Machine learning-based solution of Kepler’s equation

Maozhang Zheng[#], Jianjun Luo, Zhaohui Dang[*]
School of Astronautics, Northwestern Polytechnical University, Xi’an 710072, Shaanxi, China

ABSTRACT

In this study, an analytical solution of elliptical Kepler's equation, which gives the position of a celestial body moving in orbit as a function of time, is designed by using artificial intelligence techniques. For the eccentric anomaly, Kepler’s equation is a transcendental equation with no precise analytical solution. In this paper, a high precision approximate analytical solution is presented to determine eccentric anomaly. The proposed method is based on machine learning where a non-iterative accurate solution is learned from training data. The solution to Kepler’s solution is created using an artificial neural network based on the universal approximation theorem. Simulation results show that this solution is computationally efficient and has a constant complexity.

Keywords: Kepler’s equation, machine learning, neural network

1. INTRODUCTION

Kepler’s equation (KE), which describes how a body moves under the influence of gravity, is derived in orbital mechanics. In his book *Astronomia Nova* [1], Johannes Kepler first discovered it. The elliptical KE is

$$M = E - e\sin E \tag{1}$$

where M, E and e designate mean anomaly, eccentric anomaly and eccentricity, respectively. Both M, E are fundamental parameters for determining the position of a moving celestial body in an elliptical orbit. KE is a transcendental equation because it involves a sine function. The exact analytical solution is unknown. It’s simple to calculate M for a given value of E. However, because there is no closed-form solution, the inverse issue, which involves finding E while M and e is known, can be far more difficult. Usually, E needs to be estimated by series expansions or numerical methods. Kepler himself approximate his equation by simple iteration in 1621 in his book *Epitome of Copernican Astronomy* [1]. KE is one of the core equations and has a lot of applications in orbital mechanics, therefore even though many academics have developed several ways to solve it, this subject continues to draw attention. For KE, finding a simple, accurate, and analytical solution is still of practical importance.

In this paper, the inverse problem is transformed to a machine learning (ML) problem, more exactly a supervised learning problem. By solving the supervised learning problem, a new solution is learned from the pre-calculated data. The proposed method, called ML-based method, takes the advantage of the great flexibility of neural networks (NNs). The approach is appropriate for extensive and quick orbit propagation since the complexity of the suggested algorithm is constant and independent of the eccentricity and transition time.

The rest of this paper is structured as follows. Existing approaches to resolving KE are carefully compiled and reviewed in Section 2. The primary idea behind the suggested machine learning-based solution is described in Section 3. And in Section 4, the effectiveness of the suggested strategy is verified using numerical simulation. Finally, Section 5 provides a summary of the result.

2. METHODS FOR SOLVING KEPLER’S EQUATION

Many academics have looked into the inverse problem of KE. Finding approaches to arrive at the solution with high accuracy and minimal computational expense is the aim. Figure 1 illustrates how the various approaches to solving KE can be categorized. In this section, some very efficient and classical methods are introduced firstly as shown in Table 1. In Section 4, several of these techniques will be contrasted with our new technique in terms of their effectiveness and precision.

[#] zhengmaozhang@mail.nwpu.edu.cn

[*] zhaohuidang@nwpu.edu.cn

Third International Conference on Computer Science and Communication Technology (ICCSCT 2022)
edited by Yingfa Lu, Changbo Cheng, Proc. of SPIE Vol. 12506, 125066B
 · doi: 10.1117/12.2661776

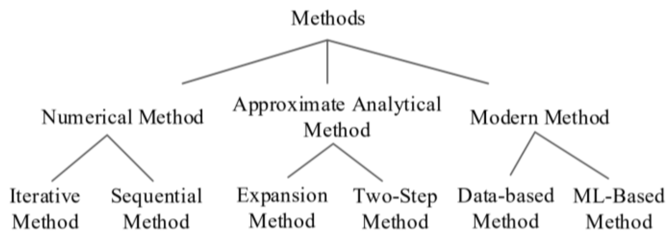

Figure 1. Classification of ways to solve KE.

Table 1. Ways to solve Kepler's equation.

Categories	Methods		Advantages	Disadvantages
Iterative method	Kepler's method [6]		The algorithm is simple	Low convergence speed; initial value required
	Newton's Method and it's variants	Newton's method [2]	Quadratic convergence	Convergence depends on initial value
		Halley's method [4]	Third degree convergence	Second-order derivative is needed; Convergence depends on initial value
		Danby's method [5]	Fourth degree convergence	Third-order derivative is needed; Convergence depends on initial value
	Conway's method [7]		Convergence is guaranteed	Convergence is slow; runtime depends on *e* and *M*
	Mortari's method [8]		High computational efficiency; convergence is guaranteed	Runtime depends on *e* and *M*
Sequential method	Davis et al. [9-11]		Taking the information of the neighborhood point.	Only suitable for the orbit propagation performed at constant time step
Expansion method	Lagrange expansion [2]		More efficient than Fourier-Bessel expansion	Diverge for some value of M when e < 0.662743419
	Fourier-Bessel expansion [2]		Convergence for all eccentricity values	Need to compute Bessel functions of the first kind of order n; many terms are needed for large e and make it very computational expensive
	AD Method [12, 13]		Convergence is faster than Fourier-Bessel expansion	nth-order derivative should be calculated; many terms are needed for large e; and make it very computational expensive
Two-steps method	Bezier curve method [14]		Complexity of the algorithm is constant	Cubic algebraic equations should be solved and select the solution satisfying the condition
	P-C method [15]		Complexity of the algorithm is constant	The expression of the solution is very complex
Data-based method	Fukushima [16] and Feinstein [17]		High computational efficiency	Need pre-computed data and iterative process
ML-based method	The method proposed in this paper		High efficiency; constant complexity; non-iterative; concise expression	

Kepler himself created the first method for resolving his equation, and Newton's method came next. The idea of Newton's method [2, 3] (or Newton-Raphson's Method) is to approximate the KE by the first two terms in a Taylor series expansion. By extending the series to the n-term, a generalized Newton's method is obtained. When $n=3$, the generated solution has an equivalent form to Halley's method [4]. When $n=4$, the generated solution is same to Danby's method [5]. Moreover, truncating the series to the first-order leads to a method which is identical to Kepler's method [6]. Thus, Kepler's method can also be seen as a variant of Newton's method. In order to finding E while M and e is known, the initial value of E should be searched first. Then the repetition is continued until accuracy is satisfied. On the other hand, Newton's technique and its variations may diverge if the original guess is not sufficiently close to the solution. According to Danby [5], as the degree of convergence increases, so do the initial value's sensitivity and the risk of divergence. Although Newton's approach and its variations perform well close to the solution, they lack a feature that would allow them to converge worldwide. The convergence of Conway's method [7] and Mortari's method [8] is guaranteed. However, these methods, like Newton's method and its variants, has different iteration steps for different M and e which means the runtime is dependent on M and e. Except the classification result as shown in Table 1, the methods can also be divided into two categories: single-point method and sequential method. KE is solved using the single-point method, but this method does not benefit from the fact that KE has been computed at the previous neighborhood point. While sequential method [9-11] calculates the present value using the value from the previous moment, making full use of the previous calculation. However, this kind of methods are only applicable to orbit propagation problems with fixed time steps. In addition, the propagation of the initial error causes such algorithms to be sensitive to the error of initial step. Expansion method, including Lagrange expansion [2], Fourier-Bessel expansion [2] and Adomian Decomposition (AD) Method [12, 13], expands the solution of KE into a polynomial consists of N terms. This kind of algorithms are more suitable for the case of small eccentricity, because when the eccentricity is large, more terms need to be reserved in order to improve the accuracy, which leads to low calculation efficiency. The two-step method is a non-iterative method which divides the problem of solving KE into two steps. First, calculate an initial estimate by using a technology, such as Bezier Curve [14] and series expansion [15]. Then correct the initial value to a high accuracy using a generalized Newton's method which is applied only once, rather than in a loop. Data-based method [16, 17] determines a start value according to the pre-computed data and then using iterative method to correct it.

For a number of issues in Celestial Mechanics, KE must be solved numerous times. Due to this, it is crucial for these applications to efficiently compute mean anomaly. For instance, in trajectory planning problems of the relative motion, the true anomaly should be calculated repeatedly which involves solving KE. The accuracy and speed of trajectory planning are significantly impacted by the effectiveness and stability of the eccentric anomaly calculation solution. Therefore, in this paper a method based on machine learning with low computation cost and constant complexity is presented.

3. MACHINE LEARNING-BASED SOLUTION

This section presents an innovative method for solving KE. E is a function of M and e, defined as $E=h(M,e)$. The function's precise formulation cannot be expressed in closed form. Since 2π radians are swept off per period T, the eccentric anomaly may be calculated for any M in an infinite interval $[0,+\infty)$ by

$$E=h([\frac{M}{2\pi}],e)+M-[\frac{M}{2\pi}] \tag{2}$$

where $[\cdot]$ denotes the remainder and $0\le[M/2\pi]\le 2\pi$. In addition, the function $E=h(M,e)$ is centrosymmetric about $E=M=\pi$, namely

$$h(2\pi-M,e)=2\pi-h(M,e)\,,\ 0\le M\le\pi\,,\ \pi\le 2\pi-M\le 2\pi \tag{3}$$

Therefore, the eccentric anomaly in the whole interval may be derived as long as the function $E=h(M,e)$ is computed in the interval $0\le M\le\pi$.

In supervised learning, a machine learning job, an input is mapped to an output using examples of input-output pairings. With each training example including an input item and the desired output value, or "labelled training examples," a function is created. A supervised learning algorithm produces an inferred function by looking at the training data. The inferred function can be used to approximate the unknown function $E = h(M, e)$ in the inverse problem.

3.1 Feedforward neural network

Artificial NNs, or NNs for short, are a mathematical algorithmic model for distributed parallel information processing that replicates the behavioral properties of animal neural networks. Such networks rely on the system's complexity to process information by altering the interaction between a large number of internally linked nodes.

Universal approximation theory (UAT) [18] pointed out that any bounded and regular function from one finite-dimension space to another can be approximated by an ordinary multilayer feedforward NN with any required accuracy. The network must have a sufficient number of neurons with a linear output layer and at least one hidden layer.

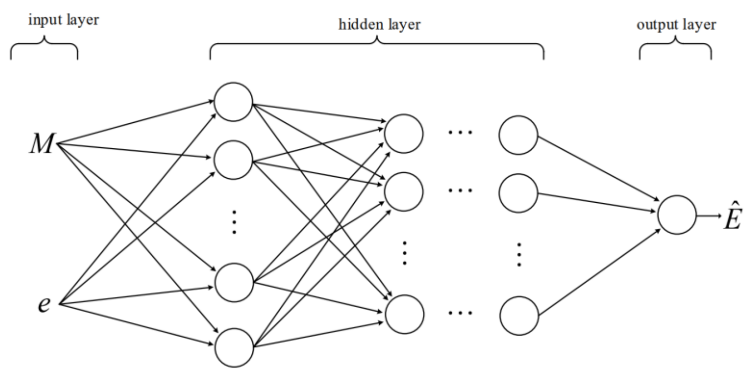

Figure 2. The NN for solving KE.

As shown in Figure 2, a feedforward NN may be used to estimate an unknown function $h(M, e)$ and produce an analytical solution with high precision. For the approximation problem of the function $E = h(M, e)$, there are two inputs and one output. One may think of the network as a composite function $g(M, e)$. Thus, the estimate of eccentric anomaly is designed as

$$\hat{E} = g(M, e) \tag{4}$$

In hidden layers, one of the squashing functions, the hyperbolic tangent sigmoid transfer function, is used in this research as the active function

$$f_l(x) = \frac{2}{1 + e^{-2x}} - 1,\ l = 1, \ldots, L-1 \tag{5}$$

and in the output layer, the activation function is

$$f_L(x) = x \tag{6}$$

3.2 Generation of training data

Due to the lack of an accurate solution to the eccentric anomaly, we are unable to get the exact value E when given M and e. Calculating M when given E and e is, fortunately, simple. As a result, Algorithm 1 is made to acquire the training data.

Algorithm 1: Data generation for training

Input: The interval of true anomaly, $E_d \le E \le E_u$, and the interval of eccentricity, $e_d \le e \le e_u$;

Output: Training data composed by N training examples $(M_k, e_k; E_k),\ k = 1, \ldots N$;

Discrete the interval of E into $N_E - 1$ equal parts;

Discrete the interval of e into $N_e - 1$ equal parts;

$k = 1$

for $i = 1; i \le N_E; i++$

  for $j = 1; j \le N_e; j++$

    $M_{i,j} \leftarrow E_i - e_j \sin E_i$

    $(M_k, e_k; E_k) \leftarrow ([M_{i,j}, e_j], E_i)$

    $k++$

  end

end

Return $(M_k, e_k; E_k), k = 1, ...N, N = N_E \cdot N_e$

3.3 Final adjustment

UAT shows that given enough neurons, we can obtain solutions to KE with arbitrary accuracy. Too many neurons can make it difficult to train the network. Therefore, in order to reduce the training difficulty of the network, a small network is chosen as the learning model in this paper. In this case, although an estimate $\hat{E}$ very close to the true value E can be obtained, the accuracy may not be high enough. To improve the accuracy, a single-step adjustment algorithm is applied

$$\hat{E}' = \hat{E} + \delta \tag{7}$$

The single-step adjustment algorithm designed by Halley's method [4] is as follows

$$\delta = -\frac{2(\hat{E} - e\sin\hat{E} - M)(1 - e\cos\hat{E})}{2(1 - e\cos\hat{E})^2 - (\hat{E} - e\sin\hat{E} - M)e\sin\hat{E}} \tag{8}$$

To increase the solution's accuracy, a single-step adjustment is performed. The analytical solution's complexity does not change because there is only one iteration.

4. NUMERICAL SIMULATION

As of October 2020, one hundred and ninety-nine elliptical orbit objects with eccentricity greater than 2 were discovered in the publicly accessible data (www.space-track.org). Excentricity is 0.898411 at its highest level. Therefore, the greatest limit of eccentricity, according to our assumptions, is 0.9. Decompose $E \in [0, \pi]$ and $e \in [0, 0.9]$ 2 into smaller intervals of equal size using eighty points respectively. 6400 training examples are obtained using Algorithm 1.

In this paper, we verify the performance of the proposed algorithm using a three-layer neural network that contains two hidden layers, each with nine neurons. Using Matlab toolbox, train the network with 6400 training examples and the residuals of the network at the examples are shown in Figure 3.

In practical engineering application, the inverse problems are solved with given eccentricities for different mean anomalies. Therefore, we take the max error in computed E under a given eccentricity, as shown in Figure 4, as index of accuracy of the algorithm. As shown in Figure 4 the algorithm proposed in this paper has high accuracy with final adjustment. When $e = 0.89$, the worst accuracy of 6.4×10^{-11} is obtained. For $e < 0.7$ the ML-based method has a precision of 10^{-15}.

The series methods and the ML-based method can be classified into one category based on the fact that these two kinds of methods both use known models to approximate unknown functions. Figures 5 and 6 show that for small eccentricity the Lagrange method and Fourier-Bessel method can obtain same precision with ML-based method even with a few terms. For large eccentricity the ML-based method has a higher precision. Additionally, the Lagrange series diverges when it surpasses 0.662743419, which suggests that adding additional terms produces poorer outcomes. For large eccentricity the number of items of Lagrange method and Fourier-Bessel method is greatly increased. This makes Lagrange method and Fourier-Bessel method very computational expensive as shown in Table 2. Table 2 shows that ML-based method is more efficient than the Lagrange method and Fourier-Bessel method. With increasing of the number of terms used for improving solutions' accuracy, the mean runtime of ML-based method is nearly constant and very small. However, the mean runtime of Lagrange method and Fourier-Bessel method increased dramatically.

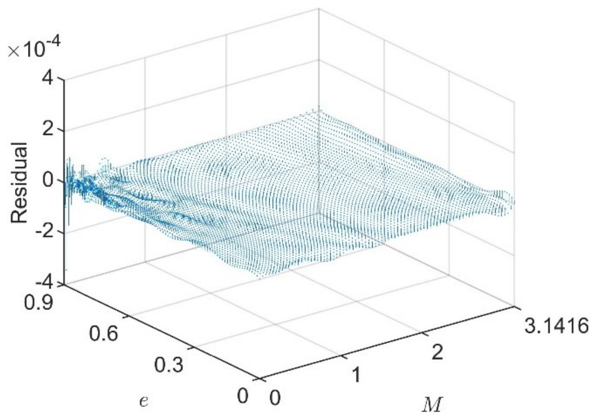

Figure 3. Errors $|\hat{E}-E|$ in computed E.

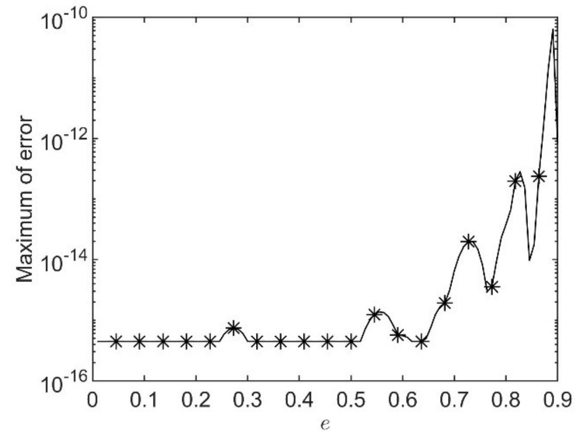

Figure 4. Maximum of errors $|\hat{E}-E|$ in computed E under given eccentricities.

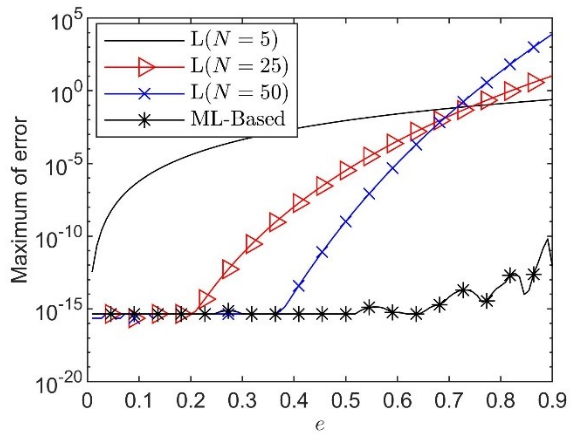

Figure 5. Comparison of the Lagrange method's accuracy with the ML-based approach.

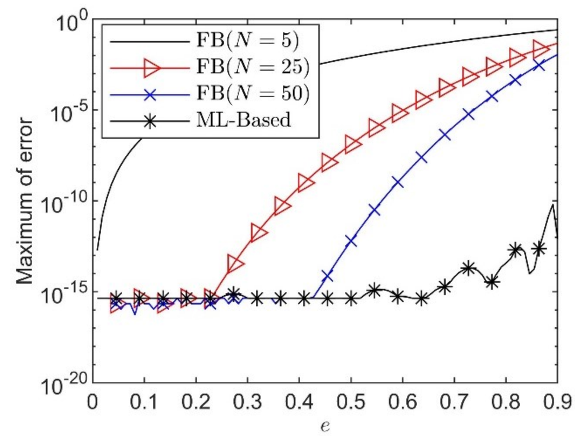

Figure 6. Comparison of the Fourier-Bessel technique's accuracy with the ML-based approach.

Table 2. Mean runtime of Lagrange method, FB method and ML-based method.

Method	**Mean runtime,** *ms*		
ML-based	0.98		
Lagrange method	38.9($N=5$)	608.8($N=25$)	2319.2($N=50$)
Fourier-Bessel method	264.2($N=5$)	1819.2($N=25$)	4744.7($N=50$)

5. CONCLUSION

In this paper a non-iterative method based on machine learning for solving KE is presented. Based on neural networks' tremendous flexibility, this technique can offer a high precision solution to KE. It can be seen as an analytical solving tool because the new method does not require any iterative computation or numerical integration. It also avoids the shortcoming of sensitivity of initial guess in some classical methods for solving KE. Numerical accuracy tests show that, the new algorithm behaves better than the most existed methods both in computational efficiency and accuracy. The approach outlined in this work, aside from the elliptical instance, may also be used to parabolic case and hyperbolic case.

REFERENCES

[1] Colwell, P., [Solving Kepler's Equation over Three Centuries], Willmann-Bell, Richmond, VA, (1993).

[2] Battin, R. H., "An introduction to the mathematics and methods of astrodynamics," AIAA, (1999).

[3] Raposo-Pulido, V. and Pelaez, J., "An efficient code to solve the Kepler's equation hyperbolic case," Astronomy & Astrophysics, 619, A129(2018).

[4] Gander, W., "On Halley's iteration method," The American Mathematical Monthly, 92(2), 131-134(1985).

[5] Danby, J. M. A. and Burkardt, T. M., "The solution of Kepler's equation," I. Celestial Mechanics, 31(2), 95-107(1983).

[6] Swerdlow, N. M., "Kepler's iterative solution to Kepler's equation," Journal for the History of Astronomy, 31(4), 339-341(2000).

[7] Conway, B. A., "An improved algorithm due to Laguerre for the solution of Kepler's equation," Celestial mechanics, 39(2), 199-211(1986).

[8] Mortari, D. and Elipe, A., "Solving Kepler's equation using implicit functions," Celestial Mechanics and Dynamical Astronomy, 118(1), 1-11(2014).

[9] Davis, J. J., Bruccoleri, C. and Mortari, D., "Quasi constant-time orbit propagation without solving Kepler's equation," Proc. of the 2008 Space Flight Mechanics Meeting Conf., (2008).

[10] Mortari, D., Davis, J. J. and Bruccoleri, C., "Fast orbit propagation without solving Kepler's equation," Proc. of the 2009 Space Flight Mechanics Meeting Conf., (2009).

[11] Davis, J. J., Mortari, D. and Bruccoleri, C., "Sequential solution to Kepler's equation," Celestial Mechanics and Dynamical Astronomy, 108(1), 59-72(2010).

[12] Alshaery, A. and Ebaid, A., "Accurate analytical periodic solution of the elliptical Kepler equation using the Adomian decomposition method," Acta Astronautica, 140, 27-33(2017).

[13] Aljohani, A. F., Rach, R., El-Zahar, E., Wazwaz, A. M. and Ebaid, A., "Solution of the hyperbolic Kepler equation by Adomian's asymptotic decomposition method," Rom. Rep. Phys. 70(2), 112(2018).

[14] Mortari, D. and Clocchiatti, A., "Solving Kepler's equation using Bézier curves," Celestial Mechanics and Dynamical Astronomy, 99(1), 45-57(2007).

[15] Pimienta-Penalver, A. R., [Accurate Kepler Equation Solver without Transcendental Function Evaluations], State University of New York at Buffalo, Buffalo, Master's Thesis, (2013).

[16] Fukushima, T., "A method solving Kepler's equation without transcendental function evaluations," Celestial Mechanics and Dynamical Astronomy, 66(3), 309-319(1996).

[17] Feinstein, S. A. and McLaughlin, C. A., "Dynamic discretization method for solving Kepler's equation," Celestial Mechanics and Dynamical Astronomy, 96(1), 49-62(2006).

[18] Hornik, K., Stinchcombe, M. and White, H., "Multilayer feedforward networks are universal approximators," Neural Networks, 2(5), 359-366(1989).