第18卷 第3期
2001年9月

华 东 交 通 大 学 学 报
Journal of East China Jiaotong University

Vol.18 No.3
Sep. 2001

文章编号:1005-0523(2001)03-0021-04

基于信息熵理论的有限元法

陈 虬, 雷震宇

(西南交通大学 应用力学与工程系, 四川 成都 610031)

摘要: 研究基于信息熵理论的有限元法19.利用随机熵和模糊熵的概念,建立了模糊变量向等效的随机变量的转换关系19.这样,可以利用随机有限元法进行模糊结构或模糊随机结构的有限元分析19.算例显示方法是有效的19.

关 键 词: 熵; 模糊; 随机; 有限元法

中图分类号: TB115 **文献标识码:** A

0 引 言

在航空、海洋、机械、土木、交通、国防等众多工程领域中19.近年来人们已不满足于对结构的确定性必分析所得的计算结果,而是要求计及随机、模糊等不确定因素的影响19.在结构分析中引入随机性和模糊性的概念,着力于研究能更加接近描述结构真实工作行为的新的分析方法[1,2] 19美国佛罗里达大西洋大学的I·Elishakoff教授在《应用力学评论》上发表了很好的综述文章[3],论述了有关模糊随机有限元法方面的研究进展和应用情况19.A·Haldar[4]认为对在役结构的可靠性分析中,应该同时考虑模糊和随机因素的影响,并提出利用熵作为一个中间量处理模糊随机结构中的有限元计算问题19.作者在[5]中国研究了一种基于信息熵概念的新的模糊有限元法19.本文通过信息熵的概念,将随机和模糊两个主要的不确定因素,给出统一的测度19在此基础上给出基于信息熵理论的模糊随机有限元法,导出相应的模糊随机有限元的递归方程组19.算例显示本文的方法是有效的19.

1 概率熵和模糊熵

信息论的建立是与概率理论紧密联系的19.信息论的奠基人Shanmon[6]认为,信息本质上是随机的,并借助热力学中"熵"的概念提出用信息熵对概率信息进行度量,不确定性越大,熵就越大,于是将这类熵称为概率熵19.

概率的定义:离散随机变量 $X=\{x_1, x_2, \cdots\cdots x_n\}$,其概率分别为 $p_1, p_2, \cdots p_n$,该随机变量的平均不确定性可以由概率熵 $H(p_1, p_2, \cdots p_n)$ 表示,定义式为

$$H=-\sum_{i=1}^{n} P_i \log P_i \tag{1}$$

概率熵 H 有如下性质

· H 总是正的;

· H 在某一事件等于1而所有其他事件为零的极端情况下等于零⒀这就是那种有关实验和数量事先均已知,而其结果并不会带来任何新的信息时的情况;

· 在所有概率相等的条件下,H 存在最大值⒀

当 X 为连续随机变量时,其概率密度函数为 $P(x)$,则概率熵 $H(x)$ 定义为

$$H=-\int_x P(x)\ln P(x)\,\mathrm{d}x \tag{2}$$

若随机变量 X 服从正态分布 $N(m, \sigma^2)$ 时,相应的概率熵为

$$H=\ln(\sqrt{2\pi e}\cdot\sigma \tag{3}$$

文[7]中,提出非概率熵的概念⒀在非概念熵中主要是讨论模糊熵,若离散模糊变量 $X=\{x_1, x_2, \cdots x_n\}$,其中元素所对应的隶属度分别为 $f(x_1)$、$f(x_2)$、$\cdots f(x_i)$ 对应的模糊熵用 $G_i(f(x_i))$ 表示,它必须满足下列条件

· $G_i=0$ 当且仅当 $f=0$ 或 $f=1$;

· G_I 达到最大值当且仅当 $f=1/2$;

收稿日期:2000-05-28
作者简介:陈虬(1940-),男,江苏苏州人,教授⒀

· $G_i \geqslant G_j$ 当 $f_j \geqslant f_i \geqslant 1/2$ 或者 $f_j \leqslant f_i \leqslant 1/2$ 时⒀

若 A 为 X 上的模糊集，即表示映射 $X \to [0,1]$⒀总的模糊熵为

$$G(A) = K\sum_{i=1}^{n} G_i \quad (4)$$

对于离散型模糊函数，其模糊熵定义为

$$G(A) = -K\sum_{i=1}^{n}[f_A(x_i)\ln f_A(x_i) + (1-f_A(x_i))\ln(1-f_A(x_i))] \quad (5)$$

对于连续型模糊函数，其模糊熵定义为

$$G(A) = -\int_x (f(x)\ln f(x) + (1-f(x))\ln(1-f(x))\,\mathrm{d}x \quad (6)$$

模糊熵的另一种形式是[7]，

$$G(A) = -\int_x f'(x)\ln f'(x)\,\mathrm{d}x \quad (7)$$

式中 $f'(x) = f(x)/\int_x f(x)\,\mathrm{d}x$，显然 $f'(x)$ 满足 $\int_x f'(x)\,\mathrm{d}x = 1$，即满足归一化的要求⒀

2 不确定信息的统一测度

概率信息和模糊信息是两类最主要的不确信息，它们右以通过引入一个新的度量概念－肯定度来统一⒀

肯定度的定义是：考虑一个抽象试验 X，它具有 N 种可能结果：$x_1, x_2, \cdots\cdots x_N$，则把 X 取某种具体可能结果 x_n 的可能性、机会或程度，称为 x_n 的肯定度，记为 $C_n(n=1,2,\cdots\cdots N)$⒀显然，如果 X 是概率型试验，此时 C_n 就是概率 P_n⒀然而，如果 X 不是概率型试验，概率量就不存在，但肯定度的概念仍然有效⒀当 X 是模糊型试验时，肯定度指的是 x_n 的隶属度⒀

肯定度 $C_n(n=1,2,\cdots\cdots N)$ 的集合称为肯定度分布，记为 C，即 $C=\{c_1, c_2, c_N\}$⒀若 $C_n = \frac{1}{N}$，$(n=1,2,\cdots\cdots N)$，则称这种分布为均匀的肯定度分布，记为 C_0；若 C_n 中某个元素为 1，其余元素为零，则称这种分布为 0－1 型的肯定度分布，记为 C_s⒀C_0 和 C_s 是经常遇到的两种极端类型的分布⒀对于给定的肯定度分布 C，可以构造一个函数

$$M\bullet c) = \bullet^{-1}\left\{\sum_{n=1}^{N} C_n \bullet c_n)\right\} \quad (8)$$

称 $M\bullet C)$ 为关于 C 的平均肯定度⒀其中 ●是一个未知的单调连续函数，$\bullet^{-1}$ 是 ●的逆函数，也是单调连续的⒀

考察由一个试验和一个观察者组成的系数，记为 $(X, c, c^*; R)$，其中 R 表示观察者，(X, c, c^*) 表示试验过程，C 是观察者 R 关于试验的先验肯定度广义分布，C^* 则是 R 关于 X 的后验肯定度广义分布⒀

进一步有

$$I(c) = \log\frac{M\bullet c)}{M\bullet c_0)} = \log N + \sum_{n=1}^{N} c_n \log c_n \quad (9)$$

$$I(c^*) = \log\frac{M\bullet c^*)}{M\bullet C_0^*)} = \log N + \sum_{n=1}^{N} c_n^* \log c_n^* \quad (10)$$

分别称 $I(c)$ 和 $I(c^*)$ 为试验系统的对数先验相对平均肯定度和对数后验相对平均肯定度⒀于是观察者 R 从试验系统 (X, c, c^*) 中得到的信息量 $I(c, c^*; R)$ 是 R 通过观察所实现的关于 X 的对数相对平均肯定度的增量，即

$$I(c, c^*; R) = I(c^*) - I(c) = \log\frac{M\bullet c^*)}{M\bullet c)} = \sum_{n=1}^{N} c_n^* \log c_n^* - \sum_{n=1}^{N} c_n \log c_n \quad (11)$$

若试验 X 是一个纯粹的概率型试验，则肯定度即为概率，即 $I(c, c^*; R) = I(P, P^*; R)$，这里 P 和 P^* 分别是 X 的先验和后验概率分布⒀另外，一旦试验完成，产生的将是一个已发生的事件，从而成为确定性的⒀因此，后验概率分布 P^* 是一个 0－1 型分布形式⒀于是观察者 R 从试验中得到的信息量为

$$I(P, P_s^*; R) = \log\frac{M\bullet p_s^*)}{M\bullet p_0^*)} - \log\frac{M\bullet P)}{M\bullet P_0)} = -\sum_{n=1}^{N} P_n \log P_n \quad (12)$$

显然这就是前述的概率熵⒀

考察一个模糊型试验 $(x, F, F^*; R)$，可以把 f_n，$(1-f_n)$ 看作是 x_n 的肯定度分布⒀于是

$$I(c, c^*; R) = \frac{1}{N}\sum_{n=1}^{N} I(c_n, c_n^*, R)$$

$$= \frac{1}{N}\sum_{n-1}^{N}[f_n^* \log f_n^* + (1-f_n^*)\log(1-f_n^*) - f_n \log f_n - (1-f_n)\log(1-f_n)]$$

即 $I(C, C_s^*; R) = I(F, F_s^*; R)$

$$= \frac{1}{N}\sum_{n=1}^{N}(-f_n \log f_n - (1-f_n)\log(1-f_n)) \quad (13)$$

显然这就是模糊熵⒀

综上所述，$I(c, c^*; R)$ 可以作为信息的统一测度，将概率熵、模糊熵等各种信息的度量统一起来⒀

3 基于信息熵的模糊随机有限元法

信息论的基本观点就是可以用信息熵来表征信息源的总体特征，以此代表信息源的不确定性；然后

通过随机和模糊两类信息熵保持不变的转换原则,实现模糊变量转化为相应的随机变量,这样模糊结构、模糊随机结构可近似变为等效的随机结构,再用研究比较充分的随机有限元等数值方法进行结构分析⒀

在保证熵不变的前提下,可以将模糊变量转变为等效的随机变量⒀转换的原则是等效随机变量的概率熵等于原来的模糊变量的模糊熵,即满足下式

$$H_{eq}=G \tag{14}$$

理论上可以将模糊变量转换成任意我们所期望的分布的随机变量⒀下面讨论将模糊变量转换为等效的正态随机变量的情况⒀假设模糊变量 Y 的隶属函数为 $f(y)$,经等效转换后成为正态随机变量 X⒀ X 的均值为 m,标准差为 σ⒀在转换过程中,保持变量的均值不改变;假设 X 的均值 m 等于不考虑变量模糊性时的取值⒀

由(3)式,可以得到 X 的概率熵为

$$H_{eq}=\ln(\sqrt{2\pi e}\,\sigma) \tag{15}$$

代入(14)式,可得到标准差 σ的表达式

$$\sigma=\frac{1}{\sqrt{2\pi}}e^{G-0.5} \tag{16}$$

于是模糊结构或模糊随机结构均可转换为等效的随机结构,可以采用摄动随机有限元法、Monte-Carlo 有限元法等随机分析方法进行结构分析[19].

将随机变量 X 表示为

$$X=m(1+\alpha) \tag{17}$$

式中 α是一个均值为零的随机参量,X 的随机性完全由 α来体现⒀对随机场 α进行离散化,得到相应的随机向量 $\{\alpha\}$⒀将等效的随机结构的有限元基本方程中的各矩阵,如刚度短阵$[K]$,结点位移向量 $\{\sigma\}$和载荷向量 $\{F\}$等,在 $\{\alpha\}$的均值处按泰勒级数展开,略去二阶以上项,则得到摄动递归方程组[2]⒀

$$[K]_o\{\delta\}_o=\{F\}_o \tag{18a}$$

$$[K]_o\{\delta\}_i=\{F\}_i-[K]_i\{\delta\}_o \tag{18b}$$

$$[K]_o\{\delta\}_{ij}=\{F\}_{ij}-[K]_i\{\delta\}_j-[K]_j\{\delta\}_i-[K]_{ij}\{\delta\}_o \tag{18c}$$

式中下标"o"表示各量的均值,下标"i,j"分别表示各量对 α_i和 α_j求偏导数⒀求解上述递归方程组,可以得到位移和应力的均值和协方差⒀

位移的均值和协方差为

$$E(\{\delta\})=\{\delta\}_o+\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\{\delta\}_{ij}Cov(\alpha_i,\alpha_j) \tag{19a}$$

$$Cov(\{\delta\},\{\delta\})=\sum_{i=1}^{n}\sum_{j=1}^{n}\{\delta\}_i\{\delta\}_j^T Cov(\alpha_i,\alpha_j) \tag{19b}$$

4　算例

图 1 所示重力坝,承受上游水压力和自重作用⒀假设坝体弹性模量 E_1 是随机的,其均值为 10^5 kN/m^2,变异系数为 0.1;地基的弹性模量 E_2 是一个模糊量,其隶属函数服从如图 2 所示的三角形分布⒀已知坝体的泊松比 $\mu=0.3$,材料比重为 $\gamma=2.4$ kN/m^3;地基的泊松比 $\mu=0.25$⒀

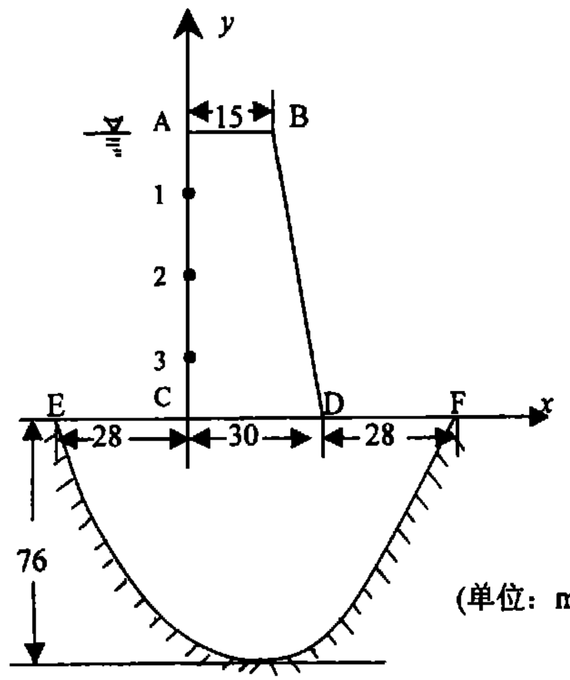

图 1　重力坝示意图

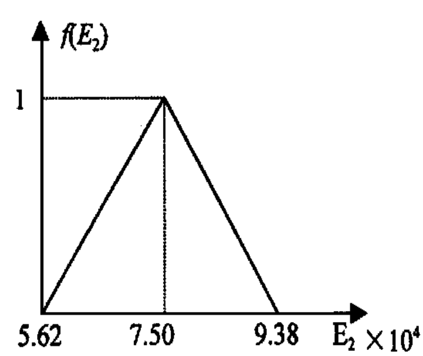

图 2　E^2 的隶属函数

首先将模糊变量 E_2 转化为随机变量⒀利用(14)、(15)和(16)式将 E_2 转化为均值 0.75×10^5 kN/m^2,变异系数为 0.1 的等效正态随机变量⒀此时,原问题转化为两种材料随机的问题,然后利用摄动随机有限元法进行结构分析⒀

计算结果给出重力坝挡水一侧 5 个节点的水平位移的均值及均方差,列于表 1⒀

表1 AC 面水平位移的均值和方差

节点	A	1	2	3	C
Y 坐标	40	28	18	8	0
u_x 均值	0.3942	0.2908	0.1786	0.0774	0.0500
u_x 均方差	0.0174	0.0119	0.0070	0.0021	0.0016

算例显示,本文研究的基于信息熵理论的模糊随机有限元法是有效的、可行的⒀本文的方法是利用信息熵理论将一个模糊结构或模糊随机结构的力学分析问题转化为通常的随机结构问题处理,从而可利用已有的随机有限元法等解法,而且得到的计算结果是一个将模糊和随机两种不确定因素统一起来的结果⒀因此这种方法是具有很好的应用前景⒀

参考文献:

[1] Klaiber M. Hien T D..The stochastic FEM:Basic Perturbation Technique and Computer Implementation,John wiley & Sons,New York,1992.

[2] 陈虬,刘先斌.随机有限元法及其应用[M].成都:西南交通大学出版社,1993.

[3] I Elishakoff,Three versions of the FEM based on concepts of either stochasticity[J].fuzziness or antioptimization,Appl.Mech.Rev.51(3),1998,209~218.

[4] A Haldar R K Reddy.A random-fuzzy analysis of existing structures[J].Fuzzy sets and systems,Vol.7,1992,201~210.

[5] 钟义信.信息科学理论[M].北京:北京邮电学院出版社,1996.

[6] E Trilles, T Rierra.Entropies in Finite Fuzzy sets[J].Inform.Sci,Vol.2,1978,159~168.

The Finite Element Method Based on Concepts of Information Entropy

CHEN Qiu, LEI Zhen-yu

(Dept.of Appl.Mech.and Eng.,Southwest Jiaotong Univ.,Chengdu 610031,China)

Abstract: In this paper,the finite element method based on concepts of information entropy is studied. Using the entropy concept, a relationship between a random-fuzzy parameter and an equivalent random parameter are presented. And hence, it can be made good result on the finite element analysis fo the fuzzy-random structures, and an example is given to show the availability and advance of the new method.

Key words: entropy;fuzziness;random;finite element method