

Sine Series Optimization with MOEA Framework

Iliyan Iliev, Victor Danev, and Gergana Mateeva

Bulgarian Academy of Sciences

Institute of Information and Communication Technologies

Acad. Georgi Bonchev Str., block 2, 1113 Sofia, Bulgaria

iliyan.iliev@iict.bas.bg, victor.danev@iict.bas.bg, gergana.mateeva@iict.bas.bg

Abstract: The MOEA Framework is an open-source and free Java library for experimenting and developing multi-objective evolutionary algorithms. Some general-purpose single optimization algorithms are also available as part of the library. The most popular are Evolution Strategy, Genetic Algorithms, and Differential Evolution. When forecasting time series, one of the ways to generate a forecast is by fitting a curve around the measured points. There is an infinite number of curves that can pass through the points, but one of the most useful is the Fourier sine series. In this study, heuristic optimization is used to estimate all coefficients for the Fourier sine series.

Keywords: curve fitting, financial forecasting, heuristic optimization

1. Introduction

Many real-world problems can be defined as time series where the sequence of ordered measurements is performed. In most cases, measurements are made at regular intervals, but this is not a prerequisite. Having the past measurements, a question arises, is it possible to predict future values using information from past values? In some situations, such forecasting is possible and many different forecasting approaches have been developed. A very efficient approach comes to be the curve fitting (Michael & Lily, 2020). It can be treated as machine learning based on big data (Dineva & Atanasova, 2020).

It is well known that there is an infinite number of curves that can be drawn over a finite number of points in a two-dimensional space. The greatest difficulty is to choose such a curve that best approximates the time series. Values of this curve calculated outside of the known range are the produced forecast (Jain et al., 2012). Selecting a proper curve for approximation is a practical problem. Often the approach for solving this problem is by trying different curves randomly. When the time series is treated as a signal with noise (Marinov et al., 2019), the Fourier decomposition (Qian et al., 2011) can be applied. In almost all cases, the time series has a trend. The trend is effectively approximated with an equation of a line (Johnson, 1935). The equation of the line comes with two parameters – slope and intercept. After extracting the trend, the rest of the signal decomposes into sinusoids. Each sine function has amplitude, angular frequency, and phase. The last open question is how many sinusoidal functions should be used. In this study, the number of sinusoids is established experimentally.

The multiple objective evolutionary algorithms (MOEA) are used to solve different real optimization problems including portfolio optimization. A survey of different evolutionary techniques able to cope with multiple objective portfolio optimization problems can be found in (Stoyanova & Guliashki, 2018). Depending on the particular application, in some cases it is better to use group decision-making taking into account uncertainty conditions and some of the well-known optimization strategies (Borissova & Dimitrova, 2021). A better understanding of heuristics could improve judgments and decisions in situations of uncertainty. To improve the efficiency of the heuristics, some authors propose the development of hybrid approaches by implementing many-objective evolutionary algorithm (Cai et al., 2020). When it comes to heuristic optimization, MOEA Framework is one of the most popular software libraries. The main target of this software library is multi-objective optimization with heuristic algorithms. It is very often a real-life problem to have multiple objectives like in economics (Petrov, 2021) or engineering (Alexandrov & Monov 2021). “The library supports genetic algorithms, differential evolution, particle swarm optimization, genetic programming, grammatical evolution, and more. A number of algorithms are provided out-of-the-box, including NSGA-II, NSGA-III, ϵ -MOEA, GDE3, PAES, PESA2, SPEA2, IBEA, SMS-EMOA, SMPSO, OMOPSO, CMA-ES, and MOEA/D. In addition, the MOEA Framework provides the tools necessary to rapidly design, develop, execute and statistically test optimization algorithms.” (MOEA, 2021) The MOEA Framework is written in Java, which makes it highly machine-independent. The programming language Java is one of the most used programming languages (TIOBE, 2021). This fact makes Java highly adopted by a wide audience of professional and non-professional users. The virtual machine

of Java is highly-optimized (Yeboah-Antwi & Baudry, 2017), which makes the code run on it with very high performance.

This research proposes financial forecasting calculation based on sine series curve fitting. All coefficients involved in the sine series are optimized with heuristic algorithms implemented in MOEA Framework. The paper is organized as follows: Section 1 describes the problems related to financial forecasting and the ideas for making forecasts by curve fitting; Section 2 proposes a practical solution for optimizing sine series coefficients; Section 3 reveals some practical experiments and related results; and Section 4 concludes with some suggestions for further work.

2. Coefficients Optimization

Decomposition of the time series, treated as a signal, can be done with line equation and series of sine functions (Eq. 1). Fitting the curve is a problem of finding optimal values for the coefficients in the equation.

$$h(t) = At + B + A_1 \sin(\omega_1 t + \varphi_1) + \dots + A_n \sin(\omega_n t + \varphi_n) \quad (1)$$

Where A is the slope of the line equation and B is the intercept. After the line function, there is a series of n sine functions. Each sine function has amplitude given by A_1 to A_n . Each sine function also has angular frequency given by ω_1 to ω_n and phase given by φ_1 to φ_n . All coefficients, presented as a vector, form an individual in the heuristic population (Eq. 2).

$$individual = \{A, B, A_1, \omega_1, \varphi_1, \dots, A_n, \omega_n, \varphi_n\} \quad (2)$$

The length of the individual varies according to the number of sine functions involved. The size of the individual is two (slope and intercept) when sine functions are not included. With n sine functions included, the size of the individual is $(3n + 2)$. The number of sine functions is determined experimentally. One way to limit the number of sinus functions is to monitor root mean square error (RMSE). When the addition of another sine function does not improve the RMSE, the addition of a sine function stops. The RMSE is expressed as follows:

$$MSE = \sqrt{\frac{\sum [f(t_j) - h(t_j)]^2}{m}} \quad (3)$$

Where $f(t_j)$ is the measured value from the original time series for time interval j . The calculated value from the fitted curve is expressed by $h(t_j)$ for the time interval j . The total number of available measurements is denoted by m .

The overall optimization goal is to find such values for the coefficients that will minimize the RMSE.

3. Experiments and Results

As input data to realize the numerical experiments, the time series of the Bitcoin price in US dollars are used as shown in Fig. 1.

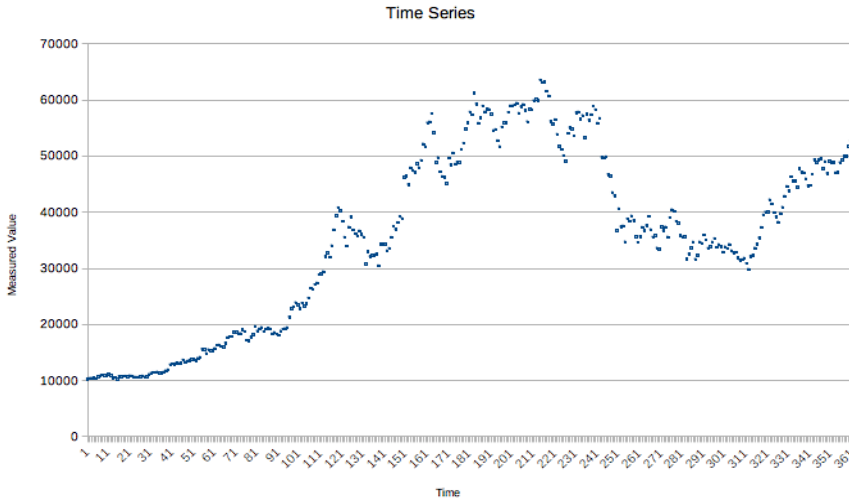


Fig. 1. Bitcoin price in USD time series

In the graph from Fig. 1, there is a clear upward trend. Apart from the trend, there are obvious price fluctuations. Calculations were done on Intel Core i5 2.3 GHz, single CPU with 2 cores, 8GB RAM, macOS High Sierra 10.13.6, Java SE 11.0.2, and MOEA Framework 2.13. Experimental results in a raw format are available at (Balabanov, 2021a). The experiments are organized in an open-source manner and the code is available at (Balabanov, 2021b).

For each optimization algorithm, there are ten optimization periods. The first optimization period looks for coefficient values only for the linear equation. For each subsequent period, one sine function is added to the sine series. The last optimization period contains a linear function and nine sine functions. Each optimization period is 10 seconds, and the total optimization time is 1 minute and 40 seconds.

The optimization with three of the algorithms in the MOEA Framework has been studied:

- Evolution Strategy (Mezura-Montes & Coello, 2005),
- Genetic Algorithms (Sivanandam & Deepa, 2008),
- Differential Evolution (Price, 2013).

The number of iterations and root mean square error over ten optimization periods when using the Genetic Algorithm are visualized in Fig. 2.

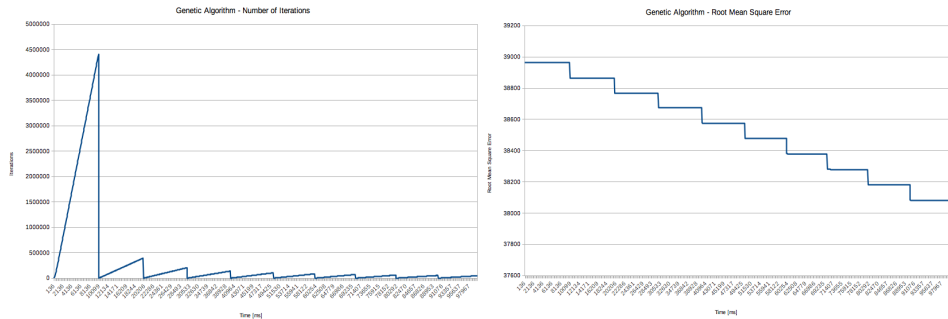


Fig. 2. Genetic Algorithm: a) number of iterations (left), b) RMSE (right)

MOEA Framework implementation of Genetic Algorithm seems not to be efficient at all in the studied problem. The convergence curve is flat for all ten optimization periods (Fig. 2b). The convergence appears stepwise because, with each new function, the RMSE sine decreases.

The number of iterations and root mean square error over ten optimization periods when using the Evolution Strategy are visualized in Fig. 3a.

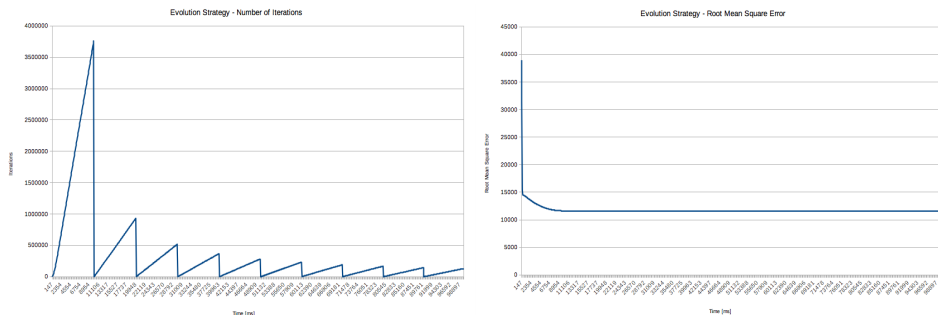


Fig. 3. Evolution Strategy: a) number of iterations (left), b) RMSE (right)

Evolution Strategy is doing less optimization cycles for a fixed period of time (Fig. 3a) than Genetic Algorithm (Fig. 2a). The convergence curve is smooth and goes lower than the Genetic Algorithm convergence curve (Fig. 3b).

The number of iterations and root mean square error over ten optimization periods when using the Differential Evolution are visualized in Fig. 4a.

Time consumption for the Differential Evolution can be estimated as average considering the needed time for performing the other algorithms (Fig. 4a). The convergence curve has smooth downs but also has a stair-like shape in the zones of sine function addition (Fig. 4b).

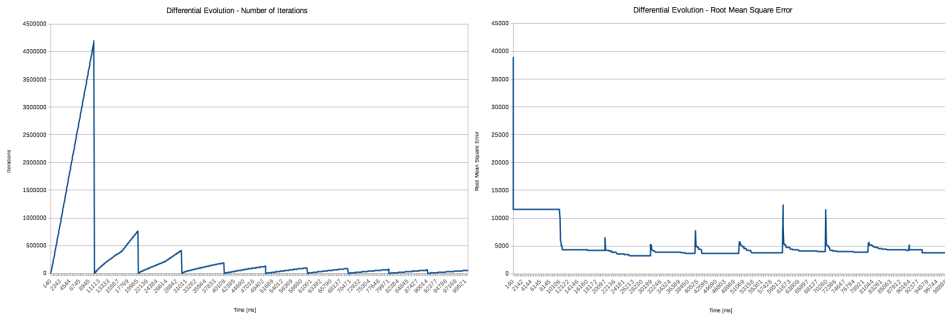


Fig. 4. Differential Evolution: a) number of iterations (left), b) RMSE (right)

The convergence curve of Differential Evolution goes lowest than the others. It means that Differential Evolution achieves the best values for RMSE, compared with the others.

During the experiments, it was noticed that Genetic Algorithm and Differential Evolution are very sensitive to the initialization of the initial population. Initializing these two algorithms with a single identical vector leads to no optimization at all. Evolution Strategy does not suffer from such miss initialization. Among the all tested algorithms, the most time-consuming is the Genetic Algorithm (Fig. 2a). Evolution Strategy is the least time-consuming algorithm than the others (Fig. 3a). Differential Evolution achieves the best minimum values for RMSE, compared with the others, which is the overall goal of optimization. The experimental results show that searching for suboptimal value is faster with shorter individuals in the heuristic population.

4. Conclusion

This study compared the performance of three heuristic optimization algorithms. The algorithms are implemented as a software library in the MOEA Framework. It should be mentioned that the Genetic Algorithm is the most time-consuming algorithm compared to the tested three algorithms, while the Differential Evolution achieves the best minimum values for RMSE that is the optimization goal. The comparison was made when solving a curve-fitting problem on time series. The resulting sine series can be used as a forecasting tool.

Application of optimization algorithms alone proves to be quite effective, but as future research, it would be interesting to experiment with a hybrid algorithm, which combines in a certain scheme presented algorithms, simultaneously.

Acknowledgment

This research is funded by Velbazhd Software LLC and it is partially supported by the Bulgarian National Science Fund by the project “Mathematical models, methods and algorithms for solving hard optimization problems to achieve high security in communications and better economic sustainability”, KP-06-N52/7/19-11-2021.

References

1. Alexandrov, A., Monov, V.: Design of a multi-objective optimization model for wireless sensor networks. In: Georgiev I., Kostadinov H., Lilkova E. (eds) *Advanced Computing in Industrial Mathematics*. BGSIAM 2018. *Studies in Computational Intelligence* 961. Springer, Cham, (2021). https://doi.org/10.1007/978-3-030-71616-5_1.
2. Balabanov, T., Encog and MOEA time series forecasting experiments, <https://github.com/TodorBalabanov/Encog-and-MOEA-Time-Series-Forecasting-Experiments>, last visited 29 Nov 2021.
3. Balabanov, T.: Line and sine functions curve fitting over bitcoin price time series. (2021). <https://doi.org/10.13140/RG.2.2.18259.09769>.
4. Borissova, D., Dimitrova, Z.: An integrated group decision-making approach considering uncertainty conditions. *Business Information Systems* 1, 307–316 (2021). <https://doi.org/10.52825/bis.v1i.52>
5. Cai, X., Hu, Zh., Zhao, P., Zhang, WSh., Chen, J.: A hybrid recommendation system with many-objective evolutionary algorithm. *Expert Systems with Applications* 159, 113648, (2020), <https://doi.org/10.1016/j.eswa.2020.113648>.
6. Dineva, K., Atanasova, T.: Machine learning solution for IoT Big data. In: *Proc. of 20th International Multidisciplinary Scientific GeoConference SGEM'20*, pp. 207–214, (2020). <http://doi.org/10.5593/sgem2020/2.1/s07.027>.
7. Jain, M., Nigam, M., Tiwari, P.: Curve fitting and regression line method based seasonal short term load forecasting. In: *World Congress on Information and Communication Technologies*, pp. 332–337, (2012). <http://doi.org/10.1109/WICT.2012.6409098>.
8. Johnson, N.: A trend line for growth series. *Journal of the American Statistical Association* 30(192), 717–717, (1935). <http://doi.org/10.1080/01621459.1935.10503297>.
9. Marinov, M., Ganev, B., Djermanova, N., Tashev, T.: Analysis of sensors noise performance using Allan deviation. In: *IEEE XXVIII International Scientific Conference Electronics (ET)*, pp. 1–4, (2019). <https://doi.org/10.1109/ET.2019.8878552>.
10. Mezura-Montes, E., Coello, C.: A simple multimembered evolution strategy to solve constrained optimization problems, in *IEEE Transactions on Evolutionary Computation* 9(1), 1–17, (2005). <http://doi.org/10.1109/TEVC.2004.836819>.

11. Michael, L., Lily, L.: A novel method of curve fitting based on optimized extreme learning machine. *Applied Artificial Intelligence* 34(12), 849–865, (2020). <http://doi.org/10.1080/08839514.2020.1787677>.
12. MOEA Framework, <http://moeaframework.org/>, last visited 29 Nov 2021.
13. Petrov, I.: Renewable energies projects selection: block criteria systematization with AHP and Entropy-MOORA methods in MCDM. In: Proc. of 26th Scientific Conference on Power Engineering and Power Machines PEPM'2021, 327, 02004, (2021). <https://doi.org/10.1051/e3sconf/202132702004>.
14. Price, K.: Differential Evolution. In: Zelinka, I., Snasel, V., Abraham, A. (eds) *Handbook of Optimization*. Intelligent Systems Reference Library, vol. 38, Springer, Heidelberg, (2013). https://doi.org/10.1007/978-3-642-30504-7_8.
15. Qian, T., Zhang, L., Li, Z.: Algorithm of Adaptive Fourier Decomposition, in *IEEE Transactions on Signal Processing* 59(12), 5899–5906, (2011). <http://doi.org/10.1109/TSP.2011.2168520>.
16. Sivanandam, S., Deepa, S.: Genetic Algorithms. In: *Introduction to Genetic Algorithms*. Springer, Berlin, Heidelberg, (2008). https://doi.org/10.1007/978-3-540-73190-0_2.
17. Stoyanova, K., Guliashki, V.: *MOEAs for Portfolio Optimization Applications*. LAP Lambert Academic Publishing, 52 pages (2018).
18. TIOBE Index, <https://www.tiobe.com/tiobe-index/>, last visited 29 Nov 2021.
19. Yeboah-Antwi, K., Baudry, B.: Online Genetic Improvement on the java virtual machine with ECSELR. *Genet Program Evolvable Mach* 18, 83–109 (2017). <https://doi.org/10.1007/s10710-016-9278-4>.