

Research Article

Novel Global Harmony Search Algorithm for Least Absolute Deviation

Longquan Yong

School of Mathematics and Computer Science, Shaanxi University of Technology, Hanzhong 723001, China

Correspondence should be addressed to Longquan Yong; yonglongquan@126.com

Received 17 April 2014; Accepted 9 July 2014; Published 21 July 2014

Academic Editor: Jinyun Yuan

Copyright © 2014 Longquan Yong. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The method of least absolute deviation (LAD) finds applications in many areas, due to its robustness compared to the least squares regression (LSR) method. LAD is robust in that it is resistant to outliers in the data. This may be helpful in studies where outliers may be ignored. Since LAD is nonsmooth optimization problem, this paper proposed a metaheuristics algorithm named novel global harmony search (NGHS) for solving. Numerical results show that the NGHS method has good convergence property and effective in solving LAD.

1. Introduction

Least squares regression (LSR) method is one of the oldest and the most widely used statistical tools for linear models. Its theoretical properties have been extensively studied and are fully understood. Despite its many superior properties, the LSR estimate can be sensitive to outliers and, therefore, nonrobust [1]. In order to overcome these problems, the researchers have recently investigated the alternative regression method which is least absolute deviation (LAD) method.

Least absolute deviation (LAD), also known as least absolute errors (LAE), least absolute value (LAV), or least absolute residual (LAR), or the L_1 norm problem, is a mathematical optimization technique similar to the LSR that attempts to find a function which closely approximates a set of data $(x_i, y_i), i = 1, 2, \dots, n$. In the simple case, the approximation function is a simple “trend line” in two-dimensional Cartesian coordinates [2]. The method minimizes the sum of absolute errors (SAE) (the sum of the absolute values of the vertical “residuals” between points generated by the function and corresponding points in the data).

The method of LAD finds applications in many areas, due to its robustness compared to the LSR method. LAD is robust in that it is resistant to outliers in the data. This may be helpful in studies where outliers may be safely and effectively ignored. Despite its long history and many ground-breaking works, the LAD has not been explored in theory as well as in

application to the extent as the LSR [3]. This is largely because computing LAD estimates are more difficult than computing LSR estimates due to the fact that algorithmic method must be employed to calculate LAD estimates. Therefore, LAD estimation method is not popular as LSR method.

Over the past few years, a number of approaches have been developed for solving the LAD using classical mathematical programming methods. Since Charnes et al. [4] reduced the LAD method to a linear programming problem, the computational difficulty is now entirely overcome by the availability of computing power and the effectiveness of linear programming. A comprehensive summary of the subject can be found in Koenker and D'Orey [5]. Large sample properties of the LAD estimates are obtained in Koenker and Bassett Jr. [6], Chen et al. [7], and Pollard [8]. Due to these developments in theoretical and computational aspects, the LAD method has become increasingly popular. In particular, it has many applications in econometrics and biomedical studies; see Bassett and Koenker [9], Powell [10], Buchinsky [11], among many others.

Since LAD is a class of nondifferentiable optimization problem, this paper proposed a metaheuristics algorithm named novel global harmony search (NGHS) [12] for solving LAD model. The NGHS algorithm includes two important operations: position updating and genetic mutation with a small probability (p_m). The former enables the worst harmony of harmony memory to move to the global best

harmony rapidly in each iteration, and the latter can effectively maintain population diversity and prevent the NGHS from trapping into the local optimum.

The remaining sections of this chapter are organized as follows. In Section 2, LAD model is shown. In Sections 3 and 4, the classical HS and NGHS are described. Numerical results are presented and compared in Section 5. Lastly, Section 6 outlines our conclusions.

2. Least Absolute Deviations Model

The LAD method is a widely known alternative to the classical LSR method for statistical analysis of linear regression models. Instead of minimizing the sum of squared errors (SSE) in LSR, it minimizes the sum of absolute errors (SAE).

Suppose that the data set consists of n points (data pairs) $(x_i, y_i), i = 1, 2, \dots, n$, where x_i is an independent variable and y_i is a dependent variable whose value is found by observation. We want to find a function f such that $f(x_i) = y_i$.

To attain this goal, we suppose that the function f is of a particular form containing some parameters which need to be determined. For instance, the simplest form would be linear: $f(x, \beta) = \beta_0 + \beta_1 x$, where β_0 and β_1 are parameters whose values are not known but which we would like to estimate.

We now seek estimated values of the unknown parameters that minimize the SAE:

$$\min S = \sum_{i=1}^n |f(x_i) - y_i|. \quad (1)$$

Let

$$A^T = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_n \end{bmatrix}, \quad \beta = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix}, \quad (2)$$

$$b^T = [y_1, y_2, \dots, y_n].$$

Then LAD problem is equivalent to following L_1 optimization problem:

$$\min S = \sum_{i=1}^n |f(x_i) - y_i| = \|A\beta - b\|_1. \quad (3)$$

Meanwhile, then LSR problem is equivalent to following L_2 optimization problem

$$\min S = \sum_{i=1}^n |f(x_i) - y_i|^2 = \|A\beta - b\|_2^2. \quad (4)$$

For a given data set, the method of LAD may produce multiple solutions, whereas the method of LSR always produces only one solution (the regression line is unique).

3. Harmony Search Algorithm

Recently, a new class of metaheuristics, named harmony search (HS), has been developed. The HS algorithm proposed

in [13] has been developed in an analogy with music improvisation process where musicians in an ensemble continue to polish their pitches in order to obtain better harmony. Jazz improvisation seeks to find musically pleasing harmony similar to the optimum design process which seeks to find optimum solution. The pitch of each musical instrument determines the aesthetic quality, just as the objective function value is determined by the set of values assigned to each decision variable. In addition, HS uses a stochastic random search instead of a gradient search.

The steps in the procedure of standard harmony search algorithm (HS) are as follows.

Step 1 (initialize the problem and algorithm parameters). The optimization problem is specified as follows:

$$\begin{aligned} &\text{Minimize} && f(x) \\ &\text{subject to} && x_i \in X_i, \quad i = 1, 2, \dots, N, \end{aligned} \quad (5)$$

where $f(x)$ is an objective function; x is the set of each decision variable x_i ; N is the number of decision variables; X_i is the set of the possible range of values for each decision variable; $X_i : x_i^L \leq x_i \leq x_i^U$. The HS algorithm parameters are also specified in this step. These are the harmony memory size (HMS) or the number of solution vectors in the harmony memory; harmony memory considering rate (HMCR); pitch adjusting rate (PAR); and the number of improvisations (Tmax) or stopping criterion.

Step 2 (initialize the harmony memory). The HM matrix is filled with as many randomly generated solution vectors as the HMS

$$\begin{aligned} \text{HM} &= \begin{bmatrix} x^1 & f(x^1) \\ x^2 & f(x^2) \\ \vdots & \vdots \\ x^{\text{HMS}} & f(x^{\text{HMS}}) \end{bmatrix} \\ &= \begin{bmatrix} x_1^1 & x_2^1 & \cdots & x_N^1 & f(x^1) \\ x_1^2 & x_2^2 & \cdots & x_N^2 & f(x^2) \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ x_1^{\text{HMS}} & x_2^{\text{HMS}} & \cdots & x_N^{\text{HMS}} & f(x^{\text{HMS}}) \end{bmatrix}. \end{aligned} \quad (6)$$

Step 3 (improvise a new harmony). Generate a new harmony is called "improvisation." A new harmony vector, $x' = (x'_1, x'_2, \dots, x'_N)$, is generated based on three rules: (1) memory consideration, (2) pitch adjustment, and (3) random selection. The procedure works as Algorithm 1.

x'_i ($i = 1, 2, \dots, N$) is the i th component of x' , and x_i^j ($j = 1, 2, \dots, \text{HMS}$) is the i th component of the j th candidate solution vector in HM. Here $r(-1, 1)$ is uniform distribution random number in the region of $(-1, 1)$, bw is an arbitrary distance bandwidth, and $\text{rand}()$ is uniform distribution random number in the region of $(0, 1)$.

Step 4 (update harmony memory). If the new harmony vector $x' = (x'_1, x'_2, \dots, x'_N)$ is better than the worst harmony

```

For each  $i \in [1, 2, \dots, N]$  do
  If  $\text{rand}() < \text{HMCR}$ 
     $x'_i = x_i^j (j = 1, 2, \dots, \text{HMS})$  //memory consideration
    If  $\text{rand}() < \text{PAR}$ 
       $x'_i = x'_i + r(-1, 1) \times bw$  // pitch adjustment
       $x'_i = \min[\max(x'_i, x_i^L), x_i^U]$  //truncation processing
    End
  Else
     $x'_i = x_i^L + \text{rand}() \times (x_i^U - x_i^L)$  //random selection
  End
End

```

ALGORITHM 1: Generating a new harmony by classical HS algorithm.

```

For each  $i \in [1, 2, \dots, N]$  do
   $x_s = 2 \times x_i^{\text{best}} - x_i^{\text{worst}}$ 
   $x_s = \min[\max(x_s, x_i^L), x_i^U]$  //truncation processing
   $x'_i = x_i^{\text{worst}} + \text{rand}() \times (x_s - x_i^{\text{worst}})$  //position updating
  If  $\text{rand}() < p_m$ 
     $x'_i = x_i^L + \text{rand}() \times (x_i^U - x_i^L)$  //genetic mutation
  End
End

```

ALGORITHM 2: Generating a new harmony by NGHS algorithm.

in the HM, judged in terms of the objective function value, the new harmony is included in the HM and the existing worst harmony is excluded from the HM.

Step 5 (check stopping criterion). If the stopping criterion (maximum number of improvisations) is satisfied, computation is terminated. Otherwise, Steps 3 and 4 are repeated.

4. Novel Global Harmony Search Algorithm

Experiments with the classical HS algorithm over the benchmark problems show that the algorithm suffers from the problem of premature and/or false convergence and slow convergence especially over multimodal fitness landscape. To enrich the searching behavior and to avoid being trapped into local optimum, more improved HS algorithms were presented. Zou et al. [12] proposed an NGHS algorithm and it has been successfully used in reliability problems [14], and 0-1 knapsack problems [15].

The NGHS algorithm modifies the improvisation step of the HS such that the new harmony can mimic the global-best harmony in the HM. In Step 3 it works as Algorithm 2.

Here, “best” and “worst” are the indexes of the best and worst harmony in HM, respectively.

The NGHS proposed algorithm has strong global search ability in the early stage of optimization and has strong local search ability in the late stage of optimization. In the early stage of optimization, all solution vectors are sporadic in feasible space, so most x_s is large, which is beneficial to the global search of the NGHS, while in the late stage of

optimization, all nonbest solution vectors are inclined to move to the global-best solution vector, so most solution vectors are close to each other. In this case, most x_s is small and most trust regions are narrow, which is beneficial to the local search of the NGHS.

In Step 4, the NGHS replaces the worst harmony x^{worst} in HM with the new harmony x' even if x' is worse than x^{worst} .

The NGHS and the HS are different in following. (i) In Step 1, harmony memory considering rate (HMCR), pitch adjusting rate (PAR), and adjusting step (bw) are excluded from the NGHS, and genetic mutation probability (p_m) is included in the NGHS. (ii) The HS carries out mutation with the probability $\text{HMCR} \times \text{PAR}$ and carries out random selection with the probability $1 - \text{HMCR}$, while the NGHS carries out genetic mutation with the probability p_m . In fact, both operations are exactly the same, and they are used to keep the individual variety better, which can effectively prevent both algorithms from being trapped into the local optimum.

5. Computational Results

In this section we solve an LAD problem in order to illustrate the implementation and efficiency of the NGHS method. All the experiments were performed on MatlabR2009a system with Intel(R) Core(TM) 4 × 3.3 GHz and 2 GB RAM.

The data pairs $(x_i, y_i), i = 1, 2, \dots, 23$, are listed in Tables 1 and 2, where x_i is an independent variable and y_i is a dependent variable whose value is found by observation.

TABLE 1: Given data.

i	1	2	3	4	5	6	7	8	9	10	11	12
x_i	20	19.6	19.6	19.4	18.4	19	19	18.3	18.2	18.6	19.2	18.2
y_i	1.0	1.2	1.1	1.4	2.3	1.7	1.7	2.4	2.1	2.1	1.2	2.3
i	13	14	15	16	17	18	19	20	21	22	23	
x_i	18.7	18.5	18	17.4	16.5	17.2	17.3	17.8	17.3	18.4	16.9	
y_i	1.9	2.4	2.6	2.9	4.0	3.3	3.0	3.4	2.9	1.9	3.9	

TABLE 2: Given data with outlier.

i	1	2	3	4	5	6	7	8	9	10	11	12
x_i	20	19.6	19.6	19.4	18.4	19	19	18.3	18.2	18.6	19.2	18.2
y_i	1.0	1.2	1.1	1.4	2.3	1.7	1.7	2.4	2.1	2.1	1.2	2.3
i	13	14	15	16	17	18	19	20	21	22	23	
x_i	18.7	18.5	10	17.4	16.5	17.2	17.3	17.8	17.3	18.4	16.9	
y_i	1.9	2.4	2.6	2.9	4.0	3.3	3.0	3.4	2.9	1.9	3.9	

TABLE 3: The Statistical results for 10 runs on given data in Table 2.

	Algorithm	Best	Mean	Worst	Standard	Mean_Time (s)
LAD	HS	1.0830E + 01	2.2618E + 01	5.9845E + 01	1.7983E + 01	2.9411E - 03
	HSCH	1.0723E + 01	2.9073E + 01	1.4305E + 02	4.0622E + 01	3.4392E - 03
	HSWB	1.1861E + 01	1.9807E + 02	7.5315E + 02	2.3679E + 02	3.5885E - 03
	NGHS	9.9041E + 00	1.2058E + 01	1.4555E + 01	1.7163E + 00	3.5500E - 03

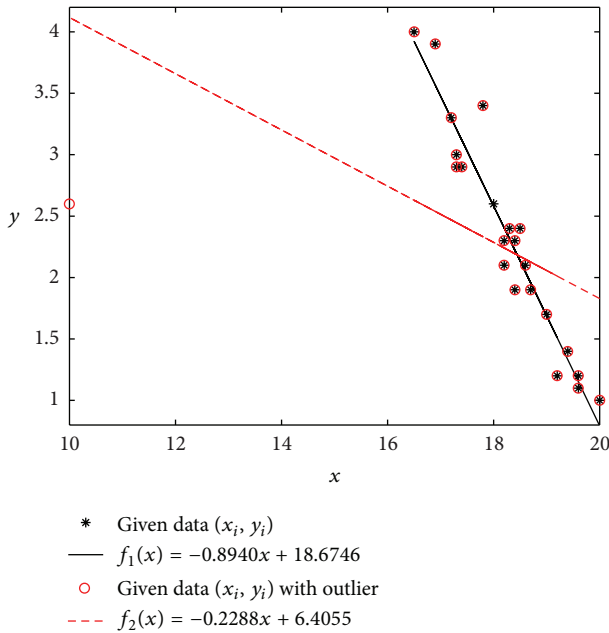


FIGURE 1: Data fitting results by LSR method. Here $f_1(x) = -0.8940x + 18.6746$ is fitting results of data in Table 1, and $f_2(x) = -0.2288x + 6.4055$ is fitting results of data in Table 2.

For data pairs listed in Tables 1 and 2, solved by LSR model [16], respectively, the results of data fitting are shown in Figure 1.

From Figure 1, LSR method is sensitive to outliers and, therefore, nonrobust. Its performance in terms of accuracy

and statistical inferences may be compromised when the errors are large and heterogeneous.

Following we use LAD model to fit given data with outlier in Table 2. Since LAD is a kind of nondifferentiable optimization problem, we solved it by NGHS. In order to show the optimization capabilities of the NGHS, the results of classical HS, HSCH [17], and HSWB [18] are listed too. To make the comparison fair, the populations for all the competitor algorithms were initialized using the same random seeds. The HS-variants algorithm parameters were set the same parameters: harmony memory size HMS = 15, harmony memory consideration rate HMCR = 0.85, pitch adjusting rate PAR = 0.35, and Tmax = 400. In NGHS, we set $p_m = 0.005$. If using a random number, we set the random-number generator to the state of 0 so that the same data can be regenerated.

To judge the accuracy of different algorithms, 10 independent runs of each of the four algorithms were carried out and the best, the mean, and the worst fitness values and the standard deviation (Std) were recorded in Table 3.

Figure 2 shows the convergence and its boxplot figure of the best fitness in the population for the different algorithms. The values plotted for every generation are averaged over 10 independent runs. As can be seen, the NGHS algorithm is the best.

Moreover, the results of data fitting by NGHS algorithm are shown in Figure 3.

Unlike the LSR method, the LAD method is not sensitive to outliers and produces robust estimates.

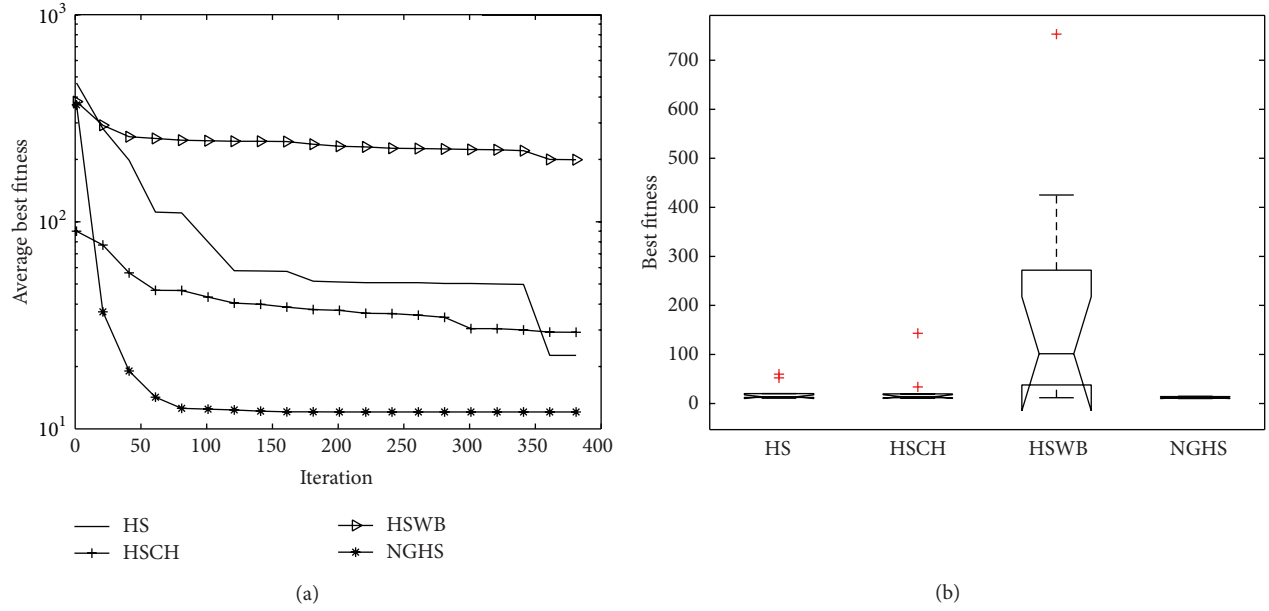
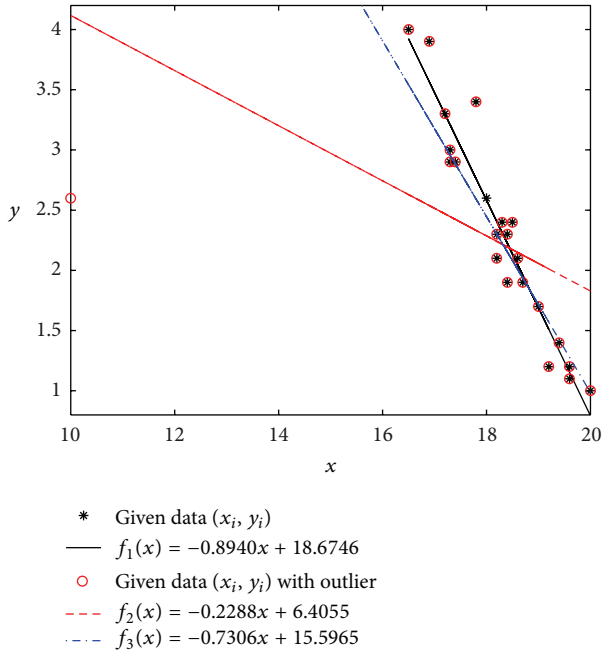


FIGURE 2: The convergence and its boxplot of the best fitness for LAD.

FIGURE 3: Data fitting results by LSR and LAD. Here $f_1(x) = -0.8940x + 18.6746$ is fitting results of data in Table 1 by LSR, $f_2(x) = -0.2288x + 6.4055$ is fitting results of data in Table 2 by LSR, and $f_3(x) = -0.7306x + 15.5965$ is fitting results of data in Table 2 by LAD.

6. Conclusion

The LAD method is a widely known alternative to the classical LSR method for statistical analysis of linear regression models. Since LAD is a class of nondifferentiable optimization problem, this paper given NGHS algorithm for solving.

Unlike the LSR, the LAD is not sensitive to outliers and produces robust estimates. Moreover, NGHS algorithm is effective for solving LAD model. The future works include investigating the application of LAD or other algorithm for solving LAD.

Conflict of Interests

The author declares that there is no conflict of interests regarding the publication of this paper.

Acknowledgment

This work is supported by Hanzhong Administration of Science and Technology under Grant no. 2013hzzx-39.

References

- [1] S. P. Ellis, "Instability of least squares, least absolute deviation and least median of squares linear regression," *Statistical Science A*, vol. 13, no. 4, pp. 337–350, 1998.
- [2] M. Caner, "A note on least absolute deviation estimation of a threshold model," *Econometric Theory*, vol. 18, no. 3, pp. 800–814, 2002.
- [3] S. Portnoy and R. Koenker, "The Gaussian hare and the Laplacian tortoise: computability of squared-error versus absolute-error estimators," *Statistical Science*, vol. 12, no. 4, pp. 279–300, 1997.
- [4] A. Charnes, W. W. Cooper, and R. O. Ferguson, "Optimal estimation of executive compensation by linear programming," *Management Science*, vol. 1, no. 2, pp. 138–151, 1955.
- [5] R. Koenker and V. D'Orey, "Algorithm AS 229: computing regression quantiles," *Journal of the Royal Statistical Society C: Applied Statistics*, vol. 36, no. 3, pp. 383–393, 1987.

- [6] R. Koenker and G. Bassett Jr., "Regression quantiles," *Econometrica*, vol. 46, no. 1, pp. 33–50, 1978.
- [7] X. R. Chen, Z. D. Bai, L. C. Zao, and Y. H. Wu, "Asymptotic normality of minimum L_1 -norm estimates in linear models," *Science in China*, vol. 33, no. 11, pp. 1311–1328, 1990.
- [8] D. Pollard, "Asymptotics for least absolute deviation regression estimators," *Econometric Theory*, vol. 7, no. 2, pp. 186–199, 1991.
- [9] J. Bassett and R. Koenker, "Asymptotic theory of least absolute error regression," *Journal of the American Statistical Association*, vol. 73, no. 363, pp. 618–622, 1978.
- [10] J. L. Powell, "Least absolute deviations estimation for the censored regression model," *Journal of Econometrics*, vol. 25, no. 3, pp. 303–325, 1984.
- [11] M. Buchinsky, "Recent advances in quantile regression models: a practical guideline for empirical research," *Journal of Human Resources*, vol. 33, no. 1, pp. 88–126, 1998.
- [12] D. X. Zou, L. Q. Gao, J. H. Wu, and S. Li, "Novel global harmony search algorithm for unconstrained problems," *Neurocomputing*, vol. 73, no. 16–18, pp. 3308–3318, 2010.
- [13] K. S. Lee and Z. W. Geem, "A new meta-heuristic algorithm for continuous engineering optimization: harmony search theory and practice," *Computer Methods in Applied Mechanics and Engineering*, vol. 194, no. 36–38, pp. 3902–3933, 2005.
- [14] D. Zou, L. Gao, J. Wu, S. Li, and Y. Li, "A novel global harmony search algorithm for reliability problems," *Computers and Industrial Engineering*, vol. 58, no. 2, pp. 307–316, 2010.
- [15] D. Zou, L. Gao, S. Li, and J. Wu, "Solving 0-1 knapsack problem by a novel global harmony search algorithm," *Applied Soft Computing Journal*, vol. 11, no. 2, pp. 1556–1564, 2011.
- [16] L. Yong, S. Liu, J. Zhang, and Q. Feng, "An improved harmony search algorithm for linear least squares," *Journal of Computational Information Systems*, vol. 8, no. 10, pp. 4217–4224, 2012.
- [17] L. Yong, S. Liu, S. Tuo, and K. Gao, "Improved harmony search algorithm with chaos for absolute value equation," *TELKOMNIKA*, vol. 11, no. 4, pp. 835–844, 2013.
- [18] L. Yong, S. Liu, S. Tuo et al., "Improved harmony search algorithm for absolute value equation," *Journal of Natural Science of Heilongjiang University*, vol. 30, no. 3, pp. 321–327, 2013.