第43卷 第2期　　气 象 学 报　　Vol.43, No. 2
1985年5月　　ACTA METEOROLOGICA SINICA　　May, 1985

关于车贝雪夫多项式和自然正交函数作气象场展开的收敛性、稳定性问题*

黄 文 杰
（国家气象局气象科学研究院）

周 家 斌
（中国科学院大气物理研究所）

车贝雪夫(Чебыщев)多项式和自然正交函数都是气象上常用的正交函数。后者从气象场本身的协方差矩阵引出正交函数，能有效地浓缩气象场信息，且可用不规则格点资料。以往国内外对车贝雪夫多项式的应用都只限于等距格点，周家斌[1]将它推广到不规则格点上，从而扩大了它的用途。

这两种正交函数各有一些优缺点，因而如何充分发挥其优点而避免其缺点就成为一个重要的研究问题。我们拟对这两种正交函数作一些比较和综合研究，本文重点放在收敛性和稳定性方面。

1. 收 敛 性

设我们有一组资料，以 $Z_{t,i,j}$ 表示。其中 t 表示资料的时次($t=1,2,\cdots,T_0$)，i,j 为资料的二维序号($i=1,2,\cdots,I_0$；$j=1,2,\cdots,J_0$)。我们可以把 $Z_{t,i,j}$ 用车贝雪夫多项式展开，即

$$\tilde{Z}_{t,i,j}=\sum_{k=0}^{K_0}\sum_{s=0}^{S_0}A_{t,k,s}\cdot\varphi_{k,i}\cdot\psi_{s,j} \tag{1}$$

其中 $\tilde{Z}_{t,i,j}$ 表示 t 时刻，格点 i，j 处的拟合值；$\varphi_{k,i}$ 为 k 阶规一化车贝雪夫多项式在 i 点的值；$\psi_{s,j}$ 为 S 阶规一化车贝雪夫多项式在 j 点的值；K_0，S_0 为拟合所取截止项数。

可以用如下相关指数表示(1)式对 $Z_{t,i,j}$ 的拟合精度

$$r_{t,K_0,S_0}=\frac{\sum\limits_{k=0}^{K_0}\sum\limits_{s=0}^{S_0}A_{t,k,s}}{\sum\limits_{i=1}^{I_0}\sum\limits_{j=1}^{J_0}Z^2_{t,i,j}} \tag{2}$$

现在将同一资料用自然正交函数展开，为了与一般的自然正交函数表达式一致，我们把二维序号(i,j)转化为一维格点表示，即 $l=1,2,\cdots,L_0$($L_0=I_0\times J_0$)。则自然正交函数展开表示为

* 本文于1982年6月10日收到，1984年1月24日收到修改稿。

$$\widetilde{Z}_{t,l}=\sum_{k=1}^{K_1} T_{t,k}\cdot X_{k,l} \qquad \begin{array}{l} t=1,\ 2,\cdots,T_0 \\ l=1,\ 2,\cdots,L_0 \end{array} \tag{3}$$

式中 K_1 为拟合所取截止项数。相应的相关指数为

$$r_{K_1}=\frac{\sum\limits_{k=1}^{K_1}\lambda_k}{\sum\limits_{l=1}^{L_0}R_{l,l}} \tag{4}$$

其中 λ_k 为协方差矩阵的第 k 个特征值。

在比较两种展开的收敛性时,取同一相对误差可分别自(2)式和(4)式定出各自展开所需的项数。需要的项数愈少,收敛愈快。

在欧亚地区(20°—70°N, 0°—160°E),我们按表1配置取35个格点,由于其分布很不规则,为适合二维车贝雪夫展开,将格点按5×7结构排列,其二维序号见表1。

表 1 欧亚35个格点位置及二维序号

二维序号 \ 东经 / 北纬	0°	10°	20°	30°	40°	50°	60°	70°	80°	90°	100°	110°	120°	130°	140°	150°	160°
70°			(1,2)				(1,4)				(1,5)				(1,6)		
60°	(1,1)		(2,2)		(1,3)		(2,4)				(2,5)		(2,6)		(2,7)		(1,7)
50°	(2,1)		(3,2)		(2,3)		(3,3)		(3,4)		(3,5)		(3,6)		(3,7)		
40°		(3,1)		(4,1)		(4,2)		(4,3)		(4,4)		(4,5)		(4,7)			
30°							(5,1)		(5,2)		(5,4)		(4,6)		(5,7)		
20°										(5,3)		(5,5)		(5,6)			

资料场 Z 取500 mb高度,时段由1951年1月至1980年12月。

从30年平均场看,每个月的 $A_{k,s}$ 根据绝对值的大小作顺序排列如表2(只给出前10个)。

表 2 1951—1980年各月平均场 $A_{k,s}$ 绝对值大小排序

月 \ 次序	1	2	3	4	5	6	7	8	9	10
1	A_{00}	A_{10}	A_{01}	A_{20}	A_{11}	A_{21}	A_{12}	A_{06}	A_{14}	A_{13}
2	A_{00}	A_{10}	A_{01}	A_{20}	A_{11}	A_{21}	A_{03}	A_{12}	A_{14}	A_{02}
3	A_{00}	A_{10}	A_{01}	A_{11}	A_{20}	A_{12}	A_{21}	A_{06}	A_{14}	A_{13}
4	A_{00}	A_{10}	A_{01}	A_{11}	A_{05}	A_{20}	A_{14}	A_{02}	A_{21}	A_{13}
5	A_{00}	A_{10}	A_{01}	A_{11}	A_{05}	A_{13}	A_{21}	A_{22}	A_{31}	A_{14}
6	A_{00}	A_{10}	A_{01}	A_{06}	A_{21}	A_{13}	A_{11}	A_{16}	A_{15}	A_{22}
7	A_{00}	A_{10}	A_{21}	A_{06}	A_{42}	A_{14}	A_{01}	A_{16}	A_{40}	A_{30}
8	A_{00}	A_{10}	A_{01}	A_{21}	A_{21}	A_{30}	A_{16}	A_{15}	A_{11}	A_{20}
9	A_{00}	A_{10}	A_{01}	A_{11}	A_{11}	A_{06}	A_{16}	A_{31}	A_{46}	A_{16}
10	A_{00}	A_{10}	A_{01}	A_{11}	A_{21}	A_{12}	A_{05}	A_{31}	A_{13}	A_{46}
11	A_{00}	A_{10}	A_{01}	A_{11}	A_{21}	A_{05}	A_{14}	A_{31}	A_{13}	A_{46}
12	A_{00}	A_{10}	A_{01}	A_{11}	A_{21}	A_{20}	A_{06}	A_{14}	A_{13}	A_{03}

从表 2 可知，A_{00}，A_{10} 这两个系数在全年各月都分别排在第一、第二位，A_{01} 则除 7 月份外皆排在第三位。这说明按表 1 的格点编号能较好地把主要信息中于 A_{00}, A_{10}, A_{01} 这三个低阶系数上。

对上述资料还计算了 $\bar{r}_{k,s}$ 值：

$$\bar{r}_{k,s}=\frac{1}{T_1}\sum_{t=1}^{T_1} r_{t,k,s} \tag{5}$$

式中 $T_1=12, k=0,1,\cdots,4,\ s=0,1,\cdots,6$

表 3 列出了各月按表 2 排序的前 8 个 $r_{t,k,s}$ 值及 $\bar{r}_{k,s}$ 值。

表 3　车贝雪夫展开与自然正交函数展开收敛性比较

$r_{t,k,s}$ 次序 / 月(t)	1	2	3	4	5	6	7	8
1	0.7752	0.9388	0.9604	0.9679	0.9745	0.9801	0.9833	0.9859
2	0.7983	0.9474	0.9643	0.9733	0.9793	0.9855	0.9855	0.9879
3	0.8276	0.9626	0.9747	0.9795	0.9837	0.9859	0.9880	0.9894
4	0.8843	0.9799	0.9855	0.9886	0.9905	0.9920	0.9930	0.9939
5	0.9365	0.9895	0.9929	0.9943	0.9954	0.9963	0.9967	0.9971
6	0.9701	0.9964	0.9973	0.9978	0.9982	0.9985	0.9987	0.9988
7	0.9824	0.9981	0.9984	0.9986	0.9987	0.9988	0.9989	0.9990
8	0.9770	0.9974	0.9977	0.9979	0.9981	0.9983	0.9985	0.9986
9	0.9521	0.9939	0.9956	0.9964	0.9970	0.9974	0.9977	0.9979
10	0.9001	0.9831	0.9898	0.9933	0.9942	0.9949	0.9956	0.9963
11	0.8305	0.9649	0.9794	0.9858	0.9885	0.9903	0.9920	0.9932
12	0.7915	0.9503	0.9685	0.9742	0.9794	0.9844	0.9868	0.9871
$\bar{r}_{k,s}$	0.8855	0.9752	0.9837	0.9873	0.9898	0.9919	0.9929	0.9939
r_k	0.8740	0.9042	0.9268	0.9465	0.9568	0.9648	0.9713	0.9762

表 3 指出，车贝雪夫展开的收敛速度是很快的，并不低于自然正交函数展开的收敛速度，这无论从 12 个月的平均收敛速度抑或任一个月的收敛速度说，都是如此。

对比表 3 末两行知，车贝雪夫展开只取前两项，收敛效果已相当于自然正交函数展开取前 7 项。即使收敛速度较慢的冬季，尤其是一月份，车贝雪夫展开取前 2 项也优于自然正交函数展开取前 3 项。

2. 稳 定 性

(1) 式中 $\varphi_{k,i}$，$\psi_{s,j}$ 只与格点分布有关，与时间没有关系，因而 $\varphi_{k,i}\cdot\psi_{s,j}$ 也只与格点有关，而与时间无关。

由于车贝雪夫展开是对原资料场 $Z_{t,i,j}$ 的各时刻($t=1,2,\cdots$)作独立展开，无论 $t=1,2$ 或任何时刻，$\varphi_{k,i}\cdot\psi_{s,j}$ 在 (i,j) 上的值始终是同样的。故应该说，在车贝雪夫展开中 $\varphi_{k,i}\cdot\psi_{s,j}$ 是绝对与时间无关的。而自然正交函数展开的空间部分 $X_{k,l}$ 虽然在形式上与时间无关，但由于它不是对原始资料场 $Z_{t,l}$ 的各时刻 ($t=1,2\cdots$) 作独立展开，而是通过 $Z_{t,l}$ 的协方差矩阵 R 作运算求得。协方差矩阵 R 是随 t 值大小有变化的，故 $X_{k,l}$，

实际上是与 l 的大小（即样本容量）有关。故应该说，自然正交函数展开的空间部分 $X_{k,l}$ 只能在样本容量取定，样本起、终点也取定情况下才是固定的。样本容量和样本起点至终点的任何变动都会影响 $X_{k,l}$ 值，因之它不是绝对与时间无关的。

总之，车贝雪夫展开的空间部分对时间说是严格稳定的；而自然正交函数展开的空间部分与样本取法相关联，因之它对时间是依赖的。自然正交函数展开的不稳定性也在于它。

上述观点，周家斌已在文献[1]中概括地提出过，现在根据资料计算作明确的比较。

表 4 a 给出 1951 年 1 月至 1980 年 12 月共 360 个月，按表 1 所示欧亚地区 500 mb 高度场的自然正交函数展开的第 5 个特征向量场。表 4 b 为同样格点，但时间为 1951 年 1 月至 1979 年 12 月共 348 个月的同一特征向量场。对比表 4 a 和 4 b 知它们的符号完全相反。同样符号相反也在第 8 个特征向量场出现。展开结果的比较表明，时间相差一年的两组展开，只有第一特征向量基本上有稳定性，第二特征向量及其它特征向量皆出现程度不同，甚至符号相反的不稳定性。由于篇幅所限，这些结果的资料均未给出。

表 4 a $l=360$ 的自然正交函数展开第 5 个特征向量场（1951—1980 年共 360 个月）

$x_{5,l}$ i \ j	1	2	3	4	5	6	7
1	−0.187	−0.443	−0.156	−0.019	−0.058	0.093	0.057
2	−0.081	−0.017	−0.039	0.157	0.149	0.062	−0.037
3	−0.019	−0.324	−0.095	−0.005	0.005	−0.046	−0.403
4	−0.200	0.035	0.010	−0.029	0.103	0.074	0.112
5	−0.104	−0.049	0.335	0.309	0.238	−0.055	−0.193

表 4 b $l=348$ 的自然正交函数展开第 5 个特征向量场（1951—1979 年共 348 个月）

$x_{5,l}$ i \ j	1	2	3	4	5	6	7
1	0.188	0.447	0.155	0.026	0.059	−0.081	−0.054
2	0.084	0.024	0.043	−0.153	−0.143	−0.159	0.040
3	0.020	0.328	0.092	−0.003	−0.005	0.051	0.399
4	0.189	−0.045	−0.010	0.031	−0.103	−0.085	−0.120
5	0.108	0.051	−0.330	−0.313	−0.237	0.057	0.198

3. 稳定性问题对天气预报的影响

稳定性问题对天气分析和预报的影响可从如下几个方面加以说明：

1) 天气分析

设有某时段 $t=1,2,\cdots,T_1$ 的资料 $Z_{t,i,j}$，用车贝雪夫多项式展开得到车贝雪夫系数

的多维时间序列 $A_{t,k,s}$。用自然正交函数展开得到时间系数的序列 $T_{t,k}^{(1)}$ 和空间函数 $X_{k,l}^{(1)}$，这两个函数都依赖于本时段的资料协方差矩阵 $R^{(1)}$。因此 $T_{t,k}^{(1)}$ 和 $X_{k,l}^{(1)}$ 皆受到 T_1 的制约。

现在把资料延长至 T_2，即补充新资料 $Z_{t,i,j}$，其中 $t=T_1+1, T_1+2, \cdots, T_2$。

若用车贝雪夫展开，我们只要将新补充的资料展开就得到车贝雪夫系数时间序列的延长，旧资料的车贝雪夫系数性质和量值都保持不变。因之，对序列的研究可只限于对新延长的车贝雪夫系数作研究。

用自然正交函数展开时，我们不能只处理新补充的资料，否则就无法对前后两个时段的结果进行统一的分析研究。因此当资料加长后需对前后两段资料合在一起重新进行分析。

2）天气预报

用车贝雪夫多项式作预报时，由于 $\varphi_{k,i}$ 和 $\psi_{s,j}$ 不变，我们只需要预报系数 $A_{T_1+n,k,s}$。因之预报误差将由两部分决定：a）$A_{T_1+n,k,s}$ 的预报误差；b）由于截断项取至 $k=K_0, s=S_0$ 引起的拟合误差。

用自然正交函数作预报时，由于 $Z_{T_1+n,k}$ 未知，因而无法求出包括 T_1+n 时刻资料的协方差矩阵，因而也求不出 $X_{k,l}$。在作预报时，一般假定 $X_{k,l}$ 不变，而仅预报时间系数 $T_{T_1+n,k}$。因此，用自然正交函数作预报取决于如下三部分误差：a）$T_{T_1+n,k}$ 的预报误差；b）截断项取至 K_1 的拟合误差；c）假设 $X_{k,l}$ 不变引起的误差。

单就拟合误差（即 b 项）说，由于拟合所取项数总少于格点数，显然有可能是某些点上自然正交函数展开的拟合好些，而另一些点则反之。至于不稳定引起的误差则仅是自然正交函数才有，因而是一部分额外的损失。

3）资料的压缩传输与存储

（1）资料传输　将原资料化为车贝雪夫系数传输由于 $\varphi_{k,i}, \psi_{s,j}$ 不随时间变化，故接收者只要知道 I_0 和 J_0 就很容易用系数求得原资料场。用自然正交函数则不可能仅根据时间系数去求得原资料场，必须同时传输 $X_{k,l}$，才能使接收者根据 $T_{t,k}$ 和 $X_{k,l}$ 去求得原资料场。

（2）资料存储　新资料出现后，自然正交函数的时间系数和空间函数都将变化，旧的展开系数和特征向量场资料将失去存储价值。车贝雪夫系数则因各时刻间有相互独立性而有永久存储价值。