Chapter 9
Curve Interpolation and Financial Curve Construction

Pengfei Huang, Haiyan Wang, Ping Wu, and Yifei Li

Abstract In this chapter, first we introduce some commonly used curve interpolation methods for interest rate curves. Then, we present a positivity-preserving piecewise rational cubic interpolation function. It is constructed to ensure positive values by adjusting the shape-control parameters. When it is applied to the interest rate curve construction, this interpolation algorithm ensures positive values. The market data has been reconstructed to restrict the fluctuation of the interest rate curve when the market data changes sharply.

9.1 Introduction

Interest rate curves are the foundation of the pricing of interest rate derivatives, and reflect the market's expectation of future interest rate development [1]. The market interest rate data is usually a set of discrete time sequences with different gaps, by which one can construct an interest rate curve with various interpolation algorithms. These algorithms should guarantee the interest rate curve satisfies continuity and smoothness conditions and accurately reflects the current market.

The cubic spline interpolation method [2] proposed by McCulloch (1975) is a popular choice for interest rate curve construction, and was developed by Fisher et al. (1995) to ensure the smoothness and goodness of fit of the interest rate curve [3]. However, the interest rate curve constructed by cubic spline interpolation is not guaranteed to be positive and is not arbitrage free. Lu Jun and Han Xuli (2005) proposed a rational cubic spline with cubic denominator and numerator which can preserve the monotonicity of the inputs [4]. Tian Meng (2006) proposed a rational cubic spline with cubic numerator and quadratic denominator which can guarantee the positivity of the constructed curve [5]. Based on the existence of rational splines,

P. Huang • H. Wang (✉)
School of Economics and Management, Southeast University, Nanjing, China
e-mail: 220141843@seu.edu.cn; hywang@seu.edu.cn

P. Wu • Y. Li
BMO Financial Group, Bank of Montreal, Toronto, ON, Canada
e-mail: Ping.Wu@bmo.com; Stephanie.li@bmo.com

L.M. Chen et al., *Mathematical Problems in Data Science*,
DOI 10.1007/978-3-319-25127-1_9

in this paper we propose a rational cubic spline with a cubic denominator and numerator that can ensure the constructed interest rate curve is positive.

9.2 Basic Interpolation Methods for Interest Rate Curves

The term structure of interest rates is fundamental to consistently value interest rate instruments. The term structure can refer to different curves: the discount curve, zero curve, or forward curve. However, we cannot obtain these curves directly from market data; instead, we use discrete dates and market quotes, and a wide range of models and numerical methods are employed to fit a continuous function to a set of discretely observed quotes.

As a brief review, we provide the mathematical background and notation for the term structure. There are three equivalent descriptions: the discount function $d(t)$, the yield curve $r(t)$, and instantaneous forward rate curve $f(t)$.

$$d(t) = e^{-r(t)t}$$

is equivalent to

$$d(t) = e^{-\int_0^t f(u)du}$$

where

$$f(t) = \frac{\partial}{\partial t} r(t)t.$$

9.2.1 Linear Interpolation

Suppose we have market quotes $r_1, r_2, \ldots, r_n$ at times $t_1, t_2, \ldots, t_n$. The objective is to determine r_t for any time $t \in [t_1, t_n]$. A straightforward method is linear interpolation. The interpolation can be linear on rates, discount factors, and the transformation of discount factors. Some common approaches are as follows:

First, we can give a formula for the zero rates. We determine i for time t such that $t \in [t_i, t_{i+1}]$, and r_t is calculated as the linear interpolation of r_i and r_{i+1}:

$$r_t = \frac{(t_{i+1} - t)}{(t_{i+1} - t_i)} r_i + \frac{(t - t_i)}{(t_{i+1} - t_i)} r_{i+1}.$$

For the instantaneous forward curve $f(t) = \frac{\partial}{\partial t} r(t)t$, we have

$$f(t) = \frac{(t_{i+1} - 2t)}{(t_{i+1} - t_i)} r_i + \frac{(2t - t_i)}{(t_{i+1} - t_i)} r_{i+1}.$$

When it is linear on the discount factors, we can apply the linear interpolation on the discount factors directly. Since the discount factors are defined as $d(t) = e^{-r(t)t}$, we have a set of discount factors $d_1, d_2, \ldots, d_n$ corresponding to $r_1, r_2, \ldots, r_n$. The formula is

$$d(t) = \frac{(t_{i+1} - t)}{(t_{i+1} - t_i)} d_i + \frac{(t - t_i)}{(t_{i+1} - t_i)} d_{i+1}.$$

According to

$$r(t) = \frac{-\ln(d(t))}{t}$$

we have

$$r(t) = \frac{-1}{t} \ln\left(\frac{(t_{i+1} - t)}{(t_{i+1} - t_i)} d_i + \frac{(t - t_i)}{(t_{i+1} - t_i)} d_{i+1}\right).$$

Then, the forward rate can be calculated as

$$f(t) = \frac{(d_i - d_{i+1})}{((t_{i+1} - t)d_i + (t - t_i)d_{i+1})}. \tag{9.1}$$

If discount factors are linear on the logarithmic rate, we can derive a formula accordingly.

9.2.2 Cubic Spline Interpolation

As well as considering the same set of assumptions as in linear interpolation, spline interpolation requires a separate cubic polynomial for each interval. The aim is to determine the coefficients and the value of the rate at time t:

$$r(t) = a_i + b_i(t - t_i) + c_i(t - t_i)^2 + d_i(t - t_i)^3, \quad t \in [t_i, t_{i+1}].$$

In order to determine the set of coefficients $\{a_i, b_i, c_i, d_i,\ i = 1, 2, \ldots, n-1\}$, constraints on the value, the first derivative, and the second derivatives at t_i should be satisfied.

First, the value of the piecewise polynomial should be identical at the knots (t_i), which gives the following system of equations:

$$r_i = a_i,$$

$$r_{i+1} = a_i + b_i(t_{i+1} - t_i) + c_i(t_{i+1} - t_i)^2 + d_i(t_{i+1} - t_i)^3.$$

Since

$$f(t) = \frac{\partial}{\partial t} r(t)t = \frac{\partial r(t)}{\partial t} + r(t),$$

and

$$\frac{\partial r(t)}{\partial t} = b_i + 2c_i(t - t_i) + 3d_i(t - t_i)^2, \quad t \in [t_i, t_{i+1}],$$

the piecewise polynomial should be differentiable to guarantee the continuity of the instantaneous forward rate curve $f(t)$. Thus, we have

$$b_{i+1} = b_i + 2c_i(t_{i+1} - t_i) + 3d_i(t_{i+1} - t_i)^2, \quad i = 1, 2, \ldots, n-2.$$

Combining the $2(n-2)$ conditions above, we now have $3n-4$ equations for $4n-4$ parameters. We still need another n constraints to obtain all of the parameters. We can use the so-called natural cubic splines to do this based on the second derivatives:

$$\frac{\partial^2 r(t)}{\partial t^2} = 2c_i + 6d_i(t - t_i), \quad t \in [t_i, t_{i+1}].$$

These cubic splines are chosen so that the second derivatives of the functions at the knots are continuous and the second derivatives of the end points equal zero, which gives n linear equations in the system:

$$c_{i+1} = c_i + 3d_i(t_{i+1} - t_i), \quad i = 1, 2, \ldots, n-2,$$
$$c_1 = 0,$$
$$c_{n-1} + 3d_{n-1}(t_n - t_{n-1}) = 0.$$

These splines guarantee the continuity and smoothness of the curve.

9.2.3 *Monotone and Convex Splines*

A simple observation is that the cubic splines in the last subsection could unexpectedly generate negative forward rates. The monotone convex spline method is introduced to build a continuous, smooth curve that ensures positive forward rates between the observations. Thus, the curve is arbitrage free. First, we calculate the forward rates $f_i^d, i = 1, 2, \ldots, n$ from the market quotes $r_1, r_2, \ldots, r_n$ at times $t_1, t_2, \ldots, t_n$, assuming forward rates are constant in each interval and are continuously compounded [6].

Various methods can be applied to fit a curve to the term structure of the interest rate; goodness of fit and smoothness of the curves are two measures of success. All such methods have strengths and weaknesses, and their selection depends on the specific requirement of the application.

In the following, we propose a rational cubic interpolation method for positive forward rates, which we call positivity preservation.

9.3 Rational Cubic Interpolation Method for Positivity Preservation

Let $a = \tau_1 < \tau_2 < \ldots < \tau_n = b$ be a set of time knots, $\theta = \frac{\tau - \tau_i}{\tau_{i+1} - \tau_i}$, $h_i = \tau_{i+1} - \tau_i$, and $f(\tau_i) = f_i (i = 1, 2, \ldots, n)$ be the input data at time knot τ_i. Define the rational cubic spline interpolation function in the i^{th} interval $[\tau_i, \tau_{i+1}]$ as

$$S_i(\tau) = \frac{P_i(\theta)}{Q_i(\theta)}, \quad i = 1, 2, \ldots, n-1, \tag{9.2}$$

where

$$P_i(\theta) = (1-\theta)^3 f_i + \theta(1-\theta)^2 A + \theta^2(1-\theta)B + \theta^3 f_{i+1} \tag{9.3}$$

$$Q_i(\theta) = (1-\theta)^3 + v_i(1-\theta)^2\theta + u_i(1-\theta)\theta^2 + \theta^3. \tag{9.4}$$

In Eq. (9.4) $v_i > 0$ and $u_i > 0$ are shape-control parameters, and d_i is the first derivative of the function $f(\tau)$ at time knot τ. To fulfill the condition $S(x) \in C^1[a, b]$, the interpolation function defined by Eq. (9.2) should satisfy the following conditions:

$$\begin{cases} S_i(\tau_i) & = f_i \\ S_i(\tau_{i+1}) & = f_{i+1} \\ S_i'(\tau_i) & = d_i \\ S_i'(\tau_{i+1}) & = d_{i+1}. \end{cases} \tag{9.5}$$

The terms A and B are given by

$$\begin{cases} A = d_i h_i + v_i f_i \\ B = -d_{i+1} h_i + u_i f_{i+1}. \end{cases} \tag{9.6}$$

The first derivative at time knot τ_i, d_i, is unknown in the market and is estimated as follows. We take two knots from both sides of τ_i (if one side does not have two, we take more from the other side). Let $\Delta\tau_j = \tau_j - \tau_i$, assume the input data at the knots is $f_{(j)} (j = 1, 2, 3, 4)$, and define

$$\begin{cases} \lambda_0 = -\sum_{j=1}^{4} \lambda_j \\ \lambda_j = \frac{\Pi_{k=1}^{4} \Delta\tau_k}{\Delta\tau_j^2 \Pi_{\substack{k=1 \\ k \neq j}}^{4} (\Delta\tau_k - \Delta\tau_j)}, \quad j = 1, 2, 3, 4. \end{cases} \tag{9.7}$$

Then, d_i can be estimated by

$$d_i = \lambda_0 f_i + \sum_{j=1}^{4} \lambda_j f_{(j)}. \tag{9.8}$$

The accuracy of the estimation is $O(\max\limits_{1 \leq j \leq 4} |\Delta\tau_j^4|)$ [7], which is good. This will improve the smoothness and help retain the monotonicity of the inputs to the interest rate curve. However, as seen in Eq. (9.8), the estimation of the first derivative at knot τ_i is achieved using four nearby points, so changes at knot τ_i affect the shape of the curve considerably. The positivity-preserving property is analyzed as follows. Since $\theta = \frac{\tau - \tau_i}{\tau_{i+1} - \tau_i}$ and $\tau \in [\tau_i, \tau_{i+1}]$, θ is in the interval $[0, 1]$. Hence, a sufficient condition for the positivity-preserving property of the interpolation function defined by Eq. (9.2) is $A > 0, B > 0$, and the value range of the shape-control parameters can be obtained as follows:

$$\begin{cases} v_i > \max(-\frac{d_i h_i}{f_i}, 0), \\ u_i > \max(\frac{d_{i+1} h_i}{f_{i+1}}, 0). \end{cases} \tag{9.9}$$

When v_i and u_i satisfy these conditions (9.9), the curve constructed with the rational cubic spline interpolation method will always be positive and arbitrage free. The error estimation of the rational cubic spline interpolation is analyzed as follows. It is sufficient to only deal with the case where the knots are equally spaced. Without loss of generality, we only consider the subintervals $[\tau_i, \tau_{i+1}]$ $(3 \leq i \leq n-2)$, and when the knots are equally spaced let $h_i = h$. By Eqs. (9.7) and (9.8), we get

$$\lambda_1 = \frac{1}{12h}, \lambda_2 = -\frac{2}{3h}, \lambda_3 = \frac{2}{3h}, \lambda_4 = -\frac{1}{12h}, \lambda_0 = 0, \tag{9.10}$$

$$d_i = \frac{1}{12h} f_{i-2} - \frac{2}{3h} f_{i-1} + \frac{2}{3h} f_{i+1} - \frac{1}{12h} f_{i+2}. \tag{9.11}$$

Then, the numerator of the interpolation function in Eq. (9.3) can be transformed into the following form:

$$P_i(\theta) = l_1 f_{i-2} + l_2 f_{i-1} + l_3 f_i + l_4 f_{i+1} + l_5 f_{i+2} + l_6 f_{i+3}, \tag{9.12}$$

where

$$\begin{cases} l_1 = \frac{1}{12}\theta(1-\theta)^2 \\ l_2 = \theta(1-\theta)[-\frac{2}{3}(1-\theta) - \frac{1}{12}\theta] \\ l_3 = (1-\theta)[(1-\theta)^2 + \theta(1-\theta)v_i + \frac{2}{3}\theta^2] \\ l_4 = \theta[\theta(1-\theta)u_i + \theta^2 + \frac{2}{3}(1-\theta)^2] \\ l_5 = -\theta(1-\theta)[\frac{1}{12}(1-\theta) + \frac{2}{3}\theta] \\ l_6 = \frac{1}{12}\theta^2(1-\theta). \end{cases} \tag{9.13}$$

When $f(t) \in C^1[a,b]$ and $S(t)$ is the rational cubic interpolation function of $f(t)$ in the subinterval $[\tau_i, \tau_{i+1}]$, using the Peano–Kernel Theorem gives

$$R[f] = f(t) - S(t) = \int_{\tau_{i-2}}^{\tau_{i+3}} f^{'}(\tau) R_t[(t-\tau)_+^0] d\tau, \tag{9.14}$$

where

$$R_t[(t-\tau)_+^0] = \begin{cases} 1 - \frac{l_2+l_3+l_4+l_5+l_6}{Q_i(\theta)}, & \tau_{i-2} < \tau < \tau_{i-1} \\ 1 - \frac{l_3+l_4+l_5+l_6}{Q_i(\theta)}, & \tau_{i-1} < \tau < \tau_i \\ 1 - \frac{l_4+l_5+l_6}{Q_i(\theta)}, & \tau_i < \tau < t \\ -\frac{l_4+l_5+l_6}{Q_i(\theta)}, & t < \tau < \tau_{i+1} \\ -\frac{l_5+l_6}{Q_i(\theta)}, & \tau_{i+1} < \tau < \tau_{i+2} \\ -\frac{l_6}{Q_i(\theta)}, & \tau_{i+2} < \tau < \tau_{i+3}. \end{cases} \tag{9.15}$$

From Eq. (9.14) we get the following result:

$$\begin{aligned} |R[f]| = |\int_{\tau_{i-2}}^{\tau_{i+3}} f^{'}(t) R_t[(t-\tau)_+^0] d\tau| \le |f^{'}(t)| [\int_{\tau_{i-2}}^{\tau_{i-1}} |1 - \frac{l_2+l_3+l_4+l_5+l_6}{Q_i(\theta)}| d\tau \\ + \int_{\tau_{i-1}}^{\tau_i} |1 - \frac{l_3+l_4+l_5+l_6}{Q_i(\theta)}| d\tau + \int_{\tau_i}^{t} |1 - \frac{l_4+l_5+l_6}{Q_i(\theta)}| d\tau \\ + \int_{t}^{\tau_{i+1}} |-\frac{l_4+l_5+l_6}{Q_i(\theta)}| d\tau + \int_{\tau_{i+1}}^{\tau_{i+2}} |-\frac{l_5+l_6}{Q_i(\theta)}| d\tau \\ + \int_{\tau_{i+2}}^{\tau_{i+3}} |-\frac{l_6}{Q_i(\theta)}| d\tau] \le h |f^{'}(t)| w(v_i, u_i, \theta), \end{aligned} \tag{9.16}$$

where

$$w(v_i, u_i, \theta) = 3 + \frac{1}{Q_i(\theta)} (|l_2+l_3+l_4+l_5+l_6| + |l_3+l_4+l_5+l_6| + 2|l_4+l_5+l_6| + |l_5+l_6| + |l_6|). \tag{9.17}$$

For a given v_i and u_i, let $e_i = \max_{0<\theta<1} w(v_i, u_i, \theta)$. This term is independent of the subinterval $[\tau_i, \tau_{i+1}]$, and only dependent on the parameters v_i and u_i. The values

Table 9.1 The values of e_i for $v_i = 0.001$ with various values of u_i

v_i	u_i	e	θ
0.001	0.001	6.9862	0.7040
0.001	0.005	6.9862	0.7040
0.001	0.010	6.9863	0.7040
0.001	0.050	6.9865	0.7039
0.001	0.100	6.9867	0.7037
0.001	1.000	6.9901	0.7015
0.001	10.000	6.9972	0.6974
0.001	100.000	6.9997	0.6961
0.001	1000.000	7.0000	0.6959

Table 9.2 The values of e_i for $u_i = 0.001$ with various values of v_i

v_i	u_i	e	θ
0.005	0.001	6.9849	0.7060
0.010	0.001	6.9833	0.7086
0.050	0.001	6.9720	0.7322
0.100	0.001	7.0000	0.9999
1.000	0.001	7.0000	1.0000
10.000	0.001	7.0000	0.9999
100.000	0.001	7.0000	0.9999
1000.000	0.001	7.000	0.9999

e_i, for various v_i and u_i, are given in Tables 9.1 and 9.2. Based on the experimental results in Tables 9.1 and 9.2, we can see the value of e_i only changes slightly for $v_i > 0$ and $u_i > 0$. This demonstrates that this interpolation method is stable.

9.4 Reconstruction of the Market Interest Rate Data

Table 9.3 details the USD market data from 04/01/2000. In Table 9.3, the first column is the time from start to expiration, the second column is the expiration date, the third column is the market rate, and the last column is the time in years (365 days=1 year). To restrict the fluctuation of the interest rate curve when the input data changes sharply, we reconstruct the market data before applying the rational cubic interpolation method. Let (τ_i, r_i) $(i = 1, 2, \ldots, n)$ be a given set of market data, where τ_i are the time knots. We assume r_i is the value corresponding to the entire subinterval $[\tau_{i_1}, \tau_i]$ $(i = 2, \ldots, n)$. The reconstructed rate is f_i at the point τ_i $(i = 2, \ldots, n-1)$:

$$f_i = \frac{\tau_i - \tau_{i-1}}{\tau_{i+1} - \tau_{i-1}} r_{i+1} + \frac{\tau_{i+1} - \tau_i}{\tau_{i+1} - \tau_{i-1}} r_i. \tag{9.18}$$

Table 9.3 USD market data in 2000

Time length	Expiration date	Market interest r (%)	Time length τ
7 day	13-Jan-00	5.53125	0.019
1 month	7-Feb-00	5.81250	0.088
3 months	6-Apr-00	6.03125	0.249
6 months	6-Jul-00	6.21875	0.499
12 months	8-Jan-01	6.59375	1.008
2 yr	7-Jan-02	6.8950	2.005
3 yr	6-Jan-03	7.0250	3.003
4 yr	6-Jan-04	7.0850	4.003
5 yr	6-Jan-05	7.1350	5.005
6 yr	6-Jan-06	7.1750	6.005
7 yr	8-Jan-07	7.2250	7.011
8 yr	7-Jan-08	7.2650	8.008
9 yr	6-Jan-09	7.2950	9.008
10 yr	6-Jan-10	7.3350	10.008
12 yr	6-Jan-12	7.3850	12.008
15 yr	6-Jan-15	7.4350	15.011
20 yr	6-Jan-20	7.4450	20.014
25 yr	6-Jan-25	7.4450	25.019
30 yr	7-Jan-30	7.4350	30.025

We add the additional subintervals $[\tau_0, \tau_1]$ and $[\tau_n, \tau_{n+1}]$ at the start and end, where

$$\begin{cases} \tau_0 &= \tau_1 - (\tau_2 - \tau_1) \\ \tau_{n+1} &= \tau_n + (\tau_n - \tau_{n-1}) \end{cases} \tag{9.19}$$

and

$$r_{n+1} = r_n + \frac{\tau_n - \tau_{n-1}}{\tau_n - \tau_{n-2}}(r_n - r_{n-1}). \tag{9.20}$$

Then, the reconstructed market data f_i $(i = 1, 2, \ldots, n)$ can be obtained from Eq. (9.18) [6]. The results of applying the reconstruction algorithm to the market data (from Table 9.3) are given in Table 9.4. The third column in Table 9.4 is the reconstructed market data.

9.5 Analysis of the Interest Curve Constructed by Rational Cubic Interpolation

The constructed interest rate curve, found by applying the rational cubic interpolation method to the reconstructed market data, is given in Fig. 9.1. In general, the term structure contains more short-term information, so the short term of the

Table 9.4 USD market data in 2000

Serial number	Time length τ	Market rare f
1	0.019	0.05671875
2	0.088	0.05878125
3	0.249	0.06104699
4	0.499	0.06342268
5	1.008	0.06695567
6	2.005	0.06959967
7	3.003	0.07054970
8	4.003	0.07109975
9	5.005	0.07155020
10	6.005	0.07199925
11	7.011	0.07245090
12	8.008	0.07279977
13	9.008	0.07315000
14	10.008	0.07351667
15	12.008	0.07404988
16	15.011	0.07438751
17	20.014	0.07445000
18	25.019	0.07440000
19	30.025	0.07432500

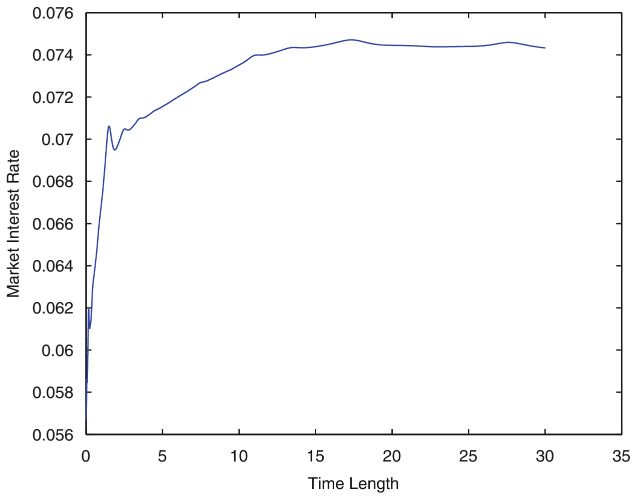

Fig. 9.1 Interest curve constructed by applying rational cubic interpolation to the reconstructed market interest rate data

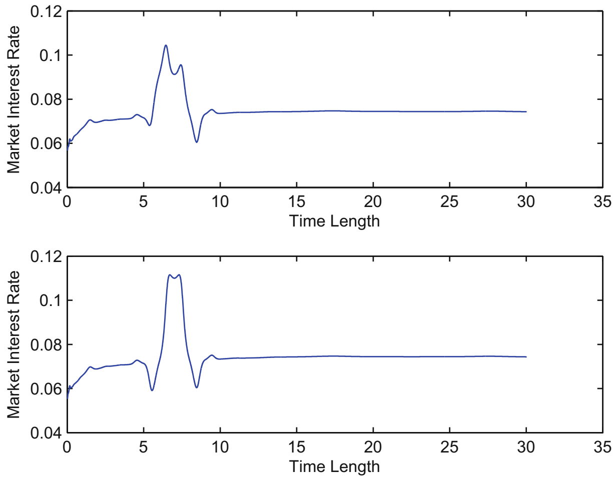

Fig. 9.2 The interest curve constructed by rational cubic interpolation with and without the reconstruction of market interest rate data

interest rate curve should focus on goodness of fit while the long term of the curve requires smoothness [8]. The interest rate curve constructed by the rational cubic interpolation method essentially satisfies these requirements. We now discuss the stability of the interest rate curve; when rational cubic interpolation is applied to the reconstructed market data the data in Table 9.3 changes from $(7.011, 7.2250\,\%)$ to $(7.011, 11.0000\,\%)$. The two interest rate curves in Fig. 9.2 are constructed by the rational cubic interpolation method; the upper curve uses reconstructed market data and the lower curve uses the original market data. It can be seen that the upper curve, from reconstructed market data, has less volatility.

9.6 Conclusion

The rational cubic interpolation algorithm can ensure that the constructed interest rate curve is positivity-preserving. The interpolation function is C^1 continuous and the constructed interest rate curve is stable and smooth. The algorithm to reconstruct the data can restrict the fluctuations of the interest rate curve when the input data jumps sharply.

References

1. Y. Gu, Z. Zhu, Y. Yan, The term structure and macro-economy of China. China Financ. **3**, 87 (2014)
2. J.H. McCulloch, The tax-adjusted yield curve. J. Financ. **30**, 811–830 (1975)

3. M. Fisher, D. Nuchka, D. Zervos, Fitting the term structure of interest rates with smoothing splines. Comput. Graph. **15**, 14–23 (1995)
4. L. Jun, H. Xuli, Monotonicity preserving rational cubic spline interpolation. Math. Theory Appl. **25**(3), 115–117 (2005)
5. T. Meng, Positivity-preserving piecewise rational cubic interpolation. J. Shandong Univ. Technol. **20**(3), 16–18 (2006)
6. P.S. Hagan, G. West, Interpolation methods for curve construction. Appl. Math. Financ. **13**(2), 89–129 (2006)
7. N. Wang, Piecewise spline interpolation. J. Huazhong Coll. Technol. **3**, 50–59 (1975)
8. H. Hu, Z. Fang, The improvement of the smoothing spline of the term structure of China. J. Manage. Sci. China **12**(1), 101–111 (2009)