Augmented D-Optimal Design for Effective Response Surface Modeling and Optimization

Min-Soo Kim*
Research Professor, Center of Innovative Design Optimization Technology, Hanyang University, 17 Haengdang-dong, Sungdong-gu, Seoul 133-791, Korea
Kyung-Jin Hong
Graduate Research Assistant, Center of Innovative Design Optimization Technology, Hanyang University, 17 Haengdang-dong, Sungdong-gu, Seoul 133-791, Korea
Dong-Hoon Choi
Director, Center of Innovative Design Optimization Technology, Hanyang University, 17 Haengdang-dong, Sungdong-gu, Seoul 133-791, Korea

For effective response surface modeling during sequential approximate optimization (SAO), the *normalized and the augmented* D-optimality criteria are presented. The *normalized* D-optimality criterion uses the normalized Fisher information matrix by its diagonal terms in order to obtain a balance among the linear-order and higher-order terms. Then, it is augmented to directly include other experimental designs or the pre-sampled designs. This augmentation enables the trust region managed sequential approximate optimization to directly use the pre-sampled designs in the overlapped trust regions in constructing the new response surface models. In order to show the effectiveness of the normalized and the *augmented* D-optimality criteria, following two comparisons are performed. First, the information surface of the *normalized* D-optimal design is compared with those of the original D-optimal design. Second, a trust-region managed sequential approximate optimizer having three D-optimal designs is developed and three design problems are solved. These comparisons show that the *normalized* D-optimal design gives more rotatable designs than the original D-optimal design, and the *augmented* D-optimal design can reduce the number of analyses by 30 % - 40 % than the original D-optimal design.

Key Words : Sequential Approximate Optimization; RSM; D-Optimality

1. Introduction

Expensive analyses and experiments being frequently encountered in the modern engineering optimizations, sequential approximate optimization (SAO) strategies have gained in popularity (Barthelemy and Haftka, 1993). Especially, response surface models (RSM) can be widely used for replacing some high fidelity computational models (Haftka et. al, 1998; Unal et. al, 1998; Roux et. al, 1998; Sobieski et. al, 2000). Thus, it is important in constructing response surface models to achieve an acceptable level of accuracy while attempting to minimize the computational effort, i.e. the number of system analyses. Although increasing the number of design points could improve the accuracy of the approximate model, many studies have concentrated on reducing the number of analyses and experiments (John, 1998).

* Corresponding Author,
E-mail : a004084@hanyang.ac.kr
TEL : +82-2-2290-1639; **FAX** : +82-2-2291-4070
Research Professor, Center of Innovative Design Optimization Technology, Hanyang University, 17 Haengdang-dong, Sungdong-gu, Seoul 133-791, Korea.
(Manuscript **Received** June 18, 2001; **Revised** November 23, 2001)

To do this, several methods have been used such as the fractional factorials, the central composite design (CCD) (Box and Wilson, 1951), and the D-optimal design (Box and Draper, 1971; Miller and Nguyen, 1993). Although the factorial approach may be optimal, as judged by the D-optimal design, in some situation that all the factors are restricted to a cuboidal region, the approximated model are often not sensible ones in practical work (Box and Draper, 1971). Especially, expensive experiments and analyses enforce one to accomplish the goal with the smallest number of experiments or analyses. Thus, Although the CCD gives more rotatable design from the viewpoint of a variance optimal design (Box and Draper, 1987), the D-optimal design is widely recommended in most of recent studies, because it can use only the equal analyses or experiments to the number of unknown-coefficients in the approximate model to be fitted.

However, many studies presented that this saturated D-optimal design made poor converge of the region of interest, especially when genetic algorithm (GA) is used in the D-optimal design. Thus the 20% to 50% *super-saturated* D-optimal designs are widely recommended for approximation model building (Unal et. al, 1998; Carpenter, 1993). This is because it leaves a good choice for response surface model building for the deterministic experiments such as numerical analyses. Although these 20% to 50% *super-saturated* D-optimal designs gave good results in some studies, these *super-saturated* D-optimal designs do not seem general but just an empirical guideline.

In order to overcome the computational burden of these *super-saturated* D-optimal designs and the excessive concentration along the perimeter of the normalized design space, this study first proposes a *normalized* D-optimality criterion, which gives equal weightings between linear-order and higher-order terms. This makes the sampled designs to be rotatable and enables one to use only the saturated designs. Also, an *augmented* D-optimal design, based on the *normalized* D-optimal design, is suggested in order to include the pre-selected design points from other experimental designs, which enables one to easily use the typical orthogonal array as the preliminary design of the D-optimal design and to directly use the pre-sampled design points in SAO with RSM.

Section 2 reviews the original D-optimal design from the view of experimental design for RSM. Section 3 fully describes the proposed a *normalized* D-optimal design and it's augmented form. In Sec. 4, the distribution and the information surface contours of the *normalized* D-Optimal design are compared with those of the original, and show the numerical performance of SAO combined with the proposed D-optimal designs such as a *normalized* D-optimal design or a *augmented* D-optimal design. The concluding remarks are presented in Sec. 5.

2. Review of the D-Optimality Criterion

2.1 D-optimality design from the view of experimental design

In order to simplify the explanation of the basis for D-optimal designs, consider the following matrix notation as:

$$\mathbf{y}=\mathbf{X}\boldsymbol{\theta}+\boldsymbol{\varepsilon} \tag{1}$$

where $\boldsymbol{\theta}$ is a vector of $\boldsymbol{n}$ parameters to be estimated, $\boldsymbol{\varepsilon}$ is a normally distributed experimental error with mean zero and constant variance σ^2. Then, $\boldsymbol{\theta}$ can be estimated using the least square method as:

$$\boldsymbol{\theta}=(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y} \tag{2}$$

The D-optimal design states that the $\boldsymbol{n}$ points are chosen, in the normalized design space, to maximize the determinant $|\mathbf{X}^T\mathbf{X}|$, which can minimizes the maximum variance of any predicted value of the function and the variance of the parameter estimates. This can be interpreted using the variance-covariance matrix defined as:

$$V(\boldsymbol{\theta})=\sigma^2(\mathbf{X}^T\mathbf{X})^{-1} \tag{3}$$

Eq. (3) is a statistical measure of the goodness-of-fit. This represents that the D-optimal design has same meaning of minimizing the asymptotic confidence regions for the maximum likelihood estimates (Haftka et. al. 1998).

2.2 D-optimality criterion from the view of response surface modeling

The information function is defined as the reciprocal of the variance; that is,

$$I_x = V(\hat{y})^{-1} = \{ n \cdot \mathbf{z}^T (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{z} \}^{-1}$$

where $\mathbf{z} = (1\ x_1\ x_2\ x_1^2\ x_2^2\ x_1x_2)^T$ for a second-order model with two design variables and n is the number of unknown coefficients. An experimental design is said to be *rotatable* if the variance of the predicted response $\hat{\mathbf{y}}$ at some point $\mathbf{x}$ is a function only of the distance of the point from the design center and is not a function of direction (Box and Draper, 1987). Furthermore, a design with this property will leave the variance of $\hat{\mathbf{y}}$ unchanged when the design is rotated about the center; hence, the name *rotatable* design.

Rotability is a very important property in the selection of a response surface design. Since the purpose of RSM is optimization and the location of the optimum is unknown prior to running the detailed experiment, it makes sense to use an experimental design that provides equal precision of estimation in all direction.

However, the original D-optimality criterion is only one of many single-valued criteria that might be used in attempts to describe some important characteristics of the Fisher information matrix $\mathbf{X}^T\mathbf{X}$.

3. Augmented D-Optimal Design

3.1 Normalized D-optimal design

In order to balance the weightings among whole the sampled designs, we normalize the original Fisher information matrix $\mathbf{M} = \mathbf{X}^T\mathbf{X}$ as

$$\mathbf{D} = \mathbf{X}^T\mathbf{S}\mathbf{X} \qquad (4)$$

where the scaling matrix $\mathbf{S}$ is a diagonal matrix whose values are $S_{ii} = w_{ii}/M_{ii}$ and Now, the *normalized* D-optimal design is obtained by solving

$$\max_{\mathbf{X}} |\mathbf{D}| \text{ or } \min_{\mathbf{X}} |\mathbf{D}|^{-1} \qquad (5)$$

Although these two formulations represent mathematically the same concept, they give somewhat different meanings from the view of numerical analysis. The maximization formulation will diverge to the infinite but the minimization formulation will be converged to the small and positive real value. Thus, the latter seems to be more observable than the former in the numerical computation. Hence, this study uses the minimization formulation.

Now, we explain the meaning of the normalized Fisher information matrix of Eq. (4). Suppose that the normalized Fisher information matrix is decomposed by $\mathbf{D} = \mathbf{L} \cdot \mathbf{U}$. Then the determinant of $\mathbf{D}$ can be defined as

$$|\mathbf{D}| = \prod_{j}^{N} \beta_{jj} \qquad (6)$$

where $\beta_{ij} = D_{ij} - \sum_{k=1}^{i-1} L_{ik} \cdot U_{kj}$ for $i = 1, 2, \cdots, j$ and $j = 2, 3, \cdots, N$. Also, the values of L_{ik} and U_{kj} are the components of $\mathbf{L}$ and $\mathbf{U}$, respectively. As all the diagonal components of $\mathbf{D}$ have unity respectively, Eq. (5) can not help reducing the interaction terms in order to maximize the determinant of $\mathbf{D}$. Hence, for the second-order response model based on the 3^k designs, maximizing the determinant of the normalized Fisher information matrix represents that all points on the corners move into the center and have equal distances from the center. However, the original D-optimal design mathematically gives the 3^k designs as the optimum design, because the linear terms are more dominant than others in its formulation.

We believe that this *normalized* D-optimal design satisfying Eq. (5) gives better near-*rotatable* than the original D-optimal design, because it can implicitly reduce inevitable correlation between coefficients terms in a quadratic response surface model.

3.2 Augmented D-optimal design including other experimental designs

Now, we augment the normalized D-optimality criterion to include other experimental designs such as orthogonal arrays or pre-sampled experimental designs. For including these pre-sampled designs, the *augmented* D-optimality criterion is suggested as

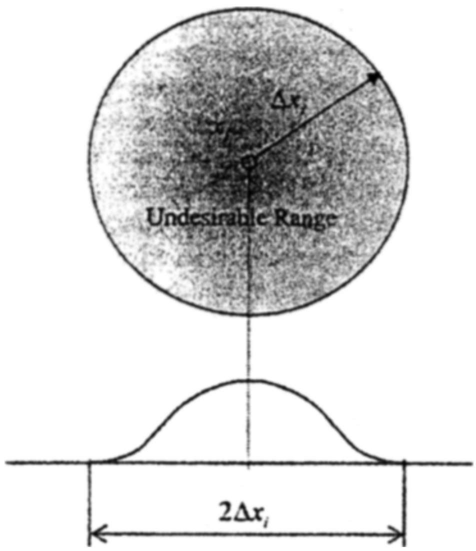


Fig. 1 Graphical representation of the second term in Eq. (8)

$$\min_{x} \frac{(|\mathbf{D}|^{-1}-|\mathbf{D}|_*^{-1})}{|\mathbf{D}|_*^{-1}}+\sum_{i=1}^{nps} w_i\left\{\cos\left[\frac{\pi}{2}\frac{\| x-x_i \|}{\Delta x_i}\right]\right\} \quad (7)$$

where represent new selecting points, x_i denotes the i^{th} pre-sampled point, Δx_i is the radius to separate new sampling points from the i^{th} pre-sampled point, and nps is the total number of the pre-sampled designs. Also, w_i denotes the weighting coefficient for the i^{th} pre-sampled design. The value of $|\mathbf{D}|_*^{-1}$ is the final value of $\min_{\mathbf{x}} |\mathbf{D}|^{-1}$ in Eq. (5). Figure 1 graphically shows the second term of Eq. (7).

When the trust-region managed sequential approximate optimization (SAO) is performed using the response surface models (Dennis & Torczon, 1996; Alexandrov, 1996; Rodriguez et. al., 1998), the trust-regions can be partially overlapped during consecutive iterations, shown in Fig. 2. Although one wants to use the pre-sampled designs in this overlapped region for constructing the new response surface models, the original or *normalized* D-optimality criteria do not use them because their design points are pre-determined. The proposed *augmented* D-optimality criterion of Eq. (7), however, can use these pre-sampled design points in the overlapped region and additionally sample only the $(np-nps)$ numbers of new designs.

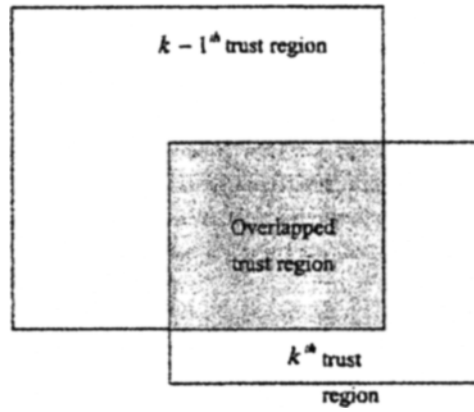
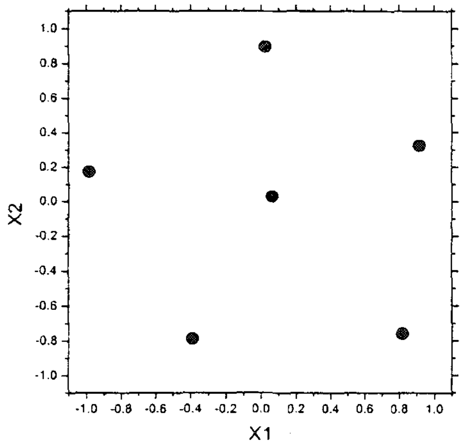


Fig. 2 The overlapped trust region between consecutive iterations in SAO with trust region model management strategy

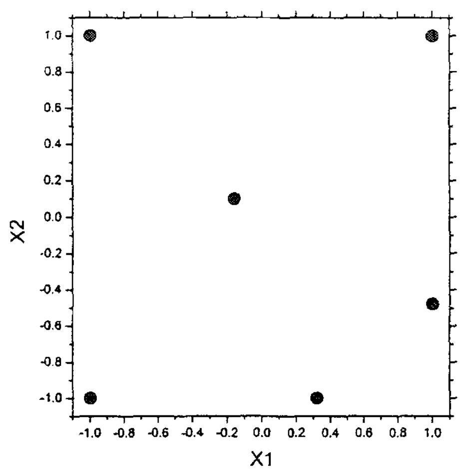
4. Numerical Studies

4.1 Comparison of the normalized and the original D-optimal designs

In order to show the effectiveness of the proposed optimality criteria, we graphically compare the sampled design points for a quadratic response surface model with those of the original D-optimality. Figure 3 shows the saturated design points of the *normalized* D-optimality criterion and the original D-optimality criterion in developing the quadratic model having two design variables. The *normalized* D-optimality criterion gives an axis-symmetric distribution, which is similar to the pentagon of the well-known equiradial designs for two variables. Now we compare the information surface for those two criteria. Figure 4 shows the information functions for those saturated designs shown in Fig. 3. The *normalized* D-optimal design seems to be better near-*rotatable* than the original. Figure 5 shows the 50 % *super-saturated* D-optimal design based on the original D-optimality criterion and their information surface. As Box and Draper (1971, 1987) had described, the optimum of this 50 % *super-saturated* D-optimal design with two variables is obtained as 3^2 factorial design. This shows that our in-house GA program (Kim and Park, 2001) is quite reliable to solve D-optimality criteria. However, It has been well



(a) The normalized D-optimal design



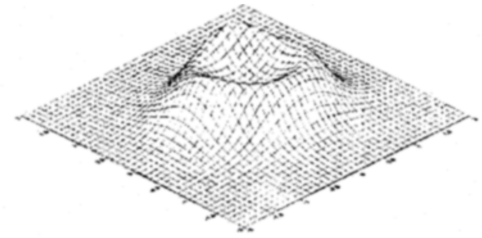
(b) The original D-optimal design

Fig. 3 Comparison of the saturated designs sampled by the normalized D-optimality criterion and the original D-optimality criterion

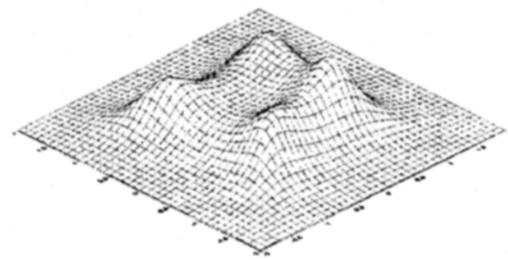
known that 3^2 factorial design was not *rotatable* and 3^k designs and their fractions were not good choices for second-order response surface designs.

4.2 Numerical comparisons of the proposed and the original D-optimality criteria during SAO

In order to show the numerical performance, the proposed and the original D-optimality criteria are embedded in the SAO (Hong, Kim and Choi, 2001), which is composed of three modules such as SAO manager, approximate

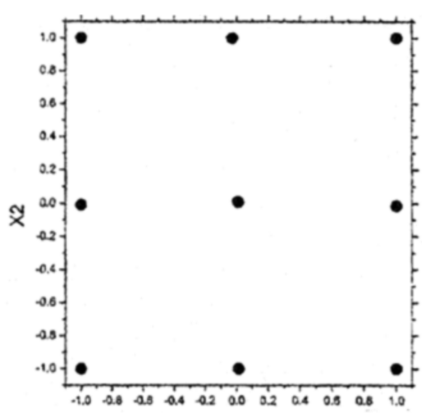


(a) The normalized D-optimal design

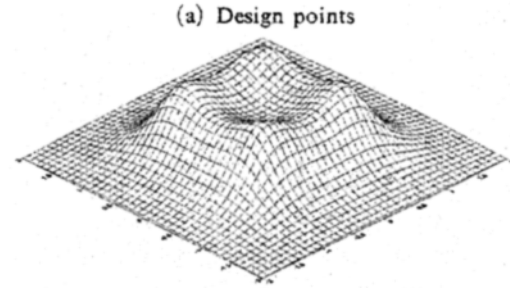


(b) The original D-optimal design

Fig. 4 Comparison of the information surface for the saturated designs sampled by the normalized D-optimality criterion an the original D-optimality criterion



(a) Design points



(b) Information surface

Fig. 5 Design points and the Information surface of the 50 % super-saturated D-optimal design

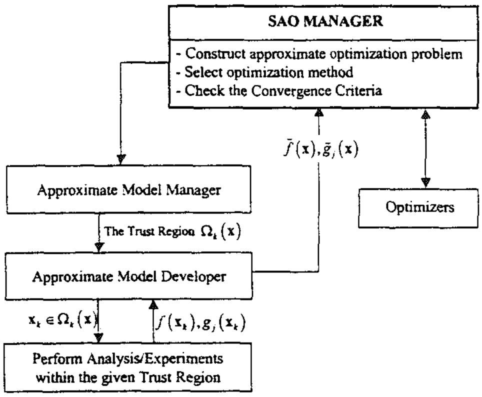


Fig. 6 Skeleton Structure of Sequential Approximate Optimizer based on RSM

model manager, and optimization process shown in Fig. 6. SAO manager constructs the approximate optimization problem and checks the convergence. Approximate model manager controls the accuracy of the response surface models with trust region concept. We use L_1 exact penalty function (Fletcher, 1987) of (8) to manage the trust region for nonlinear constrained optimization.

$$L_1(\mathbf{x}) = f(\mathbf{x}) + r\sum_{j=1}^{m} \max(g_j(\mathbf{x}), 0) \qquad (8)$$

where $f(\mathbf{x})$, $g_j(\mathbf{x})$ and r are the objective function, the j^{th} inequality constraint function, and the penalty parameter. This functional is exact in the sense that local minimums of the functional are equivalent to local minimums of the original problem to a large extent. The *augmented* D-optimality criteria are embedded in the module of Approximate Model Developer (AMD).

[Example 1] *Gear Reducer Design Problem*

This design problem has been widely used in the literature (Azam and Li, 1989). The design objective of this gear reducer system (shown in Fig. 7) is to minimize the overall volume (or weight) while satisfying the bending stress of the gear tooth, the contact stress of the gear tooth, the transverse deflection of the shafts, the stresses of the shaft and dimensional restriction of design variables. The design variables are gear face width (x_1), teeth module(x_2), number of teeth of pinion(x_3), distance between bearings I (x_4), distance between bearing II(x_5), diameter of shaft I (x_6) and diameter of shaft II(x_7).

Table 1 Optimization results of the gear reducer system design

	Initial Design	Augmented D-optimality	Normalized D-optimality	Original D-optimality
x_1	3.1	3.504	3.499	3.499
x_2	0.75	0.700	0.700	0.700
x_3	22.5	17.000	17.000	17.000
x_4	7.8	7.300	7.300	8.050
x_5	7.8	7.719	7.723	7.718
x_6	3.4	3.353	3.358	3.352
x_7	5.25	5.289	5.294	5.289
f^*	4144.83	2998.29	3000.43	3002.57
Iteration	-	3	3	3
NF	-	74	112	112

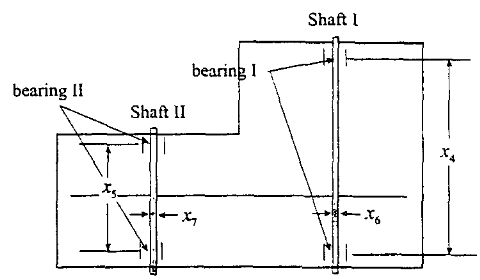


Fig. 7 Gear reducer system

The optimization results are listed in Table 1. In Table 1, f^* represents the final objective value, Iteration denotes the numbers of approximations, and NF denotes the cumulative number of function evaluations during SAO. All three methods give the similar results. However, the augmented D-optimal design can reduce the number of function evaluations by 30 % than other designs. This represents that the trust-regions are frequently overlapped as optimization progresses.

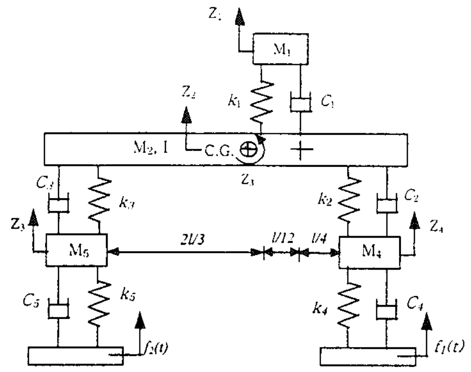
[Example 2] *Five degree-of-freedom Vehicle Suspension System Design Problem*

Figure 8(a) shows a five degree of freedom vehicle suspension system (Haug and Arora, 1979; Kim and Choi, 1998), which is to be designed to minimize the extreme acceleration of the driver's seat for a given vehicle speed and a road surface profile shown in Fig. 8(b). This profile is a combination of two sinusoidal curves with different half-wavelengths, which represents a severe bump condition. Spring constants k_1, k_2

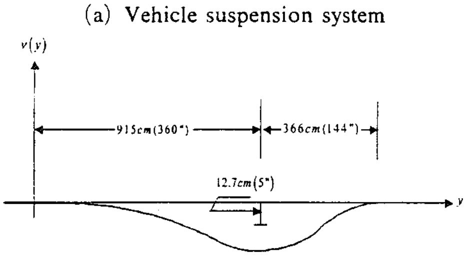
Table 2 Optimization results of the five-degree-of-freedom vehicle suspension system

	Initial Design	Augmented D-optimality*	Normalized D-optimality*	Original D-optimality*
b_1	100.00	50.00	50.00	50.00
b_2	300.00	200.00	200.00	200.00
b_3	300.00	200.00	200.00	793.70
b_4	10.00	36.56	21.88	47.23
b_5	25.00	77.38	77.04	77.34
b_6	25.00	45.08	41.16	80.00
f^*	331.79	256.28	256.74	255.03
Iteration		7	7	7
NF		173	302	302

* 50 % super-saturated sampling is employed.



(a) Vehicle suspension system



(b) Road surface profile

Fig. 8 Five-degree-of-freedom Vehicle Model

and k_3 and damping coefficients c_1, c_2 and c_3 of the system are chosen as design variables. The motion of the vehicle is constrained so that the relative displacements between the chassis and the driver's seat, the chassis and the front and rear axles, and the road surface and the front and rear axles are within given limits. The design variables are also constrained.

The optimization results are listed in Table 2. All three methods give the similar results. However, the augmented D-optimal design can reduce the number of function evaluations by 40 % than other designs. In this problem, only the normalized D-optimal design is successfully converged when the saturated sampling is employed. Hence, the optimization results of Table 2 are obtained using the 50 % super-saturated sampling points.

5. Concluding Remarks

This study proposed an *augmented* D-optimal design for effective response surface modeling during sequential approximate optimization (SAO). The proposed method fundamentally uses the normalized Fisher information matrix by its diagonal terms in order to obtain a balance among the linear-order and higher-order terms. Then, it is augmented to directly include other experimental designs or the pre-sampled designs.

The proposed method was embedded into the sequential approximate optimization framework. Then it solved the two typical examples such as a gear reducer design and a five-degree-of-freedom vehicle suspension system design. The optimization results are compared with those of the original D-optimal design. These comparisons showed that the proposed *augmented* D-optimal design could reduce the number of analyses by 30 % - 40 % than the original D-optimal design, while obtaining the similar optimum values.

Acknowledgement

This research was supported by center of Innovative Design Optimization Technology (iDOT), Korea Science and Engineering Foundation.

References

Alexandrov, N., 1996, Robustness Properties of a Trust Region Frame Work for Managing Approximations in Engineering Optimization. *Proceedings of the 6th AIAA/NASA/USAF Multidisciplinary Analysis & Optimization Symposium*, AIAA 96-4102-CP, Bellevue, Washington, September 7-9, pp. 1056 ~1059.

Azam, S and Li, W. -C., 1989, Multi-Level Design Optimization Using Global Monotonicity Analysis, *Journal of Mechanisms, Transmissions, and Automation in Design*, Vol. 111, pp. 259 ~ 263.

Barthelemy, J. F. and Haftka, R. T., 1993, Approximation Concepts for Optimum Structural Design-a Review. *Struct. Optim.* 5, pp. 129 ~144.

Box, G. P. and Wilson, K. B., 1951, On the Experimental Attainment of Optimum Conditions, *Journal of the Royal Statistical Society, Series B*, Vol. 13, pp. 1 ~ 45.

Box, G. P. and Draper, N. R., 1987, Empirical Model-Building and Response Surfaces, *John Wiley & Sons*, New York, pp. 481 ~ 489.

Box, M. J. and Draper, N. R., 1971, Fractional Designs, the $|X'X|$ Criterion, and Some Related Matters. *Technometrics* Vol. 13, No. 4, pp. 731 ~ 742.

Carpenter, W. C., 1993, Effect of Design Selection on Response Surface Performance. *NASA Contract Report4520.*

Dennis, J. E. and Torczon, T., 1996, Approximation Model Management for Optimization. *Proceedings of the 6^{th} AIAA/NASA/USAF Multidisciplinary Analysis & Optimization Symposium*, AIAA 96-4046, Bellevue, Washington, September 7-9, pp. 1044 ~1046.

Fletcher, R., 1987, *Practical Method of Optimization*. John Wiley & Sons: Chichester, pp. 296 ~ 304.

Haftka, R. T, Scott, E. P. and Cruz, J. R., 1998, Optimization and Experiments: A Survey. *ASME Appl. Mech.* Vol. 51, No. 7, pp. 435 ~ 448.

Haug, E. J. and Arora, J. S., 1979, *Applied Optimal Design*. Wiley-Interscience: New York, pp. 242 ~ 245.

Hong, K. -J, Kim, M. -S., Choi, D. -H., 2001, Efficient Approximation Method for Constructing Quadratic Response Surface Model, *KSME International Journal*, Vol. 15, No 7, pp 876 ~ 888.

John, P. W. M., 1998, *Statistical Design and Analysis of Experiments (Classics In Applied Mathematics 22)*, SIAM: Philadelpia, pp. 123 ~ 216.

Kim, M. -S. and Choi, D. -H., 1998, Min-Max Dynamic Response Optimization of Mechanical Systems Using Approximate Augmented Lagrangian, International Journal for Numerical Methods in Engineering, Vol. 43, pp. 549 ~ 564.

Kim, M. -S. and Park, H. -W., 2001, Fast and Global Convergent Genetic Program with Adaptive Fitness Target, *iDOT Technical Report OPT-01-01.*

Miller, A. J. and Nguyen, N. K., 1993, An algorithm for generating experimental designs which are close to being D-optimal designs using Fedorov's exchange algorithm. http://lib.stat.cmu.edu/design/dopt.

Rodriguez, J. F., Renaud, J. E. and Watson, L. T., 1998, Trust Region Augmented Lagrangian methods for Sequential Response Surface Approximation and Optimization. *Journal of Mechanical Design*, ASME Vol. 120, pp. 58 ~ 66.

Roux, W. J., Stander, N. and Haftka, R. T., 1998, Response Surface Approximation for Structural Optimization. *International Journal for Numerical methods in Engineering*, Vol. 42, pp. 517 ~ 534.

Unal, R., Lepsch, R. A. and McMilin, M. L., 1998, Response Surface Model Building and Multidisciplinary Optimization Using D-Optimal Designs. *Proceedings of the 7th AIAA/USAF/NASA/ISSMO Symposium on Multidisciplinary Analysis and Optimization, St. Louis, Missouri, September 2-4* AIAA-98-4759, pp. 405~411.

Sobieski, J. S., Kodiyalam, S., Yang, R. J., 2000, Optimization of Car body under Constraints of Noise, Vibration, and Harshness (NVH), and Crash. *Proceedings of the 41st AIAA/ASME/ASCE/AHS/ASC, Structures, Structural Dynamics, and Materials Conference and Exhibit*, April 3-6 AIAA-2000~1521.