分型法和分段法

以预报经验为背景配合逐步廻归的两个统计预报方法

地理系气象教研室　李德成

摘　要

在使用逐步迴归預报台风路徑的問題上，如何定量地描写預报員的分型和分段經驗，本文就此提出了一种分型和分段的方法。分型法試驗表明，文中所提出的分型因子常被收入迴归方程中；分段法指出了用文中所提出的分段因子可克服分型因子的某种缺点，但該方法尚有待于实践檢驗。

第一部分　分型经验和分型法

一、问题

預报員有把不同对象和不同情况进行分型的丰富經驗。例如，根据預报員經驗，南海台风和西太平洋台风其移行規律有所不同；付高第二次北跳前与付高第二次北跳后，台风的移行規律有所不同，等等。本文第一部分是在使用逐步迴归的前提下，提出一种定量地描写預报員分型經驗的方法。

二、分型数据矩阵

方法的思路：把所得到的样本看成是多种类型的样本，然后檢驗各类型是否有显著差异，以达到定量地描写預极員分型經驗的目的。

为了使方法的气象涵义比較明显，我們采用举例的方式进行討論。設南海台风为“1”型；太平洋台风为“2”型；付高第二次北跳前为“3”型；付高第二次北跳后为“4”型。又設x为預极因子，y为預极量的实测値。x_{it}、y_t、i为因子編号，t为样本編号。为了书写統一起見，令$x_{0t}\equiv 1$，并把x_0理解为常因子，$i=0,1,2,\cdots\cdots,n$；$t=1,2,\cdots\cdots,m$。

构造新因子，令

$$V_{it}^{k}=\begin{cases}x_{it} & \text{当第 t 組样本属于第}K\text{型时,}\\ 0 & \text{当第 t 組样本不属于第}K\text{型时,}\end{cases}$$

其中K为型的编号，$K=1,2,3,4$；$i=0,1,2,\cdots\cdots,n$；$t=1,2,\cdots\cdots,m$。当$i=0$

时，我們称 $\overset{K}{V_o}$ 为分型因子。令

$$X=\begin{pmatrix} x_{01} & x_{11} & \cdots\cdots & x_{n1} \\ x_{02} & x_{12} & \cdots\cdots & x_{n2} \\ \vdots & \vdots & & \vdots \\ x_{0m} & x_{1m} & \cdots\cdots & x_{nm} \end{pmatrix}$$

$$Y=\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix}$$

$$W^K=\begin{pmatrix} V^k_{01} & V^k_{11} & \cdots\cdots & V^k_{n1} \\ V^k_{02} & V^k_{12} & \cdots\cdots & V^k_{n2} \\ \vdots & \vdots & & \vdots \\ V^k_{0m} & V^k_{1m} & \cdots\cdots & V^k_{nm} \end{pmatrix}$$

其中 $K=1,2,3,4$。易知矩陣（X Y）、（W^1 W^2 Y）、（W^3 W^4 Y）分别为不分型、分“1”与“2”两型、分“3”与“4”两型的数据矩陣。最后令

$$U=(X\ W^1\ W^2\ W^3\ W^4\ Y)$$

U即为較一般的分型数据矩陣。容易察覺，在矩陣U中，預报因子存在以下的关系

$$\left.\begin{aligned} \overset{1}{V_i}+\overset{2}{V_i}-x_i=0 \\ \overset{3}{V_i}+\overset{4}{V_i}-x_i=0 \end{aligned}\right\}\quad (i=0,1,2,\cdots,n) \quad\cdots\cdots\cdots\cdots\cdots\cdots\cdots \quad (1)$$

（1）式表示預报因子之間的綫性相关，因此，欲得到与矩陣U相对应的全迴归解，必須采用无用参数方法。然而，在逐步迴归中，不存在采用无用参数方法的問題。因为，当吸收了 $\overset{1}{V_i}$，$\overset{2}{V_i}$，x_i（或 $\overset{3}{V_i}$，$\overset{4}{V_i}$，x_i）三个因子中某两个因子时，根据无用参数原理，剩下的未被吸收的那个因子，其对方差减小的貢献必定为零。再据根逐步迴归原理，未被吸收的那个因子，不会被吸收了，故不存在使用无用参数方法問題。因其形式証明較易，故証明在此从略。

三、矩陣 U 的解

在一定的显著水平之下，使用逐步迴归，得到与矩陣 U 相对应的解，其中吸收了某些因子，另一些因子未被吸收，从現在起約定：我們始終把解写成全迴归解或类似于全迴归解的形式，只要我們把那些未被吸收的因子的迴归系数理解为零就行了。經过統計檢驗，不外出現以下四种情况之一。

第一种情况，1型与2型，3型与4型有显著差异，相应的解为

$$Y=\sum_{i=0}^{n} b_i x_i+\sum_{i=0}^{n} b_i^{1} x_i^{1}+\sum_{i=0}^{n} b_i^{2} x_i^{2}+\sum_{i=0}^{n} b_i^{3} x_i^{3}+\sum_{i=0}^{n} b_i^{4} x_i^{4} \quad \cdots\cdots\cdots\cdots\cdots (2)$$

其中 Y 为預报值，b_i^k 为迴归系数，$i=0,1,2,\cdots\cdots,n$；$K=1,2,3,4$。把(2)式写成分枝形式：

$$Y=\sum_{i=0}^{n} b_i x_i \begin{cases} +\sum\limits_{i=0}^{n} b_i^{1} v_i^{1} \begin{cases} +\sum\limits_{i=0}^{n} b_i^{3} v_i^{3} \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots L_{13} \\ +\sum\limits_{i=0}^{n} b_i^{4} v_i^{4} \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots L_{14} \end{cases} \\ +\sum\limits_{i=0}^{n} b_i^{2} v_i^{2} \begin{cases} +\sum\limits_{i=0}^{n} b_i^{3} v_i^{3} \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots L_{23} \\ +\sum\limits_{i=0}^{n} b_i^{4} v_i^{4} \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots L_{24} \end{cases} \end{cases}$$

第二次付高北跳前，預报南海台风，用方程 L_{13}；
第二次付高北跳后，預报南海台风，用方程 L_{14}；
第二次付高北跳前，預报太平洋台风，用方程 L_{23}；
第二次付高北跳后，預报太平洋台风，用方程 L_{24}；

前面所說“1型与2型，3型与4型有显著差异”的意义現在可說明一下了。（一）如果有某个 $b_i \neq 0$，則，①只要 b_i^1 与 b_i^2 中有一个不等于零，即表示1型与2型有显著差异；②只要 b_i^3 与 b_i^4 中有一个不等于零，即表示3型与4型有显著差异。（二）如果有某个 $b_i=0$，則，①只要 b_i^1 与 b_i^2 有显著差异，即表示1型与2型有显著差异；②只要 b_i^3 和 b_i^4 有显著差异，即表示3型与4型有显著差异。

第二种情况，1型与2型有显著差异，而3型与4型无显著差异，其相应的解为

$$Y=\sum_{i=0}^{n} b_i x_i+\sum_{i=0}^{n} b_i^{1} v_i^{1}+\sum_{i=0}^{n} b_i^{2} v_i^{2} \quad \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots (3)$$

把（3）式写成分枝形式

$$
Y=\sum_{i=0}^{n} b_i x_i \begin{cases} +\sum_{i=0}^{n} \overset{1}{b}_i \overset{1}{v}_i \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots L_1 \\ +\sum_{i=0}^{n} \overset{2}{b}_i \overset{2}{v}_i \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots L_2 \end{cases}
$$

預报南海台风，用方程 L_1；

預报太平洋台风，用方程L_2。

第三种情况，3型与4型有显著差异，而1型与2型无显著差异，其相应的解为

$$
Y=\sum_{i=0}^{n} b_i x_i+\sum_{i=0}^{n} \overset{3}{b}_i \overset{3}{v}_i+\sum_{i=0}^{n} \overset{4}{b}_i \overset{4}{v}_i \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots (4)
$$

把（4）式写成分枝形式

$$
Y=\sum_{i=0}^{n} b_i x_i \begin{cases} +\sum_{i=0}^{n} \overset{3}{b}_i \overset{3}{v}_i \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots L_3 \\ +\sum_{i=0}^{n} \overset{4}{b}_i \overset{4}{v}_i \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots L_4 \end{cases}
$$

付高第二次北跳前，用方程 L_3；

付高第二次北跳后，用方程 L_4。

第四种情况，1型与2型，3型与4型都无显著差异，其相应的解为

$$
Y=\sum_{i=0}^{n} b_i x_i \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots (5)
$$

不管是什么类型，都用（5）式作預报。

关于矩陣U的解的四种情况的討論至此結束。关于矩陣U在应用上是否有实际意义，我們将在第三部分討論这个問題。下面討論一个特殊的但較重要的情况。

約定：把分型因子看作参变数，不看作自变量。

假定：各类型所对应的超迴归平面相互平行，則矩陣U退化为

$$
\acute{U}=\begin{pmatrix}
x_{01} & x_{11} & \cdots\cdots & x_{n1} & \overset{1}{v}_{01} & \overset{2}{v}_{01} & \overset{3}{v}_{01} & \overset{4}{v}_{01} & y_1 \\
x_{02} & x_{12} & \cdots\cdots & x_{n2} & \overset{1}{v}_{02} & \overset{2}{v}_{02} & \overset{3}{v}_{02} & \overset{4}{v}_{02} & y_2 \\
\vdots & \vdots & & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
x_{0m} & x_{1m} & \cdots\cdots & x_{nm} & \overset{1}{v}_{0m} & \overset{2}{v}_{0m} & \overset{3}{v}_{0m} & \overset{4}{v}_{0m} & y_m
\end{pmatrix}
$$

此时，V_0^k（k＝1，2，3，4）起分型作用。形式为 V_0 的因子在气象統計預报方面应用較广，其名称尙未取得統一，有离散因子，假因子、（1，0）因子等名称，下面我們再从另一角度討論在气象統計預报中引进因子V_0的思路。

若不管什么类型，都用方程

$$Y=\sum_{i=0}^{n}b_i x_i \quad\cdots\cdots\cdots\cdots \quad (5)$$

作預报，则不能反应出各类型之間的差异，这不符合預报員的实践經驗。为了使方程（5）能反应出各类型之間的差异，最簡单的办法是引进一个因子A，建立方程

$$Y=\sum_{i=0}^{n}b_i x_i+BA \quad\cdots\cdots\cdots\cdots \quad (6)$$

其中B为迴归系数，而

$$A=\begin{cases} A_1 & \text{当第二次付高北跳前出現南海台风时，記作〔1〕型，}\\ A_2 & \text{当第二次付高北跳后出現南海台风时，記作〔2〕型，}\\ A_3 & \text{当第二次付高北跳前出現太平洋台风时，記作〔3〕型，}\\ A_4 & \text{当第二次付高北跳后出現太平洋台风时，記作〔4〕型，}\end{cases}$$

其中 A_1、A_2、A_3、A_4 为四个待定常数。現在的問題是如何选擇 A_1，A_2，A_3，A_4之数值，使

$$Q=\sum_{t=1}^{m}\left[y_t-\left(\sum_{i=0}^{n}b_i x_i+BA_t\right)\right]^2 \longrightarrow \text{最小}$$

有两条可达到目的道路：（一）采用非綫性模型中参数确定的方法，但，这种办法的計算量很大，不宜采取；（二）采用最小二乘法，但，最小二乘法只能估計b_0，b_1，……b_n，B，而无法估計处于变量地位的 A_1、A_2、A_3、A_4 之值。为了使最小二乘法也能估計 A_1，A_2，A_3，A_4 之值，必須把 A_1，A_3，A_3，A_4 从处于变量的地位轉移到处于迴归系数的地位。为此，作变换，引进因子：

$$V_0^{(k)}=\begin{cases}1 & \text{当出現第〔k〕型时，}\\ 0 & \text{当不出現第〔k〕型时，}\end{cases}$$

其中k＝1,2,3,4。于是可把（6）式改写为：

$$Y=\sum_{i=0}^{n}b_i x_i+\sum_{i=1}^{4}BA_i V_0^{(i)}$$

令 $b^{(i)}=BA_i$，i＝1,2,3,4，则上式又可写为：

$$Y=\sum_{i=0}^{n}b_i x_i+\sum_{i=1}^{4}b^{(i)}\ V_0^{(i)}$$

于是估計A_1, A_2, A_3, A_4的数値問題轉化为估計$b^{(1)}$，$b^{(2)}$，$b^{(3)}$，$b^{(4)}$的問題，而$b^{(1)}$，$b^{(2)}$，$b^{(3)}$，$b^{(4)}$可用最小二乘法估計。$V_0^{(k)}$ $(k=1,2,3,4)$就是分型因子(或称假因子等)，它对各型的預报値起修正作用。

第二部分　分段经验和分段法

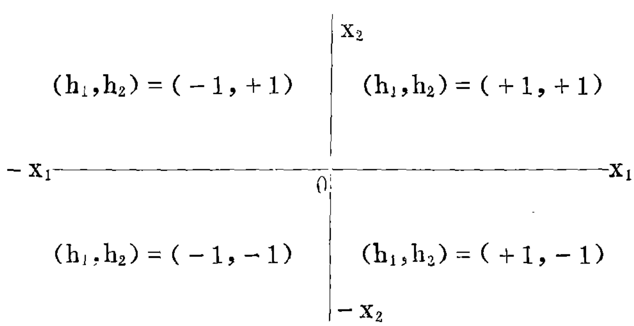
一、问题

預报員針对預报对象，有把某些預报因子分为两段考虑的丰富的分段經驗。例如，初始位置大于18°N的台风与初始位置小于18°N的台风，其移行規律有所不同；初始位置大于120°E的台风与初始位置小于120°E的台风，其移行規律有所不同，等等。本文第二部分就是在使用逐步迴归的前提下，提出一种定量地描写預报員分段經驗的方法。

二、分段数据矩阵

为了定量地描写預报員的分段經驗，我們先討論一个数学問題。如果我們只要求对每个自变量x_i，$(i=1,2,\cdots\cdots,n)$，在原点分为两段，即右段$x_i>0$，左段$x_i<0$，那末，如何把曲綫分段拟合方法，推广到n个自变量的情形？下面着手討論这个問題。

設在$n+1$維空間中的$n+1$个正交軸$-\overline{x_1ox_1}$，……，$-\overline{x_nox_n}$，$-\overline{yoy}$交于0点，其中n个正交軸$-\overline{x_1ox_1}$，……，$-\overline{x_nox_n}$把超平面$y=0$分为2^n个n維子空間（或說分为2^n个n維的子平面)。我們用序列$(h_1,\cdots\cdots,h_j,\cdots\cdots,h_n)$表示这$2^n$个子空間，其中$h_j=+1$或$-1$，$j=1,2,\cdots\cdots,n$。符号$(h_1,\cdots\cdots,h_j,\cdots\cdots h_n)$的意义見示意图。示意图是对应于$n+1=3$，



示 意 图

即$n=2$的情形，为使图形簡明，图中未画出$-\overline{yoy}$軸。符号$M(h_1,\cdots\cdots,h_j,\cdots\cdots h_n)$表示点M属于子空間$(h_1,\cdots\cdots,h_j,\cdots\cdots,h_n)$。$x_i(h_1,\cdots\cdots,h_j,\cdots\cdots,h_n)$表示第$i$个坐标值为$x_i$的那个点M属于子空間$(h_1,\cdots\cdots,h_j\cdots\cdots,h_n)$。令

$$y(h_1,\cdots\cdots,h_j,\cdots\cdots,h_n)=b_0(h_1,\cdots\cdots,h_j,\cdots\cdots,h_n)+$$
$$+b_1(h_1,\cdots\cdots,h_j,\cdots\cdots,h_n)x_1(h_1,\cdots\cdots,h_j\cdots\cdots,h_n)+$$
$$+\cdots\cdots+b_n(h_1,\cdots\cdots,h_j,\cdots\cdots,h_n)x_n(h_1,\cdots\cdots,h_j,+\cdots\cdots,h_n)\cdots\cdots \quad (1)$$

其中$b_i(h_1,\cdots\cdots,h_j,\cdots\cdots,h_n)$在子空間$(h_1,\cdots\cdots,h_j,\cdots\cdots,h_n)$上是某常数，（1）式为定义在子空間$(h_1,\cdots\cdots,h_j,\cdots\cdots,h_n)$上的綫性函数，或說是$n+1$維空間中的超平面，这种超平面共有$2^n$块。定义函数$\tilde{\tilde{y}}$：

$$\tilde{\tilde{y}}=\begin{cases} y(1,\cdots\cdots,1,\cdots\cdots,1) & \text{当 } M\epsilon(1,\cdots\cdots,1,\cdots\cdots,1)\text{ 时} \\ \cdots\cdots\cdots\cdots & \\ y(h_1,\cdots\cdots,h_j,\cdots\cdots,h_n) & \text{当 } M\epsilon(h_1,\cdots\cdots,h_j,\cdots\cdots,h_n)\text{ 时} \\ \cdots\cdots\cdots\cdots & \\ y(-1,\cdots\cdots,-1,\cdots\cdots,-1) & \text{当 } M\epsilon(-1,\cdots\cdots,-1,\cdots\cdots,-1)\text{ 时} \end{cases} \quad (2)$$

若（1）式所定义的2^n个函数$y(h_1,\cdots\cdots,h_j,\cdots\cdots,h_n)$能在边界$x_1=o$，……，边界$x_n=o$处两两滿足銜接条件（这里仅指連續条件），則（2）式就成为$n+1$維空間中的超折面。在边界处，对函数$\tilde{\tilde{y}}$的値作如下規定：令

$$y(h_1,\cdots\cdots,h_{j-1},-1,h_{j+1},\cdots\cdots,h_n)\Big|_{x_j\to-0}\equiv y(h_1,\cdots\cdots,h_{j-1},+1,$$
$$h_{j+1},\cdots\cdots,h_n)\Big|_{x_j\to+0}\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \quad (3)$$

其中$j=1,2,\cdots\cdots,n$。（3）式两边的极限値就作为函数$\tilde{\tilde{y}}$在$x_j=o$处的数値。在这种規定下，由几何知識不难知道，（3）式就是連續条件。下面我們討論一条性质。

〔性质〕 把（2）式改写成下面的（4）式

$$\tilde{\tilde{y}}=b_0^{(0)}+\sum_{i=0}^{n}\left[b_j^{(+1)}x_i^{(+1)}+b_i^{(-1)}x_i^{(-1)}\right]\cdots\cdots\cdots\cdots\cdots\cdots \quad (4)$$

的充要条件是連續条件（3）式。其中$b_0^{(0)}$，$b_i^{(+1)}$，$b_i^{(-1)}$为常数，

$$x_i^{(+1)}=\begin{cases} x_i & \text{当}x_i\geqslant0\text{时} \\ 0 & \text{当}x_i<0\text{时}\end{cases}\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \quad (5)$$

$$x_i^{(-1)}=\begin{cases} x_i & \text{当}x_i\leqslant0\text{时} \\ 0 & \text{当}x_i>0\text{时}\end{cases}\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \quad (6)$$

$i=1,2,\cdots\cdots\cdots,n$。

〔证明〕 先证明充分性。把（1）式代入（3）式，要注意（3）式是恒等式，便可

得到

$$b_i(h_1,\cdots\cdots,h_{j-1},-1,h_{j+1},\cdots\cdots,h_n)=$$
$$=b_i(h_1,\cdots\cdots,h_{j-1},+1,h_{j+1},\cdots\cdots,h_n)\cdots\cdots\cdots\cdots\cdots\cdots\cdots\quad(7)$$

$$b_0(h_1,\cdots\cdots,h_{j-1},-1,h_{j+1},\cdots\cdots,h_n)=$$
$$=b_0(h_1,\cdots\cdots,h_{j-1},+1,h_{j+1},\cdots\cdots,h_n)\cdots\cdots\cdots\cdots\cdots\cdots\cdots\quad(8)$$

其中，对于(7)式$i\neq j$；$i,j=1,2,\cdots\cdots,n$。对于(8)式$j=1,2,\cdots\cdots,n$。

現在着手簡化(7)式。为了避免被脚标搞得眼花繚乱，而又不丧失一般性，我們先考察$b_i=b_1,h_1=1$的情形，把它詳細写出来就是

$$(9)\begin{cases} b_1(1,-1,h_3,\cdots\cdots,h_n)=b_1(1,+1,h_3,\cdots\cdots h_n) \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ b_1(1,h_2,\cdots\cdots,h_{j-1}-1,h_{j+1},\cdots\cdots,h_k,\cdots\cdots,h_n)= \\ \quad =b_1(1,h_2\cdots\cdots,h_{j-1},+1,h_{j+1},\cdots\cdots,h_k,\cdots\cdots,h_n)\cdots\cdots\cdots\cdots\cdots\cdots\quad(9)_j \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ b_1(1,h_2,\cdots\cdots,h_j,\cdots\cdots,h_{k-1},-1h_{k+1},\cdots\cdots,h_n)= \\ \quad =b_1(1,h_2,\cdots\cdots,h_j,\cdots\cdots,h_{k-1},+1,h_{k+1},\cdots\cdots h_n)\cdots\cdots\cdots\cdots\cdots\cdots\quad(9)_k \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ b_1(1,h_2,\cdots\cdots;h_{n-1},-1)=b_1(1,h_2,\cdots\cdots,h_{n-1},+1) \end{cases}$$

下面我們証明由(9)式中可得到

$$b_1(1,h_2,\cdots\cdots,h_n)=b_1^{(+1)}\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\quad(10)$$

其中$b_1^{(+1)}$在所有子空$(1,h_2,\cdots\cdots,h_n)$上是常数，或說在$x_1\geqslant 0$处是常数。

証：(1)不难看出，(9)式包含了$b_1(1,h_1,\cdots\cdots,h_n)$，而无一遺漏。因此，为了証明(10)式成立，只須証明(9)式中任意两个等式$(9)_j$和$(9)_k$能够連等就行了。于是問題归結为要証明下式成立。

$$(9)_j\text{ 的某一端}=(9)_k\text{的某一端}\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\quad(11)$$

(2)細察$(9)_j$和$(9)_k$的表达形式，便知只須在

$$(i)h_j=1,h_k=1;(ii)h_j=1,h_k=-1;(iii)h_j=-1,h_k=1;(iv)h_j=-1,h_k=-1$$

四种情况之下，若能証明(11)式成立就行了。

(3)分别把(i)、(ii)、(iii)、(iv)代入$(9)_j$和$(9)_k$进行比較，便知(11)式成立，于是关于(10)式成立証毕。

同理 $$b_1(-1,h_2,\cdots\cdots h_n)=b_1^{(-1)}$$

其中 $b_1^{(-1)}$在$x_1\leqslant 0$处是常数。

同理
$$\left.\begin{array}{l} b_i(h_1,\cdots\cdots,h_{i-1},+1,h_{i+1},\cdots\cdots,h_n)=b_i^{(+1)} \\ b_i(h_1,\cdots\cdots,h_{i-1},-1,h_{i+1},\cdots\cdots,h_n)=b_i^{(-1)} \end{array}\right\}\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \quad (12)$$

其中 $b_i^{(+1)}$ 为在 $x_i\geqslant 0$ 处是常数，$b_i^{(-1)}$ 为在 $x_i\leqslant 0$ 处是常数，$i=1,2,\cdots\cdots\cdots\cdots,n$。

(12)式就是从(7)式出发所得到的最后結果。

現在着手簡化(8)式。簡化(8)式的方法与簡化(7)式的方法是一样的，其不同之处仅在于对(7)式有 $i\neq j$ 的限制，而对于(8)式无此种限制。我們仿造由(7)式得到(12)式的方法，可由(8)式得到

$$\left.\begin{array}{l} b_0(-1,h_2,\cdots\cdots\cdots\cdots\cdots,h_n)=b_0(+1,h_2,\cdots\cdots\cdots\cdots\cdots\cdots,h_n)=c_1 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ b_0(h_1,\cdots h_{j-1},-1,h_{j+1},\cdots,h_n)=b_0(h_1,\cdots,h_{j-1},+1,h_j,\cdots,h_n)=c_j \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ b_0(h_1,\cdots\cdots\cdots\cdots\cdots,h_{n-1},-1)=b_0(h_1,\cdots\cdots\cdots\cdots,h_{n-1},+1)=c_n \end{array}\right\}\cdots\cdots\cdots \quad (13)$$

其中 $c_1\cdots\cdots,c_j,\cdots\cdots,c_n$ 对于任意的 x_i 值是常数，$i=1,2,\cdots\cdots,n$。

我們再一次利用由(7)式得到(12)式的办法，由(13)式可得到

$$b_0(h_1,\cdots\cdots,h_j,\cdots\cdots,h_n)=b_0^{(0)}\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \quad (14)$$

其中 $b_0^{(0)}$ 对于任意的 x_i 值是常数，$i=1,2,\cdots\cdots,n$。(14)就是从(8)式出发所得到的最后結果。

有了(12)式和(14)式之后，可开始着手簡化(2)式。为此，先把(1)式改写为

$$\begin{aligned} y(h_1,\cdots,h_j,\cdots,h_n)=&\ b_0(h_1,\cdots,h_j,\cdots,h_n)+b_1(1,h_2,\cdots,h_n)x_1(1,h_2,\cdots,h_n)+ \\ &+b_1(-1,h_2,\cdots,h_n)x_1(-1,h_2,\cdots,h_n)+\cdots+b_j(h_1,\cdots,h_{j-1}, \\ &+1,h_{j+1},\cdots,h_n)x_j(h_1,\cdots h_{j-1},+1,h_{j+1},\cdots,h_n)+ \\ &+b_j(h_1,\cdots h_{j-1},-1,h_{j+1},\cdots,h_n)x_j(h_1,\cdots,h_{j-1},-1,h_{j+1},\cdots,h_n) \\ &+\cdots\cdots+b_n(h_1,\cdots,h_{n-1},+1)x_n(h_1,\cdots h_{n-1},+1)+ \\ &+b_n(h_1,\cdots,h_{n-1},-1)x_n(h_1,\cdots,h_{n-1},-1)\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \quad (15) \end{aligned}$$

只要我們約定：

$$x_j(h_1,\cdots,h_{j-1},+1,h_{j+1},\cdots,h_n)=$$
$$=\begin{cases} x_j(h_1,\cdots,h_{j-1},+1,h_{j+1},\cdots,h_n) & \text{当 } M\in(h_1,\cdots,h_{j-1},+1,h_{j+1},\cdots,h_n)\text{ 时} \\ 0 & \text{当 } M\overline{\in}(h_1,\cdots,h_{j-1},+1,h_{j+1},\cdots,h_n)\text{时} \end{cases}$$

$$x_j(h_1,\cdots,h_{j-1},-1,h_{j+1},\cdots,h_n)=$$
$$=\begin{cases} x_j(h_1,\cdots,h_{j-1},-1,h_{j+1},\cdots,h_n) & \text{当 } M\in(h_1,\cdots,h_{j-1},-1,h_{j+1},\cdots,h_n)\text{时} \\ 0 & \text{当 } M\overline{\in}(h_1,\cdots,h_{j-1},-1,h_{j+1},\cdots,h_n)\text{时} \end{cases}$$

那末，就可把(1)式改写成与之等价的(15)式。我們利用(12)式、(14)式、(5)式和(6)式，可把(15)式写成簡洁的形式：

$$y(h_1,\cdots,h_j,\cdots,h_n)=b_0^{(0)}+\left[b_1^{(+1)}x_1^{(+1)}+b_1^{(-1)}x_1^{(-1)}\right]+\cdots+$$
$$+\left[b_n^{(+1)}x_n^{(+1)}+b_n^{(-1)}x_n^{(-1)}\right]\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\quad(16)$$

在(16)式中的 $b_1^{(+1)}$, $b_1^{(-1)}$, $\cdots$, $b_n^{(+1)}$, $b_n^{(-1)}$ 已可认为与子空間$(h_1,\cdots,h_j,\cdots,h_n)$无关，所以

$$\widetilde{y}=b_0^{(0)}+\left[b_0^{(+1)}x_1^{(+1)}+b_1^{(-1)}x_1^{(-1)}\right]+\cdots+\left[b_n^{(+1)}x_n^{(+1)}+b_n^{(-1)}x_n^{(-1)}\right]$$

或把上式写成

$$\widetilde{y}=b_0^{(0)}+\sum_{i=1}^{n}\left[b_i^{(+1)}x_i^{(+1)}+b_i^{(-1)}x_i^{(-1)}\right]\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\quad(4)$$

于是关于連續条件(3)式是充分条件已証毕。下面着手証明(3)式也是必要条件。

我們把(4)式写成

$$\widetilde{y}=y(h_1,\cdots,h_i,\cdots,h_n)=b_0^{(0)}+\sum_{i=1}^{n}b_i^{(h_i)}x_i^{(h_i)}$$

其中 $h_i=+1$ 或 -1，$i=1,2,\cdots,n$。取左极限和右极限：

$$\lim_{x_i\to-0}y(h_1,\cdots,h_{i-1},-1,h_{i+1},\cdots,h_n)=$$
$$\lim_{x_i\to-0}\left[b_0^{(0)}+\sum_{j=1}^{i-1}b_j^{(h_j)}x_j^{(h_j)}+b_i^{(-1)}x_i^{(-1)}+\sum_{j=i+1}^{n}b_j^{(h_j)}x_j^{(h_j)}\right]=$$
$$b_0+\sum_{j=1}^{i-1}b_j^{(h_j)}x_j^{(h_j)}+\sum_{j=i+1}^{n}b_j^{(h_j)}x_j^{(h_j)}\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\quad(17)$$

$$\lim_{x_i\to+0}y(h_1,\cdots,h_{i-1},+1,h_{i+1},\cdots,h_n)=$$
$$\lim_{x_i\to+0}\left[b_0^{(0)}+\sum_{j=1}^{i-1}b_j^{(h_j)}x_j^{(h_j)}+b_i^{(+1)}x_i^{(+1)}+\sum_{j=i+1}^{n}b_j^{(h_j)}x_j^{(h_j)}\right]=$$
$$b_0^{(0)}+\sum_{j=1}^{i-1}b_j^{(h_j)}x_j^{(h_j)}+\sum_{j=i+1}^{n}b_j^{(h_j)}x_j^{(h_j)}\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\quad(18)$$

比較(17)式和(18)式的右端便知

(17)式≡(18)式 $\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$ (19)

恒等号表示对任意的 x_j 都成立，其中 $j=1,2,\cdots\cdots\cdots\cdots,n$。或把(19)式表达为

$$y(h_1,\cdots,h_{j-1},-1,h_{j+1},\cdots,h_n)\Big|_{x_j\to-0} \equiv y(h_1,\cdots,h_{j-1},+1,h_{j+1},\cdots,h_n)\Big|_{x_j\to+0} \cdots\cdots (3)$$

于是关于連續条件(3)式是必要条件証毕。至此，(3)式是充要条件証毕，

容易直观地看出$\tilde{y}$具有以下三条几何性质：

① 在$n+1$維空間中，2^n块超平面构成$\tilde{y}$；

② 这2^n块超平面在边界处連續，所以$\tilde{y}$是超折面；

③ 这2^n块超平面在$-\overline{yoy}$軸上的截距都是$b_0^{(0)}$。

我們还可把(4)式稍加推广。設$f_i(x_i)$为x_i的单值連續函数，在分点$x_i=\overset{0}{x}_i$处，$f_i\left(\overset{0}{x}_i\right)=a_i$。令

$$\overset{(+1)}{f_i}(x_i)=\begin{cases} f(x_i)-a_i & 当\ x_i\geqslant\overset{0}{x}_i\ 时, \\ 0 & 当\ x_i<\overset{0}{x}_i\ 时。\end{cases}$$

$$\overset{(-1)}{f_i}(x_i)=\begin{cases} f(x_i)-a_i & 当\ x_i\leqslant\overset{0}{x}_i\ 时, \\ 0 & 当\ x_i>\overset{0}{x}_i\ 时。\end{cases}$$

其中 $i=1,2,\cdots\cdots n$。令

$$\tilde{\tilde{y}}=b_0^{(0)}+\sum_{i=1}^{n}\left[b_i^{(+1)}\overset{(+1)}{f_i}(x_i)+b_i^{(-1)}\overset{(-1)}{f_i}(x_i)\right]\cdots\cdots (20)$$

(20)式便是(4)式的一种推广，在(20)式中，用2^n块超曲面代替了(4)式中的2^n块超平面。构造矩陣：

$$X^{(\ )}=\begin{pmatrix} X_{11}^{(+1)} & X_{11}^{(-1)} & \cdots\cdots & X_{n1}^{(+1)} & X_{n2}^{(-1)} \\ X_{12}^{(+1)} & X_{12}^{(-1)} & \cdots\cdots & X_{n2}^{(+1)} & X_{n2}^{(-1)} \\ \vdots & \vdots & & \vdots & \vdots \\ X_{1m}^{(+1)} & X_{1m}^{(-1)} & \cdots\cdots & X_{nm}^{(+1)} & X_{nm}^{(-1)} \end{pmatrix}$$

为了叙述方便，我們称矩陣

$$S=(X\ X^{(\ )}\ Y)$$

为分段数据矩陣，$x_i^{(+1)}$， $x_i^{(-1)}$为分段因子，其中X和Y与第一部分相同。

三、矩阵S的解

在使用逐步迴归的前提下，举例討論如何把矩陣S用于統計預报。

設x_1，…… x_n为n个預报因子，根据預报員經驗，令x_1为台风初始經度，以120°E为分点；x_2为台风初始緯度，以18°N为分点；x_3为台风中心风速，以35米/秒为分点；而x_4，……，x_n据預报員經驗无須分段。对x_1，x_2，x_3进行加工，令

$$x_1^{(+1)}=\begin{cases}x_1-120 & \text{当 } x_1-120\geqslant 0 \text{ 时}\\ 0 & \text{当 } x_1-120<0 \text{ 时}\end{cases}$$

$$x_1^{(-1)}=\begin{cases}x_1-120 & \text{当 } x_1-120\leqslant 0 \text{ 时}\\ 0 & \text{当 } x_1-120>0 \text{ 时}\end{cases}$$

$$x_2^{(+1)}=\begin{cases}x_2-18 & \text{当 } x_2-18\geqslant 0 \text{ 时}\\ 0 & \text{当 } x_2-18<0 \text{ 时}\end{cases}$$

$$x_2^{(-1)}=\begin{cases}x_2-18 & \text{当 } x_2-18\leqslant 0 \text{ 时}\\ 0 & \text{当 } x_2-18>0 \text{ 时}\end{cases}$$

$$x_3^{(+1)}=\begin{cases}x_3-35 & \text{当 } x_3-35\geqslant 0 \text{ 时}\\ 0 & \text{当 } x_3-35<0 \text{ 时}\end{cases}$$

$$x_3^{(-1)}=\begin{cases}x_3-35 & \text{当 } x_3-35\leqslant 0 \text{ 时}\\ 0 & \text{当 } x_3-35>0 \text{ 时}\end{cases}$$

据預报員經驗不必使用完整的S矩陣，只須使用矩陣

$$\widetilde{S}=(X\ \widetilde{X}^{(\,)}\ Y)$$

其中X，Y与第一部分相同，而

$$\widetilde{X}^{(\,)}=\begin{pmatrix} x_{11}^{(+1)} & x_{11}^{(-1)} & \cdots\cdots & x_{31}^{(+1)} & x_{31}^{(-1)}\\ x_{12}^{(+1)} & x_{12}^{(-1)} & \cdots\cdots & x_{32}^{(+1)} & x_{32}^{(-1)}\\ \vdots & \vdots & & \vdots & \vdots\\ x_{1m}^{(+1)} & x_{1m}^{(-1)} & \cdots\cdots & x_{3m}^{(+1)} & x_{3m}^{(-1)} \end{pmatrix}$$

規定显著水平，若因子x_1，x_2，x_3的分段对方差的减小都有显著貢献，則得到与矩陣$\tilde{S}$相应的解

$$Y=\sum_{i=0}^{n}b_ix_i+\sum_{i=1}^{3}\left[b_i^{(+1)}x_i^{(+1)}+b_i^{(-1)}x_i^{(-1)}\right]$$

把上式写成分枝形式

$Y=\sum_{i=0}^{n}b_ix_i$
- $+b_1^{(+1)}\ x_1^{(+1)}$
  - $+b_2^{(+1)}\ x_2^{(+1)}$
    - $+b_3^{(+1)}\ x_3^{(+1)}$ …… $L_{1,1,1}$
    - $+b_3^{(-1)}\ x_3^{(-1)}$ …… $L_{1,1,-1}$
  - $+b_2^{(-1)}\ x_2^{(-1)}$
    - $+b_3^{(+1)}\ x_3^{(+1)}$ …… $L_{1,-1,1}$
    - $+b_3^{(-1)}\ x_3^{(-1)}$ …… $L_{1,-1,-1}$,
- $+b_1^{(-1)}\ x_1^{(-1)}$
  - $+b_2^{(+1)}\ x_2^{(+1)}$
    - $+b_3^{(+1)}\ x_3^{(+1)}$ …… $L_{-1,1,1}$
    - $+b_3^{(-1)}\ x_3^{(-1)}$ …… $L_{-1,1,-1}$,
  - $+b_2^{(-1)}\ x_2^{(-1)}$
    - $+b_3^{(+1)}\ x_3^{(+1)}$ …… $L_{-1,-1,1}$
    - $+b_3^{(-1)}\ x_3^{(-1)}$ …… $L_{-1,-1,-1}$

当$x_1-120\geqslant0$，$x_2-18\geqslant0$，$x_3-35\geqslant0$时用超平面$L_{1,1,1}$作預报；

当$x_1-120\geqslant0$，$x_2-18\geqslant0$，$x_3-35\leqslant0$时用超平面$L_{1,1,-1}$作預报，等等。

矩陣$\tilde{S}$的解答共有八种情况，以上只討論了一种情况，其它七种情况的討論从略。

第三部分　试验情况和问题讨论

一、预报方程和预报实例

我們取1957年至1969年7月至9月的台风个例作为拟合样本，08^h和20^h混合。警戒区見附图一。采用的是分型法，从分析的情况来看，采用U矩陣进行拟合可能是較好的。但是，由于U矩陣龐大，受客观条件限制而不能采用U矩陣拟合，只好

采用U矩陣进行拟合，即采用分型因子的办法进行拟合。在第一警戒区用分型因子的办法进行了初步試驗，据預报員經驗，以18°N为分点，分成两型，建点了以下一套預报方程。

$$\lambda_{24}=-0.00070+1.00130[\lambda_{12}]+0.87110[\lambda_{12}-\lambda_0]-$$
$$-0.26215[\varphi_{12}-\varphi_0]-0.38242[V_1] \quad \cdots\cdots \quad (1)$$

$$\varphi_{24}=0.00060+0.94610[\phi_{12}]+0.56101[\phi_{12}-\phi_0]+$$
$$+0.04650[\lambda_{12}]-0.06067[\triangle H_{\text{福州}}]-0.00434[P_{12}]-0.01526[\psi_{SN}]$$

$$\lambda_{36}=0.00028+0.99441[\lambda_{12}]+1.63521[\lambda_{12}-\lambda_0]+$$
$$+1.03459[v_2]-0.44967[\phi_{12}-\phi_0] \quad \cdots\cdots \quad (2)$$

$$\varphi_{36}=-0.00052+0.95177[\phi_{12}]+0.91741[\phi_{12}-\phi_0]+0.01214[W_{ff}]-$$
$$-0.11047[\triangle H_{\text{福州}}]+0.01533[\lambda_{12}]-0.03520[\psi_{SN}]$$

$$\lambda_{48}=0.00074+0.98789[\lambda_{12}]+2.37548[\lambda_{12}-\lambda_0]+1.74502[v_2]+$$
$$+0.28629[\triangle H_{\text{汕头}}]+0.11291[u_{\text{汕头},1.5}] \quad \cdots\cdots \quad (3)$$

$$\varphi_{48}=0.03744+1.05228[\phi_{12}]+1.19492[\phi_{12}-\phi_0]+$$
$$+0.01962[W_{ff}]-0.13039[\triangle H_{\text{福州}}]$$

$$\lambda_{60}=0.00075+2.98105[\lambda_{12}]+2.99066[\lambda_{12}-\lambda_0]+2.29769[\lambda_2]+$$
$$+0.29205[\triangle H_{\text{马尼}}]+0.12471[u_{\text{汕头},5}]-0.12915[\triangle H_{\text{福州}}-$$
$$-\triangle H_{\text{鹿冲}}]+0.06992[\psi_{sn}] \quad \cdots\cdots \quad (4)$$

$$\varphi_{60}=0.02969+1.09077[\phi_{12}]+1.35519[\phi_{12}-\phi_0]+0.02698[W_{ff}]$$

对方程中的預报因子說明如下：

$[\lambda_0]$：初始經度。

$[\varphi_0]$：初始緯度。

$[\lambda_{12}]$：12小时后的經度。

$[\phi_{12}]$：12小时后的緯度。

$[P_{12}]$：12小时后的台风中心气压。

$[W_{ff}]$：台风中心风速。

$[v_1]$：当$\phi_{12}\geqslant 18°N$取0，当$\phi_{12}<18°N$取1，此即分型因子。

$[\triangle H_{\text{福州}}]$：福州500mb24小时变高。

$[\triangle H_{\text{汕头}}]$：汕头500mb24小时变离。

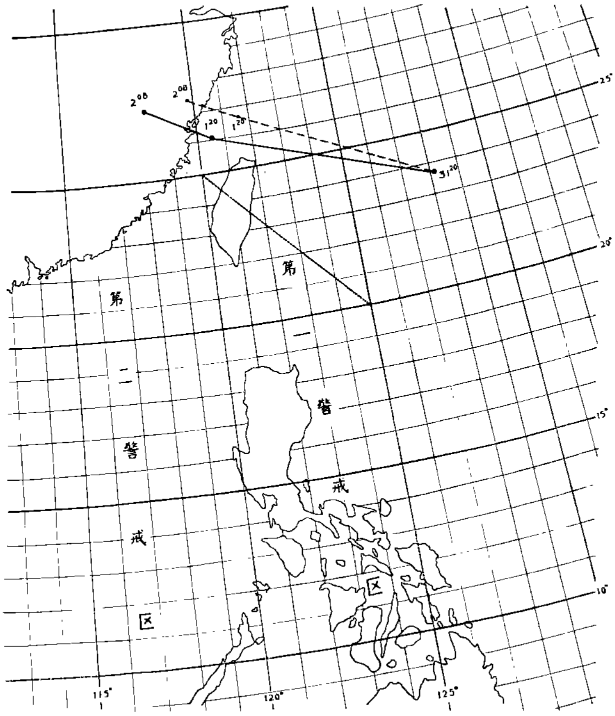
$[\triangle H_{\text{马尼}}]$：馬尼拉500mb24小时变高。

$[\triangle H_{\text{上福}}]$：上海、福州500mb高度和之24小时变高。

$[\triangle_{\text{鹿冲}}]$：鹿儿島、冲绳500mb高度和之24小时变高。

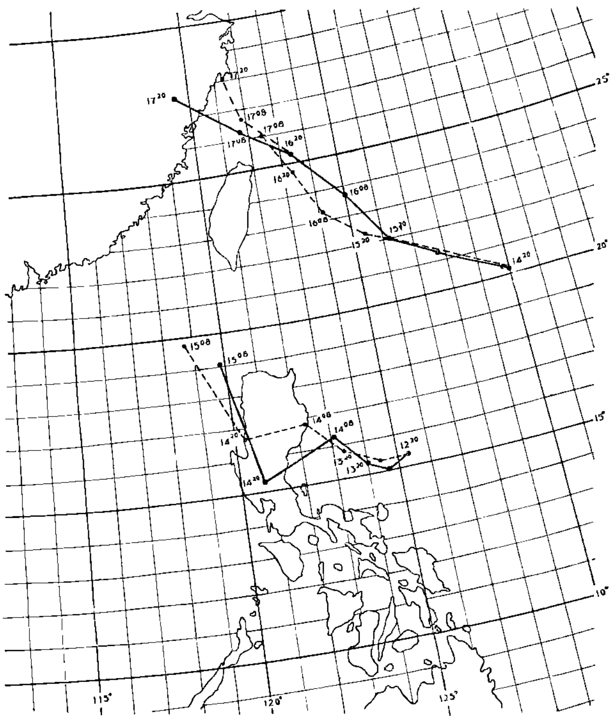
$[U_{\text{汕头},1.5}]$：汕头1500mb高空风之緯向分量。

$[U_{\text{汕头},5}]$：汕头500mb高空风之緯向分量。



（附图一）

实綫为台风实况路徑
虛綫为台风預报路徑



（附图二）

实綫为台风实况路徑

虛綫为台风預报路徑

〔φ_{sN}〕：从第一个西风糟与北緯45°緯綫交点起，沿45°綫向西移动7.5經距得到一点M；再从M点起，向南移动5个緯距处的高度，減去从M点起向北移5动个緯距处的高度。

前述一套預报方程在去年7月下旬建立，此后，至当年9月底止，只遇到72年07号、09号、13号等三个台风进入第一警戒区，对这三个台风作了試报（在实际使用时，台风未进入警戒区就起报），試报情况見附图。附图一是07号台风，附图二之上方为09号台风，下方为13号台风。其預报效果与經驗預报比較，还算令人滿意的。

二、问题讨论

（一）、若分型的标准是非数量性的，例如把南海台风作为一型，把太平洋台风作为另一型，則不存在分点，也就不存在两型的迴归方程在某分点是否連續的問題。若分型的标准是数量性的，例如把台风初始緯度大于18°N作为一型，把小于18°N作为另一型，則用分型法建立的两型方程在分点18°N是不連續的，例如分型方程（1）、（2）、（3）、（4）在分点18°N处其跳跃值（絕对值）分别为0.38242，0.44967，1.74502，2.29769，这是不合理的。为了克服这缺点，可用分段因子代替分型因子，建立分段方程，在离分点較远处用分型方程作預报，在分点邻近用分段方程作預报，这样处理，在分点邻近处的預报效果可能有所提高。

（二）、由于受客观条件限制，我們未对U矩陣进行試驗，所以沒有实例說明U矩陣在实用上是否具有意义。但是，有一点值得討論的，就是用U矩陣拟合，建立的分型方程，其中每一型都要占用其它型的資料，在目前气象資料較少的情况下，U矩陣在一定程度上克服了需要分型而又感到資料不多的困难，这可能是U矩陣的一个主要优点。

（三）、分段法实质上是提供了一种非綫性的拟合方法——超折面的拟合方法，其效果如何，有待試驗。