sigma x_{1j}\Sigma x_{1j}$$

$$Q_{12}=\Sigma x_{1j}^2-h\Sigma x_{1j}\Sigma x_{2j}$$

$$Q_{1p}=\Sigma x_{1j}^2-h\Sigma x_{1j}\Sigma x_{pj}$$

$$\cdots$$

$$Q_{pp}=\Sigma x_{pj}^2-h\Sigma x_{pj}\Sigma x_{pj}.$$

Let $A=XX'=[a_{ij}], i,j=1,2,\ldots,q,$ where $q=p+1,\ \alpha_i=h\Sigma x_{ij},$ then from equation (A11),

$$B=\begin{bmatrix} ka_{11} & ka_{12} & .. & ka_{1q} \\ a_{21}-\beta_1 a_{11}, & a_{22}-\beta_1 a_{12}, & .. & a_{2q}-\beta_1 a_{1q} \\ \cdot & \cdot & \cdot & \cdot \\ a_{q1}-\beta_p a_{11}, & a_{q2}-\beta_p a_{12}, & .. & a_{qq}-\beta_p a_{1q} \end{bmatrix} \quad \text{(A12).}$$

Note that A and B are symmetric since $A=XX'$ which is clearly symmetric and $B=XX'-hXJX'$ which is a sum of two symmetric matrices.

From equation (A12), we know that $det\ B=k\ det\ A$, since

1) the first row of B is k times the first row of A, and the i-th row of B is the i-th row of A minus, β_i times the first row of A, where i is from 2 to q.
2) if one row of a matrix is multiplied by k, the determinent of the resulting matrix is k times the determinent of the original matrix.
3) the determinent of a matrix is unchanged, if a row is subtracted from the multiple of another row.

Now, let $D=A^{-1}=[d_{ij}]$ and $C=B^{-1}=[c_{ij}]$. Then, D and C are symmetric since A and B are symmetric. Let A_{ij}, B_{ij} be the submatrices of A and B, formed by deleting the i-th row and the j-th column of A and B respectively.
Then,

$$d_{ij}=(-1)^{i+j}\,|A_{ij}|\,/det\ A,$$
$$c_{ij}=(-1)^{i+j}\,|B_{ij}|\,/det\ B. \quad \text{(A13)}$$

But, we know that $det\ B = k\cdot det\ A$ and that $|B_{ij}| = k\cdot det\ A$, except $|B_{11}|$, which is equal to $|A_{11}|$. Thus, it is now trivial that $d_{ij}=c_{ij}$ except d_{11}, so $A^{-1}=B^{-1}$ except element (1,1). Finally, we have

$$\begin{aligned}(XV^{-1}X')^{-1} &= (1-\rho)B^{-1} = (1-\rho)C \\ &\cong (1-\rho)D = (1-\rho)A^{-1} \\ &= (1-\rho)(XX')^{-1}\end{aligned} \quad \text{(A14).}$$

Here, $\cong$ means that the equality does not exactly hold since $c_{11}\neq d_{11}$, thus the relation in equation (A14) holds except element (1,1).

References

[1] 권치명, "퍼터베이션 분석을 이용한 대기 행렬 네트워크의 최적화," 한국시뮬레이션학회 논문지, 제9권, 제2호(2000), pp.89-102.

[2] 박경종, 이영해, "이산형 변수 시스템의 설계를 위한 시뮬레이션 활용 기법 연구," 한국시뮬레이션학회 논문지, 제8권, 제3호(1999), pp.1-16.

[3] 박진원, "컴퓨터 시뮬레이션을 이용한 최적화 기법" 정보과학회지, 제8권, 제1호(1990), pp.37-47.

[4] 양병희, 이영해, "시뮬레이션 최적화 기법과 절삭 공정에의 응용," 한국시뮬레이션학회 논문지, 제3권, 제2호(1994), pp.57-67.

[5] Biles, W.E. and J.J. Swain, *Optimization and Industrial Experimentation,* Wiley Interscience, New York, 1980.

[6] Box, G.E.P. and J.S. Hunter, *Multifactor Experimental Design for Exploring Response Surfaces,* Annals of Mathema -tical Statistics, Vol.28(1957), pp.195-241.

[7] Cooley, B.J. and E.C. Houck, *A Variance Reduction Strategy for RSM Simulation Studies,* Decision Sciences, Vol.13(1982), pp.303-321.

[8] Daughety, A.F. and M.A. Turnquist, *Budget Constrained Optimization of Simulation Models via Estimation of Their Response Surfaces,* Operations Research, Vol.29, No.3(1981), pp.485-500.

[9] Eglese, R.W., *Simulated Annealing: A Tool for Operational Research,* European J. of Operations Research, Vol.46(1990), pp.271-281.

[10] Glynn, P.W., *Optimization of Stochastic Systems,* Proc. of the 1986 Winter Simulation Conf.(1986), pp.52-58.

[11] Goldberg, D.E., *Genetic Algorithm in Search, Optimization and Machine Learning,* Addison-Wesely Pub., 1989.

[12] Ho, Y.C. and X.R. Cao, Pertubation Analysis Discrete Event Dynamic systems, Kluwer Academic Publishes, 1991.

[13] Lee, Y.H. and K. Iwata, *Part Ordering through Simulation Optimization in FMS,* Int. J. of Prod. Res., Vol.29, No.7(1991), pp1309-1323.

[14] Lewis, T.O and P.L. Odell, *Estimation in Linear Models,* Prentice-Hall Inc., Engle -wood Cliff, NJ, 1971.

[15] Myers, R., *Response Surface Methodology,* Allyn-Bacon, Boston, 1971.

[16] Park, J.W., *Simulation Optimization with Discrete Decision Variables and a Single Linear Constraint,* Ph.D. Dissertation, Ohio St. Univ., Columbus, Ohio, 1987.

[17] Robbins, H. and S. Monro, A Stochastic Approximation Method, Annals of Mathe -matical Statistics, Vol.22(1951), pp.400-407.

[18] Schruben, L.W. and B.H. Margolin, Pseudo random Number Assignment in Statistically Designed Simulation and Distribution Sampling Experiments, J. of Amer. Stat. Assoc., Vol.73, No.363(1978), pp.504-525.

[19] Smith, D.E., An Empirical Investigation of Optimum Seeking in the Computer Simulation Situation, Operations Research, Vol.21, No.2 (1973), pp.475-497.

[20] Tew, J.D. and J.R. Wilson, Validation of Correlation Induction Strategies for Simulation Experiments, Proc. of the 1985 Winter Simulation Conf.(1985), pp.190-195.

● 저자소개 ●

박진원

1975 서울대학교 공과대학 산업공학과 학사
1987 미국 오하이오주립대 산업공학과 박사
1987~1988 미국 남콜로라도대학 교수
1988~1999 한국전자통신연구원 책임연구원
1999~2000 영산대학교 교수
2000~현재 홍익대학교 교수(전자전기컴퓨터공학부)
관심 분야 : Computer System Simulation, 연구개발정책, 컴퓨터시스템 설계