

Statistical surrogate model based sampling criterion for stochastic global optimization of problems with constraints[†]

Su-gil Cho^{1,3}, Junyong Jang¹, Jihoon Kim¹, Minuk Lee², Jong-Su Choi⁴, Sup Hong³ and Tae Hee Lee^{1,*}

¹*Department of Automotive Engineering, Hanyang University, Seoul, 133-791, Korea* ²*Romax Technology Ltd., Seoul, 137-889, Korea*

³*Technoloty Center for Offshore Plant Industries, Korea Research Institute of Ships & Ocean Engineering, Daejeon, 305-343, Korea* ⁴*Offshore Plant Research Division, Korea Research Institute of Ships & Ocean Engineering, Daejeon, 305-343, Korea*

(Manuscript Received October 5, 2014; Revised December 30, 2014; Accepted January 21, 2015) --

Abstract

Sequential surrogate model-based global optimization algorithms, such as super-EGO, have been developed to increase the efficiency of commonly used global optimization technique as well as to ensure the accuracy of optimization. However, earlier studies have drawbacks because there are three phases in the optimization loop and empirical parameters. We propose a united sampling criterion to simplify the algorithm and to achieve the global optimum of problems with constraints without any empirical parameters. It is able to select the points located in a feasible region with high model uncertainty as well as the points along the boundary of constraint at the lowest objective value. The mean squared error determines which criterion is more dominant among the infill sampling criterion and boundary sampling criterion. Also, the method guarantees the accuracy of the surrogate model because the sample points are not located within extremely small regions like super-EGO. The performance of the proposed method, such as the solvability of a problem, convergence properties, and efficiency, are validated through nonlinear numerical examples with disconnected feasible regions.

--

Keywords: Constrained global optimization; Metamodel-based design optimization; Kriging surrogate model; Stochastic global optimization

1. Introduction

Global optimization has gained much attention because it can systematically provide not only the best design solution within a design space, but also multiple alternatives of the best designs [1-4]. However, its computational burden remains high. To overcome this difficulty and maximize the benefits of global optimization, a surrogate model, often referred to as a metamodel, has emerged as an effective alternative [5, 6]. The surrogate model is an approximate model to transfer the implicit relationship between the design variables and the response into an explicit one that can be easily expressed with simple basis functions or polynomials. One of the greatest advantages is that the surrogate model can quickly and cheaply evaluate responses at any untried points over the design domain. Therefore, the computational burden of global optimization can be remarkably reduced by the surrogate model [7].

However, metamodel-based design optimization (MBDO) can frequently lead to quite a large violation from the actual constraints even if there is only a slight error of the surrogate

Recommended by Guest Editor Joo-Ho Choi © KSME & Springer 2015

model in the neighborhood of the optimum because most optimums are located on the boundary of the active constraints [8]. Therefore, a sequential surrogate model-based global optimization algorithm has been recently developed to increase the efficiency of commonly used global optimization algorithms as well as to ensure the accuracy of optimization [9]. This approach focuses more on how to efficiently explore a global optimum with a surrogate model rather than how to globally build an accurate surrogate model over the design space. In this research, we used a surrogate model to find an initial solution, refine the solution and search for another feasible point in an optimization algorithm. Therefore, this algorithm provides an accurate optimum because it samples responses at the optimum.

For constrained global optimization, super-EGO, which is an enhanced version of EGO, is known to solve various problems including disconnected feasible regions and nonlinearity [10]. When it explores a global optimum, super-EGO uses three phases: two stochastic searching phases and a local search phase. In the two stochastic phases, the criteria are used to search the first feasible region and move to another feasible region, respectively. In the local search phase, DIRECT algorithm finds the local optimum point if the sampling point satisfies the constraints once. However, if the surrogate model

^{*}Corresponding author. Tel.: +82 2 2220 0449, Fax.: +82 2 2220 2299

E-mail address: thlee@hanyang.ac.kr

[†]This paper was presented at the 8th CJK-OSM, Gyeongju, Korea, May 2014.

varies significantly from the sampling points, super-EGO spends most of the optimization time in the local search phase without stochastic information. In this case, super-EGO is less effective than standard global optimization. Thus, it is inappropriate to classify super-EGO as a stochastic global optimization. In addition, because the local search phase locates the sample points within an extremely small region, this increases the likelihood of a singular correlation matrix of the surrogate model and the model gives inaccurate prediction. Lastly, super-EGO requires several empirical parameters, which is hard to determine suitable parameters for various problems [10]. Therefore, it is necessary to develop a stochastic and unified sampling criterion without parameters that can be easily applicable to various problems.

We propose a united sampling criterion based on a statistical surrogate model for effective global optimization of problems with constraints. The proposed method consists of feasible search parts and an expected improvement (EI) part. The feasible search part, which plays a key role of the method, is a combination of the infill sampling criterion, the boundary sampling criterion, and the mean squared error (MSE) from kriging surrogate model. The MSE determines which criterion is more dominant among two criteria. It is able to select the points located in a feasible region with high model uncertainty as well as the points along the boundary of constraint at the lowest objective value. The performance of the proposed method, including the solvability of a problem, convergence properties, and efficiency is validated through nonlinear numerical examples with disconnected feasible regions.

2. Kriging-based stochastic probability

Kriging originally comes from the field of geostatistics as a method to predict geological data, such as the thickness of ore layers [11]. Sack and coworkers exploited the kriging model as a prediction tool for engineering designs, where it is named as "Design and analysis of computer experiments (DACE)" [12]. The kriging surrogate model is an interpolation model that is appropriate for deterministic responses. Recently, a variety of sampling criteria using the kriging surrogate model have been widely developed due to its excellent prediction performance and its useful statistical quantities [10, 13]. In this section, the ways for dealing with stochastic probability in global optimization are described.

2.1 Space filling sampling criterion

Once the kriging surrogate model is generated, it can provide not only the predicted value, but also its stochastic prediction error denoted by $\hat{\sigma}^2$. The MSE is directly related to highes the uncertainty of the predicted value. At the sample data, the model passes exactly through the data and the MSE becomes zero. Meanwhile, the uncertainty of prediction at an untried point becomes high as an untried point is far away from the sample points. Thus, the sampling technique for searching the

Fig. 1. Graphical expression for the probability of the EI.

point with the highest value of the MSE is able to locate the new experimental points farthest from the sample points. That is why it is called "Space filling sampling criterion".

2.2 Expected improvement sampling criterion

Kriging theory assumes that both the sample data and the predicted values of the kriging surrogate model are normally distributed. Let us consider a response of the kriging surrogate model $Y(x_1)$ at the design point of x_1 .

We use Kushner's criterion followed by the EI probability [14]. The stochastic probability that x_I is smaller than f_{min} , the probability of improvement, is defined as

$$
P(\hat{Y}(x_1) \le f_{\min}) = \Phi(I_1) = \Phi\left(\frac{f_{\min} - \hat{Y}(x_1)}{\hat{\sigma}_Y(x_1)}\right),\tag{1}
$$

where Φ (\cdot) is the cumulative distribution function (CDF) of the standard normal distribution and I_I is the stochastic quantity depending on *x¹* . The minimum value among the observed data is denoted by *fmin*.

To illustrate the basic search strategy of the EI probability, a one-dimensional example is shown in Fig. 1. The sine-shaped solid line of the upper figure is a kriging surrogate model built from four sample data denoted by solid circles. Then, we examine the probability of EI on three arbitrary sample points, x_1 , x_2 and x_3 . The point of x_1 has the highest probability of improvement among the three different points.

2.3 Infill sampling & boundary sampling criterion

The feasible region consists of an inactive and active region.

Fig. 2. Three-hump camelback example: (a) 3-D plot; (b) contour plot and sampled points, feasible regions are shaded; (c) 3-D plot of infill sampling criterion; (d) 3-D plot of boundary sampling criterion.

First, a sampling criterion for exploring inactive region is the infill sampling defined just like probability of the EI. If a constraint is defined as $G(x) \le 0$, then the infilling probability is as follows: (c) (d) the section, the required to the results and sample with a section, the results of the constrained sampling criterion (d) 3-D plot of binding to the studying the sampling criterion and sampled points, feasible reg

$$
P(G(\mathbf{x}) \le 0) = \Phi\left(\frac{0 - \hat{G}(\mathbf{x})}{\hat{\sigma}(\mathbf{x})}\right),
$$
 the
star

where ϕ (\cdot) is the CDF of the standard normal distribution, $\hat{G}(\mathbf{x})$ is predictor of kriging surrogate model for constraint $\hat{\sigma}$ (x) is the MSE value.

To illustrate the behavior of the criterion, a two-dimensional example with a constraint function is introduced as shown in $P(G(x) \le 0) = \Phi\left(\frac{0 - \hat{G}(x)}{\hat{\sigma}(x)}\right)$, (2) the larger the larger the larger the larger the sumpling port (4) If the sumpling port (4) If the sumpling port (4) If the sumpling surrogate model for constraint That is, it pu 0) = $\Phi\left(\frac{0 - \hat{G}(x)}{\hat{\sigma}(x)}\right)$, the larger the model unc

(4) If the sampling portions there (4) If the sampling cr

is the CDF of the standard normal distribution, That is, it pursues not dictor of kriging surrogate

$$
g(\mathbf{x}) = 2x_1^2 - 1.05x_1^4 + \frac{1}{6}x_1^6 - x_1x_2 + x_2^2 - 0.5 \le 0
$$

$$
x_1 \in [-2, 2], \ x_2 \in [-1.5, 1.5]
$$
 (3)

Example the infinite sum of exponmentative region is the transformation of exponential and the summation is equilibrated on the boundary of a El. If a considered as $G(x) \le 0$, then the infilling probability is as it must This example has three disconnected feasible regions. The two feasible regions are so small that it is difficult to find the solution with other methods. The plot of constraint and a contour plot with feasible sets and sampled points are (a) and (b) as shown in Fig. 2, respectively. 21 sample points based on optimal Latin-hypercube design (OLHD) are used. In the infeasible region, the value of the infill sampling criterion becomes zero in Fig. 2(c). The larger the value this criterion has, the more the sample point locates the interior of the feasible region. Thus, because it selects the point on the inside, it is called the "Infill sampling criterion".

Second, the criterion to find the active region is boundary sampling criterion which locates the sample points around the boundary of constraint [7]. It uses the probability density function (PDF) instead of CDF. The criterion is given as follows: chnology 29 (4) (2015) 1421~1427 1423

ccond, the criterion to find the active region is boundary

bling criterion which locates the sample points around the

ddary of constraint [7]. It uses the probability density func-(1421–1427 1423)

in to find the active region is boundary

ch locates the sample points around the

[7]. It uses the probability density func-

DF. The criterion is given as follows:
 $-\hat{G}(\mathbf{x})$
 $\hat{\sigma}(\mathbf{x})$

(4)

of *Prechnology 29 (4) (2015) 1421-1427* 1423
 Becond, the criterion to find the active region is boundary
 ppling criterion which locates the sample points around the
 mdary of constraint [7]. It uses the probability d 2015) 1421-1427 1423

erion to find the active region is boundary

which locates the sample points around the

aint [7]. It uses the probability density func-

of CDF. The criterion is given as follows:
 $\left(\frac{0-\hat{G}(\mathbf{x})$ ²⁹ (4) (2015) 1421-1427 1423

the criterion to find the active region is boundary

triterion which locates the sample points around the

of constraint [7]. It uses the probability density func-

instead of CDF. The crit ²⁰¹⁵⁾ 1421-1427 1423

erion to find the active region is boundary

which locates the sample points around the

aint [7]. It uses the probability density func-

of CDF. The criterion is given as follows:
 $\left(\frac{0-\hat{G}(x)}{\$ mananology 29 (4) (2015) 1421-1427 1423

Second, the criterion to find the active region is boundary

sampling criterion which locates the sample points around the

boundary of constraint [7]. It uses the probability dens

$$
P(G(\mathbf{x}) \le 0) = \varphi \left(\frac{0 - \hat{G}(\mathbf{x})}{\hat{\sigma}(\mathbf{x})} \right),\tag{4}
$$

A plot of the criterion is Fig. 2(d). The criterion has a larger value along the boundary lines than interior of feasible region, and the value of the criterion becomes zero in an infeasible region. Thus, it is called "Boundary sampling criterion".

3. Stochastic probability based sampling criterion

In this section, the requirements of the sampling criterion for stochastic global optimization are suggested based on studying the sampling criteria as follows:

(1) Optimization technique should find the design point to improve the objective function in a feasible region.

(2) It is necessary to approximate efficiently and accurately the feasible region of the direction to improve the objective function using little sample points. Further, because the optimum point is usually located on the boundary of the constraint, it must be able to sample on the boundary of the constraint.

 $\hat{G}(\mathbf{x})$ the larger the model uncertainty. Thus, it is able to serve as the (3) The model uncertainty for global optimization should be considered. The farther the distance from the sample points, sampling portions thereof.

a sampling eriterion for exploring inactive region is the feasible region of the direction to improve the object
sampling defined just like probability of the EI. If a com-
mum point is usually located on the boundary of *g* a sampling criterion for exploring inactive region is the feasible region of the direction to improve
 g and the sampling defined just like probability of the EI. If a con-

interior is usually located on the bounda **x** ampling criterion for exploring inactive region is the feasible region of the direction to improve the objet
philing defined just like probability of the ELI fa com-
mum point is usually located on the boundary of th (4) If the sampling criterion is in cooperation with the conditions 2 and 3, it can perform optimization more efficiently. That is, it pursues not considering the model uncertainty on the overall design domain, but considering the model uncertainty in feasible regions only. **EVALUATION TOMOS CONTITY TO EXECT THE SURFACE (TO the Proprieson of the sympling criterion**
ic global optimization are suggested based on
sampling criteria as follows:
ization technique should find the design point to
ob egion of the direction to improve the objective
glittle sample points. Further, because the opti-
usually located on the boundary of the constraint,
to sample on the boundary of the constraint.
to sample on the boundary o the model uncertainty. Thus, it is able to serve as the portions thereof.

le sampling criterion is in cooperation with the con-

and 3, it can perform optimization more efficiently.

pursues not considering the model unc feasible region of the direction to improve the objective

rection using little sample points. Further, because the opti-
 m point is usually located on the boundary of the constraint,

must be able to sample on the bou ble region of the direction to improve the objective
using little sample points. Further, because the opti-
int is usually located on the boundary of the constraint,
e able to sample on the boundary of the constraint.
e m mum point is usually located on the boundary of the constraint,

it must be able to sample on the boundary of the constraint,

it must be able to sample on the boundary of the constraint

considered. The farther the dista oundary of the constraint,
lary of the constraint.
la optimization should be
from the sample points,
i, it is able to serve as the
ooperation with the con-
ization more efficiently.
le model uncertainty on
dering the mode to sample on the boundary of the constraint.

Elel uncertainty for global optimization should be

fearther the distance from the sample points,

for delay increasing the sample points,

and electratinty. Thus, it is able ont is usually located on the boundary of the constrant,
the he able to sample on the boundary of the constraint.
the model uncertainty for global optimization should be
red. The farther the distance from the sample poin

Finally, a new sampling criterion that meets the above requirements is proposed as follows:

$$
Max \sum_{j=1}^{n} (\hat{\sigma}_{G_j}(\mathbf{x}) \cdot \Phi(I_C) + \varphi(I_C)) \cdot \Phi(I_O) ,
$$
\n
$$
where I_C = \frac{0 - \hat{G}_j(\mathbf{x})}{\hat{\sigma}_{G_j}(\mathbf{x})}, I_O = \frac{f_{\min} - \hat{Y}(\mathbf{x})}{\hat{\sigma}_{Y}(\mathbf{x})} .
$$
\n(5)

Primarily, the proposed sampling criterion consists of an objective and constraint parts. The objective part is the same with the probability of the super-EGO. The key of the proposed method is the constraint parts, which are the sum of the boundary probability and the infill probability considered the uncertainty of a constraint function. When the surrogate model is inaccurate in an early stage, the infill probability will be a more dominant factor than the boundary. Thus, a new sample point is located in the feasible region. As the sample points are

added and the feasible region is approximated accurately, the inaccuracy of the surrogate model represented by the MSE is reduced. And the boundary probability becomes a more dominant factor than the infill probability. Thus, a new sample point moves to the boundary of feasible region. In other words, the proposed method performs infill sampling with space filling of feasible region in early stage of optimization. As the number of points increases, it selects a point along the boundary where the optimum point is commonly located.

The procedure of the proposed method is as follows:

Step 1. Do initial data set, **X**, and obtain (**Y**,**G**) with *p* points.

Step 2. Repeat steps $3\neg 5$ until # of data set exceeds limits.

Step 3. Fit the kriging models with the data (**X**,**Y**), (**X**,**G**).

Step 4. Find the point **P** to maximize Eq. (5).

 $Step 5$. $X \cup P \rightarrow X$, $Y \cup y(P) \rightarrow Y$, $G \cup g(P) \rightarrow G$.

Finally, minimum value of **Y**, while satisfying $\mathbf{G} < 0$, is optimum value, y^* , and **x** is the optimum solution at $y(x) = y^*$.

4. Numerical examples

To investigate the performance of the proposed method, three different global optimization problems are illustrated. For better visualization, two-variable problems are selected with a nonlinear constraint and discrete feasible regions. The seven-variable problem, Hock-Schittkowski example that has nonlinear constraint and objective function, is also solved. In this paper, the average number of function calls and determinant of correlation matrix are used to check efficiency of optimization and stability of kriging surrogate model. If the correlation matrix of the kriging surrogate model is singular, the numerical error to predict untried points is largely generated. Thus, the determinant value of the correlation matrix is needed to remain higher for stable predicting. For duplication, initial sample sets are randomly chosen by optimal Latin-hypercube design (OLHD). Genetic algorithm (GA) and super-EGO are used for comparison. Once the error of design variables becomes below 5%, algorithms are terminated and the number of function calls is calculated. The relative error at the optimum is as follows: the problem, Hock-Schittkowski example that has

constraint and objective function, is also solved. In

the average number of function calls and determi-

constraint and objective function, is also solved. In

the average ear constraint and discrete feasible regions. The

problem, Hock-Schittkowski example that has

straint and objective function, is also solved. In

a average number of function calls and determi-

a average number of func

$$
error_{opt} = \left\| \frac{\mathbf{x}_{opt} - \mathbf{x}}{\mathbf{x}_{opt}} \right\| \times 100, \tag{6}
$$

where \mathbf{x}_{out} is the point of true solution and **x** is a current point obtained by GA, super-EGO or the proposed method.

4.1 Example 1: Six-hump camelback example

The objective function is a quadratic function in Fig. 3(a). The constraint function, six-hump camelback is highly nonlinear as shown in Fig. 3(b), and it causes disconnected feasible regions as shown in Fig. 3(c). The global optimum is at the inside of a much smaller feasible region than other feasible regions. The optimum is $\mathbf{x}_{opt} = [1.815, -0.875], f(\mathbf{x}_{opt}) = -2.950$.

Table 1. Comparison of the proposed method and other optimizer for example 1.

Global optimizer	Ave. $#$ of function calls (10 runs)	Determinant of correlation matrix
GA	578.4	
Super-EGO	28.6	5.2616e-037
Proposed method	314	5.0925e-018

Fig. 3. Example 1: (a) objective function; (b) constraint function; (c) contour plot of the problem and optimum point; (d) scatter plot of 31 sampled points and contour plot of kriging predictors.

Min
$$
f(\mathbf{x}) = (x_1 - 2)^2 + (x_2 + 1)^2 - 3
$$

s.t. $g(\mathbf{x}) = (4 - 2.1x_1^2 + \frac{x_1^4}{3})x_1^2 + x_1x_2 + (-4 + 4x_2^2)x_2^2 \le 0$ (7)
 $x_1 \in [-2, 2], x_1 \in [-1, 1].$

The proposed method obtains a global optimum in small feasible region at 31 sampled points. The (d) of Fig. 3 is a scatter plot and contour plot of kriging predictors at 31 sampled points. At the beginning of optimization in bottom figures of Fig. 4, the proposed method selects the point in one of the large feasible regions due to the infill probability part. Because the infill probability part in the proposed method can consider model uncertainty, new points, red square points in solid circle are located on the interior of the feasible region during two iterations. After approximating the inner region, new sample points, red square points in dotted circles are located along the boundary of the feasible region. As the feasible region is approximated accurately, the boundary sampling criterion is more dominant because a point near the feasible region has a lower MSE value. Thus, when a feasible region is accurately approximated, the proposed method can search improved boundary point. Finally, the proposed method can find a global optimum in very small disconnected feasible

Fig. 4. Convergence history of example 1 (Red square point: last sample point, black square points: previous sample points).

region with only 31 iterations.

In this problem, super-EGO needs 29 function calls for finding the global optimum as listed in Table 1. It is more efficient than the proposed method. However, in the case of super-EGO, the correlation matrix in kriging surrogate model becomes singular because the local search phase makes many sample points extremely clustered. On the contrary, the proposed method keeps the distance between the sample points and a new point due to the MSE. Thus, it can alleviate the singularity of the correlation matrix and keep the accuracy in the large sampling points.

4.2 Example 2: Modified haupt and Schewefel example

The second example has a highly nonlinear objective function and a constraint function as shown in Figs. 5(a) and (b), respectively. It has three global optima and two global optima on the boundary of the feasible regions and the other exists inside the feasible region. In Fig. 5(c), the three global optima have identical function values of $f(\mathbf{x}_{opt}) = -1.365$ at $\mathbf{x}_{opt} =$ [6.086, 6.086], [-6.086, 6.086], [6.086, -6.086]. sampling criterion consi

ampling criterion consistes

as three global optima and two global optima

or of the feasible r example has a highly nonlinear objective tunc-
straint function as shown in Figs. 5(a) and (b),
has three global optima and two global optima
inv of the feasible regions and the other exists
ible region. In Fig. 5(c), the singular because the local search phase makes many

onions extremely clustered. On the contrary, the pro-

orions external control at the problem at the state of the control of the problem at ω is the control of the pr comes singular because the local search phase makes many

(e)

ompte points extremely clustered. On the contrary, the pro-

red method keeps the distance between the sample points

and a new point due to the MSE. Thus, it For the signal particle is university to the problem in the signal particle and method keeps the distance between the sample points contour plot of the problem & optims. (d) scatter plot at a new point due to the MSE. Thu gular because the local search phase makes many
 x is extremely clustered. On the contrary, the pro-
 x is extremely clustered. On the contrary, the pro-
 x deeps the distance between the sample points

on detail an of deceps the distance between the sample points

conour plot of the problem & optima; (d)

booint due to the MSE. Thus, it can alleviate the

points and conour plot of Kriging predictors

plug in just 41 iterations. The

Min
$$
f(\mathbf{x}) = -x_1 \sin(\frac{x_1}{3}) - 1.5x_2 \sin(\frac{x_2}{3})
$$

s.t. $g(\mathbf{x}) = -x_1 \sin(\sqrt{|x_1|}) - x_2 \sin(\sqrt{|x_2|}) \le 0$
 $x_1, x_2 \in [-15, 15].$ (8)

Upon the 23th iteration, the proposed method can find a lower right optimum. An upper right optimum inside of a feasible region is explored within 34 iterations. The last one, an upper left optimum on the boundary of constraint, is sam-

Table 2. Comparison of proposed method and other optimizer for example 2.

Global optimizer	Ave. $#$ of function calls (10 runs)	Determinant of correlation matrix
GA	440.6	None
Super-EGO	453	4.8477e-056
Proposed method	41 I	1.3996e-032

Fig. 5. Example 2: (a) objective function; (b) constraint function; (c) contour plot of the problem & optima; (d) scatter plot of 41 sampled points and contour plot of kriging predictors.

pled in just 41 iterations. The proposed method efficiently obtains global optima in the feasible region as well as along the boundary of constraints. It shows that the ratio of the infill sampling criterion considering model uncertainty and the constant boundary criterion are complementary in constraint parts.

end keeps the distance between the sample points

when the correlation particle the state of the rooten at order at a point of the problem at order plot of 41 is

when the correlation matrix and keep the accuracy in

yof The result provides evidence to support the claim that the proposed method can provide several design candidates efficiently over the entire design space. As listed in Table 2, the result of the proposed method is more efficient than the result of super-EGO. Also, the determinant of the correlation matrix of super-EGO is much smaller than that of the proposed method.

4.3 Example 3: Hock-Schittkowski problem 100

The Hock-Schittkowski problem 100 (HS100) is a test problem involving seven variables, one objective and four constraints. For this analysis, only the objective function and one of the constraints are used as response function in Eq. (9). A global optimum had function value of $f(\mathbf{x}_{opt}) = 678.68$ at \mathbf{x}_{opt} = [2.96, 1.92, -0.67, 4.26, -0.63, 1.13, 1.46] on the active boundary line of constraint.

Table 3. Comparison of proposed method and other optimizer for example 3.

example 3.	Ave. $#$ of function calls	Determinant of	by using the tota history plot of a
Global optimizer	(10 runs)	correlation matrix	is one of the rea
GA	2331.7	None	mization. An id
Super-EGO	Not converged	0 (Less than 1e-320)	mate of the mo
Proposed method	361.4	2.0874e-172	global optimum combination of
	Min $f(\mathbf{x}) = (x_1 - 10)^2 + 5(x_2 - 12)^2 + x_3^4 + 3(x_4 - 11)^2$ $+10x_5^6+7x_6^2+x_7^4-4x_6x_7-10x_6-8x_7$ s.t. $g(x) = 127 - 2x_1^2 - 3x_2^4 - x_3 - 4x_4^2 - 5x_5$	(9)	uncertainty and was proposed. T proposed metho extensively exan
	$x_i \in [-5, 5], i = 1, 2, \dots, 7.$		Acknowledger
			This study wa

Min
$$
f(\mathbf{x}) = (x_1 - 10)^2 + 5(x_2 - 12)^2 + x_3^4 + 3(x_4 - 11)^2
$$

\n+ 10x₅⁶ + 7x₆² + x₇⁴ - 4x₆x₇ - 10x₆ - 8x₇
\ns.t. $g(\mathbf{x}) = 127 - 2x_1^2 - 3x_2^4 - x_3 - 4x_4^2 - 5x_5$
\n $x_i \in [-5,5], i = 1,2,...,7$.
\n**Acknowledgement**
\n**Acknowledgement**

The results of each algorithm are listed in Table 3. Because the local search phase is continued until certain three design points are within an extremely small distance, super-EGO performs the same as usual global optimization, and it finally cannot obtain the optimum within the maximum number of function calls (1000). GA can obtain the optimum, but it needs many sample points. On the other hand, the proposed method is superior to GA with respect to the average number of function calls.

5. Conclusions and future work

We have proposed a new stochastic sampling criterion based on statistical surrogate model to resolve the several difficulties of earlier studies as well as to effectively find the global optimum of problems with nonlinear constraints. Primarily, the proposed sampling criterion consists of an objective part called as EI probability and constraint parts. The key element of the method is the constraint part, which is a combination of the infilling criterion, the boundary sampling criterion, and MSE. The MSE determines which criterion is more dominant between the infill sampling criterion and the boundary sampling criterion.

For numerical examples with disconnected feasible region, multiple optima and/or multivariable, the proposed method obtains a global optimum/s accurately as well as efficiently. In addition, due to use space filling concept, the proposed method can alleviate singularity of correlation matrix and keep the accuracy of surrogate model compared to super-EGO. Thus, the result of the proposed method in multivariate problem is much superior to that of GA and super-EGO.

In this study, only one numerical problem with seven design variables was employed. For further study, it is necessary to handle various multivariate problems. Also, in the case of a simple example with single global optimum, the proposed method was slower in finding the optimal solution compared to the super-EGO because the EI probability has the property of step function with 0 or 1, over the design domain. Speeding up the proposed method is therefore a matter of further re-

S. *Cho et al. / Journal of Mechanical Science and Technology 29 (4) (2015) 1421-1427*

(
ison of proposed method and other optimizer for
 $\frac{1}{2}$ between the contribution of the control of the control of the control of 3. Cho et al. *Journal of Mechanical Science and Technology 29 (4)* (2015) 1421-1427

Son of proposed method and other optimizer for
 $x^2 + 3x + 6$ function calls
 $x^3 + 3x + 3x + 3x + 3(x + 1)^2$

Alterminant of
 $x^4 + 3x + 3x +$ of proposed method and other optimizer for

by using the total number of function

Ave. # of function calls

(10 runs)

2331.7 None

Not converged

16 1.4 2.0874e-172

20874e-172

20874e-172

20874e-172

20874e-172

20874 by using the total number of f

(10 runs)

(10 runs)

2331.7

Note correlation matrix

2331.7

Note correlation matrix

2.0874e-172

The mization. An ideal stopping

mate of the model accuracy

361.4

2.0874e-172

2.0874e S. Cho et al. / Journal of Mechanical Science and Technology 29 (4) (2015) 1421-1427

Comparison of proposed method and other optimizer for

search, In this research, the optimization process is terminated

by using the t 5. *Cho et al. /Journal of Mechanical Science and Technology 29 (4) (2015) 1421-1427*

son of proposed method and other optimizer for

search. In this research, the optimization process is terminated

by using the total n S. Cho et al. /Journal of Mechanical Science and Technology 29 (4) (2015) 1421-1427

Comparison of proposed method and other optimizer for

search. In this research, the optimization process is te

by using the total numb **Example 18**
 Example 18 Solve and *L* Journal of *Mechanical Science and Technology 29 (4) (2015) 1421-1427

The 3. Comparison of proposed method and other optimizer for search, In this research, the optimization proce x x x x x x x s. Cho et al. / Journal of Mechanical Science and Technology 29 (4) (2015) 1421-1427*

le 3. Comparison of proposed method and other optimizer for

the sace, the optimization process is term

interaction calls
 $\frac{1}{2}$ *x. Cho et al. / Journal of Mechanical Science and Technology 29 (4) (2015) 1421-1427*

rison of proposed method and other optimizer for
 x is one of the column of function
 $x + 2331.7$

Note converged
 $x = 2341.7$

Mot 5. Cho et al. / Journal of Mechanical Science and Technology 29 (4) (2015) 1421-1427

parison of proposed method and other optimizer for

by using the total number of function calls and mointoining the

interaction calls S. Cho et al. /Journal of Mechanical Science and Technology 29 (4) (2015) 1421-1427

tison of proposed method and other optimizer for

the resumber of the optimization process is termin

by using the total number of funct 5. Cho et al. /Journal of Mechanical Science and Technology 29 (4) (2015) 1421-1427

parison of proposed method and other optimizer for

search. In this research, the optimization process is

the valid number of function search. In this research, the optimization process is terminated by using the total number of function calls and monitoring the history plot of an objective function. The termination criterion is one of the remaining challenges for stochastic global optimization. An ideal stopping rule would incorporate an estimate of the model accuracy and of the probability that the global optimum has been found. A sampling criterion with combination of expected improvement probability, model uncertainty and two constraint criteria based on probability was proposed. Therefore, a termination criterion based on the proposed method could be a good alternative and it should be extensively examined in further study.

This study was conducted as a part of a National Project, Development of Deep-seabed Mining Technology, sponsored by the Ministry of Oceans and Fisheries, Korea. The authors appreciate support for this research.

References

- [1] R. Storn and K. Price, Differential evolution: A simple and efficient heuristic for global optimization over continuous spaces, *J. of Global Optimization,* 11 (4) (1997) 341-359.
- [2] R. Christopher, A. Jeffery and G. K. Michael, A genetic algorithm for function optimization: A Matlab implementation, *ACM Transactions on Mathematical Software* (1996).
- [3] S. Kirkpatrick, C. D. Gelatt Jr. and M. P. Vecchi, Optimization by simulated annealing, *Science,* 200 (1983) 671-680.
- [4] D. R. Jones, The DIRECT global optimization algorithm, *Encyclopedia of Optimization*, 1 (2001) 431-440.
- [5] G. G. Wang and S. Shan, Review of metamodeling techniques in support of engineering design optimization, *J. of Mechanical Design,* 129 (4) (2007) 370-380.
- [6] C. Audet, J. E. Dennis, D. W. Moore, A. Booker and A. Frank, Surrogate-model-based method for constrained optimization, *Proceedings of the 8th AIAA/NASA/USAF/ ISSMO Symposium on Multidisciplinary Analysis and Optimization*, AIAA-2000-4891.
- [7] T. H. Lee and J. J. Jung, A sampling technique enhancing accuracy and efficiency metamodel-based RBOD: Constraint boundary sampling, *Computers & Structures*, 86 (13- 14) (2008) 1463-1476.
- [8] T. H. Lee, J. Y. Seung and J. J. Jung, Sequential feasible domain sampling of kriging metamodel by using penalty function, *Transaction of KSME*, 30 (6) (2006) 691-697.
- [9] D. R. Jones, M. Schonlau and W. J. Welch, Efficient global optimization of expensive black-box functions, *J. of Global Optimization*, 13 (4) (1998) 455-492.
- [10] M. J. Sasena, Flexibility and efficiency enhancements for constrained global design optimization with kriging approximations, *Ph. D. thesis*, University of Michigan (2002).
- [11] D. G. Krige, A statistical approach to some mine valuations and allied problems at the Witwatersrand, *Master's thesis*,

University of Witwatersrand (1951).

- [12] J. Sacks, W. J. Welch, T. J. Mitchell and H. P. Wynn, Design and analysis of computer experiments, *Statistical Sci ence*, 4 (4) (1989) 409-435.
- [13] A. G. Watson and R. J. Barnes, Infill sampling criteria to locate extremes, *Math. Geology*, 27 (5) (1995) 589-608.
- [14] T. H. Lee and J. J. Jung, Chapter 16: Kriging metamodel based optimization, *Optimization of Structural and Mechanical Systems*, J.S. Arora, World Scientific (2007) 445-486.

Su-gil Cho received Ph.D. degree in Automotive Engineering from the Hanyang University, Korea. He is currently a researcher at KRISO (Korea Research Institute of Ships& Ocean engineering), Korea. His research interests include design and analysis of computer experiments, uncertainty-based multidisciplinary

design optimization, and surrogate model based optimization.

Tae Hee Lee received Ph.D. degree at the University of Iowa in 1991 under supervision of Prof. J.S. Aroa. At various times during his career, he has held appointments at the University of Iowa in USA, Tokyo Denki University in Japan, Yeoungnam University in Korea,

and Georgia Institute of Technology in USA. He received an award for excellence in academic achievement in 2013 from Korean Society for Mechanical Engineers. His research interests include design optimization, design and analysis of computer experiments, design under uncertainty, and surrogate model based optimization.