复旦管理学杰出贡献奖获奖者代表成果集

2009

中国管理研究与实践

ZHONGGUO GUANLI YANJIU YU SHIJIAN

石勇 唐立新 汪寿阳 著

复旦大学出版社

复旦管理学杰出贡献奖获奖者代表成果集 2009

中国管理研究实践

ZHONGGUO GUANLI YANJIU YU SHIJIAN

石勇 唐立新 汪寿阳 著

复旦大学出版社

图书在版编目(CIP)数据

中国管理研究与实践——复旦管理学杰出贡献奖获奖者代表成果集(2009)/
石勇,唐立新,汪寿阳著. —上海:复旦大学出版社,2011.12
ISBN 978-7-309-08532-7

Ⅰ. 中… Ⅱ. ①石…②唐…③汪… Ⅲ. 管理学-中国-文集 Ⅳ. C93-53

中国版本图书馆 CIP 数据核字(2011)第 211768 号

中国管理研究与实践——复旦管理学杰出贡献奖获奖者代表成果集(2009)
石 勇 唐立新 汪寿阳 著
责任编辑/徐惠平 张咏梅

复旦大学出版社有限公司出版发行
上海市国权路 579 号 邮编:200433
网址:fupnet@fudanpress.com http://www.fudanpress.com
门市零售:86-21-65642857 团体订购:86-21-65118853
外埠邮购:86-21-65109143
上海华教印务有限公司

开本 787×1092 1/16 印张 29.25 字数 659 千
2011 年 12 月第 1 版第 1 次印刷

ISBN 978-7-309-08532-7/C·216
定价:55.00 元

序 言 一

李岚清

最近20多年来，管理学在我国日益受到人们的重视，这和我国的改革开放、经济社会快速发展有关，也和我国步入社会主义市场经济有关。其实，新中国建立以来，在经济和社会领域内都存在大量的涉及管理学的问题。我长期在大型企业、对外经济贸易部门和从事经济方面的领导工作中也都深切感受到这一点。但是由于种种原因，管理学在相当长的时期内未能得到应有的重视。

管理学真正成为一门独立的科学，走进中国人的专业视野，全面进入中国的科学研究和高等教育体系，也就是最近20多年的事情。改革开放以来，中国的经济发展突飞猛进，科学技术日新月异，经济发展和社会进步越来越离不开管理科学的支撑。社会管理、环境管理、公共管理、企业管理等等各个方面都对管理学提出了新的要求。经济社会领域改革的不断深入，在参与国际竞争中要取得持续的优势，这些都迫切需要进一步加强管理科学的研究，提高管理水平。可以说，需要管理学解决的问题越来越多，管理渗透到社会、经济生活的各个方面。当前中国管理科学正迸发出空前的生机和活力，同时也面临着空前的机遇和挑战。

管理学是一门应用性、实践性很强的学科，作为一门科学，它的一些理论和方法在世界范围内具有共性。但是管理要获得成功则必须植根于一个国家的社会组织和民族文化之中。要真正解决好中国的管理问题，要让中国人对世界范围内涉及自己的管理问题有话语权和平等的参与权，最终还是要依靠中国人自己。管理科学是一个国家软实力的重要组成部分，我们要不断地构建有中国特色的管理科学理论，要具备并不断提高解决各类实际管理问题的能力，要培养出大批有很高学养和丰富经验的管理者，要花大力气建设高质量的管理教育体系，最关键的是要有一支高水平的管理学队伍。

复旦管理学奖励基金会的宗旨在于奖励在中国管理学领域作出贡献的学者和实践工作者，推动管理学的理论和实践的结合，形成中国特色的管理科学体系，最终推动中国管理学的长远发展，促进中国管理学人才的成长，提高中国管理学的国际学术影响力。

复旦管理学杰出贡献奖到今天已经是第5个年头了，12位在管理科学、工商管理和公共管理等领域有杰出贡献的学者获得了这一奖项。这次，基金会把历届获奖人的代表性成果收录成册、公开发行，一方面是希望促进管理学研究成果在全社会的共享；另一方面也希望能够激励更多的中国管理学工作者潜心研究、勇于实践，产生高水准的学术成果，推动中国的管理创新和发展。

衷心祝愿中国管理学的明天更加美好！

序 言 二

成思危

管理学是一门应用性、实践性很强的学科，既有科学的规律可循，又有艺术的运用之妙。改革开放以来，我国管理学扎根于中国特色社会主义的实践沃土，积极回答了改革开放对理论和实践提出的新课题，适应了我国经济建设的迫切需要，并在多学科相互融合中不断发展，初步形成了比较适合我国国情的管理学科体系。

从管理科学与工程方面来看，我国的总体研究水平取得了显著提高。在分析预测方法、不确定性决策理论、群体决策理论、供应链管理、管理复杂性研究等领域，还产生了一批在国际上有影响力的优秀成果。从工商管理方面来看，改革开放实践为中国特色工商管理模式的形成提供了成长沃土，我国学者在股份制公司的组织与运作、公司治理制度的建立与评价、企业战略制定与实施、企业信息管理与电子商务、非公有制企业管理等众多领域进行了深入探索，在建立符合国情的现代企业制度、提高企业管理水平方面作出了重要贡献。在发挥市场资源配置方面的基础性作用的同时，也需要政府通过适当有效的宏观管理加以引导和调控，解决发展中产生的矛盾，维护有序的市场秩序，促进社会公平，保护生态环境，改善社会保障，实现可持续发展的和谐社会，公共管理研究为国家宏观政策制定提供了重要的理论支持。

为了推动我国管理学长远发展，促进我国管理学人才的成长，提高我国管理学在国际上的学术地位和影响力，复旦管理学奖励基金会自 2006 年起，开始奖励我国在管理学学术领域作出杰出贡献的工作者，倡导管理学理论符合中国国情，并密切与实践相结合。获奖人都是活跃在当今管理学学术领域的最优秀学者，获奖人的产生经过了学界的广泛推选，经过了严格的评议过程，始终坚持“创新性、学术性和实用性”的基本评判标准，具有较高的程序公正性和实质公正性。复旦管理学杰出贡献奖是完全由学术界独立完成推选的学术奖项，现在复旦管理学杰出贡献奖逐渐被更多的人了解，产生了一定知名度，在管理学界具有了越来越大的影响力，评选出的获奖人和他们的成果代表着目前我国管理学研究的先进水平。今后我们将持续帮助获奖人出版他们的研究成果，促进学术交流，推动理论繁荣。

“创立中国特色的管理理论、建立中国自己的管理学派”不是一朝一夕可以完成的任务。复旦管理学奖励基金会将通过对中国管理学界的长期支持，努力促成这项事业的成功。现在基金会还只是做了一点基础性的工作。我相信通过 10 年、20 年的努力，通过一代又一代管理学者的辛勤工作，通过有选择地学习和吸收国外经验，有批判地继承中国传统的管理哲学和管理思想，一定能够达到这个目标。

目　录

一、石勇学术成果汇集篇

二、唐立新学术成果汇集篇

三、汪寿阳学术成果汇集篇

一、石勇学术成果汇集篇

个人简介

石勇教授是管理科学与工程领域的国际知名学者。在模糊数学理论、多目标多资源数学规划理论、系统框架和应用、智能知识管理等领域作出了突出的贡献。主要体现在以下方面。

(1) 创立了以多目标线性规划为基础的信用评分系统与数据挖掘理论和方法用多目标最优化概念解决数据挖掘中不同数据同时分割的标准问题。该方法已被广泛应用在信用评分管理、金融信贷风险控制、信息侵扰和侦测、水资源管理、生物信息学、石油勘探及采掘等领域。

(2) 使用最优化理论对"数据—数据挖掘—智能知识管理"这一崭新的交叉性国际科学领域进行研究，首先提出了"智能知识管理"的概念，把数据挖掘的结果作为智能知识的基础，探讨了智能知识与传统知识结构之间的逻辑关系，研究了数据挖掘与智能知识内在联系的数学模型并用其展示智能知识的特征。智能知识研究不仅在数据挖掘与知识管理的交叉学科间推动并促进了传统管理科学的发展，而且具有重要的实际管理应用背景。石勇教授与他的科研团队先后将他们的理论成果应用在中国人民银行个人征信评分系统、网易公司客户管理、中国工商银行的优质客户管理、中国再保险公司的战略管理及中国投资公司的投资匹配决策等影响中国经济发展的重大项目中，其成果所带来的经济价值和社会价值正在不断显现。

(3) 创立多目标多资源线性规划理论、方法及应用的数学及理论框架，该成果以多目标多资源线性规划为基础，广泛研究了多资源水平对已有的最优线性系统设计的影响。提出了线性系统设计不应是寻求给定系统的一个最优点，而应基于资源水平设计一个最优系统。这些研究成果已在会计价格转移、资金预算、生产计划、管理信息系统和通讯管理等方面得到了成功的实际应用。

目前，石勇教授已出版著作15部，在Management Science，Operations Research，Operations Research Letters等国内外期刊发表论文150多篇，其中被SCI收录73篇、EI收录92篇。此外，石勇教授于2002年创办国际SCI学术期刊International Journal of Information Technology and Decision Making，担任主编。

石勇教授还先后获得内布拉斯加州立大学院长卓越研究奖(1993)、第一届国际运筹与数量管理大会优秀论文奖(1997)、内布拉斯加州立大学卓越研究奖(1999)、美国电气电子工程师协会卓越演讲者(1997—2000)、中国国家杰出青年科学家基金奖(2001)、第四届国际主动媒体技术大会优秀论文奖(2006)、第八届、第九届国际计算科学大会计算金融与商业智能分会优秀论文奖、国际多目标决策学会康托尔学术奖(2009)。

Several Multi-criteria Programming Methods for Classification[①]

Juliang Zhang[②], Yong Shi[③], Peng Zhang[③]

1. Introduction

Data mining refers to a set of methods and techniques that can be used to extract some useful information and hidden patterns in the huge amount of data [1]. Here data have an extensive meaning, which can be stored in various forms, such as numerical data, text data, spatial data, multimedia data and web data. Since 1980s, as the huge volumes of business and scientific data are accumulated due to the development of science and society and the need to obtain some useful knowledge from these data, the demand for development of advanced data mining techniques [2] increases significantly. This makes data mining develop quickly.

Data mining is a multi-disciplinary field and encompasses techniques from a number of fields, including information technique, statistic analysis, machine learning (ML), pattern recognition, artificial intelligence (AI) and database management. Recently, lots of authors have tried to apply optimization models to data mining and numerous models have been proposed for classification, clustering and other data mining functionalities which have enhanced both theoretical foundation and practical applications of data mining [3].

There are two objectives of this paper. The first is to propose a general mathematical programming-based model for classification, which includes some well-known methods (e.g., Multiple-Criteria Linear Programming (MCLP) developed by Shi et al. [10] and LP method developed by Freed and Glover [4, 5]) as special cases. And three new models, MCQP (multi-constrained indefinite quadratic programming), MCCQP (multi-constrained concave quadratic programming) and MCVQP (multi-constrained convex programming), are developed based on the general model. The second is to use three real-life datasets: credit card accounts, VIP mail-box and social endowment insurance classification, to test the efficiency of these models. Extensive experiments are done to

---

① 本文发表于 Computers & Operations Research, 2009 年, Vol. 36, 第 823 - 836 页。
② Institute of Logistics, School of Economics and Management, Beijing Jiaotong University, Beijing 100044, China.
③ CAS Research Center on Fictitious Economy and Data Science, Beijing 100080, China.

compare the efficiency of these methods.

This paper is organized as follows. Section 2 briefly reviews the history of applying mathematical programming to classification problems. Section 3 states the generalized multi-criteria classification model, and then formulates four specific cases. The algorithms for concave quadratic programming and the MCQP model are presented in this section. Section 4 presents the experimental results of these models over the datasets of credit card accounts, VIP mail-box and social endowment insurance classification. Finally, Section 5 summarizes the conclusions.

2. Literature review

In 1966, Kendall [6] first proposed an LP-related convex-hull procedure for classification. Since then, many authors studied linear programming method for classification problem. In 1981, Freed and Glover developed a linear programming approach to the discriminant problem [4, 5]. They discussed theoretical difficulties [7, 8], and made improvements to their original method [9] afterward. Comparing with the traditional statistical approaches, linear programming procedure shows a number of advantages, such as simplicity and directness, and free of the statistical assumptions. Shi et al. [10] proposed an MCLP to the real-life credit card accounts classification, which is the first attempt to introduce MCLP to discriminant analysis and apply it to the credit card portfolio management. Peng et al. [11] improved Shi's results and proposed a linear programming method for multiple-group classification.

Another class of optimization models for classification is known as support vector machine (SVM) [12 - 14]. This approach is a nonlinear programming-based approach and adopts convex quadratic programming to model the classification problem. By using kernel function, this method can handle linear non-separable problem. Bugera et al. [15] proposed a quadratic utility function of SVM and applied their revision to credit card accounts classification. Using a dataset of a Greek bank, they showed the efficiency of their method.

Note that the philosophy of the above two classes of models is the same: maximize the interval between the different classes and minimize the misclassification of the observed data. The only difference between these two classes of models is that different measurements of the interval and misclassification are used in these models.

Motivated by these previous studies, this paper proposes a general mathematical programming model for classification problem and develops some new methods for classification based on the general model. Then, we apply these methods to some real problems. The main purpose of this paper is to see the impact of different measurements on the efficiency of classification.

3. Problem formulations

This section introduces a general multi-criteria programming method for classification. Simply speaking, this method is to classify observations into distinct groups based on two criteria for data separation. The following models represent this concept mathematically.

Given an r-dimensional attribute vector $a = (a_1, \cdots, a_r)$, let $A_i = (A_{i1}, \cdots, A_{ir}) \in R^r$ be one of the sample records of these attributes, where $i = 1, \cdots, n$; n represents the total number of records in the dataset. Suppose two groups, G_1 and G_2, are predefined. A boundary scalar b can be set to separate these two groups. A vector $X = (x_1, \cdots, x_r)^T \in R^r$ can be identified to establish the following linear inequality [10, 16]:

$$A_iX < b, \quad \forall A_i \in G_1,$$
$$A_iX \geqslant b, \quad \forall A_i \in G_2.$$

To formulate the criteria and complete constraints for data separation, some variables need to be introduced. In the classification problem, A_iX is the score for the ith data record. Let α_i be the overlapping of two-group boundary for record A_i (external measurement) and β_i be the distance of record A_i from its adjusted boundary (internal measurement). The overlapping α_i means the distance of record A_i to the boundary b if A_i is misclassified into another group. For instance, in Fig. 1 the black dot located to the right of the boundary b belongs to G_1, but it was misclassified by the boundary b to G_2. Thus, the distance between b and the dot equals to α_i.

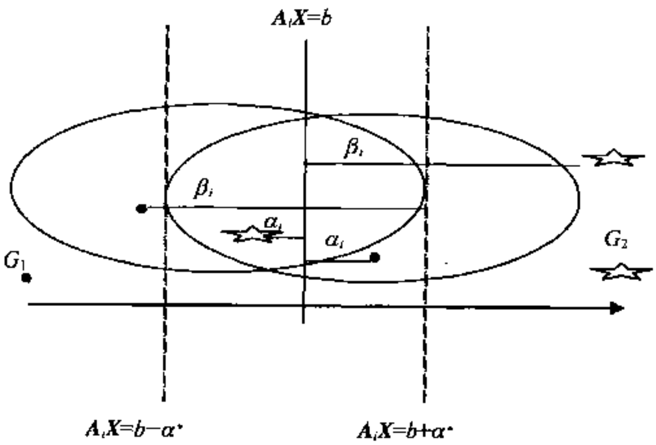

Fig. 1 Two-group classification model

Adjusted boundary is defined as $b-\alpha^*$ or $b+\alpha^*$, while α^* represents the maximum of overlapping. Then, a mathematical function $f(\alpha)$ can be used to describe the relation of all overlapping α_i; while another mathematical function $g(\beta)$ represents the aggregation of all distances β_i. The final classification accuracies depend simultaneously on minimizing $f(\alpha)$

and maximizing $g(\beta)$. Thus, a general bi-criteria programming method for classification can be formulated as

$$\begin{aligned} &\text{General model} \quad && \text{Minimize} \quad f(\alpha) \\ & && \text{Maximize} \quad g(\beta) \\ &\text{Subject to:} && A_iX - \alpha_i + \beta_i - b = 0,\ \forall A_i \in G_1, \\ & && A_iX + \alpha_i - \beta_i - b = 0,\ \forall A_i \in G_2, \end{aligned} \tag{1}$$

where A_i is given, X and b are unrestricted, and $\alpha = (\alpha_1, \cdots, \alpha_n)^T$ and $\beta = (\beta_1, \cdots, \beta_n)^T$ satisfy α_i, $\beta_i \geqslant 0$, $i = 1, \cdots, n$.

All variables and their relationships are represented in Fig. 1. There are two groups in Fig. 1: black dot indicates G_1 data objects and star indicates G_2 data objects. There is one misclassified data object from each group if the boundary scalar b is used to classify these two groups, whereas adjusted boundaries $b - \alpha^*$ and $b + \alpha^*$ separate two groups without misclassification.

We note that different forms of $f(\alpha)$ and $g(\beta)$ will affect the classification criteria. $f(\alpha)$ and $g(\beta)$ can be component-wise and non-decreasing functions. For example, in order to utilize the computational power of some existing nonlinear optimization software packages, a sub-model can be set up by using the l_p norm to represent $f(\alpha)$ and l_q norm to represent $g(\beta)$, respectively. This means $f(\alpha) = \|\alpha\|_p^p$ and $g(\beta) = \|\beta\|_q^q$. Furthermore, to transform the bi-criteria problem of the general model into a single-criterion problem, we use weights $w_\alpha > 0$ and $w_\beta > 0$ for $\|\alpha\|_p^p$ and $\|\beta\|_q^q$, respectively. The values of w_α and w_β can be predefined in the process of identifying the optimal solution. Thus, the general model can be converted into a single-criterion mathematical programming model as

$$\begin{aligned} &\text{Model 1} \quad && \text{Minimize} \quad w_\alpha \|\alpha\|_p^p - w_\beta \|\beta\|_q^q \\ &\text{Subject to:} && A_iX - \alpha_i + \beta_i - b = 0,\ \forall A_i \in G_1, \\ & && A_iX + \alpha_i - \beta_i - b = 0,\ \forall A_i \in G_2. \end{aligned} \tag{2}$$

Based on Model 1, mathematical programming models with any norm can be theoretically defined. This study is interested in formulating a linear or a quadratic programming model.

Case 1: $p = q = 1$ (MCLP)

In this case, $\|\alpha\|_1 = \sum_{i=1}^n \alpha_i$ and $\|\beta\|_1 = \sum_{i=1}^n \beta_i$. Objective function in Model 1 can now be a linear objective function:

$$\text{(linear)} \quad \text{Minimize} \quad w_\alpha \sum_{i=1}^n \alpha_i - w_\beta \sum_{i=1}^n \beta_i.$$

Model 1 turns to be a linear programming

$$\begin{aligned} &\text{Minimize} \quad w_\alpha \sum_{i=1}^{n}\alpha_i - w_\beta \sum_{i=1}^{n}\beta_i \\ &\text{Subject to:} \quad A_iX - \alpha_i + \beta_i - b = 0,\ \forall A_i \in G_1, \\ &\qquad A_iX + \alpha_i - \beta_i - b = 0,\ \forall A_i \in G_2 \\ &\qquad \alpha_i,\ \beta_i \geqslant 0,\ i = 1,\ \cdots,\ n. \end{aligned} \tag{3}$$

This model has been obtained by [4, 5, 10]. There are many softwares which can solve linear programming problem very efficiently at present. We can use some of these softwares to solve the linear programming (3).

Case 2: $p = 2$, $q = 1$ (MCVQP)

In this case, $\|\alpha\|_2^2 = \sum_{i=1}^{n}\alpha_i^2$ and $\|\beta\|_1 = \sum_{i=1}^{n}\beta_i$ Objective function in Model 1 can now be a convex quadratic objective:

$$\text{(convex quadratic)} \quad \text{Minimize} \quad w_\alpha \sum_{i=1}^{n}\alpha_i^2 - w_\beta \sum_{i=1}^{n}\beta_i.$$

Model 1 turns to be a convex quadratic programming

$$\begin{aligned} &\text{Minimize} \quad w_\alpha \sum_{i=1}^{n}\alpha_i^2 - w_\beta \sum_{i=1}^{n}\beta_i \\ &\text{Subject to:} \quad A_iX - \alpha_i + \beta_i - b = 0,\ \forall A_i \in G_1, \\ &\qquad A_iX + \alpha_i - \beta_i - b = 0,\ \forall A_i \in G_2, \\ &\qquad \alpha_i,\ \beta_i \geqslant 0,\ i = 1,\ \cdots,\ n. \end{aligned} \tag{4}$$

It is well known that convex quadratic programming can be solved easily. In this paper, we use the method developed by Powell [17] to solve problem (4).

Case 3: $p = 1$, $q = 2$ (MCCQP)

In this case, $\|\alpha\|_1 = \sum_{i=1}^{n}\alpha_i$ and $\|\beta\|_2 = \sum_{i=1}^{n}\beta_i^2$ Objective function in Model 1 can now be a concave quadratic function:

$$\text{(concave quadratic)} \quad \text{Minimize} \quad w_\alpha \sum_{i=1}^{n}\alpha_i - w_\beta \sum_{i=1}^{n}\beta_i^2.$$

Model 1 turns to be a concave quadratic programming

$$\begin{aligned} &\text{Minimize} \quad w_\alpha \sum_{i=1}^{n}\alpha_i - w_\beta \sum_{i=1}^{n}\beta_i^2 \\ &\text{Subject to:} \quad A_iX - \alpha_i + \beta_i - b = 0,\ \forall A_i \in G_1, \\ &\qquad A_iX + \alpha_i - \beta_i - b = 0,\ \forall A_i \in G_2, \\ &\qquad \alpha_i,\ \beta_i \geqslant 0,\ i = 1,\ \cdots,\ n. \end{aligned} \tag{5}$$

Concave quadratic programming is an NP-hard problem. It is very difficult to find the global minimizer, especially for large problem. In order to solve (5) efficiently, we propose an algorithm, which converges to a local minimizer of (5).

In order to describe the algorithm in detail, we introduce some notation.

Let $\omega=(X,\alpha,\beta,b)$, $f(\omega)=w_\alpha\sum_{i=1}^{n}\alpha_i-w_\beta\sum_{i=1}^{n}\beta_i^2$ and

$$\Omega=\left\{\begin{array}{c}(X,\alpha,\beta,b):A_iX-\alpha_i-\beta_i-b=0,\ \forall A_i\in G_1,\\ A_iX+\alpha_i-\beta_i-b=0,\ \forall A_i\in G_2,\\ \alpha_i\geqslant 0,\ \beta_i\geqslant 0,\ i=1,\cdots,n\end{array}\right\}$$

be the feasible region of Model 1.

Let $\chi_\Omega(\omega)$ be the index function of set Ω, i. e., $\chi_\Omega(\omega)$ is defined as follows:

$$\chi_\Omega(\omega)=\begin{cases}0, & \omega\in\Omega,\\ +\infty, & \omega\notin\Omega.\end{cases}$$

Then (5) is equivalent to the following problem

$$\min f(\omega)+\chi_\Omega(\omega). \tag{6}$$

Rewrite $f(\omega)+\chi_\Omega(\omega)$ as the following form

$$f(\omega)+\chi_\Omega(\omega)=g(\omega)-h(\omega),$$

where $g(\omega)=\frac{1}{2}\rho\|\omega\|^2+w_\alpha\sum_{i=1}^{n}\alpha_i+\chi_\Omega(\omega)$, $h(\omega)=\frac{1}{2}\rho\|\omega\|^2+w_\beta\sum_{i=1}^{n}\beta_i^2$ and $\rho>0$ is a small positive number. Then $g(\omega)$ and $h(\omega)$ are convex functions. By applying the simplified DC algorithm in [18] to problem (6), we get the following algorithm.

Algorithm 1. Given an initial point $\omega^0\in R^{3n+1}$ and a parameter $\varepsilon>0$, at each iteration $k\geqslant 1$, compute ω^{k+1} by solving the convex quadratic programming

$$(Q^k)\quad \min\left\{\frac{1}{2}\rho\|\omega\|^2+w_\alpha\sum_{i=1}^{n}\alpha_i-(h'(\omega^k),\omega),\ \omega\in\Omega\right\}.$$

The stopping criterion is $\|\omega^{k+1}-\omega^k\|\leqslant\varepsilon$.

By standard arguments, we can prove the following theorem.

Theorem 3.1. *After finite iterations, Algorithm* 1 *terminates at a local minimizer of* (5).

Case 4: $p=q=2$ (MCQP)

In this case, $\|\alpha\|_2=\sum_{i=1}^{n}\alpha_i^2$ and $\|\beta\|_2=\sum_{i=1}^{n}\beta_i^2$. The objective function in Model 1 can now be an indefinite quadratic function:

$$\text{(indefinite quadratic)}\quad \text{Minimize}\quad w_\alpha\sum_{i=1}^{n}\alpha_i^2-w_\beta\sum_{i=1}^{n}\beta_i^2.$$

Model 1 turns to be an indefinite quadratic programming