第7卷 第3期　　华侨大学学报　　Vol.7 No.3
1986年8月　　JOURNAL OF HUACHIAO UNIVERSITY　　Aug. 1986

关于分布自由的逐步回归集成预报法

吴绍敏　陈治典

（应用数学系）

摘　　要

继文〔1〕本文提出在不同信度下，建立多个以秩和为因子的逐步回归方程，尔后应用集成最优法求出最优集成预报方程，它比文〔1〕中的方法更充分地利用了大量因子的信息，从而提高了预报的精度和可靠性，其回报准确率和试报效果，均比普通逐步回归集成最优方程及文〔1〕中以秩和为因子的单个逐步回归方程更为优越.

一、引　　言

人们常用的逐步回归法要求预报对象符合正态线性模型，这使它的应用范围受到很大的限制．我们采用对历史数据进行秩变换[1]，然后建立逐步回归方程，这样的方程关于分布是自由的，从而扩大了应用范围．此外，由于使用计算机，对每一个预报对象往往可以找到几十个甚至上百个相关因子，可是，用逐步回归法只能选入其中的极少数的几个因子，因而浪费了大量因子的信息，所建立的预报方程其效果也不会很好．为了最充分地利用因子的信息，以建立效果更好的预报方程，我们首先采用适当的方法将每若干个因子组成一组，取它们的秩数之和作为一个新的因子，这样每一个新因子就集中了一组因子的信息，新因子与预报对象的相关系数有显著提高，利用新因子跟预报对象之秩在不同的信度下建立多个逐步回归方程，然后再将这些方程集成，求出最优预报方程，这样就更充分地利用了大量因子（包括线性与非线性的）信息，从而得出良好的预报效果。

二、方 法 步 骤

1、因 子 选 择

根据专业知识及实际经验选择一批跟预报对象 y 可能有关的因子，通过秩相关系数检验法，从中选出与 y 关系较为密切的因子$x_1, x_2, \cdots, x_l$。设 y 及各因子都有 n 个历史观测值，分别记为 y_1，$y_2\cdots$，y_n；x_{i_1}，x_{i_2}，$\cdots$，$x_{i_n}(i=\overline{1,l})$，以$x_{10}$，$x_{20}$，$\cdots$，$x_{l0}$，分别表示各因子的当

本文1985年7月29日收到。

前值。

2、对历史数据作秩变换

将 y 的历史数据按从小到大的顺序排列定秩[1]，以 $R(y_i)$ 表示 y_i 之秩($i=\overline{1,n}$)，若因子与 y 正相关，则将其历史数据按从小到大的顺序排列定秩；若因子与 y 负相关，则将其历史数据按从大到小的顺序排列定秩。以 $R(x_{kj})$ 表示 x_{kj} 之秩($j=\overline{1,n}$；$k=\overline{1,l}$)，至于因子当前值则以它与该因子历史数据中最接近的那个数据之秩作为它的秩的估计值，分别以 $R(x_{ko})$ 表示($k=\overline{1,l}$)。

3、构造秩和新因子

用适当的办法将全体因子每5个组成一组，取它们的秩数之和[1]，再作秩变换得一个新因子，共可得 $m=\left[\frac{l}{5}\right]$ 个新因子，记为 x_1'，x_2'，…，x'_m。每个新因子都集中了5个因子的信息，並且由于经过了秩变换，新因子与 y 之秩的线性相关系数一般都是显著地高于原因子跟 y 的相关系数(看后实例)。显然用这些新因子跟 y 之秩，可望建立更好的预报方程，並且根据文〔2〕，经秩变换后用于作显著性检验的 F-统计量亦近似地服从 F-分布，所以用新因子建立的预报方程不受正态线性的限制。

4、在不同的信度下建立多个逐步回归方程

给定一个较高的显著性水平 $\alpha_1(0<\alpha_1<1)$，用逐步回归法建立新因子与 y 之秩的回归方程：

$$R(\hat{y}^{(1)})=\hat{b}_0{}^{(1)}+\sum_{k=1}^{m_1}\hat{b}_k{}^{(1)}x_k{}^{(1)},\quad x_k{}^{(1)}\in\{x'_i,\ i=\overline{1,m}\},\quad k=\overline{1,m_1}$$

尔后，再给一个低一点的显著性水平 $\alpha_2(>\alpha_1)$，建立第二个回归方程。

$$R(\hat{y}^{(2)})=\hat{b}^{(2)}{}_0+\sum_{k=1}^{m_2}\hat{b}_k{}^{(2)}x_k{}^{(2)}$$

如此继续做下去，直到在某一低的显著性水平下已挑不出新因子来建立回归方程为止。设最后建立了 S 个预极 y 之秩的回归方程：

$$R(\hat{y}^{(i)})=\hat{b}_0^{(i)}+\sum_{k=1}^{mi}\hat{b}_k^{(i)}x_k^{(i)},\qquad i=\overline{1,S} \tag{1}$$

其中 $x_k^{(i)}\in\left\{x_1^1,\ x_2^1,\ \cdots,\ x_m^1\right\}$ 且各不相同，$\sum_{i=1}^{s}m_i<m$。

实践证明，在较高的置信度下建立的回归方程未必预报效果就是最好的，若综合多个回归方程的作用，则更充分地利用大量因子的信息，因而可得到更好的预报效果，为此采用下述集成法。

5.建立集成预报方程

以 $x_k^{(i)}(k=\overline{1,m_i})$ 的各历史数据代入上述第 i 个回归方程，即可求预报对象按第 i 个回归方

程的各历史回报值之秩

$$R(\hat{y}_j^{(i)})=b_0^{(i)}+\sum_{k=1}^{m_i}\hat{b}_k^{(i)}x_{kj}^{(i)},\quad j=\overline{1,\ n}.$$

以$Q^{(i)}$表示第i个方程的误差平方和，即

$$Q^{(i)}=\sum_{j=1}^{n}\left(R(y_j)-R(\hat{y}_j^{(i)})\right)^2$$

按式(1)令$i=1,\ 2,\ \cdots,\ S$，可求出各回归方程的误差平方和$Q^{(1)},\ Q^{(2)},\ \cdots,Q^{(s)}$。记

$$\varepsilon_j^{(i)}=R(y_j)-R(\hat{y}_j^{(i)})\quad(i=1,\ 2,\cdots,\ s;j=1,\ 2,\ \cdots,\ n)$$

建立集成预报方程为

$$R(\hat{y})=\sum_{i=1}^{s}a_iR(\hat{y}^{(i)}),\qquad(\sum_{i=1}^{s}a_i=1)\tag{2}$$

其中$a_i(i=1,\ 2,\ \cdots,\ s)$为待定常数。这样建立的预报方程之误差平方和为

$$\triangle=\sum_{j=1}^{n}\left[R(\hat{y}_j)-R(y_j)\right]^2=\sum_{j=1}^{n}\left[\sum_{i=1}^{s}a_iR(\hat{y}_j^{(i)})-\sum_{i=1}^{s}a_iR(y_j)\right]^2$$

$$=\sum_{j=1}^{n}\left\{\sum_{i=1}^{s}a_i\left[R(\hat{y}_j^{(i)})-R(y_j)\right]\right\}^2=\sum_{j=1}^{n}\left[a_i\varepsilon_j^{(i)}\right]^2.$$

用拉格朗日数乘法求解在条件

$$\varphi\equiv 1-\sum_{i=1}^{s}a_i=0$$

下使△达到极小值的a_i（$i=1,\ 2,\ \cdots,\ s$）。令

$$\frac{\partial(\triangle+\lambda\varphi)}{\partial a_i}=0,\qquad i=1,\ 2,\ \cdots,\ s$$

得

$$\sum_{j=1}^{n}2\left(\sum_{k=1}^{s}a_k\varepsilon_j^{(k)}\right)\cdot\varepsilon_j^{(i)}-\lambda=0,\ \ i=1,\ 2,\ \cdots,\ s$$

记$b_{ik}=\sum_{j=1}^{n}\varepsilon_j^{(i)}\cdot\varepsilon_j^{(k)}$（$i,\ k=1,\ 2,\cdots,s$），$\lambda'=\frac{\lambda}{2}$，则得方程组

$$\begin{cases}\sum_{k=1}^{s}a_kb_{ik}=\lambda',\ i=1,\ 2,\ \cdots,\ s,\\ \sum_{k=1}^{s}a_k=1\end{cases}\tag{3}$$

将前面s个方程写成矩形阵形式，即

$$\mathbf{BA}=\mathbf{D}\tag{4}$$

其中

$$\mathbf{A}=\begin{pmatrix}a_1\\a_2\\\vdots\\a_s\end{pmatrix},\quad \mathbf{B}=\begin{pmatrix}b_{11}&b_{12}&\cdots&b_{1s}\\b_{21}&b_{22}&\cdots&b_{2s}\\\vdots&\vdots&\vdots&\vdots\\b_{1s}&b_{2s}&\cdots&b_{ss}\end{pmatrix},\quad \mathbf{D}=\begin{pmatrix}\lambda'\\\lambda'\\\vdots\\\lambda'\end{pmatrix}.$$

显然B是对称矩阵。又由于

$$\triangle=\sum_{j=1}^{n}\left[\sum_{i=1}^{s}a_i\cdot\varepsilon_j^{(i)}\right]^2=\sum_{j=1}^{n}\left(\sum_{i=1}^{s}\sum_{k=1}^{s}a_i a_k\varepsilon_j^{(i)}\varepsilon_j^{(k)}\right)=\sum_{i=1}^{s}\sum_{k=1}^{s}b_{ik}a_ia_k>0,$$ 若记

$f(a_1,a_2,\cdots,a_s)=\sum_{i=1}^{s}\sum_{k=1}^{s}a_ia_kb_{ik}$，则它是正定二次型，故其系数矩阵的主子行列式均为正值，因此方程组(4)有唯一解：$\mathbf{A}=\mathbf{B}^{-1}\mathbf{D}$．记$\mathbf{B}^{-1}=[C_{ij}]_{s\times s}$，则

$$a_i=\lambda'\sum_{k=1}^{s}C_{ik},\quad i=1,2,\cdots,s.$$

由条件$\sum_{i=1}^{s}a_i=1$便得方程组(3)的唯一解

$$\hat{\lambda}'=\frac{1}{\sum_{l=1}^{s}\sum_{k=1}^{s}C_{lk}},\quad \hat{a}_i=\frac{\sum_{k=1}^{s}C_{ik}}{\sum_{l=1}^{s}\sum_{k=1}^{s}C_{lk}},\quad i=1,\ 2,\ \cdots,\ s \tag{5}$$

可见在最小二乘意义下按式(5)定系数的集成预报方程(3)是最优的。特别，若在式(2)中取$a_i=1$，$a_t=0$（$t\neq i$，$1\leqslant t\leqslant s$），则有

$$R(\hat{y})=R(\hat{y}^{(i)})$$
$$\triangle=Q^{(i)}.$$

由上面的证明可知它不是集成预报方程误差平方和的最小值；换句话说，按式(5)定系数的集成预报方程(2)之误差平方和$\hat{\triangle}$必满足

$$\hat{\triangle}<Q^{(i)},\quad i=1,\ 2,\ \cdots s.$$

因此由式(5)确定系数$\hat{a}_i$（$i=1,2,\cdots,s$）的集成预报方程

$$R(\hat{y})=\sum_{i=1}^{s}\hat{a}_iR(\hat{y}^{(i)}) \tag{6}$$

比以秩和为因子的单个预报方程优越，当然更比用通常的逐步回归法建立的预报方程优越。

6、求y的预报值及回报值

以诸因子的当前值$x_{ko}^{(i)}$（$k=1,2,\cdots,m_i$）代入(1)即可求得预报对象按第i个回归方程的预报值之秩$R(\hat{y}_0^{(i)})$（$i=1,2,\cdots,s$），将它们代入式(6)即求得y的集成预报值之秩：

$$R(\hat{y}_0)=\sum_{i=1}^{s}\hat{a}_i R(\hat{y}_0^{(i)}).$$

最后再将$R(\hat{y}_0)$与诸$R(y_j)$（$j=1,2,\cdots,n$）进行比较，以最接近的那个历史数据(比如$R(\hat{y}_0)\approx R(y_{j0})$)作为$y$的集成预报值$\hat{y}_0$(即$\hat{y}_0\approx y_{j0}$)。

类似地，以诸因子的历史数据$x_{kj}^{(i)}$（$j=1,2,\cdots,n$；$i=1,2,\cdots,s$；$k=1,2,\cdots,m_i$）代入式(1)，求出y的各历史回报值之秩$R(\hat{y}_j^{(i)})$（$i=1,2,\cdots,s$；$j=1,2,\cdots,n$），再代入集成预报方程(6)，即可得出y的集成回报值之秩，最后用与上面类似的方法便可求出y的集成回报值$\hat{y}_j$（$j=1,2,\cdots,n$），由此便可求出集成预报的误差平方和$\hat{\triangle}$(按秩)及$\hat{\triangle}_1$(按值)。

三、预报实例及效果比较

实例：根据福建省晋江地区气象台1959——1983年共24年的资料，预报福建省泉州市1984年2月下旬降雨量。

预报对象用y_1表示。利用秩相关系数检验法选出45个因子，它们与y_1的秩相关系数及线性相关系数见表1。对这些因子作秩变换，並用适当的办法将5个因子为一组，取其秩数之和作为一个新因子。再对新因子和y_1作秩变换求其相关系数，将结果列于表2。从表2可见新因子与y_1的秩相关系数一般显着地高于原因子跟y_1的线性相关系数。

先后在四种不同信度下用逐步回归法分别建立起新因子与y_1之秩的四个预报方程，然后按式(5)、(6)建立秩集成预报方程，结果见表3。表中的$Q(i)$、R_i分别表示第i个预报方程的误差平方和及复相关系数（$i=1,2,3,4$）$\hat{\triangle}$、R分别表示秩集成预报方程的误差平方和及复相关系数。显然$\hat{\triangle}<Q(i)$（$i=1,2,3,4$），$R>R_i$（$i=1,2,3,4$），说明集成预报方程比单个逐步回归方程为好。集成预报y_1的1984年之秩数为7.9，相当于y_1之值为6.2mm，而实际上1984年y_1之值为4.1mm，预报误差仅2.1mm。

以新因子的历史数据代入第i个预报方程，求出y_1的各历史回报秩数$R(\hat{y}_{1j}^{(i)})$($j=1,2,\cdots,24$)，再求出对应的y_1的各历史回报数值$\hat{y}_{1j}^{(i)}$（$j=1,2,\cdots 24$），其相应的误差平方和及复相关系数分别用$Q_1(i)$、R_i'表示($i=1,2,3,4$)，列入表3。按秩集成预报方程求出y_1的集成回报秩数$R(\hat{y}_{1j})$（$j=1,2,\cdots,24$），再求出对应的y_1的集成回报值$\hat{y}_{1j}$（$j=1,2,\cdots,24$），将后者列入表5，其误差平方和及复相关系数分别用$\hat{\triangle}_1$、R'表示。由于y_1历史数据波动的不均匀性，有的集成回报秩数与历史实测y_1之秩尽管相差不大，可是其相应的集成回报数值却与历史实测y_1之数值有很大的差异，从而增大了误差平方和，影响了回报准确率。为此，我们以前面已求出的y_1的各历史回报数值$\hat{y}_{1j}^{(i)}$（$j=1,2,\cdots,24$；$i=1,2,3,4$），直接代入下述集成预报方程，(称之为以秩和为因子的值集成预报方程)

$$\hat{y}_{1j}'=0.4172\,\hat{y}_{1j}^{(1)}+0.2689\,\hat{y}_{1j}^{(2)}+0.1141\,\hat{y}_{1j}^{(3)}+0.1998\,\hat{y}_{1j}^{(4)},(j=1,2,\cdots,24)$$求出

y_1的集成回报值y'_{1j}（$j=1, 2, \cdots, 24$），其误差平方和更小，回报准确率(按回报误差小于y_1历史数据极差的15%，即8.835mm为衡量标准)也有所提高(表6)。

另一方面，为了对比，直接取上述45个待选因子(不作秩变换)作逐步回归集成，计算结果列入表4(从表中可看出集成预报方程比单个回归方程优越)，将回报情况列入表5。从表3、4、5显而易见，无论从集成预报方程的误差平方和、复相关系数、预报效果或回报准确率来看，本文的方法都比文[1]的方法好(注意：在表5中以回报误差小于历史观测值之极差的10%即5.89mm作为衡量回报准确的标准)。

表 1

因　子	秩相关系数	相关系数	因　子	秩相关系数	相关系数
x_{44}	0.489	0.509	x_{437}	0.386	0.322
x_{63}	-0.417	-0.370	x_{439}	-0.375	-0.333
x_{73}	0.400	0.374	x_{450}	0.386	0.364
x_{92}	0.436	0.277	x_{454}	-0.408	-0.348
x_{142}	-0.491	-0.345	x_{466}	-0.509	-0.526
x_{159}	0.479	0.504	x_{478}	-0.479	-0.480
x_{162}	-0.485	-0.529	x_{485}	-0.342	-0.250
x_{174}	0.349	0.464	x_{492}	-0.358	-0.332
x_{222}	-0.372	-0.247	x_{498}	-0.411	-0.234
x_{233}	0.423	0.454	x_{500}	-0.413	-0.464
x_{236}	0.493	0.492	x_{536}	-0.534	-0.568
x_{264}	-0.340	-0.323	x_{553}	0.439	0.306
x_{269}	-0.421	-0.485	x_{554}	0.380	0.306
x_{271}	-0.399	-0.355	x_{556}	-0.393	-0.559
x_{284}	0.352	0.265	x_{560}	0.394	0.557
x_{310}	0.363	0.366	x_{561}	0.387	0.404
x_{347}	0.404	0.352	x_{562}	-0.422	-0.363
x_{354}	0.398	0.427	x_{563}	-0.565	-0.595
x_{355}	-0.370	-0.458	x_{586}	0.502	0.429
x_{369}	0.364	0.333	x_{608}	-0.344	-0.371
x_{379}	0.345	0.406	x_{624}	-0.393	-0.354
x_{392}	-0.601	-0.564	x_{628}	-0.421	-0.483
x_{404}	-0.581	-0.557			

表 2

新 因 子	相 关 系 数
x'_1	0.8629
(x_{562}, x_{563}, x_{404}, x_{561}, x_{553},)	(-0.36, -0.595, -0.56, 0.404, 0.306)
x'_2	0.8118
(x_{560}, x_{586}, x_{392}, x_{159}, x_{354})	(0.557, 0.429, -0.56, 0.504, 0.427)
x'_3	0.7556
(x_{628}, x_{478}, x_{44}, x_{536}, x_{142})	(-0.483, -0.48, 0.509, -0.568, -0.345)
x'_4	0.8595
(x_{174}, x_{450}, x_{162}, x_{466}, x_{264})	(0.464, 0.364, -0.53, -0.53, -0.32)
x'_5	0.8795
(x_{556}, x_{356}, x_{310}, x_{499}, x_{238})	(-0.56, -0.46, 0.37, -0.23, 0.492)
x'_6	0.8205
(x_{92}, x_{378}, x_{233}, n_{624}, x_{454})	(0.277, 0.406, 0.454, -0.354, -0.348)
x'_7	0.7551
(x_{63}, x_{222}, x_{369}, n_{492}, x_{271})	(-0.37, -0.25, 0.33, -0.33, -0.355)
x'_8	0.8613
(x_{73}, x_{347}, x_{437}, x_{429}, x_{608})	(0.374, 0.35, 0.32, -0.333, -0.371)
x'_9	0.582
(x_{485}, x_{554}, x_{500}, x_{269} x_{284})	(-0.25, 0.306, -0.46, -0.485, 0.265)

表 3

信度	回 归 方 程	误差平方和	复相关系数	预 报 值	实 际 值	误差
$\alpha_1=0.01$	$R(\hat{y}_1^{(1)})=-14.95385$ $+0.21617x'_5+0.22309x'_8$	$Q(1)=149.8$ $Q_1(1)=937.7$	$R_1=0.9325$ $R_1'=0.9438$	$R(\hat{y}_1^{(1)})=4.84$ $\hat{y}_1^{(1)}=1.1$	$R(y_1)=7$ $y_1=4.1$	-2.2 -3.0
$\alpha_2=0.025$	$R(\hat{y}_1^{(2)})=-8.90234$ $+0.16113x'_1+0.18131x'_4$	$Q(2)=190.0$ $Q_1(2)1418$	$R_2=0.9135$ $R_2'=0.9110$	$R(\hat{y}_1^{(2)})=16.8$ $\hat{y}_1^{(2)}=24.1$	$R(y_1)=7$ $y_1=4.1$	9.8 20.0
$\alpha_3=0.05$	$R(\hat{y}_1^{(3)})=-11.39657$ $+0.16025x'_2+0.22210x'_6$	$Q(3)=194.7$ $Q_1(3)=1545$	$R_3=0.9112$ $R_3'=0.9075$	$R(\hat{y}_1^{(3)})=7.6$ $\hat{y}_1^{(3)}=6.2$	$R(y_1)=7$ $y_1=4.1$	0.6 2.1
$\alpha_4=0.07$	$R(\hat{y}_1^{(4)})=-12.49191$ $+0.13190x'_3+0.18125x'_7$ $+0.08672x'_9$	$Q(4)=229.2$ $Q_1(4)=2539$	$R_4=0.8946$ $R_4'=0.8399$	$R(\hat{y}_1^{(4)})=2.6$ $\hat{y}_1^{(4)}=0.0$	$R(y_1)=7$ $y_1=4.1$	-4.4 -4.1
集成预报方程	$R(\hat{y}_1)=0.4172R(\hat{y}_1^{(1)})$ $+0.2689R(\hat{y}_1^{(2)})$ $+0.1141R(\hat{y}_1^{(3)})$ $+0.1998R(\hat{y}_1^{(4)})$ $\hat{y}'_1=0.4172\hat{y}_1^{(1)}$ $+0.2689\hat{y}_1^{(2)}$ $+0.1141\hat{y}_1^{(3)}$ $+0.1998\hat{y}_1^{(4)}$	$\Delta=105.8<Q(i)$, $i=1,2,3,4$ $\Delta_1=855.7<Q_1(i)$, $i=1,2,3,4$	$R=0.9521>R_i$, $i=1,2,3,4$ $R'=0.9557>i_i$, $i=1,2,3,4$	$R(\hat{y}_1)=7.93$ $\hat{y}_1=6.2$ $\hat{y}'_1=7.6$	$R(y_1)=7$ $y_1=4.1$ $y_1=4.1$	0.93 2.1 3.5

表 4

信　度	回　归　方　程	误差平方和	复相关系数	预报值	实际值	误　差
$a_1=0.01$	$\hat{y}_1^{(1)}=46.21484+0.14905x_{560}-0.21676x_{563}$	$Q(1)=3568$	$R_1=0.7529$	38.38	4.1	34.28
$a_2=0.025$	$\hat{y}_1^{(2)}=-35.78223-0.6106x_{536}-0.75349x_{556}+0.32306x_{585}$	$Q(2)=1762$	$R_2=0.8867$	18.11	4.1	14.01
$a_3=0.05$	$\hat{y}_1^{(3)}=-69.76489+6.21575x_{159}+23.29713x_{379}-1.50995x_{392}$	$Q(3)=2858$	$R_3=0.8082$	2.14	4.1	-1.96
$a_4=0.07$	$\hat{y}_1^{(4)}=33.52287-0.74413x_{404}-0.2391x_{562}-0.07863x_{628}$	$Q(4)=2751$	$R_4=0.8161$	8.33	4.1	4.23
$a_5=0.08$	$\hat{y}_1^{(5)}=29.75952-4.54012x_{162}+3.79872x_{174}+2.33609x_{450}$	$Q(5)=3172$	$R_5=0.7842$	38.70	4.1	34.6
$a_6=0.10$	$\hat{y}_1^{(6)}=-72.18899+48.22262x_{354}-2.72975x_{466}+0.25692x_{553}$	$Q(6)=3217$	$R_6=0.7808$	31.11	4.1	27.01
集成预报方程: $\hat{y}_1=-0.0964\hat{y}_1^{(1)}+0.5349\hat{y}_1^{(2)}+0.1372\hat{y}_1^{(3)}+0.1805\hat{y}_1^{(4)}+0.2357\hat{y}_1^{(5)}+0.0081\hat{y}_1^{(6)}$		$\hat{\Delta}=1164<Q(i)$, $i=1,2,\cdots,6$	$R=0.9426>R_i$, $i=1,2,\cdots,6$,	17.16	4.1	13.06

表 5

年份	用本文所述方法				用普通集成回归预报法			
	y_1 实测值	回报值 $\hat{y}_1$	$\hat{y}_1-y_1$	评定	y_1 实测值	回报值 $\hat{y}_1$	$\hat{y}_1-y_1$	评定
1960	0.0	0.0	0.0	✓	0.0	-2.53	-2.53	✓
61	21.9	21.9	0.0	✓	21.9	21.57	-0.33	✓
62	17.5	21.0	3.5	✓	17.5	21.64	4.14	✓
63	0.9	6.2	5.3	✓	0.9	19.16	18.26	×
64	41.9	40.0	1.9	✓	41.9	37.71	-4.19	✓
65	5.4	6.2	0.8	✓	5.4	12.27	6.87	×
66	58.9	46.1	-12.8	×	58.9	53.05	-5.85	✓
67	17.5	21.0	3.5	✓	17.5	15.98	-1.52	✓
68	40.0	24.1	-15.9	×	40.0	33.39	-6.61	×
69	55.9	58.9	3.0	✓	55.9	43.66	-12.24	×
70	6.2	5.4	-0.8	✓	6.2	12.73	6.53	×
71	22.8	22.8	0.0	✓	22.8	27.58	4.78	✓
72	24.1	21.9	-2.2	✓	24.1	13.49	-10.61	×
73	6.5	21.0	14.5	×	6.5	11.12	4.62	✓
74	41.6	38.3	-3.3	✓	41.6	35.20	-6.40	×
75	0.0	1.4	1.4	✓	0.0	3.27	3.27	✓
76	21.0	21.0	0.0	✓	21.0	27.08	6.08	×
77	0.0	0.0	0.0	✓	0.0	2.44	2.44	✓
78	1.1	5.4	4.3	✓	1.1	6.85	5.75	✓
79	1.4	1.4	0.0	✓	1.4	7.54	6.14	×
80	38.3	21.0	-17.3	×	38.3	31.80	-6.50	×
81	21.8	21.8	0.0	✓	21.8	13.55	-8.25	×
82	6.3	6.5	0.2	✓	6.3	8.27	1.97	✓
83	46.1	41.6	-4.5	✓	46.1	40.28	-5.82	✓
预　报 1984年	4.1 (实际值)	6.2 (预报值)	2.1 (误差)	回报率 83.3%	4.1 (实际值)	17.156 (预报值)	13.056 (误差)	回报率 54.2%

表 6

	按秩集成预报方程先求$R(\hat{y}_1)$再求$\hat{y}_1$						按以秩和为因子的值集成预报方程求$\hat{y}_1$			
年　份	y_1的实测值	秩$R(y_1)$	回报秩$R(\hat{y}_1)$	回报值$\hat{y}_1$	$\hat{y}_1-y_1$	评定	y_1实测值	集成回报值$\hat{y}_1$	$\hat{y}_1-y_1$	评定
1960	0.0	2	0.8952	0.0	0.0	✓	0.0	0.1798	0.1798	✓
61	21.9	15	14.5294	21.9	0.0	✓	21.9	21.8436	-0.0564	✓
62	17.5	11.5	12.9226	21.0	3.5	✓	17.5	18.5241	1.0241	✓
63	0.9	4	8.1868	6.2	5.3	✓	0.9	5.9583	5.0583	✓
64	41.9	21	19.4766	40.0	1.9	✓	41.9	40.2575	-1.6425	✓
65	5.4	7	7.9604	6.2	0.8	✓	5.4	6.1087	0.7087	✓
66	58.9	24	22.2751	46.1	-12.8	×	58.9	50.7503	-8.1497	✓
67	17.5	11.5	12.3641	21.0	3.5	✓	17.5	13.2196	-4.2804	✓
68	40.0	19	16.8751	24.1	-15.9	×	40.0	28.9981	-11.0019	×
69	55.9	23	25.3612	58.9	3.0	✓	55.9	58.9	3.0000	✓
70	6.2	8	7.4997	5.4	-0.8	✓	6.2	4.1705	-2.0295	✓
71	22.8	16	15.6987	22.8	0.0	✓	22.8	26.1639	3.3639	✓
72	24.1	17	15.3712	21.9	-2.2	✓	24.1	25.5817	1.4817	✓
73	6.5	10	12.6342	21.0	14.5	×	6.5	18.1272	11.6272	×
74	41.6	20	17.5483	38.3	-3.3	✓	41.6	35.6552	-5.9448	✓
75	0.0	2	6.1818	1.4	1.4	✓	0.0	2.8071	2.8071	✓
76	21.0	13	12.9112	21.0	0.0	✓	21.0	17.1946	-3.8054	✓
77	0.0	2	1.2639	0.0	0.0	✓	0.0	0.0000	0.0000	✓
78	1.1	5	6.7613	5.4	4.3	✓	1.1	6.4561	5.3561	✓
79	1.4	6	6.0770	1.4	0.0	✓	1.4	2.9343	1.5343	✓
80	38.3	18	12.7453	21.0	-17.3	×	38.3	19.7202	-18.5798	×
81	21.8	14	14.2405	21.8	0.0	✓	21.8	18.6971	-3.1029	✓
82	6.3	9	9.7742	6.5	0.2	✓	6.3	7.6954	1.3954	✓
83	46.1	22	20.4461	41.6	-4.5	✓	46.1	42.5295	-3.5705	✓
预报1984年	4.1(实际值)	7(实际秩数)	7.93(预报秩数)	6.2(预报值)	2.1(误差)	回报率83.3%	4.1(实际值)	7.6478(预报值)	3.5478(误差)	回报率87.5%
	误差平方和 $\Delta=1049$，复相关系数 $R'=0.9541$						误差平方和 $\Delta=855.7$ 复相关系数$R'=0.9557$			

参 考 文 献

〔1〕吴绍敏、陈治典，关于分布自由的回归预报法，待发表.
〔2〕W.J.Conover, Practical Nonparamatric Statistics, Second Edition, (1980).

A Composite Prediction Method of Stepwise Regression with Freedom of Distribution

Wu Shaumin Chen Zhidian

Abstract

As a continuation of paper [1], this paper presents an optimal predicting equation which is derived from several stepwise repression equations. These equation, with rank sums as their factors, are established at different confidence level.

As compared with paper[1], this predicting equation possesses a higher precision and reliability in prediction, for it composes more informations arising from selected factors. It is better than the common composite optimal equation of stepwise regression and the single stepwise regression equation with rank sums as its factors.