WORST CASE OPTIMIZATION USING CHEBYSHEV INEQUALITY

Kiyoharu Tagawa
School of Science and Engineering, Kindai University, Japan
tagawa@info.kindai.ac.jp

Abstract In real-world optimization problems, a wide range of uncertainties have to be taken into account. The presence of uncertainty leads to different results for repeated evaluations of the same solution. Therefore, users may not always be interested in the so-called best solutions. In order to find the robust solutions which are evaluated based on the predicted worst case, Worst Case Optimization Problem (WCOP) is formulated by using Chebyshev inequality from samples. Besides, a new evolutionary algorithm based on Differential Evolution is proposed to solve WCOP efficiently. The difference between the nominal solutions and the robust solutions is demonstrated through engineering design problems.

Keywords: Differential evolution, Prediction interval, Robust optimization.

1. Introduction

Considering the worst case is important, in particular, if the decision maker is very risk averse, or if the stakes are high. Therefore, we formulate Worst Case Optimization Problem (WCOP) to obtain the robust solution evaluated based on the predicted worst case. In order to predict the worst case from samples, we employ the upper bounds of the uncertain function values in the objective and constraints. In WCOP, we can specify the probability of risk with a significance level. We also show the minimum sample size necessary for the significance level.

We propose an Evolutionary Algorithm (EA) based on Differential Evolution (DE) [13] to solve WCOP efficiently. In order to save the number of samplings, or the number of repeated evaluations of the same solution, for calculating the upper bounds, the accumulative sampling [12] and the U-cut [21] are introduced into DE. Finally, we apply the proposed methodology to practical engineering design problems.

2. Related Work

Optimization in uncertain environments is an active research area in EAs [9]. A number of EAs have been reported to solve single- and multi-objective optimization problems under uncertainties [2, 3]. However, in most of the methods, the average performance has been considered.

Generally speaking, approaches for the worst case optimization can be categorized into two classes. Let $\boldsymbol{x} \in \Re^D$ be the vector of design variables x_j, $j = 1, \cdots, D$. In non-statistical approaches for the worst case optimization [17, 24], the uncertainty is given by the upper and lower bounds of each design variable x_j. Then the worst case of the function included in an objective or a constraint is evaluated based on vertex analysis. On the other hand, in statistical approaches for the worst case optimization, the design variable x_j is given by a random variable. WCOP described in this paper is a statistical approach.

In our previous papers [20, 21], we considered another WCOP in which the objective function $f(\boldsymbol{x})$ was subject to noise: $\mathcal{F}(\boldsymbol{x}) = f(\boldsymbol{x}) + \epsilon$. Supposing that noise ϵ was distributed normally, we predicted the upper bound of the objective function value $\mathcal{F}(\boldsymbol{x})$ with the quantile of the normal distribution. The new methodology proposed in this paper can handle two types of uncertainties: $\mathcal{F}(\boldsymbol{x}) = f(\boldsymbol{x} + \boldsymbol{\delta}) + \epsilon$, where $\boldsymbol{\delta} \in \Re^D$ denotes the vector of perturbations, or random variables. Besides, it doesn't presume particular distribution for random variables.

DE [5, 13] is a recently developed EA. DE is arguably one of the most powerful stochastic real-parameter optimization algorithms in current use. Due to its simple but powerful searching capability, DE has been used successfully in many scientific and engineering applications [19, 25]. Some variants of DE have also been reported for optimization problems under uncertainties [11, 14, 15]. However, most of them don't evaluate solutions with the worst case, but with the average performance.

3. Prediction Interval from Samples

Chebyshev inequality [23] in (1) is well known to statisticians. If $\mathcal{F}$ is a random variable with mean μ and variance σ^2, then for $\lambda > 1$,

$$P_r(|\mathcal{F} - \mu| \geq \lambda\,\sigma) \leq \frac{1}{\lambda^2}. \tag{1}$$

Chebyshev inequality has great utility because it can be applied to completely arbitrary distributions. Let $\alpha = 1/\lambda^2$ be a significance level. Then, from (1), we can derive a prediction interval as

$$P_r([\mu - \lambda\,\sigma,\ \mu + \lambda\,\sigma] \ni \mathcal{F}) \geq 1 - \alpha. \tag{2}$$

Since the mean μ and variance σ^2 in (1) are usually unknown, they have to be estimated from samples. Let $\mathcal{F}^1$, $\mathcal{F}^2$, $\cdots$, $\mathcal{F}^N$ and $\mathcal{F}^{N+1}$ be weakly exchangeable samples from some unknown distribution such that $P_r(\mathcal{F}^1 = \mathcal{F}^2 \cdots = \mathcal{F}^N = \mathcal{F}^{N+1}) = 0$. From a set of observed N samples, the sample mean $\overline{F}$ and variance s^2 are obtained as

$$\overline{F} = \frac{1}{N} \sum_{n=1}^{N} \mathcal{F}^n, \tag{3}$$

$$s^2 = \frac{1}{N-1} \sum_{n=1}^{N} (\mathcal{F}^n - \overline{F})^2. \tag{4}$$

By using the sample mean $\overline{F}$ and variance s^2, Saw et al. [18] have presented Chebyshev inequality estimated from N samples as

$$P_r\left(|\mathcal{F} - \overline{F}| \geq \lambda \sqrt{\frac{N+1}{N}}\, s\right) \leq \frac{1}{N+1} \left\lfloor \frac{(N+1)\,(N-1+\lambda^2)}{N\,\lambda^2} \right\rfloor, \tag{5}$$

where $\lfloor\cdot\rfloor$ denotes the floor function. By using the upper bound of the floor function, or its argument, we simplify the right-hand side as

$$P_r\left(|\mathcal{F} - \overline{F}| \geq \lambda \sqrt{\frac{N+1}{N}}\, s\right) \leq \frac{N-1+\lambda^2}{N\,\lambda^2}. \tag{6}$$

Now, we will derive a prediction interval from (6). We set κ as

$$\kappa = \lambda \sqrt{\frac{N+1}{N}}. \tag{7}$$

By using κ in (7), Chebyshev inequality in (6) is rewritten as

$$P_r(|\mathcal{F} - \overline{F}| \geq \kappa\, s) \leq \frac{N^2 - 1 + N\,\kappa^2}{N^2\,\kappa^2}. \tag{8}$$

Let us suppose that a significance level α is given as

$$\alpha = \frac{N^2 - 1 + N\,\kappa^2}{N^2\,\kappa^2}. \tag{9}$$

From (8) and (9), we can derive the prediction interval as

$$P_r([\overline{F} - \kappa\, s,\ \overline{F} + \kappa\, s] \ni \mathcal{F}) = P_r([\mathcal{F}^L,\ \mathcal{F}^U] \ni \mathcal{F}) \geq 1 - \alpha. \tag{10}$$

From (9), κ can be specified by the significance level α as

$$\kappa = \sqrt{\frac{N^2 - 1}{N\,(\alpha\, N - 1)}}. \tag{11}$$

Since $\kappa > 1$ holds from (7), the minimum sample size $N_{\min}$ is

$$N_{\min} = \left\lfloor \frac{1}{\alpha} + 1 \right\rfloor. \tag{12}$$

Incidentally, from (11), the lower bound of κ can be estimated as

$$\lim_{N \to \infty} \kappa = \lim_{N \to \infty} \sqrt{\frac{N^2 - 1}{N\,(\alpha\, N - 1)}} = \sqrt{\frac{1}{\alpha}}. \tag{13}$$

In this paper, we call $\mathcal{F}^U$ in (10) the upper bound of $\mathcal{F}$. If we have enough number of samples $\mathcal{F}^1, \mathcal{F}^2, \cdots, \mathcal{F}^n, \cdots, \mathcal{F}^N$ that satisfies the condition $N \geq N_{\min}$ for a given significance level α, then we can calculate the upper bound $\mathcal{F}^U = \overline{F} + \kappa\, s$ of $\mathcal{F}$ from (3), (4), and (11).

4. Worst Case Optimization Problem (WCOP)

The traditional constrained optimization problem can be stated as

$$\left[\begin{array}{ll} \text{minimize} & f(\boldsymbol{x}) = f(x_1, x_2, \cdots, x_D) \\ \text{subject to} & g_m(\boldsymbol{x}) \leq 0,\ m \in \boldsymbol{I}_M = \{1, \cdots, M\} \\ & x_j^L \leq x_j \leq x_j^U,\ j = 1, \cdots, D \end{array}\right. \tag{14}$$

with the vector $\boldsymbol{x} = (x_1, \cdots, x_D) \in \Re^D$ of D design variables.

When uncertainties are introduced in the above optimization problem, functions $f(\boldsymbol{x})$ and $g_m(\boldsymbol{x})$ included in (14) are modified as

$$\left[\begin{array}{l} \mathcal{F}(\boldsymbol{x}) = f(\boldsymbol{x} + \boldsymbol{\delta}) + \epsilon = f(x_1 + \delta_1, \cdots, x_D + \delta_D) + \epsilon, \\ \mathcal{G}_m(\boldsymbol{x}) = g_m(\boldsymbol{x} + \boldsymbol{\delta}) + \epsilon,\ m \in \boldsymbol{I}_M = \{1, \cdots, M\}. \end{array}\right. \tag{15}$$

where $\delta_j \in \Re$, $j = 1, \cdots, D$ denotes the disturbance to each $x_j \in \Re$, and $\epsilon \in \Re$ denotes a noise. Both $\boldsymbol{\delta}$ and ϵ are random variables.

The objective function value $\mathcal{F}(\boldsymbol{x})$ depends on a solution $\boldsymbol{x} \in \Re^D$. However, every time the solution is evaluated, different function values may be returned. Therefore, we predict the worst case of the solution $\boldsymbol{x} \in \Re^D$ by using the upper bound in (10). We evaluate the solution N ($N \geq N_{\min}$) times and make a sample set $\{\mathcal{F}^n(\boldsymbol{x}) \mid n = 1, \cdots, N\}$. Then, from the sample set, we calculate the upper bound $\mathcal{F}^U(\boldsymbol{x})$ of the objective function value $\mathcal{F}(\boldsymbol{x})$ for a given significance level α. In the same way, we calculate the upper bounds $\mathcal{G}_m^U(\boldsymbol{x})$, $m \in \boldsymbol{I}_M$ of the uncertain function values $\mathcal{G}_m(\boldsymbol{x})$ included in the inequality constraints.

The Worst Case Optimization Problem (WCOP) is formulated as

$$\left[\begin{array}{ll} \text{minimize} & \mathcal{F}^U(\boldsymbol{x}) = \mathcal{F}^U(x_1, x_2, \cdots, x_D) \\ \text{subject to} & \mathcal{G}_m^U(\boldsymbol{x}) \leq 0,\ m \in \boldsymbol{I}_M = \{1, \cdots, M\} \\ & x_j^L \leq x_j \leq x_j^U,\ j = 1, \cdots, D \end{array}\right. \tag{16}$$

with an arbitrary significance level α $(0 < \alpha < 1)$.

5. Differential Evolution for WCOP

5.1 Conventional Differential Evolution

Conventional DE can be applied to WCOP in (16). DE works by building a population $\boldsymbol{P}$ of vectors which is a set of possible solutions to WCOP. The initial population $\boldsymbol{x}_i \in \boldsymbol{P}$, $i = 1, \cdots, N_P$ is generated randomly. Evaluating each vector $\boldsymbol{x}_i \in \boldsymbol{P}$ N times, the upper bounds in (16), namely $\mathcal{F}^U(\boldsymbol{x})$ and $\mathcal{G}_m^U(\boldsymbol{x})$, $m \in \boldsymbol{I}_M$, are calculated. Classical DE [13] uses three control parameters: population size N_P, scale factor S_F, and crossover rate C_R. However, we employ a self-adapting mechanism of S_F and C_R [1]. Therefore, every vector $\boldsymbol{x}_i \in \boldsymbol{P}$ has its own parameter values $S_{F,i}$ and $C_{R,i}$ initialized as $S_{F,i} = 0.5$ and $C_{R,i} = 0.9$.

Within a generation of a population, each vector $\boldsymbol{x}_i \in \boldsymbol{P}$ is assigned to the target vector in turn. According to the following strategy of DE, the trial vector $\boldsymbol{u} \in \Re^D$ is generated from the target vector $\boldsymbol{x}_i \in \boldsymbol{P}$.

Through the recommendation [1], S_F and C_R are decided respectively from the parameters $S_{F,i}$ and $C_{R,i}$ of the target vector $\boldsymbol{x}_i \in \boldsymbol{P}$ as

$$S_F = \begin{cases} 0.1 + \text{rand}_1\ 0.9, & \text{if} \quad \text{rand}_2 < 0.1 \\ S_{F,i}, & \text{otherwise}, \end{cases} \tag{17}$$

$$C_R = \begin{cases} \text{rand}_3, & \text{if} \quad \text{rand}_4 < 0.1 \\ C_{R,i}, & \text{otherwise}, \end{cases} \tag{18}$$

where $\text{rand}_n \in [0,\ 1]$ is a uniformly distributed random number.

Except for the target vector $\boldsymbol{x}_i \in \boldsymbol{P}$, three other distinct vectors $\boldsymbol{x}_{r1}$, $\boldsymbol{x}_{r2}$, and $\boldsymbol{x}_{r3}$ $(i \neq r1 \neq r2 \neq r3)$ are selected randomly from $\boldsymbol{P}$. From the four vectors, the j-th element $u_j \in \Re$ of $\boldsymbol{u} \in \Re^D$ is generated as

$$u_j = \begin{cases} x_{r1,j} + S_F\,(x_{r2,j} - x_{r3,j}), & \text{if} \quad (\text{rand}_j < C_R) \vee (j = j_r) \\ x_{i,j}, & \text{otherwise}, \end{cases} \tag{19}$$

where the index of design variable $j_r \in [1,\ D]$ is selected randomly.

According to the direct constraint handling [13], the trial vector $\boldsymbol{u}$ is compared with the target vector $\boldsymbol{x}_i \in \boldsymbol{P}$. Exactly, if at least one of the following criteria is satisfied, $\boldsymbol{u}$ is judged to be better than $\boldsymbol{x}_i \in \boldsymbol{P}$.

- $\boldsymbol{u}$ is feasible $(\forall\, m \in \boldsymbol{I}_M;\, \mathcal{G}_m^U(\boldsymbol{u}) \leq 0)$ and $\mathcal{F}^U(\boldsymbol{u}) \leq \mathcal{F}^U(\boldsymbol{x}_i)$.
- $\boldsymbol{u}$ is feasible $(\forall\, m \in \boldsymbol{I}_M;\, \mathcal{G}_m^U(\boldsymbol{u}) \leq 0)$ and $\exists\, m \in \boldsymbol{I}_M;\, \mathcal{G}_m^U(\boldsymbol{x}_i) > 0$.
- $\boldsymbol{u}$ is infeasible $(\exists\, m \in \boldsymbol{I}_M;\, \mathcal{G}_m^U(\boldsymbol{u}) > 0)$, but $\forall\, m \in \boldsymbol{I}_M;\, \max\{\mathcal{G}_m^U(\boldsymbol{u}),\, 0\} \leq \max\{\mathcal{G}_m^U(\boldsymbol{x}_i),\, 0\}$.

As a result, if $\boldsymbol{u}$ is better than $\boldsymbol{x}_i \in \boldsymbol{P}$, $\boldsymbol{u}$ takes the place of $\boldsymbol{x}_i \in \boldsymbol{P}$. Besides, S_F in (17) and C_R in (18) are substituted for $S_{F,i}$ and $C_{R,i}$. If $\boldsymbol{u}$ is not better than $\boldsymbol{x}_i \in \boldsymbol{P}$, $\boldsymbol{u}$ is discarded. According to the asynchronous generation alternation model [7, 22], we use only one population $\boldsymbol{P}$.

5.2 Proposed Differential Evolution

Even though DE is applicable to WCOP, the multiple samplings of each vector to calculate the upper bounds are still expensive. Therefore, in order to examine a lot of vectors within a limited number of samplings, we introduce two techniques into the above DE, namely the accumulative sampling [12] and U-cut [21]. Both of the techniques have been contrived to solve multi-objective optimization problems. In this paper, we modify those techniques to solve WCOP efficiently. The proposed DE is called DEAU (DE with Accumulative sampling and U-cut), which can allocate the computing budget only to the promising solutions of WCOP.

Firstly, the accumulative sampling evaluates each $\boldsymbol{x}_i \in \boldsymbol{P}$ $N_{\min}$ times for calculating the upper bounds ($\mathcal{F}^U(\boldsymbol{x}_i)$ and $\mathcal{G}_m^U(\boldsymbol{x}_i)$, $m \in \boldsymbol{I}_M$), where $N_{\min}$ is the minimum sample size in (12). Thereafter, it re-evaluates $\boldsymbol{x}_i \in \boldsymbol{P}$ $N_{\min}$ times for taking additional samples and updates the upper bounds at regular generation intervals G_{int}. Let N_G be the sum total of evaluations of $\boldsymbol{x}_i \in \boldsymbol{P}$ depending on the current generation.

Secondly, the U-cut can judge hopeless trial vectors $\boldsymbol{u}$ only by few samplings and discard them. When a newborn $\boldsymbol{u}$ is compared with the target vector $\boldsymbol{x}_i \in \boldsymbol{P}$ at a generation, U-cut evaluates $\boldsymbol{u}$ N_G times for taking samples: $\mathcal{F}^n(\boldsymbol{u})$ and $\mathcal{G}_m^n(\boldsymbol{u})$, $m \in \boldsymbol{I}_M$, $n = 1 \cdots, N_G$. However, U-cut takes and examines those samples one by one. If at least one of the following criteria is satisfied along the way, $\boldsymbol{u}$ is judged to be worse than $\boldsymbol{x}_i \in \boldsymbol{P}$ and discarded immediately. That is because $\mathcal{F}^n(\boldsymbol{u}) < \mathcal{F}^U(\boldsymbol{u})$ and $\mathcal{G}_m^n(\boldsymbol{u}) < \mathcal{G}_m^U(\boldsymbol{u})$ are expected with a high probability $(1-\alpha)$.

- $\boldsymbol{x}_i$ is feasible ($\forall\, m \in \boldsymbol{I}_M;\, \mathcal{G}_m^U(\boldsymbol{x}_i) \le 0$) and $\mathcal{F}^U(\boldsymbol{x}_i) \le \mathcal{F}^n(\boldsymbol{u})$.
- $\boldsymbol{x}_i$ is feasible ($\forall\, m \in \boldsymbol{I}_M;\, \mathcal{G}_m^U(\boldsymbol{x}_i) \le 0$) and $\exists\, m \in \boldsymbol{I}_M;\, \mathcal{G}_m^n(\boldsymbol{u}) > 0$.
- $\boldsymbol{x}_i$ is infeasible ($\exists\, m \in \boldsymbol{I}_M;\, \mathcal{G}_m^U(\boldsymbol{x}_i) > 0$), but $\forall\, m \in \boldsymbol{I}_M;\, \max\{\mathcal{G}_m^U(\boldsymbol{x}_i),\, 0\} \le \max\{\mathcal{G}_m^n(\boldsymbol{u}),\, 0\}$.

If $\boldsymbol{u}$ survives, $\mathcal{F}^U(\boldsymbol{u})$ and $\mathcal{G}_m^U(\boldsymbol{u})$, $m \in \boldsymbol{I}_M$ are calculated. Then $\boldsymbol{u}$ is compared with $\boldsymbol{x}_i \in \boldsymbol{P}$ again in the same way with conventional DE.

6. Numerical Experiments

We demonstrated the proposed methodology using one test problem and two engineering design problems. Those were originally stated as

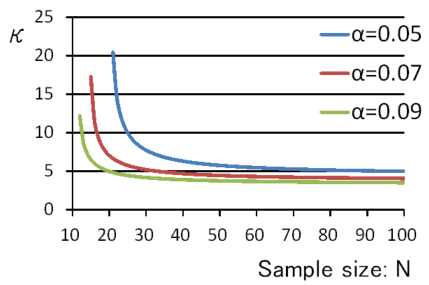

Figure 1: Change of κ for N and α.

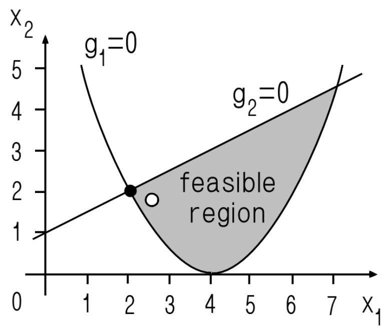

Figure 2: Feasible region of (20).

constrained optimization problems as shown in (14). We extended each of them into WCOP in which the normally distributed disturbances $\delta_j \sim \mathcal{N}(0,\, 0.01^2)$, $j = 1, \cdots, D$ were added to design variables x_j on the assumption that the machining error of tool was inevitable.

The conventional DE and the proposed DEAU were coded by the Java language. The population size was chosen as $N_P = 10\,D$. As the termination condition, the total number of samplings was limited to $D \times 10^5$. The generation interval $G_{int} = 10$ was used for DEAU.

6.1 Test Problem

The following test problem has a nominal solution $\boldsymbol{x}^\star = (2,\, 2)$ that realizes function values $f(\boldsymbol{x}^\star) = 4$, $g_1(\boldsymbol{x}^\star) = 0$, and $g_2(\boldsymbol{x}^\star) = 0$.

$$\left[\begin{array}{ll} \text{minimize} & f(\boldsymbol{x}) = f(x_1,\, x_2) = x_1^2 + (x_2 - 2)^2 \\ \text{subject to} & g_1(\boldsymbol{x}) = g_1(x_1,\, x_2) = (x_1 - 4)^2 - 2\,x_2 \le 0 \\ & g_2(\boldsymbol{x}) = g_2(x_1,\, x_2) = -x_1 + 2\,x_2 - 2 \le 0 \\ & -5 \le x_1,\; x_2 \le 10 \end{array}\right. \tag{20}$$

The test problem in (20) was extended to WCOP with $\alpha = 0.05$, 0.07, and 0.09. Figure 1 shows the value of κ in (11) that depends both on N ($N \ge N_{\min}$) and α. In order to obtain a robust solution $\boldsymbol{x}^\circ \in \Re^2$ of each WCOP, the proposed DEAU was used. Table 1 compares the robust solution $\boldsymbol{x}^\circ$ of each WCOP with the nominal solution $\boldsymbol{x}^\star$. From Table 1, $\mathcal{F}^U(\boldsymbol{x}^\star)$ is smaller than $\mathcal{F}^U(\boldsymbol{x}^\circ)$ in all WCOPs. However, the nominal solution is infeasible, while the robust ones are feasible.

Figure 2 illustrates the feasible region in the design variable space of the test problem. The nominal solution $\boldsymbol{x}^\star$ denoted by "•" exists on the boundary of the feasible region. On the other hand, every robust solution denoted by "∘" exists on the inside of the feasible region.

From Table 1 and Fig. 2, the robust solution $\boldsymbol{x}^{\circ}$ parts from the boundary of the feasible region as the significance level α becomes small.

Table 1: Nominal solution and robust solutions of test problem.

α	solution	x_1	x_2	$\mathcal{F}^U(\boldsymbol{x})$	$\mathcal{G}_1^U(\boldsymbol{x})$	$\mathcal{G}_2^U(\boldsymbol{x})$
0.05	nominal: $\boldsymbol{x}^{\star}$	2.000	2.000	4.188	0.188	0.104
	robust: $\boldsymbol{x}^{\circ}$	2.061	1.979	4.444	−0.016	−0.001
0.07	nominal: $\boldsymbol{x}^{\star}$	2.000	2.000	4.157	0.154	0.086
	robust: $\boldsymbol{x}^{\circ}$	2.048	1.980	4.359	−0.002	−0.001
0.09	nominal: $\boldsymbol{x}^{\star}$	2.000	2.000	4.138	0.133	0.075
	robust: $\boldsymbol{x}^{\circ}$	2.045	1.981	4.323	−0.009	−0.006

From the value of κ in Fig. 1, the sample size of every function was set to $N = 150$.

Table 2: Nominal solution and robust solution of pressure vessel design problem.

solution	$\mathcal{F}^U(\boldsymbol{x})$	$\mathcal{G}_1^U(\boldsymbol{x})$	$\mathcal{G}_2^U(\boldsymbol{x})$	$\mathcal{G}_3^U(\boldsymbol{x})$	$\mathcal{G}_4^U(\boldsymbol{x})$
nominal: $\boldsymbol{x}^{\star}$	6259.854	0.049	0.047	2998.767	−39.974
robust: $\boldsymbol{x}^{\circ}$	6775.558	−0.004	−0.006	−1996.695	−40.421

6.2 Pressure Vessel Design Problem

The following problem is taken from [10]. A cylindrical vessel is capped at both ends by hemispherical heads as shown in Fig. 3. There are four design variables: thickness of the vessel x_1, thickness of the head x_2, inner radius x_3, and length of the vessel without heads x_4. The objective is to minimize the total cost $f(\boldsymbol{x})$ including the cost of the material, forming, and welding. The problem is stated as follows:

$$\left[\begin{array}{ll} \text{minimize} & f(\boldsymbol{x}) = 0.6224\, x_1\, x_3\, x_4 + 1.7781\, x_2\, x_3^2 \\ & \quad +3.1661\, x_1^2\, x_4 + 19.84\, x_1^2\, x_3 \\ \text{subject to} & g_1(\boldsymbol{x}) = -x_1 + 0.0193\, x_3 \le 0 \\ & g_2(\boldsymbol{x}) = -x_2 + 0.00954\, x_3 \le 0 \\ & g_3(\boldsymbol{x}) = -\pi\, x_3^2\, x_4 - \dfrac{4}{3}\pi\, x_3^3 + 1296000 \le 0 \\ & g_4(\boldsymbol{x}) = x_4 - 240 \le 0 \\ & 0.0625 \le x_1,\ x_2 \le 6.1875,\ 10 \le x_3,\ x_4 \le 200 \end{array}\right. \tag{21}$$

Among the results of many optimization algorithms applied to the engineering design problem in (21) [4, 8], Garg [8] found the best nominal solution $\boldsymbol{x}^\star \in \Re^4$ with $f(\boldsymbol{x}^\star) \approx 5885.403$, $g_1(\boldsymbol{x}^\star) \approx -1.399 \times 10^{-6}$, $g_2(\boldsymbol{x}^\star) \approx -2.837 \times 10^{-6}$, $g_3(\boldsymbol{x}^\star) \approx -1.141$, and $g_4(\boldsymbol{x}^\star) \approx -40.019$.

The engineering design problem in (21) was extended to WCOP with $\alpha = 0.05$. By solving WCOP with DEAU, a robust solution $\boldsymbol{x}^\circ \in \Re^4$ could be obtained. Table 2 compares the robust solution $\boldsymbol{x}^\circ$ with the nominal solution $\boldsymbol{x}^\star$. From Table 2, the nominal solution is infeasible. Furthermore, since the value of $\mathcal{G}_3^U(\boldsymbol{x}^\star)$ is very large, the constraint $g_3(\boldsymbol{x}) \le 0$ in (21) is sensitive to the disturbances of design variables.

6.3 Welded Beam Design Problem

The following problem is taken from [16]. The welded beam structure is shown in Fig. 4. There are four design variables: thickness of the weld x_1, length of the welded joint x_2, width of the beam x_3, and thickness of the beam x_4. The length of beam is a constant $L = 14$. The objective is to minimize the fabricating cost of the welded beam subject to six constraints on shear stress $\tau(\boldsymbol{x})$, bending stress in the beam $h(\boldsymbol{x})$, end deflection on the beam $q(\boldsymbol{x})$, buckling load on the bar $\rho_c(\boldsymbol{x})$, and side constraints. The engineering design problem is stated as follows:

$$\left[\begin{array}{ll} \text{minimize} & f(\boldsymbol{x}) = 1.10471\, x_1^2\, x_2 + 0.04811\, x_3\, x_4\, (L + x_2) \\ \text{subject to} & g_1(\boldsymbol{x}) = \tau(\boldsymbol{x}) - 13600 \le 0 \\ & g_2(\boldsymbol{x}) = h(\boldsymbol{x}) - 3 \times 10^4 \le 0 \\ & g_3(\boldsymbol{x}) = x_1 - x_4 \le 0 \\ & g_4(\boldsymbol{x}) = 0.125 - x_1 \le 0 \\ & g_5(\boldsymbol{x}) = q(\boldsymbol{x}) - 0.25 \le 0 \\ & g_6(\boldsymbol{x}) = \rho - \rho_c(\boldsymbol{x}) \le 0 \\ & 0.1 \le x_1 \le 2,\ 0.1 \le x_2,\ x_3 \le 10,\ 0.1 \le x_4 \le 2 \end{array}\right. \tag{22}$$

$$\left[\begin{array}{l} J(\boldsymbol{x}) = \sqrt{2}\, x_1\, x_2 \left(\dfrac{x_2^2}{12} + \left(\dfrac{x_1 + x_3}{2} \right)^2 \right), \quad R(\boldsymbol{x}) = \sqrt{\dfrac{x_2^2}{4} + \left(\dfrac{x_1 + x_3}{2} \right)^2} \\ \rho = 6000, \quad \tau_1(\boldsymbol{x}) = \dfrac{\rho}{\sqrt{2}\, x_1\, x_2}, \quad \tau_2(\boldsymbol{x}) = \rho \left(L + \dfrac{x_2}{2} \right) \dfrac{R(\boldsymbol{x})}{J(\boldsymbol{x})} \\ \tau(\boldsymbol{x}) = \sqrt{\tau_1(\boldsymbol{x})^2 + \dfrac{\tau_1(\boldsymbol{x})\, \tau_2(\boldsymbol{x})\, x_2}{R(\boldsymbol{x})} + \tau_2(\boldsymbol{x})^2}, \quad q(\boldsymbol{x}) = \dfrac{\rho\, L^3}{(75 \times 10^5)\, x_3^3\, x_4} \\ h(\boldsymbol{x}) = \dfrac{6\, \rho\, L}{x_3^2\, x_4}, \quad \rho_c(\boldsymbol{x}) = \dfrac{4013000\, \sqrt{10}\, x_3\, x_4^3}{L^2} \left(1 - \dfrac{\sqrt{0.625}\, x_3}{2\, L} \right) \end{array}\right.$$

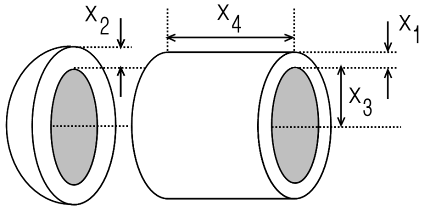

Figure 3: Pressure vessel.

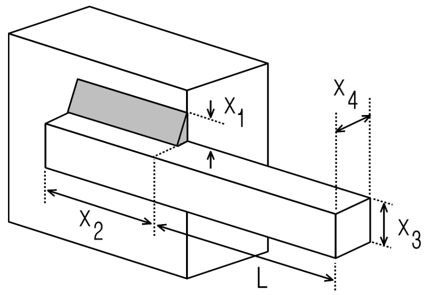

Figure 4: Welded beam.

Table 3: Nominal solution and robust solution of welded beam design problem.

	$\mathcal{F}^U(\boldsymbol{x})$	$\mathcal{G}_1^U(\boldsymbol{x})$	$\mathcal{G}_2^U(\boldsymbol{x})$	$\mathcal{G}_3^U(\boldsymbol{x})$	$\mathcal{G}_4^U(\boldsymbol{x})$	$\mathcal{G}_5^U(\boldsymbol{x})$	$\mathcal{G}_6^U(\boldsymbol{x})$
$\boldsymbol{x}^\star$	2.784	2701.887	5669.847	0.064	−0.072	−0.231	3292.867
$\boldsymbol{x}^\circ$	3.573	−85.129	−25.285	−0.007	−0.063	−0.233	−260.091

Garg [8] found the nominal solution $\boldsymbol{x}^\star$ of the problem in (22) with $f(\boldsymbol{x}^\star) \approx 2.380$, $g_1(\boldsymbol{x}^\star) \approx -0.100$, $g_2(\boldsymbol{x}^\star) \approx -1.170$, $g_3(\boldsymbol{x}^\star) \approx -6.84 \times 10^{-6}$, $g_4(\boldsymbol{x}^\star) \approx -0.119$, $g_5(\boldsymbol{x}^\star) \approx -0.234$, and $g_6(\boldsymbol{x}^\star) \approx -0.071$.

The engineering design problem in (22) was extended to WCOP with $\alpha = 0.05$. Besides disturbances $\delta_j \sim \mathcal{N}(0,\, 0.01^2)$ to design variables x_j, we assumed that the length of beam L was also a random value such as $L \sim \mathcal{N}(14,\, 0.01^2)$. By solving WCOP with DEAU, a robust solution $\boldsymbol{x}^\circ \in \Re^4$ could be obtain. Table 3 compares the robust solution $\boldsymbol{x}^\circ$ with the nominal solution $\boldsymbol{x}^\star$. From Table 3, the nominal solution is infeasible. Furthermore, the constraints $g_1(\boldsymbol{x}) \leq 0$, $g_2(\boldsymbol{x}) \leq 0$, and $g_6(\boldsymbol{x}) \leq 0$ in (22) seem to be very sensitive to the disturbances.

6.4 Comparison of DEAU with DE

The performance of DEAU was compared with the conventional DE. DEAU and DE were applied 50 times to each of the above three WCOPs with different initial populations. DE evaluated every solution $\boldsymbol{x}_i \in \boldsymbol{P}$ $N = 200$ times. The results of numerical experiments were summarized in Table 4 and Table 5. Table 4 shows the objective function value $\mathcal{F}^U(\boldsymbol{x})$ averaged over 50 runs. Similarly, Table 5 shows the number of solutions examined through the process of optimization.

From Table 4, we can confirm that DEAU outperforms conventional DE. As you can see in Table 5, by using the accumulative sampling and

Table 4: Objective function value $\mathcal{F}^U(\boldsymbol{x})$ of WCOP.

α	method	WCOP1	WCOP2	WCOP3
0.05	DEAU	4.434	6759.823	3.642
	DE	4.458	7320.667	3.796
0.07	DEAU	4.353	6615.113	3.411
	DE	4.388	7192.097	3.552
0.09	DEAU	4.306	6514.112	3.269
	DE	4.339	7003.265	3.432

- WCOP1: Test problem in (20) extended to WCOP
- WCOP2: Pressure vessel design problem in (21) extended to WCOP
- WCOP3: Welded beam design problem in (22) extended to WCOP

Table 5: Number of examined solutions.

α	method	WCOP1	WCOP2	WCOP3
0.05	DEAU	1932.0	5746.4	5224.0
	DE	1000.0	2000.0	2000.0
0.07	DEAU	2312.8	6886.4	6312.8
	DE	1000.0	2000.0	2000.0
0.09	DEAU	2683.6	8088.0	7424.0
	DE	1000.0	2000.0	2000.0

U-cut, DEAU has examined more solutions than DE in all WCOPs. As a result, DEAU could find better robust solutions than DE.

7. Conclusion

We formulated WCOP in which the worst case was estimated by using the upper bound of distributed function values. The upper bound was derived from Chebyshev inequality. Therefore, the upper bound could guarantee the worst case of unknown distributions with a significance level. For solving WCOP efficiently, we proposed a new algorithm called DEAU. We demonstrated the usefulness of the proposed methodology through one test problem and two engineering design problems.

Future work will include in-depth assessments of the methodology on a broad range of optimization problems under uncertainties. Besides, we would like to transform WCOP based on prediction intervals into various formulations including multi-objective optimization problems [6].

References

[1] J. Brest, S. Greiner, B. Bošković, M. Merink, and V. Žumer. Self-adapting control parameters in differential evolution: a comparative study on numerical benchmark problems. *IEEE Transactions on Evolutionary Computation*, 10(6):646–657, 2006.

[2] R. R. Chan and S. D. Sudhoff. An evolutionary computing approach to robust design in the presence of uncertainties. *IEEE Transactions on Evolutionary Computation*, 14(6):900–912, 2010.

[3] R. F. Coelho. Probabilistic dominance in multiobjective reliability-based optimization: theory and implementation. *IEEE Transactions on Evolutionary Computation*, 19(2):214–224, 2015.

[4] C. A. C. Coello and E. M. Montes. Constraint-handling in genetic algorithms through the use of dominance-based tournament selection. *Advanced Engineering Informatics*, 16(3):193–203, 2002.

[5] S. Das and P. N. Suganthan. Differential evolution: a survey of the state-of-the art. *IEEE Transactions on Evolutionary Computation*, 15(1):4–31, 2011.

[6] K. Deb and H. Gupta. Introducing robustness in multi-objective optimization. *Evolutionary Computation*, 14(4):463–494, 2006.

[7] V. Feoktistov. *Differential Evolution in Search of Solutions*, Springer, 2006.

[8] H. Garg. Solving structural engineering design optimization problems using an artificial bee colony algorithm. *Journal of Industrial and Management Optimization*, 10(3):777–794, 2014.

[9] Y. Jin and J. Branke. Evolutionary optimization in uncertain environments - a survey. *IEEE Transactions on Evolutionary Computation*, 9(3):303–317, 2005.

[10] B. K. Kannan and S. N. Kramer. An augmented Lagrange multiple based method for mixed integer discrete continuous optimization and its applications to mechanical design. *Journal of Mechanical Design. Transactions of the ASME*, 116:318–320, 1994.

[11] B. Liu, X. Zhang, and H. Ma. Hybrid differential evolution for noisy optimization. *Proceedings of the IEEE Congress on Evolutionary Computation (CEC)*, pages 587–592, 2008.

[12] T. Park and K. R. Ryu. Accumulative sampling for noisy evolutionary multi-objective optimization. *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO)*, pages 793–800, 2011.

[13] K. V. Price, R. M. Storn, and J. A. Lampinen. *Differential Evolution - A Practical Approach to Global Optimization*, Springer, 2005.

[14] S. Rahanamayan, H. R. Tizhoosh, and M. M. A. Salama. Opposition-based differential evolution for optimization of noisy problems. *Proceedings of the IEEE Congress on Evolutionary Computation (CEC)*, pages 6756–6763, 2006.

[15] P. Rakshit, A. Konar, S. Das, L. C. Jain, and A. K. Nagar. Uncertainty management in differential evolution induced multiobjective optimization in presence of measurement noise. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 44(7):922–937, 2013.

[16] S. S. Rao. *Engineering Optimization*. John Wiley and Sons, 3rd ed., 1996.

[17] Z. Ren, M.-T. Pham, and C. S. Koh. Robust global optimization of electromagnetic devices with uncertain design parameters: comparison of the worst case optimization methods and multiobjective optimization approach using gradient index. *IEEE Transactions on Magnetics*, 48(2):851–859, 2013.

[18] J. G. Saw, M. C. K. Yang, and T. C. Mo. Chebyshev inequality with estimated mean and variance. *The American Statistician*, 38(2):130–132, 1984.

[19] K. Tagawa, Y. Sasaki, and H. Nakamura. Optimum design of balanced SAW filters using multi-objective differential evolution. *Lecture Notes in Computer Science*, 6457:466–475, 2010.

[20] K. Tagawa and T. Suenaga. Extended differential evolution algorithm for worst-case value minimization problems. *International Journal of Mathematical Models and Methods in Applied Sciences*, 8:262–272, 2014.

[21] K. Tagawa and S. Harada. Multi-noisy-objective optimization based on prediction of worst-case performance. *Lecture Notes in Computer Science*, 8890:23–34, 2014.

[22] K. Tagawa, H. Takeuchi, and A. Kodama. Memetic differential evolutions using adaptive golden section search and their concurrent implementation techniques. *Proceedings of the IEEE Congress on Evolutionary Computation (CEC)*, pages 2532–2539, 2015.

[23] P. Tchébychef. Des valeurs moyennes. *Journal de Mathématiques Pures et Appliquées*, 2(12):177–184, 1867.

[24] J.-T. Tsai. An evolutionary approach for worst-case tolerance design. *Engineering Applications of Artificial Intelligence*, 25:917–925, 2012.

[25] A. Zamuda, J. Brest, B. Bošković, and V. Žumer. Woody plants model recognition by differential evolution. *Proceedings of the 4th International Conference on Bioinspired Optimization Methods and their Applications (BIOMA)*, pages 205–215, 2010.