气象场经验正交函数不同展开方案收敛性问题的探讨 *

丁裕国 施 能
(南京气象学院，南京，210044)

提 要

本文从理论上证明EOFs(经验正交函数)用于气象场时，三种不同计算方案特征值的相互关系．指出利用原始场积和矩阵展开的特征值 λ_j 等于用协方差阵展开的特征值 λ_j^* 加上平均场方差引起的特征值 $\overline{\lambda_j}$．进而论证了收敛速度差异的原因．最后又以实例表明，不同展开方案的典型场意义及其与时间权重系数的对应关系．

关键词：经验正交函数；气象场；特征值；收敛性．

一、引 言

在气象场的数学描述方面，许多学者都曾作过有益的探讨[1-5]．在各种正交函数用于场的数学逼近时，差不多都存在着两个至关重要的问题——稳定性与收敛性．就经验正交函数(EOFs)来说，Preisendorfer 和Barnett[6] 以及 Overland 和 Preisendorfer[7] 曾分别给出了检验随机特征值的方法；Gray 和North[8] 等曾证明小样本所估计的特征向量为不稳定结构；章基嘉等[9] 曾讨论 EOFs 特征向量的稳定性问题；作者[10] 曾就EOFs展开气象场的收敛速度建立了经验关系．除上述问题外，实际应用的普遍性也促进人们对EOFs方法及其理论意义作深入探讨．例如，对同样一种气象场来说，往往可以有三种不同的EOFs计算方案：其一是直接用样本场空间各点的积和组成积和矩阵求解相应的特征值和特征向量；其二是用样本场空间各点的协方差组成协方差矩阵求解相应的特征值和特征向量；其三是用样本场空间各点的相关系数组成的相关矩阵求解相应的特征值和特征向量．三种计算EOFs方案的特征值收敛性和稳定性差别很大．产生这些差异的原因及其应用意义是一个值得探讨的问题，而EOFs算法不同，不但对特征向量(即典型场)有影响，而且对时间权重系数也有影响．此外，样本场序列长度 n 和空间站点数 p 的大小对收敛性和稳定性究竟有什么影响，也是一个重要问题．

特别应指出的是，大范围气象场历史序列的重建和恢复，常常借助于EOFs或以它为基础的多元统计方法(如典型相关、因子分析等)，所有这些都涉及到一个重要的理论前提，即在何种情况下，样本协方差阵与相应的特征值和特征向量是稳定的．只有从理论上解决这类问题，恢复气象场历史序列才更有可靠的基础．

本文正是针对上述一系列理论问题，着重探讨气象场EOFs不同计算方案的收敛性问题．

1990年2月27日收到，12月3日收到再改稿．
* 本文为气象基金资助课题．

二、EOFs不同计算方案的特征值收敛性

设有气象场序列 $x=(x_1, x_2, \cdots, x_p)$，其中下标 p 为空间点(或网格点)数目．若以矩阵表示，可描述为

$$X=\begin{pmatrix} x_{(1)}' \\ x_{(2)}' \\ \vdots \\ x_{(n)}' \end{pmatrix}=(x_{ij}), \quad \begin{matrix} i=1, 2, \cdots, n, \\ j=1, 2, \cdots, p, \end{matrix} \tag{1}$$

上式 n 为时间点数(样本容量)，且约定 x_{ij} 的均值不等于零，方差不等于1，即非标准化资料．则协方差阵可写为

$$S=\frac{1}{n} X'\left(I-\frac{1}{n} J\right)X, \tag{2}$$

式中 I 为单位矩阵，J 为全1方阵，显然 $S=S'$为 $p\times p$ 对称阵．若存在S的 p 个特征向量 $u_1, u_2, \cdots, u_p$，利用正交变换，必有

$$S=U\Lambda U'=U\Lambda U'=U\ \mathrm{diag}(\lambda_1^*, \lambda_2^*, \cdots, \lambda_p^*)U', \tag{3}$$

其中 $U=(u_1, u_2, \cdots, u_p)$，$\Lambda=\mathrm{diag}(\lambda_1^*, \lambda_2^*, \cdots, \lambda_p^*)$．换言之，有下列关系式

$$U'SU=\Lambda=\mathrm{diag}(\lambda_1^*, \lambda_2^*, \cdots, \lambda_p^*), \tag{4}$$

引入(2)式，应有

$$U\left[\frac{1}{n} X'\left(I-\frac{1}{n} J\right)X\right]U=\mathrm{diag}(\lambda_1^*, \lambda_2^*, \cdots, \lambda_p^*). \tag{5}$$

为推导方便，令积和矩阵为

$$A=\frac{1}{n} X'X \tag{6}$$

于是(5)式又可写为

$$U'AU=U'\left(\frac{1}{n^2} X'JX\right)U+\mathrm{diag}(\lambda_1^*, \lambda_2^*, \cdots, \lambda_p^*), \tag{7}$$

上式右端第二项对角阵元素 $\lambda_1^*\geqslant\lambda_2^*\geqslant\cdots\geqslant\lambda_p^*$ 为相应于协方差阵S 的谱分解的谱，也即为相应的特征值[11]．而右端第一项则可化简为

$$\begin{aligned} U'\left(\frac{1}{n^2} X'JX\right)U &= U'\left[\frac{1}{n^2}\left(\sum_i x_{ij}\right)(x_{ij})\right]U \\ &= U'\overline{A}U, \end{aligned} \tag{8}$$

其中

$$\overline{A} = \begin{pmatrix} \overline{x}_1^2 & \overline{x}_1\overline{x}_2 & \cdots & \overline{x}_1\overline{x}_p \\ \overline{x}_2\overline{x}_1 & \overline{x}_2^2 & \cdots & \overline{x}_2\overline{x}_p \\ \cdots & \cdots & \cdots & \cdots \\ \overline{x}_p\overline{x}_1 & \overline{x}_p\overline{x}_2 & \cdots & \overline{x}_p^2 \end{pmatrix} .$$

在(8)式中，$\overline{A}$ 表示以$\overline{x}_1, \overline{x}_2, \cdots, \overline{x}_p$交叉乘积为元素的实对称矩阵.由于 U为正交向量，故(8)式仍为一个实对称方阵，因此，对于矩阵$\overline{A}$经相似变换必有

$$V'\overline{A}V = \mathrm{diag}\,(\overline{\lambda}_1, \overline{\lambda}_2, \cdots, \overline{\lambda}_p), \tag{9}$$

这里 V为相应于$\overline{A}$ 的特征向量，式中右端对角阵元素$\overline{\lambda}_1 \geqslant \overline{\lambda}_2 \geqslant \cdots \geqslant \overline{\lambda}_p$ 分别代表相应于$\overline{A}$ 的特征值. 又由(2)和(8)式可知矩阵

$$A = \overline{A} + S . \tag{10}$$

根据矩阵理论[11]，上式两端矩阵之迹应满足下列关系：$\mathrm{tr}A = \mathrm{tr}\overline{A} + \mathrm{tr}S$，又因矩阵之迹等价于其相应特征值之和，故有

$$\sum_i A_{ii} = \sum_j \lambda_j = \sum_j \overline{\lambda}_j + \sum_j \lambda_j^* , \tag{11}$$

式中 λ_j, $(j = 1, 2, \cdots, p)$即为相应于矩阵 A的特征值，并规定有 $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_p$；而$\overline{\lambda}_j$和$\lambda_j^*$ 已如前面定义. 考虑到(4)和(9)式的结果，根据相似变换和矩阵的谱分解，(11)式又可写成等价形式

$$\mathrm{diag}\,(\lambda_1, \lambda_2, \cdots, \lambda_p) = \mathrm{diag}\,(\overline{\lambda}_1, \overline{\lambda}_2, \cdots, \overline{\lambda}_p) + \mathrm{diag}\,(\lambda_1^*, \lambda_2^*, \cdots, \lambda_p^*) . \tag{12}$$

(11)和(12)式的意义表明，气象场积和矩阵的谱分解实际上包含两个独立的组成部分：(i)由平均场引起的方差对场的总平方和的贡献；(ii)由距平场引起的方差(见(4)式)对场的总平方和的贡献.

就某一特征向量场(即典型场)而言，据(12)式，一般应有

$$\lambda_j = \overline{\lambda}_j + \lambda_j^* , \qquad j = 1, 2, \cdots, p , \tag{13}$$

式中 λ_j 即为积和矩阵EOFs 的第j个典型场方差. 可见，一个典型场的方差也可认为是由平均场引起的方差$\overline{\lambda}_j$与由距平场引起的方差 λ_j^*迭加而成.由(13)式的关系，若已知A阵的 EOFs 展开的特征值λ_j以及平均场的离差阵$\overline{A}$与相应的特征值$\overline{\lambda}_j$，就可推算出协方差阵S的 EOFs 展开的特征值 λ_j^*. 由此可见，两种不同的EOFs 展开方案其特征值 λ_j 与λ_j^*可互相推算.另外指出，对应于(12)式的各相应特征向量是不一定相同的.后文实例计算已说明这一事实.

类似于上述推导，也可求得原始场相关矩阵 R 的 EOFs 与积和矩阵的 EOFs 的关系. 由(2)式可令

$$S = \frac{1}{n} X' \left(I - \frac{1}{n} J \right) X = DRD \tag{14}$$

或

$$R=D^{-1}SD^{-1}=\frac{1}{n}D^{-1}X'\left(I-\frac{1}{n}J\right)XD^{-1}, \tag{15}$$

式中 $D^{-1}=\ \mathrm{diag}(S_{11}^{-\frac{1}{2}}, S_{22}^{-\frac{1}{2}}, \cdots, S_{pp}^{-\frac{1}{2}})$.

只要令 $Y=XD^{-1}$，则 $Y'=D^{-1}X'$，于是(15)式就可写为

$$R=\frac{1}{n}Y'\left(II-\frac{1}{n}J\right)Y. \tag{16}$$

利用(2)—(12)式的方法，就可得到类似于(11)和(12)式的表达式

$$\sum_i B_{ii}=\sum_j \widetilde{\lambda}_j=\sum_j \overline{\lambda}_j'+\Sigma\lambda_j'^{*}. \tag{17}$$

$$\begin{aligned} Q'BQ&=\mathrm{diag}(\widetilde{\lambda}_1, \widetilde{\lambda}_2, \cdots, \widetilde{\lambda}_p)\\ &=\mathrm{diag}(\overline{\lambda}_1', \overline{\lambda}_2', \cdots, \overline{\lambda}_p')+\mathrm{diag}(\lambda_1'^{*}, \lambda_2'^{*}, \cdots, \lambda_p'^{*})\\ &=\mathrm{diag}(\overline{\lambda}_1'+\lambda_1'^{*}, \overline{\lambda}_2'+\lambda_2'^{*}, \cdots, \overline{\lambda}_p'+\lambda_p'^{*}), \end{aligned} \tag{18}$$

上式中 $B=\frac{1}{n}Y'Y$ 为原始场资料经尺度缩减后的相对量 $Y=XD^{-1}$ 的积和矩阵，Q 为相应于 B 阵的特征向量，$\overline{\lambda}_i'$，$(i=1, 2, \cdots, p)$ 和 $\lambda_i'^{*}$，$(i=1, 2, \cdots, p)$ 分别为相应于(经尺度缩减后的)平均场和距平场特征值．(14)—(16)式中 R 就是标准化协方差阵即相关矩阵，因而，(18)式仅是(12)式的特例．

三、实例验证

以上海、南通、杭州3站某月降水量年际序列为例，作EOFs展开．分别用 A、S、$\overline{A}$ 矩阵作EOFs计算， 求得相应的特征值和方差贡献见表1.

$$A=\begin{pmatrix} 12794.66 & 8807.56 & 13818.57\\ 8807.56 & 15954.09 & 8194.96\\ 13818.57 & 8194.96 & 18054.86 \end{pmatrix}$$

$$S=\begin{pmatrix} 2570.77 & -1123.77 & 2822.17\\ -1123.77 & 6306.92 & -2486.76\\ 2822.77 & -2486.76 & 6227.57 \end{pmatrix}$$

$$\overline{A}=\begin{pmatrix} 10223.89 & 9931.33 & 10996.40\\ 9931.33 & 9647.17 & 10681.72\\ 10996.40 & 10681.72 & 11827.29 \end{pmatrix}$$

表1 上海、南通、杭州3站某月降水量场EOFs展开结果

序号 i	A		$\overline{A}$		S		$\overline{A}+S$
	λ_i	方差%	$\overline{\lambda}_i$	方差%	λ_i^*	方差%	λ_i
1	36455.67	77.9	26602.55	83.92	9853.12	65.2	36455.67
2	9237.41	19.7	5012.18	15.81	4225.23	28.0	9237.47
3	1110.48	2.4	83.62	0.26	1026.86	6.8	1110.48
Σ	46803.56	100	31698.35	100	15105.21	100	46803.56

表1的结果验证了(11)—(13)式的正确性. 从表1还可看出，正如预期的那样，用A阵展开时收敛速度比用S阵展开的快，而A阵EOFs收敛快的原因正是由于平均场的A阵引起的. 在表1中，由于$\overline{\lambda}_i=26602.55$相对集中了平均场的绝大部分信息，因而当$S$阵展开时，$\lambda_1^*-\lambda_2^*$仅为37.2%方差百分比，而$A$阵展开时，$\lambda_1-\lambda_2$却为58.2%方差百分比. 当然，本文仅从讨论$A$阵收敛快的原因提出上述看法，并不表明作者赞成用积和矩阵A来作EOFs展开. 特别是由于通常EOFs展开所提取的信息主要是各典型场的权重系数随时间的变化特征，从时变特征的意义上说，平均场不一定能提供多少信息，因此，用积和矩阵A作EOFs展开时，收敛快是一种表面现象，后文还要从实例证明，描述场的变动性，还是以距平场或标准化变量场的EOFs为好. 很显然，假如我们消除了平均场的影响，则A与S阵的EOFs展开当然就没有收敛速度方面的差异了. 例如，众所周知的是气温场的大尺度特征要比降水场明显很多，文献[5]以及类似的许多计算都表明，气温场的收敛速度比降水场快得多，这个结果同样适用于平均场$\overline{A}$的空间结构. 如果p个空间点均值几乎相同(大尺度特征明显)，则$\overline{A}$的第一特征值特别大(对$\overline{A}$作EOFs展开时收敛快). 从而使得我们用积和矩阵A进行EOFs展开时收敛速度明显比用S阵快. 反之，若p个空间点几乎独立(或不相关)，这时A收敛很慢. 用A与用S展开的收敛速度差异也不会太大. 换言之，平均场的空间均匀性特征是影响EOFs收敛速度的一个因素.

四、不同展开方案时的特征向量场及其时间权重系数的相互关系

EOFs展开的不同方案不但特征值显示出收敛速度的差异，而且特征向量及其权重系数之间在不同展开时也有相应的意义.

文献[5]已经指出，用积和矩阵A作自然正交展开时，第一特征向量场代表平均场. 它与实测平均场之间的相似系数达到0.99以上，这与上述推证是一致的，这个结果实际上具有普遍意义. 换言之，用积和矩阵A展开时，第一典型场是平均场的主要信息(与平均场均匀性有关)，而第二、三、… 典型场则分别以不同的时间权重系数迭加到距平场上. 下面给出一个例子进一步证明不同展开方案时，时间权重系数的关系. 表2是用$p=2$，$n=11$时三种不同展开方案的EOFs结果(取两位小数). 其积和矩阵A、协方差阵S和相关矩阵R以及对应的特征向量矩阵U(行向量为特征向量)列于表3中. 由

表可见，(1)用 A 阵计算的第一特征向量的各分量几乎与 $\overline{x}_1$，$\overline{x}_2$ 成正比，即有 $\overline{x}_1/\overline{x}_2=\frac{35}{41}\doteq\frac{0.6371}{0.7708}$．这就表明，用 A 阵展开时，第一特征向量代表空间平均场的分布特征．(2)用 S 阵(协方差矩阵)展开的第一特征向量的时间权重系数 T_1 与用 A 阵展开的第二特征向量时间权重系数 T_2 的符号重合率达 11/11，其相关系数也已达到 0.9 以上，这就再次证明，积和矩阵 A 的第一特征向量代表了平均场的主要信息，而第二特征向量(及以后各特征向量)才为距平场信息．因此，对应的 S 阵展开时，由于它本身就是空间距平场展开结果，故其第一特征向量所反映的大尺度距平形势在时间上的演变特点即时间系数必然与 A 阵的第二特征向量(代表距平场)的时间系数相一致．作者计算的许多实例均表明以上结果具有普遍性．

表 2　原始场资料及不同展开方案的时间权重系数($p=2$，$n=11$)

序号	x_1	x_2	A		S		R	
			T_1	T_2	T_1	T_2	T_1	T_2
1	20	50	51.28	−16.44	−16.14	−6.75	−1.68	−0.62
2	26	45	51.25	−8.63	−8.55	−4.88	−0.93	−0.45
3	27	60	63.45	−17.42	−20.05	4.81	−1.74	0.51
4	28	50	56.38	−10.27	−11.40	−0.31	−1.07	−0.01
5	31	46	55.21	−5.41	−6.40	0.26	−0.60	−0.01
6	33	55	63.42	−9.61	−12.46	6.68	−0.98	0.67
7	39	35	51.82	7.76	7.20	−0.33	0.66	−0.05
8	41	25	45.39	15.67	16.44	−4.64	1.40	−0.49
9	42	33	52.19	11.34	10.59	0.90	1.01	0.06
10	48	24	49.08	21.71	21.40	0.40	2.00	−0.01
11	50	28	53.44	20.70	19.36	4.38	1.92	0.38
平均	35	41	53.90	0.86	0.0	0.0	0.0	0.0
方差	84.91	143.09	27.88	200.13	211.71	16.28	1.85	0.15

表3　对应于表2中 A、S、R 阵的特征向量(行向量)

种类＼项目	矩	阵	对应的特征向量	
A	1309.91	1341.73	0.6371	0.7708
	1341.73	1824.09	0.7708	−0.6371
S	84.91	−93.27	0.5926	−0.8055
	−93.27	143.09	0.8055	0.5926
R	1.00	−0.85	0.7071	−0.7071
	−0.85	1.00	0.7071	0.7071

(3) 由于 R 阵仅仅是原始资料阵经尺度缩减后所得的协方差阵(即标准化协方差阵)，它是相对量. 所以，用 R 阵 EOFs 展开所得时间系数在符号上与 S 阵展开保持一致. 由表 2 可见，R 阵 EOFs 与 S 阵 EOFs 展开的第一时间系数符号重合率也是 11/11，而第二时间系数的符号重合率则为 9/11. 至于数值大小，因受尺度缩减影响，一般说来，对应不一定很好，否则失去了尺度缩减的意义. 尺度缩减(即标准化处理)尤其在 p 个变量单位不一致时很有必要. 例如，对一些数值较小的变量来说，只有尺度缩减才能显示它对展开结果的影响. 换句话说，如果我们将变量标准化，各变量就可消除因尺度(量纲)不同对 EOFs 产生的影响，而只显示各变量本身的作用. 确定采用何种 EOFs 计算方案，一般应就具体问题而异，以上仅仅说明不同展开方案之间在本质上的联系.

顺便指出，当用 S 阵或 R 阵作 EOFs 展开时，时间权重系数实际上就是主成分分析中的主成分或主分量. 因其平均值为 0，而方差等于相应的特征值 λ_j'(见表 2). 但是，用积和矩阵 A 作 EOFs 展开时，时间权重系数并非代表主成分，因它的均值既不为 0，方差也不为 λ_j.

五、小 结

(1) 本文从理论上证明，气象场 EOFs 展开的不同算法所得特征值的相互关系表明，积和矩阵 A 的 EOFs 展开实质上包含距平场协方差阵 S 的 EOFs 展开与平均场交叉乘积矩阵 $\overline{A}$ 的 EOFs 展开两部分，其特征值也等于这两部分之和.

(2) 由于平均场迭加在距平场上而造成积和矩阵 A 的 EOFs 收敛速度增大. 而平均场本身空间均匀性程度又直接影响积和矩阵 A 的 EOFs 收敛速度增大的幅度. 场的空间均匀度愈大，A 阵 EOFs 收敛速度增加愈大.

(3) 用 A 阵展开时，第一特征向量代表了平均场的主要信息，第二特征向量及其后则代表了大尺度距平场的信息，而实际资料计算表明，用协方差阵 S 展开时的第一时间系数在数值与符号上恰好与用积和矩阵 A 展开的第二时间系数相接近，说明理论与实际的一致.

(4) 上述论证尚可推广到气象向量场的经验正交函数展开问题上，关于这一点，我们将另文讨论.

(5) 综上，可认为采用何种 EOFs 方案，应视具体问题而定. 用积和阵 A 展开，其第一典型场代表平均场信息，而从第二典型场开始的所有典型场线性迭加构成距平场；用协差阵 S 展开，其全部典型场线性迭加构成距平场. 因此，若在展开区域中没有距平意义的平均场(旱、涝；冷、暖)空间分布对未来有较大指示意义时，还是宜用积和阵展开，而若平均场空间分布(旱、涝；冷、暖)均匀一致，对未来无多大影响，则可用距平场的协差阵展开. 至于展开区域上空间点的标准差差别很大时，一般应用标准化变量(尺度缩减)，从而以相关矩阵 R 作为 EOFs 展开的基础.

致谢：程炳岩、牛涛两同志曾协助部分计算，特此致谢.

参 考 文 献

[1] 周家斌，1981，不规则格点上的车贝雪夫多项式展开问题，科学通报，26，第9期，548—550.

[2] Storch, H.V., 1983, Statistical aspects of EOFs based on small sample sizes, 2nd International meeting on statistical climatology, Lisboa Portugal, 4, No.5.1 — 4.5.7.

[3] Richman, M.B., 1983, Rotation of PCs in climatological research, Preprints Eighth conference on probability and statistics, *Atm. Sci.*, *Amer. Meteo. Soc.*, 115 — 124.

[4] 丁裕国，1985，EOFs展开气象场的几个问题，广西气象，第1期，5 — 9.

[5] 施能，1988，我国秋冬月降水、气温场的时空结构特征及其在我国初夏降水预报中的应用，大气科学，12，第3期，283 — 291.

[6] Preisendorfer, R.W. and Barnett, T.P., 1977, Significance tests for EOFs, Fifth conference on Probability and statistics in Atm.Sci., 169 — 172.

[7] Overland, J.E. and R.W.Preisendorfer, 1982, A significance test for PCs applied to a cyclone climatology, *Mon. Wea.Rev.*, **110**, No.1, 1 — 4.

[8] North, G., Bell, T.L., Cahalan, R.F., Moeng, F.G., 1982, Sampling errors in the estimation of EOFs, *Mon.Wea.Rev.*, **110**, No.7, 699 — 706.

[9] 章基嘉，孙照勃，陈松军，1979，论自然正交函数的稳定性，南京气象学院学报，第2期，89 — 98.

[10] 丁裕国，吴息，1988，经验正交函数展开气象场收敛性的研究，热带气象，4，第4期，316 — 325.

[11] 张尧庭、方开泰，1982，多元统计分析引论，科学出版社，23 — 27.

Research on Convergence of the Eigenvalues for the EOFs Expansion of Meteorological Field Among Different Schemes

Ding Yuguo　Shi Neng

(Nanjing Institute of Meteorology, Nanjing, 210044)

Abstract

The relationship of eigenvalues among three different calculation schemes is theoretically verified. It is shown that eigenvalues λ_j of the original field expansion with the product-moment matrix are equal to the eigenvalue λ_j with the covariance matrix plus the eigenvalue $\overline{\lambda}_j$ from the variance of the average field. Then, the cause of convergence difference of above three schemes is explained. Thus, the convergence speed of EOFs with the product-moment matrix depends not only on the correlation structure of the anomaly field but also on the uniformity feature of the average field. Furthermore, the corresponding relationship of time weighted coefficient and its canonical fields for different EOFs schemes are shown with several cases.

Key words: Empirical orthogonal functions (EOFs); Meteorological field; Eigenvalues; Convergence.