

改进 IVIF-BWM 的新高考数学试题 综合难度模型的构建与应用

陈艳玲¹, 谢加良^{1,2}, 刘小辉^{1,3}

(1. 集美大学理学院, 福建 厦门 361021; 2. 数字福建大数据建模与智能计算研究所, 福建 厦门 361021;
3. 厦门大学教育研究院, 福建 厦门 361005)

[摘要] 以《中国高考评价体系》为蓝本, 根据新高考数学试题的特点, 构建新高考数学试题难度指标模型。针对现有区间值直觉模糊最优最劣方法(interval value intuitionistic fuzzy best worst method, IVIF-BWM)的不足, 提出了改进的 IVIF-BWM。该方法建立了基于乘法一致的区间值直觉模糊偏好关系(interval value intuitionistic fuzzy preference relation, IVIFPR)的目标规划模型求解准则的最优区间值直觉模糊(interval value intuitionistic fuzzy, IVIF)权重, 并推导出 IVIF 环境下的一致性指标(consistency index, CI)和一致性比率(consistency ratio, CR)来验证结果的可靠性, 将改进的 IVIF-BWM 用于计算难度指标和难度水平的权重, 建立新高考试题综合难度模型。最后利用该模型评价 2022 年新高考数学试卷, 验证所构建的新高考数学试题综合难度模型的有效性。

[关键词] 区间值直觉模糊集; 最优最劣方法; 新高考; 数学试题; 综合难度模型

[中图分类号] O 159; G 633.6

Construction and Application of Comprehensive Difficulty Model for New College Entrance Examination Mathematics Questions Based on Improved IVIF-BWM

CHEN Yanling¹, XIE Jialiang^{1,2}, LIU Xiaohui^{1,3}

(1. School of Science, Jimei University, Xiamen 361021, China;

2. Institute of Big Data Modeling and Intelligent Computing, Digital Fujian, Xiamen 361021, China;

3. Institute of Education, Xiamen University, Xiamen 361005, China)

Abstract: Firstly, using the *China College Entrance Examination Evaluation System* as the method guide and according to the examination characteristics of the new college entrance examination mathematics questions, this paper constructed a comprehensive difficulty index model of mathematics in the new college entrance examination. Then, according to the shortcomings of the existing interval value intuitionistic fuzzy best worst method(IVIF-BWM), this paper developed an improved IVIF-BWM that establishes a goal programming model

[收稿日期] 2023-02-26

[基金项目] 教育部人文社会科学研究规划基金项目“应用型本科院校应用型理工科专业评价标准国际比较研究”(17YJA880070); 福建省高等教育研究院重点课题“数字化转型下福建省高校人才培养现状及提升策略”(FGJ202403); 福建省科技厅创新战略研究项目“‘福元宇宙’视域下公民数字素养影响因素与提升策略研究”(2023R0051)

[作者简介] 通信作者: 刘小辉(1972—), 教授, 博士, 硕导, 从事教育统计与数学教学研究。E-mail: 51292715@qq.com

based on the multiplicative consistent interval value intuitionistic fuzzy preference relation(IVIFPR) to derive the optimal interval value intuitionistic fuzzy(IVIF) weight of criteria, and then proposed the consistency index (CI) and consistency ratio(CR) under IVIF environment to verify the reliability of the derived results. The improved IVIF-BWM was used to calculate the weights of the difficulty indexes and the difficulty levels, so as to establish a comprehensive difficulty model for the new high exam questions. Finally, this model is used to analyze difficulty characteristics of the new college entrance exam mathematics paper in 2022, which verified the effectiveness of comprehensive difficulty model for new college entrance examination mathematics questions.

Keywords: interval value intuitionistic fuzzy; best worst method; new college entrance examination; mathematics questions; comprehensive difficulty model

0 引言

2014 年,我国启动新一轮高考改革,社会各界对高考的关注度持续上升,新高考试题的难度尤其引起社会的高度关注,因此,构建新高考背景下的试题难度评价模型尤为重要。综合难度模型^[1]作为研究试题难度的重要工具,近年来被广泛用于研究中高考数学试题的难度,并且许多学者针对不同应用场景调整了该模型。如张怡等^[2]在该模型的基础上加入思维方向指标,并将其用于高考理科数学全国卷试题;武小鹏等^[3]在五维度综合难度模型的基础上增加了思维方向、是否含参 2 个指标,形成高考试题难度的评价模型;薛欢等^[4]根据“统计与概率”试题的特点,增加了条件含量、阅读量 2 个指标,对试题进行难度分析;李保臻等^[5]结合文献 [1] 的综合难度模型和文献 [3] 的改进模型,重新划分各个难度水平,使其更适合高考试题的特点。

通过分析可以发现,上述这些模型没有考虑不同评价指标的权重,并且直接以自然数 1,2,3,⋯ 对各指标的难度水平进行划分,每个难度水平的权重呈等量递增的趋势,这显然是不合理的。基于此,武小鹏等^[6]采用层次分析法(analytic hierarchy process, AHP)计算各指标及各个难度水平的权重,完善了综合难度系数模型,形成高考数学试题难度评价模型。然而, AHP 作为经典的权重确定方法,存在两两比较矩阵中会出现不一致和指标间两两比较的工作量巨大 2 个缺点^[7]。为了克服 AHP 的不足,Rezaei^[8]提出最优最劣方法(best worst method, BWM), BWM 在减少成对比较次数和保持判断一致性方面表现良好。由于主观准则中数据的不确定性,许多学者将 BWM 拓展到不同的模糊环境,使用不同模糊集表征指标间的偏好关系,如模糊 BWM^[9]、区间值模糊 BWM^[10]、直觉模糊 BWM^[11]、区间值模糊粗糙 BWM^[12]、犹豫的模糊语言 BWM^[13]等。相较于语言变量、模糊数、区间数、直觉模糊数等,区间值直觉模糊数同时考虑了隶属度和非隶属度,并且可以更好地刻画决策者在决策过程中的不确定性,减少因信息表示模型选择不妥而导致的信息丢失或失真,进而降低决策风险,提高决策质量^[14]。因此,文献 [15-16] 将 IVIF 和经典 BWM 相结合,提出了 IVIF-BWM。然而,现有的这些 IVIF-BWM 的扩展还存在如下不足:1) 部分参考比较的 IVIF 表达不符合逻辑;2) IVIF-BWM 忽略了 IVIFPR 与乘法偏好关系之间的差异,直接将 BWM 扩展到区间直觉模糊环境;3) 在一致性比较过程中, IVIF-BWM 的扩展仍然使用了 BWM 的 CI 和 CR,并且在区间值直觉模糊环境下无法建立新的一致性指标。

基于以上分析,本文以《中国高考评价体系》为蓝本,根据新高考数学试题的考查特点,构建新高考数学试题难度指标模型,并针对现有 IVIF-BWM 的不足,提出改进的 IVIF-BWM。该方法建立了基于乘法一致 IVIFPR 的目标规划模型来求解标准化的最优 IVIF 权重,并推导出 IVIF 环境下的 CI 和 CR 来验证推导结果的可靠性;将改进的 IVIF-BWM 用于计算难度指标和难度水平的权重,建立新高考试题综合难度模型;利用该模型评价 2022 年新高考数学试卷,验证所构建的新高考数学试题综合难度模型的有效性。

1 IVIFS 和 IVIFPR 的相关理论

定义 1^[17] X 上的区间值直觉模糊集 (IVIFS) \tilde{A} 定义为 $\tilde{A} = \{ \langle x, \tilde{\mu}_{\tilde{A}}(x), \tilde{v}_{\tilde{A}}(x) \rangle | x \in X \}$, 其中 $x \in X$, 有 $\tilde{\mu}_{\tilde{A}}: X \rightarrow L([0, 1])$ 、 $\tilde{v}_{\tilde{A}}: X \rightarrow L([0, 1])$ 且 $0 \leq \sup(\tilde{\mu}_{\tilde{A}}(x)) + \sup(\tilde{v}_{\tilde{A}}(x)) \leq 1$ 。区间数 $\tilde{\mu}_{\tilde{A}}(x)$ 和 $\tilde{v}_{\tilde{A}}(x)$ 分别表示元素 x 属于区间值直觉模糊集 \tilde{A} 的隶属度和非隶属度, 记 $\tilde{\mu}_{\tilde{A}}(x) = [\mu_{\tilde{A}}^-(x), \mu_{\tilde{A}}^+(x)]$ 和 $\tilde{v}_{\tilde{A}}(x) = [v_{\tilde{A}}^-(x), v_{\tilde{A}}^+(x)]$ 。因此 \tilde{A} 可以等价表示为 $\tilde{A} = \{ \langle x, [\mu_{\tilde{A}}^-(x), \mu_{\tilde{A}}^+(x)], [v_{\tilde{A}}^-(x), v_{\tilde{A}}^+(x)] \rangle | x \in X \}$, 其中, 对于任意 $x \in X$, 满足 $0 \leq \mu_{\tilde{A}}^-(x) \leq \mu_{\tilde{A}}^+(x) \leq 1$, $0 \leq v_{\tilde{A}}^-(x) \leq v_{\tilde{A}}^+(x) \leq 1$, $\mu_{\tilde{A}}^+(x) + v_{\tilde{A}}^+(x) \leq 1$ 。同理, $\tilde{\pi}_{\tilde{A}}(x) = [\pi_{\tilde{A}}^-(x), \pi_{\tilde{A}}^+(x)]$ 表示元素 $x \in X$ 的区间直觉犹豫指标, 其中: $\pi_{\tilde{A}}^-(x) = 1 - \mu_{\tilde{A}}^+(x) - v_{\tilde{A}}^+(x)$; $\pi_{\tilde{A}}^+(x) = 1 - \mu_{\tilde{A}}^-(x) - v_{\tilde{A}}^-(x)$ 。为了方便, 称 $\tilde{\alpha} = ([\mu^-, \mu^+], [v^-, v^+])$ 为区间值直觉模糊数 (IVIFN)。

定义 2^[18] X 上的 IVIFPR 可以表示为一个矩阵 $\tilde{R} = (\tilde{r}_{ij})_{m \times m}$, 其中 $\tilde{r}_{ij} = (\tilde{\mu}_{ij}, \tilde{v}_{ij}) (i, j = 1, 2, \dots, m)$ 为区间直觉模糊值, $\tilde{\mu}_{ij}$ 表示决策者对方案 x_i 和 x_j 进行两两比较时偏爱 x_i 的程度, \tilde{v}_{ij} 表示决策者偏爱 x_i 的程度, $\tilde{\pi}_{ij} = 1 - \tilde{\mu}_{ij} - \tilde{v}_{ij}$ 为决策者偏爱 x_i 的犹豫程度。此外, $\tilde{\mu}_{ij}$ 、 \tilde{v}_{ij} 满足: $[\mu_{ij}^-, \mu_{ij}^+] \subseteq [0, 1]$, $[v_{ij}^-, v_{ij}^+] \subseteq [0, 1]$, $\tilde{\mu}_{ij} = \tilde{v}_{ji}$, $\tilde{v}_{ij} = \tilde{\mu}_{ji}$, $\tilde{\mu}_{ii} = \tilde{v}_{ii} = [0.5, 0.5]$, $\mu_{ij}^+ v_{ij}^+ \leq 1$ 。特别地, 当 $\mu_{ij}^- = \mu_{ij}^+ = \tilde{\mu}_{ij}$ 、 $v_{ij}^- = v_{ij}^+ = \tilde{v}_{ij}$ 时, IVIFPR $\tilde{R} = (\tilde{r}_{ij})_{m \times m}$ 退化为 IFPR $\tilde{R} = (\tilde{\mu}_{ij}, \tilde{v}_{ij})_{m \times m}$ 。

在实际决策中, 决策者几乎不可能给出 $([0, 0], [0, 0])$ 、 $([1, 1], [0, 0])$ 和 $([0, 0], [1, 1])$ 这样的偏好值, 因此本文只讨论满足 IVIFPR $\tilde{R} = (\tilde{r}_{ij})_{m \times m}: [\mu_{ij}^-, \mu_{ij}^+] \subseteq (0, 1)$ 且 $[v_{ij}^-, v_{ij}^+] \subseteq (0, 1)$ 。

定义 3^[19] 直觉模糊权重向量 $\tilde{\omega} = (\tilde{\omega}_1, \tilde{\omega}_2, \dots, \tilde{\omega}_m)$ 称为标准化直觉模糊权重向量, 满足

$$\sum_{j=1, j \neq i}^n \omega_j^\mu \leq \omega_i^\mu, \omega_i^\mu + m - 2 \geq \sum_{j=1, j \neq i}^n \omega_j^v, i = 1, 2, \dots, m. \quad (1)$$

其中: $\tilde{\omega} = (\tilde{\omega}_i^\mu, \tilde{\omega}_i^v)$; $\tilde{\omega}_i^\mu$ 、 $\tilde{\omega}_i^v \in [0, 1]$ 。

定义 4^[18] 已知 IVIFPR $\tilde{R} = (\tilde{\mu}_{ij}, \tilde{v}_{ij})_{m \times m}$, $\tilde{\mu}_{ij} = [\mu_{ij}^-, \mu_{ij}^+]$, $\tilde{v}_{ij} = [v_{ij}^-, v_{ij}^+]$, 称 $\tilde{R}^L = (\mu_{ij}^L, v_{ij}^L)_{m \times m}$ 为 \tilde{R} 左矩阵, $\tilde{R}^R = (\mu_{ij}^R, v_{ij}^R)_{m \times m}$ 为 \tilde{R} 右矩阵, 其中:

$$\left\{ \begin{aligned} (\mu_{ij}^L, v_{ij}^L) &= \begin{cases} (\mu_{ij}^-, v_{ij}^+), i < j, \\ (0.5, 0.5), i = j, \\ (\mu_{ij}^+, v_{ij}^-), i > j, \end{cases} \\ (\mu_{ij}^R, v_{ij}^R) &= \begin{cases} (\mu_{ij}^+, v_{ij}^-), i < j, \\ (0.5, 0.5), i = j, \\ (\mu_{ij}^-, v_{ij}^+), i > j. \end{cases} \end{aligned} \right. \quad (2)$$

定义 5^[18] 已知 IVIFPR $\tilde{R} = (\tilde{\mu}_{ij}, \tilde{v}_{ij})_{m \times m}$, 当 \tilde{R}^L 、 \tilde{R}^R 均乘法一致时, 有

$$\mu_{ij}^- \mu_{jk}^- \mu_{ki}^- = v_{ij}^+ v_{jk}^+ v_{ki}^+, \mu_{ij}^+ \mu_{jk}^+ \mu_{ki}^+ = v_{ij}^- v_{jk}^- v_{ki}^-, i, j, k = 1, 2, \dots, m. \quad (3)$$

若式 (3) 成立, 则 \tilde{R} 乘法一致。

令 $\rho_{ij}^L = \mu_{ij}^- / v_{ij}^+$, $\rho_{ij}^R = \mu_{ij}^+ / v_{ij}^-$, 则式 (3) 转化为

$$\rho_{ij}^L = \rho_{ik}^L \rho_{ki}^L, \quad \rho_{ij}^R = \rho_{ik}^R \rho_{ki}^R. \quad (4)$$

定理 1^[18] 对于 IVIFPR $\tilde{R} = (\tilde{\mu}_{ij}, \tilde{v}_{ij})_{m \times m}$ 的左矩阵 $\tilde{R}^L = (\mu_{ij}^L, v_{ij}^L)_{m \times m}$ 和右矩阵 $\tilde{R}^R = (\mu_{ij}^R, v_{ij}^R)_{m \times m}$, 如果存在标准化直觉模糊权重向量 $\omega^- = (\omega_1^-, \omega_2^-, \dots, \omega_m^-)$ 和 $\omega^+ = (\omega_1^+, \omega_2^+, \dots, \omega_m^+)$, 其中 $\omega^- = (\omega_i^{\mu^-}, \omega_i^{v^-})$, $\omega^+ = (\omega_i^{\mu^+}, \omega_i^{v^+})$, 使得

$$(\mu_{ij}^L, v_{ij}^L) = \begin{cases} (0.5, 0.5), i = j, \\ (2\omega_i^{\mu^-} / (\omega_i^{\mu^-} - \omega_i^{v^-} + \omega_i^{\mu^-} - \omega_i^{v^-} + 2), 2\omega_j^{v^-} / (\omega_i^{\mu^-} - \omega_i^{v^-} + \omega_j^{\mu^-} - \omega_j^{v^-} + 2)), i \neq j. \end{cases} \quad (5)$$

$$(\mu_{ij}^R, v_{ij}^R) = \begin{cases} (0.5, 0.5), i = j, \\ (2\omega_i^{\mu^+}/(\omega_i^{\mu^+} - \omega_i^{v^-} + \omega_i^{\mu^+} - \omega_i^{v^+} + 2), 2\omega_j^{\mu^+}/(\omega_i^{\mu^+} - \omega_i^{v^-} + \omega_j^{\mu^+} - \omega_j^{v^+} + 2)), i \neq j \end{cases} \quad (6)$$

成立, 则 \tilde{R}^L 、 \tilde{R}^R 均为乘法一致的 IFPR, 那么 $\tilde{R} = (\tilde{\mu}_{ij}, \tilde{v}_{ij})_{m \times m}$ 乘法一致。因此, IVIFPR \tilde{R} 乘法一致的 IVIF 权向量为 $\tilde{\omega} = (\tilde{\omega}_1, \tilde{\omega}_2, \dots, \tilde{\omega}_m)$, 其中: $\tilde{\omega}_i = ([\underline{\omega}_i^{\mu}, \overline{\omega}_i^{\mu}], [\underline{\omega}_i^v, \overline{\omega}_i^v])$; $\underline{\omega}_i^{\mu} = \min\{\omega_i^{\mu^-}, \omega_i^{\mu^+}\}$; $\overline{\omega}_i^{\mu} = \max\{\omega_i^{\mu^-}, \omega_i^{\mu^+}\}$; $\underline{\omega}_i^v = \min\{\omega_i^{v^-}, \omega_i^{v^+}\}$; $\overline{\omega}_i^v = \max\{\omega_i^{v^-}, \omega_i^{v^+}\}$ 。

2 新高考数学试题综合难度模型的建立

本研究以《中国高考评价体系》为方法指导, 根据新高考数学试题的考查特点, 对武小鹏等^[3]的试题综合难度模型进行调整和补充, 建立含有 9 个评价指标和 31 个难度水平的新高考数学试题综合难度模型 (见表 1)。

表 1 新高考数学试题难度指标模型

Tab. 1 Comprehensive difficulty index model of mathematics question in the new college entrance examination

指标	水平	含义
试题情境 C_1	课程学习情境 C_{11}	在数学学科背景下考查学生对基础知识掌握情况的问题情境
	探索创新情境 C_{12}	对所学数学内容深化的问题情境, 考查对数学学科内容的探究和创新
	社会实践情境 C_{13}	包括生活情境、科学情境等考查数学应用性的问题情境
参数含量 C_2	无含参 C_{21}	试题中没有相关的参数变量, 只涉及静态数值计算
	含参无需讨论 C_{22}	含有未知参数运算, 但不需要对参数进行讨论, 即包括常值参数
	含参需要讨论 C_{23}	含有未知参数运算, 但需对参数进行讨论分析, 即包括参数变量
运算水平 C_3	简单运算 C_{31}	运算涉及简单数值运算或简单符号运算
	一般运算 C_{32}	解题时同时涉及简单数值运算和简单符号运算的混合运算
	较复杂运算 C_{33}	解题时只涉及复杂符号运算或复杂数值运算
	复杂运算 C_{34}	同时涉及复杂符号运算和复杂数值运算的混合
推理能力 C_4	简单推理 C_{41}	推理涉及的数学知识背景为纯公式背景, 推理步骤 3 步之内
	一般推理 C_{42}	推理涉及对数学知识背景进行简单变形, 推理步骤 4 ~ 6 步
	复杂推理 C_{43}	推理涉及的数学知识背景比较抽象, 推理步骤多于 6 步
直观想象 C_5	无直观想象 C_{51}	题设未给出图形, 并且无需借助绘制图形就可以解答试题
	简单直观想象 C_{52}	根据题设条件对已知图形进行加工, 并借助加工后的图形才能解答试题
	复杂直观想象 C_{53}	题设未给出图形, 需要根据题设条件正确做出图形并借助图形才能解答试题
知识综合 C_6	单知识点 C_{61}	根据单个知识点就能解题
	多知识点 C_{62}	根据多个独立知识点就能作答
	知识关联 C_{63}	根据多个关联的知识点才能作答
	知识拓展 C_{64}	根据多个关联的知识点, 并需要在此基础上得出新结论才能成功解题
解题思维 C_7	简单思维 C_{71}	只需根据题设已知条件和已学过的知识进行解答, 不用进行转化
	一般思维 C_{72}	借助 3 步以内的转化将复杂问题转化为已解决的简单问题
	复杂思维 C_{73}	借助 3 步以上的转化将复杂问题转化为已解决的简单问题
认知水平 C_8	理解 C_{81}	试题涉及对数学概念或者原理的理解, 能对其过程进行描述和简单说明
	应用 C_{82}	能够将数学原理和概念应用到具体的问题解决过程中
	分析 C_{83}	解题要拓展转化或创造性运用数学知识, 深入分析和综合应用题目各个条件
开放程度 C_9	封闭题型 C_{91}	试题的已知条件完备并且结论是确定的
	条件开放型 C_{92}	试题的条件不完备或者满足结论的条件不唯一
	结论开放型 C_{93}	在给定的条件下试题的结论不唯一
	策略开放型 C_{94}	试题求解的思维策略和方法不唯一
	设计实践型 C_{95}	需要运用数学知识和方法进行计划性的预测和规划

新高考数学试题难度系数模型对表 1 中的指标和水平赋值所获得的数据进行加权聚合，主要从各个指标的难度系数 d_i 、一组试题的综合难度系数 D 方面进行说明。

各个指标的难度系数 d_i 为

$$d_i = \sum_{j=1}^{n_i} (m_{ij} \omega_{ij}) / m。 \tag{7}$$

其中： $i(i = 1, 2, \cdots, 9)$ 表示表 1 中的 9 个指标； ω_{ij} 为第 i 个指标第 j 个水平的权重； $j(j = 1, 2, \cdots)$ 表示指标的不同难度水平； n_i 表示第 i 个指标的水平数； m_{ij} 表示该组试题中编码为第 i 个指标第 j 个水平的试题数； $m = \sum_j m_{ij}$ 表示此组试题的总数量。

聚合各个指标的难度系数 d_i ，得到一组试题的综合难度系数 D 为

$$D = \sum_{i=1}^9 (\omega_i d_i) = \sum_{i=1}^9 \{ \omega_i [\sum_{j=1}^{n_i} (m_{ij} \omega_{ij}) / m] \} , \sum_j m_{ij} = m, i = 1, 2, \cdots, 9。 \tag{8}$$

其中： ω_i 表示第 i 个指标的权重。

3 改进的 IVIF-BWM 及其权重确定

3.1 改进的 IVIF-BWM

传统的 BWM^[20] 可分为 3 个阶段：1) 给出参考比较；2) 推导最优权向量；3) 计算一致性指数。本文从这 3 个方面给出 IVIF-BWM 的改进措施。

1) 给出参考比较。在此阶段，传统的 BWM 采用 1~9 标度的形式进行参考比较，而现实决策中偏好的表达存在不确定性。采用 IVIFN 能更好的表达偏好信息，现有的 IVIF-BWM^[15] 在采用 IVIFN 进行准则的参考比较时，用 $([1.0, 1.0], [0.0, 0.0])$ 表示两个准则同等重要，这显然是不符合逻辑的。

根据 IVIFPR 的定义，最优准则 C_B 对准则 C_j 的偏好可以表示为 $IVIFN \tilde{a}_{Bj} = ([a_{Bj}^{\mu-}, a_{Bj}^{\mu+}], [a_{Bj}^{v-}, a_{Bj}^{v+}])$ ，满足 $a_{Bj}^{\mu-}, a_{Bj}^{\mu+} \in [0, 1]$ 、 $a_{Bj}^{v-}, a_{Bj}^{v+} \in [0, 1]$ 、 $a_{Bj}^{\mu+} + a_{Bj}^{v+} \leq 1$ 且 $\tilde{a} = ([0.5, 0.5], [0.5, 0.5])$ 。同理，准则 C_j 对最劣准则 C_W 的偏好 $\tilde{a}_{jW} = ([a_{jW}^{\mu-}, a_{jW}^{\mu+}], [a_{jW}^{v-}, a_{jW}^{v+}])$ ，满足 $a_{jW}^{\mu-}, a_{jW}^{\mu+} \in [0, 1]$ 、 $a_{jW}^{v-}, a_{jW}^{v+} \in [0, 1]$ 、 $a_{jW}^{\mu+} + a_{jW}^{v+} \leq 1$ 且 $\tilde{a}_{WW} = ([0.5, 0.5], [0.5, 0.5])$ ，从而获得最优-其他向量 BO 和其他-最劣向量 OW 分别为 $\tilde{A}_{BO} = (a_{B1}, a_{B2}, \cdots, a_{Bn})$ 、 $\tilde{A}_{OW} = (a_{1W}, a_{2W}, \cdots, a_{nW})$ 。因此，本文用 $([0.5, 0.5], [0.5, 0.5])$ 表示同等重要，从而确定了参考比较语言术语表（见表 2）。

表 2 评价指标成对比较的语言术语
Tab. 2 Language terms for pairwise comparisons of evaluation metrics

语言术语	区间值直觉模糊数
同等重要	$([0.5, 0.5], [0.5, 0.5])$
稍微重要	$([0.5, 0.6], [0.3, 0.4])$
明显重要	$([0.6, 0.7], [0.2, 0.3])$
非常重要	$([0.7, 0.8], [0.1, 0.2])$
极其重要	$([0.8, 0.9], [0.0, 0.1])$

2) 推导最优权向量。根据定理 1，对于所有的 $j = 1, 2, \cdots, m$ ，有

$$\begin{cases} \tilde{a}_{Bj}^L = (a_{Bj}^{\mu-}, a_{Bj}^{v+}) = \begin{cases} (0.5, 0.5), B = j, \\ (2\omega_B^{\mu-} / (\omega_B^{\mu-} - \omega_B^{v-} + \omega_j^{\mu-} - \omega_j^{v-} + 2), 2\omega_j^{\mu-} / (\omega_B^{\mu-} - \omega_B^{v-} + \omega_j^{\mu-} - \omega_j^{v-} + 2)), B \neq j; \end{cases} \\ \tilde{a}_{Bj}^R = (a_{Bj}^{\mu+}, a_{Bj}^{v-}) = \begin{cases} (0.5, 0.5), B = j, \\ (2\omega_i^{\mu+} / (\omega_i^{\mu+} - \omega_i^{v+} + \omega_j^{\mu+} - \omega_j^{v+} + 2), 2\omega_j^{\mu+} / (\omega_i^{\mu+} - \omega_i^{v+} + \omega_j^{\mu+} - \omega_j^{v+} + 2)), B \neq j; \end{cases} \\ \tilde{a}_{jW}^L = (a_{jW}^{\mu-}, a_{jW}^{v+}) = \begin{cases} (0.5, 0.5), j = W, \\ (2\omega_j^{\mu-} / (\omega_j^{\mu-} - \omega_j^{v-} + \omega_W^{\mu-} - \omega_W^{v-} + 2), 2\omega_W^{\mu-} / (\omega_j^{\mu-} - \omega_j^{v-} + \omega_W^{\mu-} - \omega_W^{v-} + 2)), j \neq W; \end{cases} \\ \tilde{a}_{jW}^R = (a_{jW}^{\mu+}, a_{jW}^{v-}) = \begin{cases} (0.5, 0.5), j = W, \\ (2\omega_j^{\mu+} / (\omega_j^{\mu+} - \omega_j^{v+} + \omega_W^{\mu+} - \omega_W^{v+} + 2), 2\omega_W^{\mu+} / (\omega_j^{\mu+} - \omega_j^{v+} + \omega_W^{\mu+} - \omega_W^{v+} + 2)), B \neq j. \end{cases} \end{cases} \tag{9}$$

式 (9) 成立时, BO 向量和 OW 向量完全一致。然而, $\tilde{\omega}_B/\tilde{\omega}_j = \tilde{a}_{Bj}$ 和 $\tilde{\omega}_j/\tilde{\omega}_W = \tilde{a}_{jW}$ 并不意味着 \tilde{A}_{BO} 和 \tilde{A}_{OW} 是完全一致的。因此, 传统 BWM 的规划模型不适合用于求解最优 IVIF 权向量, 文献 [15] 中对于 BWM 的扩展是不合适的, 并且该模型不能保证所求得的 IVIF 权向量是标准化的。基于此, 为了有效地从参考比较向量 BO 和 OW 向量中获得标准化的最优 IVIF 权向量, 下面建立一个线性规划模型。

对于所有的 $j = 1, 2, \dots, m$, 式 (9) 很难满足, 期望得到的 IVIF 权重尽可能满足

$$\begin{cases} a_{Bj}^{\mu^-} = 2\omega_B^{\mu^-}/(\omega_B^{\mu^-} - \omega_B^{v^-} + \omega_j^{\mu^-} - \omega_j^{v^-} + 2), a_{Bj}^{v^+} = 2\omega_j^{\mu^-}/(\omega_B^{\mu^-} - \omega_B^{v^-} + \omega_j^{\mu^-} - \omega_j^{v^-} + 2), \\ a_{Bj}^{\mu^+} = 2\omega_i^{\mu^+}/(\omega_i^{\mu^+} - \omega_i^{v^+} + \omega_j^{\mu^+} - \omega_j^{v^+} + 2), a_{Bj}^{v^-} = 2\omega_j^{\mu^+}/(\omega_i^{\mu^+} - \omega_i^{v^+} + \omega_j^{\mu^+} - \omega_j^{v^+} + 2), \\ a_{jW}^{\mu^-} = 2\omega_j^{\mu^-}/(\omega_j^{\mu^-} - \omega_j^{v^-} + \omega_W^{\mu^-} - \omega_W^{v^-} + 2), a_{jW}^{v^+} = 2\omega_W^{\mu^-}/(\omega_j^{\mu^-} - \omega_j^{v^-} + \omega_W^{\mu^-} - \omega_W^{v^-} + 2), \\ a_{jW}^{\mu^+} = 2\omega_j^{\mu^+}/(\omega_j^{\mu^+} - \omega_j^{v^+} + \omega_W^{\mu^+} - \omega_W^{v^+} + 2), a_{jW}^{v^-} = 2\omega_W^{\mu^+}/(\omega_j^{\mu^+} - \omega_j^{v^+} + \omega_W^{\mu^+} - \omega_W^{v^+} + 2). \end{cases} \quad (10)$$

由于 $0 < a_{Bj}^{\mu^-}, a_{Bj}^{v^+}, a_{Bj}^{\mu^+}, a_{Bj}^{v^-}, a_{jW}^{\mu^-}, a_{jW}^{v^+}, a_{jW}^{\mu^+}, a_{jW}^{v^-} < 1$, 式 (10) 可以转化为

$$\begin{cases} a_{Bj}^{\mu^-}(\omega_B^{\mu^-} - \omega_B^{v^-} + \omega_j^{\mu^-} - \omega_j^{v^-} + 2) = 2\omega_B^{\mu^-}, a_{Bj}^{v^+}(\omega_B^{\mu^-} - \omega_B^{v^-} + \omega_j^{\mu^-} - \omega_j^{v^-} + 2) = 2\omega_j^{\mu^-}, \\ a_{Bj}^{\mu^+}(\omega_B^{\mu^+} - \omega_B^{v^+} + \omega_j^{\mu^+} - \omega_j^{v^+} + 2) = 2\omega_B^{\mu^+}, a_{Bj}^{v^-}(\omega_B^{\mu^+} - \omega_B^{v^+} + \omega_j^{\mu^+} - \omega_j^{v^+} + 2) = 2\omega_j^{\mu^+}, \\ a_{jW}^{\mu^-}(\omega_j^{\mu^-} - \omega_j^{v^-} + \omega_W^{\mu^-} - \omega_W^{v^-} + 2) = 2\omega_j^{\mu^-}, a_{jW}^{v^+}(\omega_j^{\mu^-} - \omega_j^{v^-} + \omega_W^{\mu^-} - \omega_W^{v^-} + 2) = 2\omega_W^{\mu^-}, \\ a_{jW}^{\mu^+}(\omega_j^{\mu^+} - \omega_j^{v^+} + \omega_W^{\mu^+} - \omega_W^{v^+} + 2) = 2\omega_j^{\mu^+}, a_{jW}^{v^-}(\omega_j^{\mu^+} - \omega_j^{v^+} + \omega_W^{\mu^+} - \omega_W^{v^+} + 2) = 2\omega_W^{\mu^+}. \end{cases} \quad (11)$$

但式 (11) 的左右两边都存在偏差, 为了确定最优的 IVIF 权重, 引入一些偏差变量, 建立一个最小化总偏差的目标规划模型为

$$\begin{aligned} \min \sum_{j=1}^m (\xi_{Bj}^+ + \xi_{Bj}^- + \varphi_{Bj}^+ + \varphi_{Bj}^- + \beta_{Bj}^+ + \beta_{Bj}^- + \gamma_{Bj}^+ + \gamma_{Bj}^- + \eta_{jW}^+ + \eta_{jW}^- + \lambda_{jW}^+ + \lambda_{jW}^- + \theta_{jW}^+ + \theta_{jW}^- + \sigma_{jW}^+ + \sigma_{jW}^-), \\ \begin{cases} 2\omega_B^{\mu^-} - a_{Bj}^{\mu^-}(\omega_B^{\mu^-} - \omega_B^{v^-} + \omega_j^{\mu^-} - \omega_j^{v^-} + 2) + \xi_{Bj}^+ - \xi_{Bj}^- = 0, 2\omega_j^{\mu^-} - a_{Bj}^{v^+}(\omega_B^{\mu^-} - \omega_B^{v^-} + \omega_j^{\mu^-} - \omega_j^{v^-} + 2) + \varphi_{Bj}^+ - \varphi_{Bj}^- = 0, \\ 2\omega_B^{\mu^+} - a_{Bj}^{\mu^+}(\omega_B^{\mu^+} - \omega_B^{v^+} + \omega_j^{\mu^+} - \omega_j^{v^+} + 2) + \beta_{Bj}^+ - \beta_{Bj}^- = 0, 2\omega_j^{\mu^+} - a_{Bj}^{v^-}(\omega_B^{\mu^+} - \omega_B^{v^+} + \omega_j^{\mu^+} - \omega_j^{v^+} + 2) + \gamma_{Bj}^+ - \gamma_{Bj}^- = 0, \\ 2\omega_j^{\mu^-} - a_{jW}^{\mu^-}(\omega_j^{\mu^-} - \omega_j^{v^-} + \omega_W^{\mu^-} - \omega_W^{v^-} + 2) + \eta_{jW}^+ - \eta_{jW}^- = 0, 2\omega_W^{\mu^-} - a_{jW}^{v^+}(\omega_j^{\mu^-} - \omega_j^{v^-} + \omega_W^{\mu^-} - \omega_W^{v^-} + 2) + \lambda_{jW}^+ - \lambda_{jW}^- = 0, \\ 2\omega_j^{\mu^+} - a_{jW}^{\mu^+}(\omega_j^{\mu^+} - \omega_j^{v^+} + \omega_W^{\mu^+} - \omega_W^{v^+} + 2) + \theta_{jW}^+ - \theta_{jW}^- = 0, 2\omega_W^{\mu^+} - a_{jW}^{v^-}(\omega_j^{\mu^+} - \omega_j^{v^+} + \omega_W^{\mu^+} - \omega_W^{v^+} + 2) + \sigma_{jW}^+ - \sigma_{jW}^- = 0, \\ \omega_i^{\mu^-} + \omega_i^{v^-} \leq 1, \sum_{j=1}^m \omega_j^{\mu^-} \leq \omega_i^{v^-}, \omega_i^{\mu^-} + n - 2 \geq \sum_{j=1}^m \omega_j^{v^-}, i = 1, 2, \dots, n, \\ \omega_i^{\mu^+} + \omega_i^{v^+} \leq 1, \sum_{j=1}^m \omega_j^{\mu^+} \leq \omega_i^{v^+}, \omega_i^{\mu^+} + n - 2 \geq \sum_{j=1}^m \omega_j^{v^+}, i = 1, 2, \dots, n, \\ \max\{\omega_i^{\mu^-}, \omega_i^{\mu^+}\} + \max\{\omega_i^{v^-}, \omega_i^{v^+}\} \leq 1, \\ \omega_i^{\mu^-}, \omega_i^{\mu^+}, \omega_i^{v^-}, \omega_i^{v^+} \geq 1. \end{cases} \end{aligned}$$

其中,

$$\begin{cases} \xi_{Bj}^+ = (2\omega_B^{\mu^-} - a_{Bj}^{\mu^-}(\omega_B^{\mu^-} - \omega_B^{v^-} + \omega_j^{\mu^-} - \omega_j^{v^-} + 2)) \vee 0, \xi_{Bj}^- = (a_{Bj}^{\mu^-}(\omega_B^{\mu^-} - \omega_B^{v^-} + \omega_j^{\mu^-} - \omega_j^{v^-} + 2) - 2\omega_B^{\mu^-}) \vee 0, \\ \varphi_{Bj}^+ = (2\omega_j^{\mu^-} - a_{Bj}^{v^+}(\omega_B^{\mu^-} - \omega_B^{v^-} + \omega_j^{\mu^-} - \omega_j^{v^-} + 2)) \vee 0, \varphi_{Bj}^- = (a_{Bj}^{v^+}(\omega_B^{\mu^-} - \omega_B^{v^-} + \omega_j^{\mu^-} - \omega_j^{v^-} + 2) - 2\omega_j^{\mu^-}) \vee 0, \\ \beta_{Bj}^+ = (2\omega_B^{\mu^+} - a_{Bj}^{\mu^+}(\omega_B^{\mu^+} - \omega_B^{v^+} + \omega_j^{\mu^+} - \omega_j^{v^+} + 2)) \vee 0, \beta_{Bj}^- = (a_{Bj}^{\mu^+}(\omega_B^{\mu^+} - \omega_B^{v^+} + \omega_j^{\mu^+} - \omega_j^{v^+} + 2) - 2\omega_B^{\mu^+}) \vee 0, \\ \gamma_{Bj}^+ = (2\omega_j^{\mu^+} - a_{Bj}^{v^-}(\omega_B^{\mu^+} - \omega_B^{v^+} + \omega_j^{\mu^+} - \omega_j^{v^+} + 2)) \vee 0, \gamma_{Bj}^- = (a_{Bj}^{v^-}(\omega_B^{\mu^+} - \omega_B^{v^+} + \omega_j^{\mu^+} - \omega_j^{v^+} + 2) - 2\omega_j^{\mu^+}) \vee 0, \\ \eta_{jW}^+ = (2\omega_j^{\mu^-} - a_{jW}^{\mu^-}(\omega_j^{\mu^-} - \omega_j^{v^-} + \omega_W^{\mu^-} - \omega_W^{v^-} + 2)) \vee 0, \eta_{jW}^- = (a_{jW}^{\mu^-}(\omega_j^{\mu^-} - \omega_j^{v^-} + \omega_W^{\mu^-} - \omega_W^{v^-} + 2) - 2\omega_j^{\mu^-}) \vee 0, \\ \lambda_{jW}^+ = (2\omega_W^{\mu^-} - a_{jW}^{v^+}(\omega_j^{\mu^-} - \omega_j^{v^-} + \omega_W^{\mu^-} - \omega_W^{v^-} + 2)) \vee 0, \lambda_{jW}^- = (a_{jW}^{v^+}(\omega_j^{\mu^-} - \omega_j^{v^-} + \omega_W^{\mu^-} - \omega_W^{v^-} + 2) - 2\omega_j^{\mu^-}) \vee 0, \\ \theta_{jW}^+ = (2\omega_j^{\mu^+} - a_{jW}^{\mu^+}(\omega_j^{\mu^+} - \omega_j^{v^+} + \omega_W^{\mu^+} - \omega_W^{v^+} + 2)) \vee 0, \theta_{jW}^- = (a_{jW}^{\mu^+}(\omega_j^{\mu^+} - \omega_j^{v^+} + \omega_W^{\mu^+} - \omega_W^{v^+} + 2) - 2\omega_j^{\mu^+}) \vee 0, \\ \sigma_{jW}^+ = (2\omega_W^{\mu^+} - a_{jW}^{v^-}(\omega_j^{\mu^+} - \omega_j^{v^+} + \omega_W^{\mu^+} - \omega_W^{v^+} + 2)) \vee 0, \sigma_{jW}^- = (a_{jW}^{v^-}(\omega_j^{\mu^+} - \omega_j^{v^+} + \omega_W^{\mu^+} - \omega_W^{v^+} + 2) - 2\omega_j^{\mu^+}) \vee 0. \end{cases}$$

通过 LINGO 软件求解该模型，可以得到各评价指标的最优标准化 IVIF 权重向量 $\omega = \{\omega_1, \omega_2, \dots, \omega_n\}$ ，其中， $\omega_i = ([\omega_i^{\mu-}, \omega_i^{\mu+}], [\omega_i^{v-}, \omega_i^{v+}])$ 。

3) 计算一致性指数。在此阶段，对于 IVIF、BWM 的 CI 直接模糊化为

$$(\tilde{a}_{Bj} - \zeta) \times (\tilde{a}_{jW} - \zeta) = \tilde{a}_{BW} + \zeta \tag{12}$$

是错误的，因为当 IVIFPR 符合式 (4) 而不是式 (12) 时，它是乘法一致的。

根据定义 5，给出所有 IVIF 参考比较的完全一致比较的定义。

定义 6 所有的 IVIF 参考比较是完全一致的，如果 $\rho_{Bj}^l \rho_{jW}^l = \rho_{BW}^l$ ， $\rho_{Bj}^r \rho_{jW}^r = \rho_{BW}^r$ ，其中： $\rho_{Bj}^l = a_{Bj}^{\mu-}/a_{Bj}^{v+}$ ， $\rho_{jW}^l = a_{jW}^{\mu-}/a_{jW}^{v+}$ ， $\rho_{Bj}^r = a_{Bj}^{\mu+}/a_{Bj}^{v-}$ ， $\rho_{jW}^r = a_{jW}^{\mu+}/a_{jW}^{v-}$ 。

但是，对于某些 j 可能不是完全一致，即

$$\rho_{Bj}^l \rho_{jW}^l \neq \rho_{BW}^l, \tag{13}$$

$$\rho_{Bj}^r \rho_{jW}^r \neq \rho_{BW}^r. \tag{14}$$

当 $\rho_{Bj}^l = \rho_{jW}^l = \rho_{BW}^l$ 、 $\rho_{Bj}^r = \rho_{jW}^r = \rho_{BW}^r$ 时，式 (13) ~ 式 (14) 达到最大不等式。设 ζ_1 为 ρ_{Bj}^l 和 ρ_{jW}^l 分别减去的值，将 ζ_1 加到 ρ_{BW}^l 中，达到式 (13) 的最大不等式； ζ_2 为 ρ_{Bj}^r 和 ρ_{jW}^r 分别减去的值，将 ζ_2 加到 ρ_{BW}^r 中，达到式 (14) 的最大不等式，则有

$$(\rho_{Bj}^l - \zeta_1) \times (\rho_{jW}^l - \zeta_1) = \rho_{BW}^l + \zeta_1, \tag{15}$$

$$(\rho_{Bj}^r - \zeta_2) \times (\rho_{jW}^r - \zeta_2) = \rho_{BW}^r + \zeta_2. \tag{16}$$

对于最小一致性 $\rho_{Bj}^l = \rho_{jW}^l = \rho_{BW}^l$ 、 $\rho_{Bj}^r = \rho_{jW}^r = \rho_{BW}^r$ ，有

$$\zeta_1^2 - (1 + \rho_{BW}^l)\zeta_1 + (\rho_{BW}^{l^2} - \rho_{BW}^l) = 0, \tag{17}$$

$$\zeta_2^2 - (1 + \rho_{BW}^r)\zeta_2 + (\rho_{BW}^{r^2} - \rho_{BW}^r) = 0. \tag{18}$$

对于不同的 ρ_{BW}^l 、 ρ_{BW}^r ，通过求解式 (17) ~ 式 (18)，得到 ζ_1 、 ζ_2 最大可能值，其中， ζ_1 、 ζ_2 的最大可能值定义为一致性指数 CI_1 和 CI_2 。表 3 给出部分不同 \tilde{a}_{BW} 对应的 ρ_{BW}^l 、 ρ_{BW}^r 的 CI_1 和 CI_2 。

表 3 IVIF-BWM 的一致性指数 CI_1 和 CI_2
Tab. 3 CI_1 and CI_2 of IVIF-BWM

指标	a_{BW}				
	$([0.5, 0.5], [0.5, 0.5])$	$([0.5, 0.6], [0.3, 0.4])$	$([0.6, 0.7], [0.2, 0.3])$	$([0.7, 0.8], [0.1, 0.2])$	$([0.8, 0.9], [0.0, 0.1])$
ρ_{BW}^l	1	1.250 0	2.000 0	3.500 0	8.000 0
CI_1	0	0.091 7	0.438 4	1.307 4	4.468 9
ρ_{BW}^r	1	2.000 0	3.500 0	8.000 0	9.000 0
CI_2	0	0.438 4	1.307 4	4.468 9	5.228 0

根据式 (9) 和定义 2，如果所有的 IVIF 参考比较是完全一致的，则 $\rho_{Bj}^l = a_{Bj}^{\mu-}/a_{Bj}^{v+}$ 、 $\rho_{jW}^l = a_{jW}^{\mu-}/a_{jW}^{v+}$ 、 $\rho_{Bj}^r = a_{Bj}^{\mu+}/a_{Bj}^{v-}$ 、 $\rho_{jW}^r = a_{jW}^{\mu+}/a_{jW}^{v-}$ 。

与文献 [20] 中的式 (1) 同理，建立公式

$$\begin{cases} \zeta_1^* = \min \max \{ |a_{Bj}^{\mu-}/a_{Bj}^{v+} - \omega_B^{\mu-}/\omega_j^{\mu-}|, |a_{jW}^{\mu-}/a_{jW}^{v+} - \omega_j^{\mu-}/\omega_W^{\mu-}| \}, \\ \zeta_2^* = \min \max \{ |a_{Bj}^{\mu+}/a_{Bj}^{v-} - \omega_B^{\mu+}/\omega_j^{\mu+}|, |a_{jW}^{\mu+}/a_{jW}^{v-} - \omega_j^{\mu+}/\omega_W^{\mu+}| \}. \end{cases} \tag{19}$$

定义 7 一致性比率 (CR) 被定义为

$$CR = \max \{ \zeta_1^*/CI_1, \zeta_2^*/CI_2 \}. \tag{20}$$

其中， ζ_1^* 和 ζ_2^* 由式 (19) 得到。

根据式 (20) 计算 CR，若 $CR \leq 0.1$ ，则 IVIF 比较向量是可接受一致的。

3.2 改进 IVIF-BWM 的权重确定

为了科学合理地确定权重, 本研究通过问卷形式收集关于高考数学试题难度评价的数据。结合数据, 权重确定过程分为难度指标的权重确定和难度水平的权重确定。

3.2.1 难度指标的权重确定

- 1) 建立评价指标体系。本研究采用所建立的新高考数学试题难度指标模型, 记为 $\{C_1, C_2, C_3, C_4, C_5, C_6, C_7, C_8, C_9\}$ 。
- 2) 确定最优指标 C_B 和最劣指标 C_W 。通过问卷及对调查对象的讨论, 最终将“推理能力 C_4 ”确定为最优指标, 将“试题情境 C_1 ”确定为最劣指标。
- 3) 使用表 2 的语言术语确定最优指标相对于其他指标的偏好 BO 和其他指标相对于最劣指标的偏好 OW, 具体偏好度如表 4 所示。

表 4 各评价指标的偏好度和子指标的最优 IVIFN 权重
Tab. 4 Preference of each evaluation index and optimal IVIFN weights for sub-indicators

指标	最优指标 C_7	最劣指标 C_1	最优权重
C_1	$([0.8, 0.9], [0.1, 0.1])$	$([0.5, 0.5], [0.5, 0.5])$	$([0.034, 0.034], [0.966, 0.966])$
C_2	$([0.5, 0.6], [0.4, 0.4])$	$([0.7, 0.8], [0.2, 0.2])$	$([0.137, 0.141], [0.805, 0.859])$
C_3	$([0.6, 0.7], [0.3, 0.3])$	$([0.6, 0.7], [0.3, 0.3])$	$([0.085, 0.085], [0.869, 0.910])$
C_4	$([0.5, 0.5], [0.5, 0.5])$	$([0.8, 0.9], [0.1, 0.1])$	$([0.181, 0.204], [0.796, 0.796])$
C_5	$([0.7, 0.8], [0.2, 0.2])$	$([0.5, 0.6], [0.4, 0.4])$	$([0.05, 0.051], [0.919, 0.949])$
C_6	$([0.5, 0.6], [0.4, 0.4])$	$([0.7, 0.8], [0.2, 0.2])$	$([0.137, 0.141], [0.805, 0.859])$
C_7	$([0.5, 0.5], [0.5, 0.5])$	$([0.8, 0.9], [0.1, 0.1])$	$([0.18, 0.204], [0.796, 0.796])$
C_8	$([0.6, 0.7], [0.3, 0.3])$	$([0.6, 0.7], [0.3, 0.3])$	$([0.085, 0.085], [0.869, 0.910])$
C_9	$([0.7, 0.8], [0.2, 0.2])$	$([0.5, 0.6], [0.4, 0.4])$	$([0.05, 0.051], [0.919, 0.949])$

4) 确定指标的最优权重。构建目标规划模型, 运用 LINGO 软件求解模型, 得到各指标的最优标准化 IVIF 权重向量 $\omega = \{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6, \omega_7, \omega_8, \omega_9\}$, 如表 4 所示。同时可得 $\xi_1^* = 0.221$ 、 $\xi_2^* = 0$ 、 $CR = 0.049$ 。

- 5) 一致性检验得 $CR = 0.049 < 0.1$, 故表 4 的比较向量是可接受一致的。
- 6) 用公式

$$\omega_i = (\omega_i^{\mu-} + \omega_i^{\mu+} - \omega_i^{v-} - \omega_i^{v+} + 2)/4$$

(21)

将各评价指标的最优标准化 IVIF 权重向量转化为精确值权重, 分别为 $C_1 = 0.034$ 、 $C_2 = 0.154$ 、 $C_3 = 0.098$ 、 $C_4 = 0.198$ 、 $C_5 = 0.058$ 、 $C_6 = 0.154$ 、 $C_7 = 0.198$ 、 $C_8 = 0.098$ 、 $C_9 = 0.058$ 。

3.2.2 难度水平的权重确定

- 1) 建立评价指标体系。新高考数学试题综合难度模型的 9 个指标分别对应不同难度水平, 因此, 本研究将每个指标 $C_i (i = 1, 2, \dots, 9)$ 的难度水平作为相应的子指标体系。
- 2) 确定各指标 C_i 的最优子指标 C_{iB} 和最劣子指标 C_{iW} 。
- 3) 确定最优子指标相对于其他子指标的偏好和其他子指标相对于最劣子指标的偏好, 得到的 BO_i 和 OW_i 如表 5 所示。

表 5 各难度水平的偏好度
Tab. 5 Difficulty level preference

指标	子指标	C_{11}		C_{12}		C_{13}	
C_1	最优子指标 C_{12}	([0.8,0.85],[0.1,0.1])		([0.5,0.5],[0.5,0.5])		([0.5,0.55],[0.4,0.4])	
	最劣子指标 C_{11}	([0.5,0.5],[0.5,0.5])		([0.8,0.85],[0.1,0.1])		([0.7,0.75],[0.2,0.2])	
指标	子指标	C_{21}		C_{22}		C_{23}	
C_2	最优子指标 C_{23}	([0.8,0.9],[0.1,0.1])		([0.6,0.7],[0.3,0.3])		([0.5,0.5],[0.5,0.5])	
	最劣子指标 C_{21}	([0.5,0.5],[0.5,0.5])		([0.6,0.7],[0.3,0.3])		([0.8,0.9],[0.1,0.1])	
指标	子指标	C_{31}		C_{32}	C_{33}	C_{34}	
C_3	最优子指标 C_{34}	([0.7,0.75],[0.2,0.2])		([0.6,0.65],[0.3,0.3])		([0.5,0.55],[0.4,0.4])	
	最劣子指标 C_{31}	([0.5,0.5],[0.5,0.5])		([0.5,0.55],[0.4,0.4])		([0.6,0.65],[0.3,0.3])	
指标	子指标	C_{41}		C_{42}		C_{43}	
C_4	最优子指标 C_{43}	([0.7,0.75],[0.2,0.2])		([0.6,0.65],[0.3,0.3])		([0.5,0.5],[0.5,0.5])	
	最劣子指标 C_{41}	([0.5,0.5],[0.5,0.5])		([0.6,0.65],[0.3,0.3])		([0.7,0.75],[0.2,0.2])	
指标	子指标	C_{51}		C_{52}		C_{53}	
C_5	最优子指标 C_{54}	([0.8,0.9],[0.1,0.1])		([0.6,0.7],[0.3,0.3])		([0.5,0.5],[0.5,0.5])	
	最劣子指标 C_{51}	([0.5,0.5],[0.5,0.5])		([0.6,0.7],[0.3,0.3])		([0.8,0.9],[0.1,0.1])	
指标	子指标	C_{61}		C_{62}	C_{63}	C_{64}	
C_6	最优子指标 C_{64}	([0.8,0.85],[0.1,0.1])		([0.7,0.75],[0.2,0.2])		([0.5,0.55],[0.4,0.4])	
	最劣子指标 C_{61}	([0.5,0.5],[0.5,0.5])		([0.5,0.55],[0.4,0.4])		([0.7,0.75],[0.2,0.2])	
指标	子指标	C_{71}		C_{72}		C_{73}	
C_7	最优子指标 C_{73}	([0.75,0.8],[0.2,0.2])		([0.55,0.6],[0.4,0.4])		([0.5,0.5],[0.5,0.5])	
	最劣子指标 C_{71}	([0.5,0.5],[0.5,0.5])		([0.75,0.8],[0.2,0.2])		([0.55,0.6],[0.4,0.4])	
指标	子指标	C_{81}		C_{82}		C_{83}	
C_8	最优子指标 C_{83}	([0.75,0.8],[0.2,0.2])		([0.65,0.7],[0.3,0.3])		([0.5,0.5],[0.5,0.5])	
	最劣子指标 C_{81}	([0.5,0.5],[0.5,0.5])		([0.65,0.7],[0.3,0.3])		([0.55,0.6],[0.4,0.4])	
指标	子指标	C_{91}	C_{92}	C_{93}	C_{94}	C_{95}	
C_9	最优子指标 C_{95}	([0.8,0.85],[0.1,0.1])	([0.7,0.75],[0.2,0.2])	([0.6,0.65],[0.3,0.3])	([0.5,0.55],[0.4,0.4])	([0.5,0.5],[0.5,0.5])	
	最劣子指标 C_{91}	([0.5,0.5],[0.5,0.5])	([0.5,0.55],[0.4,0.4])	([0.6,0.65],[0.3,0.3])	([0.7,0.75],[0.2,0.2])	([0.8,0.85],[0.1,0.1])	

4) 确定权重。构建目标规划模型，运用 LINGO 求解模型，得到各指标不同难度水平的最优标准化 IVIF 权重向量 $\omega_i = \{\omega_1, \omega_2, \cdots, \omega_j\}$, $j(j = 1, 2, \cdots)$ 代表各指标不同难度水平数。

5) 用式 (5) 进行一致性检验，得到 $CR_1 = 0.015\ 7 < 0.1$, $CR_2 = 0.099\ 5 < 0.1$, $CR_3 = 0.061\ 6 < 0.1$, $CR_4 = 0.060\ 5 < 0.1$, $CR_5 = 0.099\ 5 < 0.1$, $CR_6 = 0.014\ 6 < 0.1$, $CR_7 = 0.055\ 9 < 0.1$, $CR_8 = 0.029\ 5 < 0.1$, $CR_9 = 0.017\ 3 < 0.1$ ，故表 6 中的比较向量是可接受一致的。

6) 用式 (21) 将各指标不同水平的 IVIF 权重向量化为精确值权重，如表 6 所示。

表 6 难度水平的精确值权重
Tab. 6 Exact value weighting of the difficulty level

指标	C_{11}	C_{12}	C_{13}	C_{21}	C_{22}	C_{23}	C_{31}	C_{32}	C_{33}	C_{34}	
精确值权重	0.076 7	0.510 2	0.413 3	0.077 5	0.293 8	0.622 8	0.121 0	0.199 2	0.306 0	0.364 5	
指标	C_{41}	C_{42}	C_{43}	C_{51}	C_{52}	C_{53}	C_{61}	C_{62}	C_{63}	C_{64}	
精确值权重	0.141 2	0.297 8	0.560 0	0.077 5	0.293 8	0.622 8	0.087 0	0.142 0	0.354 0	0.413 5	
指标	C_{71}	C_{72}	C_{73}	C_{81}	C_{82}	C_{83}	C_{91}	C_{92}	C_{93}	C_{94}	C_{95}
精确值权重	0.130 0	0.358 8	0.511 0	0.149 0	0.265 8	0.585 3	0.070 5	0.106 2	0.178 7	0.277 2	0.337 2

说明: 表中部分权重因四舍五入的原因, 存在分项合计不等于 1 的情况。

4 新高考数学试题综合难度模型的应用

为验证构建的新高考数学试题综合难度模型的有效性, 本研究结合《中国高考丛书——高考试题分析》, 对 2022 年新高考 I 卷和 II 卷的试题编码赋值, 从各难度指标不同水平、各个指标难度和试题综合难度 3 方面分析 2022 年新高考试卷的难度特征。

4.1 各难度指标不同水平

根据编码结果, 统计每份试卷每道试题的各指标所处难度水平的试题数, 计算各指标不同难度水平试题数的比例, 绘制成的折线图如图 1 所示。由图 1 可以看出, 2022 年新高考 I 卷和 II 卷对“试题情境”“推理水平”“开放程度”的考查情况基本一致。在“试题情境”中, 课程学习情境考查比例最高, I 卷和 II 卷均达 50%, 其次是课程创新情境, 而社会实践情境考查最少, 体现新高考注重考查围绕数学知识展开和拓展。“推理水平”对 3 个难度水平的考查比例较均匀, 复杂推理比例略高一些, 体现出新高考数学注重考查学生的逻辑思维能力。“开放程度”以考查封闭题型为主, 对于开放性题型, 结论开放型和策略开放型考查较多, 体现了 2022 年新高考数学在保证基础题考查的前提下, 注重创新试题形式, 增强试题结论的开放性, 鼓励学生运用创造性、发散性思维分析和解决问题。

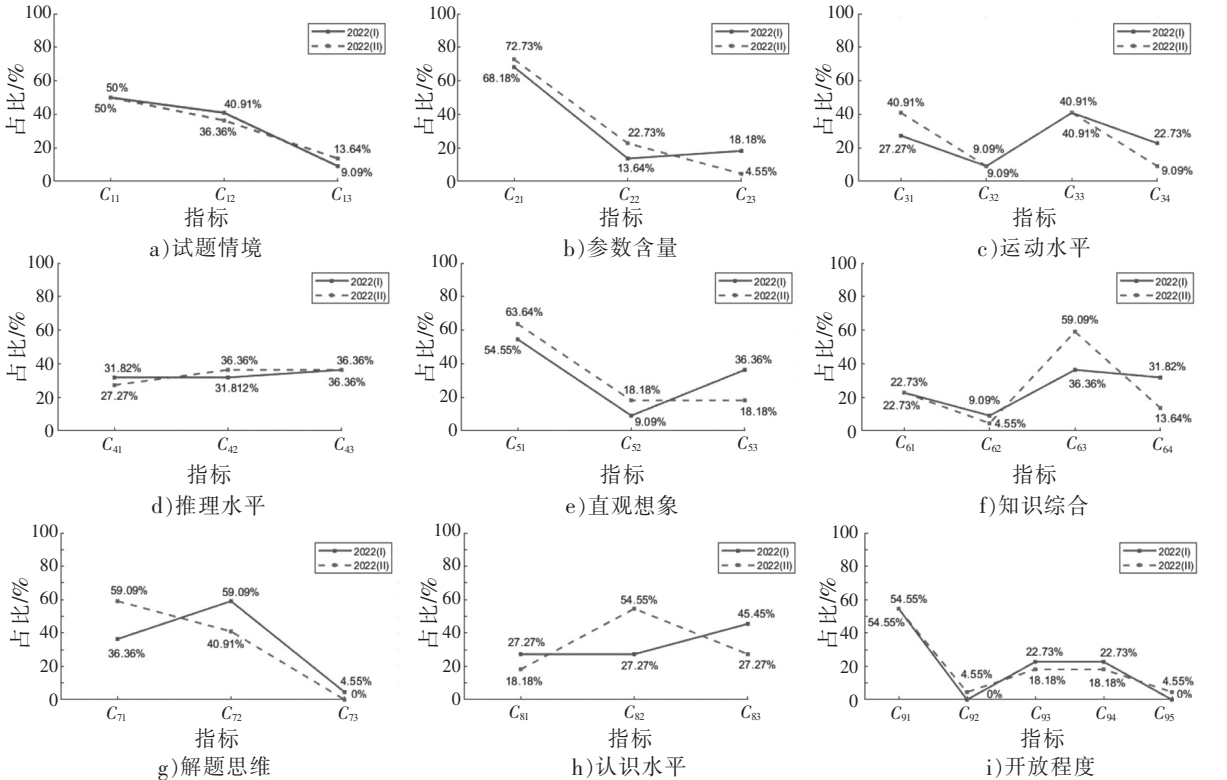


图 1 各个指标难度水平的折线图
Fig.1 Line chart of the difficulty level of each indicator

2022 年新高考 I 卷和 II 卷对“参数含量”“运算水平”“直观想象”“知识综合”“解题思维”“认知水平”的考查差异较大。在“参数含量”中, II 卷考查无参数、含参无需讨论的比例高于 I 卷, 但 I 卷含参需讨论的比例高于 II 卷, 总体不含参数的试题比含参数试题多, 体现出 2022 年新高考数学在考查基础知识和方法的前提下, 注重对理性思维的考查。对于“运算水平”, I 卷和 II 卷对一般运算、较复杂运算的考查比例一致, 而 I 卷的复杂运算考查比例高于 II 卷, 这反映了 2022 年新高考以关键能力为重点, I 卷更注重综合考查逻辑思维和运算求解能力。对于“直观想象”, I 卷考查复杂直观想象的比例高于 II 卷。在“知识综合”中, I 卷和 II 卷对单知识点的比例一致, I 卷对多知识点、知识点拓展比例均高于 II 卷, 而在知识点关联比例低于 II 卷, 但总体以知识关联、知识拓展为考查重点, 体现了新高考注重考查知识体系的综合性和知识间的联系。在“解题思维”上, I 卷在一般思维、复杂思维水平上的考查比例均高于 II 卷。在“认知水平”中, I 卷对理解、分析水平上的考查比例高于 II 卷, 但 II 卷在应用水平的比例高出 I 卷 2 倍。

4.2 各指标难度

利用式 (7) 计算各个指标的难度系数 d_i , 并绘制 2022 年新高考 I 卷和 II 卷综合难度雷达图 (见图 2)。由图 2 可见, 2022 年新高考 I 卷和 II 卷在试题情境、认知水平、开放程度的综合难度系数基本保持一致, 而 I 卷在参数含量、运算水平、解题思维和认知水平的难度高于 II 卷, 在推理能力和直观想象低于 II 卷。从各指标难度系数来看, 2022 年新高考难度主要集中在试题情境、推理水平、知识综合、解题思维、认知水平。

高考是基于学生所学知识对能力素养的考查, 通过知识的考查深度来反映学生的认知能力, 而推理能力是数学学科的关键能力之一, 因此知识综合、认知水平及推理能力成为影响试卷难度的重要指标。试题情境是考查数学知识、能力、素养的承载因素, 解题思维的曲折性直接影响考生对试题的求解, 均对试卷难度具有重要作用。开放程度的难度系数最小, 而参数含量、推理水平、直观想象处于中间地位, 一定程度上影响了试卷的难度。

4.3 综合难度

利用式 (8) 聚合各个指标的难度系数, 可得 2022 年新高考 I 卷和 II 卷的综合难度系数分别为 $D_I = 0.29$ 、 $D_{II} = 0.27$, 明显 $D_I > D_{II}$, 即 2022 年新高考 I 卷难度高于 II 卷。这一难度分布契合了我国不同区域之间经济、文化等方面发展水平不一的特点。

5 结论

基于新高考改革背景, 本研究以《中国高考评价体系》为指导, 根据新高考数学试题的特点, 建立含有 9 个评价指标和 31 个难度水平的新高考数学试题综合难度模型。针对现有 IVIF-BWM 的不足, 本文提出一种改进的 IVIF-BWM, 建立了基于乘法一致 IVIFPR 的目标规划模型, 求解标准化的最优 IVIF 权重, 并推导出区间值直觉模糊环境下的 CI 和 CR, 验证了推导结果的可靠性。同时, 将

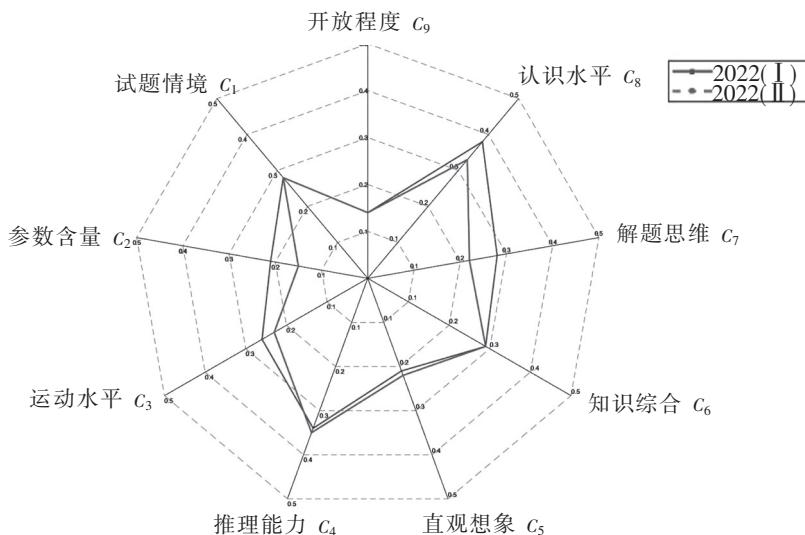


图 2 2022 年新高考数学试卷 I 卷和 II 卷综合难度雷达图
Fig.2 Comprehensive difficulty radar diagram for paper I and II of the college exam mathematics questions in 2022

改进的 IVIF-BWM 用于计算新高考试题难度指标和难度水平的权重,验证了所改进的 IVIF-BWM 的准确性和可行性。最后运用所提出的新高考数学试题综合难度模型评价 2022 年新高考数学试卷的难度,结果验证了该模型具有的有效性,进一步丰富了新高考数学试题评价方法。然而,本文只讨论了 IVIFPR 的乘法一致性,将 BWM 扩展到具有加法一致性的 IVIFPR 中还需进一步研究。

[参考文献]

- [1] 鲍建生. 中英两国初中数学期望课程综合难度的比较[J]. 全球教育展望,2002,31(9):48-52.
- [2] 张怡,武小鹏,彭乃霞. 综合难度系数模型在 2016 年高考数学试题评价中的应用[J]. 教育测量与评价,2016(12):47-53.
- [3] 武小鹏,张怡. 中国和韩国高考数学试题综合难度比较研究[J]. 数学教育学报,2018,27(3):19-24,29.
- [4] 薛欢,杜剑南,路江江. 2016—2020 年高考数学(理科)全国卷“统计与概率”试题探析:基于综合难度模型[J]. 教育测量与评价,2020(12):30-40.
- [5] 李保臻,石烨. 中国大陆与台湾地区高考数学试题难度比较研究:以 2016—2018 年大陆全国卷 I 与台湾指考试题为例[J]. 数学教育学报,2020,29(1):58-64.
- [6] 武小鹏,孔企平. 基于 AHP 理论的数学高考试题综合难度模型构建与应用[J]. 数学教育学报,2020,29(2):29-34.
- [7] MI X M, TANG M, LIAO H C, et al. The state-of-the-art survey on integrations and applications of the best worst method in decision making: why, what, what for and what's next? [J]. Omega, 2019, 87: 205-225.
- [8] REZAEI J. Best-worst multi-criteria decision-making method[J]. Omega, 2015, 53: 49-57.
- [9] GAO F, WANG W X, BI C C, et al. Prioritization of used aircraft acquisition criteria: a fuzzy best-worst method (BWM)-based approach[J]. Journal of Air Transport Management, 2023, 107: 102359.
- [10] YOU X S, CHEN T, YANG Q. Approach to multi-criteria group decision-making problems based on the best-worst-method and electre method[J]. Symmetry, 2016, 8(9): 95.
- [11] MOU Q, XU Z S, LIAO H C. An intuitionistic fuzzy multiplicative best-worst method for multi-criteria group decision making [J]. Information Sciences, 2016, 374: 224-239.
- [12] PAMUČAR D, PETROVIKĆI, ČIROVIĆ G. Modification of the best-worst and MABAC methods: a novel approach based on interval-valued fuzzy-rough numbers[J]. Expert Systems with Applications, 2018, 91: 89-106.
- [13] LIAO H C, MI X M, YU Q, et al. Hospital performance evaluation by a hesitant fuzzy linguistic best worst method with inconsistency repairing[J]. Journal of Cleaner Production, 2019, 232: 657-671.
- [14] CHEN Z H, WAN S P, DONG J Y. An integrated interval-valued intuitionistic fuzzy technique for resumption risk assessment amid COVID-19 prevention[J]. Information Sciences, 2023, 619: 695-721.
- [15] WANG J, MA Q Q, LIU H C. A meta-evaluation model on science and technology project review experts using IVIF-BWM and MULTIMOORA[J]. Expert Systems with Applications, 2021, 168: 114236.
- [16] ALKAN N, KAHRAMAN C. Prioritization of factors affecting the digitalization of quality management using interval-valued intuitionistic fuzzy best-worst method[C]//INFUS. Intelligent and Fuzzy Techniques for Emerging Conditions and Digital Transformation. New Delhi, India: Springer International Publishing, 2022: 28-39.
- [17] ATANASSOV K T, ATANASSOV K T. Interval valued intuitionistic fuzzy sets[J]. Intuitionistic Fuzzy Sets: Theory and Applications, 1989, 31(3): 343-349.
- [18] WAN S P, DONG J Y. A novel method for group decision making with interval-valued Atanassov intuitionistic fuzzy preference relations[J]. Information Sciences, 2016, 372: 53-71.
- [19] WANG Z J. Derivation of intuitionistic fuzzy weights based on intuitionistic fuzzy preference relations[J]. Applied Mathematical Modelling, 2013, 37(9): 6377-6388.
- [20] REZAEI J. Best-worst multi-criteria decision-making method: some properties and a linear model[J]. Omega, 2016, 64: 126-130.

(责任编辑 马建华 英文审校 黄振坤)