国家自然科学基金资助项目研究成果

预测理论及其应用

唐小我 著

电子科技大学出版社

国家自然科学基金资助项目研究成果

预测理论及其应用

唐小我 著

电子科技大学出版社

·1992·

内 容 提 要

本书是国家自然科学基金资助项目“关于建立经济预测模型的算法研究”的研究成果。其主要内容包括：组合预测的基本理论，最优组合预测的计算方法，递归等权组合预测方法；多元组合证券投资的风险预测；投入产出系统的计算、分解及模型系数修订；非线性盈亏平衡分析的代数方法，盈亏平衡点的变动预测；马尔科夫链转移概率矩阵的确定，稳态市场占有率向量的简便计算；多层递阶预测方法，变系数线性回归分析，变权数组合预测模型等。

本书理论联系实际，内容新颖、丰富，适用于预测、经济、管理、系统工程等领域的理论工作者和实际工作者，还可作为大专院校有关专业高年级大学生及研究生教材或教学参考书。

国家自然科学基金资助项目

预测理论及其应用

唐小我 著

*

电子科技大学出版社

（中国成都建设北路二段四号）

电子科技大学出版社激光照排中心照排

成都国营农垦印刷厂胶印

四川省新华书店经销

*

开本 787×1092 1/32 印张 9.1875 字数 190 千字

版次 1992 年 11 月第一版 印次 1992 年 11 月第一次印刷

印数 1—1500 册

中国标准书号 ISBN 7-81016-456-2/F·35

[川]016 (4452·36) 定价：5.50元

前　　言

本书是我完成国家自然科学基金资助项目“关于建立经济预测模型的算法研究”所取得的研究成果，是我1989年出版的《经济预测方法》一书的继续深入和扩展。

本书对组合预测方法，组合证券投资的风险预测，投入产出系统的计算、分解及模型系数修订，盈亏平衡点的变动预测，马尔科夫链转移概率矩阵的建立，稳态市场占有率向量的简便计算，变系数预测方法等进行了深入研究，给出了一系列新方法、改进方法和应用实例。我相信，本书对于我国预测理论的完善、发展与应用，具有重要的促进作用。愿广大读者喜欢这本书，并从中得到某种启迪。

在研究过程中，《预测》杂志编辑部陈玉祥教授和李丰、赵定涛老师等对我给予了大力支持，在此向他们表示衷心的感谢。

我的导师曹长修教授仔细审阅了本书，并提出了宝贵的意见。借此机会，向曹长修教授表示由衷的谢意。

最后，对大力支持、鼓励和培养我的国家自然科学基金会表示诚挚的谢意。

唐小我

1992年9月于电子科技大学

目　录

第一章　组合预测方法

第一节　组合预测方法简述

在预测实践中，对于同一个问题，我们常采用不同的预测方法。不同的预测方法其预测精度往往也不相同。一般是以预测误差平方和作为评价预测方法优劣的标准，从各种预测方法中选取预测误差平方和最小的那种预测方法。不同的预测方法往往能提供不同的有用信息，如果简单地将预测误差平方和较大的一些方法舍弃掉，将会丢失一些有用的信息，这是一种浪费，应予以避免。一种更为科学的作法是，将不同的预测方法进行适当的组合，形成所谓的组合预测方法。组合的主要目的是综合利用各种预测方法所提供的信息，尽可能地提高预测精度。只要组合得当，这一目的是完全可以实现的。

组合预测在国外常称为 Combination forecasting，Combined forecasting 或 Combined forecast 等，在国内也被称为结合预测，综合预测或复合预测。本书统一采用组合预测这一名称。

早在 1954 年，美国人 Schmitt 就曾经用组合预测方法对美国 37 个最大城市的人口数字进行过预测，使预测精度有所提高。1959 年，J・M・Bates 和 C・W・J・Granger 对组合预测方法进行了比较系统的研究，其研究成果引起了预测学者的重视。进入 70 年代以来，组合预测的研究更被预测工作者

所重视，发表了一系列关于组合预测的论文，大有方兴未艾之势。1989 年，国际预测领域的权威性学术刊物《Journal of Forecasting》还出版了组合预测专辑，充分说明了组合预测在预测研究中的重要地位。

近几年来，我国在组合预测方法研究方面也取得一些研究成果。这些研究成果主要发表于《预测》、《财经科学》、《电子科技大学学报》等学术刊物。特别是我国预测领域中的权威性学术刊物《预测》杂志，对组合预测研究一直予以高度重视，自 1984 年以来，每年都要发表关于组合预测方面的研究论文，为促进我国组合预测的理论研究与应用作出了重要的贡献。

本章将结合实例介绍 最优组合预测的计算方法，递归等权组合预测方法及其改进方法等，对最优组合预测的基本理论也将进行系统的阐述。

第二节　最优组合预测方法的计算公式

一、最优加权系数向量的计算公式

采用组合预测的关键是确定各个单项预测方法的加权系数。下面我们来推导最优加权系数的计算公式。

设对于同一预测问题，我们有 $n(\geqslant 2)$ 种预测方法。

设　y_t——实际观测值，$t=1,\cdots,N$；

f_{it}——第 i 种方法的预测值，$i=1,\cdots,n$，$t=1,\cdots,N$；

$e_{it}=y_t-f_{it}$——第 i 种方法的预测误差，$i=1,\cdots,n$，$t=1,\cdots,N$；

k_i——第 i 种方法的加权系数，$i=1,2,\cdots,n$，$\sum\limits_{i=1}^{n} k_i=1$；

$f_t=\sum\limits_{i=1}^{n} k_i f_{it}$——加权平均预测值，即组合预测方法的预测值，$t=1,\cdots,N$；

$e_t=y_t-f_t$——组合预测方法的预测误差，$t=1,\cdots,N$；

e_t 可以表示为

$$
\begin{aligned}
e_t &= y_t - f_t \\
&= \sum_{i=1}^{n}(k_i y_t) - \sum_{i=1}^{n} k_i f_{it} \\
&= \sum_{i=1}^{n} k_i (y_t - f_{it}) \\
&= \sum_{i=1}^{n} k_i e_{it}
\end{aligned}
$$

$$
\begin{aligned}
e_t^2 &= \left(\sum_{i=1}^{n} k_i e_{it}\right)^2 \\
&= \left([k_1,k_2,\cdots,k_n][e_{1t},e_{2t},\cdots,e_{nt}]^T\right)^2 \\
&= \begin{bmatrix} k_1 \\ k_2 \\ \vdots \\ k_n \end{bmatrix}^T \begin{bmatrix} e_{1t} \\ e_{2t} \\ \vdots \\ e_{nt} \end{bmatrix} \begin{bmatrix} e_{1t} \\ e_{2t} \\ \vdots \\ e_{nt} \end{bmatrix}^T \begin{bmatrix} k_1 \\ k_2 \\ \vdots \\ k_n \end{bmatrix} \\
&= \begin{bmatrix} k_1 \\ k_2 \\ \vdots \\ k_n \end{bmatrix}^T \begin{bmatrix} e_{1t}^2 & e_{1t}e_{2t} & \cdots & e_{1t}e_{nt} \\ e_{2t}e_{1t} & e_{2t}^2 & \cdots & e_{2t}e_{nt} \\ \vdots & \vdots & & \vdots \\ e_{nt}e_{1t} & e_{nt}e_{2t} & \cdots & e_{nt}e_{nt} \end{bmatrix} \begin{bmatrix} k_1 \\ k_2 \\ \vdots \\ k_n \end{bmatrix}
\end{aligned}
$$

记组合预测方法的预测误差平方和为 $J=\sum\limits_{t=1}^{N} e_t^2$，则有

$$
\begin{aligned}
J &= \sum_{t=1}^{N} e_t^2 \\
&= \sum_{i=1}^{n} \sum_{j=1}^{n} \left[k_i k_j \left(\sum_{t=1}^{N} e_{it} e_{jt} \right) \right] \\
&= \begin{bmatrix} k_1 \\ k_2 \\ \vdots \\ k_n \end{bmatrix}^T
\begin{bmatrix}
\sum_{t=1}^{N} e_{1t}^2 & \sum_{t=1}^{N} e_{1t}e_{2t} & \cdots & \sum_{t=1}^{N} e_{1t}e_{nt} \\
\sum_{t=1}^{N} e_{2t}e_{1t} & \sum_{t=1}^{N} e_{2t}^2 & \cdots & \sum_{t=1}^{N} e_{2t}e_{nt} \\
\vdots & \vdots & & \vdots \\
\sum_{t=1}^{N} e_{nt}e_{1t} & \sum_{t=1}^{N} e_{nt}e_{2t} & \cdots & \sum_{t=1}^{N} e_{nt}^2
\end{bmatrix}
\begin{bmatrix} k_1 \\ k_2 \\ \vdots \\ k_n \end{bmatrix}
\end{aligned}
\tag{1.1}
$$

记组合预测方法的加权系数向量为 $K=[k_1,k_2,\cdots,k_n]^T$，第 i 种预测方法的预测误差向量为 $E_i=[e_{i1},e_{i2},\cdots,e_{iN}]^T$，预测误差矩阵为 $e=[E_1E_2\cdots E_n]$，则组合预测方法的预测误差平方和 J 也可以表示为

$$
J = e^T e = K^T E_{(n)} K \tag{1.2}
$$

其中

$$
E_{(n)} = \begin{bmatrix}
E_{11} & E_{12} & \cdots & E_{1n} \\
E_{21} & E_{22} & \cdots & E_{2n} \\
\vdots & \vdots & & \vdots \\
E_{n1} & E_{n2} & \cdots & E_{nn}
\end{bmatrix} \tag{1.3}
$$

式(1.3)中， $E_{ij}=E_{ji}=E_i^T E_j=\sum_{t=1}^{N} e_{it}e_{jt}, E_{ii}=E_i^T E_i=\sum_{t=1}^{N} e_{it}^2$。

不难看出，E_{ii}即为第 i 种预测方法的预测误差平方和。$E_{(n)}$集中反映了各种预测方法所提供的预测误差信息，本书称之为预测误差信息矩阵。如果预测误差向量 $E_1,E_2,\cdots,E_n$ 是线性无关的，则 $E_{(n)}$是对称正定阵，从而其逆矩阵存在。我们

总假定 $E_{(n)}$是可逆阵。

记 R_n 表示分量全为 1 的 n 维列向量，即 $R_n=[1,1,\cdots,1]^T$，则加权系数的约束条件 $\sum_{i=1}^{n}k_i=1$ 可以改写为

$$R_n^T K = 1 \tag{1.4}$$

采用组合预测方法的主要目的是减少预测误差平方和。如果某一加权系数向量 K_* 使组合预测方法的预测误差平方和 J 达到极小值 J_*，即 $J_*=\min\limits_{R_n^TK=1}\{J\}=\min\limits_{R_n^TK=1}\{K^TE_{(n)}K\}$，则称 K_* 为最优加权系数向量，其所对应的组合预测方法称为最优组合预测方法。下面我们给出计算 K_* 和 J_* 的有关公式。

定理 1 设 $E_1,E_2,\cdots,E_n$ 线性无关，则有

$$K_* = E_{(n)}^{-1}R_n/R_n^TE_{(n)}^{-1}R_n \tag{1.5}$$

$$J_* = 1/R_n^TE_{(n)}^{-1}R_n \tag{1.6}$$

证明　加权系数的约束条件为式(1.4)。引入拉格朗日乘数 λ 后，J 可以表示为

$$J = K^TE_{(n)}K + \lambda(R_n^TK - 1)$$

要使 J 取极小值，J 对 K 的一阶导数必为零。根据标量对向量的求导公式可得

$$\partial J/\partial K = 2E_{(n)}K + \lambda R_n = 0$$

最优加权系数向量 K_* 必须满足上式，即

$$2E_{(n)}K_* + \lambda R_n = 0 \tag{1.7}$$

将 $R_n^TE_{(n)}^{-1}$左乘式(1.7)的两端可得

$$2R_n^TK_* + \lambda R_n^TE_{(n)}^{-1}R_n = 0 \tag{1.8}$$

由式(1.4)可得 $R_n^TK_*=1$，将此结果代入式(1.8)得 $2+\lambda R_n^TE_{(n)}^{-1}R_n=0$，即 $\lambda=-2/R_n^TE_{(n)}^{-1}R_n$。将此结果代入式(1.7)可解

出 K_n 为

$$K_n = E_{(n)}^{-1}R_n / R_n^T E_{(n)}^{-1} R_n \tag{1.9}$$

显然有 $R_n^T K_n = R_n^T E_{(n)}^{-1} R_n / R_n^T E_{(n)}^{-1} R_n = 1$，即约束条件式(1.4)得到满足。

$$\begin{aligned} J_n = K_n^T E_{(n)} K_n &= (E_{(n)}^{-1}R_n / R_n^T E_{(n)}^{-1} R_n)^T E_{(n)} (E_{(n)}^{-1} R_n / R_n^T E_{(n)}^{-1} R_n) \\ &= (R_n^T / R_n^T E_{(n)}^{-1} R_n)(E_{(n)}^{-1} R_n / R_n^T E_{(n)}^{-1} R_n) \\ &= 1/R_n^T E_{(n)}^{-1} R_n \end{aligned}$$

定理 1 中计算最优加权系数向量 K_n 的公式还可以作一些简化。

由矩阵的求逆公式可知，$E_{(n)}^{-1} = \dfrac{E_{(n)}^*}{\det E_{(n)}}$，其中 $E_{(n)}^*$ 为 $E_{(n)}$ 的伴随矩阵，$\det E_{(n)}$ 为 $E_{(n)}$ 行列式。不难得到

$$K_n = \frac{(E_{(n)}^* / \det E_{(n)}) R_n}{R_n^T (E_{(n)}^* / \det E_{(n)}) R_n} = \frac{E_{(n)}^* R_n}{R_n^T E_{(n)}^* R_n} \tag{1.10}$$

根据式(1.10)计算 $K_{(n)}$ 就不必计算 $E_{(n)}$ 的行列式了。这时，J_n 可按 $J_n = K_n^T E_{(n)} K_n$ 来进行计算。

不过应该强调指出，公式 $J_n = 1/R_n^T E_{(n)}^{-1} R_n$ 在理论分析方面是很有用的，以后将多次用到这一公式。

在预测实践中，用得较多的是由两种预测方法所构成的组合预测方法。下面我们给出两种方法组合的有关公式。

定理 2 对于两种方法组合的情形，我们有下述结论：

结论 1 最优加权系数为

$$k_1 = \frac{E_{22} - E_{12}}{E_{11} + E_{22} - 2E_{12}} = \frac{\sum_{t=1}^{N} e_{2t}^2 - \sum_{t=1}^{N} e_{1t}e_{2t}}{\sum_{t=1}^{N} e_{1t}^2 + \sum_{t=1}^{N} e_{2t}^2 - 2\sum_{t=1}^{N} e_{1t}e_{2t}} \tag{1.11}$$

$$k_2 = \frac{E_{11} - E_{12}}{E_{11} + E_{22} - 2E_{12}} = \frac{\sum_{t=1}^{N} e_{1t}^2 - \sum_{t=1}^{N} e_{1t}e_{2t}}{\sum_{t=1}^{N} e_{1t}^2 + \sum_{t=1}^{N} e_{2t}^2 - 2\sum_{t=1}^{N} e_{1t}e_{2t}} \tag{1.12}$$

其中 $E_{11} = \sum_{t=1}^{N} e_{1t}^2, E_{22} = \sum_{t=1}^{N} e_{2t}^2, E_{12} = \sum_{t=1}^{N} e_{1t}e_{2t}$

结论 2 最优组合预测方法的预测误差平方和为

$$J_2 = \frac{E_{11}E_{22} - E_{12}^2}{E_{11} + E_{22} - 2E_{12}} = \frac{\sum_{t=1}^{N} e_{1t}^2 \sum_{t=1}^{N} e_{2t}^2 - \left(\sum_{t=1}^{N} e_{1t}e_{2t}\right)^2}{\sum_{t=1}^{N} e_{1t}^2 + \sum_{t=1}^{N} e_{2t}^2 - 2\sum_{t=1}^{N} e_{1t}e_{2t}} \tag{1.13}$$

证明(略)。

二、简单平均法是最优组合预测方法的充分必要条件

简单平均法又叫做等权平均法，是目前讨论得较多的组合预测方法之一。简单平均法的特点是对各种方法同等看待，对每种方法都赋予相同的权数。简单平均法的加权系数向量记为 K_A，K_A 为同分量向量，每个分量均为 $1/n$，即 $K_A=[1/n, 1/n,\cdots,1/n]^T=1/n[1,1,\cdots,1]^T=(1/n)R_n$。简单平均法的预测误差平方和记为 J_A，容易证明

$$J_A = K_A^T E_{(n)} K_A = \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} E_{ij}$$

简单平均法通常不是最优组合预测方法，但是我们不能排除简单平均法正好就是最优组合预测方法的可能性。关于这个问题，我们有如下结论。

定理 3 简单平均法是最优组合预测方法的充分必要条件是：预测误差信息矩阵 $E_{(n)}$ 的每一行元素之和为常数，即

$$h_i = \sum_{j=1}^{n} E_{ij} = h = 常数, i = 1,2,\cdots,n \tag{1.14}$$

证明 先来证充分性。设式(1.14)成立，则有

$$[E_{i1}, E_{i2}, \cdots, E_{in}]R_n = h, \qquad i = 1,2,\cdots n \tag{1.15}$$

将式(1.15)写成矩阵形式可得 $E_{(n)}R_n = hR_n$ (1.16)

由式(1.16)可得

$$E_{(n)}^{-1}R_n = R_n/h \tag{1.17}$$

最优加权系数向量为

$$K_n = \frac{E_{(n)}^{-1}R_n}{R_n^T E_{(n)}^{-1} R_n} = \frac{\dfrac{R_n}{h}}{R_n^T \dfrac{R_n}{h}} = \frac{R_n}{R_n^T R_n} = \frac{R_n}{n} = K_A \tag{1.18}$$

由式(1.18)可知，简单平均法是最优组合预测方法。这就证明了充分性。

设简单平均法是最优组合预测方法，则有

$$K_A = \frac{1}{n}R_n = \frac{E_{(n)}^{-1}R_n}{R_n^T E_{(n)}^{-1} R_n} \tag{1.19}$$

由式(1.19)可得

$$E_{(n)}R_n = \frac{nR_n}{R_n^T E_{(n)}^{-1} R_n} \tag{1.20}$$

记 $n/R_n^T E_{(n)}^{-1} R_n = h$，则 $E_{(n)}R_n = hR_n$，也即 $\sum\limits_{j=1}^{n} E_{ij} = h, i = 1,2,\cdots,n$。这就证明了必要性。

下面我们给出关于 J_A 的一个结论。

定理 4 设 $E_{(n)}$ 的行和为 h，则最优组合预测方法的预测误差

平方和为

$$J_A = \frac{h}{n} \tag{1.21}$$

证明 $J_A=\frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}E_{ij}=\frac{1}{n^2}\sum_{i=1}^{n}h=\frac{nh}{n^2}=\frac{h}{n}$。

由于在多数情况下，式(1.14)很难得到满足，故简单平均法通常不是最优组合预测方法。由于最优组合预测方法所涉及的计算量并不大，因此，我们应尽量采用最优组合预测方法而不是简单平均法。

三、计算实例

例 1 孔庆凯用最小平方法和三次指数平滑法对 1971 年至 1982 年乡村居民标准储蓄存款余额进行了预测。两种预测方法的预测误差平方和分别为 553.9793 和 188.8769。采用简单平均法后，预测误差平方和减小到 $J_A=141.2411$。试对同一问题采用最优组合预测方法。有关数据如表 1-1 所示。

解 我们用 e_{1t}和 e_{2t}分别表示最小平方法和三次指数平滑法的预测误差。

$$\begin{aligned} E_{11} &= \sum_{t=1}^{12} e_{1t}^2 \\ &= (-6.98)^2 + (-1.48)^2 + \cdots + 12.51^2 \\ &= 553.9707 \end{aligned}$$

$$\begin{aligned} E_{22} &= \sum_{t=1}^{12} e_{2t}^2 \\ &= 1.46^2 + 1.83^2 + \cdots + 1.31^2 = 188.8769 \end{aligned}$$

$$E_{12} = E_{21}$$

$$= \sum_{t=1}^{12} e_{1t}e_{2t}$$

$$= (-6.98) \times 1.46 + (-1.48) \times 1.83 + \cdots + 12.51 \times 1.31$$

$$= -88.9283$$

表 1-1　对至上年末乡村居民标准储蓄余额的预测

序号	年份	至上年末乡村居民标准储蓄存款余款（亿元）y_t	最小平方法预测值（亿元）f_{1t}	预测误差（亿元）$e_{1t}=y_t-f_{1t}$	三次指数平滑法预测值（亿元）f_{2t}	预测误差（亿元）$e_{2t}=y_t-f_{2t}$
1	71	11.49	18.47	−6.98	10.03	1.46
2	72	13.06	14.54	−1.48	11.23	1.83
3	73	15.34	12.84	2.5	15.24	0.10
4	74	20.58	13.38	7.2	18.67	1.91
5	75	23.28	16.15	7.13	27.78	−4.50
6	76	26.46	21.16	5.3	26.36	0.10
7	77	27.33	28.40	−1.07	29.67	−2.34
8	78	34.22	37.87	−3.65	27.40	6.82
9	79	40.19	49.58	−9.39	42.73	−2.54
10	80	53.37	63.53	−10.16	47.36	6.01
11	81	77.79	79.71	−1.92	69.84	7.95
12	82	10.63	98.12	12.51	109.32	1.31

预测误差信息矩阵为

$$E_{(2)}=\begin{bmatrix}E_{11} & E_{12}\\ E_{21} & E_{22}\end{bmatrix}=\begin{bmatrix}553.9707 & -88.9283\\ -88.9283 & 188.8769\end{bmatrix}$$

$E_{(2)}$的第一行元素之和为 $h_1=553.9707-88.9283$

$=465.051$

$E_{(2)}$的第二行元素之和为 $h_2=-88.9293+188.8769$

$=99.9486$

由于 $h_1\neq h_2$,故简单平均法不是最优组合预测方法。

最优组合预测方法的加权系数为

$$k_1=\frac{E_{22}-E_{12}}{E_{11}+E_{22}-2E_{12}}$$

$$=\frac{188.8769-(-88.9283)}{553.9293+188.8769-2\times(-88.9283)}=0.3017$$

$$k_2=\frac{E_{11}-E_{12}}{E_{11}+E_{22}-2E_{12}}$$

$$=\frac{553.9707-(-88.9283)}{553.9707+188.8769-2\times(-88.9283)}=0.6983$$

计算结果表明,最小平方法的加权系数为 0.3017,三次指数平滑法的加权系数为 0.6983,即预测误差平方和越大的方法,其权数越小。

最优组合预测方法的预测误差平方和为 J_2

$$J_2=\frac{E_{11}E_{22}-E_{12}^2}{E_{11}+E_{22}-2E_{12}}$$

$$\frac{553.9707\times188.8769-(-88.9293)^2}{553.9707+188.8769-2\times(-88.9293)}$$

$$=105.0552$$

与简单平均法相比,最优组合预测方法的预测误差平方和减少了 141.2411-105.0552=36.1859。可见改进是很明