Worst-Case-Scenario Robust Optimization Assisted by Dynamic Kriging Using Differential Evolution

Ziyan Ren, Bin Xia, Nyambayar Baatar, and Chang-Seop Koh, *Senior Member, IEEE*
College of Electrical and Computer Engineering, Chungbuk National University, Chungbuk, 361-763, Korea
E-mail: renziyan@chungbuk.ac.kr

***Abstract*—For the robust optimal design of electromagnetic problems under uncertainties, the robustness evaluation is the critical problem. This paper presents a surrogate model based worst-case-scenario optimization algorithm, where the dynamic Kriging is incorporated to construct a higher accurate surrogate model. Finally, an improved differential evolution algorithm, DE/λ-best/1/bin, is adopted to search for the global robust optimal solution.**

***Index Terms*—Differential evolution, dynamic Kriging, robust optimal design, worst-case-scenario optimization, uncertainty.**

I. Introduction

Uncertainties in design variables such as manufacturing tolerance and material property cannot be avoidable in engineering problems; so that robust optimizations have attracted increasing attention in electrical engineering [1].

The obvious feature of electromagnetic problems is that the expensive computing cost is required for performance analysis by finite element method (FEM), which hinders application of robust analysis. There have been many efforts to deal with this problem such as the worst vertex prediction based worst case optimization (WV-WCO) [2] and gradient index (GI) method [3]. The existing algorithms, however, are not available for a general application since the computational complexity of the WV-WCO is strongly proportional to the number of uncertain variables and constraints; furthermore, not all engineering problems are available for accurate GI by sensitivity analysis.

It is observed that surrogate models constructed by the response surface method and Kriging technique have been widely used to carry out design optimization [4]. Once the accurate meta-model is constructed, the robustness evaluation can be directly applied without intensive computational burden. In addition, a highly efficient optimization strategy is strongly requisite for the true robust optimal design.

Therefore, the target of this paper is to present a surrogate model based robust optimization algorithm. The performance robustness is measured through applying the WV-WCO to the accurate surrogate model constructed by the dynamic Kriging. Finally, the global robust optimum can be found by one simple and efficient differential evolution algorithm.

II. Surrogate-model Based Robust Optimization

A. Review of Worst Case Scenario Optimization

With the help of the worst vertex prediction, the worst case scenario optimization (WCO) problem subject to a set of constraint functions, $g_i(\mathbf{x})$, i=1,…,m, is formulated as follows:

$$\begin{aligned} &\text{minimize} \quad f_w(\mathbf{x}) \equiv f(\mathbf{x}_w) \\ &\text{subject to } g_{w,i}(\mathbf{x}) \equiv g_i(\mathbf{x}_{w,i}) \le 0, \quad i = 1,\cdots,m \end{aligned} \tag{1}$$

where $f_w(\mathbf{x})$ and $g_{w,i}(\mathbf{x})$ are the worst objective and the ith worst constraint values in the uncertainty set [2], respectively. If all uncertain design variables independently following Gaussian distribution with a standard deviation σ, the worst vertex ($\mathbf{x}_w$ or $\mathbf{x}_{w,i}$) of performance function $\Psi(\mathbf{x})$ (either objective function or constraint function) are decided as:

$$\mathbf{x}_w = \mathbf{x} + \begin{Bmatrix} \text{sign}\left(\Psi(\mathbf{x}+k\sigma_1\mathbf{e}_1) - \Psi(\mathbf{x}-k\sigma_1\mathbf{e}_1)\right)\cdot k\sigma_1 \\ \vdots \\ \text{sign}\left(\Psi(\mathbf{x}+k\sigma_n\mathbf{e}_n) - \Psi(\mathbf{x}-k\sigma_n\mathbf{e}_n)\right)\cdot k\sigma_n \end{Bmatrix} \tag{2}$$

where $\mathbf{e}_i$ is a n-dimension unit vector with 1 of the ith element.

Based on (2), the robustness evaluation in (1) consumes at most (2n+m+1) times of performance analysis by the FEM [2].

B. Surrogate Modeling by Dynamic Kriging

Based on N sampling points $[\mathbf{x}_1,\ldots,\mathbf{x}_N]^T$ with their corresponding responses $\mathbf{z}=[z(\mathbf{x}_1),\ldots,z(\mathbf{x}_N)]^T$, the response prediction at the point $\mathbf{x}_0$ is predicted by the Kriging technique as follows:

$$\hat{z}(\mathbf{x}_0) = \mathbf{f}_0^T\boldsymbol{\beta} + \mathbf{r}_0^T\mathbf{R}^{-1}(\mathbf{z}-\mathbf{F}\boldsymbol{\beta}) \tag{3a}$$

$$\boldsymbol{\beta} = \left(\mathbf{F}^T\mathbf{R}^{-1}\mathbf{F}\right)^{-1}\mathbf{F}^T\mathbf{R}^{-1}\mathbf{z} \tag{3b}$$

where the design matrix $\mathbf{F}=[f_k(\mathbf{x}_i)]$, ($i$=1,..,$N$; k=1,…,K) consists of K-term basis functions $f_k(\mathbf{x})$ and $\mathbf{f}_0$ is the basis function value at point $\mathbf{x}_0$. $\boldsymbol{\beta}$ is regression coefficient vector; $\mathbf{R}$ is the symmetric correlation matrix with the component $R_{ij}=R(\theta, \mathbf{x}_i, \mathbf{x}_j)$, i,j=1,..,N; and $\mathbf{r}_0=[R(\theta, \mathbf{x}_1, \mathbf{x}_0), \ldots, R(\theta, \mathbf{x}_N, \mathbf{x}_0)]^T$ is the correlation vector between $\mathbf{x}_0$ and N samples. The Gaussian covariance function $R(\theta, \mathbf{x}_i, \mathbf{x}_j)$ is shown as follows:

$$R(\boldsymbol{\theta}, \mathbf{x}_i, \mathbf{x}_j) = \exp\left[-\sum_{d=1}^{n}\theta_d (x_i^d - x_j^d)^2\right] \tag{4}$$

where the covariance parameter $\boldsymbol{\theta}=[\theta_1,\ldots,\theta_n]^T$ is obtained by the maximum likelihood estimation [5].

The dynamic Kriging dynamically selects the optimal basis function set at each design point [5]. For example, the set of a 2-D problem may be {1, x_1, x_2, x_1x_2, x_1^2, x_2^2,…}. The prediction error can be limited in a smaller interval as follows:

$$\left|z(\mathbf{x}_0) - \hat{z}(\mathbf{x}_0)\right| \le z_{1-\alpha/2}\sigma_p(\mathbf{x}_0) \tag{5a}$$

$$\sigma_p^2(\mathbf{x}_0) = \sigma^2(1 + \mathbf{w}_0^T\mathbf{R}\mathbf{w}_0 - 2\mathbf{w}_0^T\mathbf{r}_0) \tag{5b}$$

$$\mathbf{w}_0 = \mathbf{R}^{-1}\left(\mathbf{r}_0 + 1/2\sigma^2\cdot\mathbf{F}\delta\right) \tag{5c}$$

where $z_{1-\alpha/2}$ is the 1-α level quantile of the standard normal distribution, $\sigma_p^2(\mathbf{x}_0)$ is the predicted variance, σ^2 is the process variance, and δ is Lagrange multiplier. The $d(\mathbf{x}_0)=2z_{1-\alpha/2}\sigma_p(\mathbf{x}_0)$ is called the bandwidth of the prediction interval at point $\mathbf{x}_0$. The model accuracy generated with N samples is measured as:

$$mean\left[\left(d(\mathbf{x}_i)/2z_{1-\alpha/2}\right)^2\right]\Big/Var\left[z(\mathbf{x}_j)\right] \le \varepsilon_a, i=1,...T \quad (6)$$

where T is the number of test points and $Var()$ is the variance of N true responses (j=1,…,N). The ε_a is a predefined value such as 0.005 for low dimensional problems and 0.01 for high dimensional problems [5].

The highest order P of the polynomial basis function is decided by using the following condition:

$$C_{n+P}^{P} = (n+P)!/(P!n!) \le N \,. \quad (7)$$

Once the dynamic Kriging accurately approximates the responses, there is no further approximation in the estimation of the robustness, which can guide to search the true robust optimal design under uncertainty.

C. *Differential Evolution Algorithm—DE/λ-best/1/bin*

The DE/λ-best/1/bin algorithm is proposed in [6], in this paper, we give some improvements about survival condition and convergence condition. The numerical implementation for constrained optimization problem is summarized as follows:

Step 1: Randomly generate M individuals in design space $P_g = \{\mathbf{x}_1^g, \cdots, \mathbf{x}_M^g\}^T$ and set the maximum iteration $g_{\max}$.

Step 2: Check all the constraints and count the number of feasible individuals as N_f.

Step 3: Update mutant population by λ-best strategy as:
For the ith target vector $\mathbf{x}_i$,
- Remove $\mathbf{x}_i$ and select difference vectors ($\mathbf{x}^g_{r1}$, $\mathbf{x}^g_{r2}$),
- Sort current population P_g based on fitness value,
- Select the top λ individuals with better fitness values, where the λ value is determined as follows:
  1) λ=round($\varepsilon_1 M$) with $\varepsilon_1 \in$Ran(0,1) if N_f is bigger than round($\varepsilon_1 M$).
  2) λ= N_f if N_f is between 1 and λ.
  3) If N_f is equal to zero, then λ-best individuals will be selected among the most feasible vectors or those having the lowest constraint violation.
- Randomly select one base vector among the λ-best and generate mutant vector $\mathbf{v}_i$:

$$\mathbf{v}_i^g = \mathbf{x}_{l_k}^g + \alpha \cdot (\mathbf{x}_{r1}^g - \mathbf{x}_{r2}^g), \quad k = 1,...,\lambda \,. \quad (8)$$

where α is a scaling factor and l_k is index of the kth element of λ-best.

Step 4: Crossover by binominal strategy as:
Generate the new trial vector $\mathbf{u}_i^g$ using:

$$u_{i,j}^g = \begin{cases} v_{i,j}^g & \text{if}\,(rand_j(0,1) \le C_r \text{ or } j = j_{rand}) \\ x_{i,j}^g & \text{otherwise} \end{cases} \quad (9)$$

where C_r and j_{rand} are crossover factor and random integer from 1 to n, respectively.

Step 5: Evaluate objective function, and check the feasibility of each trial vector and update N_f.

Step 6: Survival criterion.
For the ith trial vector $\mathbf{u}^g_i$ and target vector $\mathbf{x}^g_i$,
- If both are feasible, select the survivor as follows:

$$\mathbf{x}_i^{g+1} = \begin{cases} \mathbf{u}_i^g & f(\mathbf{u}_i^g) \le f(\mathbf{x}_i^g) \\ \mathbf{x}_i^g & f(\mathbf{u}_i^g) > f(\mathbf{x}_i^g) \end{cases}. \quad (10)$$

- If one is feasible, select the feasible one;
- If both are infeasible, select that with lower sum of constraint violations.
- Update the global best solution $\mathbf{x}_b^{g+1}$;

Step 7: Termination condition.
Stop if condition (11) is satisfied for all individuals during 10 consecutive iterations, or if the iteration counter g reaches $g_{\max}$. Otherwise, go to *Step* 3.

$$\left| f(\mathbf{x}_b^{g+1}) - f(\mathbf{x}_b^g) \right| \le \varepsilon_2 \quad (11)$$

where ε_2 is set as 0.0001.

Based on the above analysis, it is obvious the surrogate model based WV-WCO does not rely on any gradient information and can be coupled to arbitrary simulation tools. The whole flowchart of the proposed robust optimization algorithm is shown in Fig.1.

III. Optimization Result

The TEAM Workshop problem 22, which is one design application about the superconducting magnetic energy storage system, is selected to validate the proposed algorithm [7]. The optimization results will be shown in the full paper.

References

[1] X. J. Meng, *et al.*, "Robust multilevel optimization of PMSM using design for six sigma," *IEEE Trans. Magn.*, vol.47, no.10, pp. 3248-3251, Oct 2011.

[2] G. Steiner, A. Weber, and C. Magele, "Managing uncertainties in electromagnetic design problems with robust optimization," *IEEE Trans. Magn.*, vol.40, no.2, pp. 1094-1099, Mar 2004.

[3] J. S. Han and B. M. Kwak, "Robust optimization using a gradient index: MEMS applications,"*Struct Multidisc Optim*, vol.27, pp.469-478, 2004.

[4] N. V. Queipo, *et al*., "Surrogate-based analysis and optimization," *Progress in Aerospace Sciences*, vol. 41, pp.1-28, 2005.

[5] I. Lee, K. K. Choi, and L. Zhao, "Sampling-based RBDO using the stochastic sensitivity analysis and dynamic Kriging method," *Struct Multidisc Optim*, vol.44, pp.299-317, 2011.

[6] N. Baatar, D. H. Zhang, and C. S. Koh, "An improved differential evolution algorithm adopting λ-best mutation strategy for a global optimization of electromagnetic devices," to be published.

[7] P. G. Alotto, et al., "SMES optimization benchmark, TEAM workshop problem 22," Available: http://www.igte.tugraz.at/team 22.

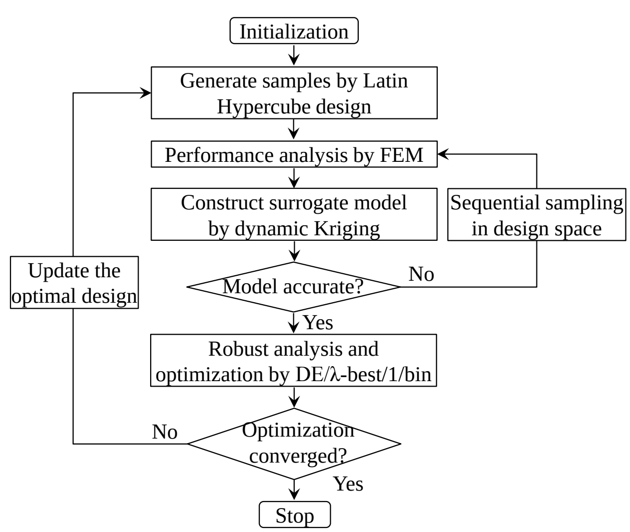

Fig.1 Flowchart of the proposed robust algorithm.