

# Multiple Criteria Mathematical Programming and Data Mining

Yong Shi<sup>1,2</sup>, Rong Liu<sup>1,3</sup>, Nian Yan<sup>2</sup>, and Zhenxing Chen<sup>2</sup>

<sup>1</sup> Research Center on Fictitious Economy and Data Sciences, Chinese Academy of Sciences  
100080 Beijing, China

<sup>2</sup> College of Information Science and Technology, University of Nebraska at Omaha  
Omaha NE 68182, USA

<sup>3</sup> School of Mathematical Science, Graduate University of Chinese Academy of Sciences  
100049 Beijing, China

liu.rong@163.com, yshi@gucas.ac.cn,  
{nyan, zchen}@mail.unomaha.edu

**Abstract.** Recently, researchers have extensively applied quadratic programming into classification, known as V. Vapnik's Support Vector Machine, as well as various applications. However, using optimization techniques to deal with data separation and data analysis goes back to more than forty years ago. Since 1998, the authors and their colleagues extended such a research idea into classification via multiple criteria linear programming (MCLP) and multiple criteria quadratic programming (MQLP). The purpose of the paper is to share our research results and promote the research interests in the community of computational sciences. These methods are different from statistics, decision tree induction, and neural networks. In this paper, starting from the basics of Multiple Criteria Linear Programming (MCLP), we further discuss penalized MCLP Multiple Criteria Quadratic Programming (MCQP), Multiple Criteria Fuzzy Linear Programming, Multi-Group Multiple Criteria Mathematical Programming, as well as regression method by Multiple Criteria Linear Programming. A brief summary of applications of Multiple Criteria Mathematical Programming is also provided.

**Keywords:** Multi-criteria programming, MCLP, MCQP, data mining, classification, regression, fuzzy programming.

## 1 Introduction

Recently, researchers have extensively applied quadratic programming into classification, known as V. Vapnik's Support Vector Machine [1], as well as various applications. However, using optimization techniques to deal with data separation and data analysis goes back to more than forty years ago. In 1960's, O.L. Mangasarian's group formulated linear programming as a large margin classifier [2]. Later in 1970's, A. Charnes and W.W. Cooper initiated Data Envelopment Analysis where a fractional programming is used to evaluate decision making units, which is economic representative data in a given training dataset [3]. From 1980's to 1990's, F. Glover proposed

a number of linear programming models to solve discriminant problems with a small sample size of data [4]. Then, since 1998, the authors and their colleagues extended such a research idea into classification via multiple criteria linear programming (MCLP) and multiple criteria quadratic programming (MQLP). These methods are different from statistics, decision tree induction, and neural networks. The purpose of the paper is to share our research results and promote the research interests in the community of computational sciences.

The data mining task which will be investigated in this paper is the classification or the so-called discriminate analysis in statistical inference. The purpose of classification is to separate data according to some criteria. There are two commonly used criteria among them. The first one is the overlapping degree with respect to the discriminate boundary. The lower of this degree the better the classification is. Another one is the distance from a point to the discriminate boundary. The larger the sum of these distances the better the classification is. Accordingly, the objective of a classification is to minimize the sum of the overlapping degree and maximize the sum of the distances [4]. Note that these two criteria can not be optimized simultaneously because they are contradictory to each other. Fortunately, the multi-criteria mathematical programming can be used to overcome this kind of problems in a systematical way.

It has been thirty years since the first appearance of the multi-criteria linear programming. During these years, the multi-criteria programming has been not only improved in theoretical foundations but also applied successfully in real world problems. The data mining is such an area where the multi-criteria program has achieved a great deal. Initialed by Shi et al. [5], the model and ideal of multi-criteria programming have been widely adopted by the researches for classification, regression, etc. To handle the unbalanced training set problem, Li et al. [6] proposed the penalized multi-criteria linear programming method. He et al. [7] introduced the fuzzy approach in the multi-criteria programming to address the uncertainty in criteria of data separation. Using a different norm to measure the overlapping degree and distance, Kou [8] presented the Multiple Criteria Quadratic Programming for data mining. Kou et al. [9] proposed Multi-Group Multiple Criteria Mathematical Programming aimed to handle the multi-group classification. To extend the application of multi-criteria programming, Zhang et al. [10] developed a regressing method based on this technique. Some important characteristics of these variations of the multi-criteria data mining technique are summarized in Table 1.

In respect of the abundance of the variations of multiple criteria mathematical programming and the diversity of applications, a comprehensive review of related methods would benefit the research in data mining. In this paper, several multi-criteria linear programming methods in data mining are reviewed and analyzed. The remaining part of the paper is organized as follows. First, we present the basics of Multiple Criteria Linear Programming (MCLP) (Section 2). Since the training set could be unbalanced, penalized MCLP method has been proposed to deal with this problem (Section 3). Furthermore, in order to achieve better classification performance and stability, Multiple Criteria Quadratic Programming (MCQP) has been developed (Section 4). Instead of identifying a compromise solution for the separation of data in MCLP, an alternative Multiple Criteria Fuzzy Linear Programming approach has also been studied (Section 5). In addition, two-group Multiple Criteria Mathematical

**Table 1.** Some Important Characteristics of MCLP (M.1) Variations

		Linear Objective	Non-Linear Objective	Hard Constraints	Soft Constraints	Unbalance Constraints	Multi-group classification	Classification	Regression
MCLP	M.2	√		√				√	
PMCLP	M.3	√		√		√		√	
MCQP	M.5		√	√				√	
	M.6		√		√			√	
FMCLP	M.7	√			√			√	
Multi-group MCLP	M.9	√	√	√			√	√	
	M.10	√	√		√		√	√	
MCLP Reg.	M.11	√		√					√

Programming has been extended to Multi-Group Multiple Criteria Mathematical Programming (Section 6). We also review how to apply MCLP to regression problem (Section 7). A brief summary of applications of multiple criteria mathematical programming is provided in Section 8. We conclude the paper in Section 9.

## 2 Multiple Criteria Linear Programming (MCLP)

In linear discriminate analysis, the data separation can be achieved by two opposite objectives. The first one is to maximize the minimum distances of observations from the critical value. The second objective separates the observations by minimizing the sum of the deviations (the overlapping) among the observations [4]. However, it is theoretically impossible to optimize MMD and MSD simultaneously, the best tradeoff of two measurements is difficult to find. This shortcoming has been coped with by the technique of multiple criteria linear programming (MCLP) [5, 11, 12]. The first MCLP model can be described as follows:

$$\begin{aligned}
 & \text{Min} \quad \sum_{i=1}^n \alpha_i \\
 & \text{Max} \quad \sum_{i=1}^n \beta_i \quad \text{(Model 1)} \\
 & \text{S.T.} \quad (\mathbf{x}_i, \mathbf{w}) = b + y_i(\alpha_i - \beta_i), \quad i = 1, \dots, n \\
 & \quad \quad \alpha, \beta \geq 0
 \end{aligned}$$

Here,  $\alpha_i$  is the overlapping and  $\beta_i$  the distance from the training sample  $\mathbf{x}_i$  to the discriminator  $(\mathbf{x}_i, \mathbf{w}) = b$  (classification boundary).  $y_i \in \{1, -1\}$  denotes the label of  $\mathbf{x}_i$  and  $n$  is the number of samples. The weights vector  $\mathbf{w}$  and the bias  $b$  are the

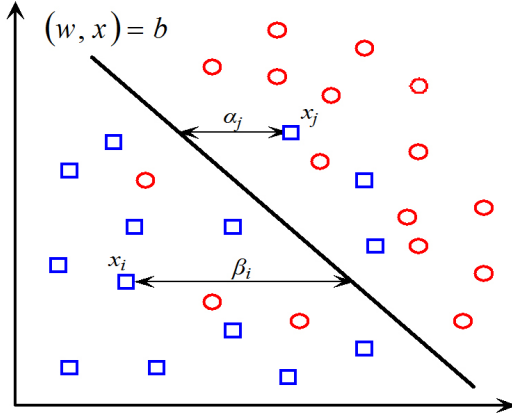


Fig. 1. The two criteria of classification

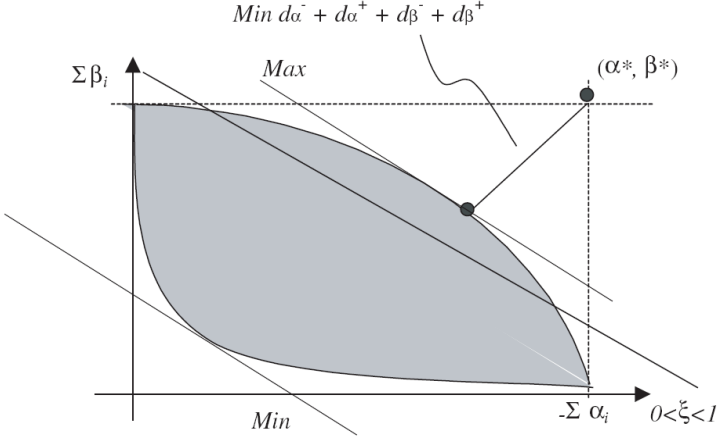
unknown variables to be optimized for the two objectives. A visual description of this model is shown in Fig. 1.

**Model 1** is formulized as Multiple Criteria Linear Programming which is difficult to optimize. In order to facilitate the computation, the compromise solution approach [5, 13] can be employed to reform the above model so that we can systematically identify the best trade-off between  $-\Sigma\alpha_i$  and  $\Sigma\beta_i$  for an optimal solution. The “ideal value” of  $-\Sigma\alpha_i$  and  $\Sigma\beta_i$  are assumed to be  $\alpha^* > 0$  and  $\beta^* > 0$  respectively. Then, if  $-\Sigma\alpha_i > \alpha^*$ , we define the regret measure as  $-d_\alpha^- = \Sigma\alpha_i + \alpha^*$ ; otherwise, it is 0. If  $-\Sigma\alpha_i < \alpha^*$ , the regret measure is defined as  $d_\alpha^+ = \alpha^* + \Sigma\alpha_i$ ; otherwise, it is 0. Thus, we have (i)  $\alpha^* + \Sigma\alpha_i = d_\alpha^- - d_\alpha^+$ , (ii)  $|\alpha^* + \Sigma\alpha_i| = d_\alpha^- + d_\alpha^+$ , and (iii)  $d_\alpha^-, d_\alpha^+ \geq 0$ . Similarly, we derive  $\beta^* - \Sigma\beta_i = d_\beta^- - d_\beta^+$ ,  $|\beta^* - \Sigma\beta_i| = d_\beta^- + d_\beta^+$ , and  $d_\beta^-, d_\beta^+ \geq 0$ . The two-class MCLP model has been gradually evolved as **Model 2**:

$$\begin{aligned}
 \text{Min} \quad & d_\alpha^+ + d_\alpha^- + d_\beta^+ + d_\beta^- \\
 \text{S.T.} \quad & \alpha^* + \sum_{i=1}^n \alpha_i = d_\alpha^- - d_\alpha^+ \\
 & \beta^* - \sum_{i=1}^n \beta_i = d_\beta^- - d_\beta^+ \tag{Model 2} \\
 & (\mathbf{x}_i, \mathbf{w}) = b + y_i(\alpha_i - \beta_i), \quad i = 1, \dots, n \\
 & \alpha, \beta \geq 0, d_\alpha^+, d_\alpha^-, d_\beta^+, d_\beta^- \geq 0
 \end{aligned}$$

Here  $\alpha^*$  and  $\beta^*$  are given,  $w$  and  $b$  are unrestricted. The geometric meaning of the model is shown as in Fig. 2.

In order to calculate a large data set, the Linux-based MCLP classification algorithm was developed to implement the above Model 2 (Kou and Shi, 2002).



**Fig. 2.** Model 2 and Model 7 formulations

### 3 Penalized MCLP

Usually, the sample sizes of different groups vary; namely, the training set is unbalanced. To handle this problem with the MCLP model, Li et al. [6] proposed the following penalized MCLP method (**Model 3**) for credit scoring.

$$\begin{aligned}
 \text{Min} \quad & d_{\alpha}^{+} + d_{\alpha}^{-} + d_{\beta}^{+} + d_{\beta}^{-} \\
 \text{S.T.} \quad & \alpha^{*} + p \times \frac{n_2}{n_1} \sum_{i \in B} \alpha_i + \sum_{i \in G} \alpha_i = d_{\alpha}^{-} - d_{\alpha}^{+} \\
 & \beta^{*} - p \times \frac{n_2}{n_1} \sum_{i \in B} \beta_i - \sum_{i \in G} \beta_i = d_{\beta}^{-} - d_{\beta}^{+} \\
 & (\mathbf{x}_i, \mathbf{w}) = b + \alpha_i - \beta_i, \quad \mathbf{x}_i \in B \\
 & (\mathbf{x}_i, \mathbf{w}) = b - \alpha_i + \beta_i, \quad \mathbf{x}_i \in G \\
 & \boldsymbol{\alpha}, \boldsymbol{\beta} \geq 0, \quad d_{\alpha}^{+}, d_{\alpha}^{-}, d_{\beta}^{+}, d_{\beta}^{-} \geq 0
 \end{aligned} \tag{Model 3}$$

Here, “Bad” and “Good” denote different groups,  $n_1$  and  $n_2$  are the number of samples corresponding to the two groups, and  $p \geq 1$  is the penalized parameter.

In this model the distance is balanced on the two sides of  $b$  with the parameter  $n_1/n_2$ , even there are less “Bad” records on the left of the credit score boundary  $b$ . The value of  $p$  enhances the effect of “Bad” distance and penalizes much more if we wish more “Bad” records on the left of the boundary.

If  $n_1 = n_2$ ,  $p = 1$ , the model above degenerates to the original MCLP model (**Model 1**). If  $n_1 < n_2$ , then exist  $p \geq 1$  to make “Bad” catching rate of PMCLP higher than that of MCLP with the same  $n_1, n_2$ .

#### 4 Multiple Criteria Quadratic Programming (MCQP)

Based on MCLP, the Multiple Criteria Quadratic Programming is later developed to achieve better classification performance and stability. The overlapping and distance are respectively represented by the nonlinear functions  $f(\alpha)$  and  $g(\beta)$ . Given weights  $\omega_\alpha$  and  $\omega_\beta$ , let  $f(\alpha) = \|\alpha\|^p$  and  $g(\beta) = \|\beta\|^p$ , the two criteria basic **Model 1** can be converted into a single criterion general non-linear classification model (**Model 4**):

$$\begin{aligned} \text{Min} \quad & \omega_\alpha \|\alpha\|^p - \omega_\beta \|\beta\|^p \\ \text{S.T.} \quad & (\mathbf{x}_i, \mathbf{w}) = b + y_i(\alpha_i - \beta_i), \quad i = 1, \dots, n \\ & \alpha, \beta \geq 0 \end{aligned} \quad (\text{Model 4})$$

On the basis of Model 4, non-linear classification models with any norm can be defined theoretically. Let

$$f(\alpha) = \alpha^T H \alpha = \sum_{i=1}^m \alpha_i^2 \quad \text{and} \quad f(\beta) = \beta^T Q \beta = \sum_{i=1}^m \beta_i^2$$

where  $H$  and  $Q$  are predefined as identity matrices. We add the term  $\frac{1}{2} \|\mathbf{w}\|_2^2$  into the objective function and formulate a simple quadratic programming with 2-norm as in **Model 5**:

$$\begin{aligned} \text{Min} \quad & \frac{1}{2} \|\mathbf{w}\|_2^2 + \omega_\alpha \sum_{i=1}^n \alpha_i^2 - \omega_\beta \sum_{i=1}^n \beta_i^2 \\ \text{S.T.} \quad & (\mathbf{x}_i, \mathbf{w}) = b + y_i(\alpha_i - \beta_i), \quad i = 1, \dots, n \\ & \alpha, \beta \geq 0 \end{aligned} \quad (\text{Model 5})$$

In order to reduce the number of variables involved in our model and thus simplify computation. Let  $\eta_i = \alpha_i - \beta_i$ . According to our definition,  $\eta_i = \alpha_i$  for all misclassified records and  $\eta_i = -\beta_i$  for all correctly separated records. To obtain strong convexity to the objection function, we add  $\frac{\omega_b}{2} b^2$  to **Model 5**'s objective function.

The weight  $W_b$  is an arbitrary positive number and  $\omega_b \ll \omega_\beta$ . **Model 6** becomes [8]:

$$\begin{aligned}
\text{Min} \quad & \frac{1}{2} \|\mathbf{w}\|_2^2 + \frac{\omega_\alpha}{2} \sum_{i=1}^n \eta_i^2 + \omega_\beta \sum_{i=1}^n \eta_i + \frac{W_b}{2} b^2 \\
\text{S.T.} \quad & (\mathbf{x}_i, \mathbf{w}) = b + y_i \eta_i, \quad i = 1, \dots, n \\
& \boldsymbol{\eta} \geq 0
\end{aligned} \tag{Model 6}$$

## 5 Multiple Criteria Fuzzy Linear Programming

Instead of identifying a compromise solution for the separation of data in MCLP, the fuzzy approach classifies the data by seeking a fuzzy (satisfying) solution obtained from a fuzzy linear program (FLP) [7]. Let  $y_{1L}$  be MSD and  $y_{2U}$  be MMD, then one can assume that the value of Maximize  $\sum \alpha_i$  to be  $y_{1U}$  and that of Minimize  $\sum \alpha_i$  to be  $y_{2L}$ . The classification problem is equivalent to the following fuzzy linear program (Model 7):

$$\begin{aligned}
\text{Min} \quad & \xi \\
\text{S.T.} \quad & \xi \leq \frac{\sum \alpha_i - y_{1L}}{y_{1U} - y_{1L}} \\
& \xi \leq \frac{\sum \beta_i - y_{2L}}{y_{2U} - y_{2L}} \\
& (\mathbf{x}_i, \mathbf{w}) = b + y_i (\alpha_i - \beta_i), \quad i = 1, \dots, n \\
& \boldsymbol{\alpha}, \boldsymbol{\beta} \geq 0
\end{aligned} \tag{Model 7}$$

Note that Model 7 will produce a value of  $\xi$  with  $0 \leq \xi < 1$ . To avoid the trivial solution, one can set up  $0 \leq \varepsilon < \xi$ , for a given  $\varepsilon$ . Therefore, seeking Maximum  $\xi$  in the FLP approach becomes the standard of determining the classifications between Good and Bad records in the database. A graphical illustration of this approach can be seen from Fig. 2, any point of hyper plane  $0 < \xi < 1$  over the shadow area represents the possible determination of classifications by the FLP method.

## 6 Multi-group Multiple Criteria Mathematical Programming

The above models are concerned with two groups' case. Now suppose we have  $k$  groups,  $G_1, G_2, \dots, G_k$ , are predefined.  $G_i \cap G_j = \Phi, i \neq j, 1 \leq i, j \leq k$  and  $x_i \in \{G_1 \cup G_2 \cup \dots \cup G_k\}$ . A series of boundary scalars  $b_1 < b_2 < \dots < b_{k-1}$ , can be set to separate these  $k$  groups. The boundary  $b_j$  is used to separate  $G_j$  and  $G_{j+1}$ .

Let  $\mathbf{w} = (w_1, \dots, w_m)^T \in \mathbf{R}^m$  be a vector of real number to be determined. Thus, we can establish the following linear inequations [14; 8]:

$$(\mathbf{x}_i, \mathbf{w}) < b_l, \quad \forall \mathbf{x}_i \in G_1; \quad (1)$$

$$b_{j-1} \leq (\mathbf{x}_i, \mathbf{w}) < b_j, \quad \forall \mathbf{x}_i \in G_j; \quad (2)$$

$$(\mathbf{x}_i, \mathbf{w}) \geq b_{k-1}, \quad \forall \mathbf{x}_i \in G_k; \quad (3)$$

$$2 \leq j \leq k-1, 1 \leq i \leq n.$$

A mathematical function  $f$  can be used to describe the summation of total overlapping while another mathematical function  $g$  represents the aggregation of all distances. The final classification accuracies of this multi-group classification problem depend on simultaneously minimize  $f$  and maximize  $g$ . Thus, a generalized bi-criteria programming method for classification can be formulated as **Model 8**:

$$\begin{aligned} & \text{Min } f \\ & \text{Max } g \quad \quad \quad \text{(Model 8)} \\ & \text{S.T. } (1), (2), \text{ and } (3) \end{aligned}$$

Furthermore, to transform the bi-criteria problems of the generalized model into a single-criterion problem, weights  $\omega_\alpha > 0$  and  $\omega_\zeta > 0$  are introduced for  $f(\alpha)$  and  $g(\zeta)$ , respectively. The values of  $\omega_\alpha$  and  $\omega_\zeta$  can be pre-defined in the process of identifying the optimal solution. As a result, the generalized model can be converted into a single-criterion mathematical programming model as **Model 9**:

$$\text{Min } \omega_\alpha \sum_{j=1}^k \sum_{i=1}^n |\alpha_{i,j}|_p - \omega_\zeta \left( \sum_{j=1 \text{ or } j=k}^n \sum_{i=1}^n |\zeta_{i,j}|_p - \sum_{j=2}^{k-1} \sum_{i=1}^n \left| \frac{b_j - b_{j-1}}{2} - \zeta_{i,j} \right|_p \right)$$

$$\text{S.T. } (\mathbf{x}_i, \mathbf{w}) = b_j + \alpha_{i,j} - \zeta_{i,j}, \quad 1 \leq j \leq k-1 \quad (4)$$

$$(\mathbf{x}_i, \mathbf{w}) = b_{j-1} + \alpha_{i,j-1} - \zeta_{i,j-1}, \quad 2 \leq j \leq k \quad (5)$$

$$\zeta_{i,j-1} \leq b_j - b_{j-1} \quad 2 \leq j \leq k \quad (6)$$

$$\zeta_{i,j} \leq b_{j+1} - b_j \quad 1 \leq j \leq k-1 \quad (7)$$

Here  $\mathbf{x}_i$  is given,  $\mathbf{w}$  and  $b_j$  are unrestricted, and  $\alpha_{i,j}, \zeta_{i,j} \geq 0, 1 \leq i \leq n$ .

(6) and (7) are defined as such due to the fact that the distances from any correctly classified data ( $\mathbf{x}_i \in G_j, 2 \leq j \leq k-1$ ) to two adjunct boundaries  $b_{j-1}$  and  $b_j$  must be less than  $b_j - b_{j-1}$ . A better separation of two adjunct groups may be achieved by the following constraints which have stronger limitation on  $\zeta_i^j$ :



$$\zeta_{i,j} \leq (b_j - b_{j-1})/2 + \varepsilon, 2 \leq j \leq k \quad (8)$$

$$\zeta_{i,j} \leq (b_{j+1} - b_j)/2 + \varepsilon, 1 \leq j \leq k-1 \quad (9)$$

$\varepsilon \in \mathfrak{R}^+$  is a small positive real number.

Let  $p = 2$ , then objective function in Model 1 can now be a quadratic objective and we have **Model 10** as shown below:

$$\text{Min} \quad \omega_\alpha \sum_{j=1}^k \sum_{i=1}^n |\alpha_{i,j}|_p - \omega_\zeta \left( \sum_{j=1 \text{ or } j=k}^n \sum_{i=1}^n |\zeta_{i,j}|_p - \sum_{j=2}^{k-1} \sum_{i=1}^n \left| \frac{b_j - b_{j-1}}{2} - \zeta_{i,j} \right|_p \right)$$

$$\text{S.T.} \quad (4), (5), (8), \text{ and } (9)$$

Note that the constant  $(\frac{b_j - b_{j-1}}{2})^2$  is omitted from the **Model 6** without any effect to the solution.

## 7 Regression Method by Multiple Criteria Linear Programming

MCLP can also be applied to regression problem. The data set of the regression problem is  $T = \{(x_1^T, y_1), (x_2^T, y_2), \dots, (x_n^T, y_n)\}$ , where  $x_i \in R^m$  is the input variable,  $y_i \in R$  is the output variable, which can be any real number. Define the  $G$  and  $B$  as “Good” and “Bad”, respectively, then the  $D_{MCLP}^+$  and  $D_{MCLP}^-$  data sets for MCLP regression model are constructed. With these data sets, MCLP regression model can be written as **Model 11** [10]:

$$\text{Min} \quad \sum_{i=1}^n (\alpha_i - \alpha_i') - \text{Max} \quad \sum_{i=1}^n (\beta_i - \beta_i')$$

$$\text{S.T.} \quad x_{i1}w_1 + \dots + x_{im}w_m + (y_i + \varepsilon)w_{m+1} = b - \alpha_i + \beta_i$$

...

for all  $\in G$

$$x_{n1}w_1 + \dots + x_{nm}w_m + (y_n + \varepsilon)w_{m+1} = b - \alpha_n + \beta_n$$

$$x_{i1}w_1 + \dots + x_{im}w_m + (y_i - \varepsilon)w_{m+1} = b + \alpha_i' - \beta_i'$$

...

for all  $\in B$

$$x_{n1}w_1 + \dots + x_{nm}w_m + (y_n - \varepsilon)w_{m+1} = b + \alpha_n' - \beta_n'$$

$$\alpha, \alpha', \beta, \beta' \geq 0$$

Aggregation of Good samples:

$$D_{MCLP}^+ = \{(x_i^T, y_i + \varepsilon)^T, +1\}, i = 1, \dots, l\}$$

Aggregation of Bad samples:

$$D_{MCLP}^- = \{((x_i^T, y_i - \varepsilon)^T, -1), i = 1, \dots, l\}$$

## 8 Applications

The multi-criteria data mining techniques reviewed above have yielded fruitful results in diverse applications. Kou [8] applied the Multiple Criteria Quadratic Programming to credit card risk analysis and obtained comparable results with some sophisticated methods. Classification of HIV-1 mediated neuronal dendritic and synaptic damage is another successful example of the multi-criteria data mining techniques [15]. Kou et al. [16] introduced this technique to network surveillance and intrusion detection system. This approach has also been applied to predict firm bankruptcies [17, 18]. Zhang et al. [19] employed the both Multiple-Criteria Linear and Quadratic Programming in VIP e-Mail behavior analysis. In addition to these applications, some of the models mentioned above have played important roles in building the national credit scoring system in China as well as an insurance fraud detection system in USA, in which tera-bytes of data have been handled for business intelligence.

## 9 Conclusions

This paper has reviewed various multi-criteria programming data mining models. These methods are different from statistics, decision tree induction, and neural networks. We have discussed 11 models related to basic Multiple Criteria Linear Programming (MCLP), MCLP Multiple Criteria Quadratic Programming (MCQP), Multiple Criteria Fuzzy Linear Programming, Multi-Group Multiple Criteria Mathematical Programming, as well as regression method by Multiple Criteria Linear Programming. These models have been successfully applied in many real-life applications, such as credit assessment management, information intrusion, bio-informatics, etc. The purpose of the paper is to share the research results and promote research interests in the international community of computational sciences.

**Acknowledgments.** This work was partially supported by National Natural Science Foundation of China (Grant No. 70621001, 70531040, 70501030, 10601064), National Natural Science Foundation of Beijing (Grant No. 9073020), 973 Project of Chinese Ministry of Science and Technology (Grant No. 2004CB720103), and BHP Billiton Cooperation of Australia.

## References

1. Cortes, C., Vapnik, V.: Support-vector Network. *Machine Learning* 20, 273–279 (1995)
2. Mangasarian, O.L.: Linear and Nonlinear Separation of Patterns by Linear Programming. *Operations Research* 13, 444–452 (1965)
3. Charnes, A., Cooper, W.W., Rhodes, E.: Measuring the Efficiency of Decision-making Units. *European Journal of Operations Research* 3(4), 339 (1979)

4. Freed, N., Glover, F.: Simple but Powerful Goal Programming Models for Discriminant Problems. *European Journal of Operational Research* 7, 44–60 (1981)
5. Shi, Y., Wise, M., Luo, M., Lin, Y.: Data Mining in Credit Card Portfolio Management: A Multiple Criteria Decision Making Approach. In: *Proceedings of International Conference on Multiple Criteria Decision Making*, Ankara, Turkey (2000)
6. Li, A.H., Shi, Y., He, J.: MCLP-based Methods for Improving “Bad” Catching Rate in Credit Cardholder Behavior Analysis. *Applied Soft Computing* 8(3), 1259–1265 (2008)
7. He, J., Liu, X., Shi, Y., Xu, W., Yan, N.: Classifications of Credit Cardholder Behavior by using Fuzzy Linear Programming. *International Journal Of Information Technology And Decision Making* 3(4), 633–650 (2004)
8. Kou, G.: *Multi-Class Multi-Criteria Mathematical Programming and its Applications in Large Scale Data Mining Problems*, PhD Dissertation, University of Nebraska Omaha (2006)
9. Kou, G., Liu, X., Peng, Y., Shi, Y., Wise, M., Xu, W.: *Multiple Criteria Linear Programming to Data Mining: Models, Algorithm Designs and Software Developments. Optimization Methods and Software* 18, 453–473 (2003)
10. Zhang, D.L., Tian, Y.J., Shi, Y.: A Regression Method by Multiple Criteria Linear Programming. In: *19th International Conference on Multiple Criteria Decision Making* (2008)
11. Shi, Y.: *Multiple Criteria and Multiple Constraint Level Linear Programming: Concepts*. World Scientific Publishing Co., Singapore (2001)
12. Shi, Y., Peng, Y., Xu, W., Tang, X.: Data Mining Via Multiple Criteria Linear Programming: Applications in Credit Card Portfolio Management. *International Journal of Information Technology and Decision Making* 1, 131–151 (2002)
13. Shi, Y., Yu, P.L.: Goal Setting and Compromise Solutions. In: Karpak, B., Zionts, S. (eds.) *Multiple Criteria Decision Making and Risk Analysis Using Microcomputers*, pp. 165–204. Springer, Berlin (1989)
14. Kou, G., Peng, Y., Shi, Y., Wise, M., Xu, W.: Discovering Credit Cardholders Behavior by Multiple Criteria Linear Programming. *Annals of Operations Research* 135(1), 261–274 (2005)
15. Zheng, J., Zhuang, W., Yan, N., Kou, G., Peng, H., McNally, C., Erichsen, D., Cheloha, A., Herek, S., Shi, C., Shi, Y.: Classification of HIV-1 Mediated Neuronal Dendritic and Synaptic Damage Using Multiple Criteria Linear Programming. *Neuroinformatics* 2, 303–326 (2003)
16. Kou, G., Peng, Y., Yan, N., Shi, Y., Chen, Z., Zhu, Q., Huff, J., McCartney, S.: Network Intrusion Detection by Using Multiple-Criteria Linear Programming. In: Chen, J. (ed.) *Proceedings of 2004 International Conference on Service Systems and Service Management*, Beijing, China, July 19-21, pp. 806–809 (2004)
17. Kwak, W., Shi, Y., Cheh, J.J.: Firm Bankruptcy Prediction Using Multiple Criteria Linear Programming Data Mining Approach. *Advances in Financial Planning and Fore-casting* 2, 27–49 (2006)
18. Kwak, W., Shi, Y., Eldridge, S., Kou, G.: Bankruptcy prediction for Japanese firms: using Multiple Criteria Linear Programming data mining approach. *International Journal of Business Intelligence and Data Mining* 1(4), 401–416 (2006)
19. Zhang, P., Zhang, J.L., Shi, Y.: A New Multi-Criteria Quadratic-Programming Linear Classification Model for VIP E-Mail Analysis. In: Shi, Y., van Albada, G.D., Dongarra, J., Sloot, P.M.A. (eds.) *ICCS 2007. LNCS*, vol. 4488, pp. 499–502. Springer, Heidelberg (2007)