

气象场相关结构对 EOFs 展开稳定性的影响*

丁裕国 江志红

(南京气象学院, 210044)

提 要

本文从矩阵扰动理论出发, 提出利用矩阵的范数(norm)作为度量气象场随机扰动的稳定性指标, 并由此间接推估 EOFs 展开的稳定性。经理论论证、数值试验和实例计算表明, 气象场的相关性越好, 达到稳定相关结构所需样本越小, 由此得到的 EOFs 稳定性也越好, 反之则不然。上述规律又直接受样本大小 n 和站点数目 p 的影响。对于不同的气象场来说, 达到稳定 EOFs 的样本临界值不同, 必须警惕 EOFs 展开有可能不是稳定的。

关键词: 经验正交函数, 矩阵扰动, 抽样分布, 气象场。

1 引 言

经验正交函数(EOFs)展开气象场的稳定性直接影响其应用的可靠性。无论是采用 EOFs 作诊断分析和外推预报, 或是采用 EOFs 分量对气象场作某些区域资料的插补、延长, 人们都希望 EOFs 展开的特征向量具有稳定性。很难设想一个特征向量随样本变化并不稳定的气象场能有较为可靠的外推外延效果。因此, 从实际和理论两个方面来论证 EOFs 展开气象场的稳定性条件, 探讨在何种情况下气象场 EOFs 属于稳定展开, 何种情况下气象场 EOFs 属于不稳定展开, 很有必要。

文献[1]曾从真实数据场附加随机场的观点针对特征向量的抽样误差问题进行数值试验, 提出一种简单的抽样误差估计方法; 文献[2]针对实际气象场讨论过 EOFs 稳定的条件; 文献[3]和[4]都曾由实际气象场序列的分段计算中寻求历史时期不同阶段特征向量的稳定性规律, 论证了 EOFs 展开的稳定性问题。但是, 从理论的高度, 寻求广泛意义下的 EOFs 稳定性条件尚缺少深入研究。

本文基于上述目的, 试图结合相关系数随机抽样理论和矩阵扰动理论, 探讨 EOFs 展开气象场的稳定性影响因素, 以便建立 EOFs 展开气象场稳定性的理论基础。

2 相关矩阵的随机扰动

设有气象场 $X = (x_1, x_2, \dots, x_p)'$, 其样本量为 n 的相关矩阵记为

$$R_n = \begin{pmatrix} r_{11}^{(n)} & r_{12}^{(n)} & \dots & r_{1p}^{(n)} \\ r_{21}^{(n)} & r_{22}^{(n)} & \dots & r_{2p}^{(n)} \\ \dots & \dots & \dots & \dots \\ r_{p1}^{(n)} & r_{p2}^{(n)} & \dots & r_{pp}^{(n)} \end{pmatrix} \quad (1)$$

若有

* 1991 年 7 月 29 日收到原稿, 1991 年 11 月 28 日收到修改稿。该文为气象基金资助项目论文之一。

$$\lim_{n \rightarrow \infty} r_{ij}^{(n)} = \rho_{ij} \quad (i, j = 1, 2, \dots, p) \quad (2)$$

则称相关矩阵序列 $\{R_n\}$ 有极限 $R = (\rho_{ij}) \in C^{p \times p}$ (实对称矩阵集合), 又称 R_n 收敛到 R , 并记为

$$\lim_{n \rightarrow \infty} R_n = R \quad (3)$$

根据矩阵扰动理论^[5], 对 $\{R_n\}$ 取极限 R , 等价于对它们的范数取极限, 记为

$$\lim_{n \rightarrow \infty} \nu(R_n - R) = 0 \quad (4)$$

式(4)表明, 当 $n \rightarrow \infty$ 时, R_n 与极限 R 的差阵的范数趋于零。式中 $\nu(\cdot)$ 代表范数符号。由于其表达式有多种, 本文为讨论方便仅取其中的一种^[5], 即 α 范数

$$\|R_n\|_\alpha = \frac{1}{p} \sum_{i,j=1}^p |r_{ij}^{(n)}| \quad (5)$$

对于 $\{R_n\}$, 实际应视为随机矩阵序列, 由于 R_n 的元素 $r_{ij}^{(n)}$ 有其抽样分布^[6], 因此, 若设 R 为其总体相关矩阵, 并假定 R_n 与 R 具有相同的选站, 且 p 为定值, 则式(4)应推广为均方意义下的极限^[7]。即当 $n \rightarrow \infty$ 时, R_n 满足

$$\lim_{n \rightarrow \infty} P\{\nu(R_n - R) \leq \varepsilon\} = 1 \quad (6)$$

上式表明, 当 n 充分大时, 由 r_{ij} 组成的样本矩阵 $(r_{ij}^{(n)}) = R_n$ 与由 ρ_{ij} 组成的总体矩阵 $(\rho_{ij}) = R$ 之间的差值的范数小于任一给定微量的可能性随 n 增大而增大直至概率为 1。因此称 R_n 有均方极限 R 。

众所周知, 在给定 ρ_{ij} 的情形下, r_{ij} 的抽样分布仅与 n 的大小有关^[6]。因此, 各个元素 $r_{ij}^{(n)}$ 的抽样振动必引起整个相关矩阵 R_n 的随机扰动。虽然实际气象场序列的相关矩阵还可能受气候变化所导致的总体相关特征改变的影响, 不过, 为了研究的方便, 本文暂不考虑这种影响, 而仅限于讨论资料来自同一总体这一特殊情况。至于长期气候趋势对气象场所造成的影响, 作者将做另外的探讨。

就相关矩阵的抽样分布而论, 许多精确分布或渐近分布如 Wishat 分布、Wilks 统计量、 T^2 统计量以及 χ^2 , F 分布等都可用于数学描述^[8]。但由于大多数多元抽样分布并不便于数学处理, 有的分布又十分复杂, 对于解决本文提出的理论问题有一定难度。而从文献[5]出发, 结合相关系数抽样分布的整体效应, 从另一种途径讨论 EOFs 稳定性及其有关问题, 则具有明显的优点。

3 谱改变量与 EOFs 稳定性度量

据式(1)至(3)的约定, 设其相应特征值分别为 $\lambda(R_n) = \{\lambda_i\}$, $\lambda(R) = \{\mu_i\}$, $i = 1, 2, \dots, p$, 并以 $d(\lambda(R_n), \lambda(R))$ 表示 R_n 和 R 的特征值差异的某种度量。因为实对称

阵的谱分解有 $R_n = \sum_{i=1}^p \lambda_i u_i u_i'$, 其特征值 λ_i 又称为 R_n 的谱, 故将 $d(\lambda(R_n), \lambda(R))$ 又称

为谱的改变量。若记 $D(R_n, R)$ 表示 R_n 和 R 之间的某种差距, 则通常有下列关系成立^[5]

$$d(\lambda(R_n), \lambda(R)) \leq F(D(R_n, R)) \quad (7)$$

或

$$f\{d(\lambda(R_n), \lambda(R))\} \leq D(R_n, R) \quad (8)$$

式中 $F(\cdot)$ 或 $f(\cdot)$ 分别为某种简单函数。公式(7)或(8)表明, 矩阵 R 因受随机扰动而使其谱(即特征值)发生某种改变, 它一般与矩阵 R 所受的扰动程度有关。根据文献[5], 关于 $d(\lambda(R_n), \lambda(R))$ 可具体引用多种谱改变量公式, 为研究方便, 本文仅引用下列公式:

$$S_R(R_n) \equiv \max_{1 \leq i \leq p} \{ \min_{1 \leq j \leq p} |\lambda_i - \mu_j| \} \quad (9)$$

由于 R_n 或 R 为实对称非负定阵, 其特征值通常总有 $\lambda_i \geq 0$, 因此, 式(9)实质上表明两个矩阵 R_n 和 R 的由大到小排列的特征值之差的某种函数, 它代表了特征值的变动水平。

由式(7)至(9), 利用文献[5]提供的理论公式就可估计相关矩阵 R_n 因抽样扰动而引起的特征值(谱)可能最大改变量, 其理论公式为

$$S_R(R_n) \leq P^{1/2} (2m)^{1-1/2} \|R_n - R\|^{1/2} \quad (10)$$

其中 $m = \max\{\|R_n\|, \|R\|\}$, $\|R_n - R\|$ 为两矩阵差值的范数。

另一方面, 从实际问题出发, 我们希望提供达到 EOFs 稳定展开的最少样本量 N , 使得当 $n \rightarrow N$ 时, 样本相关矩阵 R_n 近似于总体相关矩阵 R 。根据矩阵理论和二阶矩随机性, 由均方极限的性质知, 只要式(4)成立, 必有式(3)成立。但是, 我们不可能得到无限大样本的抽样, 而只能得到有限样本的抽样。换言之, 当样本量 n 达到某一临界值 N 时, 总可定义一个近似于均方收敛状态的条件, 即有

$$\lim_{n \rightarrow N} \nu(R_n - R) \leq \varepsilon (\text{小量}) \quad (11)$$

并称满足式(11)的相关矩阵 R_n 为在给定精度 ε 下, 达到均方意义下的稳定性或均方稳定。显然, 在式(11)中, ε 愈小, R_n 愈稳定。

又由式(7)和(8), 气象场相关矩阵愈稳定, 则其范数 $\nu(R_n - R)$ 愈小, 相应的 EOFs 谱改变量也愈小。由式(5)知, $\nu(R_n - R)$ 实质上代表一个气象场的平均相关状况的差异, 当样本量 n 增大时, 由于相关系数抽样方差减小, 而且据相关系数抽样分布理论^[6], 平均相关高的场比平均相关低的场其抽样方差减小更为明显。因此, 我们可以定性地推论, 凡气象场平均相关水平愈高, 其 EOFs 随样本增大愈容易达到稳定, 反之, 凡气象场平均相关水平愈低, 其 EOFs 随样本增大愈不容易达到稳定。事实上, 文献[9]中提出的协方差行列式(即广义方差)的抽样矩正是相关矩阵行列式的函数, 这与式(5)定义的 α 范数是一致的, 因此, 上文的推论与文献[9]的提法也是吻合的。联系到作者曾在文献[10]中论证 EOFs 收敛性的影响因素也与场内相关结构有关, 更可推论, 收敛快的 EOFs 在一定的的前提下其稳定性也愈好。

4 相关结构对 EOFs 稳定性影响的数值试验

4.1 不同临界值 ε 对应的 EOFs 谱的可能最大变动

利用式(10)和(11), 可以考察 ε 和 p 值所对应的 EOFs 谱变动情况。例如, 给定不同的 ε , 则由式(11)就可相应地获得大样本(很大 n 值)前提下, R_n 和 R 两者差值的范数最大允许限(即式(11)中 $\nu(R_n - R)$), 又由式(10), 不难给出各种 ε 和 p 组合下的 EOFs 谱的可能最大变动 $S_R(R_n)$ 。在表 1 中, 我们列出一些典型结果, 以便比较不同的 ε 和 p

表 1 不同 ε 和 p 值对应的 EOFs 谱变动可能最大值

ε	p				
	3	5	10	20	30
0.001	0.2289	0.6034	1.1774	1.5887	1.7387
0.01	0.4932	0.9564	1.4823	1.7825	1.8774
0.1	1.0627	1.5157	1.8661	2.0000	2.0272
0.2	1.3389	1.7411	2.0000	2.0705	2.0746
0.3	1.5326	1.8882	2.0828	2.1129	2.1028
0.4	1.6869	2.0000	2.1435	2.1435	2.1231
0.5	1.8171	2.0913	2.1919	2.1675	2.1389
0.6	1.9309	2.1689	2.2322	2.1874	2.1520
0.7	2.0328	2.2369	2.2669	2.2044	2.1631
0.8	2.1255	2.2974	2.2974	2.2191	2.1727
0.9	2.2104	2.3522	2.3246	2.2322	2.1813
1.0	2.2895	2.4022	2.3492	2.2440	2.1889
2.0	2.8845	2.7595	2.5178	2.3232	2.2401
3.0	3.3019	2.9925	2.6220	2.3708	2.2706
4.0	3.6344	3.1698	2.6986	2.4051	2.2925
5.0	3.9149	3.3145	2.7595	2.4321	2.3096
6.0	4.1602	3.4375	2.8102	2.4544	2.3237
7.0	4.3796	3.5452	2.8539	2.4733	2.3356
8.0	4.5789	3.6411	2.8922	2.4899	2.3460
9.0	4.7623	3.7279	2.9265	2.5046	2.3553
10.0	4.9325	3.8073	2.9575	2.5178	2.3636

值对于 EOFs 谱变动的影 响。由表 1 可见,当气象场相关结构的变动很小时(ε 小),其特征值的可能最大变动也很小,相对而言,当相关结构的变动很大时,其特征值的可能最大变动增大。与此同时,由表 1 还可看到,气象场空间站点数 p 越大,相关矩阵扰动所造成的影响也越大,例如,当 $p=3, \varepsilon=0.001$ 时, EOFs 谱改变量仅只为 0.2289,但当同样的 ε 下, $p=30$ 时, EOFs 谱改变量竟高达 1.7387,可见在相关矩阵发生同样扰动的情况下,站点数 p 越大, EOFs 谱所受到的影响也越大。此外,当相关矩阵扰动较大时,站点数 p 的影响随 p 的增大而减小。换言之,站点数 p 越少,相关矩阵扰动所引起的谱变动越明显,而 p 很大时,其造成的影响比 p 很小时要小得多。例如,当 $p=3, \varepsilon=10.0$ 时,谱改变量上限达到 4.9325,而在同样的 p 值下, $\varepsilon=0.1$ 时,谱改变量仅只有 1.0627;相反,当 p 增高到 30 时, $\varepsilon=0.1$ 和 $\varepsilon=10.0$ 其相应的谱改变量差异很小(前者 2.0272,后者 2.3636),可见,气象场的站点数越少,由相关矩阵扰动引起的 EOFs 谱变动的敏感性越高,而站点数越多,则 EOFs 谱变动的敏感性降低。这正如通常我们考察个别测站气象观测值与气象场(多个测站)整体气象观测值两者的变动情况一样,前者对某种扰动的响应一般总是比后者更为敏感。

在式(10)或(11)中, R_n 与 R 差值的范数实质上就是平均相关水平的某种改变量(见式(5)定义),因此,气象场相关结构的变动实际上就可用其平均相关为度量。显然,若 ε 大,则等价于场的平均相关有很大变动幅度,在此情况下,站点数 p 越少, EOFs 展开的特征值(谱)可能变动越大,反之,站点数 p 越多, EOFs 展开的特征值(谱)的可能变动越小。

由此可见,当气象场的站点数 p 较大时,只要平均相关水平变化不大,其 EOFs 特征值(谱)可能最大变动也不会很大。

4.2 给定置信水平 $(1-\alpha)$, 气象场 EOFs 谱的可能变动随样本量 n 的变化

由 Fisher 变换^[6], 计算不同总体相关系数 ρ 值的置信区间, 从而可给出相应样本容量 n (气象场序列长度)下, EOFs 谱变动的估计值。根据 Fisher 变换, 可估算出不同 n 所对应的相关系数置信区间限界以及它们与真值 ρ 的差异。例如, 当 $\rho=0.90$, 可计算得表 2。由表 2 可见, 不同 n 所对应的 $\rho=0.90$ 置信限随 n 增大而缩减。

表 2 不同 n 对应 $\rho=0.90$ 的置信度 $(1-\alpha=0.95)$

界限	n									
	10	20	30	40	50	60	70	80	90	100
上限	0.98	0.96	0.95	0.95	0.94	0.94	0.94	0.94	0.94	0.93
下限	0.62	0.76	0.80	0.82	0.83	0.84	0.84	0.85	0.85	0.85
最大扰动量*	0.28	0.14	0.10	0.08	0.07	0.06	0.06	0.05	0.05	0.05

* 由实际相关置信限与总体相关 ρ 的最大差值的绝对值表示。

由于范数本身的意义可等价于相关矩阵的平均相关水平(见式(5)), 因此, 假定气象场总体的平均相关水平(如表 2)为 $\rho=0.90$, 那么, 可进一步以表 2 中不同 n 所对应的样本相关置信限和最大扰动量代表气象场实际相关水平的随机扰动状况, 由此计算出表 3。显然, 表 3 是不同样本量 n 值所对应的相关矩阵范数改变量上限的一种估计。即由表 2 计算的最大扰动量近似作为相关矩阵中非主对角元素的平均差值, 利用式(5)的定义, 按 $\|R_n - R\|$ 估计范数的可能最大改变量, 这就构成表 3。由表 3 可进一步应用表 1 的结果来估计出 EOFs 谱的可能最大变动。例如, 当 $p=30, n=30$ 时, 据表 2 和表 3 知道它的范

表 3 对应于表 2 的范数改变量上限

n	p				
	3	5	10	20	30
10	0.56	1.12	2.52	5.32	8.12
20	0.28	0.56	1.26	2.66	4.06
30	0.20	0.40	0.90	1.90	2.90
40	0.16	0.32	0.72	1.52	2.32
50	0.14	0.28	0.63	1.33	2.03
60	0.12	0.24	0.54	1.14	1.74
70	0.11	0.22	0.54	1.14	1.45
80	0.10	0.20	0.45	0.95	1.45

数改变量为 2.9, 而从表 1 可查得相对应的 EOFs 谱最大可能改变量约为 2.27 左右。这个例子仅表明高相关水平 ($\rho=0.9$) 的气象场的情况, 因其相关矩阵抽样振动一般都较小, 所以相应的谱变动也不大; 但是, 不难想象, 若气象场总体的平均相关水平较低, 例如某气象场接近于白噪声场, 其抽样方差必然增大, 因而导致相关矩阵的随机扰动必然也增大, 以致其 EOFs 谱的改变量也将增大。例如, 假定某气象场总体的平均相关 $\rho=0.3$, 比

照类似的表 2, 可估计得表 4。由表 4 可见, 气象场的总体平均相关水平如果比较低, 抽样扰动上下限必然加大, 因而使相关矩阵改变量的范数也必然加大。我们也可类似于表 3, 由表 4 估算出相应的范数改变量上限。很显然, 表 4 中相关系数为 $\rho=0.3$, 其置信限幅度远比表 2 中 $\rho=0.9$ 时要大得多, 不难想象其相应的相关阵变动幅度及其范数改变量也较大。这就意味着, 气象场的总体相关水平较低时, 倘若样本场序列较短 (n 小), 且站点数较少 (p 小), 则相关矩阵变动的范数也较大。再据表 1 的规律不难推知, 它们的 EOFs 谱改变量也必将会较大。例如, 同是 $p=5, n=10$ 的气象场, 如为高相关场 (平均相关水平, 以 $\rho=0.90$ 代表), 它的谱改变量仅为 2.3 左右 (指有随机扰动而言), 但若为低相关

表 4 不同 n 对应 $\rho=0.3$ 的置信限 ($1-\alpha=0.95$)

界限	n									
	10	20	30	40	50	60	70	80	90	100
上限	0.7819	0.6555	0.5959	0.5592	0.5338	0.5147	0.4997	0.4876	0.4774	0.4688
下限	-0.4064	-0.1643	-0.0256	-0.0126	0.0236	0.0498	0.0700	0.0861	0.0991	0.1101
最大扰动量*	0.7064	0.4643	0.3256	0.3126	0.2764	0.2502	0.2300	0.2139	0.2009	0.1899

* 说明同表 2。

场 (平均相关水平, 以 $\rho=0.3$ 代表), 它的谱改变量按式 (10) 计算, 可推得为 4.39。相比之下, 可见: 气象场总体相关小, 而取样本场序列又较短时, 其 EOFs 特征值 (谱) 的稳定性必然较差, 反之则不然。并且, 随着 p 的增大或 n 的增大, 相关结构因随机扰动而引起变动的影 响将会减小, 从而使 EOFs 谱的稳定性少受影响。这一点不难从表 1 中得到验证。

5 计算实例

选取长江中下游地区 ($29^{\circ}\text{N}-33^{\circ}\text{N}, 111^{\circ}\text{E}-122^{\circ}\text{E}$) 逐月平均气温和降水量距平场序列进行 EOFs 展开, 其站点数 $p=14$, 样本总量 $n=412$ 。分别计算不同 n 值所对应的气温距平场和降水距平场的相关矩阵改变量的范数变化值。在本例中, 我们近似地以大样本 $n=412$ 所对应的相关矩阵 R 作为总体相关矩阵, 以下凡涉及 R 和 R_n 时均按此近似处理。

图 1 a, b 分别给出气温、降水场 $\|R_n - R\|$ 的两种范数 (α, F 范数)^[5] 随样本量 n 大小的变化趋势。显然, 图 1 表明, 式 (11) 所定义的均方意义下近似收敛性或稳定性是有实际基础的。进一步利用图 2 a, b 绘制气温场和降水场前三个特征值随样本量 n 的变化 (图中纵坐标以 $\lambda_i / \sum_i \lambda_i$ 百分比作为度量)。由图可见, 在一定的样本量 n 下, 特征值 $\lambda_1, \lambda_2, \lambda_3$ 各自趋于稳定状态, 但随着特征值序号的增大, 其达到稳定的临界样本量 $n \geq N$ 略有变动, 似有后移的趋势。显然, 从图 1 与图 2 可见, $\|R_n - R\|_{\alpha}$ 及 $\|R_n - R\|_F$ 与特征值稳定域基本上是互相匹配的。尤其是 λ_1 与范数的变化配合最佳。如在图 1 a 中, 气温场相关矩阵差值的范数大约在 $n > 45$ 时渐近稳定, 而相应特征值 λ_1 也大致如此 (见图

2a);而在图1b中,降水场相关矩阵差值的范数则大约在 $n > 100$ 时渐近稳定,而相应特征值 λ_1 也大致如此(见图2b)。综上所述,对任一气象场 EOFs 展开来说,假如仅考虑抽样引起的随机扰动,则当样本量(场的序列长度) n 超过某一定临界值时, EOFs 必将趋于稳定。当然,由于不同的气象场其随机扰动的程度有所不同,相关结构也各有差异,因而它们各自都可能有不同的样本量临界值。如前所证,这一临界值 N 的大小往往取决于气象场总体相关水平的高低(即相关矩阵范数)。例如,气温场相关性优于降水场,其范数大于降水场,所以气温场相关阵差值的范数较之降水场相关阵差值的范数更易达到稳定。图1和图2中的曲线所反映的规律就是典型例证。相反,当样本场序列长度 n 达不到某一临界值 N 时, EOFs 展开具有不稳定状况,因此,当人们借助于 EOFs 展开来研究气象场序列的插补、延长或外推预报时,必须注意保持足够的样本量 n , 以免出现不稳定的 EOFs 展开。

就一个气象场而言, EOFs 各不同特征值及其特征向量达到某稳定值的样本量 $n \geq N$ (临界样本量)的大小并不相同。例如图3绘制了各

不同 n 值所对应的特征向量与总体特征向量(用大样本 $n = 412$ 的特征向量近似代表)之间的相似系数随样本量 n 的变化趋势。由图3a可见,当 $n = N = 42$ 时,第1特征向量即趋于稳定,而由图3b、3c可见,当 $n = N = 70$ 时,第2特征向量趋于稳定,当 $n = N = 200$ 左右时,第3特征向量才趋于稳定。显然,这一特点与图2a的相应特征值的稳定样本量也是相当吻合的。关于这一点,又恰恰与文献[4]对旱涝指数 EOFs 展开的稳定性的探讨所证实的规律相一致。

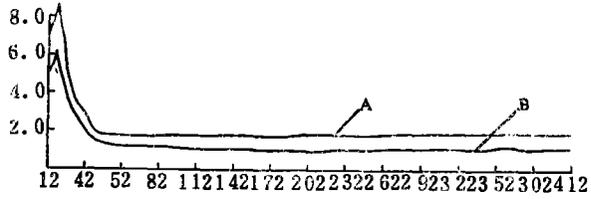


图 1a 气温场范数 $\nu(R_n - R)$ 随样本量 n 的变化 (A 为 $\|R_n - R\|_F$, B 为 $\|R_n - R\|_0$)

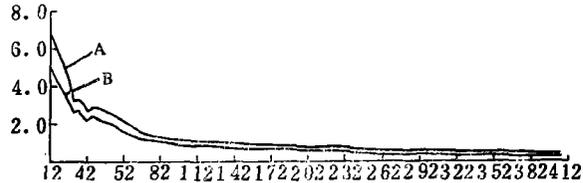


图 1b 降水场范数 $\nu(R_n - R)$ 随样本量 n 的变化 (说明同图 1a)

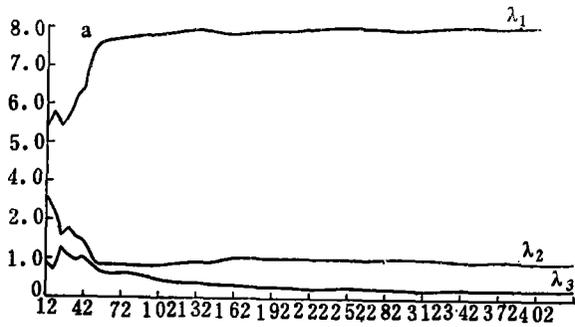


图 2a 气温场 EOFs 前三个特征值随样本量 n 的变化

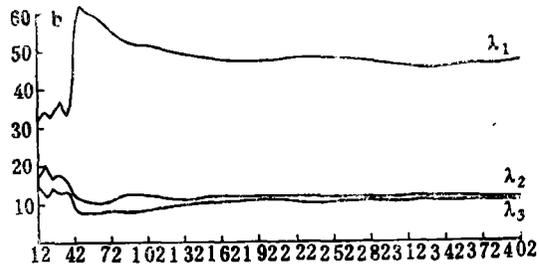


图 2b 降水场 EOFs 的前三个特征值随样本量 n 的变化

6 结 论

1. 利用矩阵范数度量相关矩阵稳定性, 间接推估 EOFs 稳定性, 从理论上证明是行之有效的。

2. 气象场相关结构及其扰动对 EOFs 稳定性的影响是多方面的。(i) 当站点数 p 较大, 只要场的平均相关水平变化不大, 其特征值(谱)可能最大变动不会很大, 但当 p 较小时, 相关矩阵随机扰动引起的 EOFs 谱变动的敏感性增高; (ii) 一般说来, 平均相关水平越高, 相关矩阵随机扰动越小, 其 EOFs 越近于稳定, 而平均相关水平越低, 相关矩阵随机扰动越大, 其 EOFs 越不易稳定; (iii) 上述规律又直接受样本量即气象场序列长度 n 和站点数 p 的影响。

3. 数值试验和实例计算都证明上述结论。而实例尤其表明, 不同气象场相关结构下达到稳定 EOFs 展开的样本长度 n 的临界值不同。当样本达不到某一临界值 N 时 ($n \geq N$ 时, EOFs 稳定), 必须警惕 EOFs 展开有可能不是稳定的。

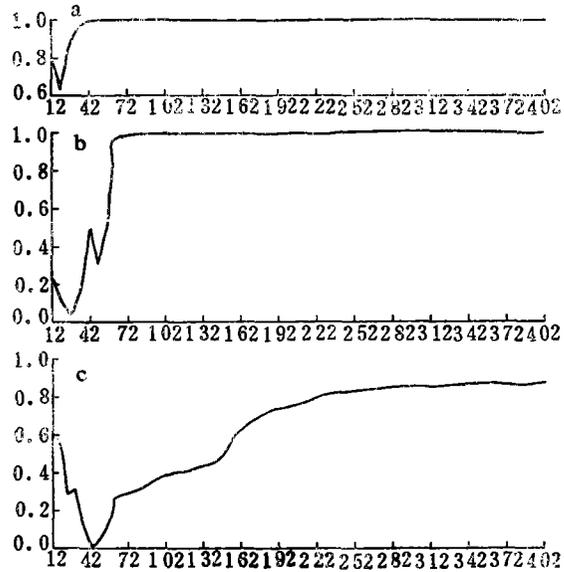


图 3 a 气温场第 1 特征向量与总体第 1 特征向量相似系数随 n 的变化, b 气温场第 2 特征向量与总体第 2 特征向量相似系数随 n 的变化, c 气温场第 3 特征向量与总体第 3 特征向量相似系数随 n 的变化

参 考 文 献

- [1] North G, Bell T L, Cahalan R F, Moeng F J. Sampling errors in the estimation of EOFs. *Mon Wea Rev*, 1982, 110:699—706.
- [2] 章基嘉, 孙照渤, 陈松军. 对自然正交函数稳定性条件的讨论. *气象学报*, 1981, 39(1):82—89.
- [3] Gray B M. On the stability of temperature eigenvector patterns, Sixth conference on Prob. and statis. in *Atm. Sci.*, 1979.
- [4] 屠其璜. 旱涝指数经验正交展开的稳定性分析. *南京气象学院学报*, 1990, 13(3):266—273.
- [5] 孙继广. 矩阵扰动分析. 科学出版社, 1987.
- [6] 么枕生. 气候统计学基础. 科学出版社, 1984.
- [7] 复旦大学. 概率论(第三册). 人民教育出版社, 1981.
- [8] Muirherd R J. Aspects of multivariate statistical theory. John Wiley & Sons INC, 1982.
- [9] 肯德尔 M. 多元分析. 中科院计算中心概率统计组译. 科学出版社, 1983.
- [10] 丁裕国, 吴 息. 经验正交函数展开气象场收敛性的研究. *热带气象*, 1988, 4(4):316—326.

THE EFFECT OF THE CORRELATION STRUCTURE OF METEOROLOGICAL FIELDS ON STABILITY OF EXPANSION OF EOFs

Ding Yuguo Jiang Zhihong

(Nanjing Institute of Meteorology, Nanjing 210044)

Abstract

According to the matrix disturbance theory, a stability index of stochastic disturbance of meteorological field series is presented by using the norm of matrix and then the stability index is applied to estimate indirectly the stability of EOFs. It is shown that the higher the correlation of the meteorological field series, the smaller the sample it needs to reach stability correlation structure by means of the theoretical derivation and numerical experiment and case calculation. As soon as a meteorological field series satisfied above, the more stable EOFs is obtained and vice versa. The regularity depends directly upon the sample length n and the number of stations p . For different meteorological field series there are different criterion values of sample length which satisfies the stable EOFs, thus it must be on guard against the instability of EOFs when expansion.

Key words: Empirical orthogonal function, Matrix disturbance, Sampling distribution, Meteorological field.