Development of efficient global optimization method for discontinuous optimization problems with infeasible regions using classification method

Naohiko BAN* and Wataru YAMAZAKI*
*Department of Energy and Environment Science, Nagaoka University of Technology
1603-1 Kamitomioka-machi, Nagaoka, Niigata 940-2188, Japan
E-mail: n_ban@stn.nagaokaut.ac.jp

Received: 5 September 2018; Revised: 30 December 2018; Accepted: 17 February 2019

Abstract
The objective of this research is to efficiently solve discontinuous optimization problems as well as optimization problems with large infeasible regions in the design variables space. Recently, major optimization targets have been changed to more complicated ones such as topology optimization problem, discontinuous optimization problem, robust optimization problem and high dimensional optimization problem. The aim of this research is to efficiently solve the complicated optimization problems by using machine learning technologies. In aerodynamic optimization problems at supersonic flow conditions, it is confirmed that aerodynamic objective functions have discontinuity due to shock waves and it needs to treat the discontinuous functions and large infeasible regions via strong shock waves. In this research, therefore, we develop an efficient global optimization method for discontinuous optimization problems with infeasible regions using classification method (EGODISC). The developed method is compared with a Bayesian optimization method using the Matern 5/2 kernel Gaussian process regression and a genetic algorithm to verify the usefulness of the developed method. The Bayesian optimization falls into an infinite loop in its optimization process by selecting an additional sample point in the infeasible regions. On the other hand, the developed method can work well with the infeasible regions in the design variables space. It is confirmed that EGODISC can be effectively used with discontinuous aerodynamic objective functions. It is also confirmed that EGODISC can be effectively used for a shape optimization problem with large infeasible regions by the negative thickness of airfoil.

Keywords: Efficient global optimization, Machine learning, Support vector machine, Discontinuous optimization, Infeasible region

1. Introduction

Design optimization techniques have already been used in actual designs such as a winglet shape design of MRJ and a nose shape design of the Bullet Train 700 series. Recently, major optimization targets have been changed to more complicated ones such as topology optimization problem, discontinuous optimization problem, robust optimization problem, high dimensional optimization problem and so on.

The optimization methods in numerical simulation fields can be roughly divided into two groups, invasive and noninvasive types. The noninvasive optimization methods include evolutionary strategies such as genetic algorithms (GA) and differential evolution. GA is one of the metaheuristic methods inspired by the process of natural generation such as selection, crossover and mutation. It has the feature to explore global optimal solutions even with complicated objective functions. The invasive optimization method includes adjoint methods. In the adjoint methods, the gradient of the objective function with respect to all design variables can be obtained at once by solving the adjoint equation. It has the feature to be able to explore a local optimal solution even in high dimensional optimization problems. However, there are also demerits in both methods. GA requires an enormous number of evaluations which indicates a very large calculation cost. The adjoint approach requires to develop the program code to solve the adjoint equation and is not
possible to search global optimal solutions. To tackle these issues, there are efficient optimization methods to solve the complicated optimization problems by using noninvasive approaches and machine learning techniques. For example, there is an optimization method using the Bayesian approach such as Gaussian process (Kriging). The Bayesian optimization method has already been applied to a design optimization problem of supersonic transport, and it could obtain design knowledge of a twin-body / biplane-wing configuration (Ban et al., 2018). There are many other methods for efficient optimization, such as GPEME and ALOS. The Gaussian Process surrogate model assisted Evolutionary algorithm for Medium-scale computationally Expensive optimization problems (GPEME) uses dimension reduction machine learning techniques for tackling the “curse of dimensionality”. In GPEME, the Sammon mapping is introduced to transform the design variables space to take advantage of Gaussian process surrogate modeling in a low-dimensional space (Liu et al., 2014). However, since the dimension reduction method may lose some neighborhood information of the training data points in the original space, GPEME is difficult to obtain comparable results with the direct Bayesian optimization. The Agglomeration of Locally Optimized Surrogates (ALOS) uses multiple local surrogate models to overcome the limited modeling flexibility of a single global surrogate model when there is a heterogeneity in the set of sample points (Rumpfkeil and Beran, 2016). However, since it requires many sample points to extract discontinuous boundaries, ALOS is difficult to apply to multiple dimensional optimization problems. Furthermore, several optimization methods for constrained optimization problems have already been proposed using machine learning technologies such as random forest classifiers (Lee et al., 2010) or probabilistic Support Vector Machine model (Basudhar et al., 2012). In these approaches, the probability of feasibility is calculated by the classification methods, and then the objective function is defined as the product of EI value and the probability of feasibility. In these approaches, therefore, probabilistic classification methods can only be adopted as the classification method.

In optimization problems at supersonic flow conditions, it was confirmed that the objective functions had discontinuity due to shock waves and it needs to treat the discontinuous aerodynamic functions and large infeasible regions via strong shock waves. However, the conventional optimization methods such as the Bayesian optimization are difficult to efficiently solve the discontinuous optimization problem since the effect of the discontinuous change propagates in the design variables space and the accuracy of the response surface model will be reduced. In addition, the negative thickness of airfoil occurs depending on the definition of the ranges of design variables. Since the negative thickness of airfoil leads to an infinite loop in the optimization process by iteratively selecting a same additional sample point in the infeasible regions, the wing shape is difficult to optimize in detail. Therefore, it needs to solve the issues relating to the discontinuous optimization problem and the infeasible regions to efficiently obtain optimal solutions in the supersonic flow conditions. The aim of this research is to efficiently solve the discontinuous optimization problems using a machine learning approach, which can also solve the issues relating to the infeasible regions. The present paper is organized as follows. The optimization methods and performance evaluation methods used in this research are concisely described in section 2. Then, optimization results obtained by the conventional Bayesian optimization and the developed approach are discussed in section 3. Finally, concluding remarks are provided in section 4.

2. Optimization and evaluation methods
2.1 Bayesian optimization

Surrogate model-based optimization methods are known as efficient global optimization methods to optimize blackbox functions. In this approach, firstly, initial sample points are generated in the design variables space by a Latin Hypercube Sampling (LHS) method, and then the objective function values are evaluated. By using the information of the initial sample points, initial surrogate models are constructed. In the Bayesian optimization approach, Gaussian process regression methods (GPs, Kriging) are usually used for the construction of the surrogate model (Williams and Carl, 1996). The GPs can predict the objective function values considering the uncertainty at unknown sample point locations giving input / output relationships. Promising locations in the design variables space are explored by a real-coded genetic algorithm on the surrogate models. The promising locations are explored by the criteria of expected improvement (EI) (Jones et al., 1998). The function of EI expresses a potential for improvement in design variables space which considers both estimated function value as well as the uncertainty of the surrogate model. The objective function values are evaluated at the explored promising locations where EI is maximal, and then new surrogate models are created by adding its information. By the iterative process described above, the accuracy of the surrogate models is
progressively improved around the promising locations in the design variables space. The flowchart of the Bayesian optimization is summarized in Fig. 1.

![Flowchart of Bayesian optimization](image)

**Fig. 1** Flowchart of Bayesian optimization, which corresponds to conventional global optimization method in this research.

### 2.1.1 Gaussian process regression

The Gaussian process regression was introduced by (Williams and Carl, 1996, Rasmussen and Williams, 2006). The regression problem is considered for an unknown blackbox function giving the input / output relationship when the observation values \( y = \{y_1, y_2, y_3, ..., y_n\} \) and the input values \( x = \{x_1, x_2, x_3, ..., x_n\} \) are given. The regression model is expressed by the linear combination of basis functions as the following equation.

\[
f(x) = w^T \phi
\]

(2.1.1)

where \( \phi \) represents a basis function and \( w \) represents a parameter vector. When the prior distribution of unknown function follows the Gaussian distribution \( N(w|0, \sigma_w^2) \) from the definition of the Gaussian process, the prior distribution of \( y \) also follows the Gaussian distribution and is expressed by the following equations.

\[
E[y] = E[w^T \phi] = \phi E[w^T] = \phi \mu_w = 0
\]

(2.1.2)

\[
cov[y] = E[yy^T] = E[w^T \phi \phi^T w] = \phi \phi^T E[w^T w] = \sigma_w^2 \phi \phi^T = C
\]

(2.1.3)

where \( C \) is a gram matrix having \( C_{ij} = k(x_i, x_j) = \sigma_w^2 \phi(x_i) \phi(x_j) \) as an element, and \( k(x_i, x_j) \) is referred to as a kernel function. The mean vector \( M^* \) and the covariance matrix \( C^* \) of conditional probability distribution with an unknown sample point \( x^* \) can be expressed by the following equations.

\[
M^* = 0, \quad C^* = \begin{bmatrix} C & k \\ k^T & c \end{bmatrix}
\]

(2.1.4)

\[
c = k(x^*, x^*), \quad k = [k(x_1, x^*) k(x_2, x^*) ... k(x_n, x^*)]^T
\]

(2.1.5)

The conditional distribution of an unknown sample point \( y^* \) is obtained as follows from the conditional distributions of multivariate normal distribution.

\[
\mu_{y^*|y} = k^T C^{-1} y = \mu, \quad \Sigma_{y^*|y} = c - k^T C^{-1} k = \sigma^2
\]

(2.1.6)

where \( \mu_{y^*|y} \) and \( \Sigma_{y^*|y} \) are respectively mean and variance of the posterior distribution of Gaussian distribution. The improvement amount \( I(x) \) is defined as follows in function minimization problems.

\[
I(x) = N(f_{min} - \mu, \sigma^2)
\]

(2.1.7)
where \( f_{\text{min}} \) represents the objective function value of current optimal solution. The expected value of the improvement amount \( E(I(x)) \) is expressed by the following equation.

\[
E(I(x^*)) = \int_{\sigma}^{\infty} I \frac{1}{\sqrt{2\pi} \sigma} \exp \left( -\frac{1}{2} \left( \frac{I - f_{\text{min}} + \mu}{\sigma} \right)^2 \right) \\
= \begin{cases} 
\sigma \phi \left( \frac{f_{\text{min}} - \mu}{\sigma} \right) + (f_{\text{min}} - \mu) \Psi \left( \frac{f_{\text{min}} - \mu}{\sigma} \right) & \text{if } \sigma > 0 \\
0 & \text{if } \sigma = 0
\end{cases}
\]

(2.1.8)

where \( \phi(\cdot) \) represents the probability density function of the standard normal distribution, and \( \Psi(\cdot) \) represents the cumulative distribution function of the standard normal distribution.

With respect to the kernel function, nonlinear regression problems can be solved by mapping data to a high dimensional feature space and executing linear regression in the high dimensional feature space. However, the mapping of data to higher dimensions leads an enormous computational cost. On the other hand, the kernel method enables to directly calculate the inner product in the high dimensional feature space without mapping. Gaussian kernel, ARD squared exponent kernel, ARD Matern 5/2 kernel are usually used as the kernel function (Rasmussen and Williams, 2006). In the Bayesian optimization, it is often assumed that the blackbox function is complex, and the following ARD Matern 5/2 kernel is often used in such cases.

\[
k(x_i, x_j) = \theta_0 \left( 1 + \sqrt{5 \sigma^2 (x_i, x_j)} + \frac{5}{3} \sigma^2 (x_i, x_j) \right) e^{-\sqrt{5 \sigma^2 (x_i, x_j)}}
\]

(2.1.9)

where \( D \) represents the number of design variables. The log-likelihood function is maximized to obtain optimal hyperparameters \( (\theta_0, \theta_1, \ldots, \theta_D) \) of the kernel function. The logarithmic type of the simultaneous probability density function of the multivariate Gaussian distribution (log likelihood function) is expressed as follows.

\[
\ln p(y|\theta) = -\frac{1}{2} \ln |\mathcal{C}| - \frac{n}{2} \ln (2\pi) - \frac{1}{2} y^T \mathcal{C}^{-1} y
\]

(2.1.10)

The differentiated equation of \( \ln p(y|\theta) \) with respect to the parameters \( \theta \) are expressed by the following equation, and the parameters \( \theta \) are optimized by a gradient based method in this research for maximizing the log likelihood function (Rasmussen and Williams, 2006).

\[
\frac{\partial}{\partial \theta_d} \ln p(y|\theta) = -\frac{1}{2} \text{Tr} \left( \mathcal{C}^{-1} \frac{\partial \mathcal{C}}{\partial \theta_d} \right) + \frac{1}{2} y^T \mathcal{C}^{-1} \mathcal{C} \frac{\partial \mathcal{C}}{\partial \theta_d} \mathcal{C}^{-1} y
\]

(2.1.11)

\[\text{2.1.2 Genetic algorithm}\]

GA is one of the metaheuristic methods inspired by the process of natural selection such as selection, crossover and mutation. GA simulates the selection of the individuals by considering the relationship between the applicability to the environment and the probability to survive (Selection). Then, GA simulates the natural evolution of the individuals by generating child individuals from the selected parents individuals (Crossover). In addition, GA maintains the diversity of the population by altering a part of the genes of the child individuals (Mutation). In this research, a tournament selection is used as the selection method. In the tournament selection, some individuals are randomly selected from the population and an individual with the highest fitness among the selected individuals is selected as the parent individual. Since lower fitness individuals are possible to be selected as the parent individual for the next generation, the tournament selection is a useful method from the viewpoint of preserving the diversity of population. Then, the blend crossover (BLX-\( \alpha \)) is used as the crossover method. In the BLX-\( \alpha \), child individuals are generated randomly according to a uniform distribution in a hyper rectangular region which surrounds the two parents individuals (Eshleman and
Schaffer, 1993). The child individuals are generated in a wide range when the distance between the parents individuals is large, and vice versa, so that the BLX-α is a useful method from the viewpoint of both local / global search. In this research, a polynomial mutation is used as the mutation method. In this method, a design variable of a child individual is randomly generated based on the following equations (Deb and Deb, 2014).

\[
x_{ch,d} = x_{par,d} + \left(2r + (1 - r)(1 - \min(x_{par,d}, 1 - x_{par,d}))^{\eta_{m+1}}\right)^{\frac{1}{\eta_{m+1}}} - 1 \quad \text{if } r \leq 0.5
\]

\[
x_{ch,d} = x_{par,d} + 1 - \left(2(1 - r) + (2r - 1)(1 - \min(x_{par,d}, 1 - x_{par,d}))^{\eta_{m+1}}\right)^{\frac{1}{\eta_{m+1}}} \quad \text{if } r > 0.5
\]

where \( r \), \( x_{ch,d} \), \( x_{par,d} \) and \( \eta_m \) are respectively a random number between 0 to 1, design variables of a generated child individual and a parent individual, and a distribution index for the mutation.

2.2 EGODISC

2.2.1 Algorithm of EGODISC

In the conventional Bayesian optimization, an additional sample point is generated by executing GA on the surrogate model. However, when the evaluation of the additional sample point fails, the surrogate model can not be updated. In the next iteration, almost the same sample point is chosen as the additional sample point, which results in the infinite loop in the optimization processes. In this research, this problem is solved by using the SVM classification method (Corinna and Vapnik, 1995). The flowchart of the EGODISC is summarized in Fig. 2. The SVM is one of the supervised learning algorithms that can be used for the binary classification. The SVM classifier is constructed to classify arbitrary locations in the design variables space as being feasible or infeasible regions. The GPs surrogate model is constructed using only the feasible sample points. Then, the SVM classifier is used to predict whether new individuals are feasible or infeasible in the process of GA. When the new individual is predicted as the infeasible one, the infeasible individual is eliminated in the process of GA. The additional sample point obtained by GA with SVM is expected to be feasible by the elimination. Since the SVM classifier is a prediction method, the additional sample point may be in the infeasible regions. However, by the iterative process described above, the accuracy of the SVM classifier is progressively improved around the promising locations in the design variables space, and then the additional sample point will finally be generated in feasible regions. The main novelty of the present algorithm is to use a classification method only to predict whether new individuals are feasible or infeasible in the process of GA. This algorithm can adopt arbitrary classification methods such as neural network / decision tree as well as advanced methods as LightGBM / XGBoost.

![Flowchart of EGODISC](image)

Fig. 2 Flowchart of EGODISC, which corresponds to proposed global optimization method using SVM. Blue and red star points respectively indicate additional sample points that will be explored by conventional Bayesian optimization and proposed EGODISC.
2.2.2 Support vector machine

The support vector machine (SVM) is one of the two-classes classification methods proposed by (Corinna and Vapnik, 1995). It is possible to obtain the optimal separating hyperplane even in the nonlinear data by mapping the data to a high dimensional space by using a kernel method. In addition, SVM separates the two-classes data in the idea of the maximum-margin hyperplane, and it is known to have high accuracy to predict for unknown data. Its separating optimization problem is expressed as follows.

\[
\min \frac{1}{2} \| \omega \|^2 + C \sum_{i=1}^{n} \xi_i \\
\text{s.t.} \quad t_i (\omega^T \varphi(x_i) - b) \geq 1 - \xi_i, \quad \xi_i \geq 0, \quad i = 1, \ldots, n
\]  

(2.2.1)

where \( \omega^T \varphi(x_i) - b, t, \varphi, \omega, b \) and \( \xi \) respectively represent the separating hyperplane, training label, nonlinear separation function, weight vector of the classifier, bias term and penalty of misclassification. \( C \) is the regularization coefficient which adjusts the effect of the penalty of misclassification. Finally, the Lagrange dual problem is given as follows.

\[
\max - \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \alpha_j t_i t_j \varphi(x_i)^T \varphi(x_j) + \sum_{i=1}^{n} \alpha_i \\
\text{s.t.} \quad \sum_{i=1}^{n} \alpha_i t_i = 0, \quad 0 \leq \alpha_i \leq C, \quad i = 1, \ldots, n
\]  

(2.2.2)

where \( \alpha \) represents the Lagrangian multiplier. The inner product of the nonlinear separation function \( \varphi \) can be calculated by using the kernel method without mapping. In this research, the gaussian kernel is used as the kernel method. The regularization coefficient \( C \) and the kernel parameter are determined by using the k-fold cross validation method (Kohavi, 1995).

2.3 Computational fluid dynamics analysis method

An inhouse Computational Fluid Dynamics (CFD) code using a gridless method is used to evaluate aerodynamic objective functions in shape optimization problems. In the gridless method (Ma et al., 2008, Suga and Yamazaki, 2015), computational points are distributed in the calculation domain and a spatial differential term of a physical quantity around the computational point is approximated by a least squares method. Therefore, this method is unnecessary to construct the grid connectivity information and can analyze flow fields around complicated shapes. Two-dimensional inviscid compressible Euler equations are solved by the gridless method. The Euler equations express the conservation law of mass, momentum and energy in inviscid and compressible fluids. The two-dimensional Euler equations are expressed as follows.

\[
\begin{align*}
\frac{\partial W}{\partial t} + \frac{\partial E}{\partial x} + \frac{\partial F}{\partial y} &= 0 \\
W &= \begin{bmatrix} \rho \\ \rho u \\ \rho v \\ e \end{bmatrix}, \quad E = \begin{bmatrix} \rho u \\ \rho u^2 + p \\ \rho uv \\ u(e + p) \end{bmatrix}, \quad F = \begin{bmatrix} \rho v \\ \rho uv \\ \rho v^2 + p \\ v(e + p) \end{bmatrix}
\end{align*}
\]  

(2.3.1)

where \( W \) represents the vector of conservative variables, and \( E \) and \( F \) represent flux vectors. \( \rho, u, v, p, e \) respectively represent fluid density, velocity components in X / Y directions, pressure and total energy per unit volume. The specific heat ratio \( \gamma \) of the ideal gas is set to 1.4.
3. Results and discussion

We solve three optimization problems to verify the usefulness of the developed method of EGODISC. The optimization problems are solved by both of the Bayesian optimization and the EGODISC. First optimization is executed to demonstrate the validity of the EGODISC in a two-dimensional analytic function problem. Second optimization is executed to examine the effect of the EGODISC on a discontinuous optimization problem. Third optimization is executed to examine the effect of the EGODISC on a shape optimization problem with large infeasible regions.

3.1 Optimization problem of Rosenbrock analytic function

In order to verify the usefulness of the EGODISC, infeasible regions are manually generated within the design variables space of the two-dimensional Rosenbrock function (Fig. 3-h) as seven cases indicated in Fig. 3-a - Fig. 3-g. The Rosenbrock function is expressed as follows.

\[
f(x, y) = 100(x' - y'^2)^2 + (1 - x')^2, \quad x' = -2.048 + 4.096x, \quad y' = -2.048 + 4.096y
\]
\[
f_{\text{min}} = f(x, y) = f(0.74414, 0.74414) = 0.0, \quad 0 \leq x, y \leq 1 \tag{3.1.1}
\]

The case 1 – case 4 have different large infeasible regions at a corner of the design variables space. The case 5 and case 6 have two small infeasible regions at different corners of the design variables space. The case 7 has an infeasible region at the center of the design variables space. At the case 2, the true optimal point is included in the infeasible region, and then the feasible optimal point is located at the boundary between the feasible / infeasible regions. The optimization was performed 10 times by both optimization methods of Bayesian and EGODISC. The number of initial sample points is set to 8, and then 200 additional sample points are evaluated. The setting of major parameters of GA is the mutation rate of 0.1, population size of 100 and number of generations of 50. The geometric mean of 10 optimization histories is summarized in Fig. 4. The horizontal axis represents the number of optimization iterations, and the vertical axis represents the difference between the minimum objective function value and the exact solution. The results of all optimizations are summarized in Fig 5. In the Bayesian optimization, additional sample points were generated in the infeasible regions in all the trials, which results in the infinite loop of the optimization process. Therefore, the Bayesian optimizations cannot be performed sufficiently. On the other hand, EGODISC could converge to the exact solution within a small tolerance of 10^{-4} for all the trials by utilizing the additional sample points at the infeasible regions to improve the classification accuracy of SVM. It was confirmed that EGODISC is a useful algorithm to obtain global optimal solutions robustly for optimization problems with various infeasible regions in the design variables space since the accuracy of the SVM classifier is progressively improved in the optimization process.

![Fig. 3](image)

Fig. 3 Infeasible regions on two-dimensional Rosenbrock function. Black filled regions indicate infeasible regions within the design variables space.
3.2 Discontinuous shape optimization problem with bow shock waves

A discontinuous shape optimization problem in two-dimensional supersonic flow is solved in this section. In the supersonic flow condition, it is known that a discontinuous change of aerodynamic objective function occurs due to the effect of bow shock waves. In the surrogate model-based optimization methods, the effect of the discontinuous change propagates in the design variables space and the accuracy of the response surface model will be reduced. To solve this problem, SVM is used to classify the design variables space into the regions with / without bow shock waves, and the sample points with bow shock waves are not used for constructing a surrogate model to prevent the reduction of the accuracy. The generation of bow shock waves is judged by whether subsonic regions exist or not in the flowfield around the airfoil. The airfoil shape is expressed as a symmetrical airfoil, and the shape of one side surface is expressed by a Bezier curve having eight control points arranged in even intervals. The control points at the leading edge and the trailing edge are fixed, and the Y coordinates of the other control points are treated as design variables. The range of all design variables is set to 0<Y<0.12 where the chord length of the airfoil is set as a unit. The negative thickness of airfoil never occurs with this definition of design variables. In this problem, the drag coefficient is minimized with a geometrical constraint for the area of airfoil, which has to be greater than 0.05. The aerodynamic performance is
evaluated at the freestream Mach number of 1.7, angle of attack of 0 degree by using inviscid CFD computations. This optimization problem is solved three times with different sets of initial sample points by both optimization methods of Bayesian and EGODISC. The number of initial sample points is 100 and then new sample points are iteratively added until the number of total evaluations reaches 200. The setting of major parameters of GA is the mutation rate of 0.1, population size of 200 and number of generations of 200. Table 1 shows the objective and constraint function values of the optimal solutions. It is confirmed that the obtained solutions with bow shock waves have much larger drag coefficients which results in discontinuous objective function. Fig. 6 and Fig. 7 show its shapes and pressure distributions. The obtained optimal airfoils have a sharp nose shape and a blunt tail shape which is similar to the airfoil shape minimizing the supersonic drag in (Payot et al., 2017). EGODISC can classify / remove the discontinuous change due to the bow shock waves while the Bayesian optimization has no choice but to deal with the discontinuous change which propagates in the design variables space and reduces the accuracy of the response surface model. Therefore, the performance of the optimal solutions obtained by using EGODISC is better than that obtained by the conventional Bayesian approach. It was confirmed that EGODISC can be effectively used in the present discontinuous shape optimization problem with bow shock waves. In the supersonic flow conditions, there are many other discontinuous objective functions such as sonic boom strength (e.g. A-weighted sound exposure level) which is a noise caused by shock waves generated from an aircraft. EGODISC can also be used for optimization problems to reduce sonic boom strength, and then it is considered that the optimization problems can be efficiently solved.

Table 1 Performance values obtained in discontinuous shape optimization problem

| Trial   | Bayesian optimization | EGODISC | Bow shock waves |
|---------|----------------------|---------|-----------------|
|         | $C_d$ | Area  | $C_d$ | Area  | $C_d$ | Area  |
| Trial 1 | 0.0162 | 0.050 | 0.0157 | 0.050 | 0.0786 | 0.070 |
| Trial 2 | 0.0163 | 0.050 | 0.0152 | 0.050 | 0.0851 | 0.064 |
| Trial 3 | 0.0162 | 0.050 | 0.0156 | 0.050 | 0.0839 | 0.076 |

Fig. 6 Optimal shapes obtained in discontinuous shape optimization problem. Dashed lines indicate control points of Bezier curves. Solid lines indicate airfoil shapes.
3.3 Shape optimization problem with large infeasible regions

A shape optimization problem with large infeasible regions is discussed in this section. The optimization problem is almost same with that of previous second optimization problem. The airfoil shape is expressed in the same way as the second optimization problem while the range of design variables is set as \(-0.02 < Y < 0.12\) in this problem. To solve this optimization problem, SVM is used to classify the design variables space into two regions. One is the infeasible region with bow shock waves or negative thickness cases, and the other is the feasible region. Sample points with bow shock waves or negative thickness are not used for constructing a surrogate model to prevent the reduction of the accuracy as well as to prevent the infinite loop in the optimization process. Table 2 shows the objective and constraint function values of the optimal solutions and the success rate of the CFD evaluations. Although same sets of initial sample points are used in both methods, the numbers of the successful initial sample points are different since the cases with bow shock waves are eliminated in EGODISC. Fig. 8 shows the shapes of optimal designs and those pressure distributions. In the Bayesian optimization, the infinite loops occurred due to the negative thickness of airfoil at the generation of additional sample points in all trials, so that the optimizations could not be performed sufficiently. On the other hand, EGODISC could perform the shape optimization even with large infeasible regions in the design variables space thanks to the appropriate classification via SVM. The performance of the optimal solutions obtained by EGODISC was better than the second optimization results of section 3.2. The shape optimization in this section was performed in wider design variables space than section 3.2 to explore better optimal solutions. Since EGODISC could effectively explore optimal solutions while preventing the problems of bow shock waves and negative thickness airfoil, the performance of the optimal solutions obtained in this section was better than the second optimization results of section 3.2. It was confirmed that EGODISC can be effectively used for the shape optimization problem with large

![Fig. 7: Pressure distributions around obtained optimal shapes. Cases of bow shock waves (right) indicate obtained solutions with maximum drag coefficient](image)
infeasible regions. In aerodynamic shape optimization problems, we can assume many other examples for the large infeasible regions in the design variables space such as the failure of grid generation/deformation and the divergence of CFD evaluations. EGODISC can also be used in such cases, and then it is considered that the optimization problems can be efficiently solved while avoiding the infinite loop in the optimization process.

Table 2 Performance values obtained in shape optimization problem with large infeasible regions

|        | Bayesian optimization |         | EGODISC |
|--------|-----------------------|---------|---------|
|        | $C_d$  | Area   | CFD eval. (init. + addt.) | $C_d$  | Area   | CFD eval. (init. + addt.) |
| Trial 1| 0.0207 | 0.052  | 69/100 + 0/200               | 0.0152 | 0.050  | 47/100 + 153/265           |
| Trial 2| 0.0214 | 0.050  | 72/100 + 0/200               | 0.0149 | 0.050  | 53/100 + 147/355           |
| Trial 3| 0.0173 | 0.050  | 73/100 + 1/200               | 0.0152 | 0.050  | 51/100 + 149/257           |

Fig. 8 Optimal shapes (left) and pressure distributions (right) obtained in shape optimization problem with large infeasible regions. Dashed lines indicate control points of Bezier curves. Solid lines indicate optimal airfoil shapes.
4. Conclusion

In this research, we have developed an efficient global optimization method for discontinuous optimization problems with infeasible regions using classification method (EGODISC) to efficiently solve discontinuous optimization problems as well as optimization problems with large infeasible regions. The conventional Bayesian optimization method falls into an infinite loop in its optimization process by iteratively selecting an additional sample point in the infeasible regions. To tackle this problem, the SVM classifier is constructed to classify arbitrary locations in the design variables space as being feasible or infeasible regions.

We solved three optimization problems to verify the usefulness of the developed method of EGODISC. The optimization problems were solved by both of the Bayesian optimization and the EGODISC for comparison purposes. First optimization was executed to demonstrate the validity of the EGODISC in a two-dimensional analytic function problem. In the Bayesian optimization, additional sample points were generated in the infeasible regions in all the trials, which results in the infinite loop of the optimization process. Therefore, the Bayesian optimizations could not be performed sufficiently. On the other hand, EGODISC could converge to the exact solution within a small tolerance of $10^{-4}$ for all the trials by utilizing the additional sample points at the infeasible regions to improve the classification accuracy of SVM. Second optimization was executed to examine the effect of the EGODISC on a discontinuous optimization problem in a supersonic flow condition. The performance of the optimal solutions obtained by using EGODISC was better than that obtained by the conventional Bayesian approach. It was confirmed that EGODISC could be effectively used in the discontinuous shape optimization problem with bow shock waves. Third optimization was executed to examine the effect of the EGODISC on a shape optimization problem with large infeasible regions, which is defined by taking wider ranges of design variables than the second optimization problem. In the Bayesian optimization, the infinite loops occurred at the exploration of additional sample points in all trials, so that the optimizations could not be performed sufficiently. On the other hand, EGODISC could perform the optimizations successfully even with large infeasible regions in the design variables space. The performance of the optimal solutions obtained by EGODISC was better than that of the second optimization problem. It was confirmed that EGODISC could be effectively used for the shape optimization problem with large infeasible regions.

SVM is widely used as the two-classes classification methods in high dimensional design variables space while it is said that the upper limit of the number of dimensions the Bayesian method can efficiently explore global optimal solutions is about 10 (Wang et al., 2013). When the developed method is straightforwardly applied to high-dimensional optimization problems, it is considered that the Bayesian response surface method will firstly reach the limitation. For the application to high dimensional optimization problems, therefore, the response surface method used in EGODISC has to be modified or newly proposed which can efficiently explore global optimal solutions in high dimensional design variables space. Then, the developed method will apply to more complicated discontinuous and high dimensional optimization problems such as topology optimization problems in supersonic flow conditions. In addition, EGODISC can adopt arbitrary classification methods such as neural network / decision tree as well as advanced methods such as LightGBM / XGBoost. We will also discuss the relationship between the classification methods and complicated optimization problems. Then, the developed method will be applied to other engineering problems such as sonic boom reduction problems, problems with divergence/failure of numerical simulations, and so on.
References

Ban, N., Yamazaki, W., and Kusunose, K., Low-Boom / Low-Drag Design Optimization of Innovative Supersonic Transport Configuration, Journal of Aircraft, Vol. 55, No. 3, (2018), pp. 1071-1081.

Basudhar, A., Dribusch, C., Lacaze, S., and Missoum, S., Constrained Efficient Global Optimization with Support Vector Machines, Structural and Multidisciplinary Optimization, Vol. 46, (2012), pp.201-221.

Corinna, C. and Vladimir, V., Support-Vector Networks, Machine Learning, 20 (1995), pp.273-297

Deb, K. and Deb, D. Analysing mutation schemes for real-parameter genetic algorithms, Int. J. Artif. Intell. Soft. Comput., vol. 4, no. 1, (2014), pp. 1–28.

Esheleman, L. and Schaffer, J. D., Real-Coded Genetic Algorithms and Interval-Schemata, Foundations of Genetic Algorithms 2, (1993), pp.187-202.

Jones, D. R., Schonlau, M., and Welch, W. J., Efficient Global Optimization of Expensive Black-Box Functions, Journal of Global Optimization, Vol.13, (1998), pp.455-492.

Kohavi, R., A study of cross-validation and bootstrap for accuracy estimation and model selection, Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence, 2 (12): 1137–1143, (1995).

Lee, H., Gramacy, R., Linkletter, C., and Gray, G., Optimization Subject to Hidden Constraints via Statistical Emulation, Technical report, Tech Rep UCSC-SOE-10-10, University of California, Santa Cruz, Department of Applied Mathematics and Statistics, (2010).

Liu, B., Zhang, Q., and Gielen, G. G. E., A Gaussian Process Surrogate Model Assisted Evolutionary Algorithm for Medium Scale Expensive Optimization Problems, IEEE Trans. Evol. Comput., vol. 18, no. 2, Apr. (2014), pp. 180–192.

Ma, Z., Chen, H., and Zhou, C., A study of point moving adaptivity in gridless method, Computer methods in applied mechanics and engineering, 197, (21-24), 1926-1937, (2008).

Payot, A. D. J., Rendall, T. C. S. and Allen, C. B., Mixing and Refinement of Design Variables for Geometry and Topology Optimization in Aerodynamics, AIAA Paper 2017-3577, (2017).

Rasmussen, C. E. and Williams, C. K. I., Gaussian Processes for Machine Learning. MIT Press. Cambridge, Massachusetts, (2006).

Rumpfkeil, M. P. and Beran, P., Construction of Multi-Fidelity Surrogate Models for Aerodynamic Databases. Proceedings of the Ninth International Conference on Computational Fluid Dynamics, ICCFD9, Istanbul, Turkey, July 11-15, (2016).

Suga, Y. and Yamazaki, W., Aerodynamic Uncertainty Quantification of Supersonic Biplane Airfoil via Polynomial Chaos Approach, AIAA Paper 2015-1815, (2015).

Wang, Z., Zoghi, M., Hutter, F., Matheson, D., and Freitas, N., Bayesian Optimization in High Dimensions via Random Embeddings, In International Joint Conference on Artificial Intelligence, (2013).

Williams, C. K. I., and Carl, E. R., Gaussian Processes for Regression, MIT press, Advances in Neural Information Processing Systems 8, (1996), pp.514-520.