All Sciences Proceedings
All Sciences Proceedings
http://as-proceeding.com/

1st International Conference on Scientific and Innovative Studies

April 18-20, 2023 : Konya, Turkey

ICSIS

https://as-proceeding.com/index.php/icsis

A Noise Reduction Method for Semi-Noisy Multiobjective Optimization Problems

Tolga Altinoz*[1]

[1]*Department of Electrical and Electronics Engineering, Ankara University, Turkey*

*(taltinoz@ankara.edu.tr) *Email of the corresponding author*

Abstract – In engineering problems, variables such as temperature, speed, location are noisy variables that are included in the system, and they become objective function variables. Because these variables are noisy, the objective functions are also noisy. Because there is more than one objective in multi-objective optimization problems, these variables may not affect each objective. Not all variables may be included for each objective function as variables. Therefore, in multi-objective optimization problems, it may be known to know both the noisy and noiseless states of one or more purposes. In this case, noise of the objective function may be extracted. In this case, the noise of other objective functions can be reduced by using the statistical properties of the known noise signal. The aim of this study is to reduce the noise in the objective functions as explained by using the statistical properties of the noise. For this purpose, two optimization algorithms and eight test problems will be used. In addition, statistical properties will be obtained from the data recorded with different window sizes.

Keywords –Optimization, Multiobjective Optimization, Noise, Mean, Standard Deviation, Noise

I. Introduction

The Multiobjective optimization problems have more than one objective which is needed to optimize simultaneously. The problems may have some constraints which are the equivalent or non-equal functions that needed to be succeeded by the algorithms. In this research un-constraint problems are considered with the decision variable boundaries. The objectives may be contaminated with the noise and become noisy Multiobjective optimization problems. The source of the noise is generally the disturbance and/or the measurement noise which are inherent in engineering problems. Generally, the noise will be in the decision variables and since the noise is in decision variables, they have an influence on the objective functions. However, it is possible to know both noisy and noiseless data in engineering problems for example at the output of the controller in control system it is possible to know the exact controller output from the microcontroller and the noise is influenced on this signal and the noisy variable is obtained. The noise Multiobjective optimization problem is given in Eq. 1 and 2, for Multiobjective and noisy Multiobjective problems.

$$F = \{f_1, f_2, \dots, f_M\} \tag{1}$$

$$F_n = \{f_1 + n_1, f_2 + n_2, \dots, f_M + n_M\} \tag{2}$$

where F is the objective function f set with M number of objectives and n is the additive noise so that at each objective a different value is added with the same properties. The best solution of the Multiobjective optimization is called the Pareto set and the values of the pareto set on the objective space is called the Pareto Front [7,8,9].

The noise from one of the objectives can be obtained or known, by this way the statistical properties of the noise can be obtained. The question is how these properties help to reduce the noise at other objectives is the main motivation of this study. For this purpose, two and three objective benchmark problems are considered and noise data from one objective value is used to extract the statistical properties. Two optimization algorithms are used for this purpose which are NSGA-II and MOEAD algorithms.

This paper is organized into five sections. After the introduction the techniques and tools are given with the statistics of the noise data follows that. In this subsection the solution method with optimization algorithms and the performance measurement function is presented. Then the implementation results and discussion of the solutions are given. Finally, the conclusion of these results is presented as the final section. Also, in that section the future study issues are discussed.

II. Techniques and Tools

In this research, the semi-noisy multiobjective optimization problem is considered so that at least one (for this research just one objective is considered however two or more noiseless objectives case leases as future study) objective is recorded both noisy and noiseless data. Since both data is already known it is easy the get the noise data as $e(n)=v(n)-s(n)$. However, unlike the speech signal or an image or video signals, the amount of data at optimization problem is limited and equals to the number of generations. For this reason, it is not easy to get a true statistical property of the error since the whole error data is not available. Therefore, a sliding-window-based error recording method looks applicable for this problem. In this section this idea and optimization problems with the algorithms are briefly explained.

III. Statistics of the Noise Data

The issue related to the extracting of the statistical features of the error data is the size of the error data. When compared to other digital signal processing applications like speech processing, the data considered on this problem is limited, therefore the windows size is more important. For this reason, the window size for this research needs to be determined by empirical studies. This is another motivation of this research. Figure 1 gives the graphical demonstration of the error data package. There are N number of data in the array. When a new generation in constructed a new error data is received and the windows given in Figure 1 shifts.

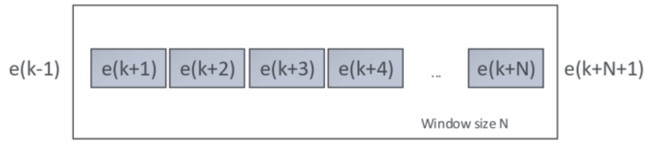

Figure 1. The graphical demonstration of the window size

The size of data N is the case study of this research. This size is selected as $N=\{5,10,15,20,25\}$. However, for all cases the windows are shifted at Case 1. In this case, until the generation is reached to N, the statistical properties of the noise signal are not extracted. For Case 2, all the constructed error data is recorded, and statistical properties are extracted from the data hence the size of error data is same as the generation. Two statistical properties are mean and standard deviation of the data which are mathematically given in Equation 3, and 4.

$$e_m = \frac{1}{N}\sum_{i=1}^{N} e(i) \quad (3)$$

$$e_{std} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(e(i) - e_m)^2} \quad (4)$$

The windows size is critical to get a suitable statistical property. After the statistics of the error noise is getting a random number is generated and subtracted from the other objective values. As the performance measurement IGD metric will be used and it is explained and mathematically given in the following sub-section. The performance of the proposed method is compared with each other.

A. *Benchmark Problems*

In this research, two types of benchmark problem set will be considered to give a hint about the effect of number of objective functions. The first set of problems are two objective benchmark problems. For this problem the ZDT benchmark problems (ZDT1,2,3, and 4) are selected [1]. The mathematical formulations for the ZDT problems are given in Eqs. 5-8 for ZDT1-4, respectively.

$$f_1 = x_1, f_2 = 1 - \sqrt{\frac{x_1}{g}}, g = 1 + \frac{9}{n-1}\sum_{i=2}^{n} x_i \quad (5)$$

$$f_1 = x_1, f_2 = 1 - \left(\frac{x_1}{g}\right)^2, g = 1 + \frac{9}{n-1}\sum_{i=2}^{n} x_i \quad (6)$$

$$f_1 = x_1, f_2 = 1 - \sqrt{\frac{x_1}{g}} - \frac{x_1}{g} sin(10\pi x_1), g = 1 + \frac{9}{n-1}\sum_{i=2}^{n} x_i \quad (7)$$

$$f_1 = x_1, f_2 = 1 - \sqrt{\frac{x_1}{g}}, g = 1 + 10(n-1) + \sum_{i=2}^{n}\left(x_i^2 - 10cos(4\pi x_i)\right) \quad (8)$$

Table 1. The IGD+ metric for the noisy ZDT and noisy three objective DTLZ problems

Problem	M	D	NSGAII	MOEAD	Problem	M	D	NSGAII	MOEAD
ZDT1N	2	30	3.4141e-2 (3.33e-2)	3.5933e-2 (2.23e-2)	DTLZ1N	3	7	4.6330e+0 (2.31e+0)	9.1942e-1 (6.27e-1)
ZDT2N	2	30	4.2504e-1 (2.77e-1)	4.7480e-1 (1.04e-1)	DTLZ 2N	3	12	4.2049e-5 (1.17e-4)	3.3360e-6 (1.05e-5)
ZDT3N	2	30	3.3528e-2 (4.01e-2)	7.0719e-3 (8.36e-3)	DTLZ 3N	3	12	4.5827e+1 (2.15e+1)	3.0768e+1 (8.55e+0)
ZDT4N	2	10	1.4857e+0 (7.51e-1)	2.2543e+0 (6.25e-1)	DTLZ 4N	3	12	5.5457e-3 (1.74e-2)	1.4010e-4 (4.43e-4)

Table 2. The IGD+ metric value for the Case 1 with different window size for ZDT benchmark problems

		N=10		N=20		N=30	
Problem	M	NSGAII	MOEAD	NSGAII	MOEAD	NSGAII	MOEAD
ZDT1C1	2	5.3357e-3 (5.64e-3)	6.7384e-3 (8.32e-3)	2.1883e-2 (3.33e-2)	6.7833e-3 (1.08e-2)	9.7891e-3 (5.71e-3)	1.3866e-3 (1.58e-3)
ZDT2C1	2	8.9576e-2 (9.08e-2)	8.5592e-2 (7.94e-2)	1.9427e-1 (1.82e-1)	5.8424e-2 (6.24e-2)	1.7730e-1 (1.17e-1)	8.1422e-2 (7.13e-2)
ZDT3C1	2	2.9771e-2 (1.78e-2)	1.3458e-2 (1.59e-2)	3.1115e-2 (3.04e-2)	1.1194e-2 (9.95e-3)	3.7169e-2 (2.66e-2)	1.4082e-2 (1.33e-2)
ZDT4C1	2	2.3571e-1 (1.80e-1)	4.0586e-1 (3.72e-1)	1.4476e-1 (1.59e-1)	3.6844e-1 (2.43e-1)	2.6531e-1 (2.26e-1)	2.5861e-1 (1.87e-1)

These benchmark problems are selected since the first objective of these problems is the first decision variable x1. This supports the motivation of the research so that it may be possible to get both noisy and noiseless data from the first objective. For this reason, it is assumed that for ZDT problems, the noisy and noiseless values of the first objective function are known.

The second class of benchmark problems are selected from the DTLZ problem set [2, 3]. The DTLZ problem set is defined so that it is possible to generate different number of objective functions. In this research DTLZ problem set is generated as three objective problems set. To evaluate DTLZ1-4 are selected as the second class of the benchmark problems with three objectives. The mathematical description of the DTLZ benchmark problems is given in Eqs. 9-12 for DTLZ1, Eqs. 13-16 for DTLZ2, Eqs. 17-20 for DTLZ3, and Eqs. 21-24 for DTLZ4 problems.

$$f_1 = \frac{1}{2}x_1 x_2 (1 + g(x_3)) \quad (9)$$

$$f_2 = \frac{1}{2}x_1(1 - x_2)(1 + g(x_3)) \quad (10)$$

$$f_3 = \frac{1}{2}(1 - x_1)(1 + g(x_3)) \quad (11)$$

$$g(x_3) = 100\left[|x_3| + \sum\left((x_i - 0.5)^2 - cos(20\pi(x_i - 0.5))\right)\right] \quad (12)$$

$$f_1 = (1 + g(x_3))cos\left(x_1\frac{\pi}{2}\right)cos\left(x_2\frac{\pi}{2}\right) \quad (13)$$

$$f_2 = (1 + g(x_3))cos\left(x_1\frac{\pi}{2}\right)sin\left(x_2\frac{\pi}{2}\right) \quad (14)$$

$$f_3 = (1 + g(x_3))sin\left(x_1\frac{\pi}{2}\right) \quad (15)$$

$$g(x_3) = \sum(x_i - 0.5)^2 \quad (16)$$

$$f_1 = (1 + g(x_3))cos\left(x_1\frac{\pi}{2}\right)cos\left(x_2\frac{\pi}{2}\right) \quad (17)$$

$$f_2 = (1 + g(x_3))cos\left(x_1\frac{\pi}{2}\right)sin\left(x_2\frac{\pi}{2}\right) \quad (18)$$

$$f_3 = (1 + g(x_3))sin\left(x_1\frac{\pi}{2}\right) \quad (19)$$

$$g(x_3) = 100\left[|x_3| + \sum\left((x_i - 0.5)^2 - cos(20\pi(x_i - 0.5))\right)\right] \quad (20)$$

$$f_1 = (1 + g(x_3))cos\left(x_1^{100}\frac{\pi}{2}\right)cos\left(x_2^{100}\frac{\pi}{2}\right) \quad (21)$$

$$f_2 = (1 + g(x_3))cos\left(x_1^{100}\frac{\pi}{2}\right)sin\left(x_2^{100}\frac{\pi}{2}\right) \quad (22)$$

$$f_3 = (1 + g(x_3))sin\left(x_1^{100}\frac{\pi}{2}\right) \quad (23)$$

$$g(x_3) = \sum(x_i - 0.5)^2 \quad (24)$$

Table 3. The IGD+ metric value for the Case 1 with different window size for DTLZ benchmark problems

		N=10		N=20		N=30	
Problem	M	NSGAII	MOEAD	NSGAII	MOEAD	NSGAII	MOEAD
DTLZ1C1	3	1.3478e+0 (8.19e-1)	6.3445e+0 (2.37e+0)	1.4341e+0 (7.90e-1)	7.9848e+0 (2.93e+0)	1.3264e+0 (4.02e-1)	7.7789e+0 (2.93e+0)
DTLZ2C1	3	3.7030e-5 (7.12e-5)	2.1669e-5 (4.05e-5)	1.3248e-4 (2.12e-4)	6.6505e-5 (1.44e-4)	1.4546e-5 (3.02e-5)	5.9132e-5 (1.30e-4)
DTLZ3C1	3	3.3537e+1 (1.42e+1)	1.4989e+2 (7.28e+1)	2.4253e+1 (7.73e+0)	1.4435e+2 (3.78e+1)	3.2369e+1 (1.27e+1)	1.4856e+2 (3.40e+1)
DTLZ4C1	3	2.9749e-3 (8.27e-3)	8.2865e-3 (9.59e-3)	5.9227e-3 (1.43e-2)	1.8167e-2 (1.18e-2)	8.3742e-3 (1.98e-2)	1.1042e-2 (1.24e-2)

Table 4. The IGD+ metric value for the Case 2 for ZDT and DTLZ benchmark problems

Problem	M	NSGAII	MOEAD	Problem	M	NSGAII	MOEAD
ZDT1C2	2	1.2493e-2 (1.64e-2)	2.9025e-3 (4.23e-3)	DTLZ1C2	3	8.8797e-1 (6.36e-1)	8.4269e+0 (2.95e+0)
ZDT2C2	2	3.0498e-1 (1.47e-1)	8.1500e-2 (8.16e-2)	DTLZ2C2	3	7.7965e-5 (2.01e-4)	5.2514e-5 (8.78e-5)
ZDT3C2	2	2.9601e-2 (2.22e-2)	1.6433e-2 (1.50e-2)	DTLZ3C2	3	3.2280e+1 (1.16e+1)	1.5678e+2 (3.73e+1)
ZDT4C2	2	2.7074e-1 (2.53e-1)	4.3763e-1 (3.17e-1)	DTLZ4C2	3	2.4240e-2 (4.00e-2)	2.0613e-2 (1.95e-2)

Unlike ZDT problems, in DTLZ problems set the final objective function contains single variable multiplication. Therefore, in DTLZ problems it is assumed that the noiseless value of the last objective is known. As brief, in this research four benchmark problems with two objectives and four benchmark problems with three objectives are selected as test problems. A Gaussian noise with zero mean and 0.15 standard deviations are added to all the objective functions. However, among them only one of the objective's noiseless objective function values is known. Therefore, it is possible to get noise data from this objective.

B. *Optimization Algorithms*

In this research two Multiobjective optimization algorithms are considered for the given problem definition. The first algorithm which is called as NSGA-II, proposed by Deb et al. in 2002 [5]. The algorithm is based on the nondominated sorting of the solution candidates on the objective space. The algorithm is composed of three genetic operators' crossover (SBX is preferred), mutation (polynomial mutation is used) and selection operators. As the selection operator population (both parents and offspring) are distributed to the ranks with respect to the nondominated sorting algorithm. The smaller ranked members survive the next generation. The remaining members are compared with each other with a method called crowding distance. By this way the more spread distributed members will be in the solution set.

The second optimization algorithm is called MOEAD which is proposed by Zhang et al. in 2007 [6]. This algorithm is another algorithm which implemented genetic operators same as NSGA-II. However, the selection operator is different than NSGA-II. In this algorithm the decomposition method, which is an aggregation method, is used to select the best members for the next generation. In addition to the decomposition-based selection operator the neighbourhood is recorded with the weight vector of the decomposition method. Thebycheff is selected as the aggregation function. Based on the competition with respect to the weight vector, the best members are selected. After the selection operator is terminated if nadir point is used for aggregation, it is updated based on the new position of the solutions on the objective space.

C. *Performance Measurement*

In this research the performance measurement metric modified inverted generational distanced

(IGD+) metric which was proposed by Ishibuchi et al. in 2015, is used as performance indicator. The difference between IGD and IGD+ metric is the definition of the distance between Pareto set and the obtained approximate solution set. In IGD metric the distance is calculated as the Euclidean distance. However, in IGD+ the weak Pareto compliant is preferred with is the positive square summation calculation which is given in Eq. 25.

$$d = \sqrt{\sum(max\{r_k - o_k, 0\})^2} \qquad (25)$$

where r is the sampled data from Pareto front and o is the data from obtained solution set. The difference between data in the Pareto front and the obtained solution is calculated. If this calculation gives negative value than it is rounded to zero, else the square of this value is calculated and summing of all the data and square root of this summation is given as the metric (IGD+) for this research.

IV. IMPLEMENTATION

In this research the statistical properties of the known noise signal are extracted to reduce the noise at the objectives. For this purpose, two Multiobjective optimization algorithms are implemented on eight benchmark problems and their performance is measured by using the IGD+ metric. All algorithms are implemented with 100 population size, 10^5 maximum number of function evaluation and 10 independent runs.

Initially, the noise is added to the objective functions with 0.15 standard deviation and zero mean Gaussian noise. The zero mean is selected because at the calculation of the mean of the collected data there will be an additional bias at the data which increases the complexity of the problem. Table 1 presents the performance of the noise benchmark problems.

First, as explained in section “Statistics of the Noise Data,” the different size of windows for calculating the mean and standard deviation is compared with each other. The windows sizes are selected as 10, 20, and 30 for this research. Table 2 gives the performance metric values for the given algorithms with respect to the windows size. When Table 2 is compared with Table 1, the proposed method increases the performance and reduces the noise effect almost x10, which shows the effect of the proposed methodology.

Table 3 gives the same implementation Case 1 for three objective DTLZ problems. When the results in Table 3 compared with Table 1, the performance of the proposed noise reduction methods falls behind the noise problems. This is one of the important conclusions of the research. Because for three objective cases the selection objective function has a noise with not very effective to reduce noise at other objectives. Even if the noise is zero mean, the extracted statistical feature has a value for a mean. That decreases the impact of the reduction method, on the contrary it increases the effect of the noise.

Finally, for Case 2, the whole noise data is recorded, and statistical features are calculated. Similar results are obtained from Table 4. For ZDT problems the same performance improvement is obtained, however the performances are very close to each other. For DTLZ the same performance is still reached but still worse than the noisy data given in Table 1.

V. CONCLUSION

In this research, it is assumed that one of the objective’s noise data is known and by extracting its statistical properties like mean and standard deviation, the inverse noise is generated and subtracted from the noisy objective functions. To measure the performance of this proposed method, two Multiobjective optimization and eight benchmark problems are considered. These benchmark problems have two and three objectives. Also, the statistical properties are extracted from the constant data size where different sizes are considered. This is considered the first case. As the second case all the error data is recorded until end of the generations, as Case 2. When all these implementations are investigated, it can be concluded that as the number of objectives increase the improvement by extracting the noise feature does not help reduce the noise. But for two objective cases it is increased the performance of the noisy objectives with respect to the noise reduction. As the future study the many objective problems may be considered and different types of the noise should be considered to get a general conclusion for this methodology. However, it can be indicated that more advanced digital signal processing methods may be implemented to reduce the noise.

REFERENCES

[1] E. Zitzler, K. Deb, and L. Thiele. "Comparison of multiobjective evolutionary algorithms: Empirical results." Evolutionary computation, vol. 8, no. 2, pp. 173-195, 2000.

[2] K. Deb, L. Thiele, M. Laumanns, E. Zitzler, "Scalable Test Problems for Evolutionary Multi-Objective Optimization," Technical Report, Computer Engineering and Networks Laboratory (TIK), Swiss Federal Institute of Technology (ETH). https://dx.doi.org/10.3929/ethz-a-004284199, 2001.

[3] K. Deb, L. Thiele, M. Laumanns, E. Zitzler, "Scalable multi-objective optimization test problems," IEEE Congress on Evolutionary Computation, pp. 825-830, 2002.

[4] H. Ishibuchi, H. Masuda, Y. Tanigaki, and Y. Nojima, "Modified distance calculation in generational distance and inverted generational distance," International Conference on Evolutionary Multi-Criterion Optimization, pp. 110-125, 2015.

[5] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: NSGA-II," Transactions on Evolutionary Computation, vol. 6, no. 2, pp. 182–197, 2002.

[6] Q. Zhang and H. Li, "MOEA/D: A multiobjective evolutionary algorithm based on decomposition," IEEE Transactions on Evolutionary Computation, vol. 11, no. 6, pp. 712-731, 2007.

[7] L. Painton, Laura, U. Diwekar, "Stochastic annealing for synthesis under uncertainty," Eur. J. Oper. Res., vol. 83, no. 3, pp. 489–502, 1995.

[8] J. Branke, S. Meisel, C. Schmidt, "Simulated annealing in the presence of noise," J. Heuristics, vol. 14, no. 6, pp. 627–654, 2008.

[9] S.B. Gelfand, S.K. Mitter, "Simulated annealing with noisy or imprecise energy measurements," J. Optim.Theory Appl., vol. 62, no. 1, pp. 49–62, 1989.