A Combinatorial Optimization Model of Interest Rate Term Structure Using Genetic Algorithms

Rongxi Zhou, Xiaoguang Wang, Guanqun Tong

School of Economics and Management, Beijing University of Chemical Technology, China, zhourx@buct.edu.cn

Corresponding Author
Department of Mathematics, Hong Kong University of Science & Technology, Hong Kong, coveragesmile@sina.com

School of Economics and Management, Beijing University of Chemical Technology, China, beverytong@hotmail.com

Abstract

In this paper, we develop a new term structure model of interest rates with combinatorial optimization method based on four classical models: polynomial spline model, exponential spline model, Nelson-Siegel model and Svensson model. Genetic algorithms are employed to solve the combinatorial optimization model. Then, we make some empirical comparisons of five models using daily bond data from Shanghai Stock Exchange in China covering the years from 2004 to 2009. The results show that the combinatorial optimization model outperforms the other models in most of the statistical indicators. Besides, the combinatorial model has good adaptability and robustness which are applicable in Chinese bond market.

Keywords: Term Structure of Interest Rates, Combinatorial Optimization Model, Genetic Algorithms, Classical Model

1. Introduction

The term structure of interest rates is the relationship between the yield to maturity and the time to maturity for pure discount bonds. It is important for many tasks, including pricing financial assets and their derivatives, managing financial risk, allocating portfolios, structuring fiscal debt, conducting monetary policy and valuing capital goods [1,2]. Traditionally, the interest rate term structure model is namely the curve-fitting models which directly fit the yield curve at a point in time with simple functional forms. The seminal work in this direction was McCulloch [3,4] who introduced the methodology of fitting the discount function by polynomial splines as a continuous function of maturity time. Later, there were several extensions of McCulloch’s spline methodology. For example, Vasicek and Fong [5] suggested exponential splines whereas Steely [6] preferred basis splines. Chiu et al.[7] proposed cubic spline model. Laurini and Moura [8] applied the constrained smoothing-splines to interpolate and construct measures associated with the term structure of interest rates. An important shortcoming in estimating the term structure using polynomial splines is that estimations obtained by spline methods are sensitive to the choice of the number and location of the break points (knots) separating the splines. So the knot selection can seriously affect the fit of the term structure of interest rates. Vasicek and Fong [5] claimed that the exponential spline functions are superior to polynomial spline models. Nevertheless, Shea [9] suggested that exponential splines yield forward rates that are as unstable as those obtained from polynomial splines. He recommended to use ordinary spline techniques in preference to exponential splines. Fernando [10] proposed free-knot splines to model interest rate term structure. Nelson and Siegel [11] proposed another kind of parametric yield curve, which was extended by Svensson (SV) [12] and Björk and Christensen [13]. Nelson and Siegel’s (NS) models, due to their parsimonious structure and remarkable efficiency in capturing the general shapes of the yield curves, have been widely used by market practitioners and central banks in different countries. However, the NS models are subject to assumptions of the optimal value of parameters in the model which are difficult to find. In fact, Tilley's results on polynomial curves for the U.S. market cast doubt on whether the Nelson & Siegel form is any more efficient than the polynomial approach [14]. Therefore, lots of references [15-19] conducted empirical studies on the disadvantages and advantages of the above models. For example, by comparing seven models Michalis [15] found that NS, SV and
VRP (variable roughness penalty) models were better than the others. Zhou and Qiu [16] made an empirical comparison of the term structure of interest rates in the bond market in China based on polynomial spline functions. Polychronis and Michalis [17] also found that the fitting error of extended NS models was lower than NS models. However, the existing studies are limited to single models. How to take full advantage of each model? Based on the famous models in references [3,5,11,12], we develop an integrated model using combinatorial optimization method, and estimate the model parameters using genetic algorithms. Then we compare the fitting efficiency of the five models employing daily data from Shanghai Stock Exchange in China from 2004 to 2009.

The remainder of the paper is as follows: in Section 2, we explain the basic definitions of abbreviations and introduce four classical term structure models of interest rates. In Section 3, the combinatorial optimization model is developed. Genetic algorithms are used to estimate the model parameters. In Section 4, empirical comparisons are presented. Section 5 concludes.

2. Four Classical Term Structure Models of Interest Rates

Some basic abbreviations adopted in this paper are defined as follows:

- $n$: number of coupon bonds;
- $T_i$: maturity of the $i$-th coupon bond, $i = 1, \ldots, n$;
- $P_i^t$: market price at time 0 of the $i$-th coupon bond;
- $\vec{P} = (P_i^t)$: vector of market price;
- $P_i^t$: theoretical price at time 0 of the $i$-th coupon bond;
- $\vec{P} = (P_i^t)$: vector of theoretical price;
- $F_i^t$: interest or principal and interest of the $i$-th coupon bond at time $t$;
- $\alpha_i^l$: time $l$ of cash flow of the $i$-th coupon bond, $l = 1, \ldots, M_i$;
- $M_i$: number of cash flows and $\alpha_i^{M_i} = T_i$;
- $D(t)$: discount function, the price at time 0 of 1 RMB paid at time $t$ and $D(0) = 1$;
- $\beta$: vector of parameter in the discount function;
- $\hat{\beta}$: estimation of $\beta$.

The value of a bond $i$ is defined as the sum of the value of the interests and principal, so we have

$$P_i = \sum_{l=0}^{T_i} F_i^l D(t), \quad i = 1, \ldots, n$$

We can obtain the best estimation of $\hat{\beta}$ by solving the least square of discrepancy between the theoretical price and market price, that is

$$\min_{\beta} \sum_{i=1}^{n} (P_i - P_i^t)^2$$

Actually, the main difference among the models is the different definition of discount functions. In the following section, we introduce some important discount functions based on four classical term structure models of interest rates.

2.1. Polynomial spline model

McCulloch [3] assumed the general form of cubic polynomial spline discount functions as follows:
These functions are subject to the assumptions of smoothness and derivativeness. Obviously, the number of spline functions needs to be finite when the model is used to estimate the term structure of interest rates.

### 2.2. Exponential spline model

Vasicek and Fong [5] proposed exponential spline functions. They defined the discount functions as follows:

\[
D(t) = \begin{cases} 
D_1(t) = a_1 + b_1 t + c_1 t^2 + d_1 t^3 & t \in [0, T_1] \\
D_2(t) = a_2 + b_2 t + c_2 t^2 + d_2 t^3 & t \in [T_1, T_2] \\
D_3(t) = a_3 + b_3 t + c_3 t^2 + d_3 t^3 & t \in [T_2, T_3] 
\end{cases} 
\]

The forward rate curve is unstable in the exponential spline model and the parameters are estimated exclusively by nonlinear optimization methods.

### 2.3. Nelson-Siegel model

Nelson and Siegel [11] proposed a parametric model, in which the instantaneous forward rate at time \( t \) has the following exponential expansion form:

\[
f(t) = \beta_0 + \beta_1 e^{-\frac{t}{\tau_1}} + \beta_2 \frac{t}{\tau_1} e^{-\frac{t}{\tau_1}}
\]

Given the instantaneous forward rates, the spot rate can be determined by taking the integration over the forward rates. So the spot rate is

\[
r(t) = \beta_0 + \beta_1 \left( \frac{1 - e^{-\frac{t}{\tau_1}}}{t/\tau_1} \right) + \beta_2 \left( \frac{1 - e^{-\frac{t}{\tau_1}}}{t/\tau_1} - e^{-\frac{t}{\tau_1}} \right)
\]

Therefore

\[
f(0) = r(0) = \beta_0 + \beta_1, \quad f(\infty) = r(\infty) = \beta_0
\]

And we have the discount function \( D(t) = e^{-r(t)} \).

### 2.4. Svensson model

Due to the lack of ability to describe the complicated interest rates curve, an extended form of Nelson-Siegel model was proposed by Svensson [12]. The model introduces two parameters \( \beta_3 \) and \( \tau_2 \), and the instantaneous forward rate is as follows:

\[
f(t) = \beta_0 + \beta_1 e^{-\frac{t}{\tau_1}} + \beta_2 \frac{t}{\tau_1} e^{-\frac{t}{\tau_1}} + \beta_3 \frac{t}{\tau_2} e^{-\frac{t}{\tau_2}}
\]

Given the instantaneous forward rates, the spot rate is
and the discount function is \( D(t) = e^{-r(t)\tau} \).

Svensson model is more flexible in calculating the bond price. Since Nelson-Siegel model and Svensson model use the exponential polynomial function, the spot rate curves have the property of high order derivatives. So the two models satisfy the assumption of smoothness and the model parameters have specific meanings in economics [20,21].

3. Combinational Optimization Model of Interest Rate Term Structure

In this section, we propose a combinatorial optimization model based on the four models above, and use genetics algorithms to estimate its parameters.

3.1. Combinatorial optimization model

Many current studies show that different empirical results can be obtained even using the same sample. The reason is that the structures of these models lead to their drawbacks. However, the information fusion involves the combination of information into a new set of information towards reducing uncertainty [22, 23]. The combinatorial optimization model in information fusion can take full advantage of each model [24, 25]. It can not only overcome certain weaknesses of the other models but also improve their accuracy in fitting and forecasting.

Combinatorial optimization methods are divided into three categories: the equal-weight combination, the optimal weighted combination and uncertainty variable weight combination [25]. The optimal weighted combination is the most popular one among them. With this method, the weighted coefficients are estimated under certain criterion, multi-criteria or information aggregation operators. Following this idea, we propose a combinational optimization model of interest rate term structure under the criterion of least absolute error between theoretical price and the market price of bonds:

\[
\min e = \sum_{i=1}^{d} |P_i - \bar{P}_i| \\
\text{s.t.} \quad \sum_{j=1}^{m} W_j P_{ij} = P_i \\
0 \leq W_j \leq 1, j = 1, \ldots, m
\]

where \( \bar{P}_i \) is the theoretical price of the \( i \)-th bond estimated by the combinatorial model, \( P_{ij} \) is the price of the \( i \)-th bond calculated by the \( j \)-th single model, and \( W_j \) is the weight of the \( j \)-th model. In fact, we develop another combinational optimization model under the criterion of least square error, and make the empirical study for it. The conclusions are similar to ones of Model (1). But we do not give the results because of the length of this paper.

3.2. Genetic algorithms for combinatorial model

Model (1) is subject to a nonlinear programming problem, so we use genetic algorithms (GA) to estimate the parameters [26-28]. Here, we describe the procedure of GA as follows.
Step 1. Determine the number of models. In this paper, we use four models, so \( m = 4 \). Then we generate \((m-1)\)-th dimension vector \( W = (W_1, \ldots, W_{m-1}) \) in the interval \([0,1]\) randomly, where \( W_1 + W_2 + \cdots + W_{m-1} \leq 1 \). The \( m \)-th dimension \( W_m \) is obtained by \( 1 - (W_1 + W_2 + \cdots + W_{m-1}) \), and then an initial population is formed. Each vector represents a chromosome, and \((W_1, \ldots, W_{m-1}, W_m) \in [0,1]^m\) and \( W_i, i = 1, \ldots, m \) are the genes on chromosomes.

Step 2. Solve the single model and calculate the corresponding theoretical prices.

Step 3. Substitute chromosomes into the models, obtain and order the value of objective functions, and then take advantages of the selection operator to choose half of the chromosomes which have better fitting accuracy.

Step 4. Use the portfolio factor to restructure the \((m-1)\)-th dimension vector of any two of the remaining half chromosomes linearly. The \( m \)-th dimension vector is generated by \( 1 - (W_1 + W_2 + \cdots + W_{m-1}) \). After two new chromosomes are formed, use the variation factor to change part of the genes of the chromosomes in order that the algorithms can cover the whole parameter space.

Step 5. For \( \omega \geq 0 \), if \( \sum_{i=1}^{m} \sum_{j=1}^{m} W_i W_j P_{ij} - P_i \leq \omega \), the procedure stops; otherwise, it goes back to step 3.

4. Empirical Studies

In this section, we explore some empirical evidences based on a sample of daily bond data from the Shanghai Stock Exchange in China covering the period from 2004-01-01 to 2009-12-31. We divide daily data into in-sample data which cover 75% of the data that day to estimate the parameters, and out-of-sample data which cover 25% to forecast. First, we solve the four models: Polynomial spline model, Exponential spline model, Nelson-Siegel model and Svensson model, which may be denoted PLOYB, EXPB, NS and SV, respectively, then compare their empirical results with the combinatorial model which is denoted COMBINE.

4.1. Fitting of in-sample data

We estimate the PLOYB model by the least square method, while EXPB, NS and SV are estimated by genetic algorithms. All the parameters of each model are estimated followed by the absolute error between theoretical price and market price. The statistics analysis of the five models are shown in Table 1, which contains the mean, standard deviation, maximum and minimum of absolute errors using in-sample data.

<table>
<thead>
<tr>
<th>Model</th>
<th>Mean</th>
<th>Standard deviation</th>
<th>Maximum</th>
<th>Minimum</th>
</tr>
</thead>
<tbody>
<tr>
<td>NS</td>
<td>0.61</td>
<td>0.79</td>
<td>17.30</td>
<td>0.00</td>
</tr>
<tr>
<td>SV</td>
<td>0.64</td>
<td>0.75</td>
<td>17.20</td>
<td>0.00</td>
</tr>
<tr>
<td>EXPB</td>
<td>0.45</td>
<td>0.17</td>
<td>10.56</td>
<td>0.00</td>
</tr>
<tr>
<td>PLOYB</td>
<td>0.44</td>
<td>0.15</td>
<td>10.59</td>
<td>0.00</td>
</tr>
<tr>
<td>COMBINE</td>
<td>0.41</td>
<td>0.18</td>
<td>11.73</td>
<td>0.00</td>
</tr>
</tbody>
</table>

In terms of the mean of fitting errors, COMBINE performs much better than the other models. As for standard deviation and maximum, NS and SV are relatively higher, while EXPB, PLOYB and COMBINE don’t have significant differences. It is worth noticing that although the minimum value of each model is zero, it does not mean that an error does not exist. The reason is that the accuracy of display is limited. Figure1 shows the mean of fitting errors of the five models and it indicates that COMBINE has outstanding effectiveness. From Table 1 and Figure 2, it can be seen that the standard deviations of the first four models are close to each other and COMBINE has advantages over the others.
4.2. Fitting of out-of-sample data

In order to compare the effectiveness of the five models further, out-of-sample data are used for robust tests. Table 2 shows the mean, standard deviation, maximum and minimum of absolute errors fitting out-of-sample data. Table 2 and Figure 3 show that in a single model, the mean of absolute error of COMBINE produces the smallest one. It can be concluded from Table 2 and Figure 4 that the standard deviation of COMBINE is lower than the other single models. Comparing Table 1 and Table 2, the absolute fitting errors of out-of-sample data are larger than those of in-sample data. But COMBINE has some advantages over the others in both the mean and standard deviation of absolute errors.

<table>
<thead>
<tr>
<th>Model</th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>Maximum</th>
<th>Minimum</th>
</tr>
</thead>
<tbody>
<tr>
<td>NS</td>
<td>0.93</td>
<td>2.55</td>
<td>45.57</td>
<td>0.06</td>
</tr>
<tr>
<td>SV</td>
<td>0.95</td>
<td>1.27</td>
<td>21.32</td>
<td>0.10</td>
</tr>
<tr>
<td>EXPB</td>
<td>0.88</td>
<td>1.28</td>
<td>16.42</td>
<td>0.15</td>
</tr>
<tr>
<td>PLOYB</td>
<td>0.91</td>
<td>1.82</td>
<td>18.47</td>
<td>0.04</td>
</tr>
<tr>
<td>COMBINE</td>
<td>0.87</td>
<td>1.20</td>
<td>31.39</td>
<td>0.08</td>
</tr>
</tbody>
</table>
In order to compare the fitting effectiveness of the five models, the in-sample and out-of-sample data are divided into 20 groups according to their terms to maturity. The results are shown in Table 3. They demonstrate that COMBINE is better than the other models and the fitting effectiveness of in-sample data is better than that of out-of-sample data. Same conclusions can be obtained from Figure 5 and Figure 6 which graph the relationship between time to maturity and mean pricing error.

### Table 3. Mean of absolute errors of fitting five models under different maturities

<table>
<thead>
<tr>
<th>Time to maturity</th>
<th>In-sample data</th>
<th>Out-of-sample data</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>NS</td>
<td>SV</td>
</tr>
<tr>
<td>0-1</td>
<td>0.33</td>
<td>0.33</td>
</tr>
<tr>
<td>1-2</td>
<td>0.45</td>
<td>0.42</td>
</tr>
<tr>
<td>2-3</td>
<td>0.50</td>
<td>0.44</td>
</tr>
<tr>
<td>3-4</td>
<td>0.53</td>
<td>0.49</td>
</tr>
<tr>
<td>4-5</td>
<td>0.56</td>
<td>0.48</td>
</tr>
<tr>
<td>5-6</td>
<td>0.69</td>
<td>0.63</td>
</tr>
<tr>
<td>6-7</td>
<td>0.54</td>
<td>0.55</td>
</tr>
<tr>
<td>7-8</td>
<td>0.40</td>
<td>0.48</td>
</tr>
<tr>
<td>8-9</td>
<td>1.27</td>
<td>1.06</td>
</tr>
<tr>
<td>9-10</td>
<td>1.24</td>
<td>0.84</td>
</tr>
<tr>
<td>10-11</td>
<td>2.85</td>
<td>2.82</td>
</tr>
<tr>
<td>11-12</td>
<td>4.07</td>
<td>3.64</td>
</tr>
<tr>
<td>12-13</td>
<td>2.88</td>
<td>2.84</td>
</tr>
<tr>
<td>13-14</td>
<td>2.57</td>
<td>2.60</td>
</tr>
<tr>
<td>14-15</td>
<td>1.77</td>
<td>1.79</td>
</tr>
<tr>
<td>15-16</td>
<td>1.29</td>
<td>1.31</td>
</tr>
<tr>
<td>16-17</td>
<td>1.47</td>
<td>1.53</td>
</tr>
<tr>
<td>17-18</td>
<td>1.62</td>
<td>1.62</td>
</tr>
<tr>
<td>18-19</td>
<td>1.18</td>
<td>1.29</td>
</tr>
<tr>
<td>19-</td>
<td>1.43</td>
<td>1.80</td>
</tr>
</tbody>
</table>
5. Conclusions

This paper puts forward a new idea in modeling the term structure of interest rates. A new combination optimization model, namely COMBINE, is developed on the base of PLOYB, EXPB, NS and SV. Genetic algorithms are suggested to solve the new model. Then we compare the new model with the four models by using data from the Shanghai Stock Exchange in China. Our empirical results show that COMBINE has relative advantages over the conventional four models in fitting in-sample data as well as forecasting out-of-sample data. Besides, we find that the selection of basic models affects the performance of COMBINE. It is found that our new model obtains the best results when it is based on appropriate basic models. Therefore, in order for our new model to be more practical and flexible, models that reflect the real situations in market should be chosen as basic models.

6. Acknowledgements

The authors thank the anonymous reviewers for providing constructive feedback. The authors thank Qun Gan from Sydney University and Donghui Li from University of New South Wales. This work is supported by National Natural Science Foundation of China Grant No.71171012, and the Training Plan of Science Research of undergraduate of Ministry of Education of the People’s Republic of China Grant No.09100127, and discipline construction project of Beijing University of Chemical Technology Grant No.2010096. All errors are the authors’ own.

7. References