

第一章 绪论

1. 请问什么是气候统计诊断?

使用统计学的手段, 根据气象观测资料, 对气候的变化和异常程度以及原因进行判断。

2. 气候统计诊断研究的内容有哪些?

- (1) 应用统计方法了解区域性或全球性气候变化的时空分布特征、变化规律及气候异常的程度。主要研究月、季、年及年代 4 个时间尺度的气候变化。
- (2) 通过统计方法探索气候变量之间及其与其它物理因素之间的联系, 以此研究造成气候异常的原因, 进而探索气候异常形成的物理机制。
- (3) 对气候数值模拟结果与实际变化状况之间的差异性进行统计诊断。

3. 气候统计诊断的一般步骤有哪些?

- (1) 收集资料。选取的资料应该准确、精确, 具有均一性、代表性和可比较性;
- (2) 资料预处理。例如对资料进行距平化或者标准化(不同的气候变量量纲不同, 无法在同一水平比较)处理;
- (3) 选择诊断方法。方法选取不恰当, 会给研究和物理解释带来困难;
- (4) 科学综合和诊断。不能盲目地套用计算公式, 需要对结果进行显著性检验; 运用深厚的气候学知识, 对计算结果进行科学的综合和细致的分析; 用专业气候知识对计算结果进行判断, 识别真伪, 概括出气候系统确实存在的事实及彼此间的联系。

第二章 基本气候状态的统计量

1. 表征中心趋势的统计量有哪些?

中心趋势统计量主要包括两个: **均值和中位数**。

(1) 均值: 描述某一气候变量样本平均水平的量;

假设对变量 x 进行 n 次观测(取样), 得到一组观测序列 $[x_1, x_2, x_3, \dots, x_n]$, 则均

值的表达式为: $\bar{x} = \frac{1}{n} \sum_{j=1}^n x_j$ 。

(2) 中位数: 是按从小到大(从大到小也可以)排列的一组数据中居于中间位置的数, 它不受分布数列的极大或极小值影响, 从而在一定程度上提高了中位数对分布数列的代表性。

2. 表征气候变量变化幅度的统计量有哪些? 请分别写出这些统计量的计算式并说明其含义。

表征气候变量变化幅度的统计量有: **距平、方差/标准差、标准化变量**, 用来表示气候变化与正常情况的偏差和变化的波动。

➤ 距平的定义为: $x_{di} = x_i - \bar{x} (i=1, 2, \dots, n)$ (下标 d 代表单词“deviation”, 偏差), 它反映了原始数据相对于均值的差异程度。

➤ 方差的定义为: $s^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$; 标准差的定义为: $s = \sqrt{s^2} = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2}$ 。

方差/标准差用来描述样本中数据与平均值差异的平均状况, 可以衡量资料围绕平均值的平均振动幅度。

➤ 在气象要素中, 各要素的单位不一样, 平均值及标准差也有所不同, 为使它们能在同一水平上进行比较, 常进行标准化, 即把它们变成无量纲的变量,

这种变量称为标准化变量。定义式为 $x_{si} = \frac{x_i - \bar{x}}{s} (i=1, 2, \dots, n)$ (下标 s 代表单词

“standardization” 标准化)。标准化变量的两个性质: **均值为 0, 方差为 1**。

3. 表示气候变量分布特征的统计量有哪些?

(1) 偏度系数: 表征分布形态与平均值偏离的程度, 作为分布不对称的测度;

$$g_1 = \sqrt{\frac{1}{6n}} \sum_{i=1}^n \left(\frac{x_i - \bar{x}}{s} \right)^3 \begin{cases} > 0, \text{正偏度、顶峰偏左} \\ = 0, \text{分布图形对称} \\ < 0, \text{负偏度、顶峰偏右} \end{cases}$$

(2) 峰度系数: 表征分布形态图形顶峰的凸平度。

$$g_2 = \sqrt{\frac{1}{24n}} \left[\frac{1}{n} \sum_{i=1}^n \left(\frac{x_i - \bar{x}}{s} \right)^4 - 3 \right] \begin{cases} > 0, \text{坡度偏陡} \\ = 0, \text{坡度正好} \\ < 0, \text{坡度平缓} \end{cases}$$

偏度和峰度系数均为 0 则表示所研究的统计量为理想的正态分布变量。利用 g_1 和 g_2 值测定出偏离 0 的程度, 以此确定变量是否遵从正态分布。实际应用时, 对 g_1 和 g_2 进行统计检验, 以判断变量是否近似服从正态分布。

4. 什么是 Pearson 相关系数? 什么是自相关系数? 什么是滑动相关系数? 什么是偏相关系数? 请分别写出它们的计算公式及适用的问题, 并简要说明结果如何分析(提示: 在考试、写作业的时候, 还需要涉及第三章相关系数的显著性水平检验, 才是对本道题的完整回答)。

(1) **Pearson 相关系数**(相关系数、点相关系数)是用来描述两个随机变量线性相关的统计量。假设有两个变量 x, y 进行 n 次观测, 得到样本 $[x_1, x_2, \dots, x_n]; [y_1, y_2, \dots, y_n]$, 则相关系数的计算公式为:

$$r_{xy} = \frac{s_{xy}}{s_x s_y} = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})^2}} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}}$$

相关系数适用于分析两个变量间的线性相关程度与相关方向, 取值范围在 $[-1, 1]$ 之间, -1 表示完全负相关, 0 表示无关(相互独立), 1 表示完全正相关, 但是变量之间是否存在相关, 还需要进行显著性水平检验, 结合物理背景解释其中的机制, 避免虚假相关。

一般来讲, 样本量大于 30 计算 Pearson 相关系数才有意义。当样本量较小时, 计算所得相关系数可能会离总体相关系数甚远。这时, 用计算无偏相关系数

加以校正—— $r^* = r \left[1 + \frac{1-r^2}{2(n-4)} \right]$ 。

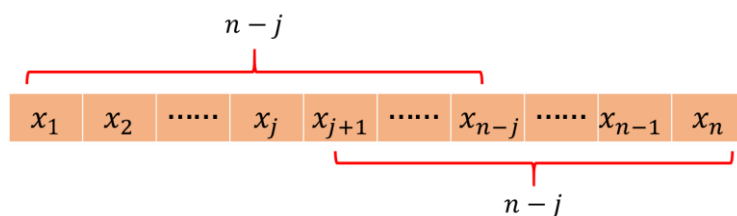
(2) 自相关系数是用来描述某一变量不同时刻之间相关程度的统计量。设变量 x

的观测序列为 $[x_1, x_2, \dots, x_n]$ ，均值 $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$ ，标准差 $s = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2}$ ，设滞

后长度为 j ，则可以将原来的样本分成两个序列：

$$\underbrace{[x_1, x_2, \dots, x_{n-j}]}_{n-j}, \underbrace{[x_{j+1}, x_{j+2}, \dots, x_n]}_{n-j}$$

用图形表示为：



根据相关系数的定义，有：

$$r(j) = \frac{\frac{1}{n-j} \sum_{i=1}^{n-j} (x_i - \bar{x})(x_{j+i} - \bar{x})}{\sqrt{\frac{1}{n-j} \sum_{i=1}^{n-j} (x_i - \bar{x})^2} \sqrt{\frac{1}{n-j} \sum_{i=1}^{n-j} (x_{j+i} - \bar{x})^2}} = \frac{1}{n-j} \sum_{i=1}^{n-j} \left(\frac{x_i - \bar{x}}{s} \right) \left(\frac{x_{j+i} - \bar{x}}{s} \right)$$

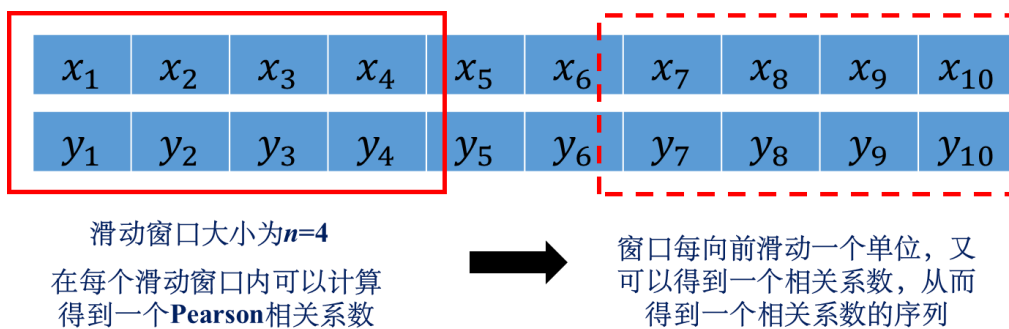
自相关系数可以帮助我们了解前 j 时刻的信息与其后时刻变化间的联系，从而判断由 x_i 预测 x_{i+j} 的可能性。

在计算时可以连续滞后长度，即 $j=1, 2, 3, \dots, k$ ，在每个滞后长度下，都可以计算得到一个相关系数，这样便可以得到 $r(1), r(2), \dots, r(k)$ 一共 k 个自相关系数，进而可以判断在滞后多少个时间长度后相关最强。

【区别】自相关 vs 超前滞后相关 自相关针对同一变量不同时刻的数据，超前滞后相关针对不同变量对应不同时刻的数据。自相关反映序列自己的内在规律，超前滞后相关判断不同变量间的潜在相关。

(3) 滑动相关系数可以用来分析两个变量在时间上的动态关系，可以观察两个变

量之间相关性随时间的变化。



计算公式为: $R_n^m = \frac{\sum_{i=t-n+1}^t (x_i - \bar{x}_m)(y_i - \bar{y}_m)}{\sqrt{\sum_{i=t-n+1}^t (x_i - \bar{x}_m)^2} \sqrt{\sum_{i=t-n+1}^t (y_i - \bar{y}_m)^2}}$, 其中

- n 表示滑动窗口的长度(即计算相关系数的子序列的长度);
- $\bar{x}_m = \frac{1}{n} \sum_{i=t-n+1}^t x_i, \bar{y}_m = \frac{1}{n} \sum_{i=t-n+1}^t y_i$ ($m=1, 2, \dots, N-n+1$) 分别表示在第 m 个滑动窗口的子序列均值, 其中 N 表示两个时间序列的总样本数量;
- $t = n, n+1, n+2, \dots, N$

根据滑动相关系数的时间序列, 可以根据其正负的变化将长时间序列划分为不同阶段, 从而可以对不同的阶段做进一步分析 (分析气候跃变)。

(4) 实际上大气的某一变量常常会受到多种因子的协同影响, 为了更好的反映单个因子对该变量的纯影响, 需要使用偏相关系数。设要素 a 受到要素 b 和 c 的影响, 则剔除 c 的影响后, a 和 b 的偏相关系数的计算式为

$$R_{ab,c} = \frac{R_{ab} - R_{ca}R_{cb}}{\sqrt{(1-R_{ca}^2)(1-R_{cb}^2)}}$$

其中 R_{ab}, R_{ca}, R_{cb} 分别为两要素之间的简单相关系数。下面是龚道溢和王绍武(2003)中的一个例子:

表 1 AO 指数与我国冬季气温及降水的相关系数统计 (资料 1958/59~1994/95), 显著水平超过 95% 的值以黑体标出

Tab. 1 Correlation statistics for the AO, the intensity of Siberian High and wintertime climate in China. The considered epoch is 1958/59~1994/95. Correlation coefficients above the significant at the 95% confidence level are bold

	AO 指数	西伯利亚高压强度
气温		
相关系数	+0.43	-0.67
偏相关系数	+0.14	-0.58
降水		
相关系数	+0.47	-0.36
偏相关系数	+0.36	-0.16
R (AO, 西伯利亚高压) = -0.51		

当西伯利亚高压的影响排除了以后, AO 指数与我国平均气温的偏相关系数只有 0.14, 达不到显著信度水平。如果把 AO 的影响排除掉以后, 西伯利亚高压与我国平均气温之间的偏相关系数高达-0.58。这说明西伯利亚高压对我国冬季气温的年际变化影响的确是非常重要的, 与 AO 相比较其影响要强得多。不过 AO 与降水的关系却明显不同: 当西伯利亚高压的影响排除了以后, AO 指数与 160 站平均降水的偏相关系数仍然有 0.36, 而当 AO 的影响排除了以后西伯利亚高压与降水的偏相关系数却只有-0.16。这说明 AO 对我国冬季降水的影响要强于西伯利亚。

第三章 基本气候状态的统计检验

1. 为什么要进行统计检验?

- (1) 我们通过某一气候变量序列的均值和方差了解其变化平均状态和变化幅度, 但不清楚这种状况是否稳定, 变化是否显著, 因此, 需要进行统计检验;
- (2) 相关系数仅仅显示气候变量间的相关程度, 到底变量间是否存在显著的相关, 必须经过统计检验。

2. 什么是统计检验与统计假设?

统计检验的基本思想是针对要检验的实际问题, 提出统计假设, 即用统计语言表达出期望得出结论的问题。由于所有统计检验都针对总体而言, 因此统计假设也必须与总体有关。例如, 统计假设必须表述为两总体均值相同或两样本来自均值相同的总体。

统计假设包括相互对立的两方面, 即原假设和对立假设, 原假设是统计检验的直接对象, 常用 H_0 表示, 对立假设是检验结果拒绝原假设时必然接受的结论, 用 H_1 表示。

3. 统计检验的一般流程有哪些?

- (1) 明确要检验的问题, 提出统计假设;
- (2) 确定显著性水平 α ;
- (3) 选取适当的统计量;
- (4) 根据观测样本计算有关统计量;
- (5) 对给定的 α , 从表上查出与 α 对应的数值, 即确定临界值;
- (6) 比较统计量的计算值与临界值, 看是否落入拒绝域, 如果落入拒绝域则拒绝原假设。

4. 显著性水平和置信度水平有什么区别?

- (1) 显著性水平, 英文为 “significance level”, 它是指在假设检验中用于判断拒绝原假设的概率水平, 通常用符号 α 来表示, 常见的表述有“显著性水平为 0.05”

“通过了显著性水平为 0.01 的检验”等。

- (2) 置信度水平, 英文为 “confidence level”, 它一般用在区间估计中, 表示对参数估计的信心程度, 常用百分比来表示 $(1-\alpha)$ 。当我们通过样本数据计算出一个参数的估计值时, 置信度水平表示在重复抽样的情况下, 得到的区间包含真实的参数值的概率有多大。

5. 总体方差已知、未知时, 均值检验的统计量如何构造?

单个总体均值的检验, 可用于检验一地气候是否稳定; 两个总体均值的检验, 用于检验两地气候变化是否存在显著差异。 单个总体的均值检验, 就是检验样本均值 \bar{x} 和总体均值 μ_0 无偏估计之间的差异是否显著; 两个总体的均值检验, 则是检验两个总体的均值是否相等。

(1) 总体方差已知时, 使用 u 检验

① 单个总体

来自正态总体容量为 n 的样本, 若有 100 次抽样(每次抽 n 个), 算得 100 个均值, 其均值 \bar{x} 显然会存在不同, 在概率统计中, 有如下定理

定理一 设 X_1, X_2, \dots, X_n 是来自正态总体 $N(\mu, \sigma^2)$ 的样本, \bar{X} 是样本的均值, 则有

$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$

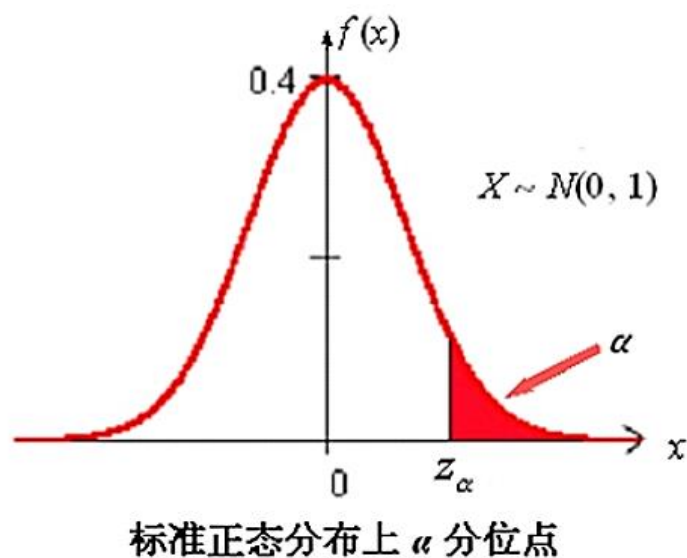
根据上述定理, 可构造统计量 u

$$u = \frac{\bar{x} - \mu_0}{\sqrt{\frac{\sigma^2}{n}}} = \frac{\bar{x} - \mu_0}{\sigma} \sqrt{n} \sim N(0, 1)$$

服从标准正态分布。当 $|u| \geq u_{\alpha/2}$ 时, 表示落在拒绝域内, 拒绝原假设, 即样本与总体有显著差异。

【注】

- 对于给定的 $\alpha (0 < \alpha < 1)$, 满足条件 $P(z > z_\alpha) = \int_{z_\alpha}^{\infty} f(x) dx = \alpha$ 的点 z_α 为标准正态分布的上 α 分位点。标准正态分布具有对称性。
- 均值是否相等属于双边检验, 给定显著性水平 α 需要除以 2 找临界值。



② 两个总体

来自两个总体的变量 x 和 y 分别遵从正态分布 $N(\mu_1, \sigma_1^2), N(\mu_2, \sigma_2^2)$, 两个变量的样本量分别为 n_1, n_2 , 均值分别为 \bar{x}, \bar{y} , 则统计量

$$u = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0, 1)$$

满足标准正态分布。当 $|u| \geq u_\alpha$ 时, 表示落在拒绝域内, 拒绝原假设, 即两总体均值之间存在显著差异。

(2) 总体方差未知时, 使用 t 检验

t 检验的原理与 u 检验类似用样本方差 s^2 代替总体方差 σ^2 。

① 单个总体

$$t = \frac{\bar{x} - \mu_0}{s} \sqrt{n} \sim t(n-1)$$

② 两个总体

$$t = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \sqrt{\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1 + n_2 - 2}}} \sim t(n_1 + n_2 - 2)$$

6. 方差检验需要构造什么统计量?

(1) 单个总体(总体方差已知)

概率统计中, 有定理

定理二 设 X_1, X_2, \dots, X_n 是来自正态总体 $N(\mu, \sigma^2)$ 的样本, \bar{X}, S^2 是样本的均值和方差, 则有

$$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$$

在确定显著性水平 α 后, 通过查卡方分布表, 若 $\chi^2 > \chi_{\alpha/2}^2 / \chi^2 < \chi_{1-\alpha/2}^2$, 则认为总体的方差有显著变化。

(2) 两个总体(方差未知)

构造统计量

$$F = \frac{\frac{n_1}{n_1-1} s_1^2}{\frac{n_2}{n_2-1} s_2^2} \sim F(n_1-1, n_2-1)$$

7. 相关性的检验有哪些方法?

(1) **t 检验法**: 构造统计量 $t = \sqrt{n-2} \frac{r}{\sqrt{1-r^2}} \sim t(n-2)$

(2) **u 检验法**: (样本足够大时)自相关系数的统计量为 $u(j) = r(j)\sqrt{n-j}$ (j 为滞后的时间)

(3) **临界相关系数法**: 根据(1)中的统计量, 可以反解得到临界相关系数为

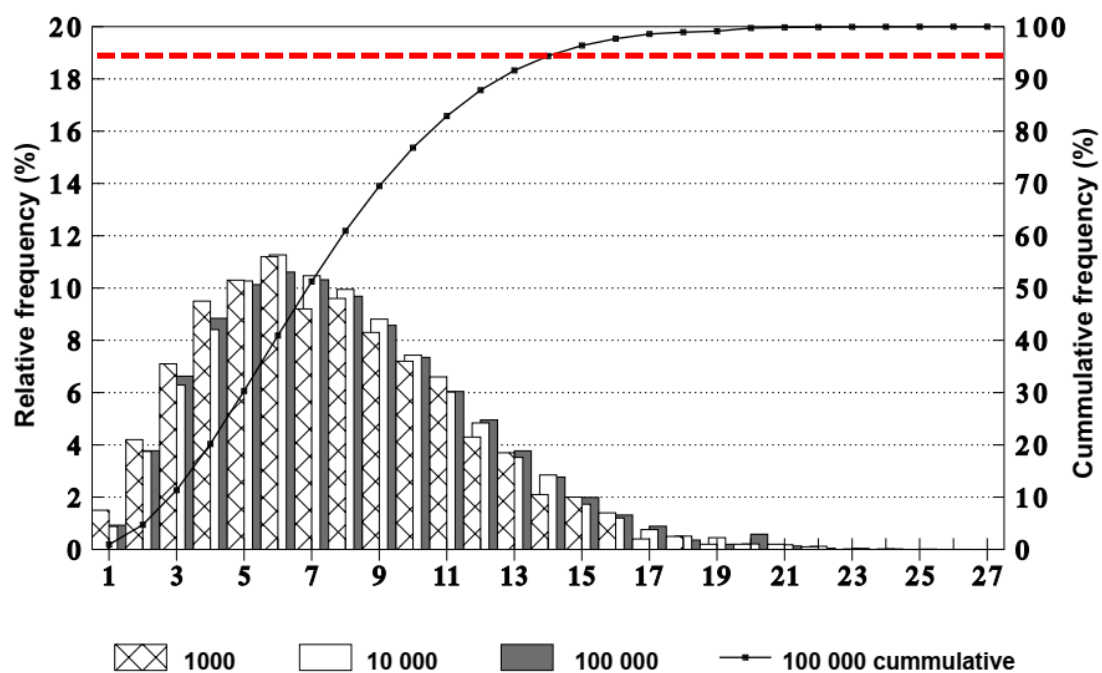
$$r_c = \sqrt{\frac{t_{\alpha/2}^2}{n-2+t_{\alpha/2}^2}}, \text{ 若通过样本计算得到相关系数 } |r| \geq r_c, \text{ 则拒绝原假设, 即相}$$

关是显著的。

(4) Monte Carlo 方法:

- 两个样本序列(一般是时间序列), 对其中的一个序列进行随机排序, 然后计算相关系数 r_i , r_i 是第 i 次调乱后的 x 和 y 所得的相关系数。
- 重复上一步 1000 次, 可得到 1000 个相关系数 $r_i (i = 1, 2, \dots, 1000)$ 。
- 计算相关系数的概率分布图(如下图), 而这一概率分布作为显著性的参考分

布。在给出显著性水平 α 后，在概率分布图上寻找累计概率密度为 $1-\alpha$ 对应的临界相关系数，若样本计算得出的相关系数 $|r| \geq r_c$ ，则认为相关是显著的。



(Hans von Storch and Francis W. Zwiers - *Statistical Analysis in Climate Research*)

第四章 气候变化趋势分析

1. 请问气候时间序列一般具有哪些特征？

随时间变化的一系列气候数据构成了一个气候时间序列，气候时间序列一般具有以下特征：

- (1) 数据的取值随时间变化；
- (2) 每一时刻取值的随机性；
- (3) 前后时刻数据之间存在相关性和持续性；
- (4) 序列整体上有上升或下降趋势，并呈现周期振荡；
- (5) 在某一时刻的数据取值出现转折或突变。

前两种特征是一般规律，后几种则在不同的序列有不同的表现。对任一气候时间序列都可以看成是由以下几个分量构成：

$$x_t = H_t + P_t + C_t + S_t + a_t$$

- H_t 为气候趋势分量，指几十年的时间尺度显示出的气候变量上升下降趋势，是一种相对序列长度的气候波动；
- P_t 为气候序列存在的一种固有的周期性变化，例如年、月变化；
- C_t 为循环变化分量，代表气候序列周期长度不严格的隐含周期性波动；
- S_t 是平稳时间序列分量（数学期望、方差不随时间变化；不同时刻之间的相关函数只是这两个时刻之差的函数，与时间起点无关）；
- a_t 是随机扰动项。

2. 请写出线性趋势分析的计算公式并说明结果如何分析，包括显著性检验。

趋势分析使用的方法为一元线性回归，用 x_i 表示样本量为 n 的某一气候变量，用 t_i 表示所对应的时刻，建立 x_i 与 t_i 之间的一元线性回归：

$$\hat{x}_i = bt_i + a (i = 1, 2, \dots, n)$$

其中 a 为回归常数， b 为回归系数。 a 和 b 可以用最小二乘法进行估计：

$$b = \frac{\sum_{i=1}^n x_i t_i - \frac{1}{n} \sum_{i=1}^n x_i \sum_{i=1}^n t_i}{\sum_{i=1}^n t_i^2 - \frac{1}{n} \left(\sum_{i=1}^n t_i \right)^2}, a = \bar{x} - b\bar{t} \left(\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i, \bar{t} = \frac{1}{n} \sum_{i=1}^n t_i \right)$$

▲ 回归系数的其他表达形式

$$b = \frac{\sum_{i=1}^n (x_i - \bar{x})(t_i - \bar{t})}{\sum_{i=1}^n (t_i - \bar{t})^2} = \frac{S_{xt}}{S_{tt}} = \frac{s_{xt}}{s_t^2} = \frac{Cov(x, t)}{Var(t)}$$

$$b = \frac{\sum_{i=1}^n (x_i t_i - x_i \bar{t} - \bar{x} t_i + \bar{x} \bar{t})}{\sum_{i=1}^n (t_i^2 - 2\bar{t} t_i + \bar{t}^2)} = \frac{\sum_{i=1}^n x_i t_i - \bar{t} \sum_{i=1}^n x_i - \bar{x} \sum_{i=1}^n t_i + n\bar{x}\bar{t}}{\sum_{i=1}^n t_i^2 - 2\bar{t} \sum_{i=1}^n t_i + n\bar{t}^2} = \frac{\sum_{i=1}^n x_i t_i - n\bar{x}\bar{t}}{\sum_{i=1}^n t_i^2 - n\bar{t}^2}$$

* S_{xt} 表示变量 x 和 t 的离差乘积和, S_{xx} 则表示变量 x 的离差平方和, 除以样本容量 n 则分别表示协方差和方差。所以对于两个变量之间的一元线性回归, 回归系数 b 的估计可以记为两个变量的协方差除以自变量的方差。

▲ 回归系数与相关系数的关系

二者表达式的比较:

$$b = \frac{Cov(x, t)}{Var(t)} = \frac{s_{xt}}{s_t^2}, r_{xt} = \frac{Cov(x, t)}{D(x)D(t)} = \frac{s_{xt}}{s_x s_t}$$

容易推导得出二者的关系为 $r_{xt} = \frac{s_t}{s_x} b$ 。

b 的符号表示气候变量的倾向趋势, $b > 0$ 表明随时间增加 x 呈上升趋势, $b < 0$ 表示随时间增加 x 呈下降趋势, 另外, b 的大小反映上升或下降的速率, 即表示上升或下降的倾向程度。因此, 通常将 b 称为倾向值, 将这种方法叫做线性倾向估计。根据相关系数与回归系数之间的关系可知, r 的符号与 b 相同, 因此 r 也可以说明 x 的上升下降趋势。

线性回归趋势的是否显著, 可以通过对相关系数的显著性检验来判断

$$t = \frac{r}{\sqrt{1-r^2}} \sqrt{n-2} \sim t(n-2)$$

服从自由度为 $n-2$ 的 t 分布, 给定显著性水平 α , 通过查表或者计算机, 求得 $t_{\frac{\alpha}{2}}$,

如果 $|t| > t_{\frac{\alpha}{2}}$, 则落在了拒绝域 (我们的原假设一般假定二者不相关), 故拒绝原假设, 从而认为两个变量显著相关, 即回归趋势显著。

3. 除了线性回归进行趋势估计, 其他的估计方法有哪些 (简要概述)?

(1) 滑动平均

滑动平均用已知时间序列的平滑值来显示变化趋势, 它相当于低通滤波器, 即保留低频部分、滤掉高频部分。对于样本量为 n 的序列 x , 其滑动平均序列表示为:

$$\hat{x}_j = \frac{1}{k} \sum_{i=-\frac{k-1}{2}}^{\frac{k-1}{2}} x_{j+i} \left(j = \frac{k+1}{2}, \frac{k+1}{2} + 1, \dots, n - \frac{k-1}{2} \right)$$

k 为滑动长度, 一般取奇数, 以使平均值可以加到时间序列中项的时间坐标上。显然, 经过滑动平均后, 原始时间序列中短于滑动长度的周期大大削弱, 从而可以看出变化趋势。

▲ 简单的理解

例如, 取 $k=3$, 即滑动长度为 3, 则第 j 个样本的滑动平均值为该点与前后两个点求和后的平均值, 即 $\hat{x}_j = \frac{1}{3}(x_{j-1} + x_j + x_{j+1})$, 这也可以理解为一种等权重的滑动平均 (此外还有权重滑动平均、指数滑动平均等)。由于滑动平均涉及前后的样本点, 所以对于一个原始的时间序列, 首尾的一些点是不能进行滑动平均的, 例如, 对于三点滑动平均, 最开始的第一个点与最后一个点没有滑动平均值; 对于五点滑动平均, 则最开始的两个点与最后两个点没有滑动平均值 (没有的数量即为 $\frac{k-1}{2}$)。

(2) 累计距平

累计距平是一种曲线直观判断变化趋势的方法。对于序列 x , 其某一时刻 t 的累积距平表示为

$$\hat{x}_t = \sum_{i=1}^t (x_i - \bar{x}) \quad (t=1, 2, \dots, n)$$

累计距平曲线呈上升趋势, 表示距平值增加, 呈现下降趋势则表示距平值减小。

从曲线明显的上下起伏, 可以判断其长期显著的演变趋势及持续性变化, 甚至还可以诊断出发生突变的大致时间。另外从曲线小的波动变化也可以考察其短期的距平值变化。

(3) 五、七和九点二次平滑

也是起到低通滤波的作用, 它可以克服滑动平均削弱过多波幅的缺点。滑动系数的求解采用最小二乘法求解。

(4) 五点三次平滑

(5) 三次样条函数

这些方法的变化趋势是否显著, 主要用到了 Mann-Kendal 非参数统计检验方法。对气候序列 x_i , 在 i 时刻, $i=1, 2, \dots, n-1$, 有

$$m_j = \begin{cases} +1, & x_j > x_i \\ 0, & x_j \leq x_i \end{cases} (j=i+1, i+2, \dots, n)$$

(因为涉及到 i 时刻后面的值要与该时刻的值进行比较, 所以前面的 i 取到 $n-1$)

令 $r_i = \sum_{j=i+1}^n m_j$, 则 r_i 表示第 i 时刻以后的所有数值 x_j 大于该时刻 x_i 的样本个数。令

$S = \sum_{i=1}^{n-1} r_i$ 对其进行标准化, 得到统计量

$$U = \frac{S - E(S)}{\sqrt{\text{Var}(S)}} \sim N(0,1)$$

在时间序列随机独立的假定下, S 的均值和方差分别为

$$E(S) = \frac{n(n-1)}{4}, \text{Var}(S) = \frac{n(n-1)(2n+5)}{72}$$

从而

$$U = \frac{S - \frac{n(n-1)}{4}}{\sqrt{\frac{n(n-1)(2n+5)}{72}}} = \frac{\frac{4}{n(n-1)}S - 1}{\sqrt{\frac{4n+10}{9n(n-1)}}}$$

其中 $Z = \frac{4}{n(n-1)}S - 1$ 即为课本上给出的统计量, 于是, 在给定显著性水平 α 时,

判据即为课本上给出的:

$$Z_{\alpha} = U_{\alpha} \sqrt{\frac{4n+10}{9n(n-1)}}$$

当 $|Z| > Z_{\alpha}$ 时, 认为变化趋势时显著的。

【计算统计量 $Z = \frac{4}{n(n-1)} \sum_{i=1}^{n-1} r_i - 1$ 。对于递增直线, r 序列为 $n-1, n-2, \dots, 1$,

这时 $Z=1$, 对于递减直线 $Z=-1$, 则 Z 值在-1 至 1 之间变化。】

参考: http://journal.iwhr.com/ch/reader/create_pdf.aspx?file_no=20190302&flag=1

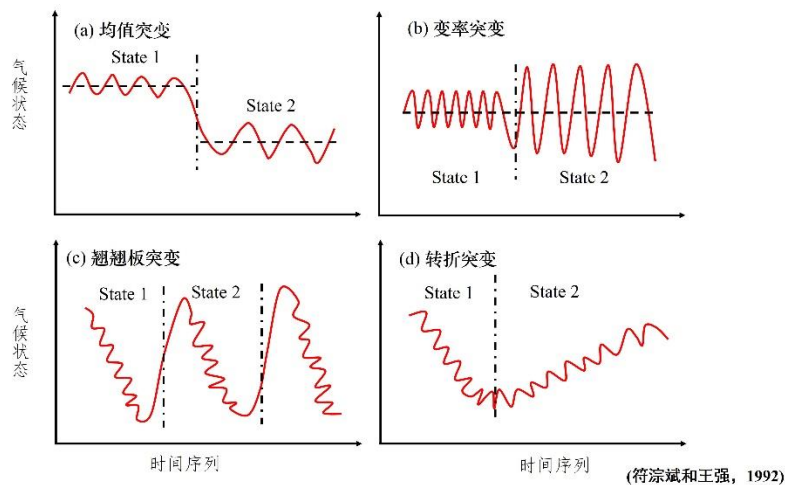
第五章 气候变化趋势分析

1. 请问气候突变有哪几种类型，对每种类型进行简要说明并画出简单的示意图。

气候突变通常是指从一个平均状态到另一个平均状态的急剧变化。根据符淙斌和王强(1992)的论文，气候突变可以分四种类型：

- (1) **均值突变**：气候从一个平均值到另一个平均值的变化，表现为气候变化的不连续性；
- (2) **变率突变**：气候从一个方差状态到另一个方差状态的急剧变化，其均值无明显变化；
- (3) **翘翘板(seesaw)突变**：气候态的变化形势如同一个翘翘板从最高(低)点到最低(高)点的转变。
- (4) **转折突变**：在某一时间持续减少(增加)，然后突然在某点开始持续增加(减少)。

简要示意图如下：



*符淙斌,王强.[气候突变的定义和检测方法](#)[J].大气科学,1992(04):482-493.

气候突变的普适定义：从一种稳定态（或稳定的持续变化趋势）跳跃式地转变到另一种稳定态（或稳定的持续变化趋势）的现象，表现为气候在时空上从一个统计特性到另一统计特性的急剧变化。

2. 请问时间序列的气候突变检测方法主要有哪些？请写出两种气候突变检测方法的基本公式、计算步骤、显著性检验并简要说明结果如何分析，比较两种方法的优缺点。

主要方法有：滑动 t 检验、克拉默 (Cramer) 方法、山本 (Yamamoto) 法、曼-肯德尔 (Mann-Kendall) 法、Pettitt 方法、勒帕热 (Le Page) 法。

(1) 滑动 t 检验

① 基本思想：考察两组样本平均值的差异是否显著。

② 基本公式

对于具有 n 个样本量的时间序列 x ，人为设置某一时刻为基准点，基准点前后两段子序列 x_1 和 x_2 的样本分别为 n_1 和 n_2 ，两段子序列平均值分别为 \bar{x}_1, \bar{x}_2 ，方差分别为 s_1^2, s_2^2 ，构造统计量

$$t = \frac{\bar{x}_1 - \bar{x}_2}{s \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim t(n_1 + n_2 - 2) \left(s = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}} \right)$$

③ 计算步骤与分析

- 确定基准点前后两子序列的长度，一般取相同长度 ($n_1 = n_2$)；
- 采取滑动办法连续设置基准点，分别计算统计量 t_i ；
- 给定显著性水平，查 t 分布表临界值 $t_{\frac{\alpha}{2}}$ ，若 $|t_i| > t_{\frac{\alpha}{2}}$ ，则认为基准点前后的两

子序列均值有显著差异，即出现突变，否则认为在基准点时刻未出现突变。

【注】选定滑动长度后，原始的时间序列长度将会有所缩短。例如，某时间序列的长度为 1911-1995，若选择 $n_1 = n_2 = 10$ ，则前 9 个样本和后 9 个样本则不能进行滑动平均，即最后画图的时候时间范围应该是从 1920-1986。

④ 缺点

子序列的选择带有人为性。因此实际使用时反复变动子序列的选取进行试验比较，提高可靠性。另外，突变时间也不能精确确定，只能确定大值范围，一般认为超过临界值的（绝对值）最大的 t_i 对应的年份为突变年份。

(2) Cramer 方法

① 基本思想：Cramer 法的原理与滑动 t 检验类似，区别仅在于它是通过比较某一个子序列与总序列的平均值的显著差异来检测突变。

② 基本公式

设总序列 x 和子序列 x_1 的均值分别为 \bar{x} 和 \bar{x}_1 ，两序列长度分别为 n, n_1 ，总序列方差为 s 。构造统计量

$$t = \tau \sqrt{\frac{n_1(n-2)}{n-n_1(1+\tau)}} \sim t(n-2) \left(\tau = \frac{\bar{x}_1 - \bar{x}}{s} \right)$$

③ 计算步骤与分析

- 确定子序列的长度 n_1 ；
- 以滑动的方式计算 t 统计量，得到 t 统计量序列 t_i ， $i = 1, 2, \dots, n - n_1 + 1$ ；
- 给定显著性水平，查 t 分布临界值，若 $|t_i| < t_\alpha$ ，则认为子序列的均值与总体序列均值无显著差异，否则认为在 t_i 对应时刻发生了突变。

④ 缺点

也需要人为确定子序列长度，在使用时应采取反复变动子序列的方法来提高计算结果的可靠性。

(3) M-K 方法

① 基本公式

设气候序列为 x_1, x_2, \dots, x_N ，定义 m_i 表示第 i 时刻 x_i 大于第 j 时刻 x_j ($1 \leq j \leq i$) 的数值个数，即

$$m_i = \begin{cases} +1, & x_i > x_j \\ 0, & x_i \leq x_j \end{cases} (j = 1, 2, \dots, i)$$

(即将当前时刻与前面时刻的值进行比较，看之前时刻当前时刻小的值有几个)
定义统计量

$$d_k = \sum_{i=1}^k m_i (k = 2, 3, \dots, n)$$

(即表示第 i 时刻数值大于第 j 时刻数值个数的 **累计数**)

在时间序列随机独立的假定下， d_k 的均值和方差分别为

$$E(d_k) = \frac{k(k-1)}{4}, \text{Var}(d_k) = \frac{k(k-1)(2k+5)}{72}$$

将 d_k 标准化后满足标准正态分布

$$UF_k(d_k) = \frac{d_k - E(d_k)}{\sqrt{\text{Var}(d_k)}} \sim N(0,1)$$

给定显著性水平 α , 若 $|UF_k| > u_\alpha$, 则表明序列存在明显的趋势变化, 所有 UF_k 值可组成一条曲线。

$$\text{【} UF_1 = UB_1 = 0 \text{】}$$

② 计算步骤与分析

- 计算顺序时间序列的秩序列 s_k , 计算统计量 UF_k 的值;
- 将此方法引用到反序列中, 计算统计量为 $UB_k (UB_k = -UF_k)$;
- UB_k, UF_k 的值大于 0 表明序列呈现上升趋势, 小于 0 则表明呈下降趋势, 超过临界值则表明上升或下降趋势显著。
- 将 UF_k, UB_k 的时间序列以及给定显著性水平的临界线画在同一张图上, 如果两条曲线的交叉点在信度线之间, 这点便是突变点的开始。

③ 优点

M-K 法计算简便, 可以确定突变的确切年份, 检测范围宽, 人为性少, 定量化程度高。

3. 请问用什么方法可检测空间分布型是否存在突变? 请写出基本公式、计算步骤、显著性检验并简要说明结果如何分析。

引入在空间 (θ, λ, p) 上随时间 t 变化的某一场变量函数

$$F = F(\theta, \lambda, p, t)$$

F 可以是标量函数, 如位势高度, 风速的分量等, 也可以是矢量 (如风速矢量)。

为了简单起见, 讨论在某一等压面上、给定区域的情形, 即 $p = \text{const}, (\theta, \lambda) \in S$ 。

记 t_1, t_2 时刻的场函数分别为

$$F_1 = F(\theta, \lambda, p, t_1), F_2 = F(\theta, \lambda, p, t_2)$$

引入内积 (F_1, F_2) 和范数 $\|F\|$ 如下:

$$(F_1, F_2) = \frac{1}{S} \iint_S F_1 \cdot F_2 dS, \quad \|F\|^2 = (F, F)$$

如果 F 为矢量, 则内积即为矢量的标量积 (数量积、点乘)。于是, 场 F_1, F_2 (假定二者不恒为 0) 的相似程度可以定义如下的相关系数来衡量:

$$R = \frac{(F_1, F_2)}{\|F_1\| \cdot \|F_2\|} \quad (-1 \leq R \leq 1)$$

- 两个场完全相同或者相似, 即 $F_1 = aF_2$ ($a = \text{const}, a > 0$), 有 $R = 1$;
- 两个场相反或者反相似, 即 $a < 0$, 有 $R = -1$;
- 两个场完全不像, 即正交, 有 $R = 0$;
- 当 F 表示行波时, 两时刻同位相有 $R = 1$, 反位相有 $R = -1$ 。

而两个场的差异性度量可以表示为: $d^2 = \frac{\|F_1 - F_2\|^2}{\|F_1\|^2 + \|F_2\|^2}$, 它与相关系数之间的

关系为

$$1 - d^2 = R \frac{2\|F_1\| \cdot \|F_2\|}{\|F_1\|^2 + \|F_2\|^2}$$

根据 $|2ab| \leq a^2 + b^2$, 得到 $-1 \leq 1 - d^2 \leq 1, 0 \leq d^2 \leq 2$ 。当 $\|F_1\| = \|F_2\|$ 时, 有 $1 - d^2 = R$,

于是

$$d = \begin{cases} 0, & \text{完全相同} \\ 1, & \text{正交} \\ 2, & \text{完全相反} \end{cases}$$

我们仅讨论空间函数场随时间变化的持续性和连续性, 则上述的相似性、差异性可以用滞后相关、滞后离差平方来表示。时间的变化是连续的, 因此可以通过积分来表示。函数 $F = F(\theta, \lambda, p, t)$ 在给定 p 的情况下的滞后相关 $R_\tau(t)$ 和滞后离差平方 $d_\tau^2(t)$ 的定义如下:

$$R_{\tau}(t) = \frac{1}{S} \iint_S \frac{F(\theta, \lambda, p, t) F(\theta, \lambda, p, t - \tau)}{\|F(t)\| \|F(t - \tau)\|} dS$$

$$d_{\tau}^2(t) = \frac{1}{S} \iint_S \frac{|F(\theta, \lambda, p, t) - F(\theta, \lambda, p, t - \tau)|^2}{\|F(t)\|^2 + \|F(t - \tau)\|^2} dS$$

对于理想情况, 有

$$F(\theta, \lambda, p, t) = \begin{cases} F_1(\theta, \lambda, p), & t < t_0 \\ F_2(\theta, \lambda, p), & t > t_0 \end{cases} \quad [F_1(\theta, \lambda, p) \equiv F_2(\theta, \lambda, p)]$$

即认为在 $t = t_0$ 前后分别有稳定态, 但在 $t = t_0$ 发生了突变, 从而有

$$R_{\tau}(t) = \begin{cases} 1, & t < t_0 \\ c(\text{常数}), & t_0 < t < t_0 + \tau \\ 1, & t > t_0 + \tau \end{cases} \quad c \equiv \frac{1}{S} \iint_S \frac{F_1 F_2}{\|F_1\| \|F_2\|} dS$$

实际上的空间场的演变并不是理想情况, 即在 t_0 与 $t_0 + \tau$ 之外时刻的空间场之间的

的相似系数不会是 1。实际的判断空间突变的方法如下:

- (1) 对于某一随时间变化的二维变量场, 确定滞后时间 τ ;
- (2) 采用滑动的方式, 计算空间相关系数, 得到序列 $R_i (i = 1, 2, \dots, n - \tau + 1)$;
- (3) 观察该序列, 如果有出现相关系数小于等于 0 的情形, 则可以认为空间分布型发生了突变;
- (4) 如果整个序列均为正值, 则需要给定显著性水平 α , 构建统计量

$$t = \frac{R}{\sqrt{1 - R^2}} \sqrt{n - 2} \sim t(n - 2), \text{ 进而求出临界相关系数 } R_c = \sqrt{\frac{t_{\alpha}^2}{n - 2 + t_{\alpha}^2}}, \text{ 当}$$

$R_i < R_c$ 时, 认为空间分布型发生了突变。

第六章 气候序列的周期分析

1. 离散功率谱分析的基本原理及显著性检验。

(1) 傅立叶变换与逆变换公式推导部分

任何一个时间函数 $x(t)$ 在满足狄氏条件下可以展开成傅立叶级数

$$x(t) = a_0 + \sum_{k=1}^{\infty} (a_k \cos \omega_k t + b_k \sin \omega_k t) \left(\omega_k = \frac{2\pi k}{T} \right)$$

利用三角函数系的正交性, 可以求得各项系数为

$$\begin{cases} a_0 = \frac{1}{T} \int_{-T/2}^{T/2} x(t) dt \\ a_k = \frac{2}{T} \int_{-T/2}^{T/2} x(t) \cos \omega_k t dt \\ b_k = \frac{2}{T} \int_{-T/2}^{T/2} x(t) \sin \omega_k t dt \end{cases}$$

根据欧拉公式 $e^{i\theta} = \cos \theta + i \sin \theta$, 有

$$\cos \omega_k t = \frac{e^{i\omega_k t} + e^{-i\omega_k t}}{2}, \sin \omega_k t = \frac{e^{i\omega_k t} - e^{-i\omega_k t}}{2i}$$

代入傅立叶展开式中, 得到

$$x(t) = a_0 + \sum_{k=1}^{\infty} \left(\frac{a_k - ib_k}{2} e^{i\omega_k t} + \frac{a_k + ib_k}{2} e^{-i\omega_k t} \right)$$

令 $c_0 = a_0, c_k = \frac{a_k - ib_k}{2}$, 因为 $a_k = a_{-k}, b_{-k} = -b_k$, 从而

$$c_{-k} = \frac{a_{-k} - ib_{-k}}{2} = \frac{a_k + ib_k}{2} = c_k^*$$

$$x(t) = a_0 + \sum_{k=1}^{\infty} (c_k e^{i\omega_k t} + c_{-k} e^{-i\omega_k t}) = \sum_{k=-\infty}^{\infty} c_k e^{i\omega_k t}$$

其中 c_k 称为 $x(t)$ 的复谱, 若设其形式为 $c_k = A_k e^{-i\theta_k}$, 则

$$\begin{cases} \text{振幅谱: } A_k = \frac{1}{2} \sqrt{a_k^2 + b_k^2} \\ \text{位相谱: } \theta_k = \arctan \frac{b_k}{a_k} \end{cases}$$

(2) 功率谱的定义与概念

若电阻为一个单位，瞬时电压用 $x(t)$ 表示，则瞬时功率为 $x^2(t)$ ($P = I^2 R = \frac{U^2}{R}$)，它的总能量为 $\int_{-\infty}^{\infty} x^2(t) dt$ 。在统计学上，这表示数学期望为 0 的方差。对于离散功率，因为

$$x(t) = \sum_{k=-\infty}^{\infty} c_k e^{i\omega_k t}$$

从而

$$\frac{1}{T} \int_{-T/2}^{T/2} x^2(t) dt = \frac{1}{T} \int_{-T/2}^{T/2} \sum_{k_1=-\infty}^{\infty} \sum_{k_2=-\infty}^{\infty} c_{k_1} c_{k_2} e^{i(\omega_{k_1} + \omega_{k_2})t} dt = \sum_{k_1=-\infty}^{\infty} \sum_{k_2=-\infty}^{\infty} c_{k_1} c_{k_2} \cdot \frac{1}{T} \int_{-T/2}^{T/2} e^{i(\omega_{k_1} + \omega_{k_2})t} dt$$

【注】 $x^2(t)$ 表示相对应的项相乘。

而积分 $\frac{1}{T} \int_{-T/2}^{T/2} e^{i(\omega_{k_1} + \omega_{k_2})t} dt = \begin{cases} 1, k_1 + k_2 = 0 \\ 0, k_1 + k_2 \neq 0 \end{cases}$ ，将下标都用 k 来表示，上述积分又可以

写成

$$\frac{1}{T} \int_{-T/2}^{T/2} x^2(t) dt = \sum_{k=-\infty}^{\infty} c_k c_{-k} = \sum_{k=-\infty}^{\infty} c_k c_k^* = |c_k|^2 = \frac{1}{4} (a_k^2 + b_k^2)$$

由于振幅谱相同，即离散功率谱的正负半轴是对称的，因此，可以将谱值乘以 2，只画出离散功率谱图的正半轴一部分 ($k > 0$)，故令

$$s_k^2 = 2|c_k|^2 = \frac{1}{2} (a_k^2 + b_k^2)$$

称 s_k^2 为离散功率谱，又称能谱密度。

(3) 离散功率谱的估计与检验

- ① 首先，针对不同的波数 $k(k=1, 2, \dots, \left[\frac{n}{2}\right])$ ，计算傅立叶系数 a_k, b_k 。其中 n 为观测资料的时间序列长度，根据 Nyquist 采样定理，可以分辨的最大频率不应超过采样频率的一半，因此最高的波数为 $\left[\frac{n}{2}\right]$ ；
- ② 根据 $s_k^2 = 2|c_k|^2 = \frac{1}{2} (a_k^2 + b_k^2)$ 计算离散功率；
- ③ 绘制功率谱图，以波数为横坐标，相应的功率谱值为纵坐标。

④ 进行功率谱的显著性检验，判断信号的显著周期。构造 F 统计量如下：

$$F = \frac{\frac{1}{2}(a_k^2 + b_k^2)}{\left(s^2 - \frac{1}{2}a_k^2 - \frac{1}{2}b_k^2\right)/(n-2-1)} \sim F(2, n-2-1)$$

服从分子自由度为 2，分母自由度为 $n-2-1$ 的 F 分布，式中 s^2 为原序列的方差。

2. 连续功率谱的基本原理及显著性检验。

多数实际资料并没有周期，这就需要考虑“非周期信号”的傅立叶变换，这里的非周期信号可以看成是 $T \rightarrow \infty (k \rightarrow 0)$ 无穷的周期信号。已知连续形式的傅立叶变换与逆变换的公式为如下。

$$F(\omega) = \int_{-\infty}^{\infty} x(t)e^{-i\omega t} dt \quad (\text{傅里叶变换, 从时域到频域})$$

$$x(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega)e^{i\omega t} d\omega \quad (\text{逆变换, 从频域到时域})$$

于是

$$\begin{aligned} \int_{-\infty}^{+\infty} x^2(t) dt &= \int_{-\infty}^{+\infty} x(t) \left[\frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega)e^{i\omega t} d\omega \right] dt = \frac{1}{2\pi} \int_{-\infty}^{+\infty} F(\omega) \left[\int_{-\infty}^{\infty} x(t)e^{i\omega t} dt \right] d\omega \\ &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} F(\omega)F(-\omega) d\omega \end{aligned}$$

由于 $F(-\omega) = F^*(\omega)$ ，上述积分又可以写为

$$\int_{-\infty}^{+\infty} x^2(t) dt = \frac{1}{2\pi} \int_{-\infty}^{+\infty} F(\omega)F^*(\omega) d\omega = \frac{1}{2\pi} \int_{-\infty}^{+\infty} |F(\omega)|^2 d\omega$$

定义 $S(\omega) = |F(\omega)|^2 = F(\omega)F(-\omega) = F(\omega)F^*(\omega)$ 为连续功率谱密度。

根据维纳-辛钦定理，连续功率谱和自相关函数是傅氏变换对，即

$$\begin{aligned} \rho(\tau) &= \int_{-\infty}^{\infty} x(t)x(t+\tau) dt \\ &= \int_{-\infty}^{\infty} x(t) \left[\frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega)e^{i\omega(t+\tau)} d\omega \right] dt = \int_{-\infty}^{\infty} x(t) \left[\frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega)e^{i\omega t} e^{i\omega\tau} d\omega \right] dt \\ &= \int_{-\infty}^{\infty} F(\omega) \left[\frac{1}{2\pi} \int_{-\infty}^{\infty} x(t)e^{i\omega t} dt \right] e^{i\omega\tau} d\omega \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega)F(-\omega)e^{i\omega\tau} d\omega \\ S(\omega) &= \int_{-\infty}^{\infty} \rho(\tau)\cos\omega\tau d\tau = 2 \int_0^{\infty} \rho(\tau)\cos\omega\tau d\tau \end{aligned}$$

连续功率谱的估计与检验过程如下:

- (1) 首先计算样本的自相关系数 $r(\tau)$ ($\tau = 0, 1, 2, \dots, m$), m 为最大落后长度。关于“最大落后长度” m 的选取: m 太小时, 用于谱估计的采样点太少, 影响功率谱的估算精度; m 太大时, 用于计算自相关系数的样本长度就太小, 影响自相关系数的估算精度, 因此一般 m 取在 $n/3 \sim n/10$ 之间(n 为样本容量)。
- (2) 利用“梯形法”对 τ 求积分, 计算粗谱(除以 m 表示平均)。

$$\hat{S}(\omega_k) = \frac{2}{m} \left[\frac{1}{2} r(0) + \sum_{\tau=1}^{m-1} r(\tau) \cos \omega_k \tau + \frac{1}{2} r(m) \cos \omega_k m \right]$$

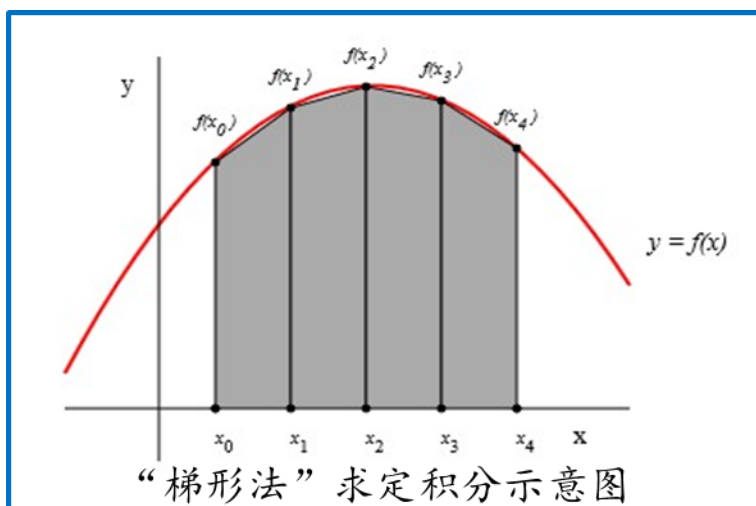
- (3) 根据自相关系数, $r(1)$, $r(2)$, $r(3)$ 确定检验谱。如果 $r(1)$, $r(2) \rightarrow 0$ 或变为负, 则用白噪音谱, 如果 $r(2) \cong [r(1)]^2$, $r(2) \cong [r(1)]^3$, 表明满足一阶自回归模型, 则用红噪音。
- (4) 根据选定的检验谱, 计算出 95% 置信限上界。对于红噪音, 其功率谱密度的

表达式为 $S(\omega) = \frac{1 - \rho_1^2}{1 - 2\rho_1 \cos \omega + \rho_1^2}$, 而对于白噪声, 其功率谱密度在整个频域内

呈均匀分布, 为常数, 即 $S(\omega) = 1$ 。

- (5) 绘制谱图, 同时绘制噪声谱的置信上限, 高于该上限的周期即为显著周期。

【注】梯形法求积分的示意图



“梯形法”近似求定积分： $f(x)$ 在区间 (x_0, x_m) 上的定积分可表示成多个小梯形面积的和：

$$\begin{aligned}\int_{x_0}^{x_n} f(x)dx &\approx \frac{1}{2}[f(x_0) + f(x_1)]\Delta x + \frac{1}{2}[f(x_1) + f(x_2)]\Delta x + \dots + \frac{1}{2}[f(x_{m-2}) + f(x_{m-1})]\Delta x + \frac{1}{2}[f(x_{m-1}) + f(x_m)]\Delta x \\ &= \frac{1}{2}f(x_0)\Delta x + \sum_{i=1}^{m-1} f(x_i)\Delta x + \frac{1}{2}f(x_m)\Delta x \\ &= \Delta x \left[\frac{1}{2}f(x_0) + \sum_{i=1}^{m-1} f(x_i) + \frac{1}{2}f(x_m) \right]\end{aligned}$$

3. 什么是窗口傅立叶变换？与传统的傅立叶变换有什么区别？该方法有什么优缺点？

- (1) 地球物理过程通常是非平稳的，把这些非平稳过程看成是一系列短时平稳信号的叠加。这种短时性则是通过时间域上加窗来实现的，即采用一个窗函数 $g(t-\tau)$ 对信号 $f(t)$ 的乘积运算实现在 τ 附近的开窗和平移，然后再进行 Fourier 变换，这就是加窗 Fourier 变换(WFT)，也称短时 Fourier 变换(STFT)。
- (2) 传统傅立叶变换的积分作用平滑了非平稳过程的突变成分，无法知道信号在突变时刻所对应的频率成分。
- (3) 优点：① 可以提取局部信息；② 能量守恒的变换；③ 没有损失 $f(t)$ 在频域上的任何信息。

缺点：① **不具有自适应性**。实际的信号过程是很复杂的，为了提取高频分量的信息，时域窗口应尽量窄；对于慢变信号或低频成分，时域窗口应适当加宽，以保证至少包含一个周期过程，因此需要窗口宽度能根据实际信号的变化来调节，即需要窗口具有自适应性；② 窗口 Fourier 变换的时-频窗口大小固定不变，只适合分析所有特征尺度大致相同的各种过程，不适于分析多尺度信号过程和突变过程。

【窗口函数的特点】

窗口函数的宽度非常有限，在某一指定宽度内窗口函数的值不为零，在指定宽度外，窗口函数的值迅速衰减为 0，故对信号加窗后，只在窗口函数不为零的信号可以显现出来，而在指定宽度以外的信号则不能显现。

4. 小波分析的基本原理([Christopher and Cilbert P., 1997](#))。

通过小波变换可以分析非平稳过程功率的时间序列在不同频率的谱值。给定

时间序列 $x_n (n=1, 2, 3, \dots, N-1)$ ，假定一个小波函数 $\Psi_0(\eta)$ ，它依赖于无量纲的时间参数 η ，同时，为了允许作为小波，该函数满足两个条件：(1) 均值为 0；(2) 具有时-频局部性。那么序列 x_n 的小波变换为

$$W_n(s) = \sum_{n'=0}^{N-1} x_{n'} \Psi^* \left[\frac{(n'-n)\delta t}{s} \right]$$

其中：

- Ψ^* 表示 Ψ 的共轭；
- s 为小波尺度，它表示在小波函数在时间轴上的缩放参数，反映了小波函数在时间域上的扩展和压缩程度。尺度越大(小)，小波函数在时间上的宽度越宽(窄)，频率越低(高)；
- n 为局部时间指数，表示小波函数在时间轴上的平移情况。

因此通过变换小波尺度 s 及其在局部时间点 n 上进行滑动，可以得到相对于每个尺度的振幅及其振幅随时间的变化。

注意到小波函数的下标 0 在上述公式中没有出现，表示已经进行了标准化，其相应的公式为

$$\Psi \left[\frac{(n'-n)\delta t}{s} \right] = \left(\frac{\delta t}{s} \right)^{\frac{1}{2}} \Psi_0 \left[\frac{(n'-n)\delta t}{s} \right]$$

常用小波函数 $\Psi(\eta)$ 为复数，其小波变换 $W_n(s)$ 也为复数，将它的模长平方 $|W_n(s)|^2$ 定义为小波功率谱。为了方便对不同的的小波功率谱进行比较，需要对小

波功率谱进行标准化，即 $\frac{|W_n(s)|^2}{\sigma^2}$ ，其中 σ^2 为原序列的方差。

小波变换分析(这里为定性阐释)：利用(红/白)噪声谱对变换得到的小波功率谱进行显著性检验，绘制其时间频率分布图，判断显著的周期，并分析事件发生的程度。如下图 b 所示，通过显著性检验的大部分功率主要集中在 2-8 年的 ENSO 波段内。通过小波分析，可以看到 ENSO 时间发生的频率和幅度的变化。在 1875-1920 年期间，出现了许多较强(功率谱值大)的 ENSO 事件，而在 1920-1960 年间，这些事件很少出现。

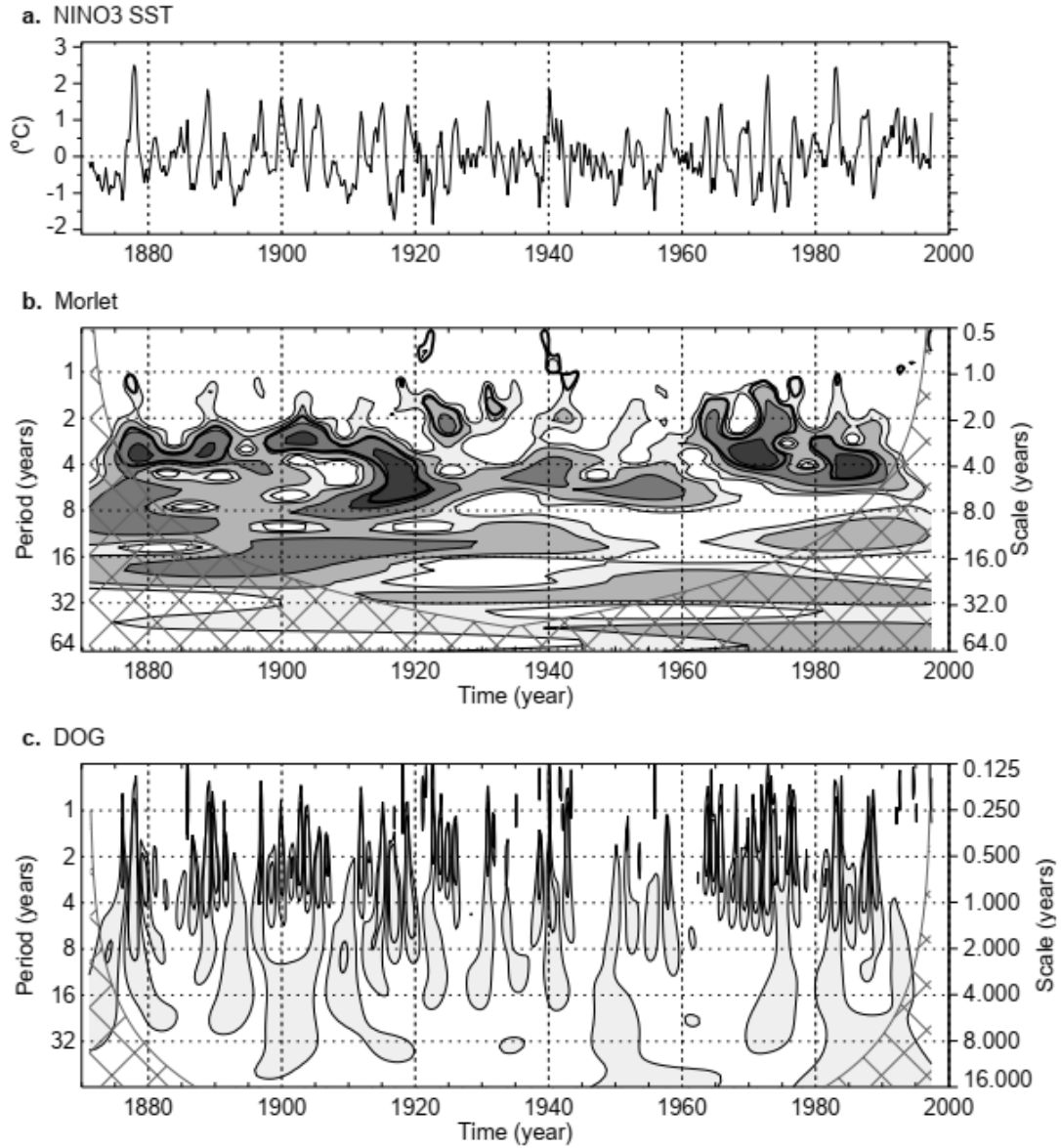


FIG. 1. (a) The Niño3 SST time series used for the wavelet analysis. (b) The local wavelet power spectrum of (a) using the Morlet wavelet, normalized by $1/\sigma^2$ ($\sigma^2 = 0.54^\circ\text{C}^2$). The left axis is the Fourier period (in yr) corresponding to the wavelet scale on the right axis. The bottom axis is time (yr). The shaded contours are at normalized variances of 1, 2, 5, and 10. The thick contour encloses regions of greater than 95% confidence for a red-noise process with a lag-1 coefficient of 0.72. Cross-hatched regions on either end indicate the “cone of influence,” where edge effects become important. (c) Same as (b) but using the real-valued Mexican hat wavelet (derivative of a Gaussian; DOG $m = 2$). The shaded contour is at normalized variance of 2.0.

第七章 气候变量场的时空结构的分离

1. EOF 的适用什么问题？其基本原理是什么？

EOF 方法可以把气象要素的时空变量场分解为经验正交函数的组合，实现时间变化和空间变化的分离，从而构成数量较少的几个的不相关典型模态，代替原变量场。基本原理如下：

- (1) 假设空间个点的数量有 p 个，每个格点均有一个 n 次的观测序列，得到原始距平资料矩阵

$$X_{p \times n} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{p1} & x_{p2} & \cdots & x_{pn} \end{bmatrix}$$

- (2) 根据线性代数的知识，对该距平资料阵进行分解

$$X_{p \times n} = V_{p \times m} Y_{m \times n} = \begin{bmatrix} v_{11} & v_{12} & \cdots & v_{1m} \\ v_{21} & v_{22} & \cdots & v_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ v_{p1} & v_{p2} & \cdots & v_{pm} \end{bmatrix} \begin{bmatrix} y_{11} & y_{12} & \cdots & y_{1n} \\ y_{21} & y_{22} & \cdots & y_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ y_{m1} & y_{m2} & \cdots & y_{mn} \end{bmatrix}$$

或者也可以写成(向量的表示形式均为列向量)

$$\begin{aligned} X_{(p \times n)} &= V_{(p \times m)} Y_{(m \times n)} = [v_1, v_2, \dots, v_m] \begin{bmatrix} y_1^T \\ y_2^T \\ \vdots \\ y_m^T \end{bmatrix} = \sum_{k=1}^m v_k y_k^T \\ &= v_1 y_1^T + v_2 y_2^T + \cdots + v_m y_m^T \\ &= X^{(1)} + X^{(2)} + \cdots + X^{(m)} \end{aligned}$$

其中 $V_{p \times m}$ 称为空间函数矩阵， $Y_{m \times n}$ 称为时间函数矩阵， $m = \min\{p, n\}$ 是矩阵

XX^T ($p \times p$) 的秩。 $V_{p \times m}$ 是矩阵 XX^T 的特征向量，它的每个列向量是相互正交的

单位向量，即 $VV^T = V^T V = I$ ，且每个特征向量对应矩阵的一个特征值 λ ，代表对

应模态的方差贡献，进而可以求得主成分 $Y_{p \times n} = V_{m \times p}^T X_{p \times n}$ 。

(3) 方差贡献

将特征向量按 λ 从大到小的顺序排列，从而代表了第一、第二、...、第 m 主

成分, 其中第 k 主成分的方差贡献以及前 k 主成分的累计方差贡献分别为

$$R_k = \frac{\lambda_k}{\sum_{i=1}^m \lambda_i}, G(k) = \frac{\sum_{i=1}^k \lambda_i}{\sum_{i=1}^m \lambda_i}$$

(4) 显著性检验

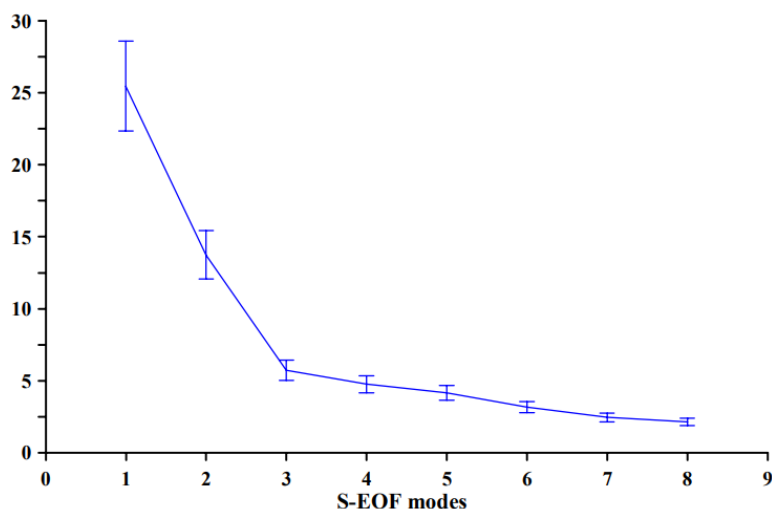
➤ 方法一: North(1982)准则

第 k 模态特征值的典型取样误差为 $\Delta\lambda_k = \lambda_k \sqrt{\frac{2}{n}}$, 要使得第 $k+1$ 模态显著区别

于第 k 模态, 则需要

$$\frac{|\lambda_k - \lambda_{k+1}|}{\Delta\lambda_k} > 1 \Leftrightarrow \lambda_{k+1} + \Delta\lambda_{k+1} < \lambda_k - \Delta\lambda_k$$

即前一模态的误差下限与后一模态的误差上限不相交。



➤ 方法二: Monte Carlo 方法

① 首先根据原始的资料阵计算某一模态的方差贡献 $R_k = \lambda_k / \sum_{i=1}^m \lambda_i$;

② 根据空间点数 p 及其样本量 n , 利用随机数发生器生成随机序列的资料矩阵, 进行 100 次模拟 EOF 计算。每次模拟后计算一次方差贡献

$$U_k^r = \delta_k^r / \sum_{i=1}^m \delta_i^r \left(\begin{matrix} k = 1, 2, 3, \dots, m \\ r = 1, 2, 3, \dots, 100 \end{matrix} \right)$$

③ 将 U_k^r 进行排序, 即 $U_k^1 \leq U_k^2 \leq \dots \leq U_k^{100}$ ($k = 1, 2, \dots, m$), 如果 $R_k > U_k^{95}$, 则认为第

k 模态在 95%置信度水平上是显著的。

(5) 优点

- EOF 没有固定的函数，不像有些分解需要以某种特殊函数为基函数；。
- 它能在有限区域对不规则分布的站点进行分解；
- 它的展开收敛速度快，很容易将变量场的信息集中在几个模态上；
- 分离出的空间结构具有一定的物理意义。

【模态方差贡献与模态局地方差贡献】

- 模态方差贡献：某一模态所有格点的方差和对原始场所有格点方差和的比值。
- 模态局地方差贡献：某一模态在某个格点上的方差与原始场该格点方差的比值。

2. 扩展 EOF 适用什么问题？其基本原理是什么？

EOF 方法可以分析固定时间形式的空间分布结构，它不能得到随时间移动的空间分布结构。然而，气候变量场在时间上存在显著的正相关及交叉相关，EEOF 考虑了变量场时间上的这种联系，因此可以得到变量场的移动性分布结构。EEOF 方法主要用在空间分布时滞性变化的分析研究中。对于逐月的资料，它的基本步骤如下：

(1) 已知原始的距平资料阵为

$$\mathbf{X}_{p \times n} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{p1} & x_{p2} & \cdots & x_{pn} \end{bmatrix}$$

选定滞后时间长度为 j ，它的大小依据具体情况决定，例如研究气候变量场准两年振荡， j 取 4 个月，此时资料阵由滞后 0, 4, 8 个月构成。构建滞后 1 个时次和 2 个时次的资料阵。此时，新资料阵的时间长度应为 $n-2j$ ，每个时次都有 p 个空间格点，所以新资料阵的大小为 $3p \times (n-2j)$ 。

$$X_{p \times n} = \begin{bmatrix} x_{11} & x_{12} & x_{13} & \cdots & x_{1,n-2} & x_{1,n-1} & x_{1n} \\ x_{21} & x_{22} & x_{23} & \cdots & x_{2,n-2} & x_{2,n-1} & x_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ x_{p1} & x_{p2} & x_{p3} & \cdots & x_{p,n-2} & x_{p,n-1} & x_{pn} \end{bmatrix}$$

 $n-2$
 $n-1$
 n

$$X_{3p \times (n-2j)} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1,n-2j} \\ x_{21} & x_{22} & \cdots & x_{2,n-2j} \\ \vdots & \vdots & \ddots & \vdots \\ x_{p1} & x_{p2} & \cdots & x_{p,n-2j} \\ x_{1,j+1} & x_{1,j+2} & \cdots & x_{1,n-j} \\ x_{2,j+1} & x_{2,j+2} & \cdots & x_{2,n-j} \\ \vdots & \vdots & \ddots & \vdots \\ x_{p,j+1} & x_{p,j+2} & \cdots & x_{p,n-j} \\ x_{1,2j+1} & x_{1,2j+2} & \cdots & x_{1n} \\ x_{2,2j+1} & x_{2,2j+2} & \cdots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{p,2j+1} & x_{p,2j+2} & \cdots & x_{pn} \end{bmatrix}$$

(2) 计算新资料阵的协方差矩阵 S ，此时 S 是 $3p$ 阶方阵。

(3) 求解特征值和特征向量，并进行显著性检验，选定主要模态。需要注意的是每个特征值对应的特征向量长度为 $3p$ 。第 k 模态的特征向量的表达式为

$$\mathbf{v}_k^T = (v_{1k}, v_{2k}, \dots, v_{pk}, v_{p+1,k}, v_{p+2,k}, \dots, v_{2p,k}, v_{2p+1,k}, v_{2p+2,k}, \dots, v_{3p,k})$$

其中红、蓝、橙对应的向量分别代表了滞后 0,1,2 时次的特征向量。

(4) 根据 $\mathbf{Y}_{p \times n} = \mathbf{V}_{m \times p}^T \mathbf{X}_{p \times n}$ 计算时间系数。每一个模态的时间系数长度也为 $3p$ ，对应着三个不同的时刻。

(5) 计算方差贡献。

【如果不是逐月资料，而是要考虑季节变化，可以采用季节 EOF 方法(S-EOF)，四个季节的距平阵构合成一个大矩阵，即 $\mathbf{X}_{4p \times n}$ 】

第八章 两气候变量场相关模态的分离

1. SVD 适用什么问题？请简述其基本原理

SVD 方法用于研究两个场之间相关系数的空间结构和它们各自对相关场的贡献，提取两个气象场的耦合信号。

- (1) 设有两个变量场，一个变量场由 p 个空间点构成，样本量(时间长度)为 n ，称作左场，记为 \mathbf{X} ；另一个变量场由 q 个空间点构成，样本量(时间长度)也为 n ，称作右场，记为 \mathbf{Y} 。对两资料矩阵做**标准化处理**。

$$\mathbf{X}_{p \times n} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{p1} & x_{p2} & \cdots & x_{pn} \end{bmatrix}, \mathbf{Y}_{q \times n} = \begin{bmatrix} y_{11} & y_{12} & \cdots & y_{1n} \\ y_{21} & y_{22} & \cdots & y_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ y_{q1} & y_{q2} & \cdots & y_{qn} \end{bmatrix}$$

- (2) 计算其交叉协方差阵 $\mathbf{S}_{xy} (p \times q) = \frac{1}{n} \mathbf{X} \mathbf{Y}^T$ ，对其做奇异值分解，得到

$$\mathbf{S} = \mathbf{L} \begin{bmatrix} \Lambda_m & 0 \\ 0 & 0 \end{bmatrix} \mathbf{R}^T$$

$$\mathbf{L}_{p \times m} = [\mathbf{l}_1, \mathbf{l}_2, \mathbf{l}_3, \dots, \mathbf{l}_m] = \begin{bmatrix} l_{11} & l_{12} & \cdots & l_{1m} \\ l_{21} & l_{22} & \cdots & l_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ l_{p1} & l_{p2} & \cdots & l_{pm} \end{bmatrix}, \mathbf{R}_{q \times m} = [\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \dots, \mathbf{r}_m] = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1m} \\ r_{21} & r_{22} & \cdots & r_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ r_{q1} & r_{q2} & \cdots & r_{qm} \end{bmatrix}$$

其中 \mathbf{L} 是矩阵 $\mathbf{S} \mathbf{S}^T$ 的特征向量，其中有 m 个向量 \mathbf{l} ，两两相互正交，每个向量包含 p 个元素，称为**左奇异向量**； \mathbf{R} 是矩阵 $\mathbf{S}^T \mathbf{S}$ 的特征向量其中有 m 个向量 \mathbf{r} ，两两相互正交，每个向量包含 q 个元素，称为**右奇异向量**。 $\mathbf{S} \mathbf{S}^T$ 和 $\mathbf{S}^T \mathbf{S}$ 具有相同的特征值，为 $\lambda_1^2, \lambda_2^2, \dots, \lambda_m^2 (\lambda_1^2 > \lambda_2^2 > \dots > \lambda_m^2)$ 。

【注】这里使用的是 λ 的平方表示 $\mathbf{S} \mathbf{S}^T / \mathbf{S}^T \mathbf{S}$ ，因为在 SVD 分解中，特征值与奇异值之间存在平方关系，即奇异值的平方为特征值。

- (3) 分别由左、右场构建新变量(计算时间系数矩阵)。

$$\mathbf{A}_{m \times n} = \mathbf{L}^T \mathbf{X} = \begin{bmatrix} \mathbf{a}_1^T \\ \mathbf{a}_2^T \\ \vdots \\ \mathbf{a}_m^T \end{bmatrix}_{(m \times p)(p \times n)} = \begin{bmatrix} l_{11} & l_{12} & \cdots & l_{1p} \\ l_{21} & l_{22} & \cdots & l_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ l_{m1} & l_{m2} & \cdots & l_{mp} \end{bmatrix} \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{p1} & x_{p2} & \cdots & x_{pn} \end{bmatrix}$$

$$\mathbf{B}_{m \times n} = \mathbf{R}^T \mathbf{Y} = \begin{bmatrix} \mathbf{b}_1^T \\ \mathbf{b}_2^T \\ \vdots \\ \mathbf{b}_m^T \end{bmatrix}_{(m \times q)(q \times n)} = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1p} \\ r_{21} & r_{22} & \cdots & r_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ r_{m1} & r_{m2} & \cdots & r_{mp} \end{bmatrix} \begin{bmatrix} y_{11} & y_{12} & \cdots & y_{1n} \\ y_{21} & y_{22} & \cdots & y_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ y_{q1} & y_{q2} & \cdots & y_{qn} \end{bmatrix}$$

因为 \mathbf{L} 、 \mathbf{R} 均为正交矩阵, 所以 $\mathbf{L}^T \mathbf{L} = \mathbf{I}$, $\mathbf{R}^T \mathbf{R} = \mathbf{I}$, 从而

$$\mathbf{X}_{p \times n} = \mathbf{L}^T \mathbf{A} = [l_1, l_2, l_3, \dots, l_m] \begin{bmatrix} \mathbf{a}_1^T \\ \mathbf{a}_2^T \\ \vdots \\ \mathbf{a}_m^T \end{bmatrix}, \mathbf{Y}_{q \times n} = \mathbf{R}^T \mathbf{B} = [r_1, r_2, r_3, \dots, r_m] \begin{bmatrix} \mathbf{b}_1^T \\ \mathbf{b}_2^T \\ \vdots \\ \mathbf{b}_m^T \end{bmatrix}$$

故 SVD 也相当于将左、右场分别进行 EOF 分解。通常称 $\mathbf{l}_k, \mathbf{r}_k$ 为一对变量型 (pattern), $\mathbf{a}_k, \mathbf{b}_k$ 分别为左右场的展开系数 (时间系数), 每一对变量型和相应的展开系数确定了一种模态 (mode)。

(4) 计算方差贡献与累计方差贡献, 并利用 Monte Carlo 方法进行显著性检验。

$$SCF_k = \frac{\lambda_k^2}{\sum_{i=1}^m \lambda_i^2}, CSCF_k = \frac{\sum_{i=1}^k \lambda_i^2}{\sum_{i=1}^m \lambda_i^2}$$

① 首先根据原始的资料阵计算某一模态的方差贡献 $SCF_k = \lambda_k^2 / \sum_{i=1}^m \lambda_i^2$;

② 根据左场空间点数 p 、右场空间点数 q 及其样本量 n , 利用随机数发生器生成随机序列的两个资料矩阵, 进行 100 次模拟 SVD 计算。每次模拟后计算一次方差贡献 (也可以只对一个场进行随机生成)

$$U_k^r = \delta_k^r / \sum_{i=1}^m \delta_i^r \left(\begin{matrix} k=1, 2, 3, \dots, m \\ r=1, 2, 3, \dots, 100 \end{matrix} \right)$$

- ③ 将 U_k^r 进行排序, 即 $U_k^1 \leq U_k^2 \leq \dots \leq U_k^{100} (k=1, 2, \dots, m)$, 如果 $SCF_k > U_k^{95}$, 则认为第 k 对 SVD 模态在 95% 置信度水平上是显著的。

(5) 分析模态相关系数

一旦由 SVD 得到时间系数矩阵 \mathbf{A} 、 \mathbf{B} , 就可以定义每对奇异向量的时间系数 \mathbf{A} 和 \mathbf{B} 之间的相关系数:

$$r(\mathbf{a}_k, \mathbf{b}_k) = \frac{\mathbf{a}_k \cdot \mathbf{b}_k}{\sqrt{|\mathbf{a}_k|^2} \cdot \sqrt{|\mathbf{b}_k|^2}} = \frac{\sum_{t=1}^n a_k(t) b_k(t)}{\sqrt{\sum_{t=1}^n a_k^2(t)} \sqrt{\sum_{t=1}^n b_k^2(t)}} \quad (\text{有点像计算相似性})$$

上式称为模态相关系数, 它反映每对奇异向量之间相互关系的密切程度, 反映的是典型变量场总体相关状况。

【其他相关系数】

➤ 异性相关系数

左场 \mathbf{X} 与右奇异向量的时间系数 \mathbf{B} 之间的相关系数为

$$r(\mathbf{X}, \mathbf{b}_k) = \frac{\sum_{t=1}^n X(t) b_k(t)}{\sqrt{\sum_{t=1}^n X^2(t)} \sqrt{\sum_{t=1}^n b_k^2(t)}}$$

右场 \mathbf{Y} 与左奇异向量的时间系数 \mathbf{A} 之间的相关系数为

$$r(\mathbf{Y}, \mathbf{a}_k) = \frac{\sum_{t=1}^n Y(t) a_k(t)}{\sqrt{\sum_{t=1}^n Y^2(t)} \sqrt{\sum_{t=1}^n a_k^2(t)}}$$

上两式计算的相关系数分布型分别表示右(左)场的时间系数所反映的左(右)场时间变化程度大小的分布, 显著相关区则是两变量相互作用的关键区。

➤ 同性相关系数

左(右)场时间系数与左(右)场的相关系数称为同性相关系数, 其分布反映时间系数序列表示自身气象场时间变化程度大小的地理分布, 在一定程度上代表了左、右场的遥相关型。