

复经验正交函数 (CEOF) 方法及其在天气过程诊断分析中的应用

黄嘉佑

(北京大学地球物理系)

提 要

复经验正交函数 (CEOF) 是在1983年由Barnett提出的。它在长期天气过程分析中是一个有用的工具。本文综合介绍和评述这一方法在这个领域上的应用。

一、前 言

复经验正交函数 (Complex Empirical Orthogonal Function, 简记为 CEOF) 是 Barnett (1983) [1] 提出的一种针对气象要素场的分析方法。它是一种能从要素场的时间变化中, 识别空间尺度行波和驻波的方法。一般在天气过程的诊断分析中, 揭示要素场的不同尺度的空间波动和时间振动的现象, 大多使用时空交叉谱 (或称时空谱) [2] 的方法。但是, 这一方法一般适用于全球或半球范围的要素场, 对于局部地区或不规则网格点分布的范围则十分困难。而且, 当非平稳现象和资料记录较短时, 对谱的估计会带来较大的误差。一般传统的经验正交函数 (EOF) 虽然能解决这些困难, 但它只能揭露空间波动的驻波振动现象, 而对行波却无能为力。复经验正交函数方法可以兼顾上述两种分析方法的优点, 因而近年来被许多气象学者用来作为天气过程诊断分析的新工具之一, 并受到广泛的重视 [4]—[12]。

由于经验正交函数与主分量分析 (PCA) 方法在原理上是一致的, 因而称呼上可以混同。类似地, 复经验正交函数也有时被叫作复主分量分析 (CPCA)。

二、序列的Hibert变换

主分量分析中的主分量, 是从气象要素场以空间点为变量的协方差阵或相关阵进行退化值分解而得到的。但是由于矩阵中的元素是落后零时刻的协方差或相关系数, 因此由该矩阵所提取的主分量只能反映要素场空间变量的同时关系。Wallace和Dickinson [3] 曾经提出用如下的变换把实时间序列变换为复时间序列

$$W(t) = x(t) + \frac{i}{2}[x(t+1) - x(t-1)] \quad (1)$$

式中 $x(t)$ 为原来时间序列, $i = \sqrt{-1}$ 。在新时间序列中的虚部就包含有不同时刻的信息。后来 Rasmusson 等人 [4] 及 Barnett [1] 提出可以对原来实的时间序列作 Hibert 变换来达到这一目的。如果记任一气象要素场中第 j 个空间点 (或网格点) 变量的时间序列为 $u_j(t)$, 用 Hibert 变换把这一时序变为它的虚部, 记为

$$\hat{u}_j(t) = \sum_{l=-\infty}^{\infty} u_j(t-l)h(l) \quad (2)$$

其中

$$h(j) = \begin{cases} \frac{2}{\pi l} \sin^2\left(\frac{\pi l}{2}\right) & l \neq 0 \\ 0 & l = 0 \end{cases} \quad (3)$$

这一过程实际上是一滤波过程, 过滤后的序列 $\hat{u}_j(t)$ 的频率响应振幅大小为 1, 而相差为 90° 。因而实际上这一变换能揭示同一频率中相差为 90° 的振动情况。由这两个时序构成一个新的复时间序列

$$U_j(t) = u_j(t) + i\hat{u}_j(t) \quad (4)$$

可以揭露原时序的不同时刻变化信息。

从频域方面来看, 据傅里叶展开可以把这一实时间序列 $u_j(t)$ 展成

$$u_j(t) = \sum_{\omega} [a_j(\omega)\cos\omega t + b_j(\omega)\sin\omega t] \quad (5)$$

式中 ω 为圆频率, $a_j(\omega)$ 及 $b_j(\omega)$ 为傅氏系数。如果展成它的复谱形式, 则有新的复数序列, 表示为

$$U_j(t) = \sum_{\omega} c_j(\omega) e^{-i\omega t} \quad (6)$$

其中

$$c_j(\omega) = a_j(\omega) + ib_j(\omega)$$

亦可表示为

$$\begin{aligned} U_j(t) &= \sum_{\omega} [a_j(\omega) + ib_j(\omega)] (\cos \omega t \\ &\quad - i \sin \omega t) \\ &= \sum_{\omega} \{ [a_j(\omega) \cos \omega t + b_j(\omega) \sin \omega t] \\ &\quad + i [b_j(\omega) \cos \omega t - a_j(\omega) \sin \omega t] \} \end{aligned}$$

$U_j(t)$ 的虚数部分可表示为

$$\hat{u}_j(t) = \sum_{\omega} [b_j(\omega) \cos \omega t - a_j(\omega) \sin \omega t] \quad (7)$$

显然易见，它与 (5) 式相差 90° 。所以

(7) 式是 Hilbert 变换在频域上的表现形式。由此，实时序 $u_j(t)$ 的 Hilbert 变换可以通过时域上的 (2) 式及频域上的 (7) 式实现。后一种方法是通过求傅氏系数 $a_j(\omega)$, $b_j(\omega)$ 基础上得到的，又称为直接方法，前一种则称为卷积方法，在该方法中过滤器的长度对实际计算不能为无穷大，故一般取截断长度 L 作为近似，即 (2) 式写为

$$\hat{u}_j(t) = \sum_{l=-L}^L u_j(t-l) h(l) \quad (8)$$

实际计算 L 取 7—25，已能很好地满足要求。

在长期天气过程分析中，有时我们只对时序某一频率（或周期）段的振动现象感兴趣。这时可以对原时序作某一频带的带通滤波，然后再作滤波后的 Hilbert 变换。这一过程在频域上变换亦可并行进行，即把时序 $u_j(t)$ 的变换表示为

$$\hat{u}_j(t) = \sum_{\omega} W(\omega) [b_j(\omega) \cos \omega t - a_j(\omega) \sin \omega t] \quad (9)$$

其中 $W(\omega)$ 为在频率 ω 上的过滤权重，如果对所有频率 $W(\omega)$ 取为 1，则 (9) 式即是

(7) 式。因此可令在滤波频带范围 $\omega_0 \pm \delta\omega$ 上 $W(\omega)$ 为 1，而不在 $\omega_0 \pm \delta\omega$ 上为 0，则可过滤出在 ω_0 附近的振动序列的 Hilbert 变换。当然，这一过滤过程也必须对原序列的实部作同样的操作。

三、复相关系数

空间点 j 及 k 上变量的复时间序列的相关系数称为复相关系数，表示为

$$r = \frac{1}{T} \int_{-T/2}^{T/2} U_j(t) U_k^*(t) dt \quad (10)$$

式中 T 为时序所定义的时间， $U_k^*(t)$ 为 $U_k(t)$

的复共轭。用 (4) 式代入上式，则

$$\begin{aligned} r &= \frac{1}{T} \int_{-T/2}^{T/2} \{ [u_j(t) u_k(t) + \hat{u}_j(t) \hat{u}_k(t)] \\ &\quad + i [u_k(t) \hat{u}_j(t) - u_j(t) \hat{u}_k(t)] \} dt \end{aligned} \quad (11)$$

$$\text{若令 } x = \frac{1}{T} \int_{-T/2}^{T/2} [u_j(t) u_k(t) + \hat{u}_j(t) \hat{u}_k(t)] dt$$

$$\begin{aligned} y &= \frac{1}{T} \int_{-T/2}^{T/2} [u_k(t) \hat{u}_j(t) \\ &\quad - u_j(t) \hat{u}_k(t)] dt \end{aligned}$$

则复相关系数可表示为

$$r = x + iy \quad (12)$$

或

$$r = A e^{i\varphi} \quad (13)$$

其中 $A = (x^2 + y^2)^{1/2}$, $\varphi = \pi^{-1} \tan^{-1}(y/x)$ 。 x 表示序列相关系数的实部， y 则表示其虚部。 A 反映相关的密切程度， φ 在 -1 到 $+1$ 之间变化，它反映相关时的位相变化。

在频域上复相关系数的表示可用 j 及 k 空间点频域表达式 (6) 式代入 (10) 式得到

$$r = \sum C_j(\omega) C_k^*(\omega) \quad (14)$$

若令

$$\begin{aligned} S_{jk}(\omega) &= C_j(\omega) C_k^*(\omega) \\ &= [a_j(\omega) a_k(\omega) + b_j(\omega) b_k(\omega)] \\ &\quad + i [a_k(\omega) b_j(\omega) - a_j(\omega) b_k(\omega)] \end{aligned} \quad (15)$$

显而易见， $S_{jk}(\omega)$ 为两个时序的交叉谱，因而在任两空间点变量的复相关系数中，包含着它们之间交叉谱的信息。

四、复相关矩阵的分解

由要素场所包含的各空间点变量两两的复相关系数组成的矩阵，称为复相关矩阵，记为 $R = (r_{ij})$ 。它是在复数意义下的对称阵，即 $R^* = R$ 。 R^* 为 R 的复共轭阵，“ $'$ ”表示矩阵的转置。满足这种关系的矩阵又称为 Hermit 矩阵。如果 R 是 m （空间点数）阶的 Hermit 矩阵，则一定存在一个 $m \times m$ 酉矩阵 V ，使

$$V^* R V = D \quad (16)$$

式中 D 是实对角阵，阵中对角线上的元素称为 R 的特征值，对应 V 阵中的列向量为其（复）特征向量。不同的复特征向量之间是正交的。即

$$V^* V = I$$

I 为单位阵。求 R 阵的分解可在一般的线性代数计算程序包中得到。由此即可得到它的（实）特征值及其对应的（复）特征向量。

类似主分量分析，对应不同的 R 的特征值的复主分量可得到。例如对第 k 个复主分量，可由各空间点变量的复时间序列与第 k 个复特征向量求得，即

$$P_k(t) = \sum_{j=1}^m U_j(t) V_k(j) \quad (17)$$

式中 $V_k(j)$ 表示第 k 个复特征向量在第 j 个空间点（即分量）的值。类似地，复主分量具有如下性质：

$$\langle P_1(t) P_k^*(t) \rangle_t = \begin{cases} \lambda_k & l=k \\ 0 & l \neq k \end{cases} \quad (18)$$

式中 $\langle \rangle_t$ 表示对序列的平均值， λ_k 为 R 阵的第 k 个特征值。因此，第 k 个复主分量的解释方差可表示为

$$EV_k = \lambda_k / \sum_{j=1}^m \lambda_j \quad (19)$$

由于得到的主分量及特征向量是复数，因而可以类似地定义相应的振幅函数和位相函数。由复特征向量定义的称为空间点 j 的位相函数为

$$\theta_k(j) = \arctan \left[\frac{I_m V_k(j)}{R_e V_k(j)} \right] \quad (20)$$

式中 I_m 及 R_e 表示取复数的虚部与实部，下标 k 表示第 k 个量。对应的空间振幅函数定义为

$$S_k(j) = [V_k(j) V_k^*(j)]^{1/2} \quad (21)$$

对于由第 k 个主分量所引出的时间位相函数及时间振幅函数分别定义为

$$\theta_k(t) = \arctan \left[\frac{I_m P_k(t)}{R_e P_k(t)} \right] \quad (22)$$

$$S_k(t) = [P_k(t) P_k^*(t)]^{1/2} \quad (23)$$

五、比较

为了说明 CEOF 方法的优越性，Trenberth^[5] 给出一个一维行波振动的简单例子。这一行波可表示为

$$Z = A \sin(kx - \omega t) \quad (24)$$

式中 A 为振幅， k 为空间波数， ω 为时间振动频率，显然，上式亦可写为

$$Z = A \cos \omega t \sin kx - A \sin \omega t \cos kx \quad (25)$$

按照传统的 EOF 方法，把随空间及时间变化的 $Z(x, t)$ 函数分解为只随时间变化和只随空间变化的时间函数及空间函数部分，那么分解结果会有两个特征值，其对应的两个特征向量与 $\sin kx$ 及 $\cos kx$ 成比例，得到的两个主分量分别与 $A \cos \omega t$ 及 $-A \sin \omega t$ 成比例。但若用复时空序列表示，(25) 式可表为

$$Z = (\sin kx - i \cos kx) (A \cos \omega t - i A \sin \omega t) \quad (26)$$

因此，用 CEOF 对 Z 作分解，可得到一个特征值，对应一个复特征向量及一个复主分量，其对应的空间振幅、时间振幅、空间位相及时间位相分别为

$$s(x) = [\sin^2 kx + \cos^2 kx]^{1/2} = 1$$

$$s(t) = [A^2 \sin^2 \omega t + A^2 \cos^2 \omega t]^{1/2} = A$$

$$\theta(x) = \text{tg}^{-1}(\cos kx / \sin kx) = \pi/2 - kx$$

$$\theta(t) = \text{tg}^{-1}(A \sin \omega t / A \cos \omega t) = \omega t$$

由此可知，空间振幅是一常数，时间振幅与 A 成比例，而空间与时间相函数具有在空间与时间上的梯度，它们与波数 k 及频率 ω 有关。

上述例子说明传统的 EOF 方法不能描述系统的行波结构，而 CEOF 则可克服这一缺点。为了说明实际要素场这两种方法分析比较，还可利用 Horel^[6] 给出的一个例子来说明。

考虑一个正弦波，使它从左至右通过由 11 个网格点所构成的场，其振幅 $A(t)$ 是时间的函数，因而所构成的时间序列是非平稳的。其函数描述为

$$u_j(t) = A(t) \sin \left[2\pi \left(\frac{j-1}{10} - \frac{t-1}{50} \right) \right] + \varepsilon(t) \quad (27)$$

其中 $1 \leq j \leq 11$ ， $1 \leq t \leq 301$ ，当 $t \leq 150$ 时，令 $A(t) = 1$ ；当 $t > 150$ 时，令 $A(t) = 2$ 。 $\varepsilon(t)$ 为白噪音，它的振幅为信号振幅的 1%。

图 1 给出第 3、6 及 9 个格点上的时间序列 (PT3、PT6 及 PT9)。其行波特征是十分明显的。图中 PC1 及 PC2 是用传统的方法求得第 1 及第 2 主分量的序列，一个具有正弦波变化的特征，另一具有余弦波变化的特征。由于它们已具有极大的荷载方差，故转动后的主分量亦具有相同的特征（图中用

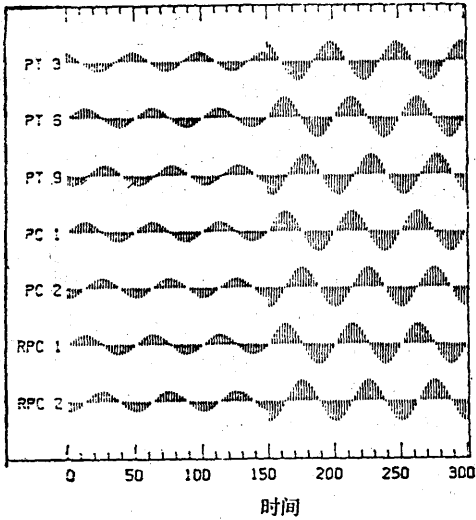


图1 非平稳正弦波序列

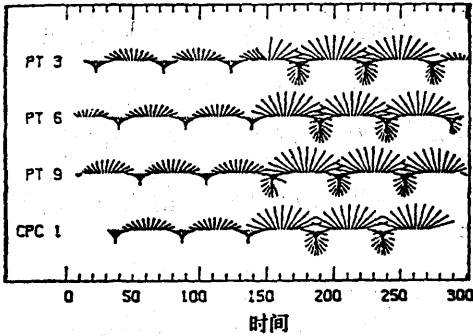


图2 复非平稳正弦波序列

RPC1 及 RPC2 表出)。图2 给出第3、6及9 网格点上包含有虚部 (Hibert 变换) 的复时间序列, 表示为复平面上向量的形式。图中CPC1 为该序列的唯一特征值所对应的复主分量, 由图可见, CPC1 随时间变化能很好地反映该时间序列时间变化的主要

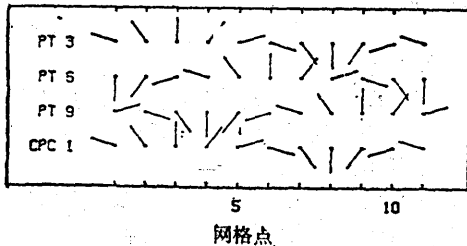


图3 序列的空间表现

特征。CPC1 同样能反映空间的正弦波变化特征, 图3 给出第3、6及9 格点与其余格点之间的复相关系数 (亦表示为复平面中向量形式), CPC1 在图中表示复特征向量在各空间点上的复平面向量, 其正弦波形是明显的。

六、应用

Barnett^[1]利用纬向与经向风在赤道附近的主要特征来讨论 Walker 环流的性质。这种主要特征用CEOF方法来提取。图4 给出对应纬向风场的第1复特征向量的空

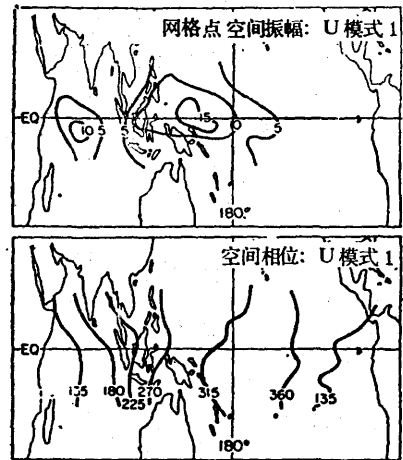


图4 纬向风的第1模式的空间振幅 (上图: 单位是相对单位) 及空间位相 (下图, 单位: 度)

间振幅及空间位相的分布模式。由于对应复主分量的解释方差达27%, 因而是具有代表性的显著模式。从图可见, 最大变化的地区出现在180°以西的赤道太平洋地区, 该地区是季风主要影响地域, 第2个变化最大地区出现在印度洋。从空间位相分布可见, 纬向风异常的变化区域是从印度洋向西太平洋传播, 由此他认为 Walker 环流扰动首先出现在印度季风带, 然后才在太平洋季风区出现。图5 给出纬向风时间振幅、位相及其向量 (用复平面中向量表示) 与秘鲁 Talara 站的海温比较。从图可见, 在1970—1976年间纬向风的变化是增加的, 这种增强与厄尔尼诺现象或多或少是有关的。从位相随时间变

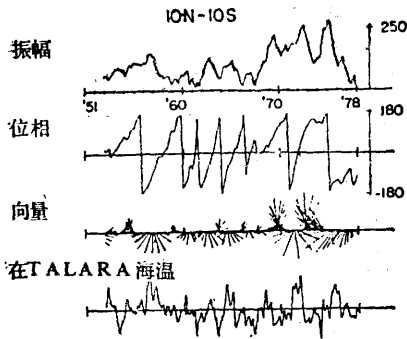


图5 第1复主分量的振幅、位相与实测海温变化比较

化来看, 纬向风的主要变化特征是有强的周期性, 但不同时期主要变化周期不同。在50年代是3—4年的变化周期, 在60年代初是2—3年, 在70年代又是4—5年周期。计算表明, 其基本周期为3.3年, 调制周期为25年。从第1复主分量随时间变化上看是有强的季节变化的, π 值大多出现在9至12月份, 且在它出现以后的2—6个月, 秘鲁Talara的海温亦出现较强的异常。

Anderson等人^[7]用CEOF研究大气角动量40—50天变化的纬度及高度的特征。

参考文献

- [1] Barnett, T. P., Interaction of the Monsoon and Pacific trade wind system at interannual time scales. Part I: The equatorial zones, Mon. Wea. Rev., 111, 756—773, 1983.
- [2] Hayashi, Y., A method of analyzing transient waves by space-time cross spectra, J. Appl. Meteor., 12, 404—408, 1973.
- [3] Wallace, J. W. and R. E. Dickinson, Empirical orthogonal representation of time series in the frequency domain, Part I: Theoretical consideration, J. Appl. Meteor., 11, 887—892, 1972.
- [4] Rasmusson, E. M., P. A. Arkin and Chen W-Y, Biennial variations in surface temperature over the United States as

他们计算经带通过滤后的风速与角动量的复相关系数, 从复相关系数的振幅及位相(其表达形式见(13式)分布中发现, 在热带地区对流层上层它们之间有强的相关, 其传播方向是向下及指向极地。对风场的复经验正交函数分析表明, 在整个对流层的中层及上层、在20°S到60°N地区内出现大范围的高凝聚。

Trenberth等人^[5]则用这一方法研究北半球海平面气压场的准两年周期振动的表现。他们用提取的前几个复主分量及复特征向量, 分别恢复相应的随时间变化的恢复场(例如第1特征值所对应恢复场的第j点为 $\bar{U}_j^{(1)}(t) = P_1(t) V_1^*(j)$)。通过恢复场研究几个主要特征随时间的变化, 从而揭示它们与北大西洋涛动、太平洋北美涛动及南方涛动表现特征的关系。

此外, 黄嘉佑^[8]也用此法研究我国降水量的准两年周期振荡的时空表现; 刘家铭等人^[9]用它研究热带对流的传播方向; Barnett除用它研究热带地区季风相互作用^[10]外, 还用于研究全球海平面气压变化特征及动力模式的敏感性上^[11]。

revealed by singular decomposition, Mon. Wea. Rev., 109, 587—598, 1981.

- [5] Trenberth, K. E. and W-T. K. Shin, Quasi-biennial fluctuations in sea level pressures over the Northern Hemisphere, Mon. Wea. Rev., 112, 761—777, 1984.
- [6] Horel, J. D., Complex principal component analysis: Theory and examples, J. Climate Appl. Meteor., 23, 1660—1673, 1984.
- [7] Anderson, J. R. and R. D. Rosen, The latitude-height structure of 40—50 day variations in atmospheric angular momentum, J. Atmos. Sci., 40, 1584—1591, 1983.
- [8] 黄嘉佑, 准两年周期振荡在我国月降水量中的表现 (将在《大气科学》上发表)。

- [9] Lau Ka-Ming and T. J. Philips, Coherent fluctuations of extratropical geopotential height and tropical convection in intraseasonal time scales, *J. Atmos. Sci.*, 43, 1164-1181, 1986.
- [10] Barnett, T. P., Interaction of the Monsoon and Pacific trade wind system at

interannual time scales. Part II: the tropical band, *Wn. Wea. Rev.*, 112, 287-238, 1984.

- [11] Barnett, T. P., Variations in near-global sea level pressure, *J. Atmos. Soc.*, 42, 478-501, 1985.

CEOF method and its application to the diagnosis on weather process

Huang Jiayou

(Department of Geophysics, Beijing University)

Abstract

Complex Empirical Orthogonal Function (CEOF) was proposed by Barnett in 1983. It is a useful tool for the analysis on long-range weather process. Its applications in the field are synthetically introduced and reviewed in this paper.