

地震预报的证据理论方法

石雅鏐¹, 周民都², 张晓东³, 宁书年¹

(1. 中国矿业大学机电与信息工程学院 北京 100083;

2. 中国地震局地震预测研究所兰州科技创新基地 甘肃 兰州 730000;

3. 中国地震局地震预测研究所 北京 100081)

摘要 概率论和数理统计方法一直是地震预报中处理不确定性的主要方法,但在很多情况下并不是特别有效。本文引进了一种新的不确定性处理方法—证据理论方法,介绍了如何用该理论表达地震预报问题。讨论了应用其 D-S 合成规则进行预报的一般过程,并根据地震综合预报的特点,提出了加权的 D-S 规则。研究表明,证据理论应用于地震预报有其独特的优点。

关键词 证据理论 不确定性 地震预报 加权 D-S 规则

中图分类号 P315.75

文献标识码 A

文章编号 1000-0844(2004)04-0289-04

0 引言

地震孕育物理过程的复杂性决定了地震预报信息的不确定性。不确定性是相对命题而言的,其类型包括信息的不可靠、不完备、不精确和不一致^[1]。在地震预报中处理不确定性一直主要依赖概率论和数理统计方法,虽然取得了很多成果,但面对不断出现的不确定性表达和推理问题也确实面临很多困难。其中主要有:

(1) 不能综合不同信息的信念。不同信息提供的概率如何予以综合,这在原来的概率理论框架内是无法处理的。用数学的语言来叙述,问题是这样的^[2]: 给定了一个可测空间 (Ω, F) , 把一种信念看成是 (Ω, F) 上的一个概率测度,如果有 K 个信息,就给出了 K 个测度 $P_i, i = 1, 2, \dots, K$ 。所谓综合,就是要从 $P_1, \dots, P_K$ 去获得一个概率测度 P , P 还是 (Ω, F) 上的测度,常用的一种办法是对 P_i 赋以权 $\omega_i, \omega_i \geq 0, \sum_{i=1}^k \omega_i = 1$, 用 $\sum_{i=1}^k \omega_i P_i$ 作为综合后的 P 。且不说如何去定 ω_i 就有困难,从直觉上看,这一类 P 是不合要求的。因为不论怎样选择 ω_i , 总有

$$\min_{1 \leq i \leq k} P_i \leq \sum_{i=1}^k \omega_i P_i \leq \max_{1 \leq i \leq k} P_i$$

人们的直觉是,如果不同的信息都肯定了某一事件出现的可能性较大,那么人的信念就会增大,它比每个信息提供的信念会更大,而上述不等式是与这个直觉不一致的。

(2) 依赖命题的独立性。这往往使推理计算受到影响。如:

已知 $P(A) = 0.75; P(\neg A) = 0.25$ 。

于是 $P(A \cap \neg A) = P(A) \cdot P(\neg A) = 0.1875$;

$P(A \cup \neg A) = P(A) + P(\neg A) - P(A) \cdot P(\neg A) = 0.8125$ 。

但依概率公理 $P(A \cap \neg A) = 0; P(A \cup \neg A) = 1$ 。

(3) 依赖先验概率。Bayes 公式是使用最广泛的概率方法,但在计算 $P(B_i/A)$ 时,首先是要求 B_i 到

B_n 相互无关难以得到保证,其次是 $P(A/B_i)$ 和 $P(B_i)$ 难以获取。

正是因为概率论的这些问题,促进了不确定性理论和方法的发展,使得与其相关的一些新的数值和非数值方法自上世纪后期以来不断涌现。其中证据理论被认为是最有前途的数据融合和知识处理方法。该理论的最为突出的特点是能表达由于“无知”而引起的不确定性。对于地震预报,前兆场产生的原因往往是不完全清楚的,即存在不同程度的无知,因此需要这一理论来表达其不确定性。证据理论的另一突出特点是融合数据能力强。地震预报的主要形式是专家会商,而如何汇总专家的意见在概率的框架内很困难。使用证据理论的合成规则就能很好地解决这一问题。总之,该理论应用于地震预报有其独特的优点。证据理论是 Dempster 在研究统计问题时提出的^[3-6],Shafer 把它推广到更一般的情形^[6-7]。这一理论虽然有三十多年的历史,但得到广泛应用则是在最近几年。本文回避其中涉及的复杂数学和算法问题,仅根据以上两个特点讨论其在地震预报的基本应用框架。

1 地震预报的证据表达

设 Ω 是一限集,它的一切子集,即 Ω 的幂集记为 2^Ω 是 Ω 到 $[0, 1]$ 上的函数,满足 (i) $m(\emptyset) = 0$; (ii) 对 $E \in 2^\Omega$ $m(E) \geq 0$ 并且 $\sum m(E) = 1$ 。 $m(\cdot)$ 称为是 Ω 上 E 的基本概率赋值。对给定的 m , 对任一 $A \in 2^\Omega$ 定义它所相应的信任函数如下:

$$\text{Bel}(A) = \sum_{B \subseteq A} m(B) \quad (1)$$

定义似真函数为

$$\text{Pl}(A) = 1 - \text{Bel}(\bar{A}) \quad (2)$$

$m(\cdot)$ 实际上是定义在 Ω 上的概率,但不是单个事件的概率,它反映了人们的信念不确定性程度,而信任函数 $\text{Bel}(A)$ 是支持 A 的最小值,似真函数 $\text{Pl}(A)$ 是支持 A 信念的最大值。这样 $[\text{Bel}(A), \text{Pl}(A)]$ 就自然形成了对 A 的信念区间(不等式 $\text{Pl}(A) \geq \text{Bel}(A)$ 是容易证明的)。

下面通过一个具体的示例说明以上框架定义的含义。由于问题的核心是确立基本概率赋值,以下的讨论集中在这一点。

设有两个罐,第一个罐中全是红球(记为 R),第二个罐中装的是红球和白球(记为 W),但不知这个罐中红白球的比例。现随机从这两个罐中选一个,再在罐中取一个球。设选第一个罐的概率是 0.7,选第二个罐的概率是 0.3。怎样用信任函数来描述呢?选 $\Omega = \{W, R\}$ 基本概率赋值 $m(\cdot)$ 定义如下:

$$m(\{R\}) = 0.7; \quad m(\{W\}) = 0.3.$$

于是 $\text{Bel}(\{R\}) = 0.7$, $\text{Pl}(\{R\}) = 1$;

$$\text{Bel}(\{W\}) = 0, \quad \text{Pl}(\{W\}) = 0.3.$$

从取到红球这一事件来看,它的概率在 $[0.7, 1]$ 之内;从取白球这一事件来看,它的概率在 $[0, 0.3]$ 之内。这一结论很符合人的直觉。如果用概率的框架来描述对于第二个罐中红、白球各占的比例不清楚时,就比较难以处理,因为这时需要多引入一个参数(白球占的比例)才能加以讨论。而信任函数框架的优点是可以对 Ω 赋以基本概率 0.3 来反映我们对白球无知的状况。这一点对习惯于在样本点上定义基本概率的人,一开始可能不习惯。

在地震预报中,对某种观测异常很多情况下并不知道它对应哪一级地震时,用这种方式赋值就易于处理。也就是说,在地震预报中基本概率赋值 m 可通过集值统计获得。例如某台站 5 年内一共观测了 10 次异常,其中有 6 次不对应地震(在这里是指一定时空范围内没有大于 5 级以上地震);有 2 次对应 5~7 级强地震;有 2 次对应 7 级以上大震,其中有一次是孤震,另一次发生了一次 5.5 级余震。用 A 表示本文的无震, B 表示强震, C 表示大震,即该问题的识别框架为 $\Omega = \{A, B, C\}$ 。则该指标异常与地震的信任关系可表示如下:

$$m(A) = 0.6, \quad m(B) = 0.2, \quad m(C) = 0.1, \quad m(\{B, C\}) = 0.1$$

其它子集的 m 值为零。这样下次观测到异常后就可根据以上基本概率赋值,利用(1)和(2)式计算各级地震的信任区间,从而进行地震预报。

2 地震预报的信息综合

信任函数之所以受到重视,与它有一个较为合理的合成规则有关。这个规则是由 Dempster 和 Shafer 给出的,所以简称为 D-S 规则。设在 $(\mathbb{H}, \Omega)$ 上有两个信任函数 $\text{Bel}_i, i = 1, 2$, 它们相应的概率赋值是 $m_i, i = 1, 2$, 如何去合成一个 $(\mathbb{H}, \Omega)$ 上的 Bel 呢? 这正是地震预报信息综合中的一个基本问题。根据 D-S 规则,令

$$m(A) = \sum_{A_1 \cap A_2} m_1(A_1) m_2(A_2) / (1 - K) \quad (3)$$

其中
$$K = \sum_{A_1 \cap A_2 = \phi} m_1(A_1) m_2(A_2)$$

则 m 所对应 Bel 就为 Bel_1 与 Bel_2 的综合。

现在来对(3)给出的规则作一些解释。由于 $\sum_{A_i \subseteq H} m_j(A_i) = 1, j = 1, 2$, 因此

$$1 = \sum_{A_1 \subseteq H} m_1(A_1) \sum_{A_2 \subseteq H} m_2(A_2) = \sum_{A_1 \subseteq H} m_1(A_1) m_2(A_2) \sum_{A_1 \cap A_2 = \phi} m_1(A_1) + m_2(A_2) + \sum_{A_1 \cap A_2 \neq \phi} m_1(A_1) m_2(A_2)$$

$$A_i \subseteq H, i = 1, 2$$

所以用 $m_1(A_1) m_2(A_2)$ 来描述合成的信念时,它会丢失了 $K = \sum_{A_1 \cap A_2 = \phi} m_1(A_1) m_2(A_2)$ 的这一部分,因而总量就成为 $1 - K, A_i \subseteq \mathbb{H}, i = 1, 2$ 。 $\sum_{A_1 \cap A_2 = A} m_1(A_1) m_2(A_2)$ 表达了两个 m_i 共同支持 A 的信念的部分,因此用它在 $1 - K$ 中的比例来反映就是合理的。

D-S 规则反映了信念的重新分配,从表达式 $\sum_{A_1 \cap A_2 = A} m_1(A_1) m_2(A_2)$ 可以看出,尽管 $m_1(A_1) > 0$ 表示 A_1 中有支持 A 信念的内容,但由于 $m_2(A_2) = 0$, 比如 $A_2 = A \cup B$, 这样乘积 $m_1(A_1) m_2(A_2)$ 就是零。所以 $m_1(A_1)$ 的内容分配给 $m_2(A_2) > 0$ 所支持的那一部分,这样就是信念的一种调整。可以证明 D-S 合成规则是可交换的,有结合律的,这样可以合成多个 Bel , 并与合成的次序无关。

D-S 合成规则为地震预报提供了一个很好的信息融合方法。设讨论的地震为三种:无震(包括弱震)、强震和大震,分别用 A, B 和 C 表示。即 $\Omega = \{A, B, C\}$ 。甲专家(指标)给的基本概率赋值为 $m_1(A) = \frac{1}{2}, m_1(B) = \frac{3}{8}, m_1(C) = \frac{1}{8}$; 乙专家(指标)给的基本概率赋值为 $m_2(A) = \frac{1}{2}, m_2(B) = \frac{3}{8}, m_2(C) = \frac{1}{8}$ 。运用 D-S 规则可得 $m(A) = \frac{8}{13}, m(B) = \frac{7}{26}, m(C) = \frac{3}{26}$ 。

由此我们看到对 A 的支持强度大大增加,这符合问题的解释。

在地震预报中,各专家的意见(或各观测指标)所起的作用是不一样的。为此我们设计了加权 D-S 规则。假设有 n 个专家,权重比为 $\lambda_1, \lambda_2, \dots, \lambda_n$, 其中的 λ_i 为正整数。则综合这 n 位专家的意见的组合规则是:先分别组合 λ_i 个独立同信任函数的第 i 位专家的组合信任函数,然后再两两组合即可。如:设识别框架 $\Omega = \{A, B, C\}$ (意义同前例),若基于两组不同指标而导出的基本概率赋值函数分别为

$$m_1(A) = 0.4, m_1(\{A, C\}) = 0.4, m_1(\{A, B, C\}) = 0.2, m_2(A) = 0.6, m_2(\{A, B, C\}) = 0.4.$$

且 m_1 与 m_2 的权重比为 3:1。则我们可用如下方法进行组合:

先由两个独立的 m_1 与 m_1 组合,可得到组合后的基本概率赋值函数为

$$m'_1(A) = 3 \times 0.4 \times 0.4 + 2 \times 0.4 \times 0.4 = 0.64,$$

$$m'_1(\{A, C\}) = 0.4 \times 0.4 + 2 \times 0.2 \times 0.4 = 0.32,$$

$$m'_1(\{A, B, C\}) = 0.2 \times 0.2 = 0.04.$$

由此综合所得的基本概率赋值函数 m_1 与 m_1 再去综合,可得

$$m''_1(A) = 2 \times 0.64 \times 0.4 + 0.64 \times 0.2 + 0.32 \times 0.4 + 0.04 \times 0.4 = 0.78,$$

$$m''_1(\{A, C\}) = 0.32 \times 0.4 + 0.32 \times 0.2 + 0.04 \times 0.4 = 0.208,$$

$$m''_1(\{A, B, C\}) = 0.04 \times 0.2 = 0.08。$$

然后 m_1 与 m_2 再组合可得

$$m(A) = 0.9136, m(\{A, C\}) = 0.0832, m(\{A, B, C\}) = 0.0032。$$

最后结果大大提高了对无震的支持度 ($Bel > 0.9136$)。符合专家定性决策。

在信息源比较大时,使用 D-S 规则可能出现组合爆炸,目前已有可行算法。在地震预报中,当涉及的指标只有五种左右时,不会产生这一问题。总之,证据理论确实为我们提供了一个有效的地震综合预报定量方法。

3 结论

根据以上的讨论,我们可得出以下结论:

(1) 证据理论能表达和处理概率框架不能表达和处理的地震预报问题,能定量地模拟专家决策。

(2) 应用证据理论进行地震预报的基本方法和步骤是:首先对要预报的地震用识别标架表示,然后对各前兆指标的对应情况(用集值统计等)建立概率分布函数,最后使用公式(1)、(2)和 D-S 规则得出各震级的单指标预报信度或多指标总信度。

[参考文献]

- [1] Bhatnager R G, Kanal L N. Handling Uncertain Information: A Review of Numeric and Non-numeric Methods[A]. in: Uncertainty in Artificial Intelligence[C]. New York: North-Holland, 2002. 3-26.
- [2] Fox J. Three Arguments for Extending the Framework of Probability[A]. in: Uncertainty in Artificial Intelligence[C]. New York: North-Holland, 1996. 447-465.
- [3] Dempster A P. Upper and Lower Probabilities Induced by a Multivalued Mapping[J]. Ann. Math. Statist., 1967, **38**: 325-339.
- [4] Dempster A P. A Generalization of Bayesian Inference[J]. J. of Royal Statistical Society(Series B), 1968, **30**(2): 205-247.
- [5] Dempster A P, Kong A. Uncertain Evidence and Artificial Analysis[J]. J. Statist. Plan. & Info., 1988, **20**: 355-368.
- [6] Shafer G. A Mathematical Theory of Evidence[M]. Princeton University Press, 1976.
- [7] Shafer G. Belief Functions and Parametric Models[J], J. Royal Statist. Soc.(Series B), 1982, **44**(3): 322-352.

METHODOLOGY OF EVIDENCE THEORY APPLYING FOR EARTHQUAKE PREDICTION

SHI Ya-liu¹, ZHOU Min-du², ZHANG Xiao-dong³, NING Shu-nian¹

(1. College of MEI, China University of Mining Technology, Beijing 100083, China;

2. Lanzhou Base of Institute of Earthquake Prediction, CEA, Lanzhou 730000, China;

3. Institute of Earthquake Prediction, CEA, Beijing 100081, China)

Abstract For handling information uncertainty in earthquake prediction, possibility theory and statistics have been the primary method. But in many cases, this method can't produce effective results. For this reason, a new uncertainty method—Evidence theory is introduced in this paper. At first, the representation of uncertainty for earthquake prediction under the frame of this theory is introduced. Then the general procedure of prediction with the D-S combination rule is discussed, and in accordance with the earthquake synthesizing prediction need, the weighted D-S rule is produced. From our study, it is concluded that the method introduced holds its own effectiveness in earthquake prediction.

Key words: Evidence theory; Uncertainty; Earthquake prediction; Weighted D-S rule