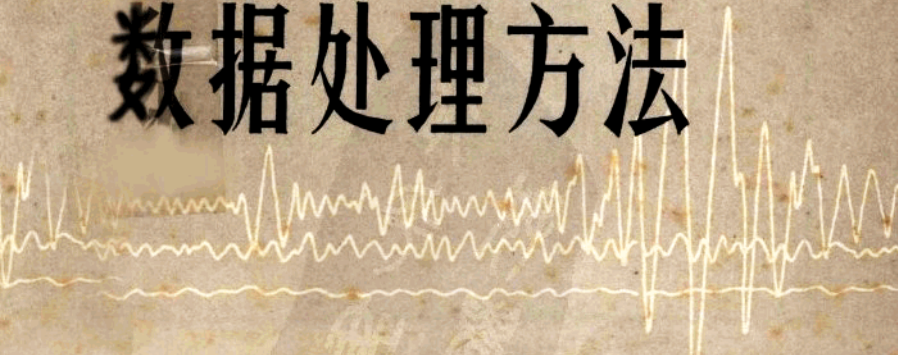


地震预报知识丛书

地震预报中的 数据处理方法



地震预报知识丛书(9)

地震预报中的数据
处理方法

国家地震局

一九七五年

毛主席语录

一个正确的认识，往往需要经过由物质到精神，由精神到物质，即由实践到认识，由认识到实践这样多次的反复，才能够完成。这就是马克思主义的认识论，就是辩证唯物论的认识论。

内 容 提 要

本书简明扼要地介绍了目前在地震预报工作中应用较多或较有成效的一些数据处理方法，并举实例加以说明。对于公式的推导尽量省略。鉴于电子计算机在国内的逐步推广，在本书最后介绍了使用电子计算机的基本知识。可供广大地震工作者进行数据处理时参考。

本书由中国科学院地质研究所四室计算组编写。

出 版 说 明

地震预报是一项艰巨而光荣的科研任务。我国地震工作者和震区广大人民,在毛主席革命路线指引下,为完成这项任务正在努力工作,并且取得了一定成绩和进展。

为了帮助广大地震工作者学习地震知识,提高业务水平,更好地为我国社会主义革命和建设事业服务,在几年来广泛实践的基础上,我们组织有关单位编写了这套《地震预报知识丛书》,介绍了地震的基本知识及各种预报方法(包括基本理论、工作状况和经验;仪器的结构、操作及维修等)。

这套书中介绍的某些方法,就各自的学科本身而言,虽然已有相当长的历史,形成了系统的知识,但是用来预报地震,仍处于探索性阶段;因此书中对某些问题的论述,是不够成熟的,仅供读者参考。

由于时间仓促,加之作者、编者的水平有限,这套书在内容的深、浅及叙述的繁、简等方面,未能做到统一。其它缺点或错误亦难免,敬请读者批评指正。

在编写这套书的过程中,得到了很多单位的大力协助与支持,在此一并表示感谢。

若以频率 $\omega_j = j\omega_t = 2\pi j/2N$ 为横座标, 对应于 ω_j 的 A_j 为纵座标, 所得图形称为该已知序列的振幅谱。若以频率为横座标, 以相应的相位为纵座标, 所得图形称为该已知序列的相位谱。

整点值幅差 y_k 可以按(5-1)式表示成:

$$\begin{aligned}
 y_k = & 45.44 + 1.552 \sin\left(\frac{2\pi \cdot 1 \cdot k}{24} + 25.8\right) \\
 & + 1.111 \sin\left(\frac{2\pi \cdot 2 \cdot k}{24} + 227.2\right) \\
 & + 1.338 \sin\left(\frac{2\pi \cdot 3 \cdot k}{24} + 6.4\right) \\
 & + \dots + 0.588 \sin\left(\frac{2\pi \cdot 11 \cdot k}{24} + 53.1\right)
 \end{aligned}$$

富氏系数 a_j 、 b_j 和分波个数无关, 也就是说各系数彼此间是独立的, 所以如果想了介某频率范围的分波振幅, 可以只求有关的 n 项。在没有电子计算机情况下, 为了节省工作量, 对不同数据个数, 有各种便于手算的计算格式。河北地震队把两地磁台整点值幅差作频谱分析, 求出日及半日波幅度, 然后再按红绿灯法进行预报地震。根据他们经验, 这样可以减少许多虚报。下面我们给出他们用于计算日波及半日波的格式。表格适用于 24 个数据情况, 其中 Z 表示原始数据, V 为两数据之和, W 为两数据之差。

富氏分析在自然科学及工程技术方面应用很广, 通常用来介微分方程或逼近非周期函数。但目前在地震及其前兆数据处理中, 往往感兴趣的是用它来研究时间序列的频率成分。例如图 5-2 为华北地区历史地震资料的振幅频谱, 按每 10 年内地震能量总和的对数作为原始数据, 组成等时间间隔的时间序列, 富氏分析的目的是要找出历史地震主要周期, 作为今

表 5-3

$V_0-44.3$	$V_1-91.2$	$V_2-90.9$	$V_3-90.8$	$V_4-89.9$	$V_5-89.6$	$V_6-89.2$	$V_7-87.1$	$V_8-90.1$	$V_9-95.5$	$V_{10}-93.4$	$V_{11}-89.7$	$V_{12}-48.9$
$Z_0-44.3$	$Z_1-45.2$	$Z_2-44.9$	$Z_3-44.5$	$Z_4-43.8$	$Z_5-43.5$	$Z_6-44.1$	$Z_7-44.3$	$Z_8-44.4$	$Z_9-43.6$	$Z_{10}-45.1$	$Z_{11}-40.8$	$Z_{12}-48.9$
	$Z_{23}-46.0$	$Z_{22}-46.0$	$Z_{21}-46.3$	$Z_{20}-46.1$	$Z_{19}-46.1$	$Z_{18}-45.1$	$Z_{17}-42.8$	$Z_{16}-45.7$	$Z_{15}-51.9$	$Z_{14}-48.3$	$Z_{13}-48.9$	
	W_1	W_2	W_3	W_4	W_5	W_6	W_7	W_8	W_9	W_{10}	W_{11}	W_{12}
	0.8	1.1	1.8	2.3	2.6	1.0	1.5	1.3	8.3	3.2	8.1	8.1

$V_0-44.3$	$V_1-91.2$	$V_2-90.9$	$V_3-90.8$	$V_4-89.9$	$V_5-89.6$
$-V_{12}$	48.9	$-V_{11}$	89.7	$-V_{10}$	93.4
	8.1	3.2	8.3	1.3	1.5
	0.259	0.500	4.307	10.1	0.866
	4.6	2.165	3.3229	0.1	0.6475
	1.1	1.8	2.3	2.6	1.0
	8.3	10.1	3.60966	1.1	1.00
	2.3051	7.1407	3.1176	1.0626	16.776

$$a_1 = \frac{\Sigma_1}{12} = 0.6743 \quad b_1 = \frac{\Sigma_2}{12} = 1.398$$

$$a_1^2 = 0.4547 \quad b_1^2 = 1.9544 \quad \Sigma_3 = 2.4096 \quad A_1 = \sqrt{\Sigma_3} = 1.552$$

$$a_2 = \frac{\Sigma_4}{12} = -0.8156 \quad b_2 = \frac{\Sigma_5}{12} = -0.7544$$

$$a_2^2 = 0.6652 \quad b_2^2 = 0.5691 \quad \Sigma_6 = 1.2343 \quad A_2 = \sqrt{\Sigma_6} = 1.111$$

$V_0-44.3$	$V_1-91.2$	$V_2-90.9$	$V_3-90.8$	$V_4-89.9$	$V_5-89.6$	$V_6-89.2$	$V_7-87.1$	$V_8-90.1$	$V_9-95.5$	$V_{10}-93.4$	$V_{11}-89.7$	$V_{12}-48.9$
$-V_6$	89.2	$-V_5$	89.6	$-V_4$	89.9	$-V_3$	90.1	$-V_2$	93.4	$-V_1$	91.2	$-V_0$
	1.00	-3.9	0.866	-4.2	0.500	-4.3						
	-3.9	$+$	-3.6372	$+$	-2.15	Σ_4	-9.7872					
	W_1	0.8	W_2	1.1	W_3	1.8						
	W_4	2.6	W_5	2.3	$-W_6$	-8.3						
	$-W_7$	1.5	$-W_8$	-1.3								
	$-W_{11}$	-8.1	$-W_{10}$	-3.2								
	0.500	-3.2	0.866	-1.1	1.00	-6.5						
	-1.6	$+$	-0.9526	$+$	-6.5	Σ_5	-9.0526					

后长期预报的参考。我们采用改变分析系列长度方法，找其稳定出现的较大振幅的周期作为主要周期。在图 5-2 中为了直观起见，我们将频率换算成周期。

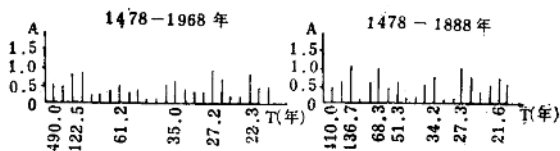


图 5-2 华北历史地震 10 年 $\lg \Sigma E$ 富氏谱

在震源机制分析中，根据有限移动源的模型，利用各台站记录的瑞利波垂直分量资料，求出其振幅谱中第一个极小值所对应的周期，可以得到断层面的空间位置、断层长度及破裂速度，利用振幅谱中长周期的振幅值，并能求得地震力距等重要参数，具体做法可参阅国家地震局组织的震源机制会战组第二阶段总结。图 5-3 是 1968 年 3 月 3 日四川理塘地震长

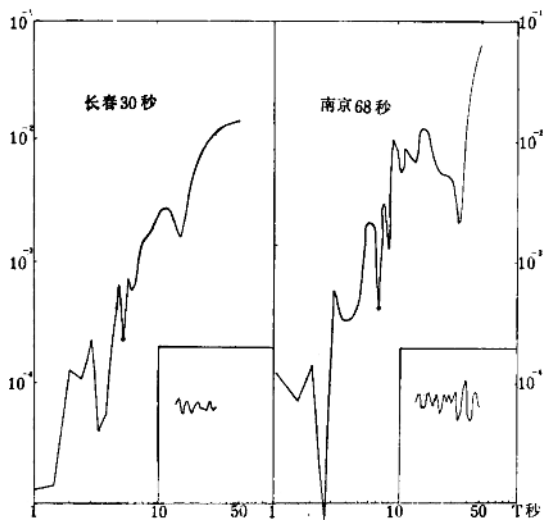


图 5-3

春台及南京台记录的端利波垂直分量的振幅谱，图中右下角方框内为端利波垂直分量的原始记录。

图 5-4 为山东某地大地电场的水平分量和正交的磁场水平分量的振幅谱，利用同一周期电及磁的谐波振幅之比，可以计算出相应周期的视电阻率，对视电阻率一周曲线与理论量板进行一定的分析对比，可以得出测点剖面的电性分层。

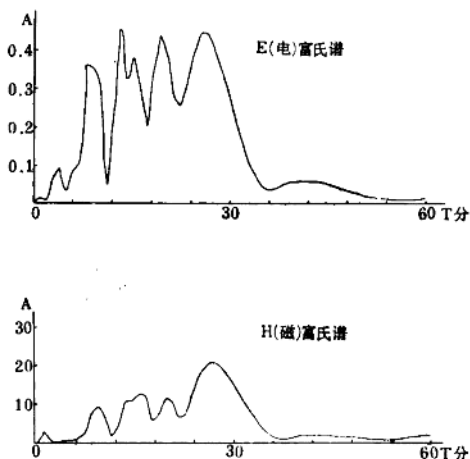


图 5-4

§5-3 功率谱

在实际资料分析中，有时用经典的富氏分析得出的效果不好，它的方差不随时间系列长度的增加而减少，其原因这是由于实际时间序列的信号和噪声的振幅相位频率变化的随机性质所造成的。因此从统计的角度出发，引进了认为是适合于时间序列随机性质的谱分析方法—功率谱，它是把富氏分析方法和统计分析方法两者结合起来。

在功率谱的计算中，假定时间序列是属于平稳随机过程，

即序列的平均值为零,并且不同时间段内方差相近,也就是说序列的分布性质不受时间起点的影响。按 1959 年 Blakman 和 Tukey 所建议的方法,计算功率谱通过自协方差函数。

对一连续函数 $x(t)$, 若其平均值为零,那么自协方差函数定义为:

$$W(p) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{+T/2} x(t)x(t+p)dt \quad (5-13)$$

其中 p 为时间滞后。当 $p=0$ 时 $W(0)$ 为序列的方差, $W(p)/W(0)$ 称为序列的自相关函数。

可以证明自协方差函数 $W(p)$ 与序列的谱分布函数 $L(f)$ 互为富氏变换。如果考虑

$$f = \frac{1}{T}, \quad \omega = \frac{2\pi}{T} = 2\pi f, \quad d\omega = 2\pi df,$$

则有

$$W(p) = \int_{-\infty}^{+\infty} L(f)e^{i2\pi fp}df \quad (5-14)$$

$$L(f) = \int_{-\infty}^{+\infty} W(p)e^{-i2\pi fp}dp \quad (5-15)$$

根据平稳随机过程的自协方差函数定义,可以证明分布函数和时间序列之间关系为:

$$L(f) = \lim_{T \rightarrow \infty} \frac{1}{T} \left| \int_{-T/2}^{T/2} x(t)e^{-i2\pi ft} dt \right|^2$$

$L(f)$ 代表了 $x(t)$ 在频率区间 f 到 Δf 之间的方差分布,所以 $L(f)$ 称为 $x(t)$ 的协方差谱,又称为 $x(t)$ 的功率谱。因为如果我们把 $x(t)$ 看作是通过一欧姆电阻所需的电压,则长期消耗于电阻的平均功率正比于 $x(t)$ 的方差。

如果 $x(t)$ 是实的, $W(p)$ 也是实的,而且对称于 $p=0$,那么 $W(p)$ 和 $L(f)$ 可以简单地表示为单边余弦变换:

$$W(p) = 2 \int_0^{+\infty} L(f) \cos 2\pi fp df \quad (5-16)$$

$$L(f) = 2 \int_0^{+\infty} W(p) \cos 2\pi f p dp \quad (5-17)$$

在实际应用中,对一离散的时间序列 $x_t, t = 1, 2, 3, \dots, N$, 计算自协方差函数 $W(p)$ 及功率谱 $L(f)$ 的近似估计值公式为:

$$W_p = \frac{1}{N-p} \sum_{t=1}^{N-p} (x_t - \bar{x})(x_{t+p} - \bar{x}) \quad (5-18)$$

这里, $\bar{x} = \sum_{t=1}^N x_t / N$ 为序列 x_t 的平均值,

N ——为等时间间隔的数据个数,

$p = 0, 1, 2, \dots, M$, 称为数据的时间滞后,

M ——为最大时间滞后,一般建议取 $M = 0.1 \sim 0.2N$

$$L(f) = W_0 + 2 \sum_{p=1}^{M-1} W_p \cos \frac{\pi j p}{M} + W_M \cos \pi j \quad (5-19)$$

$$j = 1, 2, \dots, M$$

功率谱的计算,目前除了用(5-19)式外,不少人直接应用富氏变换的平方量,对不同样本的富氏变换的平方量进行平均或光滑,作为时间序列功率谱的估计。

图 5-5 是华北历史地震资料的功率谱,按每年地震能量和的对数作为原始数据,计算中数据个数 $N = 490$, 取最大时间滞后 $M = 50$ 。先按 (5-18) 式计算出自协方差序列

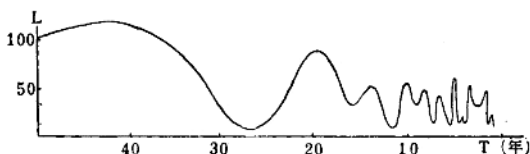


图 5-5 华北历史地震每年 $I_g \Sigma E$ 功率谱
($M = 52, 1478-1968$ 年)

$W(p)$, $p = 0, 1, 2, \dots, 50$, 再按 (5-19) 式计算功率谱 $L(f)$ 。

§5-4 问题讨论

1. 如果时间序列中包含有谱分析所能介得的频率以外的成份, 那么分析前必须采取一定的处理方法去除, 否则在频谱中带进的复杂成份将会大大降低分析精度, 给分析造成困难。图 5-6 为大地电磁测深资料去除趋势前后谱分析结果。从图上可以看出, 除了个别峰值外, 大多数峰值位置都有错动, 而且各分波振幅相差也很大。

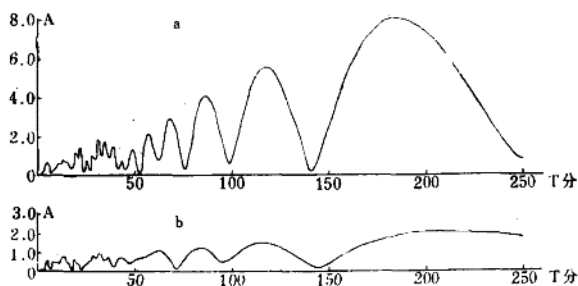


图 5-6

2. 关于分析所需的时间序列长度问题。从函数逼近角度来看, 只要分波项数足够多, 精度是可以满足要求的, 序列长度不成为问题。但从研究频率成份角度出发, 如果序列是属于平稳随机过程, 那么序列越长, 分析精度也就越高, 但往往限于客观条件, 资料长度有限, 究竟最低限度长度如何确定, 关于这方面有不少讨论, 主要取决于所要研究的频率范围, 富氏分析中所能介得的频率最低为 $1/T$, 最高为 $1/2\Delta t$, 或者最大周期为 T , 最小周期为 $2\Delta T$, 一般建议分析长度为所须研究的最大周期的 8—10 倍。但在下节讨论中将提到的, 如果

进行加密计算, 所需序列长度就可短得多。

3. 在地震工作中, 当我们研究某时间序列的频率成份或周期情况时, 往往要求知道主要分波的周期越精确越好, 例如分析历史地震资料, 如果能确切知道某地区历史地震存在那些周期, 就便于进一步研究地震成因的天文或其他背景, 并以此来作中长期预报。在前兆数据分析中也同样, 周期知道确切, 便于分析其与别的因素关系, 并可设计合适的过滤器排除干扰或提取信息。

在富氏分析中, 一般各谐波的频率是等间隔的, 而且把所取序列长度作为第一分波的周期, 因此各分波的周期带有很大的的人为因素, 如果实际存在的分波周期和谱分析所能介得的周期不一致, 例如介于某二个分波之间, 则从振幅频谱上很难找出主要周期, 或者很不准确。为了介决此矛盾, 一般文献中建议采用的方法是改变分析所用序列长度, 即每次分析时增加或减少一些数据, 对所得的一组频谱曲线中稳定出现的较大振幅进行分析, 或求其周期平均值作为主要周期。这样做一则计算量大, 二则如果低频部分存在有主要周期, 还是比较难找准。

为此, 在实际计算中, 我们试用了不是按频率等间隔求系数 a_j 、 b_j , 而是按周期等间隔, 对系数 a_j 、 b_j 按所要求的周期精度进行加密计算, 把(5-12)式改为:

$$\begin{aligned} a_{T_j} &= \frac{2}{2N} \sum_{k=0}^{2N-1} y_k \cos \frac{2\pi k}{T_j} \\ b_{T_j} &= \frac{2}{2N} \sum_{k=0}^{2N-1} y_k \sin \frac{2\pi k}{T_j} \end{aligned} \quad (5-20)$$

T_j 的变化根据所需周期精度来定, 可以是整数, 也可以是小数。

如果已知序列较短, 所要研究的是频谱分析结果中低频

部分,那么从我们计算结果来看,这样做一是可以提高效率,不须反复算;二是可以提高低频部分分析精度;三是序列长度比频率等间隔计算中相对可以短些,约为所须分析最大周期的2—3倍。

4. 在实际资料分析中,所得振幅谱往往很复杂,我们计算了周期为20的梯形脉冲试验数据的功率谱,结果除了主要周期20外,尚有10、7、5等周期,说明非正弦周期函数的振幅频谱中,不是一个峰值,而是一组峰值,所以在实际分析中,如何判断去除一些伪周期是很重要的。谱分析工作随着电子计算机的普及,近年来应用越来越广,但如何提高分析精度,仍然是个大问题,尤其是对一些复杂的时间序列,如地震及其前兆数据。上海天文台及复旦大学数学系的同志用周期图方法找寻观测数据中隐蔽周期方面已摸索了一些方法,取得了较好的效果,读者如果需要的话,可以参阅他们写的“地球自转速率不规则变化的周期分析”报告,这儿就不作介绍了。

5. 方法对比:为了比较不同方法的差异,我们把几个已知周期的余弦波迭加,作为原始数据进行计算,即:

$$y_k = \sum_{T_i} \cos \frac{2\pi k}{T_i} \quad k = 0, 1, 2, \dots, 2N - 1$$

T_i 为各分波 n 已知周期,我们取 $N = 100$,原始数据个数200, $T_i = 1.6, 22, 24, 74, 150$ 。我们分别计算了周期等间隔的功率谱,最大时间滞后取为序列长度的1/2及1/5,以及频率等间隔和周期等间隔的富氏谱。计算结果列于表5-4中。

从表5-4可以看出功率谱受 M 值的大小,也就是自协方差函数系列长短影响较大。从功率谱所能得到的最大周期约为 $2M$,所以当 M 取值较小时,如 $M = 2N/5$ 时,大于 $2M$ 的周期就不可能分析出来。当 M 取值较大时,如 $M = 2N/2$,较大周期能得到一定反映,但误差较大,如与已知周期72、150

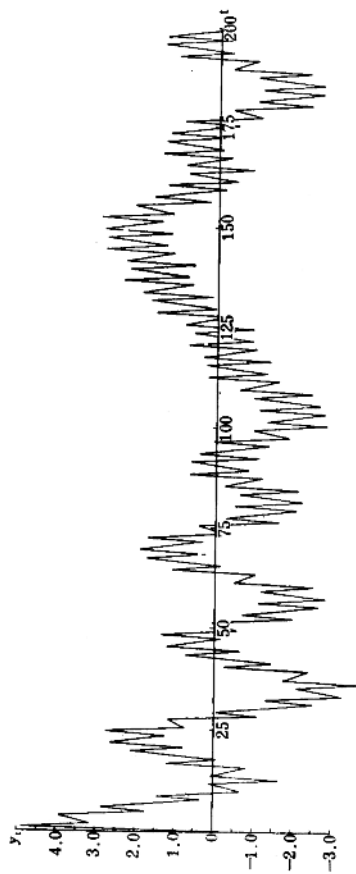


图 5-7



图 5-8

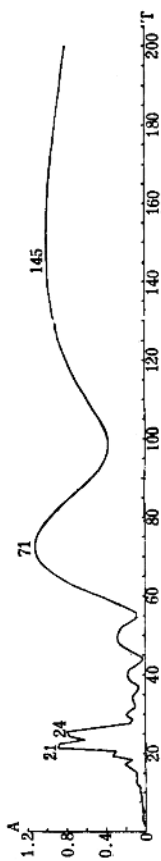


图 5-9

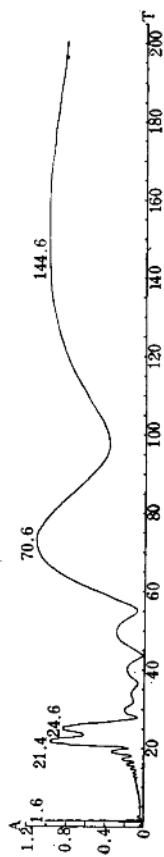


图 5-10

表 5-4

已知 周期	计算所得主要周期					
	功率谱(周期等间隔)			富氏谱		
	$M=2N/2$	$M=2N/5$		频率 等间隔	周期等间隔	
		$\Delta T=1$	$\Delta T=0.2$		$\Delta T=1$	$\Delta T=0.2$
1.6	—	—	1.6	2.6	—	1.6
22	22	23	22.8	22.2	21	21.4
24	—	—	—	25	24	24.6
72	66	—	—	66.6	71	70.6
150	157	—	—	200	145	144.6

相对应的有周期为 66 及 157 的峰值。所以在功率谱分析中, M 取多大合适, 是个比较复杂的问题, 要根据具体情况分析决定。在富氏谱的计算中, 按频率等间隔分析结果误差较大, 尤其是长周期部分, 这也说明按频率等间隔的富氏分析要求序列长度至少为所须分析周期的 8—10 倍。按周期等间隔计算的效果较好, 尤其是长周期。从功率谱及富氏谱计算中也可以看出, 如果周期间隔 ΔT 取得小些, 也可提高精度及分辨能力, 图 5-7 为本试验数据时间序列, 图 5-8 到图 5-10 为富氏分析结果。图中峰值上所标数字为各峰值所对应的周期。