

讲 座

Fuzzy 综合评判在生物科学中的应用

季 静 秋 冯 晋 臣

(南京计划生育管理干部学院)(南 京 林 学 院)

Fuzzy 数学是研究和处理模糊现象的最年轻的数学分支，它为解决生物数学中非实数特性的困难开辟了一条新途径。而模糊综合评判是 Fuzzy 数学的基本内容之一。为便于生物学工作者了解这一方法，本文首先就有关基本概念及 Fuzzy 综合评判的数学原理作一简单介绍，然后通过两个实例说明 Fuzzy 综合评判在生物科学中的应用。

一、Fuzzy 集合的基本概念

(一) 模糊数学并不“模糊”

Fuzzy 即“勿晰”、“模糊”，因此 Fuzzy 数学在国内就称之为“模糊数学”。乍一听，似乎不可思议，其实它真实的含义是使数学这一精确方法，打入模糊现象的禁区，去处理和研究模糊现象中的各种客观规律，从而获得有关模糊现象的定量描述。

在自然界和人类生活中，这种“亦此亦彼”，处于差异的中介过渡状态的模糊现象，广泛而大量地存在着。例如，骑自行车在川流不息的车辆和人群中行驶，就是一种典型的模糊现象。因为在你骑自行车行驶时，并不需要知道周围的车辆和人群的精确数量，及其每一时刻的具体位置与速度，而只要对其有个大概的了解和判断就行了。所谓“大概的了解和判断”，就是一种模糊的概念，深究起来，确是一种耐人寻味的事情，因为你就凭借着这种模糊的概念，在“模糊”中行驶，却居然可以安全、正确地到达目的地。由此可见，模糊现象并不“模糊”，模糊数学更不是“模模糊糊的数学”。

1965 年美国控制论专家 L. A. Zadeh 总结了模糊现象的规律，以通讯、模式识别等实际应用问题为背景，提出了 Fuzzy set 的概念，创立了以模糊现象为研究对象的模糊数学^[1]。在过去十九年中，它在理论研究方面和应用范围上，都获得了迅速的进展。尤其是一向被认为最难数量化的学科，如生物科学、经济与管理科学以及人文科学等，如今已成为模糊数学施展力量的重要领域。

在生物科学中，不少现象都带有模糊性：癌症病人病期的划分、癌细胞与正常细胞的识别、土壤栽培条件适宜性的综合评判等，都没有明确的界限或阈值可

以作为判定的依据。换言之，它们都是模糊现象。因此，用模糊数学去研究和处理这些现象，是既有此需要，又十分适宜的。

(二) 模糊子集及其隶属函数

现代数学的理论基础是德国数学家 G. Cantor 于 1880 年建立的集合论；而模糊数学却是以 L. A. Zadeh 提出的模糊集合论为基础的。

若用 A, B 等表示的集合是模糊的，则论域 U 上的模糊子集 A ，由其隶属函数 $\mu_A(x)$ 来表征，并得定义：

$$A \triangleq \{(x, \mu_A(x)) | x \in U\} \quad (1)$$

其中 $\mu_A(x) \in [0, 1]$ ，其值反映 U 的元素 x 对于模糊子集 A 的隶属程度，简称隶属度。

因为论域 U 的元素往往是分明的，而 U 的子集 A, B, \dots 才是模糊的，所以模糊集合一般是指模糊子集。

模糊集合的运算如下：

设 A, B 是 U 上的两个模糊子集，则它们的并集 $A \cup B$ ，交集 $A \cap B$ ，补集 \bar{A} 仍然是模糊子集；当 $x \in U$ 时，它们的隶属函数分别定义如下：

$$\begin{cases} \mu_{A \cup B}(x) = \mu_A(x) \vee \mu_B(x) = \max(\mu_A(x), \mu_B(x)); \\ \mu_{A \cap B}(x) = \mu_A(x) \wedge \mu_B(x) = \min(\mu_A(x), \mu_B(x)); \\ \mu_{\bar{A}}(x) = 1 - \mu_A(x). \end{cases} \quad (2)$$

设 A_1, A_2, \dots, A_n 是 U 中的 n 个模糊子集， x_0 是 U 中的一个元素，则隶属原则是

$$\text{若 } \mu_{A_k}(x_0) = \max(A_i(x_0)) \quad (3)$$

则就认为 x_0 相对隶属于 A_k 。这里 $A(x)$ 就是 $\mu_A(x)$ 。此法则称为隶属原则。

在 Fuzzy 综合评判中，当我们求得综合评判向量后，究竟被评判的对象属于哪个评判等级，就要用到隶属原则或入截集的概念^[3, 4]。

二、Fuzzy 综合评判与 CQ 法

(一) 模糊综合评判概述

所谓 Fuzzy 综合评判，就是某事物的着眼点的权数分配 A 对于评判矩阵 B 的模糊变换：

$$B = A \circ R \quad (4)$$

B 称为综合评判向量。

若设 $A = (a_1, a_2, \dots, a_n)$

$$E = (r_{ij})_{l \times m} \quad (i=1,2,\dots,l; j=1,2,\dots,m)$$

其中 l 为指标数, m 为评价的总等级数,

则 $B = (b_1, b_2, \dots, b_m)$

其中 $b_j = \bigvee_{i=1}^l (a_i \wedge r_{ij})$ 。

这里“ \vee ”为“取大”, “ \wedge ”为“取小”。

为了使用一类事物在评判时可以相互比较, B 的元素须要归一化, 即令

$$\sum_{j=1}^m b_j = 1 \quad (5)$$

(4) 式也可写成

$$B_k = A \circ R_k \quad (k \text{ 为事物的编号})$$

则便有

$$B_k = (b_{1k}, b_{2k}, \dots, b_{mk}) \quad (4')$$

及

$$\sum_{j=1}^m b_{jk} = 1 \quad (5)'$$

(二) CQ 法

为能实现 Fuzzy 综合评判, 必须设法将原始数据的各指标值模糊化, 从而才能求得各样点对应的评判矩阵 R_k (k 为样点编号)。因此, 关键问题在于如何模糊化, 这是个技巧问题。作者所提出的“明晰数据分段 Fuzzy 化”的新方法, 简称 CQ 法, 便是实现原始数据模糊化的有效手段。CQ 法如下(参看图 1):

(1) A, D 两界限之外, 对应地取隶属值为 $1/A$ 或 $1/D$ 。

(2) 设 E 点落入某区间如 $[B, C]$ 之间, 则 E 只产生两个隶属值即 x/B 与 y/C ; 值 x 与 y 按下公式求取:

$$x = \frac{d - |B \text{ 之界限值} - \text{实测值}|}{d};$$

$$y = \frac{d - |C \text{ 之界限值} - \text{实测值}|}{d}$$

如文献[2]中, 样地 3 的指标 2 的实测值落入 $[A, B]$ 之间, 便有 x/A 与 x/B , 而

$$x = \frac{0.107 - |1.37 - 1.45|}{0.107} = 0.252$$

$$y = \frac{0.107 - |1.477 - 1.45|}{0.107} = 0.748$$

据此便得原始数据 Fuzzy 化表示值, 详见文献[2]。

$1/A$	x/B	y/C	$1/D$
A	B	E	C

图 1

(三) 评判矩阵 R 的求取

原始数据模糊化后, 我们如何求评判矩阵 R 呢?

设样点是一 n 维向量, 即有 n 个指标; 而欲分等级数为 m , 则 R 就是一个 $(n \times m)$ 维模糊矩阵。

这样由各指标值的模糊化表示表就可求得第 k 个

样地所对应的评判矩阵 R_k 。如由文献[2]中的表 3 就可求得样点 3 的评判矩阵为:

					评价等 级	指标
	A	B	C	D		
$R_3 =$	0	0.817	0.183	0		1
	0.252	0.748	0	0		2
	1	0	0	0		3
	1	0	0	0		4
	0	0	0.647	0.353		5

R_3 右边的虚线表格, 示该矩阵的构成法则。亦可用此法求得其他样点的评判矩阵。

三、应用实例

(一) 土壤栽培条件适宜性的 Fuzzy 综合评判

1. 原始数据的来历与 E 的求得

近年来, 我国南方若干地区出现了一种杉木黄化病, 这种使杉木黄化枯死的病, 是由于不良的立地条件所引起的生理性障碍, 在这方面起作用的主要土壤条件。原始数据就是 1976—1977 年江西省德兴、广昌、全南等县的低山区 18 个标准地的土壤有关指标的理化分析值^[2,7]。

各样地对应的 R_k ($k=1,2,\dots,18$), 据文献[2]中的表 3 与评判矩阵的构成法则, 可一一求得, 上文已有交待。

2. Fuzzy 综合评判的效果

(1) 综合评判向量 B_k 的求取与评判效果

根据经验, 我们首先试取着眼点的权数分配为

$$A = (0.002, 0.009, 0.04, 0.939, 0.01) \quad (7)$$

那么, 每一样地的综合评判向量 B_k 即可据公式(4)求得。例如, 第 3 块样地的综合评判向量

$$R_3 = A \circ R_3 = (0.970, 0.009, 0.010, 0.011)$$

(其元素已归一化)

再根据隶属原则, 将样地归入隶属度较大的相应评价等级中。这里由于 R_{3k} 最大, 故第 3 块样地的评判等级是“很合适”。

如此便获得类别评判符合率为 94.4% 的良好效果。

(2) 指标进一步筛选后的综合评判效果

由(7)式可知, 指标 1, 2, 5 的权重甚小, 可以忽略, 故可改用双指标, 重新建立着眼点的权数分配为

$$A' = (0, 0, 0.041, 0.959, 0) \quad (8)$$

那么, 每块样地的综合评判向量便为

$$B'_k = A' \circ R_k \quad (k \text{ 为对应样地编号})$$

其计算结果, 类别评判符合率仍保持 94.4% 的良好效果。

再由(8)式中将指标 3 删去, 改为单指标, 则重建

表1 判别分析与综合评判结果比较表

组别(k)	实际类别	判别分析结果		综合评判结果		
		五个指标	双指标	五个指标	双指标	单指标
1	a	+	+	+	+	+
2	a	+	+	+	+	+
3	a	+	+	+	+	+
4	a	+	+	+	+	+
5	a	+	-	+	+	+
6	a	+	+	+	+	+
7	b	+	+	-	-	-
8	b	+	+	+	+	+
9	b	+	+	+	+	+
10	b	+	+	+	+	+
11	b	+	+	+	+	+
12	b	+	+	+	+	+
13	b	+	+	+	+	+
14	b	+	+	+	+	+
15	b	+	+	+	+	+
16	b	+	+	+	+	+
17	b	+	+	+	+	+
18	b	+	+	+	+	+

注：“+”表示计算结果与实际相符；“-”表示不符

的着眼点权数分配为

$$\mathcal{A}'' = (0, 0, 0, 1, 0)$$

那么,每一样地的综合评判向量便为

$$\mathcal{B}'_k = \mathcal{A}'' \circ \mathcal{R}_k, (k \text{ 为对应样地编号})$$

其计算结果,类别评判符合率与前两种处理一样,仍有 94.4% 的良好效果。详见表 1。

(3) 综合评判与判别分析的结果之比较

若以 A, B, C, D, E 代表指标 1, 2, 3, 4, 5 的实测值,那么,全指标时,用于分辨杉林土壤障碍性条件的判别函数,据文献[7]研究为:

$$\begin{aligned} R = & -2.477A - 25.656B + 5.140C \\ & + 14.324D - 0.226E \end{aligned} \quad (10)$$

分界值 $R_c = -32.88$; 大于 R_c , 判为正常土壤; 小于 R_c , 判为障碍性土壤。

原始数据,经作者自编的具有数据自动筛选功能的判别分析程序计算后,获得双指标时的对应判别函数式为:

$$R' = 2.823C + 13.420D \quad (11)$$

分界值 $R'_c = 11.38$; 大于 R'_c , 判为正常土壤; 小于 R'_c , 判为障碍性土壤。

式(10)、(11)的判别分析结果以及上述综合评判的结果,同列于表 1。可见,在全指标(五个)时,类别划分的符合率,判别分析为 100%,综合评判为 94.4%;在双指标时,两者均为 94.4%;单指标(即第 4 个指标: 土壤上层渗水深度是最关键的影响因素)

时,综合评判仍保持 94.4% 的良好效果,而此时判别分析却无能为力了。

(二) 胃溃疡良、恶性 X 线计量诊断的 Fuzzy 综合评判

医学计量诊断是模拟有经验的医师进行诊断与鉴别的定量分析方法,近二十年来已在医学领域中得到愈来愈广泛的应用。而将模糊数学应用于医学计量诊断,则是其中的新课题。

1. 原始数据的来历与 \mathcal{R} 的求取

本文目的是介绍方法的应用。因此,仅从南京铁道医学院搜集的、经病理切片证实的,并有完整 X 线检查资料的 116 个胃溃疡病例中,随机抽取 14 例来作为本文的例题。原始数据经 CQ 法模糊化后,即得表 2。

由表 2 即可求得各病例对应的评判矩阵 \mathcal{R}_k 。例如病号 75 之 \mathcal{R}_{75}^T 为

$$\mathcal{R}_{75}^T = \begin{bmatrix} 0.92 & 0 & 0.33 & 0 & 0 & 0 & 0 \\ 0.08 & 1 & 0.67 & 1 & 1 & 1 & 1 \end{bmatrix} \quad (T \text{ 表示转置})$$

2. Fuzzy 综合评判向量 \mathcal{B}_k 的求得与评判效果

根据经验与数学上的分析,分别取

$$\mathcal{A} = (0.03, 0.1, 0.25, 0.4, 0.02, 0.1, 0.1),$$

$$\text{与 } \mathcal{A}' = (0.11, 0.26, 0.42, 0.11, 0.10)$$

作七个指标与五个指标 ($x_5, x_6, x_7, x_{10}, x_{11}$) 的评判;类似上例,可求得诸病例的对应评判向量和评判类别(表 3)。

表2 各指标值的模糊化表示表

指标 病例号 \	x_3	x_5	x_6	x_7	x_9	x_{10}	x_{12}
5	$\frac{1}{A}$	$\frac{1}{A}$	$\frac{1}{A}$	$\frac{0.90}{A} + \frac{0.10}{B}$	$\frac{1}{A}$	$\frac{1}{A}$	$\frac{1}{A}$
9	$\frac{0.92}{A} + \frac{0.08}{B}$	$\frac{1}{A}$	$\frac{1}{A}$	$\frac{0.76}{A} + \frac{0.24}{B}$	$\frac{1}{A}$	$\frac{1}{A}$	$\frac{1}{A}$
20	$\frac{0.92}{A} + \frac{0.08}{B}$	$\frac{1}{A}$	$\frac{1}{A}$	$\frac{0.76}{A} + \frac{0.24}{B}$	$\frac{1}{A}$	$\frac{1}{A}$	$\frac{1}{A}$
26	$\frac{1}{A}$	$\frac{1}{A}$	$\frac{1}{A}$	$\frac{0.76}{A} + \frac{0.24}{B}$	$\frac{1}{A}$	$\frac{1}{A}$	$\frac{1}{A}$
27	$\frac{0.92}{A} + \frac{0.08}{B}$	$\frac{0.2}{A} + \frac{0.8}{B}$	$\frac{0.33}{A} + \frac{0.67}{B}$	$\frac{1}{A}$	$\frac{1}{B}$	$\frac{1}{B}$	$\frac{1}{A}$
36	$\frac{1}{A}$	$\frac{1}{A}$	$\frac{1}{A}$	$\frac{0.76}{A} + \frac{0.24}{B}$	$\frac{1}{A}$	$\frac{1}{A}$	$\frac{1}{A}$
46	$\frac{1}{A}$	$\frac{1}{A}$	$\frac{1}{A}$	$\frac{0.90}{A} + \frac{0.10}{B}$	$\frac{1}{A}$	$\frac{1}{A}$	$\frac{1}{A}$
66	$\frac{1}{B}$	$\frac{1}{B}$	$\frac{0.33}{A} + \frac{0.67}{B}$	$\frac{1}{B}$	$\frac{1}{B}$	$\frac{1}{B}$	$\frac{1}{B}$
75	$\frac{0.92}{A} + \frac{0.08}{B}$	$\frac{1}{B}$	$\frac{0.33}{A} + \frac{0.67}{B}$	$\frac{1}{B}$	$\frac{1}{B}$	$\frac{1}{B}$	$\frac{1}{B}$
82	$\frac{1}{B}$	$\frac{1}{B}$	$\frac{1}{B}$	$\frac{1}{B}$	$\frac{1}{B}$	$\frac{1}{B}$	$\frac{1}{B}$
87	$\frac{1}{B}$	$\frac{1}{B}$	$\frac{0.33}{A} + \frac{0.67}{B}$	$\frac{1}{B}$	$\frac{1}{A}$	$\frac{1}{B}$	$\frac{1}{B}$
88	$\frac{1}{B}$	$\frac{1}{B}$	$\frac{0.33}{A} + \frac{0.67}{B}$	$\frac{1}{B}$	$\frac{1}{A}$	$\frac{1}{B}$	$\frac{1}{B}$
104	$\frac{0.92}{A} + \frac{0.08}{B}$	$\frac{1}{A}$	$\frac{1}{B}$	$\frac{0.38}{A} + \frac{0.62}{B}$	$\frac{1}{B}$	$\frac{1}{A}$	$\frac{1}{B}$
116	$\frac{1}{B}$	$\frac{1}{B}$	$\frac{0.33}{A} + \frac{0.67}{B}$	$\frac{1}{B}$	$\frac{1}{B}$	$\frac{1}{B}$	$\frac{1}{B}$

注: x_3 —部位; x_5 —溃疡与胃壁关系; x_6 —龛影轮廓; x_7 —围绕壁龛圆周的透亮带; x_9 —皱壁异常;
 x_{10} —皱壁消失; x_{12} —胃壁僵硬或拐角征^[5,6]。 A全“良性”; B全“恶性”。

表3 七指标与五指标的综合评判效果表

病例号(k)	实际类型	七指 标			五指 标	
		综合评判向量(b_{1k}, b_{2k})	评判效果	疑诊(*)	评判效果	疑诊(*)
3	1. (良性)	0.80, 0.20	+		+	
9		0.63, 0.37	+		+	
20		0.63, 0.37	+		+	
26		0.63, 0.37	+		+	
27		0.62, 0.38	+		+	
36		0.63, 0.37	+		+	
46		0.80, 0.20	+		+	
66	2. (恶性)	0.38, 0.62	+		+	
75		0.38, 0.62	+		+	
82		0, 1	+		+	
87		0.38, 0.62	+		+	
88		0.38, 0.62	+		+	
104		0.49, 0.51	+	*	+	
116		0.38, 0.62	+		+	*

3. 疑诊与确诊的截取

“人命非儿戏”，因此在医学计量诊断时，还应分清疑诊与确诊。对于疑诊病例，必须进行一步检查，以便确诊治疗。

作为 Fuzzy 变换的综合评判，实质上是一种论域的变换。它把概念 α ，在论域 U 中以模糊向量 a 表示的方式转换成论域 V 中以模糊向量 $b = \alpha \circ R$ 来描述。亦即^[3]

$$\{\alpha\} \sim \xrightarrow{a \circ R} V \quad (12)$$

因此，表 3 中 B_k 的分量 b_{1k} , b_{2k} 分别表示病号 k 与类型 1, 2 的隶属度。对模糊集 B_k ，作 $\lambda = 0.56$ 的截集 $(B_k)_k$ 。规定：凡属于 $(B_k)_k$ 的病号，被认为是确诊；否则就划为疑诊，此时 $(B_k)_k = \emptyset$ （空集）。如病号 104，虽然由于 $b_{1k} = 0.51 > 0.49 = b_{2k}$ ，被判为恶性；但因 $b_{1k} = 0.51 < \lambda = 0.58$ ，故 $(B_k)_k = \emptyset$ ，据此该病号便划为疑诊，需要继续诊断。由表 3 可见，本例评判符合率为 100%。而确诊率也高达 92.86%，效果良好。

四、讨 论

(一) 本文简要地介绍了 Fuzzy 综合评判的数学方法，结合实例使其具体化，力求使读者得知模糊数学能解决一些其它方法不易解决的问题，并大略知道解决的步骤。

(二) 要能具体运用 Fuzzy 综合评判，关键有二：

(上接第 29 页)

- [34] Martin, R. P. and J. Scharff; *An Introduction to Bioinorganic Chemistry*, (ed. D. R. Williams), Charles C Thomas Pub., Springfield, USA, 120, 1975.
- [35] Prince, R. H.: *Adv. in Inorg. Chem. and Radionuclides*, (ed. H. J. Emeleus), Acad. Press, New York, 22, 349, 1979.
- [36] Ibers, I. A. and R. H. Holm; *Science*, 209, No. 4453, 223, 1980.
- [37] Martell, A. E.: in *Metal Ions in Biological Systems*, 2, 207, 1973.
- [38] Dwyer, F. P.: in *Chelating Agents and Metal Chelates* (ed. F. P. Dwyer), Acad. Press, New York, 335, 1964.
- [39] Hay R. W.: in *An Introduction to Bioinorganic Chemistry* (ed. D. R. Williams), 51, 1975.
- [40] O'Young, C.L. and S. Lippard et al.; *J. Am. Chem. Soc.*, 100, 7291, 1978.
- [41] MacCilivray, R. T. A. and E. Mendez; *Proc. Natl. Acad. Sci. USA*, 79, 2504 (1982); P. Aisen and I. Listovsky, *Ann. Rev. Biochem.*, 49, 357, 1980.
- [42] Perutz, M. F.: *Nature*, 228, 726 1970.

一是使原始数据模糊化，从而求得评判矩阵 R ，CO 法正是应这一需要而提出的有力工具；二是要确定合适的权数分配（因素权重） A ，但这属于综合评判的逆命题与 Fuzzy 关系方程研究的内容，故不在本文中讨论。

(三) 众所周知，不同的土壤栽培作物或林木，配合适当，能速生丰产，否则适得其反。在进行土地利用规划时，这一点尤为重要。过去，生物学工作者习惯于凭经验判断，结果往往因人而异，而且当影响因子增多时，很难作出客观的判断。随着生物数学的发展，一些生物学家开始用判别分析的方法来定量研究这类问题，但这种方法计算工作量大，不易为一般生物医学和农林工作者掌握。如用 Fuzzy 综合评判来处理，工作量小，效果颇好，因此值得生物医学工作者去研究和运用。

参 考 文 献

- [1] Zadeh, L. A.: *Information and Control*, 8, 338—357, 1965.
- [2] 冯晋臣、季静秋：《模糊数学》，2, 55—64, 1984。
- [3] 汪培庄：《模糊数学讲义》(上)，第 1、2 章，北京师范大学，1981。
- [4] 张文修：《模糊数学基础》，第 1、3 章，西安交通大学，1982。
- [5] 冯晋臣、季静秋：《胃溃疡良、恶性的 X 线计量诊断的 Fuzzy 模式识别之研究》(待发表)，1982。
- [6] 伍福庆、陈启光：《中华放射学杂志》，2, 130, 1981。
- [7] 罗汝英：《土壤学报》，2, 145—155, 1982。
- [43] Moore, G. R. Z., X. Huang (黄仲贤) and R. J. P. Williams et al.: *Faraday Discussion of Chem. Soc.*, 74, 311, 1982.
- [44] Colburn, R. W. and J. W. Maas; *Nature*, 208, 37, 41, 1965.
- [45] Rajan, K. S. et al., *J. Neurochem.*, 19, 1117, 1972.
- [46] Pressman, B. C.: in *Proc. Fifth FEBS Meeting: Mitochondria-Structure and Function*, eds. L. Ernster and Z. Drahota, Acad. Press, Prague 1969.
- [47] 黄仲贤，《自然杂志》，6, 26, 1983。
- [48] Albert, A.: *Selective Toxicity*, Chapman and Hall, London, 1979.
- [49] Perrin, D. D.: *Topic Curr. Chem.*, Springer-Verlag, Berlin, 64, 1979.
- [50] Chisholm, J.: *J. Pediat.*, 73, 1, 1968.
- [51] Eichhorn, G. L.: *Adv. Chem. Ser.*, 37, 37, 1963.
- [52] Wallace, A.: *A Decade of Synthetic Chelating Agents on Inorganic Plant Nutrition*, in A. Wallace (ed.): Los Angeles, 1962.
- [53] Neiland, J. B.: *Essay in Coord. Chem.*, (eds. W. Schneider et al.) Birk Hauser, Bassel, 1964.

[本文于 1984 年 3 月 25 日收到]