A Robust Optimization Using the Statistics Based on Kriging Metamodel

Kwon-Hee Lee*
Department of Mechanical Engineering, Dong-A University, 604-714, Korea
Dong-Heon Kang
Graduate student, Department of Mechanical Engineering, Dong-A University, 604-714, Korea

Robust design technology has been applied to versatile engineering problems to ensure consistency in product performance. Since 1980s, the concept of robust design has been introduced to numerical optimization field, which is called the robust optimization. The robustness in the robust optimization is determined by a measure of insensitiveness with respect to the variation of a response. However, there are significant difficulties associated with the calculation of variations represented as its mean and variance. To overcome the current limitation, this research presents an implementation of the approximate statistical moment method based on kriging metamodel. Two sampling methods are simultaneously utilized to obtain the sequential surrogate model of a response. The statistics such as mean and variance are obtained based on the reliable kriging model and the second-order statistical approximation method. Then, the simulated annealing algorithm of global optimization methods is adopted to find the global robust optimum. The mathematical problem and the two-bar design problem are investigated to show the validity of the proposed method.

Key Words : Robust Design, Kriging, Uncertainties, Global Robust Optimum

1. Introduction

Robust design is an engineering methodology for optimizing the product and process conditions. The concept of robust design was pioneered by Dr. G. Taguchi in the late 1940s, and he had a big effect on quality engineering in the 1980s and 1990s. Since 1980s, his technique has been applied to numerical optimization, complementing the deficiencies of deterministic optimization. This newly developed method is often called robust optimization, and it overcomes the limitation of deterministic optimization that neglects the effect of uncertainties in design variables and/or design parameters (Fowlkes and Creveling, 1995 ; Lee and Park, 2005).

The robustness can be classified into two categories by taking into account the conventional (or deterministic) optimization formulation. One is related to the objective function, while the other is related to the constraint function. The robustness of objective function is determined by a measure of its variation. On the contrary, the robustness of the constraint function is defined by the feasibility condition which indicates that the optimum always lies in the feasible region (Parkinson, 1995 ; Lee and Park, 2001). The uncertainties, that can be the tolerances of design variables and/or the variations in design parameters, often induce severe variations in the objective and constraint functions. It is no less dubious to relate

* Corresponding Author,
E-mail : leekh@donga.ac.kr
TEL : +82-51-200-7638; **FAX** : +82-51-200-7656
Department of Mechanical Engineering, Dong-A University, 604-714, Korea. (Manuscript **Received** November 7, 2005; **Revised** May 18, 2006)

that the robust design concept is essential to that kind of design problem.

To consider the robustness, statistics such as mean and variance (or standard deviation) of a response should be calculated in the robust optimization process. Hereafter, "statistics" imply mean and variance. Due to the computational burden associated with the calculations of true statistics, most researches (Parkinson, 1995 ; Lee and Park, 2001 ; Lee and Jung, 2002 ; Doltsinis and Kang, 2004 ; Han and Kwak, 2004 ; Gumbert et al., 2005) in robust optimization suggest approximate statistics or a newly defined robustness index in lieu of true statistics. However, existing approximate statistics are determined from the first-order statistical approximation method, so its use is limited in gradient-based optimization requiring second-order derivatives. The use of robustness index also does not provide exact statistics. Furthermore, the previous researches did not dare to adopt an algorithm for global optimization in searching robust optimum since they utilized the true function in order to evaluate the approximate statistics or the newly defined robustness index. An implementation of robust design adopting the gradient-based optimization algorithm leads to finding a local robust optimum. However, multiple local robust optima can exist even in a uni-modal or non-monotonic function.

This research implements a methodology for robust optimization overcoming the current limitations. Two main approaches are developed. The first approach is to make the reliable kriging model and the approximate statistics for a response. The second approach is to solve the robust optimization by applying the simulated annealing algorithm in order to find the global robust optimum.

Robust optimization can be formulated by capturing the design characteristics. However, regardless of the design characteristics, its objective and constraint functions are composed of their statistics. When the statistics of responses are calculated based on their kriging models, all of the functions in the robust optimization can be expressed in mathematical forms, which leads to a simple optimization problem. The critical component of the procedure is to make a reliable kriging model in order to replace the true response function. To gain more accurate kriging model, two sequential sampling methods are applied. One is to select a new sample point by maximizing MSE (mean square error) of initial kriging model, while the other is to select the new sample point as a stationary point. If the stationary point exists in the response function, the added point determined by the latter method may be one of the local robust optimum.

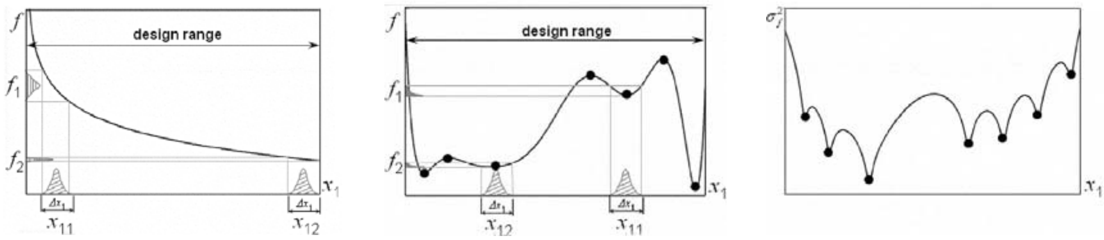
In the formulation for robust optimization, the statistics of responses are represented by the second-order statistical approximation method. However, the first-order statistical approximation method can be confused when finding the robustness. Once this is accomplished based on kriging metamodels, a global optimization method such as tabu search method, simulated annealing algorithm or genetic algorithm can be employed to solve the design formulation. In this research, the simulated algorithm is adopted. In the course of calculating the global robust optimum, the computational cost is very low since all the true functions composing robust optimization formulation are replaced by simple mathematical expressions.

The mathematical problem (Leary, 2004 ; Lee and Jung, 2005 ; Lee and Park, 2005) and the two-bar design problem (Jin et al., 2003) are investigated to show the validity of the proposed method. In the two-bar design problem, not only the robustness of objective function but also the robustness of constraint function is investigated to consider the constraint feasibility. For both of examples, the approximate statistics at the calculated global robust optimum are compared with those generated by the Monte Carlo simulations.

2. Global Robust Optimum and Kriging Model

2.1 Global robust optimum (Lee and Park, 2005)

The variations of responses are generated from the uncertainties in the design variables and/or the design parameters. The purpose of global robust optimization is to find the design with target



(a) Monotonic function (b) Multi-modal function (c) Variance of multi-modal function

Fig. 1 Concept of global robust optimum

response and smallest variation. The result shall be called the global robust optimum hereafter. To explain the global robust optimum, the monotonic and multi-modal functions are shown in Fig. 1(a) and (b), respectively. Suppose that it is a minimization problem of $f(x_1)$ with random variable x_1 normally distributed, then it can be seen in two functions that the value x_{12} is more robust than value x_{11} since the distribution of $f(x_{12})$ is smaller than that of $f(x_{11})$. Thus, the design point x_{12} is considered better for an insensitive design. As depicted in Fig. 1(a), there is only one local robust optimum in the design range. On the contrary, seven local robust optima in the design range are marked with solid points in Fig. 1(b). The seven local robust optima are also marked in Fig. 1(c), which represents the variance of $f(x_1)$. Furthermore, a monotonic function or a uni-modal function may have more than one local robust optimum. Thus, it is desirable to adopt a global optimization method to obtain the global robust optimum.

As illustrated in Figs. 1(b) and (c), a robust optimum with the smallest variance in design range is called the global robust optimum. However, the global robust optimum should be obtained around the target value that is determined by design characteristics. Thus, the formulation for global optimization is represented by the means and variances of responses.

2.2 Kriging model

Kriging method for an approximation model is well described in references (Sacks et al., 1989; Guinta and Watson, 1998; Sakata et al., 2003; Santner et al., 2003; Lee and Jung, 2005; Lee, 2005). In the kriging model, the global approximation model is represented as

$$f(\mathbf{x}) = \beta + z(\mathbf{x}) \tag{1}$$

where $\mathbf{x}=\mathbf{b}$ or $\mathbf{x}=[\mathbf{b}^T\ \mathbf{p}^T]^T$, β is a constant, and $z(\mathbf{x})$ is the realization of a stochastic process with mean zero and variance s^2, following the Gaussian distribution. In this paper, $\mathbf{b}$ means the design variable vector, and $\mathbf{p}$ means the design parameter vector. Let the number of components in $\mathbf{b}$ n, and the number of components in $\mathbf{p}$ o. Thus, the number of components in $\mathbf{x}$ is $m=n+o$. Let $\hat{f}(\mathbf{x})$ be an approximation model. Hereafter, $\hat{}$ means the estimator. When the mean squared error between $f(\mathbf{x})$ and $\hat{f}(\mathbf{x})$ is minimized, the $\hat{f}(\mathbf{x})$ becomes

$$\hat{f}(\mathbf{x}) = \hat{\beta} + \mathbf{r}^T(\mathbf{x})\mathbf{R}^{-1}(\mathbf{f} - \hat{\beta}\mathbf{q}) \tag{2}$$

where $\mathbf{R}^{-1}$ is the inverse of correlation matrix $\mathbf{R}$, $\mathbf{r}$ is the correlation vector, $\mathbf{f}$ is the observed data with n_s sample data, and $\mathbf{q}$ is the vector with n_s components of 1. The correlation matrix is defined as

$$R(\mathbf{x}^j, \mathbf{x}^k) = Exp\left[-\sum_{i=1}^{m} \theta_i |x_i^j - x_i^k|^2\right], \quad (j=1, \cdots, n_s,\ k=1, \cdots, n_s) \tag{3}$$

where θ_1 is the i-th parameter corresponding to i-th variable.

By differentiating log-likelihood function with respect to β and s^2, respectively, and letting them be equal to 0, the maximum likelihood estimators of β and s^2 are determined as Eqs. (4) and (5).

$$\hat{\beta} = (\mathbf{q}^T\mathbf{R}^{-1}\mathbf{q})^{-1}\mathbf{q}^T\mathbf{R}^{-1}\mathbf{f} \tag{4}$$

$$\hat{s}^2 = \frac{(\mathbf{f} - \hat{\beta}\mathbf{q})^T\mathbf{R}^{-1}(\mathbf{f} - \hat{\beta}\mathbf{q})}{n_s} \tag{5}$$

Similarly to previous estimators, the unknown parameters of θ_1, θ_2, $\cdots$, θ_m are calculated from the formulation as follows

$$maximize - \frac{[n_s \ln(\hat{s}^2) + \ln|\mathbf{R}|]}{2} \tag{6}$$

where θ_i $(i=1, 2, \cdots, m) > 0$. In this study, the method of modified feasible direction is utilized to determine the optimum parameters. The mean squared error of the predictor is derived as Eq. (7).

$$\hat{t}^2 = \hat{s}^2 \left[1 - \mathbf{r}^T \mathbf{R}^{-1} \mathbf{r} + \left[\frac{(1 - \mathbf{1}^T \mathbf{R}^{-1} \mathbf{r})^2}{\mathbf{1}^T \mathbf{R}^{-1} \mathbf{1}} \right] \right] \tag{7}$$

2.3 Sequential sampling

The reliability of a metamodel depends on the sampling strategy and the number of sample points. When the sampling strategy and the number of sampling points are fixed, the sequential sampling approach can be utilized to improve the previous metamodel. Despite the steady growth of hardware, a single evaluation of response can take minutes to hours. For such a design problem, the accuracy of approximation model can be improved efficiently by sequential sampling approach. Every sequential sampling method has its strong and weak points, and the details are described in the references (Jin et al., 2002 ; Turner et al., 2003 ; Lee and Jung, 2005). In this research, two sampling strategies are simultaneously adopted to make a sequential kriging model.

The first strategy is to find the points with minimum magnitude of gradient vector. In non-monotonic function, the points coincide with the stationary points. As illustrated in Figs. 1(b) and (c), each stationary point is identified as one of local robust optima. If we make a prediction at the j-th sample point using kriging model, the estimator gives the j-th response of $\mathbf{f}$. That is clearly proved by Eq. (8).

$$\begin{aligned} \hat{f}(\mathbf{x}^j) &= \hat{\beta} + \mathbf{r}^T(\mathbf{x}^j)\mathbf{R}^{-1}(\mathbf{f} - \hat{\beta}\mathbf{q}) \\ &= \hat{\beta} + \mathbf{i}^T(\mathbf{f} - \hat{\beta}\mathbf{q}) = f^j \end{aligned} \tag{8}$$

where $\mathbf{i}$ is the vector with 1 in the j-th component and 0s in other components, and f^j is the j-th component of $\mathbf{f}$. It is known that the error of estimator is very low around any selected sample point in kriging model. Thus, this fact can improve the kriging model for robust optimization by including the stationary points as the sample points, though they are determined from the approximation model. The formulation to find those points is defined as

$$\begin{aligned} &Minimize \quad \|\nabla \hat{f}(\mathbf{x})\| \\ &\mathbf{x}_L \le \mathbf{x} \le \mathbf{x}_U \text{ with } \mathbf{x}_o^k (k=1, \cdots, n_k) \end{aligned} \tag{9}$$

where $\mathbf{x}_o^k$ means the initial design and n_k is the number of initial designs. In short, the formulation of Eq. (9) is solved as many as n_k, using a gradient-based optimization. For the example problems, n_k is set up as 100, and the initial designs are randomly determined by the Latin hypercube design method. In this study, the method of modified feasible direction is utilized to solve Eq. (9). This process is not a computational burden since the optimization is performed by using simple kriging model.

The second strategy is to apply the MSE approach like Eq. (10).

$$\begin{aligned} &\underset{\mathbf{x}}{Maximize} \quad \hat{t}^2 = \hat{s}^2 \left[1 - \mathbf{r}^T \mathbf{R}^{-1} \mathbf{r} + \frac{(1 - \mathbf{q}^T \mathbf{R}^{-1} \mathbf{r})^2}{\mathbf{q}^T \mathbf{R}^{-1} \mathbf{q}} \right] \\ &\mathbf{x}_L \le \mathbf{x} \le \mathbf{x}_U \end{aligned} \tag{10}$$

The formulation of Eq. (10) (Sacks et al., 1989 ; Jin et al., 2002) is solved by using simulated annealing algorithm. From this formulation, one point is added to the existing sample points.

For the conventional optimization, the reliability of kriging model can be increased when any point near the optimum is selected as the sample point. On the contrary, for the robust optimization problem, it is important to maintain the reliability of kriging model over the entire design space since the formulation of robust optimization is defined as the statistics of response functions. Thus, it is recommended that the more sample points than those of conventional optimization be selected to make a prediction model.

2.4 Approximate statistics

Regardless of the design characteristics, the objective and constraint functions in robust optimization are composed of the statistics of responses. If all the functions in design formulation can be expressed as simple mathematical forms, it is not hard to apply any optimization algorithm to find the robust optimum. Thus, the introduction of kriging model and approximate statistics can facilitate the robust optimization.

The mean and variance of a response function are regarded as the statistics to measure its robustness. The mean μ_f called the first statistical moment and the variance σ_f^2 called the second statistical moment of function f are represented as

$$\begin{aligned}\mu_f &= E[f(\mathbf{b}+\mathbf{z}^b, \mathbf{p}+\mathbf{z}^p)] \\ &= \iint\cdots\int f(\mathbf{b}+\mathbf{z}^b, \mathbf{p}+\mathbf{z}^p)\, u_1(z_1^b)\cdots \\ &\quad u_n(z_n^b)\, v_1(z_1^p)\cdots v_o(z_o^p)\, dz_i^b\cdots dz_n^b dz_1^p\cdots dz_o^p\end{aligned} \tag{11}$$

$$\begin{aligned}\sigma_f^2 &= E[f(\mathbf{b}+\mathbf{z}^b, \mathbf{p}+\mathbf{z}^p)-\mu_f]^2 \\ &= \iint\cdots\int \{f(\mathbf{b}+\mathbf{z}^b, \mathbf{p}+\mathbf{z}^p)-\mu_f\}^2 u_1(z_1^b)\cdots \\ &\quad u_n(z_n^b)\, v_1(z_1^p)\cdots v_o(z_o^p)\, dz_1^b\cdots dz_n^b dz_1^p\cdots dz_o^p\end{aligned} \tag{12}$$

where $\mathbf{z}^{\mathrm{b}}$ and $\mathbf{z}^{\mathrm{p}}$ represent the noises of design variables and design parameters, and $u_{\mathrm{i}}(z_{\mathrm{i}}^{\mathrm{b}})$ and $v_{\mathrm{i}}(z_{\mathrm{i}}^{\mathrm{p}})$ are the probability density functions of noise factors $z_{\mathrm{i}}^{\mathrm{b}}$ and $z_{\mathrm{i}}^{\mathrm{p}}$, respectively. The noises, $\mathbf{z}^{\mathrm{b}}$ and $\mathbf{z}^{\mathrm{p}}$, induce the variation of a response, f (Lee and Park, 2005; Park et al., 2006). The integrations represented as Eqs. (11) ~ (12) are very expensive or sometimes impossible to perform. Thus, most researches have utilized the approximate statistics using the first-order statistical approximation method, overcoming the time consuming calculations. That is,

$$\mu_f \cong f(\mathbf{x})_{\bar{\mathrm{x}}} \tag{13}$$

$$\sigma_f^2 \cong \sum_{i=1}^{m}\left(\frac{\partial f}{\partial x_i}\right)_{\bar{\mathrm{x}}}^2 \sigma_{x_i}^2 \tag{14}$$

where $\sigma_{x_i}^2$ represents the variances of i-th variable x_{i}, and $\bar{\mathbf{x}}$ represents the mean vector of variables, respectively. However, the use of Eqs. (13) and (14) in the highly nonlinear function with large input deviations can be confused when measuring the robustness. Figs. 1(b) and (c) illustrate that the global robust optimum is x_{12}. When the Eq. (14) is adopted to obtain σ_f^2, any point marked with a solid circle has the same variance as zero, which gives rise to an unpredictable result. On the contrary, for a monotonic function with small input deviations, Eqs. (13) and (14) can be utilized as an index of robustness even though it cannot give the exact variance. Thus, Eqs. (13) and (14) are useful in replacing the true statistics only for the monotonic function as in Fig. 1(a) but in general, it is difficult to determine whether a function is monotonic or not in design range.

To overcome these difficulties, the approximate statistics using the second-order statistical approximation method are adopted as

$$\mu_f \cong f(\mathbf{x})_{\bar{\mathrm{x}}} + \frac{1}{2}\sum_{i=1}^{m}\left(\frac{\partial^2 f}{\partial x_i^2}\right)_{\bar{\mathrm{x}}} \sigma_{x_i}^2 \tag{15}$$

$$\sigma_f^2 \cong \sum_{i=1}^{m}\left(\frac{\partial f}{\partial x_i}\right)_{\bar{\mathrm{x}}}^2 \sigma_{x_j}^2 + \frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{m}\left(\frac{\partial^2 f}{\partial x_i \partial x_j}\right)_{\bar{\mathrm{x}}}^2 (\sigma_{x_j}^2)(\sigma_{x_j}^2) \tag{16}$$

One might ask how difficult it is to calculate these statistics. As you know, the direct use of Eqs. (15)-(16) may fail to determine the robust optimum since it is very expensive to obtain the second order terms of true functions. Furthermore, if the gradient-based optimization algorithm is selected as the optimizer to solve the formulation, that requires second derivatives for Eqs. (13) and (14), and third derivatives for (15) and (16). In practical application of complex design, we cannot accept the calculation of more than first-order derivative using finite difference method in optimization process. Thus, it is not desirable to introduce the gradient-based optimization algorithm and the true function in solving the robust optimization.

Let us replace the true function f with the kriging model $\hat{f}$ in Eqs. (15) and (16). Then, the first and second derivatives with respect to x_{i} are analytically determined as

$$\frac{\partial \hat{f}(\mathbf{x})}{\partial x_i} = \frac{\partial \mathbf{r}^T(\mathbf{x})}{\partial x_i}\mathbf{R}^{-1}(\mathbf{f}-\hat{\beta}\mathbf{q}),\ (i=1, \cdots, m) \tag{17}$$

$$\frac{\partial^2 \hat{f}(\mathbf{x})}{\partial x_i \partial x_j} = \frac{\partial^2 \mathbf{r}^T(\mathbf{x})}{\partial x_i \partial x_j}\mathbf{R}^{-1}(\mathbf{f}-\hat{\beta}\mathbf{q}),\ (i,\ j=1, \cdots, m) \tag{18}$$

where

$$\frac{\partial \mathbf{r}^T(\mathbf{x})}{\partial x_i}=[-2\theta_i(x_i-x_i^1)A^1,\ -2\theta_i(x_i-x_i^2)A^2,\ \cdots,\ -2\theta_i(x_i-x_i^{ns})A^{ns}] \tag{19}$$

$$\frac{\partial^2 \mathbf{r}^T(\mathbf{x})}{\partial x_i \partial x_j}=[(-2\theta_i(x_i-x_i^1))(-2\theta_j(x_j-x_j^1))A^1,\ (-2\theta_i(x_i-x_i^2))(-2\theta_j(x_j-x_j^2))A^2,\ \cdots,\ (-2\theta_i(x_i-x_i^{ns}))(-2\theta_j(x_j-x_j^{ns}))A^{ns}]\ (i\neq j) \tag{20}$$

$$\frac{\partial^2 \mathbf{r}^T(\mathbf{x})}{\partial x_i^2}=[-2\theta_i A^1(1+2\theta_i(x_i-x_i^1)^2),\ -2\theta_i A^2(1+2\theta_i(x_i-x_i^2)^2),\ \cdots,\ -2\theta_i A^{ns}(1+2\theta_i(x_i-x_i^{ns})^2)]\ (i=j) \tag{21}$$

$$A^k=Exp[-\theta_1((x_1-x_1^k)^2)-\theta_2((x_2-x_2^k)^2)\cdots-\theta_m((x_m-x_m^k)^2)] \tag{22}$$

Substituting Eqs. (17) ~ (22) into Eqs. (15) ~ (16), the statistics such as mean and variance are simply and mathematically defined, facilitating global optimization as well as gradient-based optimization.

3. A Strategy for Robust Optimization

A strategy for robust optimization is proposed by using sequential kriging model and simulated annealing algorithm. The steps of the proposed algorithm are as follows :

Step 1 : Construction of primitive kriging model in lieu of true response function

First, calculate the response functions of $f(\boldsymbol{x})$ with respect to n_s sampled design points. Design of experiments strategies is often used to sample the design space. Depending on analysis time, full combination, orthogonal array or Latin hypercube design can be selected as the sampling method. For an example, the sample points are generated by Latin hypercube design that minimizes Eq. (23) (Leary et al., 2004).

$$\sum_{i=1}^{n_s}\sum_{j=i+1}^{n_s}\frac{1}{d_{ij}} \tag{23}$$

where d_{ij} refers to the distance between points i and j. To assess the kriging model, the error in surrogate model is characterized by using a few metrics defined as

$$RMSE=\sqrt{\frac{1}{n_t}\sum_{i=1}^{n_t}(f_i-\hat{f}_i)^2} \tag{24}$$

$$MAXERR=MAX[|f_i-\hat{f}_i|,\ i=1,2,\cdots,n_t] \tag{25}$$

$$Ave.\ \%\ error=\frac{1}{n_t}\sum_{i=1}^{n_t}\left|\frac{\hat{f}_i-f_i}{f_i}\right|\times 100 \tag{26}$$

$$CV=\sqrt{\frac{1}{n_s}\sum_{i=1}^{n_t}(f_i-\hat{f}_{-i})^2} \tag{27}$$

where n_t is the number of sample points for validation, and $\hat{f}_{-i}$ is the i-th estimator of kriging model constructed without the i-th observation. The metrics such as Eqs. (24)-(26) have the shortcomings that they need n_t additional function calculations. On the contrary, the metric called the cross validation defined by Eq. (27) does not require any additional function calculation. But, CV should construct the kriging models as many as n_s, which is a time consuming process. In the reference (Jones et al., 1998), this process is reduced by using the calculated $\hat{\beta}$ and θ, but by calculating $\mathbf{R}$, $\mathbf{r}$ and $\mathbf{f}$ with respect to n_s-1 sample points. However, this reduction is valid under the assumption that elimination of one sample data has a negligible effect on the maximum likelihood estimates.

Step 2 : Construction of improved kriging model using sequential sampling

The more accurate kriging model than one made by Step 1 is constructed by two sequential sampling approaches. The details are explained in Sec. 2.3. Thus, two optimization problems should be solved in Step 2, of which one is performed by modified feasible direction method and the other is performed by simulated annealing algorithm. When a new selected point is very close to any existing sample point or another new selected point, it is excluded from the sample points. To describe the distance constraint, Eqs. (28)-(29) are introduced as

$$\|\mathbf{x}^{new,i}-\mathbf{x}^j\|\geq\varepsilon',\ (i=1,\cdots,n_n,\ j=1,\cdots,n_s) \tag{28}$$

$$\|\mathbf{x}^{new,i}-\mathbf{x}_{-i}^{new,j}\|\geq\varepsilon',\ (i=1,\cdots,n_n,\ j=1,\cdots,n_n) \tag{29}$$

where n_n is the number of sample points supplemented by Eqs. (9)-(10), $\mathbf{x}^{new,i}$ is a new selected point, $\mathrm{x}_{-i}^{new,j}$ is the selected point except the

i-th point, and ε' is a small number. That is, the new selected points determined from Eqs. (9) ~ (10) should satisfy Eqs. (28) ~ (29) to be augmented to the existing sample points.

Step 3: Formulation for robust optimization

The formulation for robust optimization can be defined by the design characteristics. Nevertheless, the objective and constraint functions are described as the statistics of responses. For an example with minimization, let us consider the following formulation.

$$\begin{aligned} &\textit{Minimize} \quad \hat{\mu}_f(\mathbf{x}) + k\hat{\sigma}_f(\mathbf{x}) \\ &\textit{Subject to} \quad \hat{\mu}_{gj}(\mathbf{x}) + k\hat{\sigma}_{gj}(\mathbf{x}) \le 0, \ j=1, \cdots, Nc \\ &\qquad \mathbf{x}_L \le \mathbf{x} \le \mathbf{x}_U \end{aligned} \tag{30}$$

where N_c is the number of constraints, and $\hat{\mu}_{gj}(\mathbf{x})$ and $\hat{\sigma}_{gj}(\mathbf{x})$ are the mean and the standard deviation of j-th constraint function $\hat{g}(\mathbf{x})$ defined in conventional optimization, respectively. The constant k is introduced to consider the robustness. When the distribution of variables $\mathbf{x}$ is normal and its functions $f(\mathbf{x})$ and $g(\mathbf{x})$ can be approximated as linear with respect to $\mathbf{x}$, the distributions of $f(\mathbf{x})$ and $g(\mathbf{x})$ can be regarded as the normal distribution. Then, if $k=3$, it means that the constraint of Eq. (30) imposes the feasibility related to 99.73% of its distribution.

In view of objective function in Eq. (30), the worst case of response $f(\mathbf{x})$ is set up as the objective function. It means that the objective is to minimize the worst case of original objective function. In design process, it is not guaranteed that their distributions are always assumed to be normal even though they can be tested by Shapiro-Wilk statistic for a fixed design point. As an example for Figs. 1(b) and (c), the distribution of $f(x_1)$ at any stationary point is apparently non-normal. However, though the distribution of any function is not normal, Eq. (30) is also meaningful. At that case, it seems reasonable to assume that the functions defined by Eq. (30) are considered as the multi-objective or penalty function with weighting factor k, respectively, which does not offer severe results.

The validation of kriging model is performed at Step 1. The accuracy of statistics also depends on the accuracy of kriging model for a response. To assess the statistic more directly than Eqs. (24) ~ (27), the metric can be defined as Eq. (31) which is acceptable for the case that sensitivity information is easy to obtain.

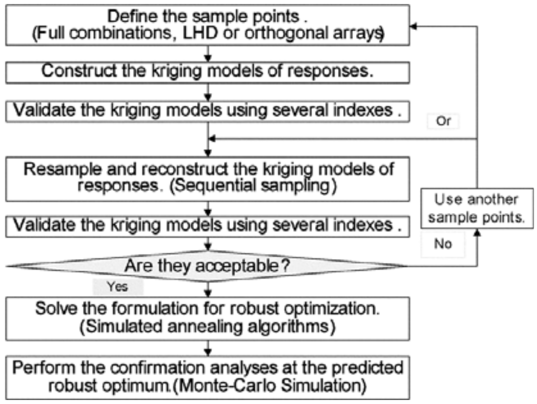


Fig. 2 Suggested design procedures

$$GI = \frac{1}{n_t}\sum_{i=1}^{n_s} \big| \|\nabla f(\mathbf{x}^i)\| - \|\nabla \hat{f}(\mathbf{x}^i)\| \big| / \|\nabla f(\mathbf{x}^i)\| \times 100 \tag{31}$$

Step 4: Determination of global robust optimum

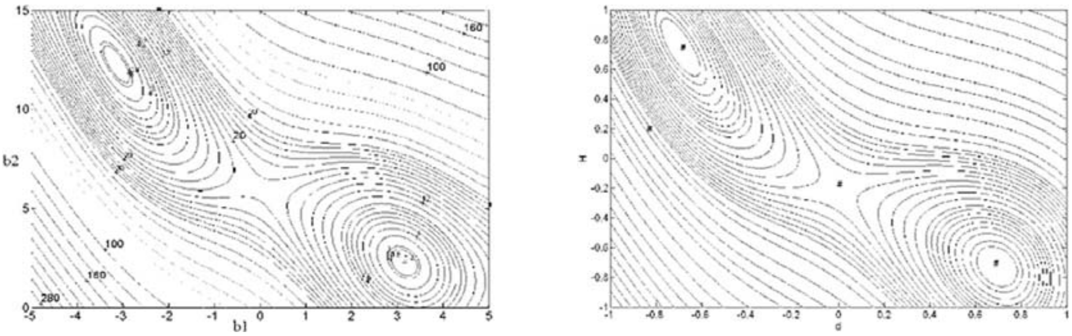
Even in a response function with only one local optimum, its variance function may have more than one local optimum. Thus, the gradient-based optimization algorithm cannot supply the global robust optimum though it is superior to any other global optimization algorithm. Fortunately, all the functions composing Eq. (30) are described as simple mathematical expressions, making the application to global optimization algorithm possible. In this research, the simulated annealing algorithm is adopted to solve the formulation.

The statistics at the predicted robust optimum determined by solving Eq. (30) can be compared with those generated by Monte Carlo simulations. The overall design process is represented in Fig. 2.

4. Example Problems

4.1 Mathematical problem

The object of this example is to determine the



(a) Original function (b) Initial kriging model of f

Fig. 3 Response function of mathematical problem

global robust optimum in the Branin's function (Leary et al., 2004 ; Lee and Jung, 2005 ; Lee and Park, 2005). It is assumed that the tolerance of x_i Δx_i (i=1, 2) is 1.0 or 0.5, its standard deviation σ_{xi} is $\Delta x_i/6$, each variable is statistically independent, random and normally distributed, and $\mathbf{p}$ does not exist. That is, $\mathbf{x}=\mathbf{b}$. The minimization problem of the Branin's function is represented as :

$$\begin{aligned} \textit{Minimize} \quad & f(x_1,x_2)=(x_2-0.12918x_1^2+1.59155x_1-6.0)^2 \\ & \quad +9.60212\cos(x_1)+10.0 \\ & (-5\leq x_1\leq 5,\ 0\leq x_2\leq 15) \end{aligned} \tag{32}$$

Its two global optima are $[-\pi\ 12.275]^{\mathrm{T}}$ and $[\pi\ 2.275]^{\mathrm{T}}$, and their function values are the same as 0.398. The contour plots of true function and initial kriging model are represented in Fig. 3. The dimension in each axis of Fig. 3(b) is scaled to $[-1\ 1]^2$, which is consistently represented in any plot for kriging model.

The initial kriging model is constructed by adopting 20 point LHD. Then, by the rule of Step 2, 4 points are supplemented to enhance the accuracy of initial kriging model. The Eq. (9) produces three points, while Eq. (10) does one point, which are marked with # in Fig. 3(b). From the looks of Fig. 3(a) or (b), it can be seen that each point generated from Eq. (9) is near the stationary point. The indexes for validation are summarized in Table 1. The kriging model using 24 point LHD is greatly improved as compared to initial kriging model, which has sufficient flexibility to fit the highly nonlinear function.

The robust optimization for the Branin's function can be formulated like Eq. (33).

Table 1 Validations of kriging models (math. Problem)

	n_s	RMSE	MAXERR	Ave.% error	CV	GI
$\hat{f}$	20	0.934	7.023	2.087	1.377	2.477
	24	0.438	3.903	0.566	0.389	1.430

$$\begin{aligned} \textit{Minimize} \quad & \hat{\mu}_f(\mathbf{x})+3\cdot\hat{\sigma}_f(\mathbf{x}) \\ & \mathbf{x}_L\leq\mathbf{x}\leq\mathbf{x}_U \end{aligned} \tag{33}$$

By applying the suggested procedures for the cases with Δx_i=1.0 and Δx_i=0.5, the global robust optima are determined to be $[3.141\ 2.283]^{\mathrm{T}}$ and $[3.141\ 2.277]^{\mathrm{T}}$, respectively. They are compared with the conventional optima and another local robust optimum in Tables 2 and 3. The another local robust optimum is obtained by minimizing $\hat{\sigma}_f(\mathbf{x})$ in lieu of $\hat{\mu}_f(\mathrm{x})+3\cdot\hat{\sigma}_f(\mathrm{x})$, which is identical to the concave stationary point of response function. In Tables 2 (Δx_i=1.0) and 3 (Δx_i=1.0), the true statistics are calculated from 50,000 Monte Calro samples, while the approximate statistics are calculated from the first-order and second-order statistical approximation methods based on the kriging model for f.

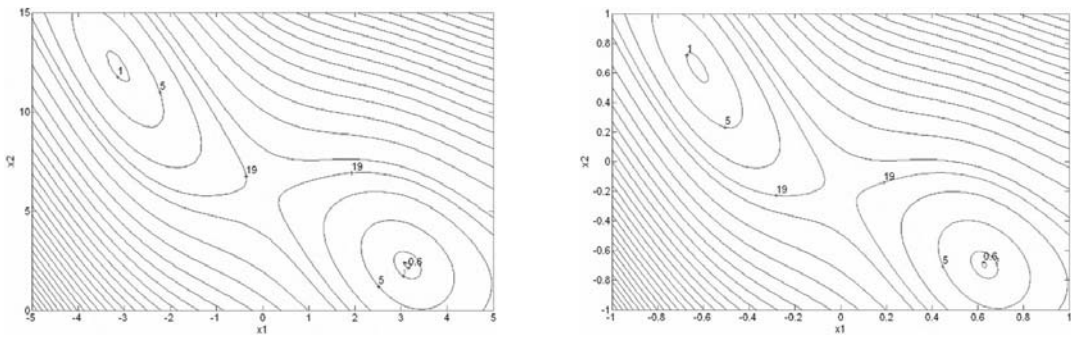
The sharp contrast between the first-order and second-order statistical approximations is enhanced by their percentage error metric. Note that the approximate statistics in Table 3 agree better with the Monte Calro simulations than those in Table 2. At this point, it should be mentioned that the accuracy of approximation method is intimately linked with standard deviations of variables and

Table 2 Statistics using the first-order and second-order statistical approximation methods ($\Delta x_i = 1.0$)

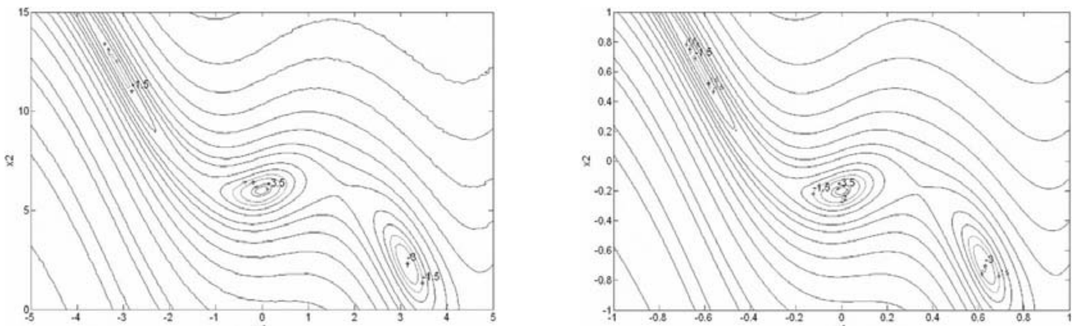
statistics	global robust opt. $[3.141\ 2.283]^T$		true local opt. of f $[\pi\ 2.272]^T$		true local opt. of f $[-\pi\ 12.272]^T$		local robust optimum $[-0.005\ 5.995]^T$	
	value	% error	value	% error	value	% error	value	% error
$\hat{\mu}_f$ (1st)	0.39775	30.650	0.39770	30.841	0.39826	44.498	19.6025	0.584
$\hat{\mu}$ (2nd)	0.57610	0.446	0.57612	0.187	0.71938	0.255	19.5676	0.405
$\hat{\sigma}$ (1st)	0.00372	98.242	0.00098	99.544	0.00223	99.500	0.0045	96.670
$\hat{\sigma}$ (2nd)	0.22083	4.408	0.22090	3.191	0.43751	1.664	0.1315	2.864
μ_f (true)	0.57354	—	0.57505	—	0.71755	—	19.4887	—
σ (true)	0.21150	—	0.21407	—	0.44492	—	0.1354	—

Table 3 Statistics using the first-order and second-order statistical approximation methods ($\Delta x_i = 0.5$)

statistics	global robust opt. $[3.141\ 2.277]^T$		true local opt. of f $[\pi\ 2.272]^T$		true local opt. of f $[-\pi\ 12.272]^T$		local robust optimum $[-0.001\ 5.999]^T$	
	value	% error	value	% error	value	% error	value	% error
$\hat{\mu}_f$ (1st)	0.39769	10.138	0.39770	10.045	0.39826	16.727	19.6025	0.048
$\hat{\mu}_f$ (2nd)	0.44229	0.061	0.44230	0.044	0.47853	0.060	19.5937	0.003
$\hat{\sigma}$ (1st)	0.00046	99.148	0.00049	99.095	0.00111	98.959	0.0005	98.328
$\hat{\sigma}$ (2nd)	0.05522	1.595	0.05523	2.405	0.10938	2.362	0.0328	0.889
μ_f (true)	0.44256	—	0.44211	—	0.47825	—	19.59307	—
σ (true)	0.05435	—	0.05393	—	0.10686	—	0.03318	—



(a) True mean (μ) of f (b) Approximate mean ($\hat{\mu}$) of f



(c) Log transformed true variance (σ_f^2) of f (d) Log transformed approximate variance ($\hat{\sigma}_f^2$) of f

Fig. 4 Mean and variance of Branin's function f

function's nonlinearity. In general, a design problem requiring robust design has large perturbation and highly nonlinear response. Thus, it is desirable to employ the second-order statistical approximation method rather than the first-order statistical approximation method for robust optimization. As mentioned in Sec. 2.4, the first-order statistical approximation method has other shortcomings that it cannot find the global robust optimum. The means and variances of true function and kriging model are represented in Fig. 4 (a), (b), (c) and (d). From the Fig. 4(a)～(d), it is clear that the contours of approximate statistics are almost identical with those of true statistics, which are calculated through 50,000 Monte Carlo samples for each of 50×50 equally spaced points.

4.2 Two-bar structure

A two-bar structure design problem for conventional optimization is defined as (Jin et al., 2003)

$$\begin{aligned} &\text{Minimize} \quad V \\ &\text{Subject to} \quad S \leq S_{max} \\ &\qquad\qquad\quad S \leq S_{crit} \qquad (34) \\ &20\text{ mm} \leq d \leq 80\text{ mm},\ 200\text{ mm} \leq H \leq 1000\text{ mm},\ t = 2.5\text{ mm} \\ &V = 2\pi dt\sqrt{B^2+H^2},\ S = \frac{P\sqrt{B^2+H^2}}{2\pi dtH},\ S_{crit} = \frac{\pi^2 E d^2}{8(B^2+H^2)} \end{aligned}$$

where V is the volume, S is the normal stress, S_{max} is 400 MPa, S_{crit} is the critical buckling stress, d is the mean diameter of thin tube, t is the thickness, $2B$ is the width, H is the height, P is the acting force, and E is Young's modulus. The two-bar structure is represented in Fig. 5. The random variables are set up as $\mathbf{x}=[d\ H\ B\ E\ P]^T$, which are separated as the design variables $\mathbf{b}=[d\ H]^T$ and the design parameters $\mathbf{p}=[B\ E\ P]^T$. The mean values of design parameters μ_B, μ_E and μ_P are described in Fig. 5. It is assumed that each variable is statistically independent, random and normally distributed, the tolerance $\Delta\mathbf{x}$ is [3.0 mm 60.0 mm 60.0 mm 30,000 MPa 30,000 N]^T, and Δx_i is $6\sigma_{xi}$.

First, to build three kriging models for volume, normal stress and critical buckling stress, the sample points should be determined. In this example, an orthogonal array OA(64,2,8,9) (Sherwood, 2005) is adopted in which the numbers in parenthesis represent number of experiments, number of strength, number of levels and number of columns. Since we treat five random variables, the last 4 columns are deleted. The kriging models for volume, stress and critical buckling stress are constructed by solving Eq. (6) three times. Secondly, additional sample points should be determined by applying Eqs. (9) and (10). However, since there are three responses, one of the three responses should be selected to apply Step 2. In this study, the volume response is arbitrarily selected, then one point by Eq. (9) and the other point by Eq. (10) are determined. The additional point by Eq. (9) is identical to one of the existing sample points. Thus, the sequential kriging model determined from 65 sample points is constructed.

From Table 4, it is seen that the accuracy of volume model using 65 sample points is greatly improved, as compared to 64 sample point case. However, the CVs for normal stress and critical buckling stress in 65 sample points are slightly worse than those in 64 sample points, which pro-

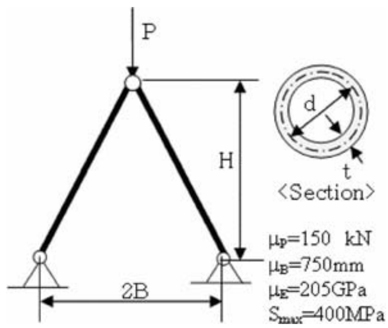
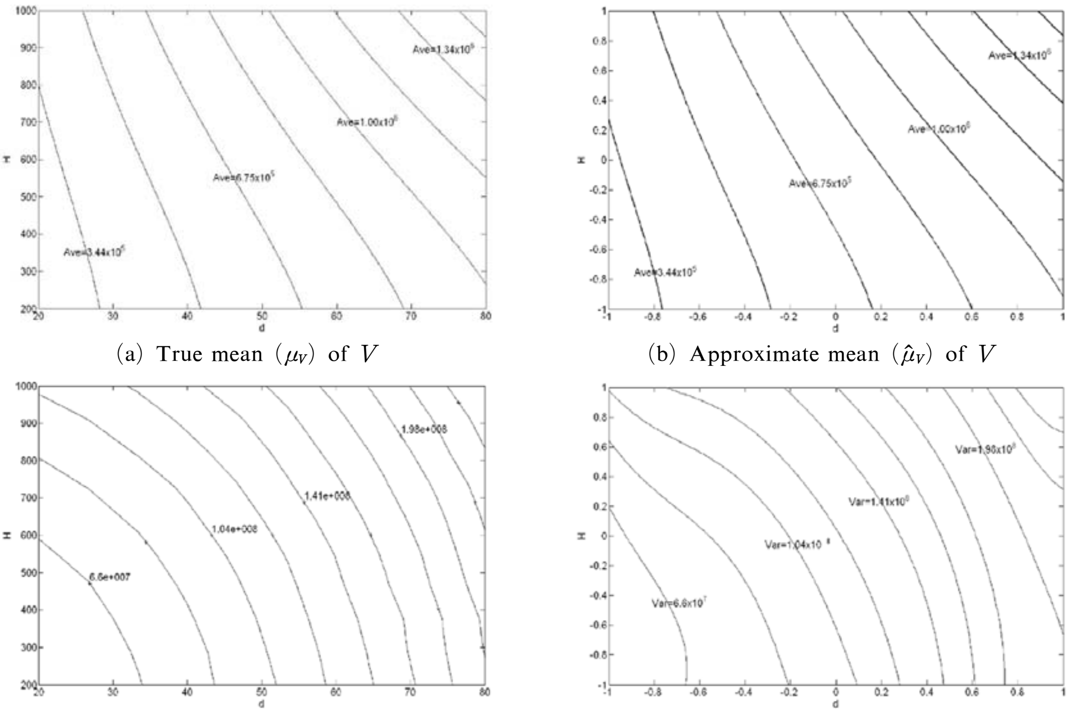


Fig. 5 Two-bar structure

Table 4 Validations of kriging models (2-bar structure)

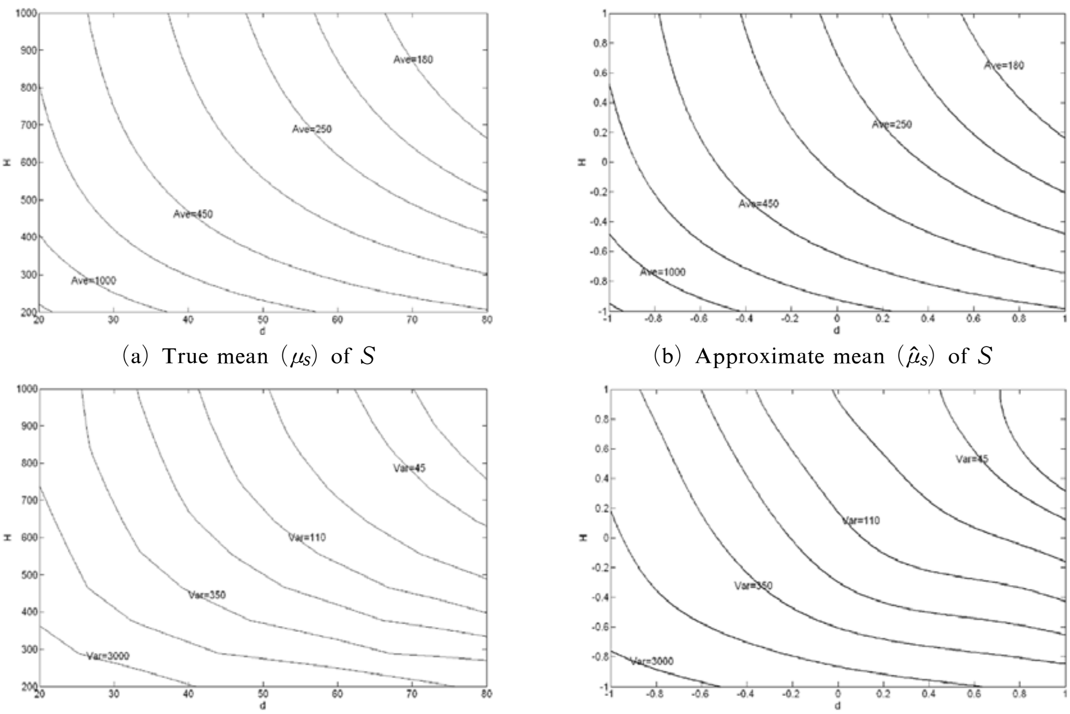
	n_s	RMSE	MAXERR	Ave.% error	CV	GI
$\hat{V}$	64	2791.6	24938.0	0.203	1806.5	0.982
	65	809.8	3159.2	0.102	308.5	0.781
$\hat{S}$	64	9.748	126.055	0.811	10.496	3.312
	65	9.092	117.160	0.660	10.631	2.634
$\hat{S}_{crit}$	64	2.963	21.720	0.403	5.645	1.180
	65	2.650	18.287	0.294	6.630	1.114



(a) True mean (μ_V) of V (b) Approximate mean ($\hat{\mu}_V$) of V

(c) True variance (σ_V^2) of V (d) Approximate variance ($\hat{\sigma}_V^2$) of V

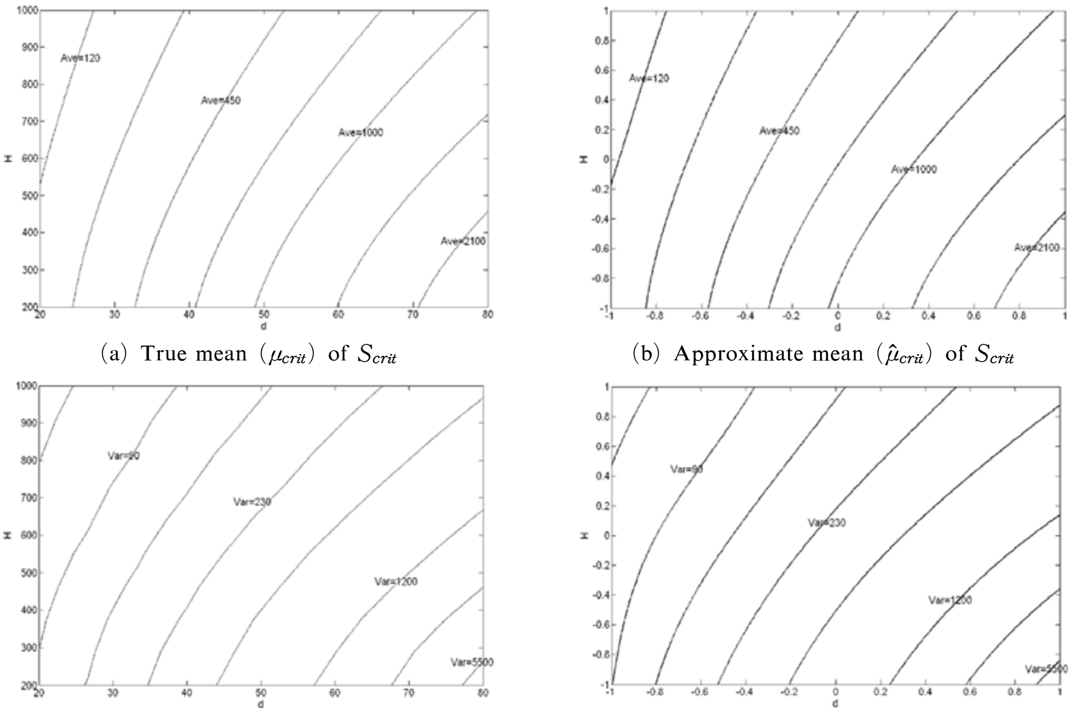
Fig. 6 Mean and variance of volume



(a) True mean (μ_S) of S (b) Approximate mean ($\hat{\mu}_S$) of S

(c) True variance (σ_S^2) of S (d) Approximate variance ($\hat{\sigma}_S^2$) of S

Fig. 7 Mean and variance of normal stress



(a) True mean (μ_{crit}) of S_{crit} (b) Approximate mean ($\hat{\mu}_{crit}$) of S_{crit}

(c) True variance (σ^2_{crit}) of S_{crit} (d) Approximate variance ($\hat{\sigma}^2_{crit}$) of S_{crit}

Fig. 8 Mean and variance of critical buckling stress

duces an effect contrary to our intention. That would be error in calculating CV. The error comes from the fact that the CV is obtained based on $\hat{\beta}$ and $\boldsymbol{\theta}$ calculated with n_s samples.

By Step 3, a robust optimization formulation using the approximate statistics based on kriging models can be defined like Eq. (35).

$$\begin{aligned}&\textit{Minimze} \quad \hat{\mu}_V + 3\hat{\sigma}_V \\ &\textit{Subject to } \hat{\mu}_S + 3\hat{\sigma}_S \le S_{\max} \\ &\qquad\qquad \hat{\mu}_S + 3\hat{\sigma}_S \le \hat{\mu}_{crit} - 3\hat{\sigma}_{crit} \\ &20\text{ mm} \le d \le 80\text{ mm},\ 200\text{ mm} \le H \le 1000\text{ mm},\ t = 2.5\text{ mm}\end{aligned} \tag{35}$$

In Eq. (35), all the statistics are calculated by using Eqs. (15)-(16). Even though the kriging models for three responses are expressed as five random variables, Eq. (35) has two variables of d and H. For each response, the contour plots of approximate statistics are represented in Figs. 6-8. These figures prove clearly that the plots of approximate statistics are very much like those of true statistics over the design space. The plots for true statistics are created over 50×50 equally spaced points with 50,000 Monte Calro samples.

Finally, after the formulation of Eq. (35) is defined, the simulated annealing algorithm is applied to find the global robust optimum, considering constraint robustness. This process does not offer severe computational burden since the functions composing Eq. (35) are expressed as simple and explicit mathematical equations. Through Step 4, the robust optimum is determined as $\mathbf{b}^* = [41.383\ 626.679]^T$, while the conventional optimum obtained from Eq. (34) is determined as $\mathbf{b}^* = [38.200\ 600.304]^T$. In Fig. 9, the conventional and robust optima and their distributions are indicated together with the objective and imposed constraint functions defined in Eq. (34). The distribution of each optimum is obtained through 5,000 Monte Carlo simulations. It is noted that a lot of samples in the conventional optimum violate the constraints, taking into account the uncertainties of random variables. On the contrary, the robust optimum has few violated samples. Table 5 summarizes the statistics of conventional and robust optima and the probability of feasibility. As Fig. 9 indicates, 51.0% of 5,000 sample

Table 5 Conventional and robust optimum (2-bar structure)

	Conventional optimum $\mathbf{b}^*=[38.200\ 600.304]^T$					Robust optimum $\mathbf{b}^*=[41.383\ 626.679]^T$				
	$\hat{\mu}$	μ	$\hat{\sigma}$	σ	Prob.	$\hat{\mu}$	μ	$\hat{\sigma}$	σ	Prob.
V	577086.378	576462.793	9644.993	9519.463	—	635863.412	635369.321	10051.177	10209.940	—
S	398.104	400.223	15.806	15.178	52.0%	358.491	360.009	13.755	13.263	99.6%
S_{crit}	401.310	400.068	16.608	15.178	51.0%	454.643	453.560	18.215	18.649	100.0%

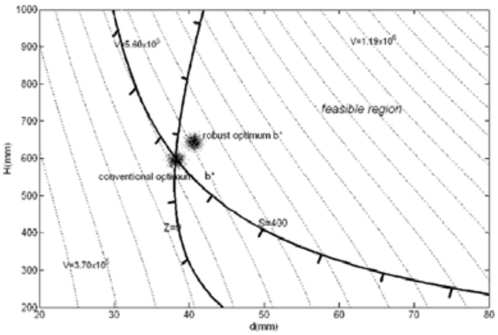


Fig. 9 Conventional optimum and robust optimum

points in the conventional optimum lie in feasible region, while 99.6% in the robust optimum lie in feasible region.

5. Conclusions

The following conclusions can be made from this study.

(1) The robust optimization is achieved through kriging approach, approximate statistics and global optimization algorithm. The kriging model using sequential sampling methods are constructed, considering the robustness of a response. The construction of the reliable kriging model facilitates the calculations of second-order approximate statistics. These approximate statistics enable one to solve the formulation for robust optimization with simulated annealing algorithm.

(2) A design problem requiring robust design has relatively large perturbation of random variable and highly nonlinear response. Through the mathematical problem with multi-modal function, we can see that the suggested optimization procedure is successfully applied to that kind of design problem. This study has shown that it is desirable to employ the second-order statistical approximation method rather than the first-order statistical approximation method for robust optimization.

(3) A robust optimization formulation has been presented. That combines robust design and approximation concepts, complementing the deficiencies of deterministic optimization. For an example, the robust optimization for two-bar structure is successfully performed. The suggested procedures can be applied to a variety of structural design for robust optimization. In terms of computational efficiency, it is seen that the suggested method can be applicable to structural design.

(4) An implementation deriving mean and variance of a response can be applied to a design methodology using stochastic approach. Especially, the results of this study point to several promising applications for reliability based analysis and design. That topic will be left as the future work.

Acknowledgments

This work was supported by the CANSMC from the Regional Innovation Center (RIC) Program of the Ministry of Commerce, Industry and Energy (MOCIE).

References

Doltsinis, I. and Kang, Z., 2004, "Robust Design of Structures Optimization Methods," *Computer Methods in Applied Mechanics and Engineering*, Vol. 193, pp. 2221～2237.

Fowlkes, W. Y. and Creveling, C. M., 1995, *Engineering Method for Robust Product Design*, Addison-Wesley, Massachusetts.

Guinta, A. and Watson, L., 1998, " A Comparison of Approximation Modeling Techniques :

Polynomial Versus Interpolating Models," *Proceedings of the 7th AIAA/USAF/NASA/ISSMO Symposium on Multidisciplinary Analysis and Optimization*, St. Louis, MO, AIAA, Vol. 2, Sept. 2-4, pp. 392~440 (AIAA-98-4758).

Gumbert, C.R., Newman, P.A. and Hou, G.J.-W., 2005, "High-Fidelity Computational Optimization for 3-D Flexible Wings: Part II-Effect of Random Geometric Uncertainty on Design," *Optimization and Engineering*, Vol. 6, No. 1, pp. 139~156.

Han, J. S. and Kwak, B. M., 2004, "Robust Optimization Using a Gradient Index: MEMS Applications," *Structural and Multidisciplinary Optimization*, Vol. 27, No. 6, pp. 469~478.

Jin, R., Chen, W. and Sudjianto, A., 2002, "On Sequential Sampling for Global Metamodeling in Engineering Design," DAC-34092, *Proceedings of the 2002 ASME IDETC Conferences*, Montreal, Quebec, Canada, Sep. 29-Oct. 2.

Jin, R., Du, X. and Chen, W., 2003, "The Use of Metamodeling Techniques for Optimization under Uncertainty," *Structural Multidisciplinary Optimization*, Vol. 25, No. 2, pp. 99~116.

Jones, D. R., Schonlau, M. and Welch, W. J., 1998, "Efficient Global Optimization of Expensive Black-Box Functions," *Journal of Global Optimization*, Vol. 13, No. 4, pp. 455~492.

Leary, S. J., Bhaskar, A. and Keane, A. J., 2004, "A Derivative Based Surrogate Model for Approximating and Optimizing the Output of an Expensive Computer Simulations," *Journal of Global Optimization*, Vol. 30, pp. 39~58.

Lee, B. C. and Jung, D. H., 2002, "Development of a Simple and Efficient Method for Robust Optimization," *International Journal for Numerical Methods in Engineering*, Vol. 53, pp. 2201~2215.

Lee, K. H. and Park, G. J., 2001, "Robust Optimization Considering Tolerances of Design Variables," *Computers & Structures*, Vol. 79, No. 1, pp. 77~86.

Lee, K. H. and Park, G. J, 2005, "A Global Robust Optimization Using Kriging Based Approximation Model," *Accepted by JSME International Journal*, Series C, Sep.

Lee, K. H., 2005, "Optimization of a Driver-Side Aribag Using Kriging Based Approximation Model," *Journal of Mechanical Science and Technology*, Vol. 19, No. 1, pp. 116~126.

Lee, T. H. and Jung, J. J., 2005, "Sensitivity Approach of Sequential Sampling Using Adaptive Distance Criterion," *Transactions of the KSME A (in Korean)*, Vol. 29, No. 9, pp. 1217~1224.

Park, G. J., Lee, T. H., Lee, K. H. and Hwang K. H., 2006, "Robust Design: An Overview," *AIAA J.*, Vol. 44, No. 1, pp. 181~191.

Parkinson, A., 1995, "Robust Mechanical Design Using Engineering Models," *Journal of Mechanical Design*, Vol. 117, pp. 48~54.

Sacks, J., Welch, W. J., Mitchell, T. J. and Wynn, H. P., 1989, "Design and Analysis of Computer Experiments," Statistical Science, Vol. 4, No. 4, pp. 409~435.

Sakata, S., Ashida, F. and Zako, M., 2003, "Eigenfrequency Optimization of Stiffened Plate Using Kriging Estimation," *Computational Mechanics*, Vol. 31, pp. 409~418.

Santner, T. J., Williams, B. J. and Notz, W. I., 2003, *The Design and Analysis of Computer Experiments*, Springer, New York.

Sherwood, G, 2005, http://home.att.net/~gsherwood/cover.htm.

Turner, C. J., Campbell, M. I. and Crawford, R. H., 2003, "Generic Sequential Sampling for Metamodel Approximations," DETC2003/CIE-48230, *Proceedings of DETC'03*, pp. 555~564.