A Novel Stock Market Risk Prediction Method Based on Support Vector Machine Optimized by Multi-Objective Optimization

Leilei Wu  
Ningbo Institute of Education, Ningbo315010, Zhejiang, China  
Yue Li*  
School of Information, GuangDong Communication Polytechnic, GuangZhou510650, GuangDong, China  
*Corresponding author

Abstract  
Stock market risk is attracting more and more attention nowadays, because of the non-predictability characteristics. Accurate stock market risk prediction can provide necessary technical support and guidance in the assessment of stock market risk. In order to improve the accuracy of stock market risk prediction, a combination genetic algorithm and support vector machine (SVM) model has been proposed to forecast stock market risk. Firstly, grey model was adopted to accumulate the original stock market risk data and weaken the randomness of data sequence, and then SVM model is used to predict the stock market risk. Furthermore, using the regressive features of grey model to reduce the prediction results and obtain the final predicted value of stock market risk. The comparison results with other popular predicting algorithms, BP and standard SVM, show that, the GA-SVM model can improve the forecasting accuracy of short term stock market risk and is of a certain practical value.

Keywords: Stock Market, Risk Prediction, Support Vector Machine, Multi-Objective Optimization

1. INTRODUCTION

In recent years, researchers have distributed their important work in the stock market risk predictions. The majority of previous approaches use modern regression techniques. (Mattia, Hufkens and Van, 2015) studied the stock market risk prediction in numerical method, which improved remarkable forecasting accuracy. In 2014, (Harish, Monica and Sharma, 2014) used autoregressive moving average process and persistence models to predict the short-term stock market risk in advance. The researchers proposed long-term stock market risk forecasting method based on probabilistic Bayesian networks (BNs) in Spain (Seyedali and Lewis, 2015). Besides that, most adopted methods are intelligent prediction method based on neural network due to the advantage of self-learning, self-organizing and highly adaptive properties (Barbounis and Theocharis, 2007). Such as multi-layer perceptrons, fuzzy-based neural approaches, two-hidden layer neural networks and fast training neural approaches (Sheela and Deepa, 2013). In spite of this huge work on modern methods for stock market risk prediction from stock market, there are still some margin for improvement, coming from methodologies that have been under-explored in this problem.

As one of the machine learning research topics, support vector machine (SVM) is a novel general method for machine learning based on the statistical theory, which was proposed in the 1990s. It is designed to target learning of small samples. It is so theoretically robust that it avoids such problems as over learning, insufficient learning, high dimensionality, non-linearity and local minuteness, which exist in other methods (Serinivas and Patnaik, 2014). Moreover, it has great learning and generalization abilities and has been successfully applied to many domains (Bhunia and Samanta, 2014). In order to improve the prediction accuracy of stock market risk prediction, a model of multi-objective optimized support vector machine model is proposed to predict the stock market risk.

2. SUPPORT VECTOR MACHINE

Support vector machine is a major breakthrough in machine learning in recent years based on the structural risk minimization criterion (Karmakar and Bhunia, 2014). Consider the data set of n samples \( \{ X_i, y_i, \quad i=1,2,\cdots, n \} \), where \( X_i \in \mathbb{R}^m \) is an m-dimension vector, \( y_i \in \mathbb{R} \) is a real number, \( X_i \) is the input data, \( y_i \) is the output data. In SVM, the sample space is mapped from the original space \( \mathbb{R}^m \) to the high-dimensional feature space \( \mathbb{R}^h \) using the non-linear mapping \( \varphi(x) \), constructing the optimization decision function in the high-dimension feature space:
\[ y = \omega \phi(X) + b \]  
\[(1)\]

where \( \omega \) is the weight vector, \( \omega \in \mathbb{R}^n \) and \( b \) is the offset quantity.

While solving the regression problem, the high-dimension space linear function fitting problem based on the above equation can be formulated as the following optimization problem

\[
\begin{aligned}
& \min \phi(\omega, b, e) = \frac{1}{2} \omega^T \omega + \frac{1}{2} C \sum_{i=1}^{n} e_i^2 \\
& \text{s.t.} \\
& y_i = \omega^T \phi(X_i) + e_i \\
& k = 1, 2, \ldots, n
\end{aligned}
\]

\[(2)\]

where \( e_k \) is the error, \( C \) is the punitive parameter. By using the Lagrange's method of multipliers, the above constrained problem can be regarded as the following unrestricted optimization problem.

\[
L(\omega, b, e, \alpha) = \frac{1}{2} \omega^T \omega + \frac{1}{2} C \sum_{i=1}^{n} e_i^2 - \sum_{i=1}^{n} \alpha_i \{ \omega^T \phi(X_i) + e_i + b - y_i \}
\]

\[(3)\]

where \( \alpha_i \) is the Lagrange multiplier. To solve for the saddle point of the above equation, let the partial derivative of its sum be zero, then we have:

\[
\begin{bmatrix}
0 \\
l \Omega + \frac{1}{\gamma} I
\end{bmatrix} \begin{bmatrix}
b \\
\alpha
\end{bmatrix} = \begin{bmatrix}
0 \\
y
\end{bmatrix}
\]

\[(4)\]

where \( y = y_1, y_2, \ldots, y_n \), \( l = 1, 2, \ldots, n \), \( \alpha = \alpha_1, \alpha_2, \ldots, \alpha_n \), \( \Omega = \phi(X_k)^T \phi(X_k), k = 1, 2, \ldots, n \). Considering Mercer conditions, there exist the mapping function \( \phi \) and the kernel function \( k(\cdot, \cdot) \) which satisfy:

\[ K(X_i, X_j) = \phi(X_i)^T \phi(X_j) \]

\[(5)\]

\[ K(X_i, X_j) = \exp \left( \frac{- (X_i - X_j)^2}{2\sigma^2} \right) \]

\[(6)\]

where \( \sigma > 0 \) is the parameter of the kernel function. If the regression parameter and the offset \( b \) are computed, then the non-linear regression model of SVM is:

\[ f(X) = \sum_{i=1}^{n} x_i K(X, X_i) + b \]

\[(6)\]

From the principles of SVM, it can be known that the optimal values of parameters \( C \) and \( \sigma \) must be determined to construct a model for accurately predicting stock market risk. Hence, the genetic algorithm capable of effectively finding the optimum is chosen in this paper to optimize SVM parameters.

3. MULTI-OBJECTIVE OPTIMIZATION BASED SVM

Generally, an unconstraint minimum multi-objective optimization problem can be described as (Li, Luo and Rong, 2013)

\[
\min F(x) = [f_1(x), f_2(x), \ldots, f_m(x)]^T
\]

\[(7)\]

where, \( F(x) \) represents a multi-objective optimization problem with \( m \) sub-objectives, \( x = [x_1, x_2, \ldots, x_n]^T \in \Omega \) represents \( n \)-dimensional decision vector and \( \Omega \) denotes \( n \)-dimensional decision space; \( f_i(x) \{i = 1, 2, \ldots, m\} \) represents single objective function. In the following, some important definitions for multi-objective optimization problems are given.

Decision vector \( x_1 \) is said to Pareto dominance another vector \( x_2 \) (denote \( x_1 \succ x_2 \)) if and only if
\((\forall i \in [1, 2, ..., m] : f_i(x_i) \leq f_i(x_j)) \land (\exists j \in [1, 2, ..., m] : f_i(x_i) < f_j(x_j))\)  

(8)

The Pareto front (PF) is the mapping of PS in objective space and is defined as

\[ PF = \{ F(x) \mid x \in PS \} \]

(9)

The velocity and position of the \(i\)th particle of \(d\)th dimension are updated as follows

\[ v_{id}^{t+1} = w v_{id}^t + c_1 r_1 (p_{id}^t - x_{id}^t) + c_2 r_2 (p_{gd}^t - x_{id}^t) \]

(10)

\[ x_{id}^{t+1} = x_{id}^t + v_{id}^{t+1} \]

(11)

where \(i = 1, 2, \cdots, n\), \(n\) is the size of population; \(t\) is iteration number; \(x_{id}\) and \(v_{id}\) respectively represent \(d\)-dimensional position and velocity of \(i\)th particle; \(p_{id}\) and \(p_{gd}\) respectively represent history individual optimal position and global optimal position; \(c_1\) and \(c_2\) represent acceleration coefficients; \(r_1\) and \(r_2\) are uniform random numbers in \([0, 1]\). \(w\) represents inertia weight which used to balance global search and local search ability of PSO algorithm.

A class of plane or hyperplane is defined as

\[ wx + b = 0 \]

(12)

This collection is able to converge the sample matrix into \(R_1\) and \(R \neq R_1\), indicating this matrix has neighborhood criterion with clear rules, that is

\[ \begin{cases} \quad w_{x_i} + b \geq 1 \\ \quad w_{x_i} + b \leq -1 \end{cases} \]

(13)

Weight of the link between an element \(x_i\) of the matrix and a certain neighborhood of formula(13) is

\[ \varepsilon_{i} = y_i (wx_i + b) = |wx_i + b| \]

(14)

Map-Min-Max treatment is conducted on \(w\) and \(b\) of formula(8). Substituting \(w\) and \(b\) on \( ||w||\) and \( \frac{b}{||w||} \), the Euclidean space is defined as

\[ \delta_i = \frac{wx_i + b}{||w||} \]

(15)

This matrix is defined to converge into \(R_1\) and \(R \neq R_1\), and the definition of allowed Euclidean space is

\[ \delta = \min \delta_i , \quad i = 1, 2, \cdots, l \]

(16)

The mapping between the error cardinality \(N\) of sample matrix (matrix elements) and allowed Euclidean space definition of converging into \(R_1\) and \(R \neq R_1\) is

\[ N \leq \left( \frac{2R}{\delta} \right)^2 \]

(17)

In the formula, \(R = \max \| x_i \| i = 1, 2, \cdots, l \).

It can be seen from formula(11) that the error cardinality of sample matrix (matrix elements) is determined by \(\delta\). There is negative correlation trend between \(\delta\) and matrix elements. Therefore, the search for the optimal set is the key factor of whether \(\delta\) is the biggest. For example, \(\varepsilon = |wx_i + b| = 1\), so the link weight of two elements is \(2 \left| \frac{wx_i + b}{||w||} \right| = 2 \left| \frac{wx_i + b}{||w||} \right| \). Applying logic to portray above expression
This problem can be transformed into the problem of the Lagrangian function, namely:

$$\Phi(w,b,a_i) = \frac{1}{2} ||w||^2 - \sum_{i=1}^{l} a_i [y_i(wx_i + b) - 1]$$  \hspace{1cm} (19)$$

Due to the limitation of calculated limit, the problem is transformed into another form:

$$\max Q(\alpha) = \sum_{i=1}^{l} a_i - \frac{1}{2} \sum_{i=1}^{l} \sum_{j=1}^{l} a_i a_j y_i y_j (x_i x_j)$$  \hspace{1cm} (20)$$

The function of the optimal vector value ultimately obtained is

$$f(x) = \operatorname{sgn}[\sum_{i=1}^{l} a_i^* y_i (x x_i^*) + b^*]$$  \hspace{1cm} (21)$$

The process for predicting the short-term stock market risk via the multi-objective optimization-based SVM is given in Figure 1.

**Figure 1.** Process for predicting the stock market risk via the multi-objective optimization-based SVM

4. PERFORMANCE METRICS OF THE PREDICTION MODEL

There are many methods for evaluating the performance of the prediction model. The mean absolute error (MAE) and mean standard error (MSE) are chosen as the metrics in this paper. And they can be expressed as:

$$MAE = \frac{1}{k} \sum_{i=1}^{n+k} \left| \frac{s_i - \hat{s}_i}{s_i} \right| \times 100\%$$  \hspace{1cm} (22)$$

$$MSE = \frac{1}{k} \sum_{i=1}^{n+k} \left( \frac{s_i - \hat{s}_i}{s_i} \right)^2$$  \hspace{1cm} (23)$$
The experiment adopted multiple sets of data from different sampling scales. The first dataset is sampling for 3 days to test the short-term forecasting capability, and obtain 48 stock market risk points. The second dataset is sampling for a day, and get 24 stock market risk points. Finally, the third dataset is sampling by one day for a week to observe the long-term scale prediction capability.

Figure 2 and Figure 3 show the MAE and MSE of different models for a week. It can be obviously observed that the proposed GA-SVM method has lowest MAE and MSE value than that of other models.

![MSE value of different prediction methods for a week](image1)

![MAE value of different prediction methods for a week](image2)

5. CONCLUSION

The stock market risk prediction is important for finance system operation research, which can provide important reference for the stock market forecasting. This paper proposes a model for predicting the stock market risk in different time scales by using the grey multi-objective optimization-based SVM. The SVM parameters are optimized through multi-objective optimization to solve the defects of SVM. The strengths of the grey model and SVM are combined to improve the accuracy of stock market risk prediction. Experimental results show that the proposed model outperforms other prediction methods in terms of accuracy, demonstrating the effectiveness and feasibility of the proposed model.

REFERENCES


