

经验正交函数在两个气象场相关分析中的应用*

徐瑞珍 张先恭

(中央气象局气象科学研究所)

五十年代以来,经验正交函数的应用愈来愈广泛。气象方面,大多用于气象场的定量描述。许多人用特征向量表示个别气象场的特征,也有人用组合特征向量表示多个气象场的特征^[1],这种方法可以大大减少信息量,给天气气候的分析和预报带来方便。

我们的兴趣不在于气象场本身的时间变化和空间分布特征,而在于成对气象场之间相关系数场集合的空间结构,希望了解每一个气象场对相关场集合的贡献以及全部相关的分布状况。为此,我们参考文献[2]、[3]将经验正交函数直接应用于两个气象场的相关系数场的分析,分别使两个气象场对它们的相关系数场贡献显著的部分最大限度地反映出来,而显著地减少相关系数场的自由度。对应于每一个特征根,均有一组特征向量和相应的振幅系数,分别依赖于这两个要素场的空间特征,它们的线性组合为相关场集合某种精度的近似结果。

我们用这种方法分析了北半球500毫巴6—8月平均高度场和同时期我国东部地区地面温度场的关系。作为初步的试验,得到了一些有意义的结果。

1. 分析方法

假定两个气象要素场的集合分别为 $x(t, s)$, $y(t, g)$, 它们的相关系数场的集合为 $R(s, g)$ 。 t 为时间, $t=1, 2, \dots, T$, T 为序列长度。 s, g 分别为两个场站点或格点的序号。 $s=1, 2, \dots, N$, $g=1, 2, \dots, M$ 。 N, M 分别为两个场的站点数和格点数。

我们的目的是要求得特征向量 $X_i(s)$ 和振幅系数 $Y_i(g)$,使

$$R(s, g) = \sum_{i=1}^N Y_i(g) X_i(s) \quad (1)$$

这就是相关场按特征向量 $X_i(s)$ 的展开式。显然, $X_i(s)$ 随第一个场的空间而变化, $Y_i(g)$ 随第二个场的空间而变化。 $X_i(s)$ 称为相关场集合的特征函数。

特征向量和相应的振幅系数是通过交叉相关的平均乘积矩阵求特征根的方法得来的。我们感兴趣的是平均乘积而不是协方差,因为我们期望得到的是全部相关的分布情况而不是平均相关的变化。

以下用矩阵形式表示相关阵、特征向量和振幅系数。

将两个多维时间序列 $x(t, s)$, $y(t, g)$ 标准化以后,求得它们的相关阵 $R_{s, g}$, 其平均乘积矩阵

* 本文于1980年4月8日收到, 1980年10月20日收到修改稿。

$$A_{ij} = \frac{1}{M} \sum_{g=1}^M R_{ig} R_{jg} \quad i, j = \overline{1, N} \quad (2)$$

显然, A 是对称矩阵, 只有对称矩阵的特征向量才形成正交系。

矩阵方程

$$(A - \lambda I)X = 0 \quad (3)$$

I 为单位矩阵, λ 和 X 分别为特征根和特征向量。由(3)式可以解得 N 个特征向量 $X_{i\sigma}$ 和 N 个特征根 λ_i 。 λ_i 由大到小排列, 它说明了均方相关系数的大小。

由此可得振幅系数矩阵

$$Y_{i\sigma} = \sum R_{i\sigma} X_{\sigma i} \quad (4)$$

和交叉相关阵

$$R_{i\sigma} = \sum_{i=1}^n Y_{i\sigma} X_{\sigma i} \quad (5)$$

如果将(5)式矩阵形式改写为函数形式

$$R(s, \sigma) = \sum_{i=1}^N Y_i(\sigma) X_i(s)$$

这便是我们所要求的(1)式。

相关系数就这样分成了成对的空间型式。 $Y_i(\sigma)$ 和 $X_i(s)$ 是成对相配的, 某一对空间型相应的特征根表示了该空间型对整个相关场的贡献。相应于最大特征根的这一对空间型是均方相关的最大部分。特征根小, 对均方相关的贡献相应减少。由(3)和(5)式可以得到特征根的表达式

$$\lambda_i = \left(\sum_{\sigma=1}^M Y_{i\sigma}^2 \right) / M \quad (6)$$

由此, 特征根等于振幅系数的平方在 M 个格点上的平均值。由(5)和(6)式得均方根相关系数

$$(\overline{R_{i\sigma}^2})^{\frac{1}{2}} = \left[\left(\sum_{i=1}^N \lambda_i \right) / N \right]^{\frac{1}{2}}$$

因此, 仅特征根就完全反映了均方相关的大小。

由于经验正交函数具有收敛快的特点, λ_i 随 i 的增加递减很快, (1)式的近似结果只要极少数项就可以了。定义个别函数 $I_i = \lambda_i / \sum_{i=1}^N \lambda_i$, 它表示第 i 对空间型的权重。对任一特征根 λ_i 所对应的空间型均有一定的权重, 但相关系数的大小和符号, 由它所对应的振幅系数和特征向量乘积的大小和符号决定。分解相关场集合的过程就是逐次确定两个要素场最常见的相关型的过程, 也是确定这种相关型在这个集合中的权重的过程。如果权重 I_i 太小, 它相应的振幅系数和特征向量的乘积的大小必然达不到一定信度要求, 自然就不予考虑。

这个方法的基本优点在于只要少数项就能反映多点场的主要信息, 且这少数项在某

种意义上最佳地反映了两个气象场的线性相关场的主要特征。

2. 计算结果

作为对以上方法的一个试验, 我们取我国东部地区 1951—1978 年夏季(6—8 月)地面平均气温与同期北半球 500 毫巴高度场的相关场作一尝试, 高度场范围为 10° — 80° N。格点间距从 10° — 50° N, 每 10 个经纬度取一点, 60° N 纬圈上每隔 20 个经度取一点, 在 70° N 和 80° N 上, 格点间距为 40 个经度, 一共 216 个点。温度取 41 个站, 均匀分布在我国东部地区(如图 1 所示)。

下表为这两个要素场的相关场分解后的前六个个别函数和它们的累积函数($\sum I_i$)。

i	1	2	3	4	5	6
I_i	0.569	0.134	0.091	0.064	0.035	0.023
$\sum I_i$		0.703	0.794	0.858	0.894	0.917

累积函数可以表示精度。由此表可见, 前面少数几个特征根相应的振幅系数和特征向量乘积的线性组合就足以表示相关场集合分解的近似结果。其中又以第一个特征根所相应的部分(即第一对空间型)占绝对优势。相对第一对空间型来说, 第二、第三、……均不必予以考虑了。根据 Craddock、Colgate(1974)^[4]和 Farmer^[5]的分析, 认为前几个函数决非随机的, 相应的振幅系数和特征向量显然也不是随机的。这里, 个别函数衰减得很快, 第一个个别函数 0.569, 而第二个仅 0.134, 其它更小。所以, 仅第一个个别函数所对应的两个气象场的相关型就可以反映这两个场的关系了。这正是我们所期望的。一般最好不超过两对相关型, 也就是个别函数衰减越快越好。

为了分析两个气象场之间的相关型, 对任何给定信度的相关系数水平, 可以调整特征向量场和振幅系数场上满足信度要求的边界值, 来平衡每一个场上显著高相关区范围的大小。

我们取信度为 5%, 自由度为 $28-2=26$, 查表得相关系数为 0.37。在特征向量场和振幅系数场上, 分别可以确定两组边界值(见图 1、2)。第一组 $X(s)$ 场 0.203 与 $Y(g)$ 场 1.82 相配; 第二组 $X(s)$ 场 0.161 与 $Y(g)$ 场 2.30 相配, 这里 $0.203 \times 1.82 \doteq 0.161 \times 2.30 \doteq 0.37$ 。边界值的大小是任意确定的, 但是 $X(s)$ 场和 $Y(g)$ 场上对应边界值的乘积必须等于给定信度的相关系数, 因此, 它们是互相制约的。在一定的显著性水平下, 扩大 $Y(g)$ 场满足信度的范围, 必然要缩小 $X(s)$ 场上相应的范围, 反之亦然。它们可以相互调整。在每一对特征向量值和振幅系数值所划分的范围内, 信度均超过 5%。分析时, 这两个场的图形结构需要成对地考虑。边界值太多会给分析带来困难, 一般用两组似乎已足够了。

由于第二型的权重太小, 这里仅取第一个型来分析。该取几个型应由计算结果来定。一般可能是 1—2 个。如果权重(即个别函数)太小, 在特征向量场和振幅系数场上均很难找到满足一定信度要求的高相关区, 当然也就没有意义了。

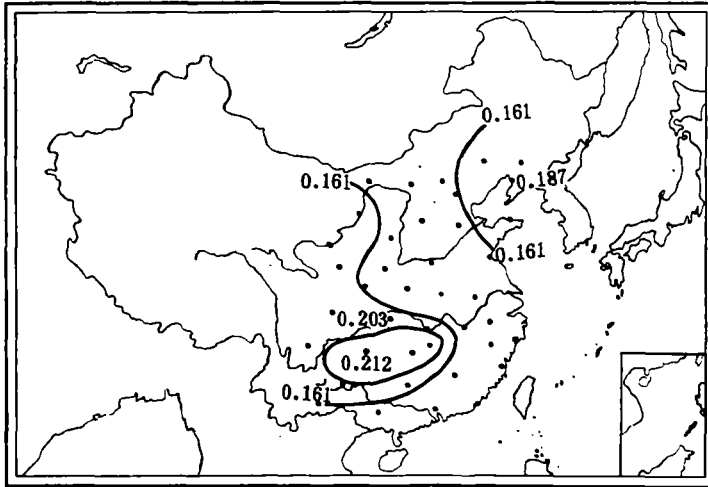


图 1 特征向量 $X(s)$ (图中黑点为站点分布)

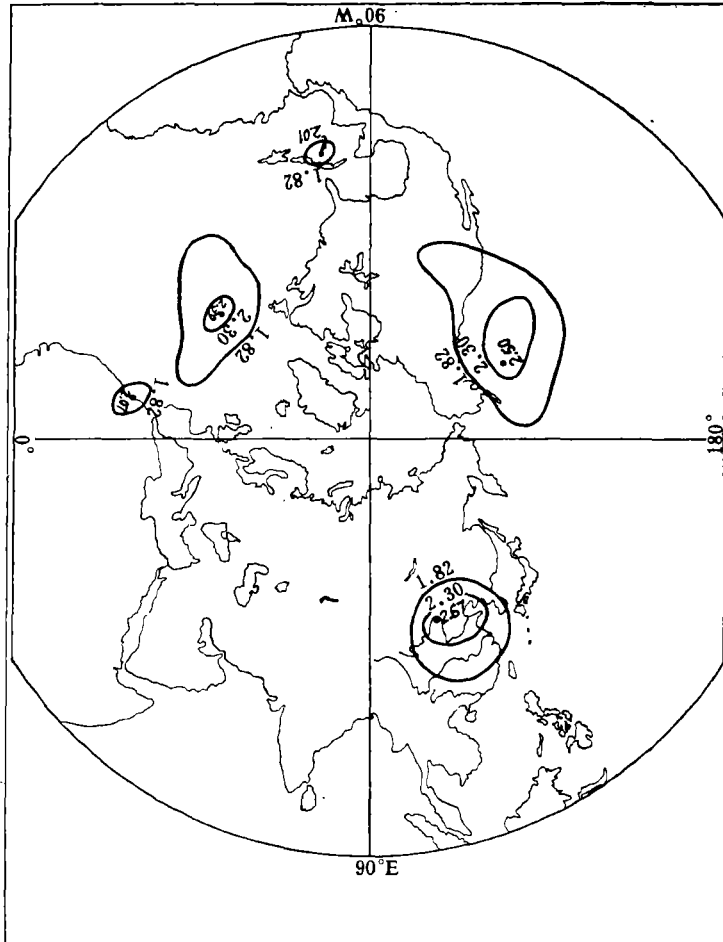


图 2 振幅系数 $Y(s)$

(图中黑点为最大值所在格点; $Y \times X =$ 相关系数(信度为 5%); $1.82 \times 0.203 = 0.37$; $2.30 \times 0.161 = 0.37$)

图 1 和图 2 边界值以内的区域相关系数的信度超过 5%, 区域内标明的数值为该区域内的最大值, 即高相关中心所在处。

3. 讨 论

在分析讨论以上结果以前, 必须强调两点:

1. 相关系数的大小和符号依赖于 Y 场格点和 X 场站点数值的乘积;
2. 因为相关系数只是线性关系的一种量度, 达到一定信度要求即表示一个极端的气压异常相应于一个极端的温度异常。

我们用北半球格点上的振幅系数型和我国东部地区特征向量型的搭配来说明 500 毫巴高度场——温度场的关系。图 1 为特征向量场, 图 2 为振幅系数场。

图 1 有两个超过信度的地区, 分别在东北和西南; 图 2 主要有三个中心: 分别在我国东部(110° — 120° E, 30° — 40° N)、东太平洋和大西洋; 这两张图说明, 这三个地区夏季 500 毫巴高度的变化与我国东北和西南地区的温度变化有明显的关联, 数值均为正, 都是正相关。东亚、东太平洋和大西洋高度上升(或下降), 伴随着我国东北和西南地区温度的升高(或降低)。图 2 振幅系数场上, 三个中心都在 40° N 附近, 500 毫巴等压面上是西风急流所在的平均纬度, 在海平面图上, 正是高压活动频繁的地区。夏季, 北半球海上副热带高压(以下简称副高)位置偏北, 西风偏强, 偏南西风偏弱。所以 40° N 这个纬度带上, 高度的变化对天气的反映较敏感。

有趣的是, 东亚地区的高相关区不象其它两个那样在海上, 而是在大陆上。有人在分析 100 毫巴环流与我国天气的关系时, 也发现这个地区(110° — 120° E, 30° — 40° N)所反映的环流变化与我国天气关系明显, 有人称之为“震源”^[6]。在对流层上部 100 毫巴等压面上, 这地区正是西太平洋副高和南亚高压东西摆动相重叠或分离的地方。南亚高压维持东部型, 西太平洋副高西伸加强; 南亚高压维持西部型, 西太平洋副高东退减弱^[7]。这个所谓震源既可能处在西太平洋副高的西端, 又可能处在南亚高压的东端。该处高度的抬升(或降低)意味着南亚高压东伸(或西退)和西太平洋副高西进(或东退)以及它们的增强(或减弱)。就这个意义而言, 该地区高度的变化可能反映了直接影响我国天气气候的这两个环流系统的位置和强度的变化。

100 毫巴上的这个震源与 500 毫巴有什么关系呢? 朱福康等人^[8]在研究夏季 100 毫巴的环流与 500 毫巴环流的关系时, 发现这两个等压面上的环流在这个地区存在着同时性变化。他们将 100 毫巴和 500 毫巴对应格点的高度求相关, 得出两层等压面高度的同时相关图。东亚出现了显著的正相关区, 其中心恰与图 2 吻合, 就是说, 这两层等压面上的震源有对应关系。这说明, 在东亚, 100 毫巴和 500 毫巴上的环流与我国天气有明显关联的地理区域是一致的。

图 1 有两个高值区, 一个在东北(最大值在营口; 一个在西南(最大值在遵义)。这两个高值区与上面所述的所谓震源相配合表明: 南亚高压维持西部型, 高压中心位于我国西部, 与此同时, 往往西太平洋副高可能向东南退缩, 110° — 120° E, 30° — 40° N 这一区域的高度随之下降, 我国东北和西南地区的温度就偏低。反之, 南亚高压维持东部型, 高压中心移至我国东部平原上, 西太平洋副高可能北进西伸, 110° — 120° E, 30° — 40° N

这一区域的高度抬升,这种环流形势非常有利于我国东北、西南甚至西北地区出现高温天气。近年来,在东北夏季低温冷害的预报中,也注意到了南亚高压和西太平洋副高的作用^[9]。

值得研究的是,为什么东北、西南地区的温度还与东北太平洋和大西洋上的高度有关,如何解释这种遥相关及其机理呢?我们认为,与其说我国东部地区的温度场与东北太平洋和大西洋上的高度有关,还不如说与我国温度场有密切关系的我国东部(110°—120°E,30°—40°N)的高度与东北太平洋和大西洋的高度有关。也就是说,对流层中部西太平洋副高的活动与东太平洋和大西洋上环流系统的变化可能有同步性。图2这三个中心的位置,基本上是等距的(间隔110—120个经度)。这种半球性的波动,似乎说明大气象驻波,在这三个区域内存在着半球性的相互作用,它们往往是同时加强或同时减弱的。环流变化的这种同步性是否可以说明北半球各大陆温度场的变化也可能具有某种同步性呢?

以上只是我们应用两个气象场线性相关分析技术的一个例子。它说明可以用经验正交函数的分析方法来分离两个气象场的高相关区,屏弃相关性不显著的区域。这对于分析任何两个场任意后延的相互关系是很好的工具。它可以客观地最大限度地游离出成对气象场相互关系密切的地理区域。这对于天气、气候的分析和预报因子的选择等诊断性研究是很有帮助的。

全部计算是在108机上完成的。由于机器内存的限制,高度场从576个格点¹⁾中取216个,温度场从160个站点¹⁾中取41个。尽管这样,整个计算分两步进行才能实现。资料工作量很大。如果使用内存量足够大的机器,这种方法可以很方便地解决成对气象场的相关问题。如果需要讨论多个场的关系,可以通过两两相关的步骤解决。不过这些均限于线性关系。

参 考 文 献

- [1] Kutzbach, J. E., *J. Appl. Meteor.*, 6, 791—802, 1967.
- [2] Prohaska, J. T., *Mon. Wea. Rev.* 104, 1345—1354, 1976.
- [3] Баров, Н. А., *Труды ЧИПа, вып.* 74, 1959.
- [4] Craddock, J. M. and M. G. Colgate, *Bull. Inst. Math. Appl.*, 10, 152—160, 1974.
- [5] Farmer, S. A., *The Statistician*, 20 63—72, 1971.
- [6] 朱福康等,青藏高原气象论文集, 148—157, 1977.
- [7] 陶诗言、朱福康, *气象学报*, 34 385—395, 1964.
- [8] 朱福康等,南亚高压,科学出版社, 1980.
- [9] 东北低温科研协作组, *气象学报*, 37 44—58, 1979.

1) 中央气象台长期组的资料纸带。